# Constraining the QSO Luminosity Function Using Gravitational Lensing Statistics

by

Onsi Joe Fakhouri

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Bachelor of Science in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2004 [June 2004]



Author .....
Department of Physics
May 7th, 2004

Certified by .....
Scott Burles
Assistant Professor
Thesis Supervisor

Accepted by .....
Professor David E. Pritchard
Senior Thesis Coordinator, Department of Physics

# Constraining the QSO Luminosity Function Using Gravitational Lensing Statistics

by
Onsi Joe Fakhouri

Submitted to the Department of Physics
on May 7th, 2004, in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Physics

## Abstract

In this thesis we use gravitational lensing statistics to constrain the QSO luminosity function at a variety of redshifts. We present a theoretical discussion of gravitational lensing statistics and illustrate how high resolution QSO imagery can be used to constrain the QSO luminosity function. We then discuss the selection and observation of the 1073 QSO exposures in our sample. The sample covers a redshift range of $0.7<z<5.5$ and may include up to 10 multiply imaged QSOs. We discuss the QSO analysis pipeline developed to compute the gravitational lensing probabilities of each QSO and then present the constraints on the QSO luminosity function and compare them to results in the literature. Our results confirm the suspected fall off in the high-end QSO luminosity function slope at high redshift and agree with modern literature results. We conclude with a brief discussion of improvements that can be made to our analysis process.

Thesis Supervisor: Scott Burles
Title: Assistant Professor

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction and Structure

This thesis represents the culmination of two and a half years of research under the guidance of Professor Scott Burles in the MIT Physics department. In it, we shall detail a holistic approach to the problem of constraining high redshift QSO statistics using gravitational lensing probabilities. Our approach incorporates several interrelated theoretical, observational, and computational elements and we shall not attempt to describe them in this brief introduction. Rather, we shall motivate and connect the different elements together throughout the thesis.

We will, instead, present a quick overview of the thesis' structure:

**Chapter 2** presents a brief history of QSOs, describes the QSO Luminosity Function, and offers motivation for the thesis' ultimate goal: constraining the QSO Luminosity Function.

**Chapter 3** describes gravitational lensing theoretically. In it we present the Singular Isothermal Sphere (SIS) gravitational lens model and derive the distribution for the gravitational lens population. We then present derivations of the lensing probability equations and describe the means by which high resolution QSO imagery can be used to compute the lensing probabilities. We end the chapter by describing how the lensing probability equations can be used to constrain the QSO Luminosity Function.

**Chapter 4** describes the selection criteria used to select the 1073 QSOs in our high resolution QSO sample. It also discusses the instrumentation used to obtain the imagery.

**Chapter 5** describes the QSO analysis pipeline. It motivates and presents the $\chi^2$ statistic and describes, in detail, the fitting procedure used to compare objects within exposures to identify lensed QSOs. It then describes the simulation techniques used to quantify the empirical lensing probabilities and the limitations of our technique's lens detection capabilities.

**Chapter 6** concludes by presenting the results of the analysis process. It includes a presentation of the identified lensed QSOs in our sample, and the resulting QSO Luminosity Function constraints. The chapter ends with a discussion of the limitations of our approach and offers potential improvements to the theoretical and computational techniques we have used.

The thesis also contains three appendices. The first of these presents an analytical fit to the lens distribution derived in chapter 3. The second describes an alternate normalization of one of the theoretical probabilities derived in chapter 3. The motivation for this renormalization will be

explained in that chapter. The third appendix presents some screenshots and a brief description of the Coordinator and the Analyzer, two software packages written to organize and analyze the QSO sample.

# Chapter 2

# QSOs

## 2.1 A Brief History

The 1940s and 50s saw a boom in radio astronomy. With the development of new technologies and, more importantly, the growing academic interest in radio astronomy, the number of known radio-loud objects in the universe began to grow rapidly. The new radio observations were carried out hand in hand with complimentary optical follow ups of the radio-loud objects; more often than not, the radio sources were identified optically as galaxies.

In 1960, however, things changed when Thomas Mathews and Allan Sandage went in search for an optical counterpart to radio source 3C48, an uncharacteristically small radio source, less than an arc-second in diameter [Thorne]. Using the 5-meter Palomar optical telescope, Sandage was surprised to find a single, star-like, blue point at the radio source's coordinates. Sandage recalls, "I took a spectrum the next night and it was the weirdest spectrum I'd ever seen." [Thorne]

It would only be a few months before the Dutch astronomer Maarten Schmidt would recognize the heavily redshifted Hydrogen Balmer lines in the spectra of these newly-christened **Quasi Stellar Radio Sources (quasars)**. According to Hubble's law these objects, which appear to be moving away from the Earth at relativistic speeds, must be extremely far away and, by virtue of the fact that they can be observed today, must be exceedingly bright. In fact, quasars can be up to $10^5$ times more luminous than normal galaxies [Carroll]. While there is no direct evidence, most astrophysicists are convinced that these objects must be driven by large black holes in their cores; quasars are probably extreme examples of Active Galactic Nuclei (AGNs).

Optical astronomers quickly began actively searching for quasars and it was soon discovered that only about 10% [Carroll] of the spectroscopically confirmed quasars they found were radio-loud. Thus, "quasar" is a misnomer and these point-like high redshift objects are now referred to as **Quasi-Stellar Objects** or **QSOs**. Recent surveys, such as the Sloan Digital Sky Survey (SDSS), have led to an exponential growth in the number of known QSOs. According to [Carroll] there were about 5000 QSOs known in 1996. Today, SDSS alone reports 32,241 QSOs with redshift $< 2.3$ and 3,791 QSOs with redshift $> 2.3$ [SDSS]!

## 2.2 The QSO Luminosity Function

The fact that QSOs are extremely far away and extremely bright makes them excellent probes of the early universe: By understanding the distribution of QSOs we can test the assumption of isotropy and homogeneity in the universe, attempt to understand the underlying physics behind the evolution of QSOs with redshift, and constrain the black hole accretion models that attempt to describe QSOs [Pei].

The most commonly used QSO distribution is the **QSO Luminosity Function (QSO LF)**. We define the QSO LF, $\Psi(L,z)\,dL$, as the number density (per comoving volume) of QSOs at redshift $z$ with luminosity $L < l_{qso} < L + dL$. The commonly accepted empirical model for the QSO LF at $z < 2.2$ is a double power law shaped LF [Boyle et. al.]

$$\Psi(L,z) = \frac{\Psi_0(z)}{\left(\frac{L}{L_\star(z)}\right)^{\beta_l} + \left(\frac{L}{L_\star(z)}\right)^{\beta_h}}. \tag{2.1}$$

Here, $\Psi_0(z)$ is a redshift dependent overall normalization that has units of number density (per comoving volume) per luminosity and $L_\star(z)$ is referred to as the break luminosity; it parameterizes the kink observed in the QSO distributions and also depends on redshift (see figure 2 in [Pei]). $\beta_l$ and $\beta_h$ are, respectively, the low end (below the kink, towards low luminosities) and high end (above the kink, towards high luminosities) slopes of the QSO LF with $\beta_l < \beta_h$. These parameters may also depend on redshift, though a quick look through the literature [Comerford et. al.], indicates that most authors ignore such dependences, even at high redshifts. The literature also points to a fair degree of disagreement regarding the evolution of $L_\star$with redshift [Comerford et. al.].

The QSO parameters are generally obtained by fitting equation 2.1 to an observed QSO distribution. We shall take a different approach; our goal, in this thesis, is to constrain the $L_\star$ and $\beta_h$ QSO LF parameters at a variety of redshifts using gravitational lensing statistics. We will compare our constraints with the best-fit estimates of [Pei], [Madau et. al.], and [Wyithe and Loeb] as tabulated by [Comerford et. al.]. Figure 2-1 presents two representative normalized QSO LFs.

## 2.3 Cosmology

Perhaps it is now appropriate to discuss the assumptions regarding cosmology that will be made throughout this thesis. Some authors (notably [Mitchell]) have used QSO gravitational lensing statistics to constrain cosmology. While this is certainly an excellent use of QSO lensing statistics that offers independent verification of the values of cosmological parameters, the accurate results of WMAP set far stronger and more robust limits on cosmology. Thus, in this thesis, we will assume the most recent cosmological parameters, as observed by WMAP [Bennett et. al.] (see table 2.1) and will choose to concentrate our use of gravitational lensing statistics to better constrain the QSO LF.

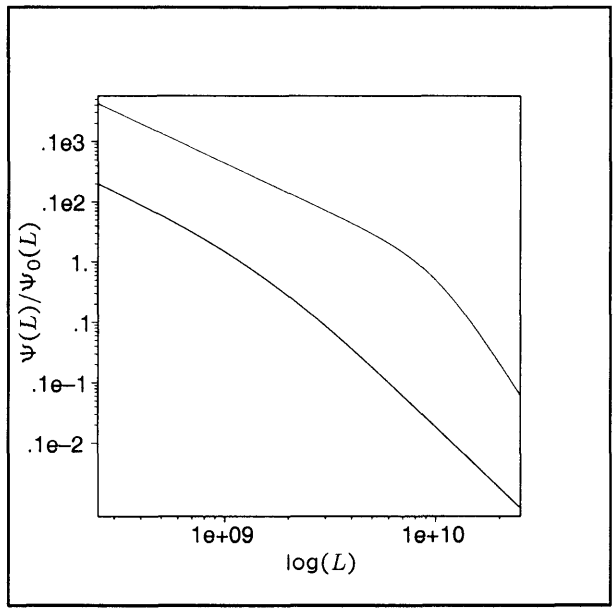


Figure 2-1: The QSO Luminosity Function for $\beta_h = 5.5$, $L_\star = 10^{10}$ (top, red) and $\beta_h = 3.4$, $L_\star = 10^{8.4}$ (bottom, blue). $\beta_l = 1.64$ for both QSO LFs.

| Parameter | Value | Assumed Value |
|---|---|---|
| $\Omega$– Total density | $1.02 \pm 0.02$ | 1 |
| $\Omega_\Lambda$– Dark energy density | $0.73 \pm 0.04$ | 0.7 |
| $\Omega_m$– Matter density | $0.27 \pm 0.04$ | 0.3 |
| $\Omega_k$– Curvature (effective) density | 0 | 0 |
| $h$ – Hubble constant | $0.71\,(+0.04, -0.03)$ | 0.71 |
| $\sigma_8$ – Fluctuation amplitude in $8h^{-1}$ Mpc Spheres | $0.84 \pm 0.04$ | 0.84 |

Table 2.1: Relevant WMAP Cosmological Parameters, the "Assumed Value" column presents the parameter values adopted throughout this thesis.

# Chapter 3

# Gravitational Lensing

## 3.1 Gravitational Lensing and QSOs

That gravity should affect the trajectories of photons is not a new idea. As early as 1783, by treating "corpuscles" of light as particles with velocity $c$, John Michell showed using Newton's theory of gravitation that a star could conceivably be sufficiently compact that no light corpuscles could escape to infinity. It would be impossible for observers to see such a "dark star" [Thorne]. Today we have Einstein's formulation of the General Theory of Relativity to guide our calculations and, while the modern day black hole is fundamentally different from Michell's "dark star," the essential idea behind both phenomena is the same: gravity bends light.

In 1919 the English physicist Sir Oliver Lodge proposed that light could be focused by massive bodies in a manner similar to the focusing effects of optical lenses. This phenomenon is called **gravitational lensing**, and the massive body that causes light to bend is called a **gravitational lens**. While all massive bodies can act as gravitational lenses, Fritz Zwicky proposed in 1937 that the majority of observable gravitational lenses would be galaxies [Carroll].

In order to observe a lensing event it is necessary to place the lens *between* a light source and the observer (in our case, the Earth); as QSOs are both far away and bright, one would expect the chances of observing a QSO lensed by a galaxy to exceed the chances of observing other lensed objects. For this reason, QSOs have become popular targets in the search for gravitational lensing phenomena.

In the next few sections in this chapter we will discuss the population of gravitational lenses that are most likely to lens QSOs. We will describe the distribution of these gravitational lenses and use these distributions, along with the QSO LF, to compute lensing probabilities. We will then show how the resulting probabilities can be applied to observational data to set constraints on the QSO LF. To proceed, however, we will first need to understand the conditions under which gravitational lensing occurs, and choose an analytical model to describe the lensing phenomenon quantitatively.

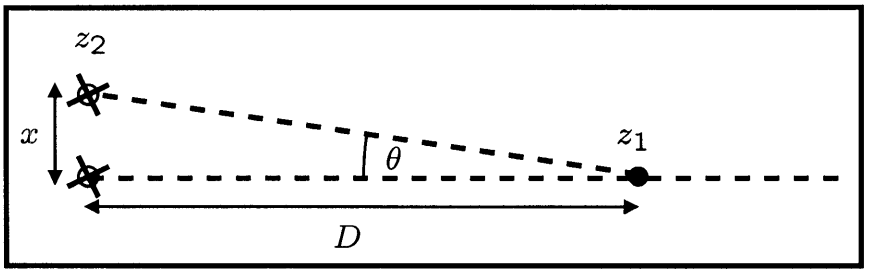


Figure 3-1: The definition of Angular-Diameter distance, $D$.

## 3.2 The Gravitational Lensing Setup

We will be studying systems involving cosmological distances but there are several different distance measures in cosmology [Hogg]. In geometries relevant to gravitational lensing systems, however, the most useful distance measure is the **angular-diameter distance**: Given two objects at the same redshift $z_2$ separated by a small angle $\theta$ on the sky as seen by an observer at $z_1$, we define the angular-diameter distance between the observer and the objects to be $D$ such that

$$x \approx D\theta \tag{3.1}$$

where $x$ is the physical separation between the two objects (see figure 3-1). To compute the angular-diameter distance between two objects at redshifts $z_1$and $z_2 > z_1$ we use

$$D = \frac{D_M(z_2) - D_M(z_1)}{1 + z_2} \tag{3.2}$$

where $D_M(z)$ is the **comoving distance** given by

$$D_M(z) = D_H \int_0^z \frac{dz'}{E(z')} \tag{3.3}$$

with $E(z)$ given by

$$E(z) = \sqrt{\Omega_M (1+z)^3 + \Omega_\Lambda} \tag{3.4}$$

and $D_H = 3000h^{-1}$Mpc is the **Hubble distance**. These equations are taken from [Hogg] and have been simplified by assuming that $\Omega_k = 0$.

Now, consider the setup in figure 3-2. The distances $D_l$, $D_s$, and $D_{ls}$ are the angular-diameter distances between the observer and the lens, the observer and the source, and the lens and the source respectively. The lensing equation [Narayan and Bartelmann], in terms of the parameters in figure 3-2, is simply

$$\beta = \theta - \alpha(\theta). \tag{3.5}$$

To compute $\alpha$ we need the surface-mass distribution $\Sigma$ of the lens (assumed to be thin and well

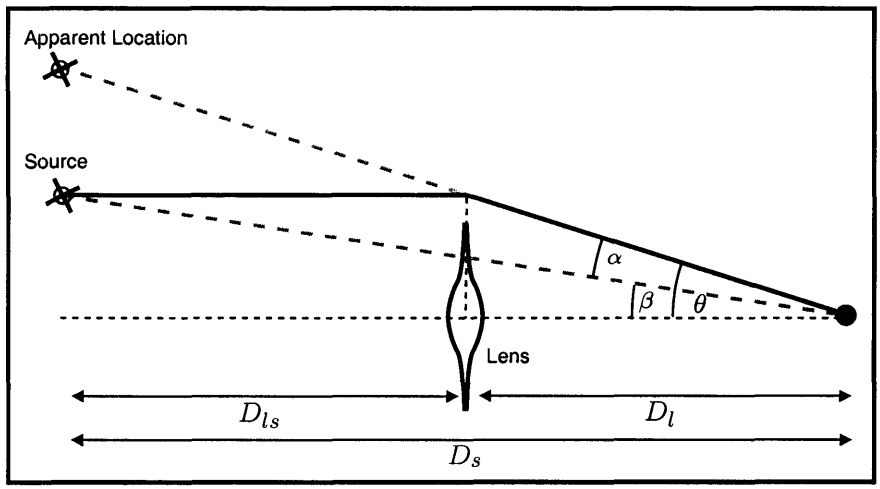


Figure 3-2: A typical gravitational lensing setup.

localized at a single redshift). If we assume a circularly symmetric mass distribution $\Sigma(r)$ then it can be shown that (see [Narayan and Bartelmann])

$$\alpha = \frac{D_{ls}}{D_l D_s} \frac{4G}{c^2 \theta} \left[ 2\pi \int_0^{D_l \theta} \Sigma(r)\, r dr \right]. \tag{3.6}$$

The integral is simply the lens mass enclosed by a circle of radius $D_l\theta$. To proceed we'll need to select an appropriate surface-mass distribution; the most commonly used distribution is the SIS model which we discuss next.

## 3.3 The SIS Model

Gravitational lenses bend light gravitationally, not optically. While the refractive properties of an optical lens depend on the material and geometry of the lens, the refractive properties of gravitational lenses depend on the distribution of mass in the lens plane. In statistical applications of lensing the most commonly used mass distribution is the **Singularly Isothermal Sphere** mass profile **(SIS)** as it yields simple, analytical, expressions that can be easily incorporated into probability computations while offering a good approximation to actual mass distributions observed in galaxies[1].

The SIS model is not parameterized in terms of the total galactic mass, as one might expect. Rather, since most of the mass in galactic halos is due to dark matter, which astronomers cannot observe directly, it is more convenient to use $\sigma$, the **velocity dispersion**. By definition, $\sigma$ is the one dimensional velocity dispersion of the stars in the galactic halo and, since the stars are influenced

[1][Li and Ostriker] disagree with this statement, they claim an NFW profile (first described by Navarro, Frenk & White) yields better fits to CDM simulation results.

gravitationally by both luminous *and* dark matter, $\sigma$ is a measure of the *total* mass density profile of the galaxy.

The SIS model treats the mass components in the galactic halo like ideal gas particles under the influence of their own spherically symmetric gravitational potential. A simple thermodynamic derivation, relating temperature to $\sigma$, (see [Narayan and Bartelmann]) then yields the SIS surface-mass density

$$\Sigma(r) = \frac{\sigma^2}{2G}\frac{1}{r}. \tag{3.7}$$

This mass distribution, when coupled with equation 3.6, yields

$$\alpha = 4\pi\frac{\sigma^2}{c^2}\frac{D_{ls}}{D_s} \equiv \theta_E \tag{3.8}$$

where $\theta_E$ is the **Einstein radius** of the lens. Note that $\theta_E$ is independent of $\beta$ but does depend on the distances (and therefore redshifts) of the lens and source. The lensing equation yields two solutions for $\theta$, the observed position of the source, if $\beta \leq \theta_E$ (see [Narayan and Bartelmann])

$$\theta_\pm = \beta \pm \theta_E. \tag{3.9}$$

In this regime the source is said to be **strongly lensed**: the lens both magnifies the source and produces multiple images. We can compute the separation $\Delta\theta$ and magnification $\mu_\pm$ [Narayan and Bartelmann] of these two images to find

$$\Delta\theta = \theta_+ - \theta_- = 2\theta_E \tag{3.10}$$

$$\mu_\pm = \frac{\theta_\pm}{\beta} = 1 \pm \frac{\theta_E}{\beta} \tag{3.11}$$

If $\beta > \theta_E$ then we only observe one image at a position $\theta_+$ with magnification $\mu_+$ and the source is said to be **weakly lensed**. Weak lensing is very hard to identify in observational data.

Combining these results we can compute the total magnification of the source

$$\mu(\beta) = \begin{cases} \mu_+ - \mu_- = \frac{2\theta_E}{\beta} & \text{for } \beta \leq \theta_E \\ 1 + \frac{\theta_E}{\beta} & \text{for } \beta > \theta_E \end{cases}. \tag{3.12}$$

The separation $\Delta\theta(\sigma)$ and magnification $\mu(\beta, \sigma)$ are the only two results we will need from the SIS model.

## 3.4 The Lens Population

We now have a gravitational lens model, the SIS model, determined completely by the redshifts of the source and lens and the velocity dispersion of the lens. The next step towards computing lensing probabilities involves understanding the distribution of the lens population. Ideally, we would like

to have at our disposal a **lens distribution**, $\phi$, of the form

$$n\left(\sigma', z\right) = \phi\left(\sigma', z\right) d\sigma \tag{3.13}$$

where $n$ is the number of lenses per comoving volume with velocity dispersion $\sigma' - \frac{d\sigma}{2} \leq \sigma \leq \sigma' + \frac{d\sigma}{2}$ at redshift $z$ for vanishingly small $d\sigma$.

A wide variety of objects might contribute to the lens distribution and examples vary from the extreme, such as cosmic string loops [A de Laix et. al.], to the ordinary late-type and early-type galaxies. In our discussion we shall follow [Mitchell] and focus on the dominant contributions made by early-type galaxies.

Most lensing statistics studies compute the lens distribution using an empirically determined early-type galaxy luminosity function combined with the Faber-Jackson relationship, which relates luminosity to the velocity dispersion $\sigma$. [Wyithe] points out, however, that this approach is unreliable as it ignores the intrinsic scatter in the Faber-Jackson relationship. Also, this approach generally ignores the evolution of the lensing population with redshift. We will present an alternative technique for obtaining $\phi$, based on [Mitchell], in the remainder of this section.

### 3.4.1 An Empirically Determined Local Lens Distribution

Recently, [Sheth et. al.] fit a modified version of the Press-Schechter ([PS]) mass function of dark matter haloes (first presented in [Sheth and Tormen]) to SDSS velocity dispersion data. They obtained an empirical fit to the distribution of $\sigma$ among 30,000 local ($z \approx 0$) early-type galaxies given analytically by

$$\phi_{ST}\left(\sigma\right) d\sigma = \phi_\star \left(\frac{\sigma}{\sigma_\star}\right)^\alpha \exp\left[-\left(\frac{\sigma}{\sigma_\star}\right)^\beta\right] \frac{\beta}{\Gamma\left(\alpha/\beta\right)} \frac{d\sigma}{\sigma} \tag{3.14}$$

with

$$\begin{aligned} \phi_\star &= (1.4 \pm 0.1) \times 10^{-3}\mathrm{Mpc}^{-1} \\ \sigma_\star &= 88.8 \pm 17.7\,\mathrm{km/sec} \\ \alpha &= 6.5 \pm 1.0 \\ \beta &= 1.93 \pm 0.22 \end{aligned}$$

The Sheth-Tormen distribution is only valid at $z = 0$ and one would expect that $\phi\left(\sigma, z\right)$ should evolve with redshift. This is due to the fact that dark matter haloes merge, and that the velocity dispersion of the resulting halo generally exceeds the velocity dispersions of either constituent halo. Thus, we would expect the number of high $\sigma$ haloes to decrease with redshift.

There is little empirical data describing the evolution of $\phi_{ST}\left(\sigma\right)$ with redshift. Several authors (e.g. [Wyithe]) assume a simple $\sigma$ independent redshift evolution of the lens distribution; [Mitchell], however, makes use of N-body galaxy formation simulations to produce a fully $\sigma$-dependent redshift

evolution. We shall do the same.

### 3.4.2 Simulated Redshift Dependence

Our approach will be to use N-body simulation results to compute the ratio

$$\phi_R(\sigma, z) = \frac{\phi_N(\sigma, z)}{\phi_N(\sigma, 0)} \tag{3.15}$$

where, $\phi_N$ is the lens distribution computed by the simulations; in this equation $N$ stands for N-body simulation while $R$ stands for ratio. We will then combine this *theoretical* result with Sheth's [Sheth et. al.] *empirical* $\phi_{ST}(\sigma, z = 0)$ lens distribution by defining

$$\phi(\sigma, z) \equiv \phi_{ST}(\sigma)\, \phi_R(\sigma, z)\,. \tag{3.16}$$

The N-body simulations used to compute $\phi_R(\sigma, z)$ were performed by [Jenkins]. Professor Jenkins has graciously provided us with software that returns the computer simulation results for different cosmological models. The software takes several input parameters which we briefly describe now.

The first set of parameters specify the simulated universe's cosmology; we set $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$. The software also requires a power spectrum. The power spectrum describes the initial distribution of perturbations in the universe; these perturbations eventually collapse into galaxies. Following [Li and Ostriker] we use the analytical power spectrum fit proposed by [Eisenstein and Hu], normalized by WMAP's value of the normalization constant: $\sigma_8 = 0.84$ (see table 2.1).

Professor Jenkin's software returns the number density (per comoving volume) of galaxies at a given redshift as a function of $M$, the mass of the galaxy haloes: $\phi_N(M, z)$. To proceed we must first convert this to a function of $\sigma$. [Mitchell] outlines the procedure that must be followed to relate $\sigma$ to $M$; it requires taking into account the fact that $\sigma(M)$ relies on the formation redshift of the halo ($z_f$, the epoch when the halo collapsed). The formation redshift is discussed in [Lacey and Cole] where the authors obtain (essentially) the probability distribution of $z_f$ as a function of $z$ and $M$. [Mitchell] simplifies the distribution by setting $z_f$ to be the mean $\langle z_f \rangle$ obtained from the [Lacey and Cole] distribution. [Mitchell] checks, and ensures, that ignoring the dispersion in the $z_f$ distribution does not significantly affect $\phi_R(\sigma, z)$. Thus, we can speak of a function $z_f(M, z)$ which allows us to define $\sigma(M, z_f)$. All the relevant equations are available in [Mitchell], we simply present the results of this computation in figure 3-3.

We can now rescale the $\phi_N(M, z)$ distribution to find

$$\phi_N(\sigma(M, z), z) = \phi_N(M, z) \left.\frac{dM}{d\sigma}\right|_{M,z}. \tag{3.17}$$

Professor Jenkin's software returns $\phi_N$ at discrete values of $M$ and $z$. Equation 3.15 then gives

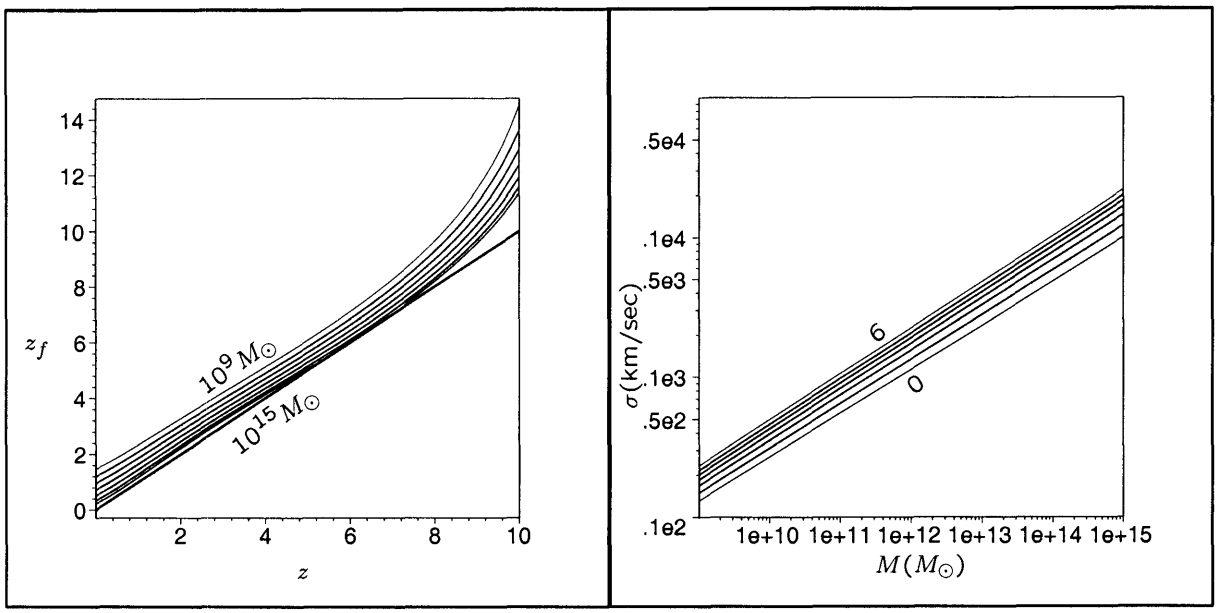


Figure 3-3: Left: A plot of $z_f$ vs $z$ for $M \in [10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14}, 10^{15}]$ $M_\odot$ (top to bottom). The thick black line is $z_f = z$. Right: A plot of $\sigma$ vs $M$ for $z \in [0, 1, 2, 3, 4, 5, 6]$ (bottom to top).

us $\phi_R(\sigma(M, z), z)$ at corresponding discrete values of $\sigma$. We will soon see, however, that we need to be capable of integrating over $\phi_R$ quickly. To achieve this we have put together a series of fitting functions, $\phi_R^{\text{fit}}$, that match the $\phi_R(\sigma, z)$ distribution well (see figure 3-4). We present the form of the fitting functions, with the fitting parameter values, in appendix A. The fits are valid for $20 \lesssim \sigma \lesssim 700$ and $0 < z < 6$.

The relative deviations between the fit and the data are presented in figure 3-4. The high deviation in the far right quadrant is due to the fact that $\phi_R$ is very close to zero in that region and even the slightest fit error is largely magnified. However, since $\phi_R$ is so close to zero in that quadrant, the contributions of the errors will be quite insignificant when we ultimately use $\phi_R$ in lensing probability computations.

Figure 3-5 presents the final $\phi$ distribution as obtained by combining $\phi_R$ with $\phi_{ST}$ (equation 3.16). Figure 3-6 is a two-dimensional presentation of $\phi$ for $0 \leq z \leq 6$; note the marked drop in the number of high $\sigma$ lenses with increasing $z$.

## 3.5 Multiple Imaging Cross Section

We have presented the SIS lens model and have obtained the redshift dependent lens distribution we were after. It is now time to combine both results to obtain **multiple imaging cross sections**. These cross sections are to be interpreted as the *geometric* probability that a point source at some redshift $z_s$ is multiply imaged by a foreground lens. To be precise, we seek a function $d\tau(z_s, z, \sigma)$

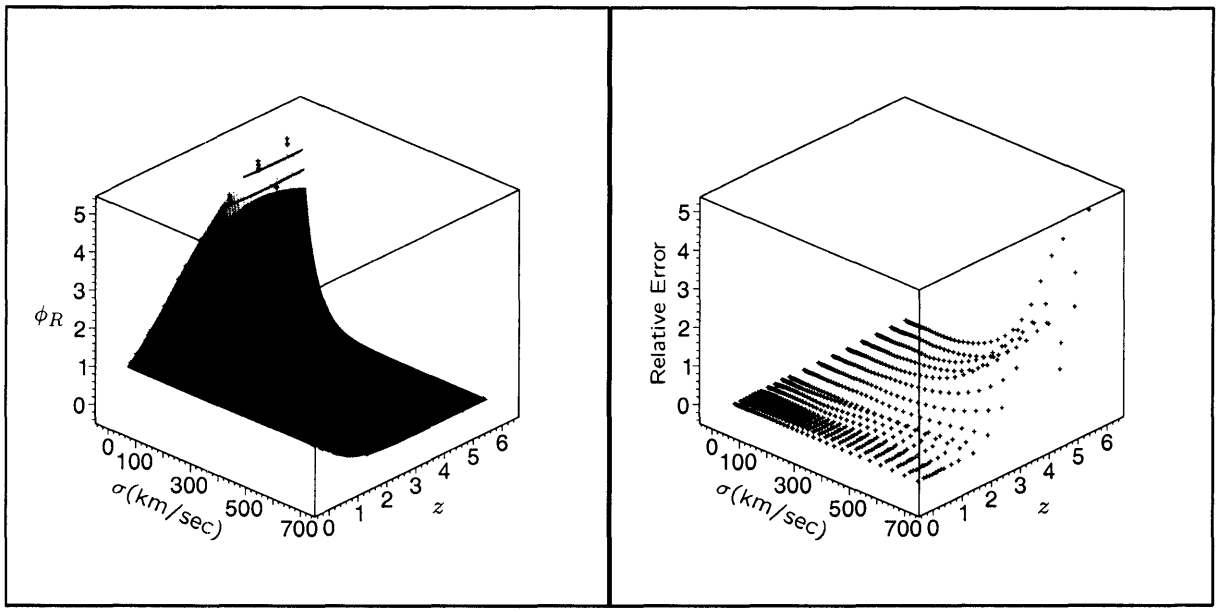


Figure 3-4: Left: $\phi_R$– the points are from the data generated by the simulation software, the contour lines are on the fitted mesh. Note that the visible points are slightly above the mesh, and that there are points slightly below the mesh that can't be seen. Right: The relative error $\left|\left(\phi_R^{\text{fit}} - \phi_R\right)/\phi_R\right|$.

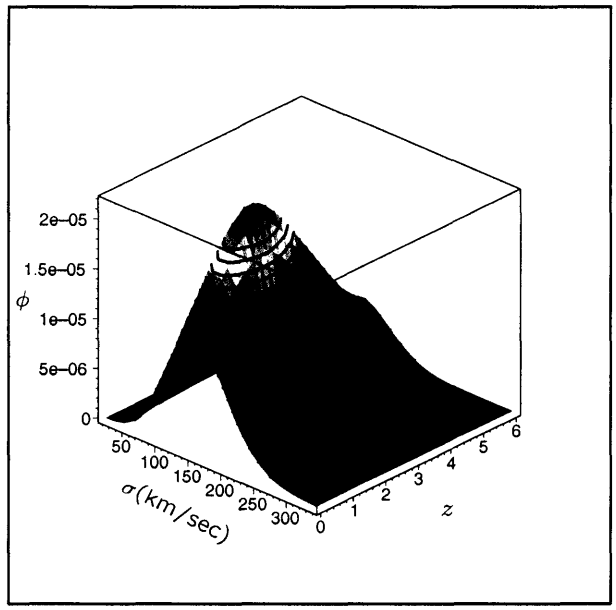


Figure 3-5: $\phi$ (number density per comoving volume ($\text{Mpc}^3$) per $\sigma$ interval) as defined in equation 3.16.

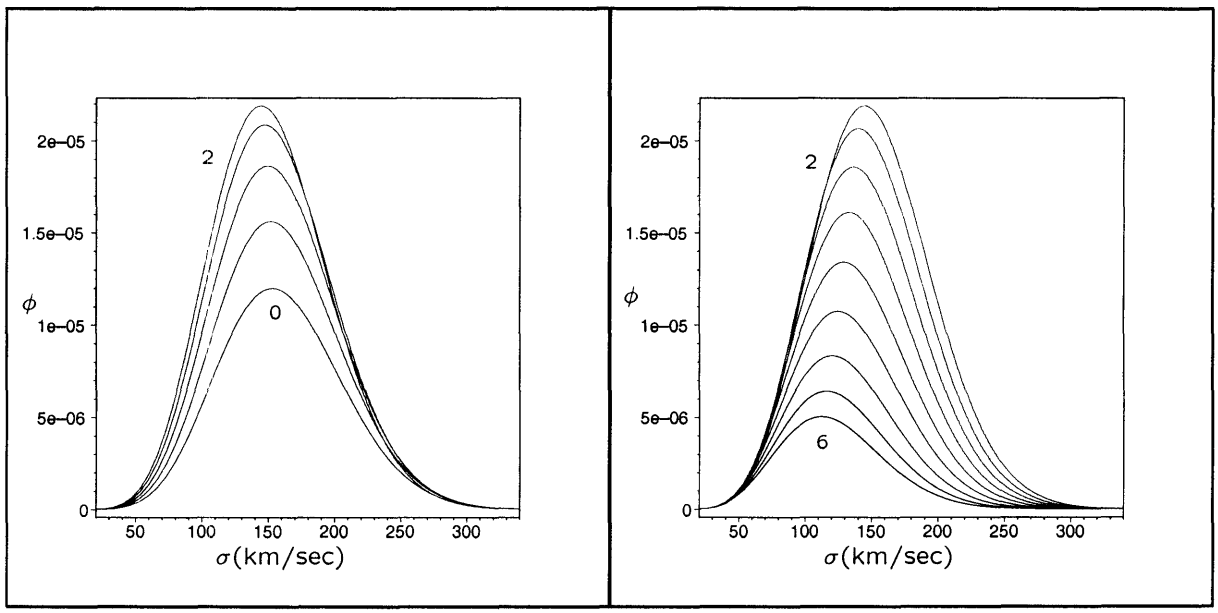


Figure 3-6: $\phi$ (number density per comoving volume ($\mathrm{Mpc}^3$) per $\sigma$ interval) presented in two dimensions. Left: $0 \leq z \leq 2$ bottom to top. Right: $2 \leq z \leq 6$ top to bottom.

such that

$$d\tau\left(z_s, z', \sigma'\right) dz d\sigma \tag{3.18}$$

is the probability that a source at redshift $z_s$ is multiply imaged by lenses living in the phase space: $z' - \frac{dz}{2} \leq z \leq z' + \frac{dz}{2}$ and $\sigma' - \frac{d\sigma}{2} \leq \sigma \leq \sigma' + \frac{d\sigma}{2}$. We can then define

$$\tau\left(z_s\right) \equiv \int_0^{z_s} \int_0^{\infty} d\tau\left(z_s, z, \sigma\right) d\sigma dz \tag{3.19}$$

to be the geometric probability that a source at redshift $z_s$ is multiply imaged. We use the term *geometric* probability to emphasize the fact that the probability does not depend on any properties of the source other than its position (i.e. redshift), we will see later that we can define true *a priori* lensing likelihoods that take into account the luminosity of QSO sources.

The SIS model tells us that the multiple imaging cross section of a lens at redshift $z$, with velocity dispersion $\sigma$, and lensing a source at redshift $z_s$ is given by

$$A\left(\sigma, z, z_s\right) = \pi \left(D_l \theta_E\right)^2 = 16\pi^3 \left(\frac{\sigma}{c}\right)^4 \left(\frac{D_l D_{ls}}{D_s}\right)^2. \tag{3.20}$$

Now, $\phi\left(\sigma, z\right) d\sigma$ is the comoving number density of lenses with multiple imaging cross sections $A\left(\sigma, z, z_s\right)$ in a dispersion velocity interval $d\sigma$. According to [Hogg] the incremental probability that

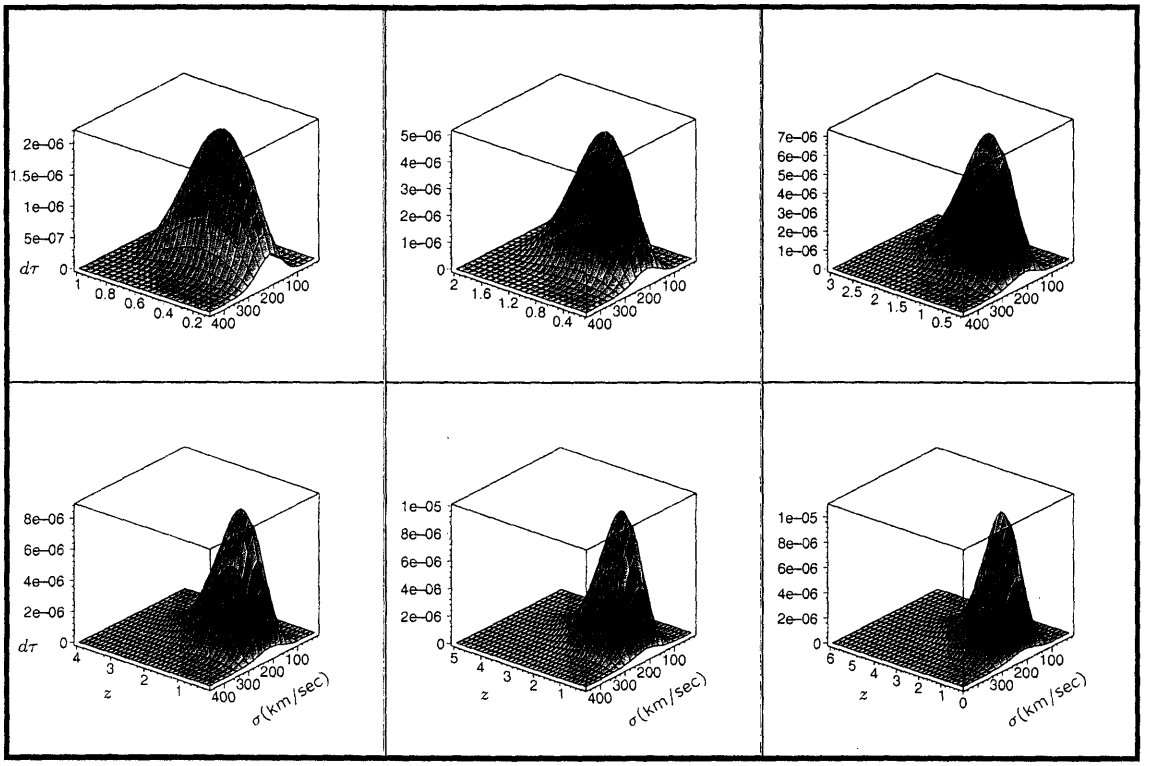


Figure 3-7: $d\tau$ vs $\sigma$(km/sec) and $z$ for (clockwise from topleft) $z_s \in [1, 2, 3, 4, 5, 6]$.

a line of sight will intersect one of these objects at redshift $z$ in a redshift interval $dz$ is

$$d\tau\left(z_s, z, \sigma\right) d\sigma dz = \phi\left(\sigma, z\right) A\left(\sigma, z, z_s\right) D_H \frac{\left(1+z\right)^2}{E\left(z\right)} d\sigma dz \tag{3.21}$$

(refer to section 3.2 for the definitions of $D_H$and $E\left(z\right)$). This is precisely the expression we are looking for.

Figure 3-7 presents six plots for $d\tau$ at source redshifts $z_s \in [1, 2, 3, 4, 5, 6]$. Note the behavior at large $z_s$ where there is little contribution to the multiple imaging cross section beyond $z \approx 3.5$, this is due primarily to the lensing geometry encapsulated in the $D\left(z, z_s\right)$ term, but is accentuated by the redshift evolution of the lens population.

Figure 3-8 presents two plots of $\tau\left(z_s\right)$. The blue (top) curve is $\tau\left(z_s\right)$ as defined in equation 3.19, complete with the simulated redshift evolution. The red (bottom) curve ignores redshift evolution completely. Interestingly, the redshift evolved lens population has a higher multiple imaging cross section. This can be understood by noting the formation of a peak in the lens distribution at $z \approx 2$ in figure 3-5 which amplifies the multiple imaging cross section contribution by lenses at $z \approx 2$. The effects of the peak are manifested in the steep slope of $\tau\left(z_s\right)$ at $z_s \approx 2$.

We now perform one more refinement to our analysis and present the multiple imaging cross

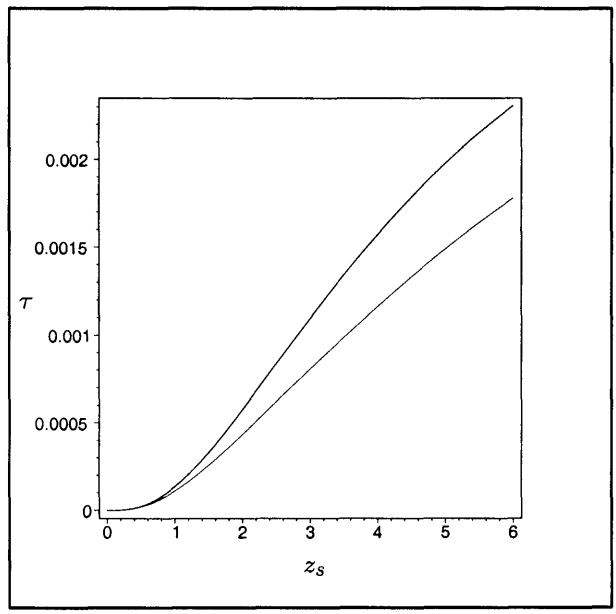


Figure 3-8: $\tau(z_s)$ as a function of $z_s$. The blue (top) curve takes into account the simulated redshift evolution. The red (bottom) curve does not.

section, $d\tau$, as a function of the image separation $\Delta\theta$ instead of $\sigma$. To do this we invert equation 3.10 to find

$$\sigma(\Delta\theta, z, z_s) = \sqrt{\frac{c^2}{8\pi}\frac{D_s}{D_{ls}}\Delta\theta}. \tag{3.22}$$

We then insert this expression into $d\tau(\sigma, z, z_s)$ and rescale the distribution by $\frac{d\Delta\theta}{d\sigma}$ (which we can compute analytically) to find:

$$d\tau(\Delta\theta, z, z_s) = d\tau(\sigma(\Delta\theta, z, z_s), z, z_s)\left.\frac{d\Delta\theta}{d\sigma}\right|_{\Delta\theta, z, z_s} \tag{3.23}$$

We can now use equation 3.23 to define $P(\Delta\theta, z_s)$ (as in [Barkana and Loeb]), the (cumulative) probability that a lensed source at redshift $z_s$ has an image separation $\geq \Delta\theta$

$$P(\Delta\theta, z_s) = \int_{\Delta\theta}^{\infty}\int_{0}^{z_s} d\tau\left(\Delta\theta', z, z_s\right) dz d\Delta\theta'. \tag{3.24}$$

We can also define the differential probability

$$\frac{dP}{d\Delta\theta} = \int_{0}^{z_s} d\tau\left(\Delta\theta', z, z_s\right) dz. \tag{3.25}$$

Figure 3-9 is a plot of $d\tau(\Delta\theta, z, z_s)$ for $z_s \in [1, 3, 6]$ and figure 3-10 is a plot of $P(\Delta\theta), z_s$ for $z_s \in [1, 3, 6]$.

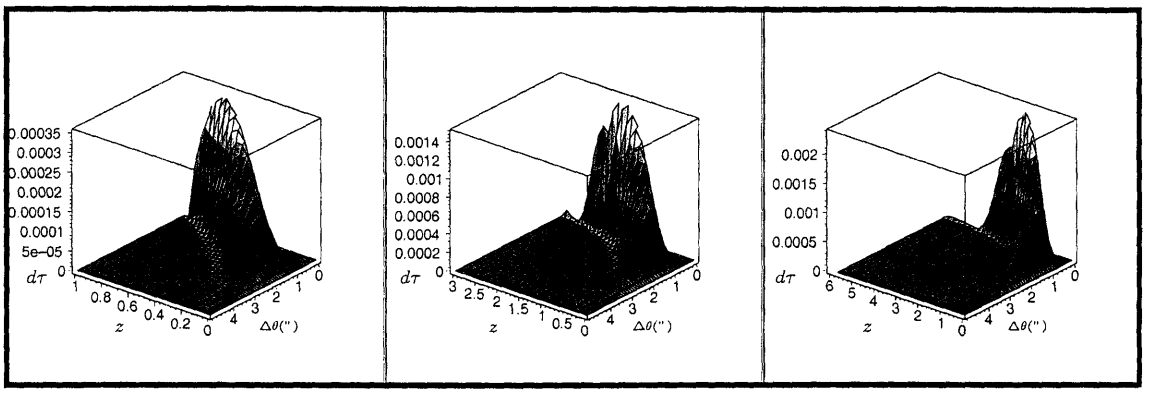


Figure 3-9: $d\tau\,(\Delta\theta, z, z_s)$ as a function of $\Delta\theta$ and $z$ for (from left to right) $z_s \in [1, 3, 6]$.

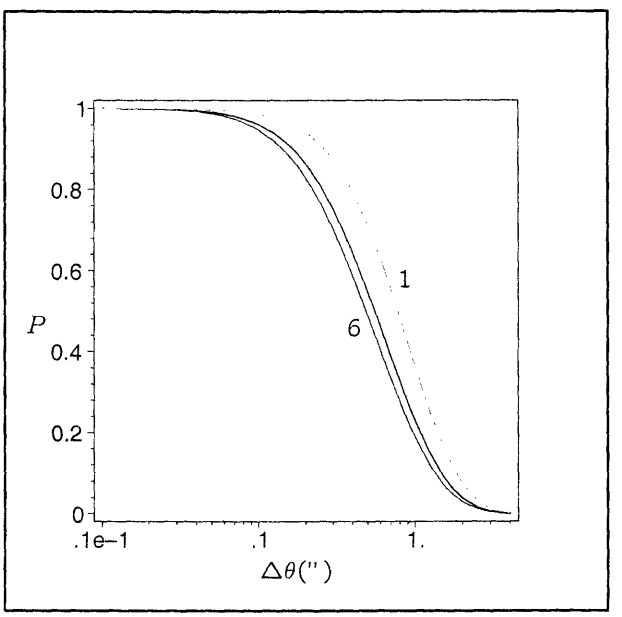


Figure 3-10: $P\,(\Delta\theta, z_s)$ for (from top to bottom) $z_s \in [1, 3, 6]$.

## 3.6 Lensing Likelihoods

We can now determine the probability that a QSO, with luminosity $L_{qso}$, is multiply imaged. This will involve taking into account the information provided by the QSO LF to weight the probability that the QSO is magnified by a factor $\mu$. For brevity, we shall refer to multiply imaged QSOs as **mQSOs**.

### 3.6.1 The Magnification Probability Distribution

To begin we will need to determine the **magnification probability distribution** for a given SIS lens. We first rewrite equation 3.12 in terms of $\mu$

$$\beta = \begin{cases} \frac{2\theta_E}{\mu} & \text{for } \mu \geq 2 \\ \frac{\theta_E}{\mu - 1} & \text{for } 1 + \frac{\theta_E}{\pi} < \mu < 2 \end{cases} . \tag{3.26}$$

Now, $\beta$ is an angle on the sky defined such that $\beta = 0$ at the center of the lens (see figure 3-2) and the probability distribution of $\beta$ is given by

$$\frac{dP}{d\beta} = \frac{1}{4\pi} 2\pi \sin\beta = \frac{1}{2}\sin\beta. \tag{3.27}$$

But using equation 3.26 we can compute the corresponding magnification probabilities as $\frac{dP}{d\mu} = \left|\frac{dP}{d\beta}\frac{d\beta}{d\mu}\right|$

$$\begin{array}{ll} \frac{dP_m}{d\mu} = \frac{\theta_E}{\mu^2}\sin\frac{2\theta_E}{\mu} & \text{for } \mu \geq 2 \\ \frac{dP_s}{d\mu} = \frac{\theta_E}{2(\mu-1)^2}\sin\left(\frac{\theta_E}{\mu-1}\right) & \text{for } 1 + \frac{\theta_E}{\pi} < \mu < 2 \end{array} . \tag{3.28}$$

Here $\frac{dP_m}{d\mu}$ and $\frac{dP_s}{d\mu}$ represent the multiple imaging and single imaging magnification probability distributions respectively. Plots of these distributions are presented in figure 3-11.

Now, note that the integral $\int_2^\infty \frac{dP_m}{d\mu} d\mu$ is the multiple imaging cross section of the lens and that it evaluates to the appropriate value

$$\int_2^\infty \frac{\theta_E}{\mu^2}\sin\frac{2\theta_E}{\mu} d\mu = -\frac{1}{2}\cos(\theta_E) + \frac{1}{2} \approx \frac{\pi\theta_E^2}{4\pi} \tag{3.29}$$

(assuming $\theta_E$ is small).

This result allows us to replace equation 3.20 with

$$\begin{aligned} A(\sigma, z, z_s) &= \pi D_l^2 \theta_E^2 = 4\pi D_l^2 \int_2^\infty \frac{\theta_E}{\mu^2}\sin\frac{2\theta_E}{\mu} d\mu \\ &= 16\pi^2 D_l^2 \int_2^\infty \frac{1}{\mu^2}\frac{\sigma^2}{c^2}\frac{D_{ls}}{D_s}\sin\left[\frac{8\pi}{\mu}\frac{\sigma^2}{c^2}\frac{D_{ls}}{D_s}\right] d\mu \end{aligned} \tag{3.30}$$

where we have substituted $\theta_E$ using equation 3.8. We can now substitute equation 3.30 into the

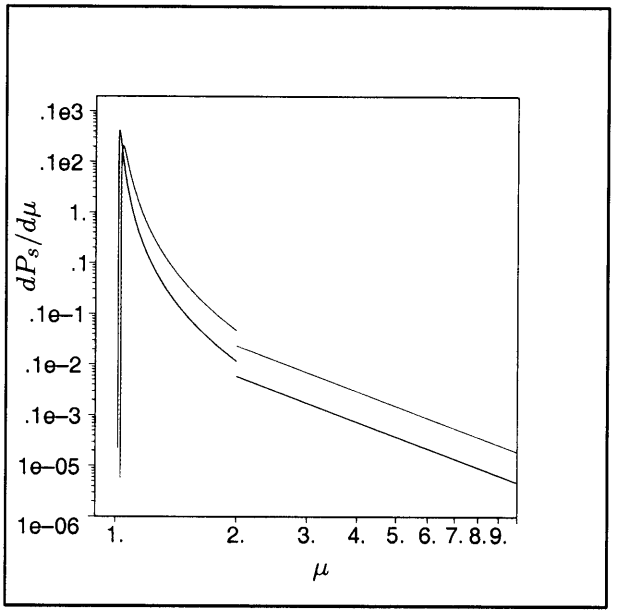


Figure 3-11: $\frac{dP_s}{d\mu}$ for $\theta_e \approx 0.1$ (red curve, top) and $\theta_e \approx 0.05$ (blue curve, bottom)

expression for $\tau(z_s)$ (equation 3.19) to obtain an integral expression for $\tau$ of the form

$$\begin{aligned}\tau(z_s) &= \int_0^{z_s}\int_0^{\infty} d\tau(z_s, z, \sigma)\, d\sigma dz = D_H \int_0^{z_s}\int_0^{\infty} \phi(\sigma, z)\, A(\sigma, z, z_s)\, \frac{(1+z)^2}{E(z)} d\sigma dz \\ &= 16\pi^2 D_H \int_0^{z_s}\int_0^{\infty}\int_2^{\infty} D_l^2 \phi(\sigma, z)\, \frac{(1+z)^2}{E(z)}\, \frac{1}{\mu^2}\, \frac{\sigma^2}{c^2}\, \frac{D_{ls}}{D_s} \sin\left[\frac{8\pi}{\mu}\, \frac{\sigma^2}{c^2}\, \frac{D_{ls}}{D_s}\right] d\mu d\sigma dz. \quad (3.31)\end{aligned}$$

This allows us to define $d\tau_\mu(z_s, z, \sigma, \mu)$ as a function of $\mu$

$$d\tau_\mu(\sigma, z, z_s, \mu) = 16\pi^2 D_H D_l^2 \phi(\sigma, z)\, \frac{(1+z)^2}{E(z)}\, \frac{1}{\mu^2}\, \frac{\sigma^2}{c^2}\, \frac{D_{ls}}{D_s} \sin\left[\frac{8\pi}{\mu}\, \frac{\sigma^2}{c^2}\, \frac{D_{ls}}{D_s}\right]. \quad (3.32)$$

### 3.6.2 The QSO LF as a Weighting Function

We now seek to incorporate the QSO LF with the multiple imaging cross sections. If we consider an mQSO with observed magnitude $i_{qso}$, and naively compute its intrinsic luminosity[2] $L_{qso}$, then the actual intrinsic luminosity of the QSO is $\frac{L_{qso}}{\mu}$ where $\mu$ is the magnification due to gravitational lensing.

Assuming our QSO LF describes the intrinsic distribution of QSO luminosities, we can describe the relative likelihoods of observing a QSO at redshift $z_s$ with different magnifications $\mu$. All we need to do is evaluate the luminosity function (equation 2.1) at $\Psi\left(\frac{L_{qso}}{\mu}, z_s\right)$ and the larger the result, the more likely the QSO is magnified by $\mu$. Note, however that the actual probability of finding a QSO with intrinsic luminosity $L$ within some $\Delta L$ interval is $\Psi(L, z_s)\, \Delta L$, if we replace $L$

[2]We will discuss how this is done in section 5.2.8.

with $L = \frac{L_{qso}}{\mu}$ we must replace $\Delta L$ with some $\frac{\Delta L_{qso}}{\mu}$ where $\Delta L_{qso}$ is a constant independent of $\mu$. This, then, allows us to write the magnification probability weight factor

$$\Psi\left(\frac{L_{qso}}{\mu}, z_s\right)\frac{1}{\mu}. \tag{3.33}$$

### 3.6.3 The Multiply Imaged Likelihood

We can use equation 3.33 to weight the multiple imaging cross section element, $d\tau_\mu$, and define the **multiply imaged likelihood**[3]

$$\mathcal{L}_m\left(l_{qso}, z_s\right) = \int_0^{z_s}\int_0^{\infty}\int_2^{\infty} d\tau_\mu\left(\sigma, z, z_s, \mu\right)\Psi\left(\frac{L_{qso}}{\mu}, z_s\right)\frac{1}{\mu}d\mu d\sigma dz. \tag{3.34}$$

This is a rather complicated triple integral and is not easily separable because the $\sin\left[\frac{8\pi}{\mu}\frac{\sigma^2}{c^2}\frac{D_{ls}}{D_s}\right]$ term in equation 3.32 combines all three integration variables in a nonlinear fashion. We can, however, assume that $\frac{8\pi}{\mu}\frac{\sigma^2}{c^2}\frac{D_{ls}}{D_s} = \frac{2\theta_E}{\mu} \leq \theta_E \ll 1$ and write $\sin\left[\frac{8\pi}{\mu}\frac{\sigma^2}{c^2}\frac{D_{ls}}{D_s}\right] \approx \frac{8\pi}{\mu}\frac{\sigma^2}{c^2}\frac{D_{ls}}{D_s}$. Equation 3.6.3 then separates to give

$$\begin{aligned}\mathcal{L}_m\left(l_{qso}, z_s\right) &\approx \int_0^{z_s}\int_0^{\infty} A\left(\sigma, z, z_s\right)\frac{\left(1+z\right)^2}{E\left(z\right)}\phi\left(\sigma, z\right)d\sigma dz\int_2^{\infty}\frac{8}{\mu^4}\Psi\left(\frac{L_{qso}}{\mu}, z_s\right)d\mu \\ &= \tau\left(z_s\right)\int_2^{\infty}\frac{8}{\mu^4}\Psi\left(\frac{L_{qso}}{\mu}, z_s\right)d\mu.\end{aligned} \tag{3.35}$$

We can reinterpret this result to shed light on the physics of the small angle approximation. Consider a single lens with multiple imaging cross section $\tau\left(z_s\right) = \frac{\pi\theta_E^2}{4\pi} = \frac{\theta_E^2}{4}$. The Einstein radius of this lens is

$$\theta_E \equiv 2\sqrt{\tau\left(z_s\right)} \tag{3.36}$$

and figure 3-8 assures us that $\theta_E \ll 1$ for all $z_s \leq 6$. The multiply imaged likelihood due to the lens is then given by

$$\begin{aligned}\mathcal{L}_m\left(L_{qso}, z_s\right) &= \int_2^{\infty}\frac{dP_m}{d\mu}\Psi\left(\frac{L_{qso}}{\mu}, z_s\right)\frac{1}{\mu}d\mu \approx \int_2^{\infty}\frac{2\theta_E^2}{\mu^4}\Psi\left(\frac{l}{\mu}, z_s\right)d\mu \\ &= \tau\left(z_s\right)\int_2^{\infty}\frac{8}{\mu^4}\Psi\left(\frac{L_{qso}}{\mu}, z_s\right)d\mu.\end{aligned} \tag{3.37}$$

This is precisely the result obtained in equation 3.35. Equations 3.36 and 3.37 offer an interesting physical interpretation for the approximation in equation 3.35; the individual lenses in the lens distribution can be brought together and treated like one large lens with a total multiple imaging cross section equal to the total cross section of all the individual lenses. In a sense, we are moving the lenses from their random locations on the sky together, combining them, and treating the resulting

[3]The multiply imaged likelihood is *not* a probability distribution. It is not normalized and must be combined with the corresponding singly imaged likelihood (section 3.6.3) to obtain multiple imaging probabilities (section 3.7).

lensing mass like one large SIS lens.

### 3.6.4 The Singly Imaged Likelihood

This interpretation of the multiply imaged likelihood, as codified by equation 3.37, motivates the definition of the **singly imaged likelihood**

$$\mathcal{L}_s\left(L_{qso}, z_s\right) = \int_{1+\frac{\theta_E}{\pi}}^{2} \frac{dP_s}{d\mu}\left(\theta_E, \mu\right) \Psi\left(\frac{L_{qso}}{\mu}, z_s\right) \frac{1}{\mu} d\mu \tag{3.38}$$

where $\frac{dP_s}{d\mu}\left(\theta_E, \mu\right)$ is defined in equation 3.36, and $\theta_E$ is given by equation 3.28.

The relationship between $\mathcal{L}_m$ and $\mathcal{L}_s$ can be clarified by taking a step back and removing the QSO LF weight function. Then, by definition

$$\int_{1+\frac{\theta_E}{\pi}}^{2} \frac{dP_s}{d\mu} d\mu + \int_{2}^{\infty} \frac{dP_m}{d\mu} d\mu = 1 \tag{3.39}$$

which, by equations 3.29 and 3.36, gives us

$$\int_{1+\frac{\theta_E}{\pi}}^{2} \frac{dP_s}{d\mu} d\mu = 1 - \tau\left(z_s\right). \tag{3.40}$$

Thus, equation 3.38 takes into account the magnification *outside* the multiple imaging area of the large lens constructed through equation 3.36. This interpretation leads to an interesting problem. Equation 3.38 only incorporates magnifications with

$$1 < 1 + \frac{\theta_E}{\pi} \leq \mu \leq 2 \tag{3.41}$$

implying that *every* source, (whether inside or outside the multiple imaging area), is magnified. Clearly, such a situation is unphysical – by conservation of photons alone we expect that the mean magnification, $\langle\mu\rangle$, over all sources should be unity. The SIS model does not allow for demagnification and, as such, is unrealistic far from the lens at low $\mu$, (though the approximation is quite valid for $\mu > 2$ inside the multiple imaging radius). Several authors (e.g. [Comerford et. al.] and [Wyithe and Loeb]) have attempted to bypass this problem by extending $\frac{dP_s}{d\mu}$ down to $\mu < 1$ and rescaling the resulting distribution to yield $\langle\mu\rangle = 1$ while conserving the relationship in equation 3.40. We have done the same and present the resulting renormalized $\frac{dP_s}{d\mu}$ as well as the (small) changes in the QSO LF constraints in appendix B.

## 3.7 Lensing Probabilities

The lensing likelihoods introduced in sections 3.6.3 and 3.6.4 are not, strictly speaking, probabilities. They have not been normalized, and doing so would be difficult as we do not understand the behavior of $\Psi_0\left(z_s\right)$ (see equation 2.1) with redshift very well. To get around this we now define actual lensing

probabilities. The first is the **multiply imaged probability**:

$$P_m = \frac{\mathcal{L}_m}{\mathcal{L}_m + \mathcal{L}_s} \tag{3.42}$$

and the second is the **singly imaged probability**:

$$P_s = \frac{\mathcal{L}_s}{\mathcal{L}_m + \mathcal{L}_s} \tag{3.43}$$

Thus, given a QSO with luminosity $L_{qso}$, the probability that it is an mQSO is $P_m$. Note that, while these probabilities depend on $z_s$, the redshift of the QSO, and $L_{qso}$, the observed luminosity of the QSO, there is also an implicit dependence on the parameters of the QSO LF. In particular, the probabilities may be sensitive to changes in $\beta_l$ (the low end QSO LF slope), $\beta_h$(the high end slope), and $L_\star$(the break luminosity) (see equation 2.1). By taking ratios of the lensing likelihoods, however, we have effectively canceled out the dependence on $\Psi_0(z_s)$.

Of course, equations 3.42 and 3.43 satisfy

$$P_m + P_s = 1 \tag{3.44}$$

## 3.8 Detection Probabilities

The multiply imaged probability, $P_m$, predicts whether a QSO with observed luminosity $L_{qso}$ is multiply imaged or not. It does *not*, however, take into account the fact that we cannot identify all mQSOs. It is possible, for example, that the multiple images are too close together and, therefore, unresolved even in high resolution imagery. It is also possible that the images are fairly well separated, but that the flux ratio between the two images is too far from unity to allow for a successful detection of both images.

These two factors: the flux ratio (which we will denote $f_r$ from now on) and image separation (which, following our current notation, we denote $\Delta\theta$), determine whether an mQSO can be resolved or not. More explicitly, if we consider the phase space $(f_r, \Delta\theta)$ we can define a subset of the phase space in which mQSOs will be successfully detected. We call this subset the **detection hull** and shall discuss its properties, from an observational point of view, in section 5.2.7. For the current theoretical discussion, however, let us assume we are given a detection hull for a given QSO. We must now modify our definition of $\mathcal{L}_m$ (in equation 3.34) to take the detection hull into account.

We begin by defining $f_r$ in terms of the SIS model parameters. Combining equations 3.11 and 3.26 we can write

$$\mu_\pm = 1 \pm \frac{\mu}{2} \tag{3.45}$$

and then define

$$f_r = -\frac{\mu_-}{\mu_+} = \frac{\mu - 2}{2 + \mu}. \tag{3.46}$$

Since $2 \leq \mu < \infty$ in the multiple imaging regime we have $0 \leq f_r \leq 1$. Equation 3.46 allows us to

define $\mu$ in terms of $f_r$

$$\mu = 2\frac{1+f_r}{1-f_r} \tag{3.47}$$

and this definition will allow us to convert the detection hull from $(f_r, \Delta\theta)$ phase space to $(\mu, \Delta\theta)$ phase space. With this in mind, let us reconsider equation 3.34

$$\mathcal{L}_m(z_s, l_{qso}) = \int_0^{z_s}\int_0^{\infty}\int_2^{\infty} d\tau_\mu(z_s, z, \sigma, \mu)\,\Psi\left(\frac{l_{qso}}{\mu}, z_s\right)\frac{1}{\mu}d\mu d\sigma dz.$$

Our first step will be to express the integral in terms of $\Delta\theta$ instead of $\sigma$. We've done this before, in equation 3-9. We define

$$d\tau_\mu(\Delta\theta, z, z_s, \mu) = d\tau_\mu(\sigma(\Delta\theta, z, z_s), z_s, z, \mu)\left.\frac{d\Delta\theta}{d\sigma}\right|_{\Delta\theta,z,z_s}. \tag{3.48}$$

Now, using the approximation made in equation 3.35 and the result obtained in equation 3-9 we can write $d\tau_\mu(\Delta\theta, z, z_s, \mu)$ as

$$\begin{aligned} d\tau_\mu(\Delta\theta, z, z_s, \mu) &\approx \left[A(\sigma(), z, z_s)\frac{(1+z)^2}{E(z)}\phi(\sigma(), z)\left.\frac{d\Delta\theta}{d\sigma}\right|_{\Delta\theta,z,z_s}\right]\left[\frac{8}{\mu^4}\Psi\left(\frac{l_{qso}}{\mu}, z_s\right)\right] \\ &= d\tau(\Delta\theta, z, z_s)\left[\frac{8}{\mu^4}\Psi\left(\frac{l_{qso}}{\mu}, z_s\right)\right]. \end{aligned} \tag{3.49}$$

We can now define $\mathcal{L}_{md}$, the likelihood that a QSO is multiply imaged *and detected as such* to be

$$\mathcal{L}_{md} = \iint_{DH}\int_0^{z_s} d\tau(\Delta\theta, z, z_s)\,dz\left[\frac{8}{\mu^4}\Psi\left(\frac{l_{qso}}{\mu}, z_s\right)\right]d\Delta\theta d\mu \tag{3.50}$$

where $DH$ is the detection hull delineated in $(\mu, \Delta\theta)$ phase space. Note, that the integral over redshift simply yields $\frac{dP}{d\Delta\theta}$ as defined in equation 3.25. Thus, we obtain a simplified version for our definition of $\mathcal{L}_{md}$

$$\mathcal{L}_{md} = \iint_{DH}\frac{8}{\mu^4}\frac{dP}{d\Delta\theta}(\Delta\theta, z_s)\,\Psi\left(\frac{l_{qso}}{\mu}, z_s\right)d\Delta\theta d\mu. \tag{3.51}$$

Using $\mathcal{L}_{md}$ we define the probability that an observed QSO is multiply imaged and detected as such to be

$$P_{md} = \frac{\mathcal{L}_{md}}{\mathcal{L}_m + \mathcal{L}_s}. \tag{3.52}$$

## 3.9 Constraining the QSO LF

We now have all the pieces necessary to constrain the QSO LF. Consider, now, that we have at our disposal a collection of QSOs with redshifts in some redshift range $z_1 \leq z_s \leq z_2$ and a variety of luminosities, $L_{qso}$. Let us also assume that we have determined, by analyzing high resolution imagery of the QSOs, which QSOs are mQSOs and which are not, and that we know what the

detection hull for each QSO is. In particular, let $n_{md}$ be the number of detected mQSOs. This information is sufficient to compute the likelihoods $\mathcal{L}_{md}$, $\mathcal{L}_m$, and $\mathcal{L}_s$ for each QSO, assuming some set of parameters for the QSO LF: $(\beta_l, \beta_h, L_\star)$.

We can then constrain the QSO LF (as discussed in [Comerford et. al.]) by computing the likelihoods for *each* QSO on a grid that discretizes the QSO LF parameters. We shall do this, in section 6.2, for $\beta_h$ and $L_\star$. For each point on the grid (that is, for each pair $(\beta_h, L_\star)$) we compute the mean number of expected detected mQSOs ($\langle n_{md} \rangle$) as follows

$$\langle n_{md} \rangle = \sum_{\text{QSOs}} P_{md}\left(z_s, l_{qso}, \beta_h, L_\star\right). \tag{3.53}$$

We then apply Poisson statistics to determine the probability of observing $n_{md}$ mQSOs given an expected mean of $\langle n_{md} \rangle$ detected mQSOs

$$P\left(n_{md}, \beta_h, L_\star\right) = \frac{\langle n_{md} \rangle^{n_{md}}}{n_{md}!} e^{-\langle n_{md} \rangle}. \tag{3.54}$$

This probability is a measure of the appropriateness of the QSO LF parameters. By plotting these values on a $(\beta_h, L_\star)$ grid, one can obtain probability contours to constrain the QSO LF parameters. This is our ultimate goal, and we now have all the theory in place and can proceed to the observational problem.

# Chapter 4

# QSO Observations

We have obtained 1073 high resolution QSO exposures using the MagIC imager on the Magellan telescope in Las Campanas, Chile. This chapter discusses the selection process by which the target QSOs were selected, the instrumentation used to obtain the imagery, and a brief description of the Coordinator, a software package we wrote to help organize the QSO imaging runs.

## 4.1 QSO Selection

The high resolution QSO imagery was obtained over two runs. The first run took place over two seasons, October 2001 and late December 2001 and obtained approximately 750 images. These were then analyzed using a significantly less sophisticated theoretical model than that discussed in chapter 3 and a less organized pipeline than that discussed in chapter 3.

The QSOs in the first run were predominantly selected from the SDSS database, with $\sim$ 50 QSOs coming from [Anderson et. al.]. Three distinct selection criteria were used to select the QSOs. First, a large portion of the SDSS QSOs were selected using magnification bias arguments based onSSSSSS [Kochanek]. These QSOs are unreasonably bright for their redshifts: a strong indication that they may be multiply imaged.

The second selection criteria was also used in conjunction with the SDSS database to select targets. The criteria requires that Magnesium II doublets be present in the SDSS QSO spectra. These lines are due to intervening galaxies and are almost always present in all currently known multiply imaged QSOs. We would expect targets with this spectral features to have a higher lensing probability than others.

The third selection criteria is the simplest. Motivated by figure 3-8, it simply requires that we select the highest redshift QSOs for observation. These were obtained from the [Anderson et. al.] paper which contained a plethora of high redshift QSOs.

It is not quantitatively clear what effect these selection criteria might have on the lensing statistics. It is possible, for example, to take into account the redshift of the Mg II absorption lines in the SDSS spectra to define new lensing likelihoods. This has not been done, however, and in the remainder of our analysis we treat the QSOs in our sample as if they were drawn from a random

pool.

After the first run QSOs were analyzed, the possibility of obtaining more data at another run in Chile came up. A proposal was submitted to obtain imaging for $\sim$ 500 more QSOs. The QSOs were selected from the SDSS database to fill in gaps in the redshift distribution of left by the first run (see figure 4-1). The QSOs were first selected to lie in the appropriate redshift range: $2.2 \leq z \leq 5$, and the appropriate location on the sky so as to be visible from Chile in January: declination $<$ 20°, 2h$<$ right ascension$<$14h. The resulting QSOs were checked against the QSOs obtained in the first run and duplicates were taken off the list. The multiple imaging probability, $P_m$, was then computed[1] for each remaining QSO using the SDSS database values for $z_s$ and $l_{qso}$, and the QSOs were sorted in decreasing order of $P_m$; this ordering effectively set the QSO observation priorities.

To optimize the order in which the QSOs would be observed we wrote the Coordinator, a simple database that allowed us to keep track of which QSOs were observed, to mark interesting QSOs for follow up imagery, and to make quick changes to the QSO observation ordering to accommodate for changing seeing conditions. A screenshot of the Coordinator at work is available in appendix C. The Coordinator allowed us to observe approximately 350 QSOs in about 10 hours of observation time.

Figure 4-1 presents the distribution of the redshifts and absolute magnitudes of the QSOs. The blue crosses are QSOs obtained in the first run, the green crosses are QSOs obtained in the second run. The red circles denote potential mQSOs as detected by our QSO analysis pipeline (chapter 3). The vertical lines delineate the redshift bins, each containing the same number of QSOs, which we have used to set constraints on the QSO LF.

## 4.2 Instrumentation

We obtained high resolution QSO imagery using the Baade Magellan telescope in Las Campanas, Chile. Baade is a 6.5-meter telescope complete with active optics, capable, on good nights, of attaining angular resolutions at the 0.3" $-$ 0.4" level. For imaging, we used MagIC, the Raymond and Beverly Sackler Magellan Instant Camera (MagIC), which has a 2048x2048 pixel CCD camera, at a pixel scale of 0.069"/pixel and a total field size of 2.36 square arcminutes. We obtained the images using the SDSS $i'$ band filter with 40 second exposures. The short exposure time was necessary to expedite the observation process and cover as many targets as possible. In fact, during the second run, our observation time was most strongly limited by the CCD readout time! More information about Magellan and MagIC can be found online at [Magellan].

Several calibration flats were taken to filter out CCD noise, and the pointing coordinates were modified to place the QSOs in a relatively "clean" region of the CCD. The high resolution images were stored using the FITS file format and transported by DVD to MIT where they now reside on several computer hard disks. The total, uncompressed, size of the images is at the 5 gigabyte level.

[1] $P_m$ was not computed using the theoretical model described in chapter 3 which was only developed after the data acquisition run in Chile. Instead, a simpler model was used that, nevertheless, yields results similar to those of the more complicated theory.

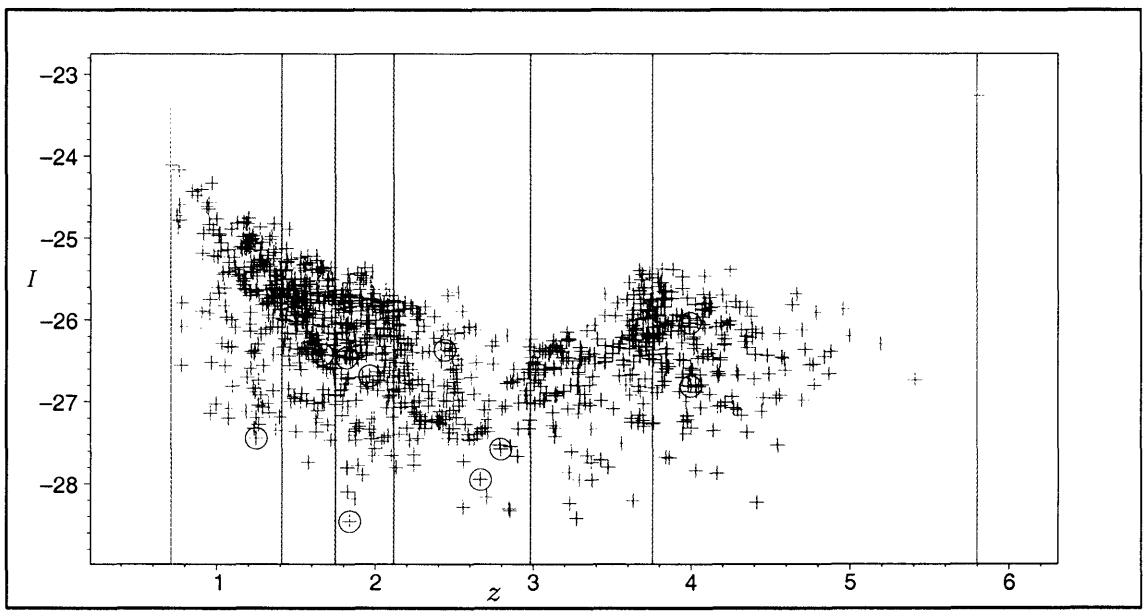


Figure 4-1: The distribution of QSOs in our sample, presented as observed absolute magnitude $I$ vs redshift $z$. The blue crosses are QSOs obtained in the first run. The green crosses were obtained in the second run. The red circles denote potential mQSOs and the vertical lines delineate the redshift bins used to constrain the QSO LF.

# Chapter 5

# The QSO Analysis Pipeline

To constrain the QSO LF we must analyze each QSO exposure to determine whether or not the QSO is an mQSO, and to quantify the shape of the QSO's detection hull. While the former question can be answered relatively easily by sifting through the data by eye, we desire a more rigorous quantitative approach that will also allow us to identify and describe the detection hull. In this chapter we will first discuss our overall approach to the problem, and then present a detailed description of the **QSO analysis pipeline** applied to each QSO exposure.

## 5.1 Quantifying Lensing Through $\chi^2$ Fitting

As we saw in section 3.3, lensing, as described by the SIS model, falls into two categories: weak and strong lensing. Weak lensing simply makes the QSO point source appear brighter or fainter and it is essentially impossible to determine whether or not a QSO is weakly lensed on the basis of a single, high resolution, image. Strong lensing, however, can be easily detected if the image separations and flux ratios are within the observing instruments resolving capabilities.

Strongly lensed mQSOs are morphologically different from simple point sources (we shall refer to the images of point sources as **point spread functions** or **PSFs** for short) and can usually be easily identified by eye. To identify these morphological anomalies by computer, however, we have used a $\chi^2$ based fitting technique.

The fitting procedure provides us with several benefits: we can automate the mQSO identification process, quantify the degree to which a QSO image is anomalous, and filter out the effects of poor and variable seeing from exposure to exposure. This last factor is especially important in our analysis as the data was obtained with short exposures and quick telescope slews. There are several images where the telescope slipped, or failed to settle completely, in which the QSOs appear slightly streaked or, in the most extreme cases multiply imaged.

Abstractly, each exposure has associated with it various smearing and smudging effects due to variable seeing. We assume that these effects are essentially constant across the exposure – a reasonable assumption given the area of the sky covered by each image (2.36 square arcminutes). Unfortunately, because of the variability in the smudging effects from frame to frame we cannot

simply assume an analytical form for the exposure PSFs and compare QSOs across exposures based on fits to a single analytical function. Instead, we define **control stars (cstars)** to be point sources in each exposure that define what is meant by a *normal* morphology. We then compare each QSO to the cstars in its exposure.

The selection of the cstars is important, and the resulting statistics are fairly sensitive to how this is carried out. Selecting a galaxy as a cstar, for example, would yield a misleadingly high $\chi^2$. Also, several exposures contain PSFs near CCD defects, or superimposed over cosmic rays; choosing such defective point sources as cstars would also overestimate $\chi^2$. For these reasons we have decided against utilizing an automated cstar selection system; we have attempted an implementation such a system, but found that the computer could not consistently select appropriate cstars.

Thus, while the $\chi^2$ fitting process can be easily automated, the selection of cstars cannot and we are faced with the task of manually identifying appropriate cstars for each exposure. This task, coupled with the problem of keeping the data organized and maintaining a decent level of quality control, motivated the development of the Analyzer, a database package that incorporates the fitting algorithms and allows the user to interact with the exposures and select cstars through a simple graphical user interface. Appendix C has several screenshots of the software in action. We shall refer to these screenshots as we describe the QSO analysis pipeline.

## 5.2 The QSO Analysis Pipeline

We begin our discussion of the QSO analysis pipeline by introducing some notation.

### 5.2.1 Notation

We will be discussing several algorithms that access pixels in the QSO exposures. We will denote the $i^{\text{th}}$ pixel in an exposure using the notation $p_i$. Alternatively, we can switch to Cartesian coordinates with the origin centered at the lower left corner of the image and denote pixels using the notation $p\,(x, y)$. When discussing pixels associated with different images we use the notation $\mathit{image}.p_i$ and $\mathit{image}.p\,(x, y)$ $(x, y \in \mathbb{Z})$ to clarify the source of each pixel.

Throughout the discussion of the pipeline we will make references to "the QSO," to be interpreted as the QSO of the exposure currently going through the pipeline. We will also refer to the selected cstars in the exposure by the names cs1 and cs2 (we will select at most two control stars per exposure). Associated with each object in the exposure (be it the QSO or cs1/cs2) are several properties such as the $x$ and $y$ coordinates of the object. When referring to the property of a specific object we will use the notation *object.property*. So, for example, the $x$ coordinate of cs1 is denoted cs1.$x$.

Finally, each exposure in the data set has its own unique ID. Exposures from the first run have IDs of the form xxxxxx.nnn where xxxxxx is the date the exposure was taken and nnn is an integer corresponding to the exposure's ID (these were assigned sequentially). Exposures from the second run have IDs of the form nnn where nnn is an integer corresponding to the QSO's $P_m$ sorted ranking.

### 5.2.2 Preparing the Exposure

Several (around 10) MagIC flat fields were taken during the observation runs to compensate for CCD defects and gradients. These flat fields were combined together to form a single calibration flat field image.

MagIC images are acquired by four amplifiers on one CCD and the resulting images must be carefully trimmed, with each CCD quadrant undergoing a bias transformation to account for the different amplifications characteristic of the different CCDs. Once trimmed the image is flatfielded using the expression

$$\text{exposure}.p_i = \text{exposure}.p_i \times \text{flatfield}.p_i$$

For the purposes of our fitting algorithm, we define the error associated with each pixel to simply be the Poisson counting error. Thus the error of the $i^{th}$ pixel is

$$p_i.\sigma = \sqrt{p_i}. \tag{5.1}$$

### 5.2.3 Target Identification

Once flatfielded, the exposure is ready for analysis. The first step involves the identification of the QSO and, ideally, of two cstars. These objects are identified manually by the user who, by clicking on the exposure, instructs the Analyzer to attempt a PSF fit centered on the clicked point. The PSF fitting function is a two dimensional Gaussian of the form

$$PSF\,(x,y) = a_0 \exp\left[-\frac{1}{2}\left(\left(\frac{\alpha}{a_4}\right)^2 + \left(\frac{\beta}{a_5}\right)^2\right)\right] + a_1 \tag{5.2}$$

where

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \cos(a_6) & \sin(a_6) \\ -\sin(a_6) & \cos(a_6) \end{bmatrix} \begin{bmatrix} x - a_2 \\ y - a_3 \end{bmatrix}. \tag{5.3}$$

Here, the $a_i$ are the fitting parameters described in table 5.1. The fitting function is compared to pixels in the exposure contained within a 60x60 pixel square. The square size is adjustable though this feature is rarely needed to allow the fits to converge. Figure C-2 is a screenshot of the Analyzer performing a PSF fit.

We define the **full width at half maximum** (**fwhm**) of these analytical PSFs in terms of the major axis fit parameter ($M$)

$$\text{fwhm} = \begin{cases} 2M\sqrt{2\ln 2} & \text{(pixels)} \\ 2M\sqrt{2\ln 2} \times 0.069 & (") \end{cases}. \tag{5.4}$$

The fwhm is a good measure of the exposure's seeing quality. We also define the PSF pixel function

$$\text{PSF}.p\,(x,y) \tag{5.5}$$

| Fitting Parameter | Description | Property Identifier |
|---|---|---|
| $a_0$ | The PSF peak | peak |
| $a_1$ | The sky background | sky |
| $a_2$ | $x$ coordinate of the PSF center | $x_c$ |
| $a_3$ | $y$ coordinate of the PSF center | $y_c$ |
| $a_4$ | Major axis (pixel units) | $M$ |
| $a_5$ | Minor axis (pixel units) | $m$ |
| $a_6$ | Orientation (radians) | $\theta$ |

Table 5.1: PSF fitting parameters (see equations 5.2 and 5.3)

such that $\mathrm{PSF}.p\,(0,0) = p\,(\mathrm{QSO}.x_c, \mathrm{QSO}.y_c)$.

Targets are selected using the PSF fitting procedure for two reasons. First, we shall soon see that the PSF fit parameters will play an integral role in running lensing simulations for each exposure. Second, the PSF fits guide the user through the target selection process. One can easily mistake extended sources for point sources, particularly in exposures with poor seeing and inappropriately set color scaling schemes; the resulting $\chi^2$ obtained are often misleadingly high. To avoid targeting such objects the fwhm, as computed by the fit parameters, can be quickly compared against the fwhm of other objects in the field – objects with relatively high fwhm are generally avoided.

Target identification is completed by cleaning up the targeted objects. Several exposures contain QSOs and cstars that are plagued by bad pixels and cosmic rays. These anomalous pixels can be masked out in the Analyzer and effectively ignored throughout the rest of the QSO analysis pipeline. Figure [coming soon!] is a screenshot of the masking function at work.

### 5.2.4 The $\chi^2$ Statistic

With targets selected we proceed with the $\chi^2$ fitting. We have not yet described the $\chi$ statistic in detail, but shall do so now.

The $\chi^2$ statistic is often used to compare theoretical models to data. We think of the data as a function $D\,(\{x_i\})$, where $\{x_i\}$ is an appropriate set of variables with the dimensionality of the data (in the case of the QSO, the data are represented by the QSO pixels: $\mathrm{QSO}.p_{(x,y)}$). Note that each data point has, associated with it, some known error which we call $\sigma_D\,(\{x_i\})$ (in the case of the QSO, this is $\mathrm{QSO}.p_{(x,y)}.\sigma$: see equation 5.1).

We can then define a theoretical fitting function $f\,(\{x_i\}, \{a_i\})$, where $\{x_i\}$ is the same set of coordinates used to describe the data and $\{a_i\}$ is a set of parameters that is, for our purposes, independent of $\{x_i\}$. We also have the error associated with the fitting function $\sigma_f\,(\{x_i\}, \{a_i\})$. We can then compute the sum

$$\chi^2 = \sum_{\text{data points}} \frac{\left[D\,(\{x_i\}) - f\,(\{x_i\}, \{a_i\})\right]^2}{\sigma_D^2\,(\{x_i\}) + \sigma_f^2\,(\{x_i\}, \{a_i\})}. \tag{5.6}$$

This defines the $\chi^2$ parameter which is a measure of the goodness of fit. Using $\chi^2$ alone, however, is misleading especially when comparing the quality of fits among different data sets. Consider two data sets describing the same system, one with more data points than the other, and both with comparable errors. Then by virtue of the fact that the denominator and numerator of the expression for $\chi^2$ are both positive and of, essentially, the same magnitude for similar data points in either data set, the $\chi^2$ for the data set with more data points will necessarily exceed the $\chi^2$ of the smaller data set. In fact, the $\chi^2$ will scale with the number of data points.

To correct for this we rescale the $\chi^2$ statistic by $\nu = n - n_c$ where $n$ is the number of data points and $n_c$ is the number of parameters (i.e. the number of elements in $\{a_i\}$) [Bevington], and $\nu$ is then the number of free parameters. This allows us to define the **reduced** $\chi^2$

$$\chi^2_\nu = \frac{\chi^2}{\nu} = \frac{\chi^2}{n - n_c}. \tag{5.7}$$

The reduced $\chi^2$ has a probability distribution associated with it that allows one to convert a given value of $\chi^2_\nu$ into the probability that the theoretical model being tested correctly describes the data set. We will not make use of this probability distribution but shall instead try to understand the behavior of the $\chi^2_\nu$ statistic within the context of the QSO fits. As a general guideline, in this context $\chi^2_\nu$ close to 1 ($\chi^2_\nu \approx 1 \pm 0.2$) is indicative of a good fit.

The $\chi^2$ statistic is often used hand in hand with optimization (in our context, minimization) algorithms to explore the $\{a_i\}$ parameter space in search of the "best fit" parameters. The algorithm we have employed in the Analyzer is the "amoeba" Levenberg-Marquardt algorithm as implemented in the C programming language by [Press et. al.]. Amoeba takes in initial guesses for the values of $\{a_i\}$ as well as initial search vectors in the $\{a_i\}$ parameter space (as many as the dimensionality of the space) along which it begins its search for the optimal parameters. The length of these vectors, in a sense, sets the maximum scale at which the algorithm traverses the $\{a_i\}$ space in search for the best fit parameters. The appropriate selection of these starting parameters and vectors is quite important in obtaining good fits; this generally tends to be more of an art than a science.

### 5.2.5 $\chi^2$ Fitting

We now describe the fitting procedure. Given a **target object** (let's use the QSO), and a **fitting object** (let's use cs1), we define the data function

$$D(x, y) = \mathrm{QSO}.p(x, y) \tag{5.8}$$

where $\mathrm{QSO}.p(x, y)$ is obtained from equation 5.5. We then define the fitting function

$$f(x, y, \{a_i\}) = a_0\, \mathrm{cs1}.p(x - a_2, y - a_3) + a_1. \tag{5.9}$$

The parameters $\{a_i\}$ simply adjust the cs1 pixels: $a_0$ rescales the pixels, $a_1$ shifts the background pixel level, and $a_2$ and $a_3$ shift the pixels about the cs1 fit centroid. The shift parameters, $a_2$ and

$a_3$, were initially allowed to take on non-integer values, and a sophisticated Gaussian averaging technique was used to incorporate sub-pixel shifts into the fitting process. This technique slowed down the fitting process significantly, however, and did not improve the $\chi^2_\nu$ results significantly. We have switched to a simpler model in which $a_2$ and $a_3$ are forced to be integers. This does not increase the resulting $\chi^2_\nu$ significantly at all but does yield a speed up factor of $\sim 30$.

The fitting object is compared to the target object by computing $\chi^2_\nu$ over a 60x60 pixel square. This square size is adjustable, though this feature has only been used in rare cases where severe sky gradients or discontinuities across different CCD quadrants impede the convergence of the amoeba algorithm. Limiting the comparison to squares of this size means that the algorithm is only sensitive to image separations with $\Delta\theta \lesssim 30\sqrt{2} \times 0.069 \approx 3$". A quick look at figure 3-10 indicates that we should only miss about 1 out of every 100 mQSOs with this square size.

With the definitions for $D$ and $f$, and an understanding of the implications of the fitting square size, we can easily compute and minimize $\chi^2_\nu$ using the amoeba algorithm. For exposures with only one cstar we are forced to use the $\chi^2_\nu$ of just one fit to test whether or not the QSO is multiply imaged. This can be problematic, particularly in exposures where the singly cstar is near a strong sky gradient, or in exposures where the QSO and cstar signal strengths are not well matched. For exposures with two cstars we compute $\chi^2_{\nu:1Q}$, the $\chi^2_\nu$ obtained by fitting cs1 to the QSO, and $\chi^2_{\nu:2Q}$, the $\chi^2_\nu$ obtained by fitting cs2 to the QSO. We then compute $\chi^2_{\nu:12}$, the $\chi^2_\nu$ obtained by fitting cs1 to cs2: this allows us to locate anomalous cstars – if the cstars are not well matched then at least one of the cstars is not representative of what "morphologically normal" means for the current exposure.

Among the noisier exposures, and the exposures with particularly poor seeing, the $\chi^2_\nu$ tend to be generally higher. We attempt to normalize against these effects by defining a new goodness of fit measurement, applicable only to the exposures with two cstars. We define $\chi^2_r$ to be the ratio

$$\chi^2_r = \frac{\chi^2_{\nu:1Q} + \chi^2_{\nu:2Q}}{2\chi^2_{\nu:12}} \tag{5.10}$$

$\chi^2_r$ is generally more robust than $\chi^2_{\nu:1Q}$. Since all QSO exposures, regardless of the number of cstars, go through the analysis pipeline, however, we shall use the notation $\chi^2_*$ to refer to the $\chi^2$ measure for the current exposure. Thus, for exposures with two cstars $\chi^2_* = \chi^2_r$, while for exposures with one cstar we set $\chi^2_* = \chi^2_{\nu:1Q}$.

It would seem that we are now equipped with a robust lens finding algorithm. Unfortunately, things are not so simple and we are faced with two remaining issues that must be dealt with. We describe these issues, and their solutions, now.

#### 5.2.5.1 Misleading Amoeba

The Analyzer allows us to visualize the residuals of the fits (the difference $f - D$ at each pixel) and a quick look through the first fitting run showed that the amoeba algorithm failed to successfully match the target and fitting objects consistently. As a result, several fits suffered from unreasonably high $\chi^2_\nu$; apparently, the parameter space $\{a_i\}$ being explored is rather complicated and filled with

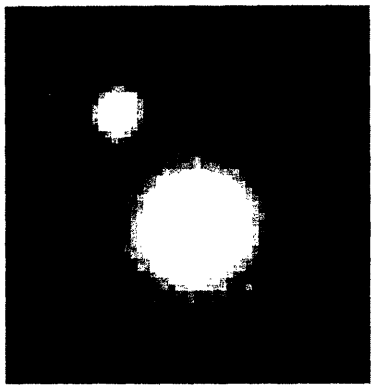

Figure 5-1: QSO Exposure 011011.243. The flux ratio between the two PSFs is 27!

regions where the amoeba algorithm is tricked into premature convergence. We approach this problem with a two-pronged solution.

On the one hand, we pick more sophisticated initial guesses for the $\{a_i\}$ parameters by refining both the initial parameters as well as expanding the initial search vectors to search more of the parameter space (thereby, hopefully, avoiding false convergences by "looking over" $\chi^2_\nu$ valleys).

Unfortunately, picking the initial values with finesse, however, only improves the $\chi^2_\nu$ results mildly. To truly force the system to converge to lower $\chi^2_\nu$ we have devised a brute force scheme. We not only expand the initial search vectors, but also allow them to be selected randomly (admittedly, there is some degree of additional finesse associated with the randomization of the vectors, but that is tangential to our discussion). We then run amoeba on the target and fitter objects several times, keeping only the fit parameters that result in the lowest $\chi^2_\nu$. In a sense, we send amoeba looking in different random directions several times and choose the "winning" fit on the basis of $\chi^2_\nu$. When the brute force method was initially implemented we were still using the sub-pixel shifts and could only repeat the fits 2-3 times. When we switched to the integer pixel shifts we began repeating the fits ~10-15 times and saw improvements in the $\chi^2_\nu$ convergence.

#### 5.2.5.2 Signal to Noise Issues

Purely by chance, a significant issue with the $\chi^2_\nu$ process we've described, was discovered. While going through a subset of the dataset by hand we noticed that the QSO in exposure 011011.243 had a nearby partner (figure 5-1) – an event that would ordinarily yield a large $\chi^2_\nu$ signal – but that the associated $\chi^2_\nu$ were well within normal values. Further investigation shows that the brighter image is 27 times brighter than the fainter one, 90 times brighter than its sky background, and 8 times brighter than the brightest cstar in the image (cs1)!

When the cstar fitting objects are rescaled to match the bright QSO image the adjacent sky is also rescaled. This, unfortunately, scales the sky error up by the same factor (8 in the particular case of 011011.243) and faint objects (with a signal, in this case, fainter than 8 times the sky) near the QSO become shrouded in the scaled sky noise (in terms of $\chi^2_\nu$, the denominator in equation 5.6

becomes unreasonably large). This problem is quite serious, as exemplified by 011011.243, and an effective solution is essential for the success of the $\chi^2_\nu$ approach.

After experimenting with several ideas, we opted for the following solution (which is only made possible by the speed increase obtained in switching to integer pixel shifts). We replace the fitting function in equation 5.9 with

$$f(x,y,\{a_i\}) = \begin{cases} a_0 \mathrm{cs1}.p(x-a_2, y-a_3) + a_1 & \text{If } (x-a_2, y-a_3) \in \Lambda \\ \mathrm{cs1}.p(x-a_2, y-a_3) + a_5 & \text{If } (x-a_2, y-a_3) \notin \Lambda \end{cases} . \tag{5.11}$$

where $\Lambda$ is the region in which cs1's fitted PSF (equation 5.2) exceeds $\frac{\sqrt{\mathrm{cs1}.sky}}{1000} + \mathrm{cs1}.sky$. Thus, the new fitting function is identical to the original fitting function where the signal due to cs1 exceeds the sky noise $\sqrt{\mathrm{cs1}.sky}$ (the factor of $\frac{1}{1000}$ expands this region slightly to take into account the fact that the PSF fits may be imperfect). In this region, rescaling predominantly affects the pixels containing data dominated by cs1 signal and not sky signal. Outside $\Lambda$, where the signal is dominated by sky, we avoid noisy rescaling completely by introducing a new parameter, $a_5$, which simply shifts the sky background level. Figure 5-2 illustrates the differences between the new fitting function (equation 5.11) and the old one (equation 5.9).

### 5.2.6 Non Lensed Simulations

Our knowledge of a QSO's $\chi^2_*$ is not enough to determine whether or not the QSO is lensed: we need to understand the distribution of $\chi^2_*$ to tackle this problem. Unfortunately, the distribution depends heavily on the conditions of each exposure: the seeing, as characterized by the fwhm of the cstars, and the flux ratios between the QSO and the cstars, all affect the $\chi^2_*$ distribution.

Our approach, then, is to generate the non-lensed $\chi^2_*$ distribution for each exposure. We do this by generating dummy PSFs (using equation 5.2) and running the amoeba fitting algorithm on the resulting dummy images. We repeat this 10 times, randomly perturbing the positions of the PSF centroids and regenerating the signal noise of the simulated image each time. These perturbations lead to a spread in the $\chi^2_*$ distribution which allows us to characterize the mean and dispersion of the non-lensed $\chi^2_*$ for the exposure.

To do this well we must carefully select the PSF parameters to match the actual QSO and cstar object parameters of the exposure. Since the cstars are chosen because of their morphological regularity we can generally trust all the fit parameters obtained in section 5.2.3, during the target selection phase, as adequate representatives of the cstar morphology. This gives us all the parameters necessary to generate dummy cstar PSFs. To generate the QSO PSFs we must be more careful. Since lensed QSOs are, by definition, morphological deviants we cannot trust the fit parameters obtained in section 5.2.3 to be good representatives of a *non-lensed* QSO morphology. Instead, we combine parameters from the QSO target selection fits with cstar target selection parameters.

To choose the appropriate combination of parameters we make the following argument: QSO PSFs carry information that comes from two sources. They carry intrinsic information about the

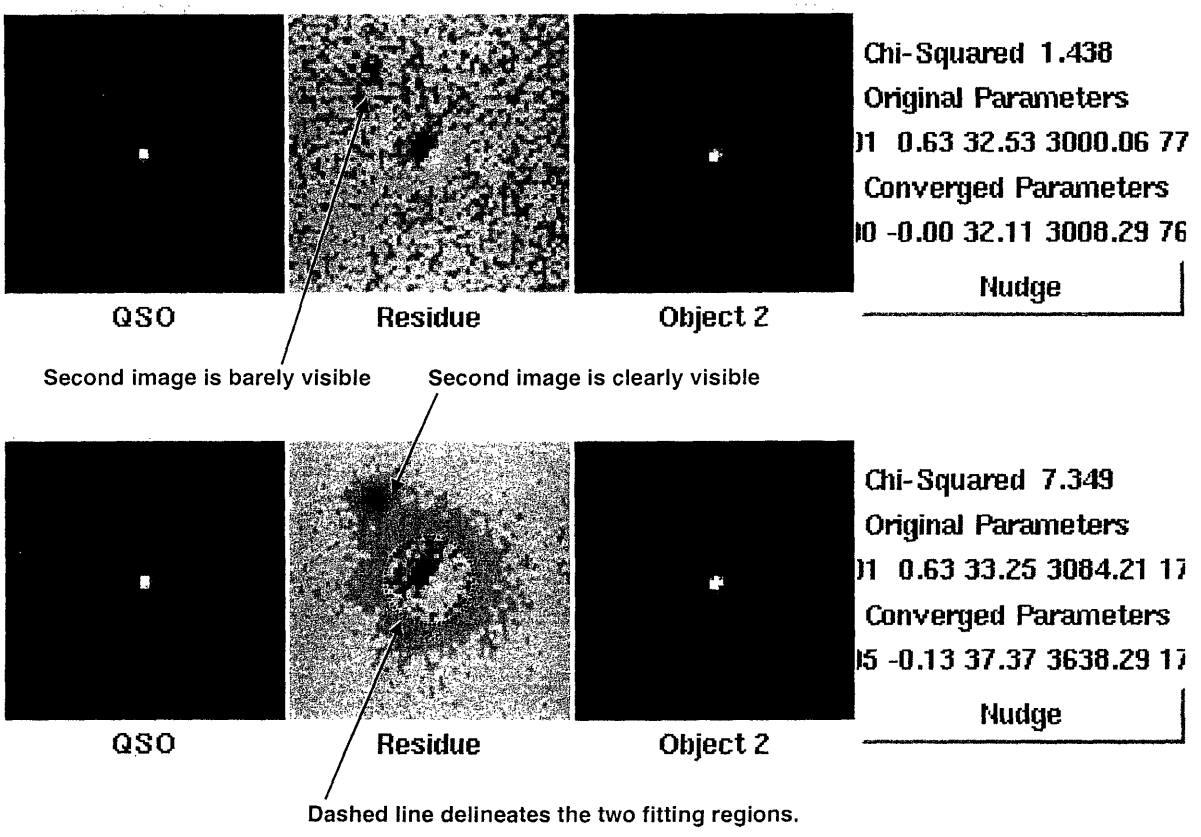


Figure 5-2: Two screenshots illustrating the difference between fit equation 5.9 (top) and equation 5.11 (bottom). In the top image, noisy sky rescaling washes out the fainter images residue (center image). In the lower image the dashed grey line delineates the boundary between $\Lambda$ and the rest of the frame. The fainter image's residue is now clearly visible and the $\chi^2_\nu$ is much higher. Finally, notice that the faint image is not visible in the QSO image (leftmost images) – this is because of the color scaling and the extreme flux ratio between both images.

QSO and sky quantified by the peak and sky parameters in table 5.1. These properties are (essentially) independent of the seeing and are unrelated to the exposure's intrinsic morphological anomalies. Rather, these intrinsic anomalies are encoded in the $M$, $m$, and $\theta$ parameters of the PSF fits and the $\chi^2_*$ fits are very sensitive to these parameters. Since the cstars are used to define morphological normality in the exposure we borrow the $M$, $m$, and $\theta$ parameters from the cstar target selection fit and combine them with the peak and sky parameters, obtained from the QSO target selection fit, to generate dummy QSO PSFs.

To determine how best to model the $\chi^2_*$ distribution analytically we generated several high-resolution distributions by carrying out the simulated fits 400 times for 10 randomly selected exposures. We then generated normalized $\chi^2_r$ histograms for each of the exposures (all of which were chosen to contain 2 cstars), and normalized $\chi^2_{\nu:1Q}$ histograms using the cs1 of each exposure. We experimented with a variety of analytical functions in search for good fits to the distributions and, after some trial and error, chose the Gumbel distribution [McLaughlin] – a distribution commonly used by economists – because it matches the high $\chi^2_\nu$ slope of the distributions well (see figure 5-3 for an example). The Gumbel distribution is given by

$$\frac{1}{B}\exp\left(\frac{A-y}{B}\right)\exp\left(-\exp\left(\frac{A-y}{B}\right)\right) \tag{5.12}$$

with mean and variance given by

$$\text{Mean} = A + \gamma B \tag{5.13}$$

$$\text{Variance} = \frac{1}{6}(\pi B)^2 \tag{5.14}$$

where $\gamma$ is the Euler gamma constant $\gamma = 0.57721566$. Figure 5-3 is obtained, *not* by fitting the parameters $A$ and $B$ to the observed histogram, but by computing $A$ and $B$ from equations 5.13 and 5.14. Note that the high $\chi^2_r$ end of the distribution fits well and that the fits are good for both $\chi^2_r$ and $\chi^2_{\nu:1Q}$ distributions.

We apply these results to the entire data sample by computing 10 $\chi^2_\nu$ fits for simulated QSO and cstar PSFs. We then compute the mean and variance of the 10 $\chi^2_\nu$ values and estimate the $\chi^2_\nu$ distribution using the Gumbel distribution (equation 5.12) and equations 5.13 and 5.14.

We can now compare the exposure's actual $\chi^2_\nu$ results (be they $\chi^2_r$ or $\chi^2_{\nu:1Q}$) against the Gumbel distribution to determine whether or not the QSO is multiply imaged. Though the Gumbel distribution allows us to compute the *probability* that the QSO is multiply imaged we do not do this but, instead, opt for a clearcut definition by setting a $\chi^2_\nu$ cutoff ($\chi^2_c$) at the 0.1% level. With this cutoff we expect no more than $\sim 1$ out of every 1000 QSOs to register as a false positive mQSO.

### 5.2.7 Detection Hull

The nonlensed distributions tell us whether the QSO is lensed or not. To constrain the QSO LF, however, we also need the QSO's detection hull. Like the $\chi^2_\nu$ distribution the detection hull

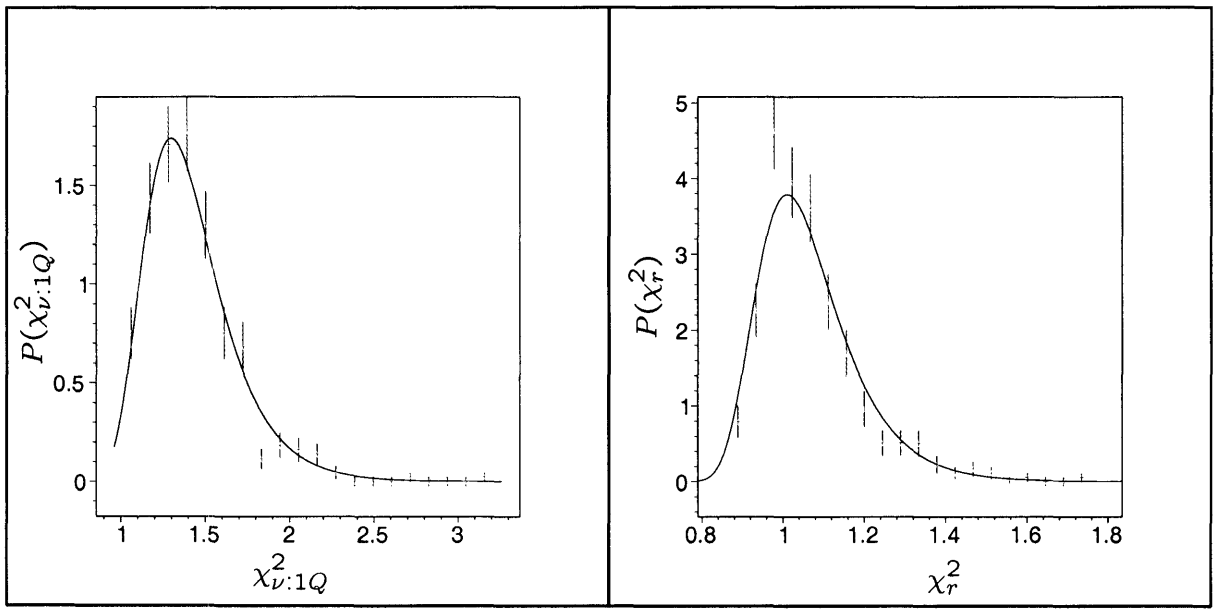


Figure 5-3: The Gumbel distribution overlaid on the $\chi^2_{\nu:1Q}$ distribution (left) and the $\chi^2_\nu$ distribution (right) for two randomly chosen exposures.

depends on many different properties of the exposure and its objects. Thus, rather than developing a general detection hull that applies to all exposures, we compute the detection hull individually for each exposure.

We do this, again, through simulation. We generate cstar fitter PSFs just as in section 5.2.6, and QSO PSFs with parameters chosen as described in section 5.2.6. We add, however, to the QSO image a second PSF separated from the core image by some $\Delta\theta$ with flux ratio $f_r$. The second image has the same $m$, $M$, and $\theta$ as the primary image, but has an adjusted peak to take the $f_r$ into account. Also, the line connecting the first and second image is at a random angle relative to the horizontal.

We run the amoeba algorithm on these setups for several values of $(f_r, \Delta\theta)$ and generate the function

$$\chi^2_* (f_r, \Delta\theta) . \tag{5.15}$$

The $(f_r, \Delta\theta)$ are chosen on a grid extending from $0 \leq f_r \leq 1$ and $0 \leq \Delta\theta \leq 1.4$ with $\Delta f_r = 0.2$ and $\Delta(\Delta\theta) = 0.35$. We then fit an analytical function of the form

$$\begin{aligned} &a_1 f_r^3 + a_2 (\Delta\theta)^3 + a_3 f_r^2 \Delta\theta + a_4 f_r (\Delta\theta)^2 + a_5 f_r^2 + a_6 (\Delta\theta)^2 + \\ &a_7 f_r \Delta\theta + a_8 f_r + a_9 \Delta\theta + a_{10} \end{aligned} \tag{5.16}$$

to the generated $\chi^2_* (f_r, \Delta\theta)$. This fitting function can be thought of as the third order two-dimensional Taylor expansion of $\chi^2_*$ and it generally fits the generated data points fairly well. Figure 5-4 presents typical $\chi^2_* (f_r, \Delta\theta)$ distributions as observed in three different exposures.

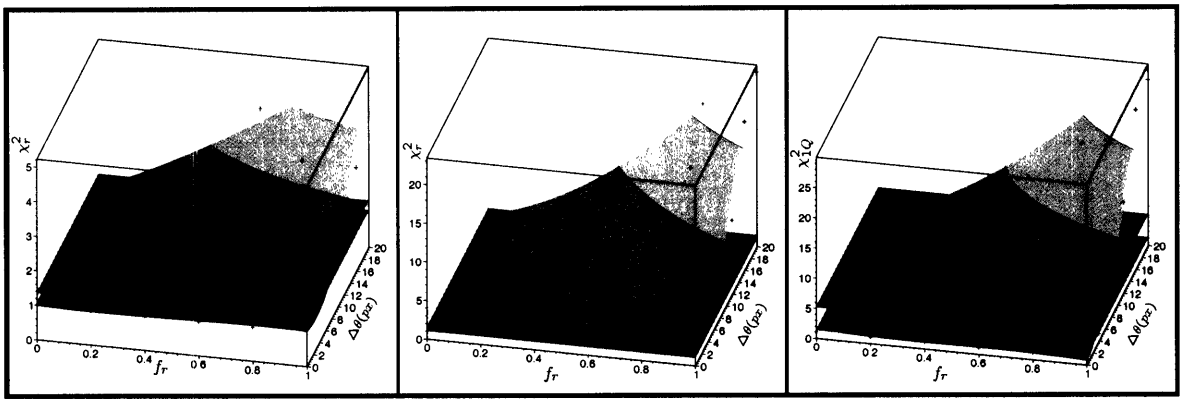


Figure 5-4: $\chi^2_\star(f_r, \Delta\theta)$ for (from left to right) exposures 70,31, and 020101.55. The blue planes indicate the measured value of $\chi^2_\star$, the green planes are the threshold $\chi^2_\star$ obtained by non-lensed simulation. Exposures 70 and 31 are not lensed and illustrate typical variations of the shape of $\chi^2_\star(f_r, \Delta\theta)$. Exposure 020101.55 is a potential mQSO. Note that $\Delta\theta$ is in units of pixels, not arcseconds.

By combining the non-lensed simulation results with $\chi^2_\star(f_r, \Delta\theta)$ we can identify the detection hull to be the region in the $(f_r, \Delta\theta)$ plane where

$$\chi^2_\star(f_r, \Delta\theta) \geq \chi^2_C. \tag{5.17}$$

Figure 5-4 illustrates the wide variation in shape that the detection hull can take. The detection hull for exposure 70, for example, is markedly different from that of exposure 31. While it is possible to combine equations 5.16 and 5.17 to determine the detection hull outline analytically such an approach is difficult to automate. Though the equations can be solved easily by an analytical mathematics package, the integral necessary to compute $P_{md}$ (equation 3.51) becomes extremely complicated.

Instead, we approximate (equation 3.51) by breaking up the detection hull into rectangular regions (see figure 5-5). We do this by first locating the boundary of the detection hull at discrete intervals in $f_r$ and then constructing rectangular regions that alternately underestimate and overestimate the actual area of the detection hull. The resulting integrals are well behaved and can be computed quickly.

### 5.2.8 Computing $L_{qso}$

With the detection hull approximation presented in figure 5-5 we almost have everything necessary to compute $P_{md}$. We still need to find $L_{qso}$ and we turn to this problem now.

The SDSS database gives us both the redshift and observed magnitude (in SDSS $i'$ band) of the QSOs. The redshift, $z_s$, is taken at face value and enters into equation 3.51 unaltered. We must, however, convert $i_{qso}$ into an observed intrinsic luminosity $L_{qso}l_{qso}$ and then compute $\frac{L_{qso}}{L_\star}$.

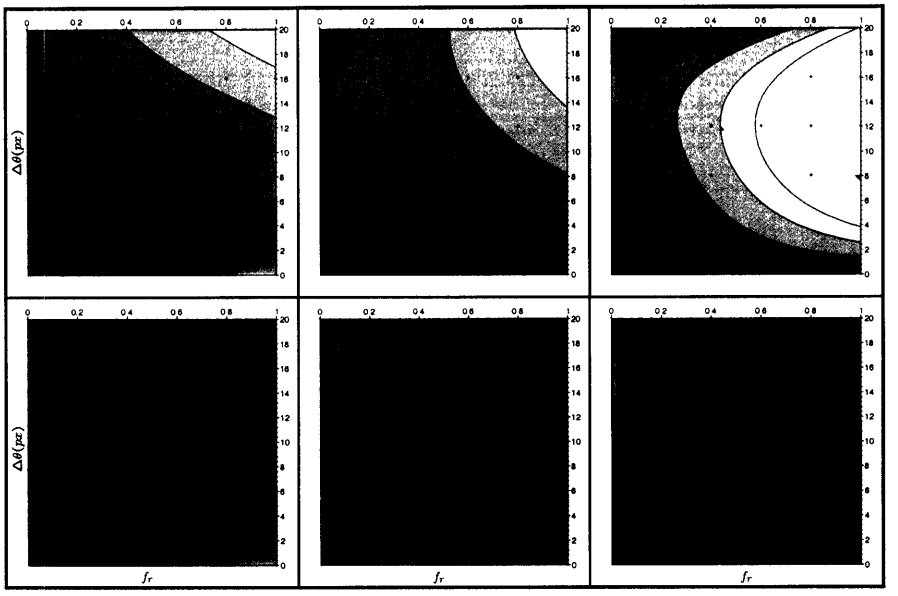


Figure 5-5: Various examples of detection hulls (top) and the approximation used to compute $P_{md}$ (bottom). Each red and blue rectangle is integrated over. These detection hulls are for exposures 315, 011229.104, and 011231.110.

The $i_{qso} \rightarrow L_{qso}$ transformation is made up of several elements. We must first compute the intrinsic magnitude of the QSO, $M_{qso}$, and then form

$$L_{qso} = 10^{-\frac{2}{5}(M_{qso})}. \tag{5.18}$$

To compute $M_{qso}$ we must take into account the QSOs distance from the Earth, and the fact that the QSO's spectra has been redshifted by the Hubble flow. The latter is of great concern: to compare the $M_{qso}$ of two QSOs at different redshifts we must compare the brightness of the *same part of the QSO spectrum.* By assuming that the QSO spectrum follows a simple power law we can compute a $k$-correction term ($k_c$) geared towards the SDSS $i'$ band [Schneider et. al.]

$$k_c = 2.5\,(0.79 - 1)\log_{10}\left(\frac{1+z_s}{5.2}\right). \tag{5.19}$$

To take the QSO's distance into account we can compute $dM$, the change in magnitude due to the QSO's redshift. This is computed in [Hogg] and is given by

$$dM = 5\log_{10}\left(\frac{(1+z_s)\,D_M\,(z_s)}{10\,\mathrm{pc}}\right) \tag{5.20}$$

where $D_M\,(z_s)$ is the comoving distance given in equation 3.3. $dM$ is known as the distance modulus and incorporates information about the assumed cosmology. To be consistent with our theoretical discussion we will use the $\Lambda - \mathrm{CDM}$ parameters in table 2.1.

we combine all these elements to find

$$M_{qso} = i_{qso} - k_c - dM. \tag{5.21}$$

It is customary to visualize QSO LF constraints in terms of the logarithm of $L_\star$, and not $L_\star$ itself [Comerford et. al.]. We shall follow suit by defining $M_\star$ so that

$$\frac{L_{qso}}{L_\star} = 10^{-\frac{2}{5}(M_{qso}-M_\star)} \tag{5.22}$$

or, more explicitly

$$L_\star = 10^{-\frac{2}{5}M_\star}. \tag{5.23}$$

To compare our QSO LF constraints with those of others we will need to compare break magnitudes: $M_\star$. Unfortunately, most QSO LF constraints in the literature are computed assuming a mass dominated cosmology with $\Omega_m = 1$, $h = 0.5$, and $\Omega_\Lambda = 0$. Thus we must rescale the $M_\star$ constraints found in the literature to compare them to our $M_\star$constraints. We do this by noting that cosmology manifests itself entirely in the distance modulus $dM$, thus we expect to obtain a different $dM^{\text{lit}}$ for the cosmology used in the literature.

There is one other difficulty that must be overcome before we can compare our QSO LF constraints with those in the literature. The literature results are presented in $B$ band magnitudes. To convert from a $B$ band $M_\star$ to an SDSS $i'$ band $M_\star$ we apply a $k$-correction with an effective redshift

$$z_{\text{eff}} = \frac{\lambda_i}{\lambda_B} - 1$$

where $\lambda_i = 7500$ and $\lambda_B = 4400$ are the central wavelengths of the two filters.

Combining the $k$-correction with the distance modulus correction we can then compute

$$M_\star^{\text{com}}(z_s) = M_\star^{\text{lit}}(z_s) - dM(z_s) + dM^{\text{lit}}(z_s) + k_c(z_{eff}). \tag{5.24}$$

We then compare $M_\star^{\text{com}}$ to our $M_\star$constraints.

With the ability to compute $L_{qso}$ and the detection hull we are finally capable of computing $P_{md}$ for all the QSOs in the data set. In the next chapter we present the results of these computations: the constraints on the QSO LF.

# Chapter 6

# Results

## 6.1 mQSOs in the QSO Sample

The theoretical framework and analysis approach presented in this thesis were applied to all 1073 high resolution QSO images. These QSOs span a redshift range $0.71 < z < 5.8$ and the distribution of redshift vs absolute magnitude is presented in figure 4-1.

Of these QSOs 15 were identified as mQSOs by the nonlensed simulation process discussed in section 5.2.6. After double checking the fitting residues by hand, however, it was found that 5 of the identified mQSOs are, in fact, false positives. Table 6.1 presents the vital statistics of each identified mQSO. The table includes, from left to right, the exposure ID, the QSO's RA, DEC, redshift, and observed magnitude, the threshold $\chi^2_C$ obtained from the Gumbel distribution, and the measured $\chi^2_*$. The last column distinguishes the mQSOs from the false positives.

The 5 false positives were identified by looking through the $\chi^2$ fit residues by eye. Note however that, while the remaining 10 candidate mQSOs appeared to be morphologically anomalous, this alone does not verify that they actually are multiply imaged gravitationally lensed objects; we discuss this further in section 6.3.

We now present images of the mQSO fits for all 15 identified mQSOs. For the false positives we also offer an explanation as to why the obtained $\chi^2_*$ were high.

| ID | RA | DEC | $z$ | $i_{qso}$ | $\chi^2_c$ | $\chi^2_*$ | mQSO? |
|---|---|---|---|---|---|---|---|
| 117 | 08:10:01.8 | +05:19:36.5 | 4.00 | 19.8 | 1.20 | 1.52 | Yes |
| 153 | 09:11:27.6 | +05:50:54.1 | 2.80 | 17.8 | 1.47 | 1.91 | Yes |
| 116 | 10:42:57.6 | +07:48:50.6 | 2.67 | 17.3 | 1.29 | 1.61 | Yes |
| 011011.243 | 03:33:20.4 | +00:07:20.6 | 1.25 | 16.6 | 2.02 | 5.94 | Yes |
| 011228.48 | 10:48:37.4 | -00:28:13.8 | 4.00 | 19.1 | 1.37 | 2.63 | Yes |
| 011229.168 | 03:48:01.2 | -07:04:16.7 | 1.97 | 18.1 | 1.40 | 1.77 | Yes |
| 011230.180 | 08:41:06.8 | +03:12:06.8 | 1.84 | 16.3 | 1.99 | 6.46 | Yes |
| 011230.318 | 11:12:25.7 | -00:51:01.8 | 1.82 | 18.2 | 1.60 | 1.84 | Yes |
| 011231.160 | 11:38:03.7 | +03:14:57.8 | 2.44 | 18.8 | 1.41 | 1.77 | Yes |
| 020101.55 | 01:34:39.3 | +00:17:31.7 | 1.67 | 18.1 | 1.33 | 5.13 | Yes |
| 4 | 09:44:9.5 | +10:06:56.7 | 4.78 | 19.3 | 1.25 | 1.32 | **NO** |
| 287 | 10:38:21.2 | +09:43:23.0 | 3.66 | 19.7 | 0.87 | 1.03 | **NO** |
| 011011.170 | 00:34:13.0 | -01:00:27.0 | 1.29 | 17.1 | 1.31 | 1.49 | **NO** |
| 011011.204 | 01:48:12.2 | +00:01:53.4 | 1.71 | 17.68 | 1.25 | 1.28 | **NO** |
| 011011.241 | 03:21:19.8 | -01:05:39.84 | 2.41 | 17.9 | 1.43 | 1.72 | **NO** |

Table 6.1: The vital statistics of the detected mQSOs

### 6.1.1 The mQSOs

We shall first describe the layout of figure 6-1, before discussing QSO 117; this description also applies to the other QSO images. The raster images on the left are (from left to right) the fit target, the residue (red indicates negative pixels, blue indicates positive pixels, each pixel is obtained by computing $D - f$), and the fitting object adjusted using the best fit parameters obtained from the $\chi^2_\nu$fit (the fit model used in all these images is described in equation 5.11). The rows of raster images, from top to bottom, are the results of the cs1 $\rightarrow$ QSO, cs2 $\rightarrow$ QSO, and cs2 $\rightarrow$ cs1 fits respectively.

On the right, the top image is the $\chi^2_*$ distribution obtained using the Gumbel distribution (equation 5.12). The red vertical line is the 0.1% threshold $\chi^2_c$ and the green line is the measure $\chi^2_*$. The figure on the bottom is $\chi^2_*(f_r, \Delta\theta)$ (equation 5.15). The blue plane is the measured $\chi^2_*$ plane, while the green plane is the threshold $\chi^2_c$ plane.

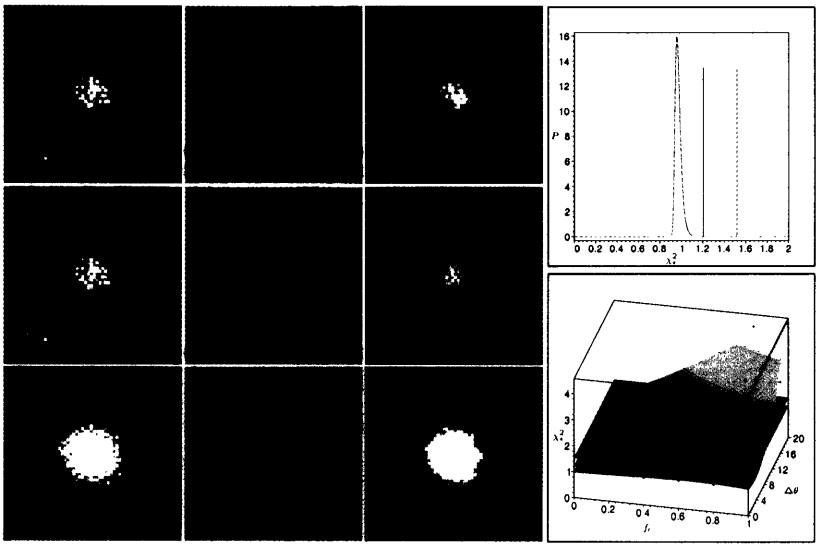


Figure 6-1: **Exposure 117**

Exposure 117: The QSO has a nearby partner ($\Delta\theta \approx 29$px $\approx 2$"). The residual plot clearly depicts the presence of the secondary image.

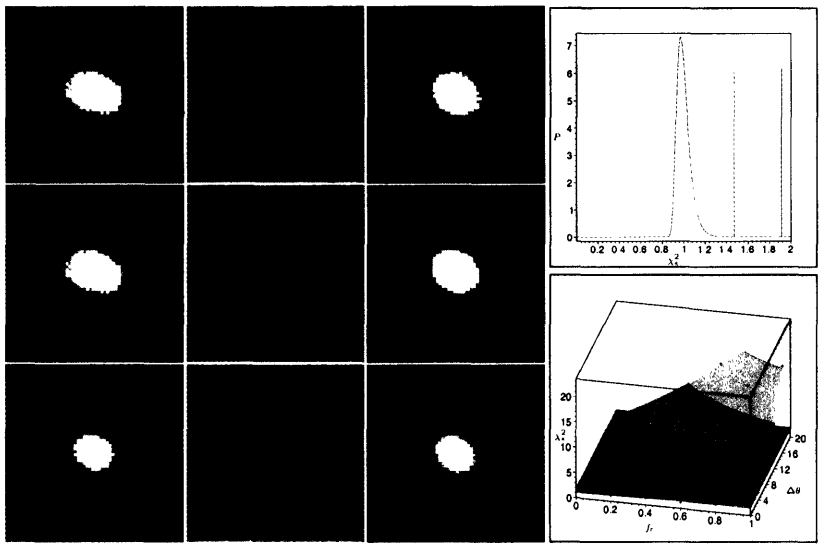

Figure 6-2: **Exposure 153**

Exposure 153: This QSO is an excellent example of the importance of using the $\chi^2_\nu$ fitting approach to identify mQSOs. The seeing quality in the exposure is poor and is incapable of resolving what is, in fact, three point sources at the QSO coordinates. The residual image, however, clearly indicates the presence of some extended structure in the QSO pixels. This QSO is the only one in our sample of mQSOs present in the CASTLES mQSO catalog: it is object RXJ0911+0551 and was first discovered in 1997!

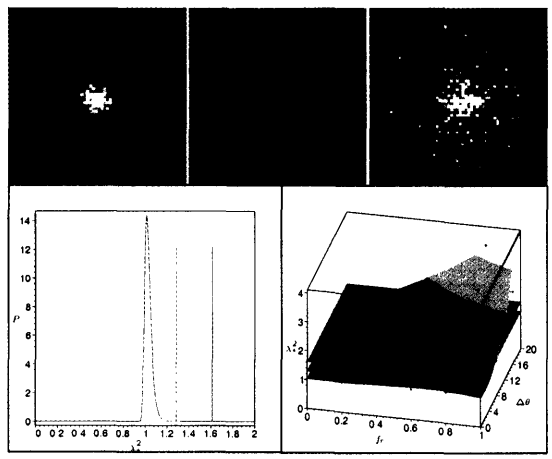

Figure 6-3: **Exposure 116**

Exposure 116: There was only one faint cstar available in this exposure. The QSO appears to be near an extended object, probably a galaxy. It is possible that higher resolution multi-band imagery might detect signs of lensing given the foreground galaxy's close proximity to the QSO.

The image at our disposal, however, does not offer any convincing proof that the QSO is, in fact, an mQSO.

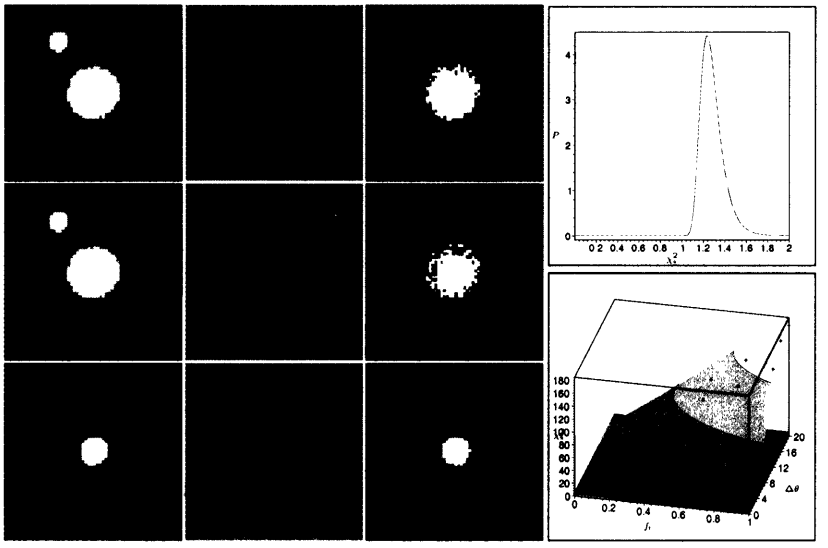

Figure 6-4: **Exposure 011011.243**

Exposure 011011.243: This exposure was discussed in section 5.2.5.2. The QSO is extremely bright relative to the other objects in the system and this has even shifted the non-lensed $\chi^2_\nu$ distribution towards significantly high $\chi^2$. The faint image is not lost in rescaled sky noise, but is, instead, strongly present in the residual images. The separation between the two objects is $\Delta\theta \approx 22.5\text{px} \approx 1.6"$.

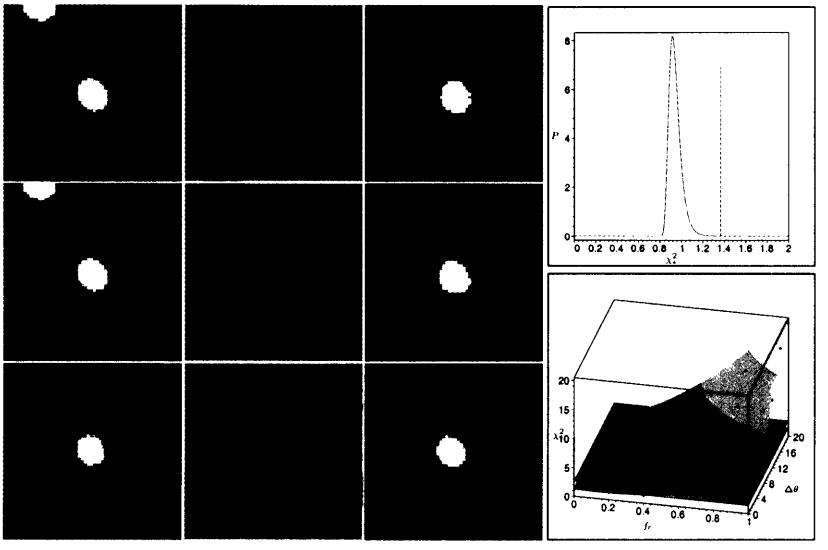

Figure 6-5: **Exposure 011228.48**

Exposure 011228.48: The QSO in this exposure has a superimposed cosmic ray (the black dot in the QSO images) which has been masked out using the Analyzer's masking feature. The QSO has a partner at a separation $\Delta\theta \approx 37.6\text{px} \approx 2.6"$.

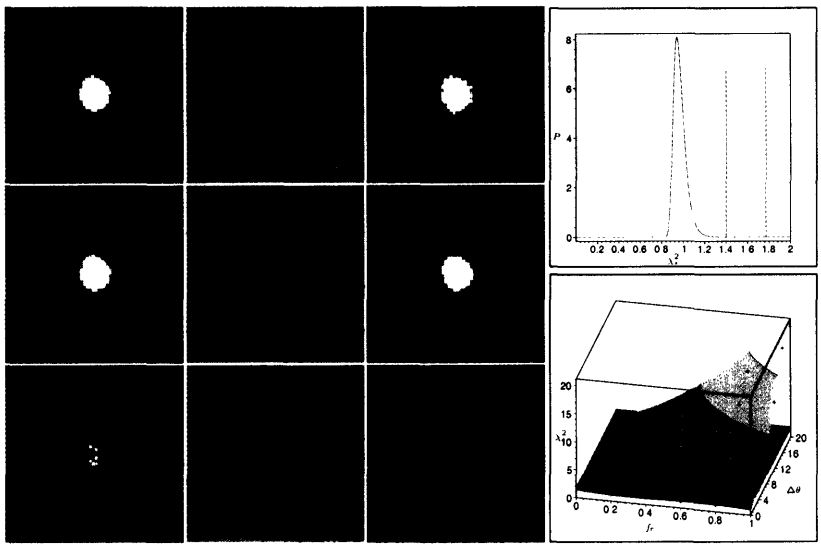

Figure 6-6: **Exposure 011229.168**

Exposure 011229.168: The QSO as a nearby partner (separation $\Delta\theta \approx 21.8\text{px} \approx 1.5"$) clearly visible in the residuals.

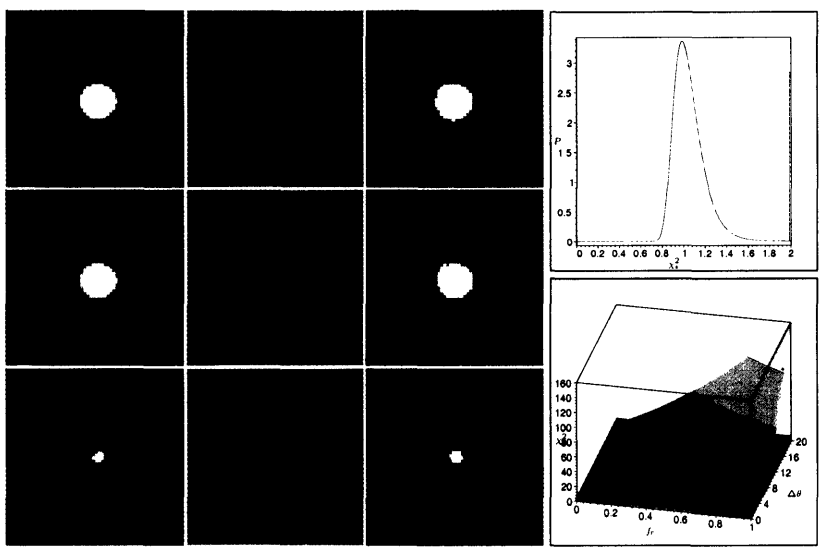

Figure 6-7: **Exposure 011230.180**

Exposure 011230.180: The QSO has a distant partner at a separation $\Delta\theta \approx 38.4\text{px} \approx 2.6"$.

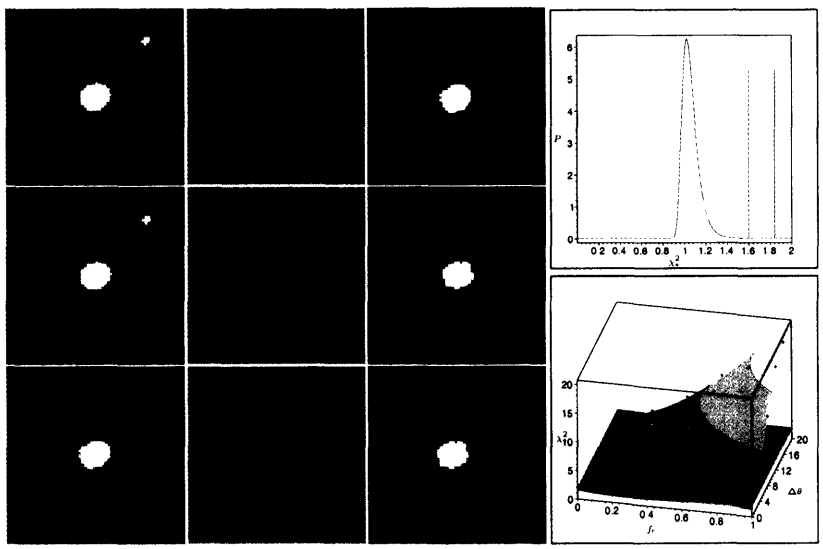

Figure 6-8: **Exposure 011230.318**

Exposure 011230.318: The QSO has a nearby partner at a separation $\Delta\theta \approx 26.6\text{px} \approx 1.8$".

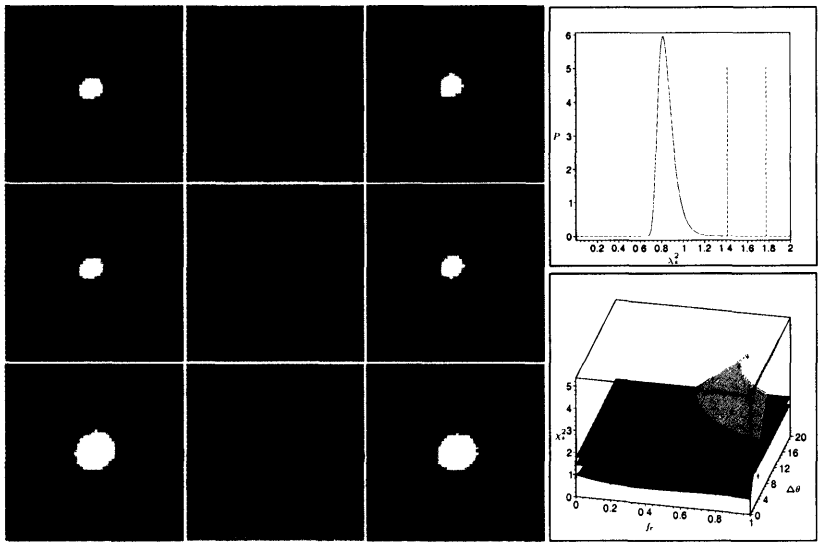

Figure 6-9: **Exposure 011231.160**

Exposure 011231.160: This QSO is an mQSO discovered by Professor Scott Burles while acquiring data for this survey. It is an excellent example of the typical quad mQSO morphology. The multiple images are clearly visible in the residual image. The blue patch at the center of the quad, as seen in the residual, may be the lensing galaxy! This QSO was selected on the basis of the presence of Mg II lines in the QSO spectrum.

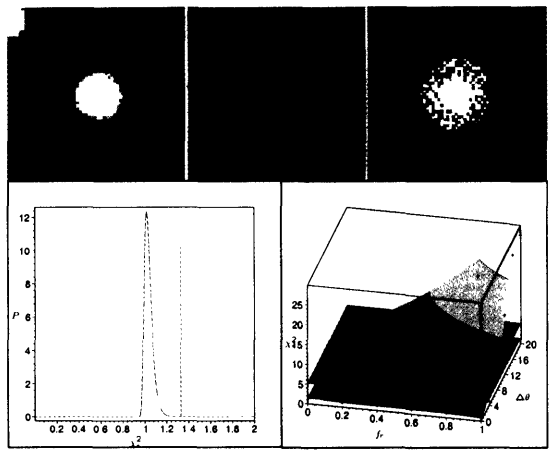

Figure 6-10: **Exposure 020101.55**

Exposure 020101.55: This exposure only has one, very faint, QSO. The $\Lambda$ region discussed in section 5.2.5.2 is clearly visible in the cstar image. This image is interesting. It shows the extremely magnified cstar core surrounded by a noisy black and white annulus. This signal in the annulus is predominantly from the sky surrounding the QSO. It is completely black and white because the rescaling of the sky to high values, followed by the background subtraction expands the sky noise so much that the sky pixels are either extremely positive (and, therefore, white on our color scale) or extremely negative (and therefore black). Just outside the $\Lambda$ region the sky coloration returns to its normal values. The residual image also illustrates the noisy effects of rescaling inside the $\Lambda$ region, and the less extreme noise outside the $\Lambda$ region.

The QSO has a partner at a separation $\Delta\theta \approx 43.5\text{px} \approx 3$".

### 6.1.2 The False Positives

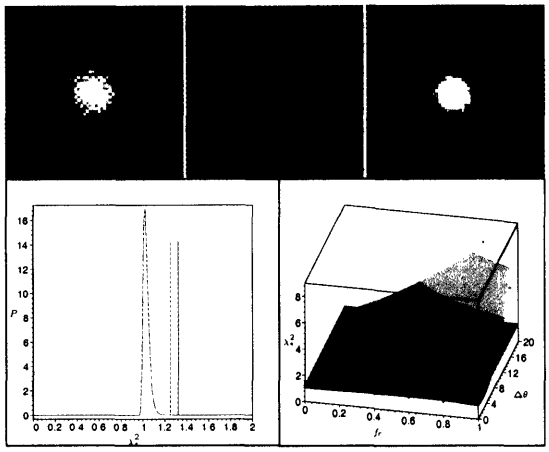

Figure 6-11: **Exposure 4**

Exposure 4: This exposure contains only one moderately bright cstar. Unfortunately, the QSO is quite faint with a signal about equal to the sky signal. The resulting $\chi^2_\nu$ is fairly high and only just exceeds the simulated nonlensed $\chi^2_\nu$ distribution. The QSO residual clearly shows that the QSO is *not* an mQSO: this exposure is a false positive.

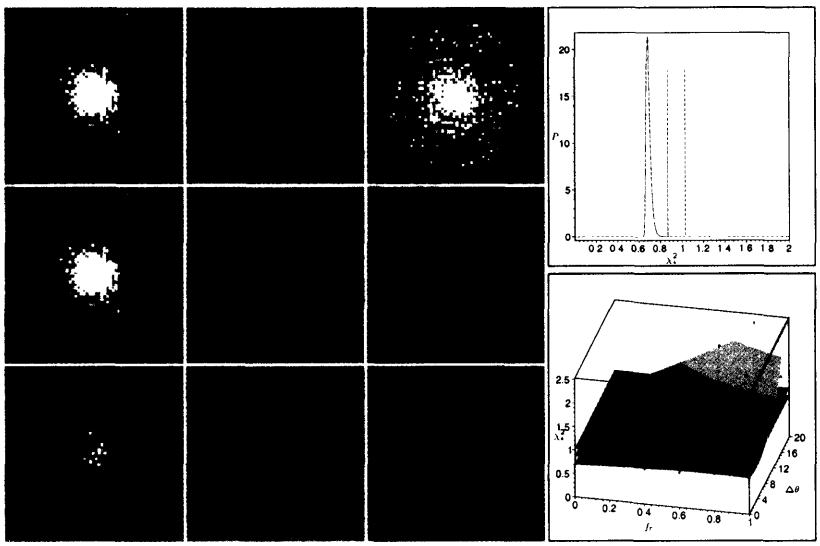

Figure 6-12: **Exposure 287**

Exposure 287: This exposure is another false positive. The residual in the cs1 fit clearly shows no sign of extended structure in the QSO. The second cstar, however, is extremely faint and fails to fit the QSO at all! In fact, the cs2 $\rightarrow$ cs1 fit also fails. This image is an example of the problems

one faces when one selects poor cstars.

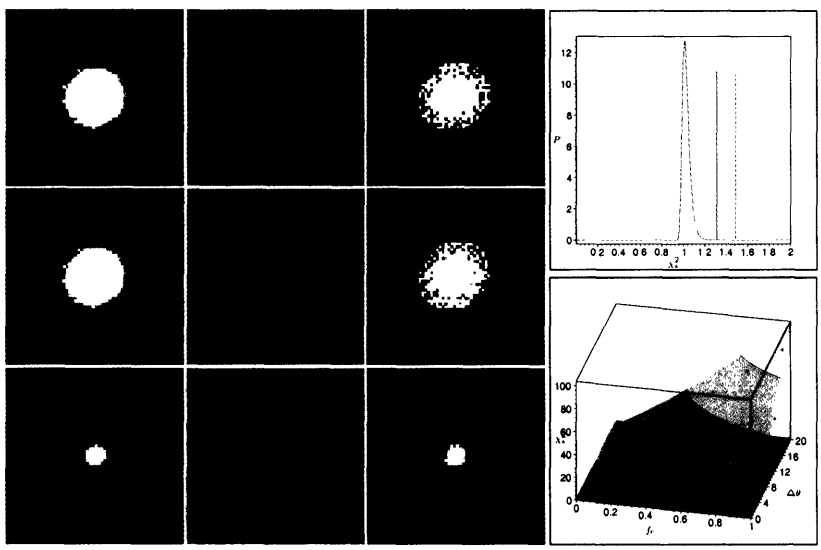


Figure 6-13: **Exposure 011011.170**

Exposure 011011.170: The QSO in this exposure is extremely bright relative to all the cstars in the exposure. While the cstars fit each other well ($\chi^2_\nu = 1.137$), the cstars both fail to fit the QSO well. In fact, this exposure, and the next remaining two false positives, illustrate the primary drawback to using the $\Lambda$ region fitting model described in section 5.2.5.2. In these two exposures the QSO is so bright that its $\Lambda$ region (the region where the signal is dominated by signal from the QSO) is larger than the cstars $\Lambda$ regions. This effect manifests itself in the blue halo surrounding the core $\Lambda$ region in each of the cstar $\rightarrow$ QSO fits and leads to higher $\chi^2_r$ values than anticipated.

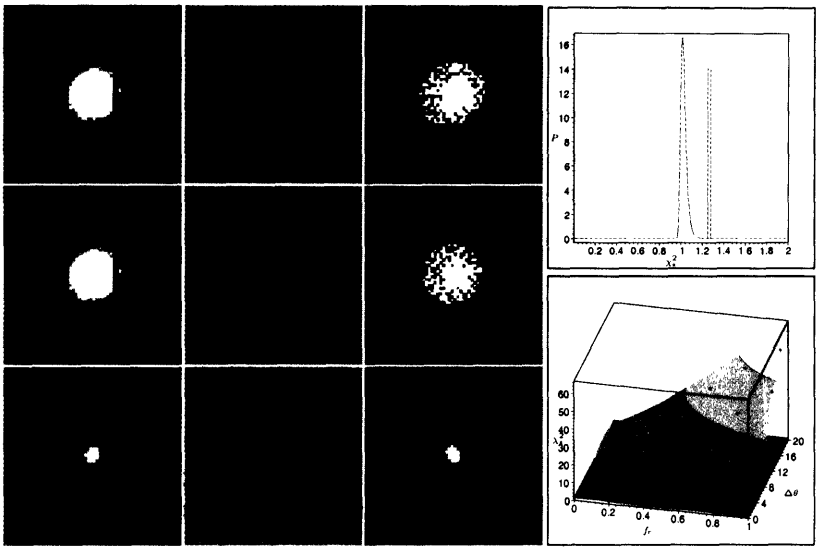

Figure 6-14: **Exposure 011011.204**

Exposure 011011.204: This exposure is essentially identical to exposure 011011.170. It also presents another example of masking. The QSO has a row of bad pixels in it which has been masked out. There are several other exposures in which this kind of masking takes place and none of those were identified as mQSOs. It appears the primary problem with this exposure is *not* the masking, it is the relative brightness of the QSO.

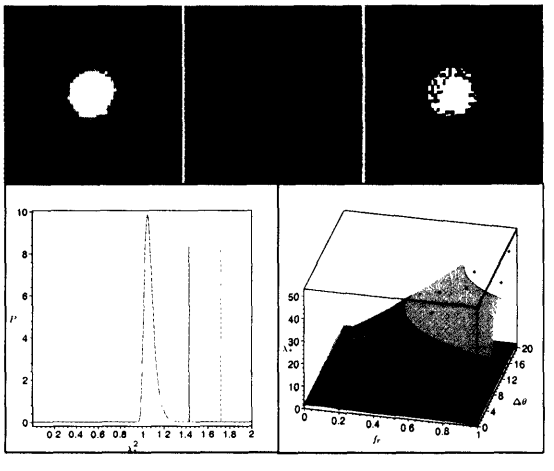

Figure 6-15: **Exposure 011011.241**

Exposure 011011.241: This exposure is also, essentially, identical to exposure 011011.170, though the pretense of only one, faint, cstar makes controlling the $\chi_r^2$ values even harder.

## 6.2 QSO LF Constraints

We constrain the QSO LF parameters in 6 redshift bins, chosen so that each bin contains the same number of QSOs. Figure 4-1 depicts the 6 bins; the redshift ranges of the bins, the average redshift of the QSOs in each bin, the number of QSOs in each bin, and the number of mQSOs in each bin, are presented in table 6.2.

| Bin | Redshift Range | Average Redshift | Number of QSOs | $n_{md}$ |
|---|---|---|---|---|
| 1 | $0.71 \leq z_s < 1.41$ | 1.18 | 178 | 1 |
| 2 | $1.41 \leq z_s < 1.75$ | 1.58 | 179 | 1 |
| 3 | $1.75 \leq z_s < 2.12$ | 1.93 | 179 | 3 |
| 4 | $2.12 \leq z_s < 2.98$ | 2.45 | 179 | 3 |
| 5 | $2.98 \leq z_s < 3.76$ | 3.42 | 178 | 0 |
| 6 | $3.76 \leq z_s < 5.8$ | 4.14 | 185 | 2 |

Table 6.2: The distribution of QSOs and mQSOs in the LF constraint bins.

Before we present the QSO LF parameter constraints, let's first consider how we expect the constraints to behave. Changing the parameters of the QSO LF changes the expected number of lenses. Intuitively, if we fix $M_\star$ and allow $\beta_h$, the high end slope, to vary then for high $\beta_h$ the number of bright QSOs drops rapidly. Since QSOs must be brighter than a certain threshold luminosity for them to be observable we would expect the number of observed QSOs that are lensed to rise. This is known as the **magnification bias**. The same reasoning holds true if we shift $M_\star$ to the fainter end (less negative). Thus, if we observe no mQSOs, then we should be able to constrain the $\beta_h$ and $M_\star$ from above[1]. We can say nothing about constraining $\beta_h$ and $M_\star$ from below if we observe no mQSOs: thus we can only constrain $\beta_h$ and $M_\star$ from above for redshift bin 5.

We *do* have mQSOs for the other bins, however, and their presence allows to set strong upper *and* lower constraints on $\beta_h$ and $M_\star$. For if $\beta_h$ and $M_\star$ are too low we would expect relatively fewer observed QSOs would need to be lensed to be observable; the magnification bias is low in such cases. All these properties are encapsulated in equation 3.54.

We shall compare our constraints to those obtained by [Pei], [Madau et. al.], [Wyithe and Loeb], and [Wyithe]. To do this we first describe the constraints found by these authors.

[Pei] does not evolve $\beta_h$ with redshift. He does, however, use an $L_\star$ (B band) evolution of the form

$$L_\star^{\text{Pei}}(z) = \tilde{L}_\star (1+z)^{-0.5} \exp\left[-\frac{(z-z_\star)^2}{2\sigma_\star^2}\right].$$

The parameters [Pei] uses are $\beta_h = 3.52 \pm 0.11$, $z_\star = 2.75 \pm 0.05$, $\sigma_\star = 0.93 \pm 0.03$, and $\tilde{L}_\star = 6.9 \times 10^{10} \pm 1.5 \times 10^{10}$.

[1]Note: "constraining $M_\star$ from above" means constraining how positive $M_\star$ can be or, equivalently, how *faint* it can be.

[Madau et. al.] also uses a fixed $\beta_h$ but adjusts Pei's $L_\star$ evolution to better fit the high redshift QSO LF

$$L_\star^{\text{Madau}}(z) = \tilde{L}_\star (1+z)^{-0.5} \frac{e^{\zeta z}\left(1+e^{\xi z_\star}\right)}{e^{\xi z}+e^{\xi z_\star}}.$$

The parameters [Madau et. al.] uses are $\beta_h = 3.52$, $z_\star = 1.9$, $\zeta = 2.58$, $\xi = 3.16$, and $\tilde{L}_\star = 8.7 \times 10^8$ (Madau does not present errors in his paper).

[Wyithe and Loeb] use Madau's $L_\star$ evolution but provide alternate values for the parameters. [Wyithe and Loeb] *do* allow for the evolution of $\beta_h$, but they also present a model that ignores the $\beta_h$ evolution. The parameters for the model that incorporates $\beta_h$ evolution are $\beta_h = 3.43$ for $z_s < 3$ and $\beta_h = 2.58$ for $z_s > 3$, $z_\star = 1.60$, $\zeta = 2.65$, $\xi = 3.30$, and $\tilde{L}_\star = 9.6 \times 10^{10}$. The nonevolving $\beta_h$ model uses $\beta_h = 3.43$, $z_\star = 1.45$, $\zeta = 2.70$, $\xi = 2.90$, and $\tilde{L}_\star = 9.6 \times 10^{10}$.

[Wyithe] constrains $\beta_h$ using the $z > 6$ QSOs found in the SDSS survey. He obtains $2.9 \leq \beta_h \leq 3.1$ at the 90% confidence level. We shall present this constraint in the high redshift bins used to compute our constraints.

Figure 6-16 presents the QSO LF constraints as color coded contour plots in the $(M_\star, \beta_h)$ plane for each redshift bin. The color scale in the figure describes the scaling: whiter regions represent more likely QSO LF models, while darker regions represent the least likely QSO LF models. Bins 1, 2, 3, 4, and 6, (the bins containing mQSOs) have maximal probability contours which we trace out in blue. The constraints in figure 6-16 assume that all the observed mQSOs in our sample are, in fact, real multiply imaged QSOs. Since this is probably not the case, we also present the constraints assuming that only the verified mQSOs are real mQSOs (exposures 153 and 011231.160). Both of these mQSOs lie in the fourth redshift bin. The resulting conservative constraints are presented in figure 6-17.

The constraint plots show good agreement between our constraints and those of [Pei], [Madau et. al.], [Wyithe and Loeb], and [Wyithe] at low redshift. At high redshift we verify the increasingly popular possibility that $\beta_h$ flattens out at higher redshift. Our results match those of [Wyithe and Loeb] (with evolving $\beta_h$), and [Wyithe] well at high redshift. Particularly important is the good agreement of our results in redshift bin #4, where we are certain of the number of mQSOs.

To compare constraints, we compute and present the $M_\star$s of [Pei], [Madau et. al.], [Wyithe and Loeb], and [Wyithe] over the redshift range of each bin; we also compute the $M_\star$s at the average redshift of the bin (see table 6.2). We present the resulting range in $M_\star$s as a line on the plot, and the value of $M_\star$ at the average redshift as a point on this line.

The slightly poorer agreement in the first three redshift bins in figure 6-16 is probably due to two factors. First, the actual number of mQSOs may have been overestimated. Second, the constraints are smeared out because of the size of the redshift bins in which they are computed. This smearing is a systematic effect and has a tendency to prematurely pull the constraint contours up to higher $\beta_h$ at relatively faint $M_\star$. Compared to the constraints, we find better agreement between the fainter (in terms of $M_\star$) end of the constraint and our contours than the brighter end (see redshift bins 1,2,

and 3 in particular; bin 6 also exhibits similar behavior).

In appendix B we present these results again using a renormalized version of $\frac{dP}{d\mu_s}$.

## 6.3 Potential Improvements

Several improvements and refinements can be made to the analysis we have described in this thesis. Perhaps the most important of these is the determination of the number of mQSOs in each redshift bin. While our algorithm is capable of picking out morphologically interesting QSOs there is no guarantee that these are actually gravitationally lensed mQSOs. The QSO may, by chance, be close to a faint foreground star and some of the mQSOs might actually be binary QSOs (an interesting phenomenon in and of itself). By far, the most convincing verification that a QSO is an mQSO is the extraction of the lensing galaxy from high resolution imagery. This can be done with long exposure times and multiple color imaging. One can also compare the spectra of the multiple images and compare the light curves of the images over time to verify that the images do, in fact, stem from the same source.

Improvements can also be made to the theoretical lensing probabilities derived in chapter 3. According to [Li and Ostriker] the NFW lens model is superior to the SIS model in its agreement to simulation results. Unfortunately, the NFW model complicates the probability integrals severely making the integrals more computationally intensive. In large statistical samples such as the one we have been working with, this can overwhelmingly increase the necessary analysis time.

In our analysis we assumed that early-type galaxies were the dominant lenses. The veracity of this statement should be determined by analyzing the contributions to the multiple imaging cross section of other potential lenses.

There are also a variety of improvements that can be made to the analysis pipeline discussed in chapter 5. The acquisition of longer exposure data would help expand the mQSO detection hull, allowing us to, potentially, identify more mQSOs. Ideally the acquisition of high resolution imagery would occur over several nights to allow for more consistent images as well as longer exposures. This would decrease the variability in the quality of the exposures and would significantly reduce the number of false positives in the analysis pipeline. Improvements could also be made to better determine the mQSO detection hull. A denser $(\Delta\theta, f_r)$ grid, with an increased $\Delta\theta$ range, would greatly improve the determination of the detection hull – though such an improvement is unlikely to greatly affect the resulting lensing probabilities.

## 6.4 Acknowledgements

The work for this thesis was partially supported by the Paul E. Gray (1954) Endowed Fund for UROP. The author is grateful for the excellent comments made by Kristin Burgess and Michael Mortenson. I am greatly indebted to Professor Scott Burles whose relentless support and excellent insight throughout the various phases of the research made this work possible.

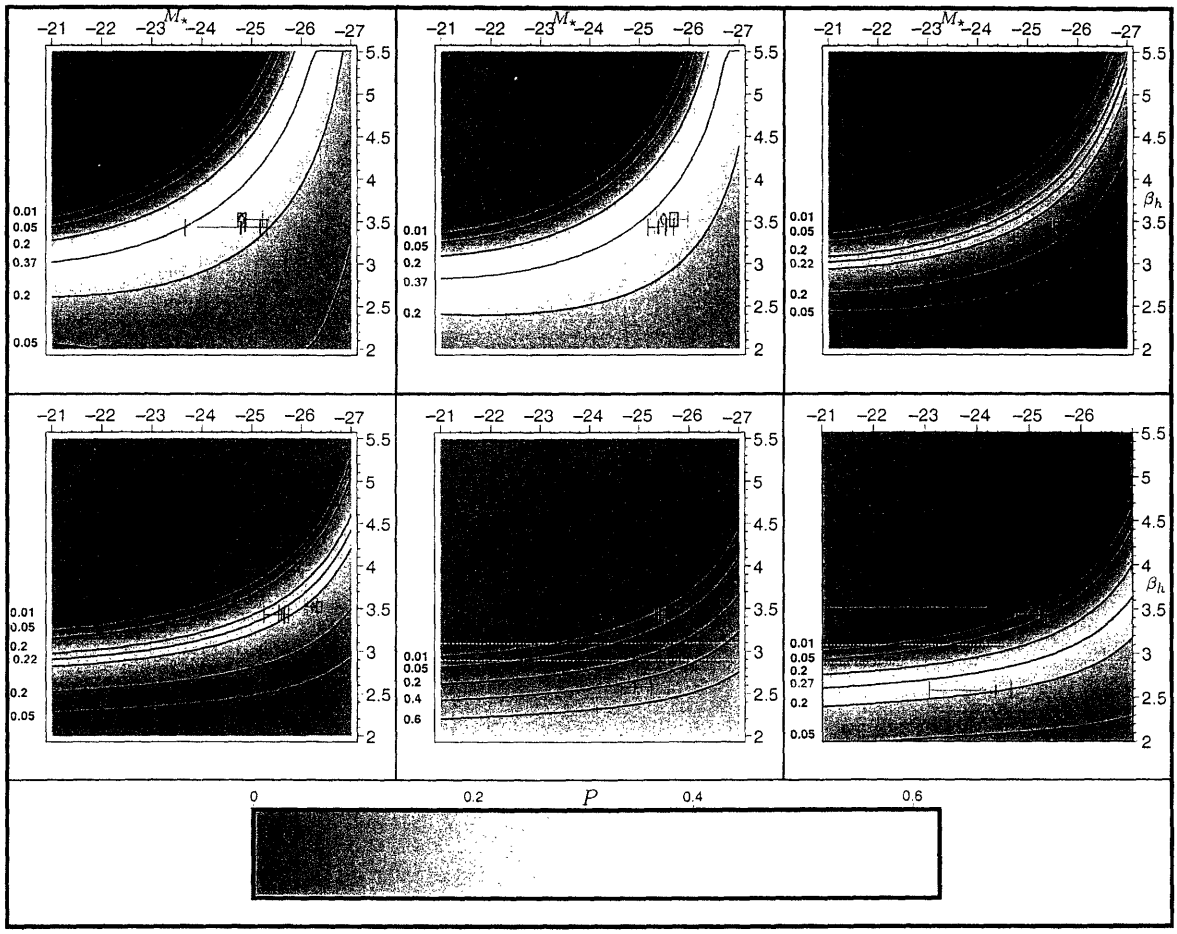


Figure 6-16: The QSO LF Constraints. Top (left to right): the constraints in bins 1,2,3. Bottom (left to right): the constraints in bins 4,5,6. The vertical axes are $\beta_h$ while the horizontal axes are $M_\star$. The plot beneath the curves is the color scaling used to generate the QSO LF constraint plots. The horizontal axis is $P$ as defined in equation 3.54. The constraints from [Pei] are presented in red, the constraints from [Madau et. al.] are presented in green, the constraints from [Wyithe and Loeb] are presented in blue (with $\beta_h$ evolution) and black (without $\beta_h$ evolution). Finally, the high redshift constraint on $\beta_h$ by [Wyithe] is bounded by the dashed white lines.

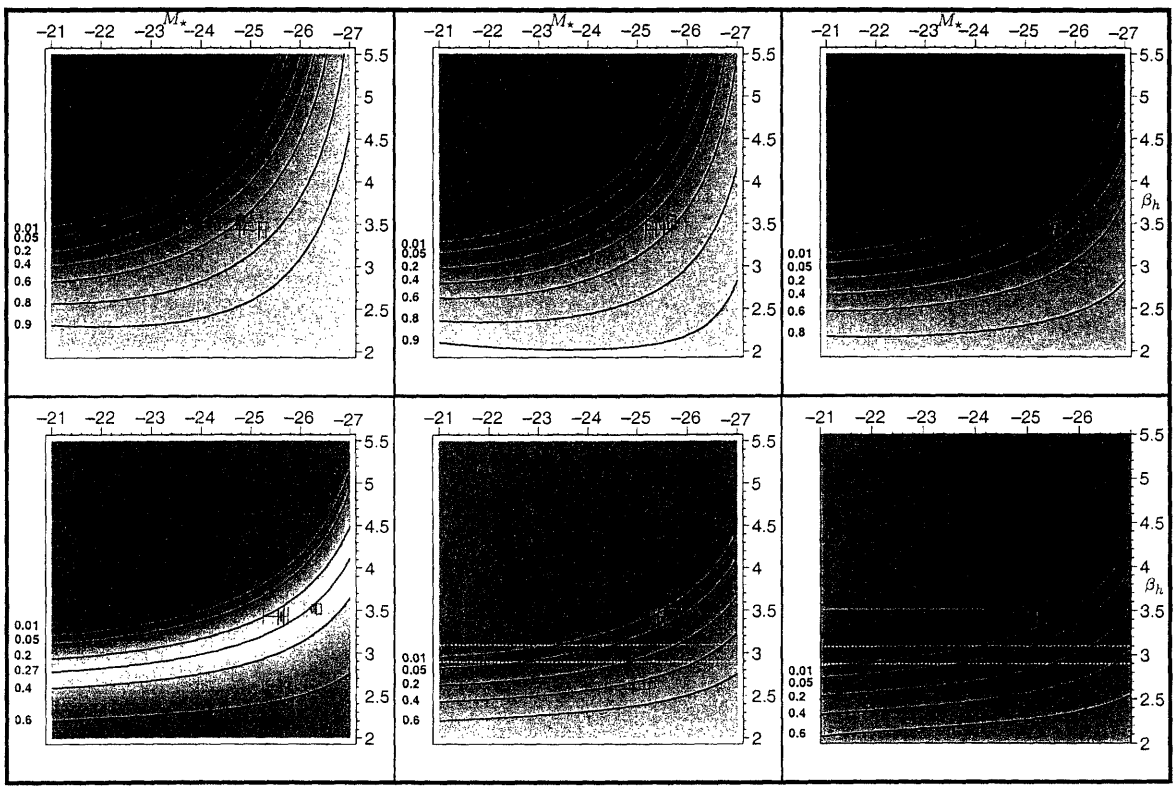


Figure 6-17: The conservative QSO LF constraints, computed assuming only the verified mQSOs are, in fact, real mQSOs.

# Appendix A

# $\phi_R$ Fitting Functions

The plot of $\phi_R(\sigma, z)$ in figure 3-4 presents the results of simulation computations (the discrete points) accompanied by an overlaid analytical fit. We shall describe the fit function and present the fit parameters in this section. Unfortunately, we were unable to obtain the errors associated with the simulation computations and, as such, we cannot quote errors to our fitted parameters. We have, however, already discussed the goodness of fit in terms of the relative error between the analytical fit function and the data points (see section 3.4.2).

We tailored an analytical fit to match the data points in a step by step process. We first fit functions to trace out the behavior of $\phi_R$ at low $\sigma$ and at high $\sigma$. We shall call these **edge functions** as they delineate the $\sigma$ edges of the $\phi_R(\sigma, z)$ distribution. To be precise, we define the low and high $\sigma$to be

$$\sigma_l = 20.07285 \tag{A.1}$$

$$\sigma_h = 686.53617 \tag{A.2}$$

We then define the edge functions

$$f_{l\sigma}(z) \approx \phi_R(\sigma_l, z) \tag{A.3}$$

$$f_{h\sigma}(z) \approx \phi_R(\sigma_h, z) \tag{A.4}$$

throughout this appendix $f_*$ will refer to analytical fitting functions. Also, when the context is not clear, we use the notation $f_*.a_i$ to denote the $i^{th}$ fit parameter of fit function $f_*$. For $f_{l\sigma}(z)$ we chose a simply polynomial analytical fit function of the form

$$f_{l\sigma}(z) = 1 + a_1 z + a_2 z^2 + a_3 z^3 + a_4 z^4. \tag{A.5}$$

The resulting fit is presented in figure A-1, and the fit parameters are presented in table A.1.

For $f_{h\sigma}(z)$ we chose a fitting function of the form

$$f_{h\sigma}(z) = \exp(a_1 z^{a_2}). \tag{A.6}$$

| $f_{l\sigma}$ Parameter | Value | $f_{h\sigma}$ Parameter | Value |
|---|---|---|---|
| $f_{l\sigma}.a_1$ | $6.823 \times 10^{-1}$ | $f_{h\sigma}.a_1$ | $-9.531 \times 10^{-1}$ |
| $f_{l\sigma}.a_2$ | $2.515 \times 10$ | $f_{h\sigma}.a_2$ | 1.408 |
| $f_{l\sigma}.a_3$ | $-6.717 \times 10^{-2}$ | | |
| $f_{l\sigma}.a_4$ | $4.086 \times 10^{-3}$ | | |

Table A.1: The fit parameters for $f_{h\sigma}$and $f_{l\sigma}$

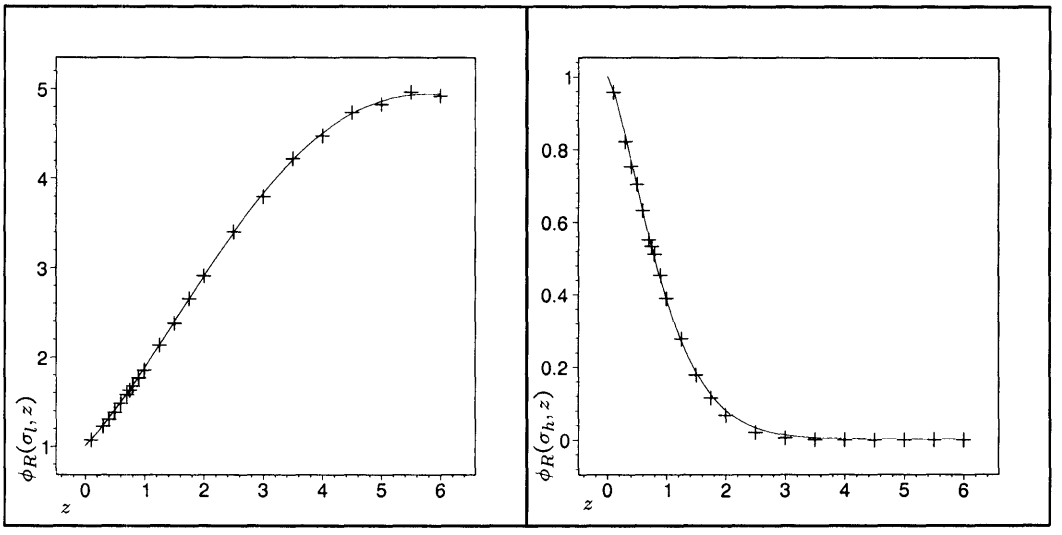


Figure A-1: Plots of the $f_{l\sigma}$ (left) and $f_{h\sigma}$ (right) fits.

The resulting fit is presented in figure A-1, and the fit parameters are presented in table A.1. Note that both $f_{h\sigma}$ and $f_{l\sigma}$ go to 1 at $z = 0$. This is a necessary condition enforced by equation 3.15.

We now use these edge functions to guide what we call **filler functions** – functions that trace out the $z$ behavior of $\phi_R$ as a function of $z$. After much experimenting it was found that one analytical function would be insufficient to completely describe the behavior of $\phi_R$ at all redshifts. We have opted, instead, to define two filler function forms: one that is valid for $z \leq 2$ and another that is valid for $z > 2$. (This splitting is unimportant: when we compute integrals involving $\phi_R$ we simply split the redshift integral into two). Thus, we define

$$f_{lz}(\sigma, z) \approx \phi_R(\sigma, z \leq 2) \tag{A.7}$$

$$f_{hz}(\sigma, z) \approx \phi_R(\sigma, z > 2). \tag{A.8}$$

The analytical form chosen for $f_{lz}(\sigma, z)$ is

$$f_{lz}(\sigma, z) = f_{l\sigma}(z) + \left[f_{h\sigma}(z) - f_{l\sigma}(z) - b_1(z)(\sigma_h - \sigma_l)^2\right] \frac{\sigma - \sigma_l}{\sigma_h - \sigma_l} + b_1(z)(\sigma - \sigma_l)^2. \tag{A.9}$$

This is not as complicated as it looks. It is simply a quadratic constrained such that $f_{lz}(\sigma_l, z) =$

| $b_1$ Parameter | Value | $b_2$ Parameter | Value | $b_3$ Parameter | Value |
|---|---|---|---|---|---|
| $b_1.a_1$ | $-3.071 \times 10^{-3}$ | $b_2.a_1$ | $-4.440 \times 10^{-3}$ | $b_3.a_1$ | 1.833 |
| $b_1.a_2$ | $-9.309 \times 10^{-4}$ | $b_2.a_2$ | $5.733 \times 10^{-3}$ | $b_3.a_2$ | $-4.201 \times 10^{-1}$ |
| $b_1.a_3$ | $1.093 \times 10^{-3}$ | $b_2.a_3$ | $-2.293 \times 10^{-3}$ | $b_3.a_3$ | $7.293 \times 10^{-2}$ |
| | | $b_2.a_4$ | $1.322 \times 10^{-4}$ | $b_3.a_4$ | $-4.420 \times 10^{-3}$ |

Table A.2: The parameters for the $b_i$ functions.

$f_{l\sigma}(z)$ and $f_{lz}(\sigma_h, z) = f_{h\sigma}(z)$. These two constraints leave only one free fitting parameter, $b_1(z)$. Unfortunately, $b_1$ cannot simply be a constant: it must change with $z$ to track the curvature of $\phi_R(\sigma, z)$. We must, therefore, find an appropriate fitting function for $b_1(z)$! To do this we first fit $f_{lz}(\sigma, z)$ to the $\phi_R$ data by holding $z$ at a constant $z_c$ and fitting the function of $\sigma$: $f_{lz}(\sigma, z_c)$ to $\phi_R(\sigma, z_c)$. This gives us the value of $b_1(z_c)$. We repeat this for all values of $z_c \leq 2$ in the $\phi_R$ data set and obtain a scatter plot of $b_1(z)$ vs $z$ (figure A-2). We then choose an appropriate function to fit to this scatter plot.

Again, after much trial and error we choose a function of the form

$$b_1(z) = \left(a_1 - a_1 \exp\left(a_2 z + a_3 z^2\right)\right). \tag{A.10}$$

This function guarantees that $b_1(z) = 0$ at $z = 0$, which guarantees that $f_{lz}(\sigma, 0) = 1$ which it must to agree with equation 3.15. The values of the parameters are presented in table A.2.

The analytical form chosen for $f_{hz}(\sigma, z)$ is

$$f_{hz}(\sigma, z) = \frac{\left(f_{h\sigma}(z) - f_{l\sigma}(z)\right)\left[1 + \exp\left(b_2(z)\left(\sigma - \sigma_m\right)^{b_3(z)}\right)\right]}{\exp\left(b_2(z)\left(\sigma - \sigma_m\right)^{b_3(z)}\right) - 1} - f_{l\sigma}(z). \tag{A.11}$$

Again, this function is not as complicated as it appears. It is simply an exponential fit equation of the form $A\left[1 + e^{b_2(z)(\sigma-\sigma_m)^{b_3(z)}}\right] + B$ where $A$ and $B$ have been determined so that $f_{hz}(\sigma_l, z) = f_{l\sigma}(z)$ and $f_{hz}(\sigma_h, z) = f_{h\sigma}(z)$. Just as in the fitting function for $f_{lz}$ the expression for $f_{hz}$ has $z$ dependent fitting parameters. We proceed as we did with $b_1(z_c)$ by finding scatter plots of $b_2(z)$ vs $z$ and $b_3(z)$ vs $z$ (figure A-2) and fitting analytical expressions for $b_2$ and $b_3$ to the resulting scatterplots.

It turns out that we can simply use polynomial fit functions for both $b_2(z)$ and $b_3(z)$. To be precise

$$b_{2(3)}(z) = a_1 + a_2 z + a_3 z^2 + a_4 z^3. \tag{A.12}$$

The values of the parameters are presented in table A.2.

The analytical forms for the $b_i$ allow us to express $f_{hz}(\sigma, z)$ and $f_l(\sigma, z)$ completely analytically. The resulting function was presented in figure 3-4 along with an image depicting the relative errors of the fit.

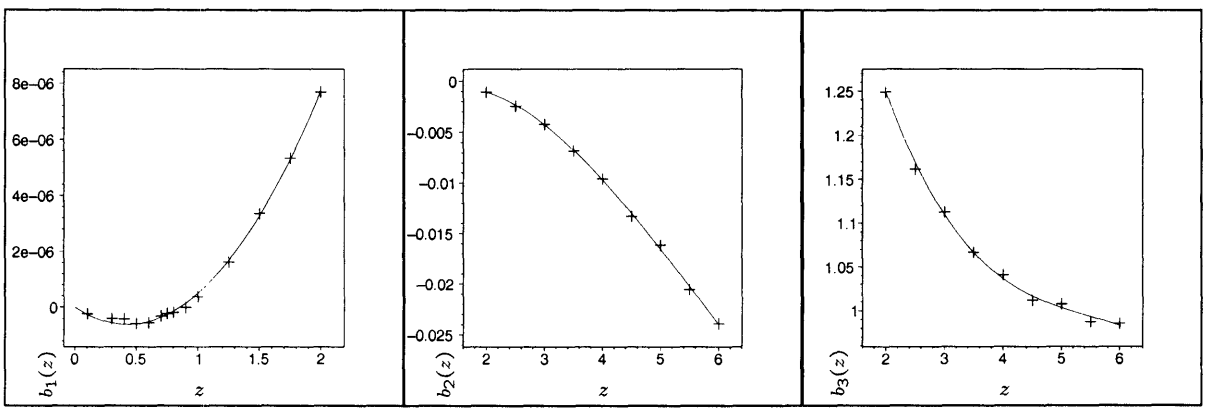


Figure A-2: The $b_i$ scatter plots with associated fit functions. From left to right: $b_1$,$b_2$, and $b_3$.

# Appendix B

# Renormalizing $\frac{dP_s}{d\mu}$

We discussed, in section 3.6.4, the physical motivation behind renormalizing $\frac{dP_s}{d\mu}$. As it stands the combined magnification probability distribution defined in equation 3.28 yields a mean magnification $\langle\mu\rangle > 1$. This cannot be right and we seek to renormalize $\frac{dP_s}{d\mu}$ so that $\langle\mu\rangle = 1$. At the same time we must retain the fact that $\int_2^\infty \frac{dP_m}{d\mu} d\mu$ is the multiple imaging cross section while keeping the total probability distribution normalized.

Our approach will be similar to that of [Comerford et. al.] and [Wyithe and Loeb]: we will add constant probability to the low $\mu$ end of $\frac{dP_s}{d\mu}$ down to a new minimum $\mu$

$$\mu_{min} < 1. \tag{B.1}$$

The value for this constant probability will be given by $\frac{dP_s}{d\mu}$ evaluated at some $\mu_c > 1+\frac{\theta_E}{\pi}$ (remember that $1 + \frac{\theta_E}{\pi}$ was the old minimum value of $\mu$). We define

$$P_c \equiv \left.\frac{dP_s}{d\mu}\right|_{\mu=\mu_c}. \tag{B.2}$$

We can write the renormalization constraints in terms of $\mu_{min}$ and $\mu_c$. To enforce the standard probability normalization we must set

$$(\mu_c - \mu_{min})\, P_c + \int_{\mu_C}^{2} \frac{dP_s}{d\mu} = 1 - \int_2^\infty \frac{dP_m}{d\mu} d\mu. \tag{B.3}$$

To force $\langle\mu\rangle = 1$ we set

$$\frac{1}{2}\left(\mu_c^2 - \mu_{min}^2\right) P_c + \int_{\mu_C}^{2} \mu \frac{dP_s}{d\mu} = 1 - \int_2^\infty \mu \frac{dP_m}{d\mu} d\mu. \tag{B.4}$$

We can solve these two equations numerically to obtain, for each value of $\theta_E$ both $\mu_{min}$ and $\mu_c$. This process is not very enlightening, it simply involves solving for $\mu_{min}$ in terms of $\mu_c$ for each equation (this just involves solving a linear equation for the normalization constraint and a quadratic equation for the $\langle\mu\rangle = 1$ constraint). We then compute the $\mu_{min}$ for a values of $\mu_c$ on a

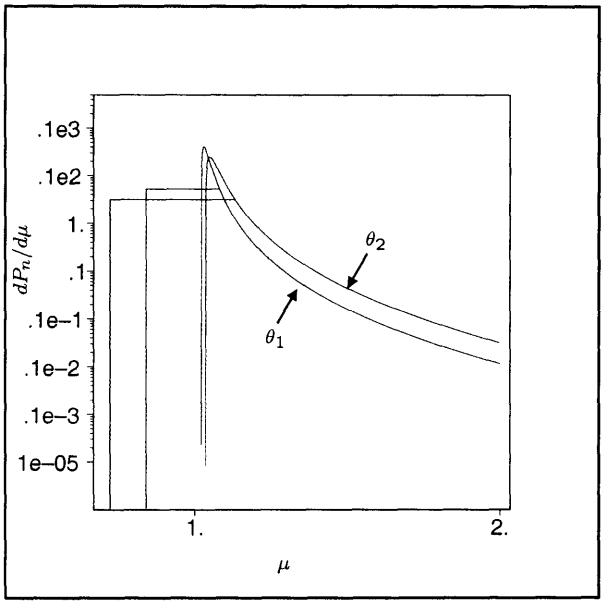


Figure B-1: $\frac{dP_n}{d\mu}$ as normalized for $\theta_1 = 0.048$ and $\theta_2 = 0.079$.

grid and find the (unique) value of $\mu_c$ that yields agreement for both expressions for $\mu_{min}$. We then define the normalized singly imaged magnification probability distribution

$$\frac{dP_n}{d\mu} = \begin{cases} P_c & \mu_{min} < \mu \leq \mu_c \\ \frac{dP_s}{d\mu} & \mu_c < \mu \leq 2 \end{cases} . \tag{B.5}$$

Figure B-1 presents plots of $\frac{dP_n}{d\mu}$ for two values of $\theta_E$: $\theta_1 = 0.048$ represents the effective Einstein radius of the lenses lensing an object at $z_s = 2$ and $\theta_2 = 0.079$ is the effective Einstein radius for an object at $z_s = 4$. The $\frac{dP_n}{d\mu}$ distribution normalized for $\theta_1$ has $\mu_{min} = 0.896$ and $\mu_c = 1.058$. The $\frac{dP_n}{d\mu}$ distribution normalized for $\theta_2$ has $\mu_{min} = 0.825$ and $\mu_C = 1.0966$.

We've used $\frac{dP_n}{d\mu}$ to compute new constraints on the QSO LF. These are presented in figure B-2, with the corresponding conservative constraints presented in figure . There are some small differences (which become more pronounced at higher redshifts). For a more direct comparison figure B-3 presents the difference in the constraints $P - P_n$ where $P$ is defined in equation 3.54 and $P_n$ is the corresponding definition involving $\frac{dP_n}{d\mu}$.

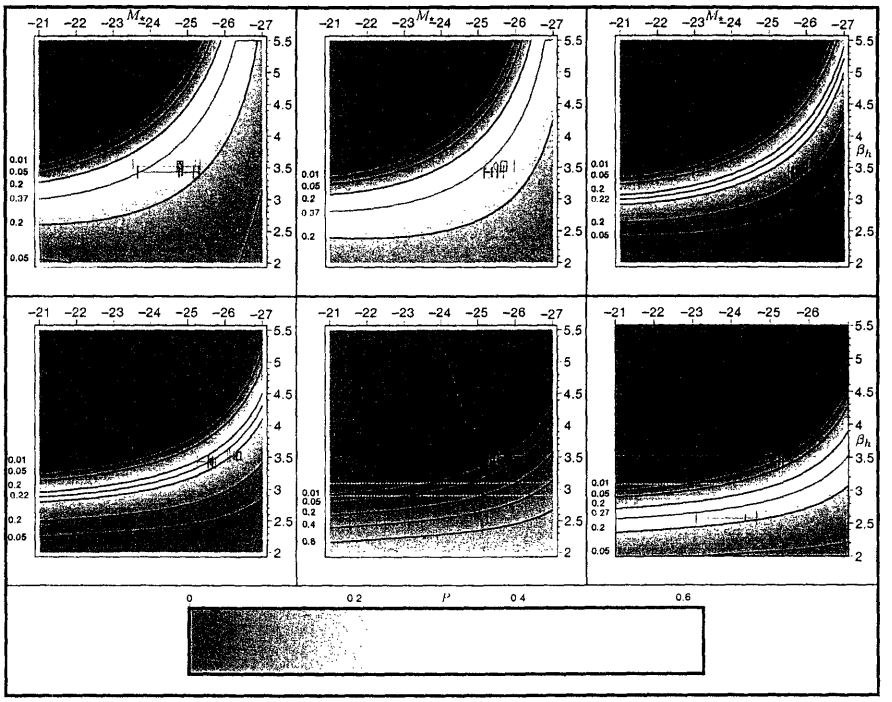


Figure B-2: The QSO LF constraints computed with $\frac{dP_n}{d\mu}$.

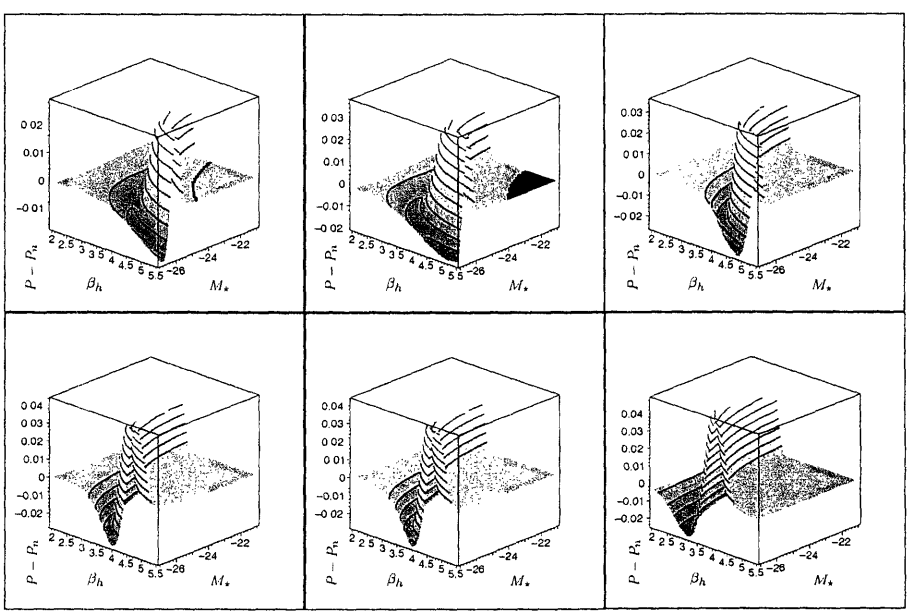


Figure B-3: The difference $P - P_n$ as computed in each redshift bin.

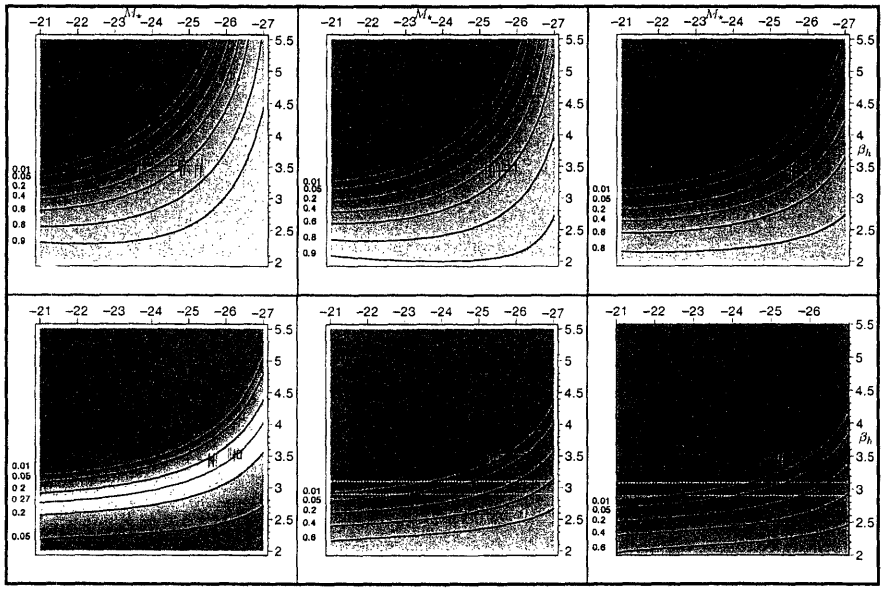


Figure B-4: The conservative QSO LF constraints computed with $\frac{dP_n}{d\mu}$.

# Appendix C

# Software Screenshots

The Analyzer and Coordinator are software packages written to analyze and organize the high resolution QSO imagery. Both packages were implemented in C++ using the Tcl/Tk library for the graphical user interface.

The Coordinator was used to organize the second run QSO target list and optimize the QSO ordering to minimize the telescope slew time between targets. Figure C-1 is a screenshot of the Coordinator in action. The labels on the figure illustrate the different functions that the Coordinator carries out.

The Analyzer was used to select targets and perform $\chi^2_\nu$ fits. Figure C-2 is an image of the Analyzer performing a PSF fit. Figure C-3 illustrates the Analyzer's masking capabilities. Figure C-4 presents the Analyzer's $\chi^2_\nu$ fitting interface.

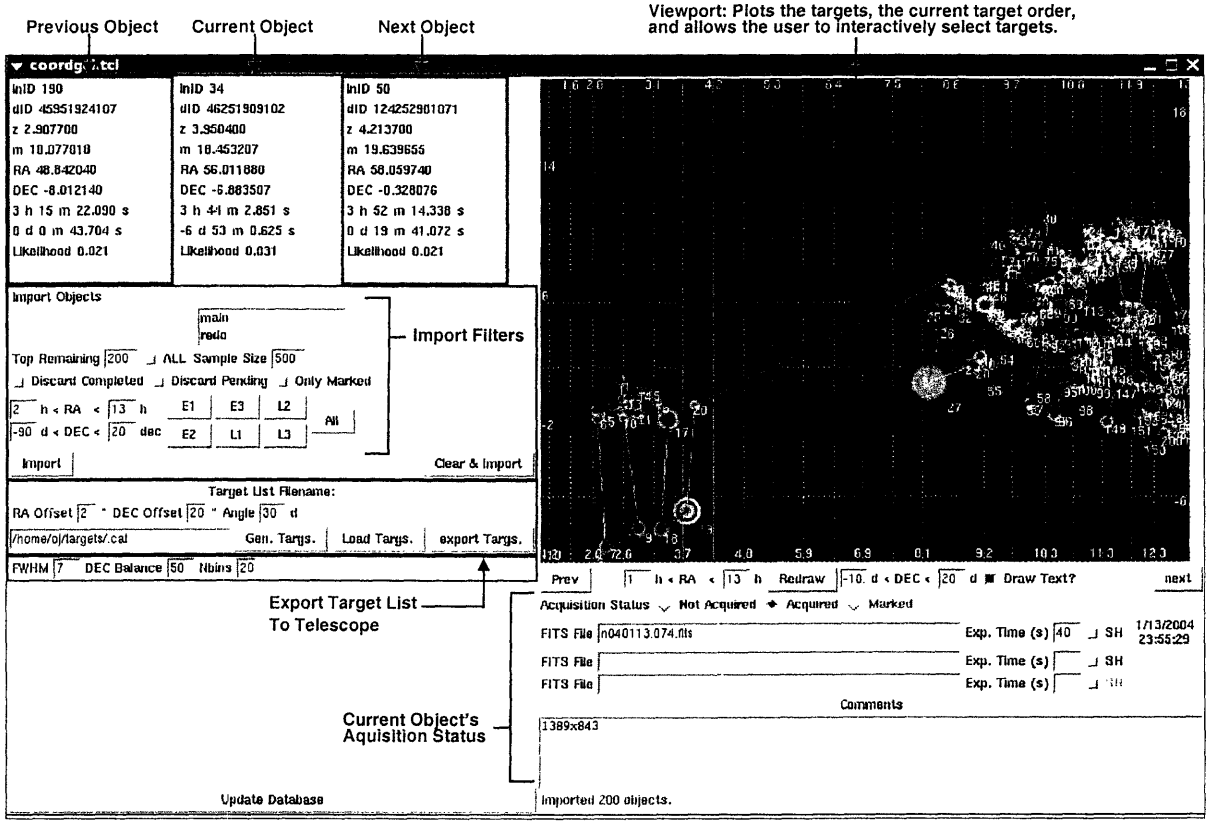


Figure C-1: The Coordinator.

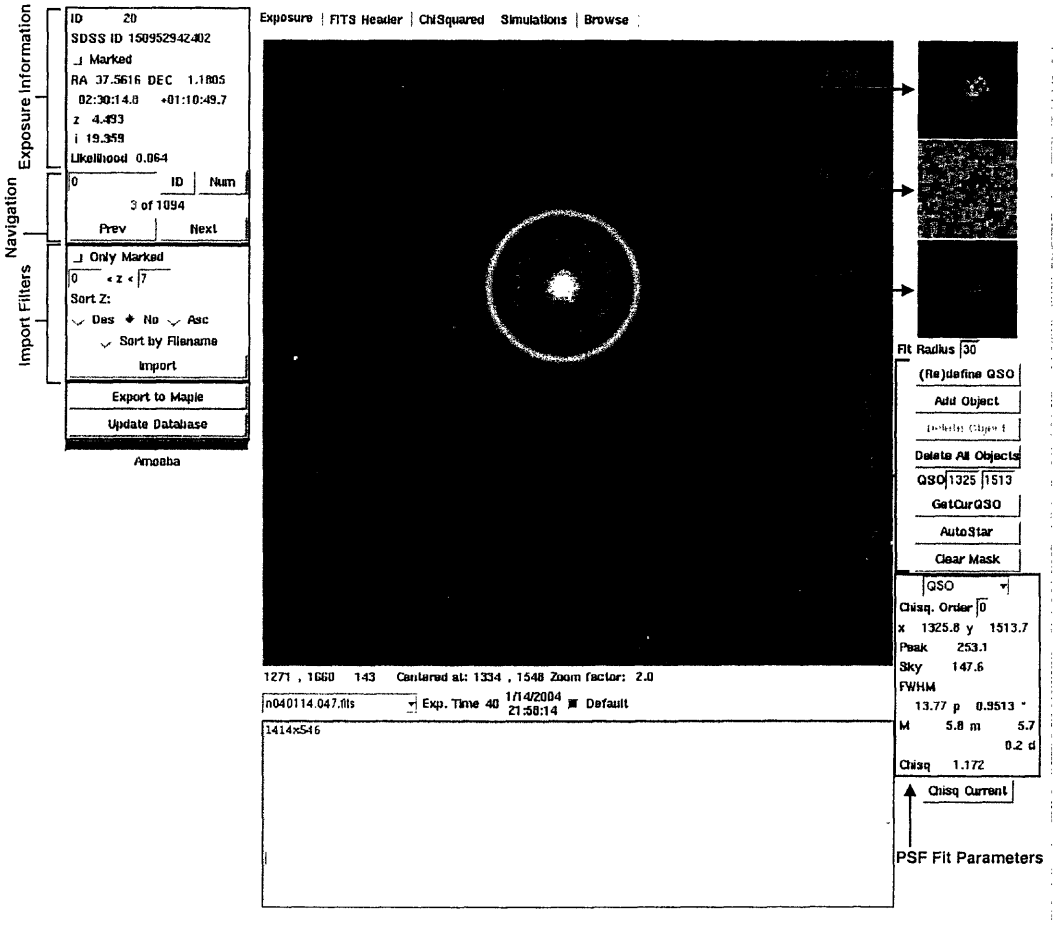


Figure C-2: A PSF fit in the Analyzer (exposure ID: 20). The three images in the top right corner are (from top to bottom) the selected target, the residue obtained by subtracting the target from the generated PSF, and the generated PSF. The resulting best fit parameters are presented in the lower right corner alongside the reduced $\chi^2$ of the fit.

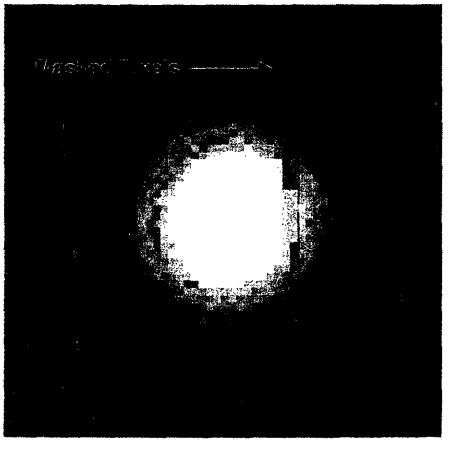

Figure C-3: An example of masking in the Analyzer.

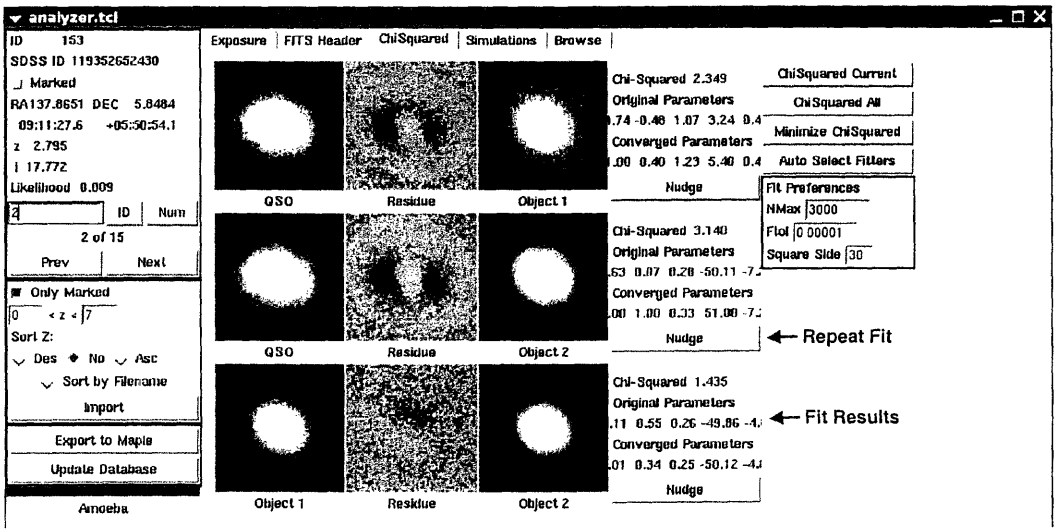


Figure C-4: The Analyzer's $\chi_\nu$ fitting interface.

# Bibliography


[Carroll] Carroll, B. W., and Ostlie, D. A., *An Introduction to Modern Astrophysics*, Addison-Wesley Publishing Company Inc. 1996

[Thorne] Thorne, K. S., *From Black Holes to Time Warps: Einstein's Outrageous Legacy*, W. W. Norton & Company, Inc. 199

[SDSS] Sloan Digital Sky Survey Data Release 2 Web Site: http://www.sdss.org/dr2/

[Pei] Pei, Y. C., The Luminosity Function of Quasars, ApJ **438**:623-631, January 10 1995

[Boyle et. al.] Boyle, B. J, Shanks, T., and Peterson, B. A. The evolution of optically selected QSOs. II, MNRAS **235**:935-948 December 1988

[Comerford et. al.] Comerford, J. M., Haiman, Z., and Schaye, J., Constraining the Redshift $z \approx 6$ Quasar Luminosity Function Using Gravitational Lensing, ApJ **580**:63-72, November 20 2002

[Madau et. al.] Madau, P., Haardt, F., and Rees, M. J., Radiative Transfer in a Clumpy Universe. III. The Nature of Cosmological Ionizing Sources, ApJ **514**: 648M 1999

[Wyithe and Loeb] Wyithe , J. S. B., and Abraham, L., Gravitational, Lensing of the Sloan Digital Sky Survey High-Redshift Quasars, ApJ **577**:57-68 September 20 2002

[Hogg] Hogg, D. W., Distance Measures in Cosmology, astro-ph/9905116

[Narayan and Bartelmann] Narayan, R., Bartelmann, M., Lectures on Gravitational Lensing

[Li and Ostriker] Li-Xin, L., Ostriker, J. P., Semianalytical Models for Lensing by Dark Halos I Splitting Angles, AJ **566**:652, February 20, 2002

[Mitchell] Mitchell, J. L., et. al., Robust Cosmological Constraints from Gravitational Lens Statistics astro-ph/0401138 v1, January 8 2004

[Bennett et. al.] Bennett, C. L., et. al., First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results, Accepted by ApJ, available at http://map.gsfc.nasa.gov/m_mm/pub_papers/firstyear.html

[A de Laix et. al.] A de Laix, A., and Vachaspati, T., Gravitational Lensing by Cosmic String Loops, Phys.Rev. **D54** (1996) 4780-4791

[PS] Schechter, P., Press, W., Method for Determining Maximum-Likelihood Distance Moduli for Groups of Galaxies,ApJ **203**:557-568, February 1 1976

[Wyithe] Wyithe, J. S. B., The Shallow Slope of the $z$ 6 Quasar Luminosity Function: Limits from the Lack of Multiple Image Gravitational Lenses, astro-ph/0308290

[Sheth et. al.] Sheth et. al., The Velocity Dispersion Function of Early-Type Galaxies, AJ **594**:225-231, September 1, 2003

[Sheth and Tormen] Sheth, R. K., and Tormen, G., Large Scale Bias and the Peak Background Split, astro-ph/9901122 v2

[Jenkins] Jenkins, A., et. al., The Mass Function of Dark Matter Haloes, MNRAS **321**, 372-384 (2001)

[Eisenstein and Hu] Eisenstein, D. J., and Hu, W., Power Spectra for Cold Dark Matter and its Variants, ApJ **511**:5-15, January 20, 1999

[Lacey and Cole] Lacey, C., and Cole, S., Merger Rates in Hierarchical Models of Galaxy Formation II. Comparison with N-body Simulations, astro-ph/9402069

[Barkana and Loeb] Barkana, R., Loeb, A., HIgh-Redshift Galaxies: Their Predict Size and Surface Brightness Distributions and their Gravitational Lensing Probability, ApJ **531**:613-623, March 10 2000

[Anderson et. al.] High-Redshift Quasars found in Sloan Digital Sky Survey Commissioning Data VI Sloan Digital Sky Survey Spectrograph Observations, AJ **122**:503-517,August 2001

[Magellan] Magellan Project and Instrumentation Information available on-line at: http://www.ociw.edu/magellan/

[Bevington] Bevington, P. R., Robinson, D. K., *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-HIll 2003

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

[McLaughlin] McLaughlin, M. P., Regress+ Appendix A: A Compendium of Common Probability Distributions, available online at http://www.causascientia.org/math_stat/Dists/Compendium.pdf

[CASTLES] C.S. Kochanek, E.E. Falco, C. Impey, J. Lehar, B. McLeod, H.-W. Rix), CASTLES Gravitational Lens Database, available online at http://cfa-www.harvard.edu/castles/

[Kochanek] Kochanek, C.S., The Analysis of Gravitatioanl Lens Surverys II. Maximum Likelihood Models and Singular Potentials. ApJ **419**:12-29, December 10 2003

[Schneider et. al.] Shneider, D. P., Schmidt, M., Gunn, J. E., A Study of Ten Quasars with Redshifts Greater than Four, AJ **98** (5)