MULTI-WAVELENGTH APPLICATIONS OF GRAVITATIONAL LENSING

BY ROSS FADELY

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Physics and Astronomy

Written under the direction of

Charles R. Keeton

and approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2010

ABSTRACT OF THE DISSERTATION

Multi-wavelength Applications of Gravitational Lensing

by Ross Fadely

Dissertation Director: Charles R. Keeton

Using an array of multi-wavelength data, we examine a variety of astrophysical problems with gravitational lensing. First, we seek to understand the mass distribution of an early-type galaxy with an analysis of the lens Q0957+561. We dissect the lens galaxy into luminous and dark components, and model the environment using results from weak lensing. Combining constraints from newly-discovered lensed images and stellar population models, we find the lens has a density profile which is shallower than isothermal, unlike those of typical early-type galaxies. Finally, using the measured time delay between the quasar images we find the Hubble constant to be $H_0 = 79.3^{+6.7}_{-8.5}$ km s^{-1} Mpc^{-1}.

One intriguing application of lensing is to exploit the lens magnification boost to study high-redshift objects in greater detail than otherwise possible. Here, we analyze the mid-infrared properties of two lensed $z \sim 2$ star-forming galaxies, SDSS J120602.09+514229.5 and SDSS J090122.37+181432.3, using *Spitzer*/IRS spectra to study their rest-frame $\sim 5 - 12\,\mu$m emission. Both systems exhibit strong polycyclic aromatic hydrocarbon (PAH) features in the spectra, indicating strong star formation

and the absence of significant AGN activity. For SDSS J090122.37+181432.3, this detection belies that inferred from optical measurements, indicating mid-IR spectroscopy provides key information needed to understand the properties of high-redshift star-forming galaxies.

While lensing provides measurements of the macroscopic properties of lens systems, it can also shed light on small-scale structure of galaxies. To identify and understand lens substructure, we examine the multi-wavelength properties of flux ratios for six lenses. Variations of the flux ratios with wavelength can be used to study the lensed quasars and the small-scale mass distribution of lens galaxies. We detect strong multi-wavelength variations in the lenses HE 0435-1223 and SDSS 0806+2006. For HE 0435-1223, we study its substructure with a series of lens models which add clumps of mass near the lensed images. We detect the presence of a clump near image A, with a mass of $\log(\frac{M_A(<R_{Ein})}{M_\odot\, h_{70}^{-1}}) = 7.68^{+0.92}_{-0.85}$. We also find support for a second clump, near image B, with mass $\log(\frac{M_B(<R_{Ein})}{M_\odot\, h_{70}^{-1}}) = 6.6^{+1.02}_{-1.52}$, although evidence for this clump is not decisive. Using Monte Carlo simulations, we connect these clumps to their associated populations, finding the mass fraction in substructure to be $f_{sub} \gtrsim 0.00092$.

Acknowledgements

I owe an enormous debt of gratitude to a large number of people, without whom this would not have been possible. Foremost, I would like extend my heartfelt thanks to my advisor Chuck Keeton. Over the course of my graduate career, Chuck has served as a constant and unwavering source of help, encouragement, and inspiration. His patience, coupled with his enthusiasm, brought forth my intellectual growth and greatly facilitated the learning process. As a scientist, his skill, demeanor, and knowledge provided a great model to aspire towards. I could not have asked for a better mentor, thank you. A special thanks is also owed to Andrew Baker, who served as my advisor for work presented in Chapter 4 here. Both in and out of the classroom, Andrew has broadened and deepened my knowledge of astrophysics and pushed me forward. Intellectually and professionally, I clearly would not be in the same place without his efforts and generosity.

I would like to thank all of the Rutgers astro group: Eric Gawiser, Jack Hughes, Saurabh Jha, Chuck Joseph, Terry Matilsky, Tad Pryor, Jerry Sellwood, and Ted Williams. Each one of you has played a role in my learning and growth over the years and made Rutgers a fantastic environment for a graduate student, thanks. I am grateful to my collaborator Gary Bernstein, who has graciously helped advance my career, and has been a pleasure to work with. A special thanks is also extended to Viviana Acquaviva, Eric Burgh, Art Congdon, Kris Eriksen, Felipe Menanteau, Naseem Rangwala, and Neelima Sehgal for being very helpful and supportive, and most of all for being great friends.

A long list of friends deserve my thanks, all of whom cannot be listed here. Instead,

let me especially thank PJ Bond, Josh Glawe, Jessica Hamilton, John Laudenberger, Tom Murchio, Greg Roehrich, Mike Schultz, and Kristin Smith for helping me through ups and downs and making the journey spectacularly fun. A deep thanks is also owed to my entire family whose love, support, and encouragement made this possible.

In addition to the above, I would like to acknowledge the helpful people and sources of support for various parts of this thesis. For work presented in Chapter 2, support was provided by grant HST-GO-10569, under NASA contract NAS5-26555, and additional support from the National Science Foundation grant AST-0747311. For Chapter 3, I am especially grateful to collaborators Sahar Allam, Huan Lin, Dieter Lutz, Alice Shapley, Min-Su Shin, J. Allyn Smith, Michael Strauss, and Douglas Tucker, and for support from NASA/*Spitzer* grant 1380986. Lastly, I acknowledge support from National Science Foundation grant AST-0747311 for work presented in Chapters 4 and 5.

Dedication

For my Parents: Lynne, Woody, Dennis, & Sharon.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Strong gravitational lensing is one of the most unique and versatile tools which can be used to study the universe. It can provide a telescope to study distant objects, a scalpel to dissect galaxies, and a ruler to measure cosmic distances. The range of observational probes can span decades in angular size ($10^{-6\prime\prime} \lesssim \theta \lesssim 10\prime\prime$) and in mass ($1M_{\odot} \lesssim M \lesssim 10^{14}M_{\odot}$), and can be used to shed light on the nature of dark matter and dark energy. Yet in spite of the great wealth of studies possible with strong lensing, a complete understanding of its applications is not possible without a synthesis between constraints provided by lensing and the knowledge and methods from other types of astrophysical studies. From light bending alone, we can learn about a limited number of quantities, like the mass of the deflector and the luminosity distribution of the source. However, interpreting and understanding these measurements hinges on connections from knowledge not derived from lensing, such as that from N-body simulations, dynamical measurements of galaxies, and studies of stars. In Section 1.1, I introduce and discuss various studies in which the union of lensing and other analyses jointly yield conclusions which are not possible with one method alone. Following this, Section 1.2 discusses the some of the methodology of lensing and statistics which are vital for detailed studies.

1.1 Astrophysical Studies Using Strong Gravitational Lensing

1.1.1 The Structure of Massive Galaxies

Strong gravitational lensing is a powerful probe of the mass distribution of galaxies. While able to infer the properties of mass distributions on large scales, strong lensing's best measure is the mass within an aperture defined by the Einstein radius. Independent of the profile of lens mass distributions, aperture masses from lensing are precise at the few percent level (Rusin et al. 2003), making aperture masses an excellent tool for studying the mass of galaxies (e.g., Bolton et al. 2007). Alone, aperture masses can provide a means to study the mass-density structure of ensembles of galaxies (Rusin et al. 2003; Rusin & Kochanek 2005). Combined with mass measurements on other scales (e.g., dynamical measurements), aperture masses can even test the fundamental scaling relationships of galaxies, such as the "fundamental plane" (Bolton et al. 2008).

While clearly powerful, aperture masses provide only a limited amount of information about the properties of lens galaxies. By using the positions (and often flux ratios) of lensed quasar images, or the surface brightness distribution of Einstein rings or arcs, lens modeling can go beyond aperture masses to obtain constraints on the angular and radial properties of lens profiles, albeit with modest uncertainties. Discussed below are some of the various measures possible of lens galaxy profiles and their components.

Radial Profiles of Lens Galaxies

A key quantity for understanding the mass distribution of massive galaxies is the inner logarithmic slope of the mass density profile, $\gamma' \equiv d\log\rho/d\log r$. For isothermal mass distributions, $\gamma' = -2$ and the associated circular velocity of a spherical model corresponds to a flat rotation curve. Intriguingly, studies based on local kinematics (e.g., Gerhard et al. 2001) and lensing (e.g., Kochanek 1995) indicate the mass distributions of massive, early-type galaxies are close to isothermal.

Joint analyses based on lensing and stellar dynamical measurements have confirmed this to remarkable precision. By combining these techniques, two independent mass estimations of the lens are obtained, dramatically improving constraints (e.g., Treu & Koopmans 2002). Using a lensing + dynamical analysis of 58 early-type lens galaxies, the Sloan Lens ACS (SLACS) survey found an average value of $\langle\gamma'\rangle = -2.085^{+0.025}_{-0.018}$ (Koopmans et al. 2009), with an intrinsic scatter in γ' around 10%. The SLACS value of γ', however, relies on dynamical constraints derived from an averaged, single aperture measurement of the stellar velocity dispersion. Connecting such measurements relies on a series of assumptions (e.g., on radial anisotropy), which may introduce bias in both the average and scatter. Nevertheless, more detailed studies using 2D dynamical information find $\langle\gamma'\rangle = -1.98 \pm 0.05$ (Barnabè et al. 2009), in agreement with other in-depth lensing studies (e.g., Kochanek 1995).

While very intriguing, the evidence for isothermality in ellipticals should still be considered tentative. Gravitational lenses are fairly rare and unique objects whose selection may be strongly biased, both in mass and in γ' (Mandelbaum et al. 2009). Furthermore, evidence exists that some lenses are inconsistent with isothermal profiles (Treu & Koopmans 2004; Kochanek et al. 2006, see also Chapter 2 here), indicating there may in fact be a significant population of ellipticals which deviate from isothermality.

Stellar and Dark Matter Distributions in Lens Galaxies

In lens systems with complementary datasets (e.g., dynamical or stellar population measurements), parameterizations of the lens potential need not be restricted to single, power law profiles. Instead, a useful exercise is to deconstruct the lens into its luminous and dark components. Doing so provides a unique opportunity to test the universality of dark matter profiles and to study stellar populations in massive galaxies.

Numerical simulations of cold dark matter halos generically conclude their mass density follows a profile similar to that proposed by Navarro et al. (1997) (NFW), with

an inner logarithmic slope of $\gamma'_{\rm DM} = -1$. Evidence from the kinematics of local galaxies and clusters, however, call the universality of halo profiles into question, indicating the true profile may be shallower than predicted (Salucci et al. 2007; Sand et al. 2008). In addition, cosmological simulations which include the effects of baryons seem to suggest shallow or flat-cored inner profiles might better describe the dark matter distribution (Romano-Díaz et al. 2009).

Typically, the Einstein radius of early-type lens galaxies falls within two effective radii of the luminosity distribution (see e.g., van de Ven et al. 2009). At such radii, the baryonic component of the galaxy's mass plays an important role, increasing the difficulty of extracting dark matter specific measurements (due to the uncertainty in the profile decomposition). A common approach, therefore, is to apply additional mass constraints from stellar dynamics or stellar population models. Using a joint lensing + dynamics analysis, Treu & Koopmans (2004) found the slope of the inner dark matter profile to be $\langle\gamma'_{\rm DM}\rangle = -1.3^{+0.2}_{-0.4}$, consistent with N-body predictions. In stark contrast, recent work combined lensing, stellar dynamics, and stellar population synthesis (SPS) models to analyze the two-source lens, SDSSJ1538+5817, finding $\langle\gamma'_{\rm DM}\rangle < 0.7$ (Grillo et al. 2010). Clearly, more studies with larger samples of lenses are needed.

Another approach to examining stellar versus dark matter content within a galaxy is to fix the dark matter profile (typically to NFW), and instead infer the properties of the stellar population (e.g., Auger et al. 2009). Using lensing (or lensing + stellar dynamics) in concert with SPS models, two separate estimators for stellar mass can be obtained; one lensing based, the other SPS based. Here, the lensing measurement provides an absolute calibration of the stellar mass. This lens-based measurement may then be used to test SPS masses, as a function of the population parameters (e.g., metallicity, formation history, IMF). Recent works have used lensing + dynamical models to constrain the stellar IMF, finding it similar in form to Salpeter (Grillo et al. 2009; Grillo & Gobat 2010; Treu et al. 2010). In addition, Treu et al. (2010) detect

a mismatch in the lensing vesus SPS derived steller masses which varies as a function of total lens mass, indicating either non-universal dark matter profiles or non-universal stellar IMFs. Caution must be taken, however, in interpreting SPS derived results. In particular, complications due to dust, varying star formation histories and metallicities, and treatments of late-stage stellar evolution carry additional uncertainties of $\gtrsim 0.3$ dex (Maraston 2005; Conroy et al. 2009; Maraston et al. 2009). In Chapter 2, we consider and discuss the effects of more complicated treatments of SPS models, in combination with lensing.

1.1.2 Dark Matter Substructure

With an understanding of the large scale, "macroscopic" mass distribution of a lens, it is interesting to consider the effects of granularity in mass distributions. One particularly interesting source of such structure is the large collection of satellite galaxies believed to reside in distant galaxies (Diemand et al. 2008; Springel et al. 2008). Motivated by studies of the Galactic satellite population, substructure gravitational lensing provides uniques tests of the small scale structure of galaxies.

Missing Satellites

Since the inception of N-body simulations of dark matter, a tension has arisen between predictions of Cold Dark Matter (CDM) theories and observational data. On the smallest scales, CDM simulations invariably predict that small dark matter "subhalos" must exist in large quantities within the virial radius of a massive galaxy (Moore et al. 1998; Klypin et al. 1999a). Furthermore, the mass function of subhalos is predicted to follow a power law of the form $dn/dM \propto M^{-\alpha}$ ($\alpha \sim 1.9$), such that Milky Way size galaxies should have hundreds of satellites with masses $< 10^7 M_\odot$ (Diemand et al. 2008; Springel et al. 2008). In contrast to these predictions, observational censuses of Galactic companions have historically been at odds with CDM, totaling only a few tens in number

(Klypin et al. 1999; Moore et al. 1999).

In recent years, the disparity between CDM and Galactic satellite counts (dubbed "The Missing Satellites Problem") has been alleviated in part by the greater sky coverage provided by the Sloan Digital Sky Survey (SDSS). By filtering the data for stellar clusters with appropriate color-magnitude and spatial properties, studies using SDSS have begun to detect new, ultra-faint dwarf galaxies surrounding the Galaxy (e.g., Willman et al. 2005). With this, the number of Milky Satellites has more than doubled the (pre-2005) count, rising from 11 to the current 35 (for a comprehensive list, see Wadepuhl & Springel 2010). Accounting for the sky coverage and magnitude limits of SDSS, the estimated number of luminous Galactic satellites is likely around $70 - 500$, depending on assumptions based on currently known dwarfs (Tollerud et al. 2008). Thus, the missing satellites problem may be much alleviated (or even solved) once a volumetrically complete survey of the Milky way is conducted.

In conjunction with improved detection of ultra faint dwarfs, suites of hydrodynamical simulations have further improved our understanding of the Missing Satellites Problem. Focusing on the physical processes at play in low mass halos, such studies have shown that a number of astrophysical mechanisms may hamper the formation of stars (for a review, see Kravtsov 2010). In particular, photo-ionization from hard UV photons, tidal disruption, supernova and AGN feedback, cosmic rays, and ram pressure stripping are amongst the processes which may render small mass halos bereft of light (e.g., Gnedin 2000; Scannapieco et al. 2001; Strigari et al. 2007; Madau et al. 2008; Mashchenko et al. 2008; Macciò et al. 2010; Penarrubia et al. 2010; Wadepuhl & Springel 2010). If the satellites surrounding the Milky Way are indeed sub- or non-luminous below some mass threshold, it is possible that the Missing Satellites Problem is in fact close to resolution. However, other solutions exist. If, for instancen dark matter particles are not "cold" with masses in the 10s $-$ 100s of GeV but is instead "warm" with masses in the 10s $-$ 100s of keV, the free streaming length of dark matter

would increase at early times. This has the effect of washing out bound structure on small mass scales, and would suppress the expected abundance of low mass satellites around the Milky Way (e.g., Yoshida et al. 2003; Gao & Theuns 2007). Studies of the smallest mass satellites, therefore, provide an unusual opportunity to examine the interplay of astrophysical effects (e.g., feedback, tidal disruption, photo-ionization) versus the effects of varying the properties of dark matter.

Substructure Gravitational Lensing

Deciphering the behavior of dark matter on small scales clearly requires a different approach than offered by traditional optical surveys, which are limited by the luminosity of low mass dwarfs. Sensitive to mass alone, gravitational lensing offers the only technique to study substructure outside of the Local Group, regardless of luminosity.

In the strong lensing regime, image deflections are on the order of $1''$ on the sky for typical redshifts of the lens galaxy and background source. For systems with two lensed images, the various observables (flux ratios, positions, and time delays) are well produced by a single, simple parametric model representing the mass distribution of the lens (e.g., Inada et al. 2007). These smoothly varying "macro" models often include effects of the lens environment, which manifests its effects in a tidal shear of the potential (Keeton et al. 1997).

In the early 1990s, several new four image lenses were discovered, and with them, the first failures of simple parametric models (Irwin et al. 1989; Kayser et al. 1990; Kormann et al. 1994; Lawrence et al. 1995). For these new four image systems, the addition of two more images greatly restricted the freedom available to lens models, increasing the difficulty of reproducing observations. In particular, it was quickly realized that simple models could adequately describe the image positions, but often failed to explain the image flux ratios. These so-called "flux ratio anomalies" presented the first significant challenges to lens modelers, indicating models lacked some key form of freedom.

One clear omission in simple lens models is the presence of small scale structure in the mass distribution. An obvious candidate for such modification, stellar perturbations to simple models were physically well-motivated and known for some time (Irwin et al. 1989; Vanderriest et al. 1989). Once added to models, the presence of stars cause extra light bending on small scales, resulting in changes to image magnifications. Indeed, the inclusion of stellar "microlensing" in models is able to explain some flux anomalies. For example, the quasar lens Q2237+0305 is known to exhibit optical flux ratios which differ by up to ~ 10 times those predicted from smooth macro models (Woźniak et al. 2000). A comprehensive explanation, especially across multiple wavelengths, requires significant microlensing (Agol et al. 2009).

In the case of radio lenses, however, microlensing is unlikely the origin of anomalous flux ratios. At masses around $1\,M_{\odot}$, typical stellar Einstein radii are $\sim 10^{-6}\,''$. In comparison, the size of the background radio source is $\gtrsim 1$ pc, or $\gtrsim 10^{-4}\,''$, much too large to be affected by such tiny perturbations. Instead, radio quads seem to require additional structure at scales of $\sim 10^{-3}''$, corresponding to $10^5 - 10^9 M_{\odot}$ in mass.

The enterprise of detecting dark matter substructure in lens galaxies via perturbations to image flux ratios has been quite successful. Early works (Mao & Schneider 1998; Chiba 2002; Metcalf & Madau 2001; Dalal & Kochanek 2002; Metcalf & Zhao 2002; Bradač et al. 2002) provided the initial theoretical understanding of substructure "millilensing" and the first detections in individual objects. Perhaps the most influential of these works, Dalal & Kochanek (2002) examined the properties of seven radio lenses known to exhibit flux anomalies. Their analysis of the combined sample provided the first constraints on the fraction of mass in substructure; between 0.6% and 7% within the Einstein radius of the lens galaxy (90% confidence).

Unfortunately, the utility of radio selected lenses for finding substructure is limited in scope. Currently, $\lesssim 35\%$ of known quasar lenses exhibit significant radio emission,

making a rare phenomenon even more scarce. More importantly, radio images are biased towards relatively large source sizes ($\gtrsim 10^{-4\prime\prime}$) and, therefore, larger substructure masses of $\gtrsim 10^5 M_\odot$. More recent works have explored the use of mid-infrared observations to study flux ratios (Chiba et al. 2005; MacLeod et al. 2009). With rest frame emission in the near-infrared, mid-IR flux ratios originate from the dusty torus of gas believed to surround the central engine of quasars (Minezaki et al. 2004; Suganuma et al. 2006). Depending on the intrinsic luminosity of the QSO, mid-IR sources are typically $\sim 10^{-5\prime\prime}$–$10^{-4\prime\prime}$ in size and are sensitive to yet lower mass substructure than radio lenses (Chiba et al. 2005; Minezaki et al. 2009). In Chapters 4 and 5, we consider the use of infrared flux ratios, in conjunction with measurements at other wavelengths, to examine the properties of dark matter substructure.

Perturbation of lens images are not solely limited to distortions of the magnification. If a clump of dark matter lies close to the lensed image, changes to the deflection can cause milliarcsecond-scale shifts in the image positions. With precise measurements from *HST* or VLA/VLBI observations, astrometric data have been shown to provide useful and complimentary constraints to image flux ratios (e.g., Chen et al. 2007; Williams et al. 2008; More et al. 2009). Moreover, astrometric changes to image positions open up the usefulness of galaxy-galaxy lenses for substructure studies. By modeling the distortions in high signal-to-noise Einstein rings, individual galaxy-galaxy lenses can be used to detect substructure with masses as low as $\sim 10^8 M_\odot$ (Koopmans 2005; Vegetti & Koopmans 2009b; Vegetti et al. 2009, 2010).

For quasar lenses, dark matter substructure can also be studied using the time delay between lensed images. In general, the time delay between images is determined by the smooth, macroscopic distribution of the lens mass. In detail, however, the exact value of the time delay is altered by granularity in the mass distribution. Under the presence of substructure, the time delay is scattered from the macroscopic prediction, with the amount of scatter dependent on the properties of the substructure population (e.g.,

mass fraction, power-law slope, mass range) (Keeton & Moustakas 2009). Interestingly, lensing observables (including flux and position anomalies) of substructure *all* depend on different moments of the mass function of the population (Keeton 2009). While we address some of these in Chapter 4, future studies will clearly benefit from the union of all three techniques.

1.1.3 Gravitational Telescopes

Strong lenses provide the opportunity to study faint objects at high redshifts. With improved spatial resolution and magnification boosts $\gtrsim 10$, lensing allows the possibility of detailed studies of distant objects, which might otherwise be beyond our current observational capabilities. In the case of galaxy-galaxy lenses, with source redshifts $z \sim 1-3$, investigations of the physical properties of young, star-forming are possible in much finer detail than for unlensed counterparts, permitting studies of their physical size (Marshall et al. 2007), chemical compostions (Stark et al. 2008), kinematics (Riechers et al. 2008), and star formation (Teplitz et al. 2000).

Star Forming Galaxies at High Redshift

Key questions for understanding the star formation and evolution of high redshift $z \sim 2-3$ galaxies include: How rapidly are stars being formed? How large is the gas reservoir from which they can be formed? What are the physical conditions (i.e., metallicity, spectral hardness of radiation, etc.) present in high redshift ISMs? What role do AGN play in the energetics of distant galaxies? Some of these questions are addressable through rest-frame UV measurements of unlensed sources, albeit often at the high luminosity tail of galaxies (e.g., Reddy et al. 2006a). Access to the rest-frame optical, and even more so to the rest-frame infrared, is hampered by the difficulty of conducting observations redward of $1\,\mu$m, and often relies on stacking techniques for all but the brightest objects (Reddy & Steidel 2004; Reddy et al. 2006b). This

is particularly unfortunate since the bulk of the luminosity lies in the infrared, which probes the gas and dust content of star forming galaxies.

Fortunately, lensing can provide the helping hand necessary to study the gas and dust content in galaxies with sub-maximal luminosities. One extensively studied system is the lensed $z = 2.73$ Lyman break galaxy (LBG), MS 1512-CB58 (Yee et al. 1996). Unlike non-lensed LBGs, MS 1512-CB58 seems to exhibit an infrared luminosity $L_{\rm IR} \sim 10^{11} L_{\odot}$ and dust extinction of $0.0 < E(B-V) < 0.03$, similar to that of typical luminous infrared galaxies in the local universe (Baker et al. 2001; Smail et al. 2007; Siana et al. 2008). Interestingly, the star formation rate (SFR) of MS 1512-CB58 is a factor $3-5$ higher when estimated from rest frame UV versus infrared, suggesting dust reddening may be behaving differently at high redshift (Siana et al. 2008). This deviation from standard Calzetti extinction seems to be confirmed in the "Cosmic Eye", a lensed LBG at $z = 3.07$ (Siana et al. 2009).

In addition to estimations of the SFR, mid-IR observations of lensed high redshift systems have also measured the polycyclic aromatic hydrocarbon (PAH) emission present in these young galaxies (Siana et al. 2008, 2009, also Chapter 3 here). PAH emission originates from the reprocessing of more energetic photons in the radiation field, and is an indicator of rapid star formation. One caveat exists, however – if the spectrum of the radiation is very hard, PAH molecules are destroyed (see, e.g., Wu et al. 2006). Thus, PAH emission serves as a diagnostic of the energetics present in interstellar media, and can be used to distinguish between AGN versus star formation dominated galaxies (Laurent et al. 2000; Baum et al. 2010). So far, all mid-IR spectra of lensed LBG galaxies have shown significant PAH emission. While still heterogenous in properties like metallicity, SFR, and gas content, the uniform presence of strong PAH emission suggests AGN play a sub-dominant role in these systems (see Chapter 3). Unfortunately, the current sample of lensed galaxies with mid-IR observations is small and will likely not increase much until the launch of the *James Webb Space Telescope*.

Quasar Structure

To date, the majority of our understanding of AGN structure have been derived from local Seyfert galaxies, with luminosities $\lesssim 10^{45}\,\mathrm{erg\,s^{-1}}$ (Urry & Padovani 1995; Kaspi et al. 2005; Suganuma et al. 2006). At higher luminosities, AGN are quite rare in the local universe until redshifts $\gtrsim 1$ are reached. At these cosmological distances, understanding the properties and structure of highly luminous AGN (quasars) is hindered by their small angular size and (often) faint nature. Lensing, however, not only provides a magnification boost but also has observables which are dependent on the angular size of the background source (Dobler & Keeton 2006; Dai et al. 2010). In particular, the same small structure which is necessary to explain flux anomalies can also be used to study AGN, since it is sensitive the size of the emitting region.

At rest frame UV to optical wavelengths, AGN emission originates thermally from a hot accretion disk surrounding a supermassive black hole (Shakura & Sunyaev 1973). For quasars, the expected size of typical accretion disks are $10^{-3} - 10^{-2}$ pc, corresponding to angular scales of $10^{-7\prime\prime} - 10^{-6\prime\prime}$ (Pooley et al. 2007; Morgan et al. 2010). In the case of strong lensing, however, the Einstein radii of lens galaxies are of the order $\sim 1''$. The huge discrepancy in these sizes means that, effectively, QSOs are point sources for lensing in X-ray through optical wavelengths, and their sizes have no effects on lens observables. However, this is not true in the case of microlensing by stars. With Einstein radii of order $10^{-6\prime\prime}$, stars can alter the magnification of lens accretion disks. Moreover, since microlensing effects are comparable to the source size of disks, variations from microlensing can be used to model the size of the emission region (Dai et al. 2010; Poindexter & Kochanek 2010). By studying this effect across multiple wavelengths, microlensing studies have found that disk sizes are somewhat larger than expected from Shakura & Sunyaev (1973), with somewhat under-luminous brightnesses (Poindexter & Kochanek 2010; Morgan et al. 2010). In lenses with extensive monitoring, variations of the source size have been shown to provide improved fits to the data, unveiling the

act of accretion in progress (Blackburne & Kochanek 2010).

Just as stellar microlensing is sensitive to the size of the accretion disk of the background AGN, substructure millilensing is equally sensitive to other parts of the AGN structure. Substructure with mass ranges of $10^3 - 10^{10} M_\odot$ have Einstein radii of order $\sim 10^{-4\prime\prime} - 10^{-2\prime\prime}$, and are insensitive to the size of accretion disks (Minezaki et al. 2009). However, more extended structures like the surrounding dust torus and broad/narrow line regions have sizes which are comparable to the Einstein radii of substructure clumps (e.g., Chiba et al. 2005). Thus, substructure studies using infrared or broad-line flux measurements can not only address flux ratio anomalies, but can also be used to study AGN structure through the size dependence of millilensing. In Chapters 4 and 5, we explore these possibilities.

1.2 Methodology

1.2.1 Basic Gravitational Lensing Theory and Notation

In this section, the basic description and notation of strong gravitational lensing is described. For more detailed and in depth treatments of lensing theory, some excellent reviews are presented in Schneider et al. (1992) and Kochanek (2004).

Figure 1.1 presents the basic geometry for strong lensing, corresponding to the presence of multiple images. We first define the optical axis as the line connecting the observer to the center of mass in the lens plane. Relative to this line, we define the angular position of the images, $\mathbf{x} = (x, y)$, and angular position of the source, $\mathbf{u} = (u, v)$. The light bending which connects source to the images is then defined by the lens equation:

$$\mathbf{u} = \mathbf{x} - \nabla\phi \tag{1.1}$$

where ϕ is the scaled gravitational potential. In order to relate the angular positions

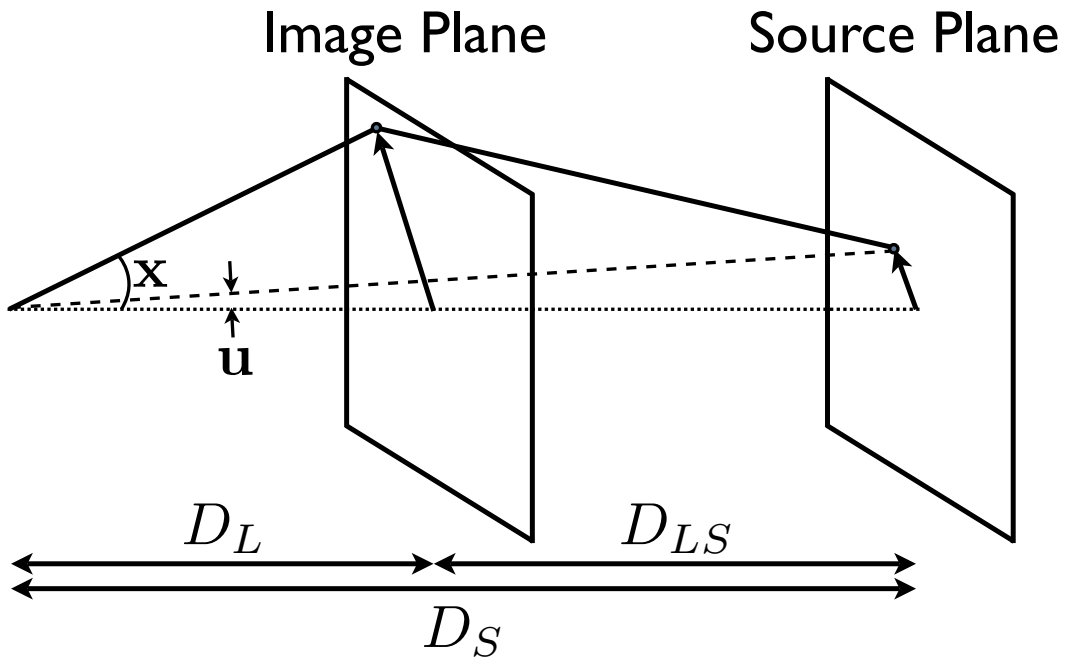

Figure 1.1 The basic geometrical scheme for strong lensing. The dotted central line, called the optical axis, represents the vector connecting the observer to the center of the lens mass. A background source lies at an angular position **u** from the optical axis. Light from the source is deflected by the lens along the solid line, forming images at an angular position **x**. Angular diameter distances D_L, D_{LS}, and D_S correspond to the distance from the observer to the lens, from the lens to the source, and from observer to the source, respectively.

in Equation 1.1 to physical sizes, it is necessary to know the relative angular diameter distances D_L, D_{LS}, D_S, which correspond to the distance from the observer to the lens, from the lens to the source, and from the observer to the source, respectively.

For a lens with surface mass density Σ, the potential ϕ is defined by

$$\nabla^2\phi(\mathbf{x}) = 2\kappa(\mathbf{x}) = 2\Sigma/\Sigma_{crit}, \tag{1.2}$$

where κ is the convergence and $\Sigma_{crit} = \frac{c^2}{4\pi G}\frac{D_S}{D_L D_{LS}}$ is the critical density for lensing. In the case of strong gravitational lensing, the criterion for multiple images is $\kappa \gtrsim 1$ or $\Sigma \gtrsim \Sigma_{crit}$.

While the lens equation (Equation 1.1) defines the positions of lensed images on the sky, another key observable is the image flux ratios, which are determined by the magnification. Since lensing conserves surface brightness, we can characterize the magnification using the matrix:

$$\mathcal{A} = \frac{\partial \mathbf{u}}{\partial \mathbf{x}}. \tag{1.3}$$

From Equation 1.1, it follows that

$$\mathcal{A} = \begin{bmatrix} 1 - \phi_{xx} & -\phi_{xy} \\ -\phi_{yx} & 1 - \phi_{yy} \end{bmatrix} \tag{1.4}$$

In turn, we can define the magnification matrix as $\mathbf{M} = \mathcal{A}^{-1}$, such that the magnification of a lensed image is then

$$\mu(\mathbf{x}) = \frac{1}{\det \mathcal{A}} \tag{1.5}$$

With a lens model in hand we can now define the image positions and magnifications for any position of the background source. Equation 1.5 indicates that certain positions have divergent magnifications, where $\det \mathcal{A} = 0$. The locations of such points are referred to as "critical curves" in the image plane and "caustic curves" in the source plane. In Section 1.2.2 examples of typical critical and caustic curves are given along with a discussion of their utility.

For images of lensed quasars, the flux observed may change over time due to variations in brightness of the background AGN. While this complicates the interpretation of flux ratios, it presents an additional observable through the differences in light travel time between images. Differences in the arrival time between images originate from two effects. The first is the difference in the path length between images and the second is

the gravitational delay from the Shapiro effect. Combining these two effects, the time delay between images is then

$$\Delta t_{ij} = \frac{1+z_l}{c}\frac{D_L D_S}{D_{LS}}\left\{\frac{1}{2}\left(|\mathbf{x}_i-\mathbf{u}|^2-|\mathbf{x}_j-\mathbf{u}|^2\right)-\left[\phi(\mathbf{x}_i)-\phi(\mathbf{x}_j)\right]\right\}. \tag{1.6}$$

1.2.2 Lens Configurations

As an example, let us consider a lens mass distribution corresponding to a singular isothermal sphere (SIS). While a simple "toy" model, the SIS profile has been used extensively in lensing since it easy to understand, and can be solve analytically. Additionally, SIS profiles are a somewhat fair treatment of lens galaxies, since most are found to be close to isothermal (Koopmans et al. 2009).

Behind this SIS lens, let's consider a background source with a spherical, Gaussian surface brightness distribution. Figure 1.2 shows the configuration of the lensed images and background source. Overplotted in the panels are the critical and caustic curves, defining the regions of very high magnification and distortion. In the top row, the source sits on the optical axis and produces an Einstein ring (coincident with the critical curve). In the lower panels, the background source is moved away from the optical axis and image arcs are produced.

Realistic lenses, however, are never completely spherical. Ellipticity in the lens mass distribution, as well as tidal shear produced by the surrounding environment, destroy the symmetry. In Figure 1.3, the lens is now replaced by a singular isothermal ellipsoid (SIE), with the same total mass as the lens in Figure 1.2. Once spherical symmetry is broken, the inner, "tangential" or "astroid" caustic can been seen. As the lens becomes increasingly elliptical, the tangential caustic grows. When the size of the tangential caustic is bigger than the size of the source, an Einstein ring is no longer produced, resulting in four lensed images.[1]

[1]Formally, five lensed images are produced. However, the fifth image lies very close to the optical

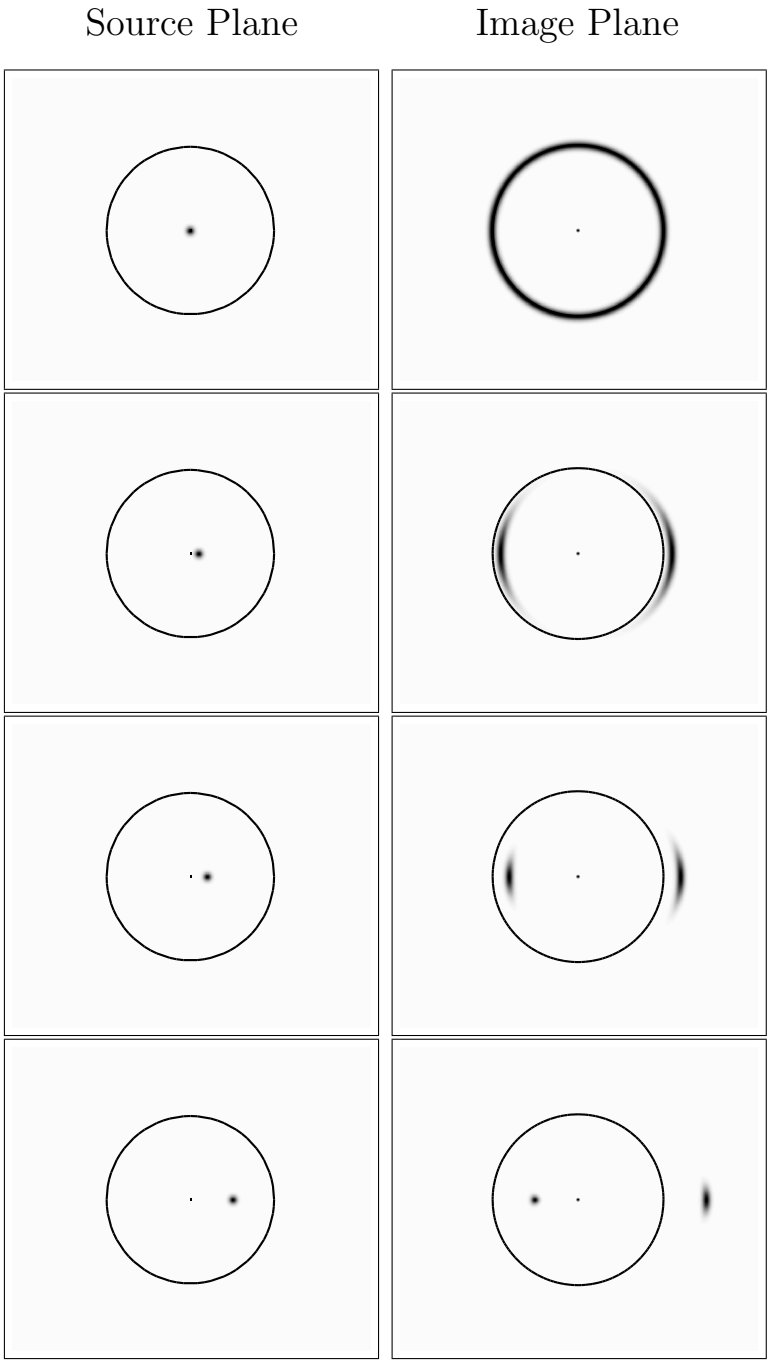

Figure 1.2 Lensing by a singular isothermal sphere (SIS). The left hand column represents the source plane, with a small spherical, gaussian source visible. The right hand column shows the corresponding images. The solid black lines represent the caustics in the source plane, and the critical curves in the image plane. In the top row, the source sits on the optical axis and a perfect Einstein ring is formed. As the source is moved away (lower rows), the ring is broken into two arc images.

Figure 1.4 shows the effects of moving the source within the tangential caustic. Typical lens configurations called "folds", "crosses", and "cusps" are produced. The sign of the eigenvalues of the magnification matrix (Equation 1.4) determines the parity of the lensed images, and is useful in describing the lensed images. For instance, images of opposite parity straddle the critical curve. Additionally, positive parity images correspond to minima of the potential, negative parity images correspond to saddle points, and double negative parity images correspond to maxima. From top to bottom in Figure 1.4, the source position is varied. As the source moves outside of the tangental caustic, the number of images decreases by two and a two image configuration is produced.

1.2.3 Statistical Analysis in Lensing

A main task in astrophysical studies is to connect a set of observations to a physically justified model. In strong lensing studies, this involves taking observables (like image positions, fluxes, and time delays) and using them to infer properties of the projected mass distribution and the background source. Practically, this involves constructing a model for the lens + source, computing the properties of the images, and comparing them to the data. The ultimate goal of this procedure is to assess the likelihood of the model and estimate range of acceptable model parameters.

The simplest estimator of the likelihood of a model is the chi-squared goodness-of-fit statistic:

$$\chi^2 = \left(\frac{x_M - x_d}{\sigma_{x_d}} \right)^2 \tag{1.7}$$

where x_d is the data values, x_M is the model values, and σ_{x_d} is the uncertainty associated with the data values (see, e.g., Wall & Jenkins 2003). The χ^2 statistic may

axis and is usually highly demagnified, making it hard to detect.

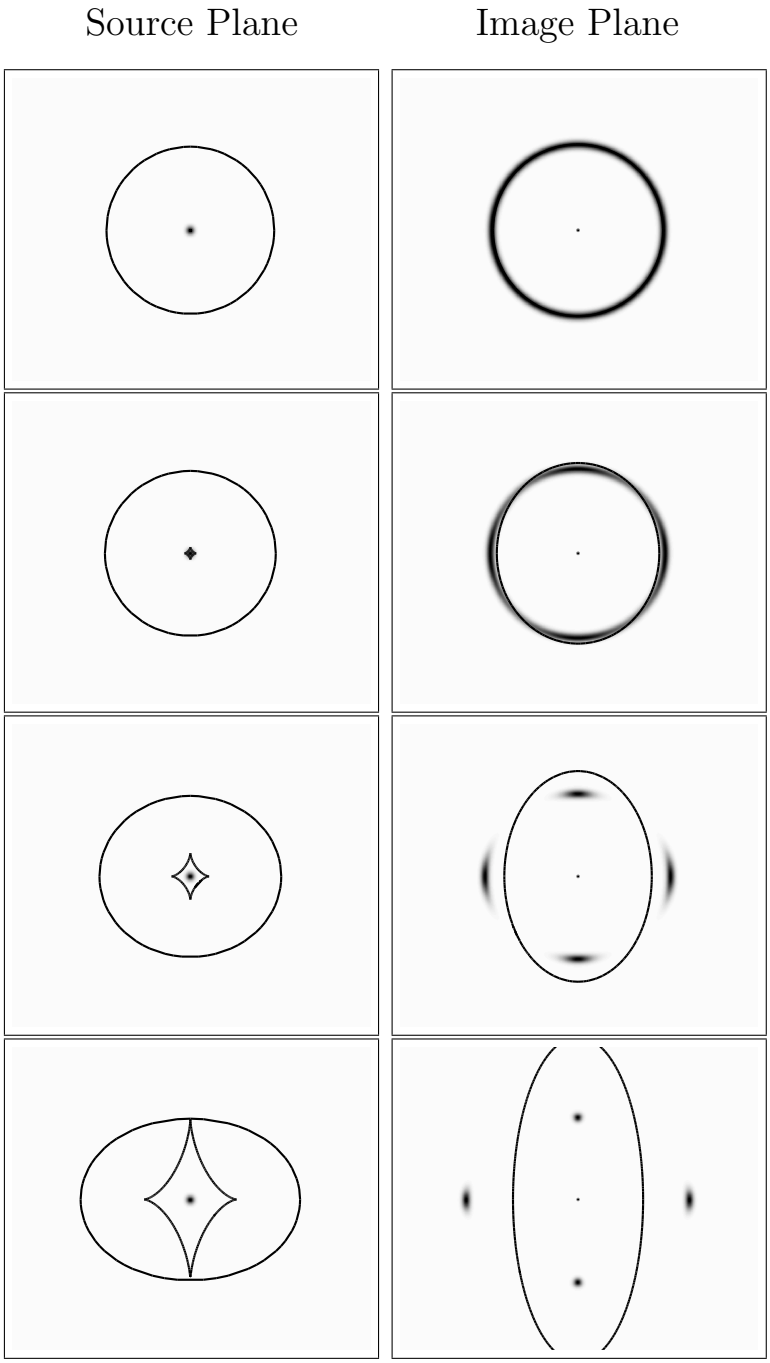

Figure 1.3 Lensing by a SIE mass distribution. The left hand column represents the source plane, with a small spherical, gaussian source visible. The right hand column shows the corresponding images. The solid black lines represent the caustics in the source plane, and the critical curves in the image plane. From top to bottom, the ellipticity of the lens corresponds to 0.01, 0.1, 0.3, 0.6. At low ellipticity, the tangential (astroid) caustic is small compared to the source, and an Einstein ring is formed. As the ellipticity increases, the tangential caustic grows and a ring can no longer be formed.

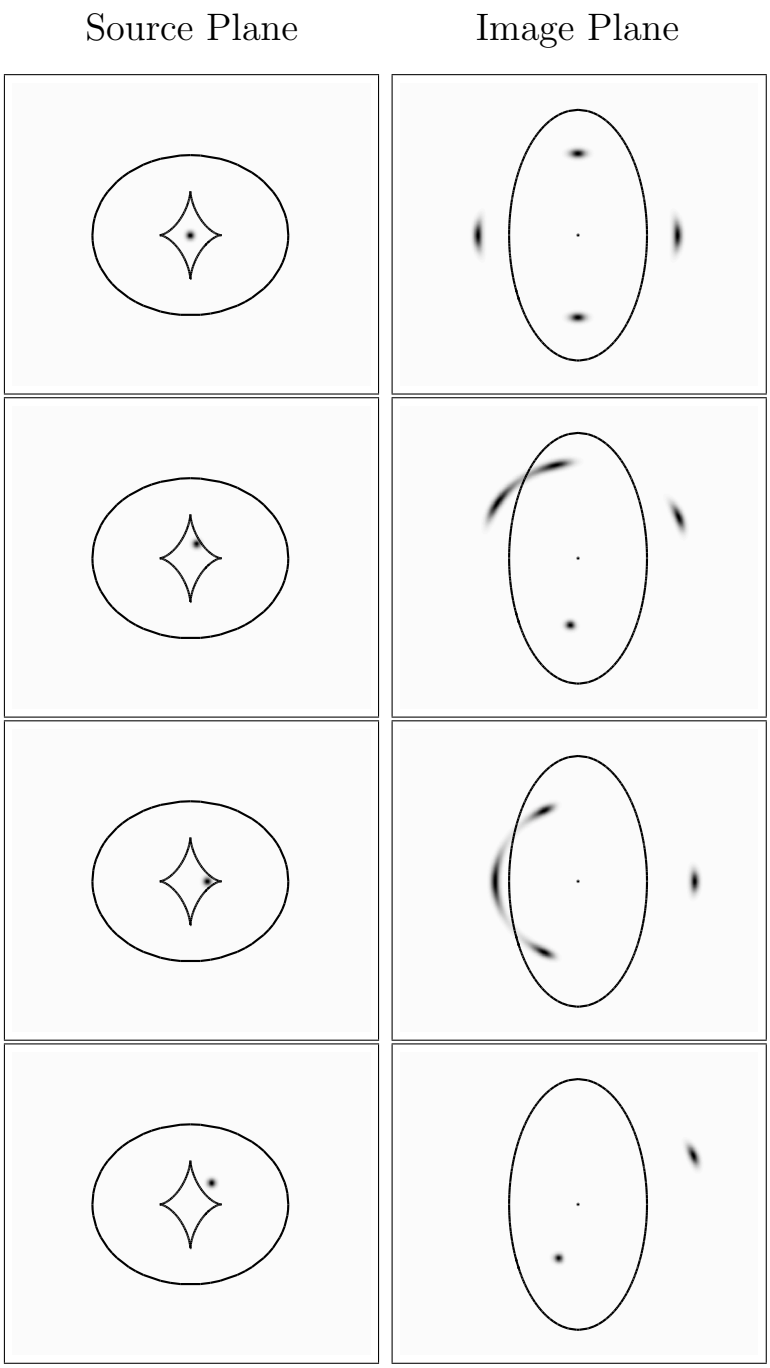

Figure 1.4 Typical lens configurations, produced by a SIE mass distribution. The left hand column represents the source plane, with a small spherical, gaussian source visible. The right hand column shows the corresponding images. The solid black lines represent the caustics in the source plane, and the critical curves in the image plane. In the first three rows, common four-image configurations referred to as "cross", "fold", and "cusp" lenses, are shown from top to bottom. The bottom row shows an example of a two-image lens. The position of the source, relative to the caustics, determines the configuration. As the source crosses a caustic, the number of images decreases by two.

be used to estimate the likelihood $\mathcal{L}$ of the model, with $\mathcal{L} \propto e^{-\chi^2/2}$. Thus, by using standard optimization routines one can easily compute the peak value of the likelihood by computing χ^2 for the best fitting model. If the likelihood is Gaussian in form, well fitting models should have a *reduced* $\chi^2_r \sim 1$ (Wall & Jenkins 2003).[2]

While the χ^2 statistic gives a good estimate of the peak of the likelihood, the task of estimating the range of allowed model parameters is not as simple. For this we want to know the posterior probability distribution $P(\theta|d, M)$ of parameters θ, given the data d and model M.

Parameter Estimation with Markov Chain Monte Carlo

Bayesian probability states

$$P(\theta|d, M) = \frac{P(d|\theta, M)P(\theta|M)}{P(d|M)} \tag{1.8}$$

$$= Z^{-1}P(d|\theta, M)P(\theta|M) \tag{1.9}$$

where the normalization

$$Z = \int P(d|\theta, M)P(\theta|M)d\theta. \tag{1.10}$$

Since Z does not change the shape of $P(\theta|d, M)$, we can ignore it for the purpose of parameter estimation. Therefore, what is needed is a technique which can compute the combination of the likelihood $\mathcal{L} = P(d|\theta, M)$ and the prior distribution $P(d|M)$.

Markov Chain Monte Carlo (MCMC) provides an accurate and relatively efficient method of sampling the shape of the posterior distribution (see, e.g., Gelman et al. 2003). The basic premise of MCMC is to use a random walk to generate many samples of points which represent the desired distribution. The generic algorithm starts with

[2]The reduced chi-squared, χ^2_r, is simply that of Equation 1.7 divided by the number of degrees of freedom ($N_{data} - N_{parameters}$).

a Markov chain at a random point in the parameter space. The chain is then stepped randomly in the space and the value of the likelihood times the prior is calculated in the form of a χ^2 value. If the χ^2 goes down the step is accepted as the new location of the chain. If it goes up, the step may be accepted with a probability $P = e^{-\Delta\chi^2/2}$. As the chain continues, it steps through the posterior distribution, eventually moving from lower to higher likelihood regions (see Figure 1.5). To insure a robust sampling, multiple such chains are run and compared to see if they recover similar properties of the target distribution. Final convergence is reached once the multiple chains have similar variances between and within themselves. See Chapter 2 (Section 2.3.4) for details regarding MCMC algorithms, convergence, and methods.

Model Evidence and Nested Sampling

Standard chi-squared analysis allows one to determine how well a model can fit the data. The next level of inference, the estimation of model parameter ranges, is tractable with Monte Carlo techniques (e.g., MCMC). Difficulties arise, however, when there are multiple models which can describe a given set of data. In particular, models with different priors, different numbers of parameters, or disparate properties can often provide similar and adequate fits. Which model is then judged to be more preferable?

Model comparison relies on a true calculation of the posterior distribution. Earlier in this section, it was sufficient to ignore the normalization of the posterior, since our focus was on single model fits to the data. To compare various models, however, this normalization is needed since it provides a scale of the overall probability of the model. This normalization, Z, is called the Bayesian Evidence. From Equation 1.10, the evidence is the combination of the likelihood and the priors, integrated over the volume of parameter space. Thus, the evidence is a measure not only of the peak of the likelihood but how broadly a model fits the data for the range of parameters. With the evidence in hand, models can directly be compared through the ratio of their evidences,

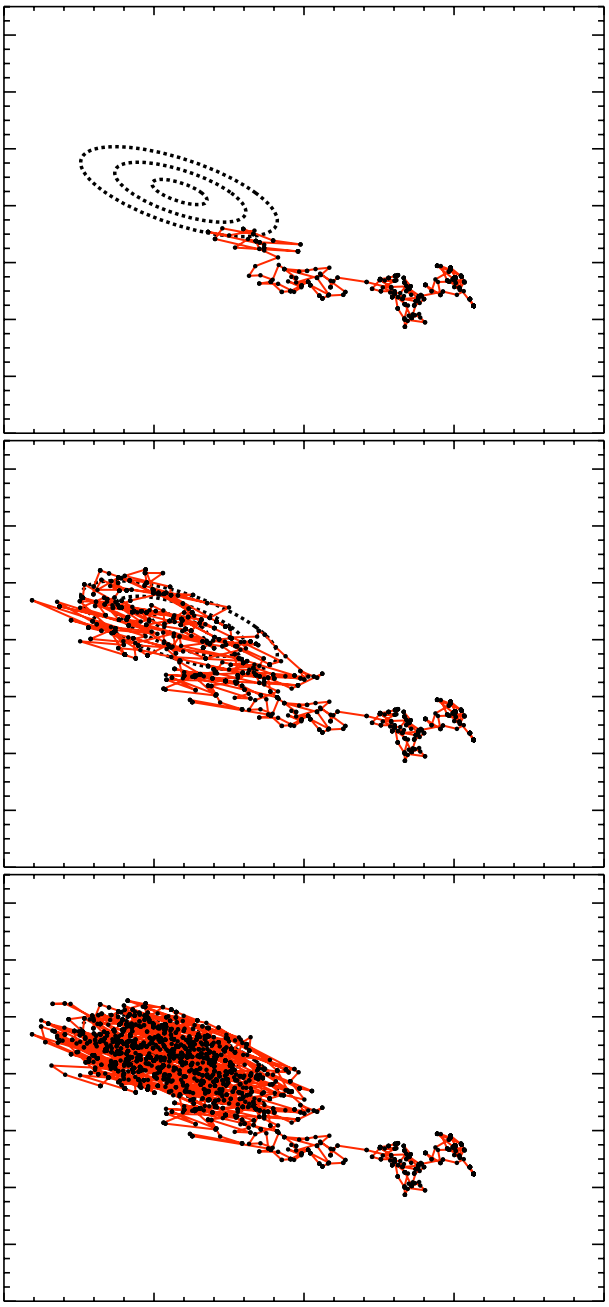

Figure 1.5 Three panels showing the various stages of a MCMC exploration of a joint posterior distribution. In this example, a single chain (red) makes steps in the parameter space defined by the x and y axis here. Steps of the chain (black points) are accepted if the χ^2 value decreases. If the χ^2 increases, the step is taken with probability $P = e^{-\Delta\chi^2/2}$. The dashed lines represent various confidence intervals of the target distribution. Initially (top panel) the chain is far from the region(s) of high likelihood. Though any particular step can move up or down in likelihood, eventually the chain moves from lower to higher likelihood regions (middle panel). At the end of the MCMC run, the chain has thoroughly sampled the posterior (bottom panel). For a given MCMC analysis, multiple such chains are run to check convergence and to provide large numbers of samples.

called the Bayes Factor (see example case below).

Many techniques exist which can calculate the Bayesian Evidence for a particular model. One particularly useful method is called Nested Sampling (Skilling 2004). Let us write

$$Z = \int P(d|\theta, M)P(\theta|M)d\theta \tag{1.11}$$

$$= \int \mathcal{L}(X)dX \tag{1.12}$$

where $dX = P(\theta|M)d\theta$ and $X = \int_{\mathcal{L}>\mathcal{L}'} P(\theta|M)d\theta$ is the prior volume within an iso-likelihood surface with likelihood $\mathcal{L}'$. Since $\mathcal{L}(X)$ is monotonically decreasing with X, we can then consider a numerical evaluation of Equation 1.12 on a sequence of decreasing values of X:

$$0 < X_N < ... < X_1 < X_0 < 1. \tag{1.13}$$

With this, the integral of Z is simply a sum $Z = \sum_{i=1}^{N} \mathcal{L}_i w_i$, where the weights w_i are given by, for example, the trapazoid rule $w_i = \frac{1}{2}(X_{i-1} - X_{i+1})$. Figure 1.6 provides an illustration of the Nested Sampling scheme.

In practice, the algorithm is executed in the following way. First, a Monte Carlo sample of the full prior volume is drawn with N 'live' points, defining the initial prior volume X_0 which is set to 1. The initial sample is then sorted such that the point of *lowest* likelihood is selected to defined the first 'nested' point, and is replaced by another point drawn such that $\mathcal{L}_{\rm new} > \mathcal{L}_{\rm old}$. This procedure is repeated, with the remaining amount of unexplored prior volume decreasing towards zero. A running estimate of the evidence is calculated from the nested points ($Z_{\rm nest} = \sum_{i=1}^{\rm Current} \mathcal{L}_i w_i$) and of the live points ($Z_{\rm live} \sim \langle \mathcal{L}_{\rm live} \rangle X_i$). Convergence is achieved when $Z_{\rm nest}$ is within some tolerance of $Z_{\rm total} \sim Z_{\rm nest} + Z_{\rm live}$. In all, calculating the total evidence is achieved by traversing

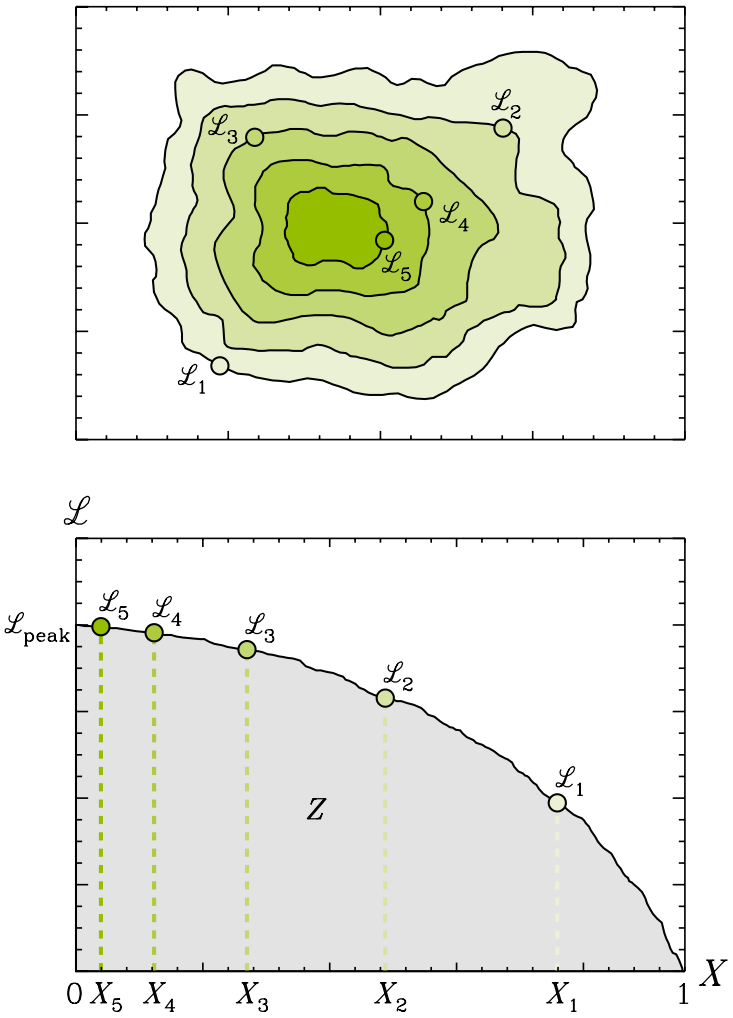

Figure 1.6 *Top panel* A hypothetical joint posterior probability distribution of a 2D parameter space. As the Nested Sampling algorithm runs, live points sample the remaining fraction of prior volume X_i (denoted by shaded regions). The live point with the lowest likelihood $\mathcal{L}_i$ (colored circles) then defines the iso-likelihood surface (black lines) which corresponds the particular volume X_i. This procedure is repeated many times, with $X_i \to 0$. In this panel, five such iterations are shown. *Bottom panel* The information in the top panel, projected into the function $\mathcal{L}(X)$. As the algorithm proceeds, $\mathcal{L}_i$ is defined for each X_i, producing the black curve. The Bayesian evidence Z is then the integral of the $\mathcal{L}(X)$ curve. In this example, it is clear that (even at fixed peak likelihood) the evidence may be dramatically higher/lower if the likelihoods $\mathcal{L}_i$ are higher/lower for each X_i.

the prior volume in a series of nested shells (see Figure 1.6). Below, we demonstrate the utility of calculating the evidence by examining the lens PG 1115+080.

Example Case: PG 1115+080

Consider the case of the lens PG 1115+080, a four image quasar lens which has been well studied previously (see e.g., the CASTLES database[3]). Using positions and $F814W$ flux ratios (with 25% uncertainties) from the CASTLES website, we can model the lens using a projected mass distribution of the form:

$$\kappa(\xi) = \frac{1}{2}\frac{b^{2-\alpha}}{(a^2+\xi^2)^{1-\alpha/2}} \tag{1.14}$$

where $\xi = (x^2 + y^2/q^2)^{1/2}$ is the ellipse coordinate with axis ratio q, a is the core radius, b is a normalization parameter, and α is the power law index of the projected mass profile. With additional terms which represent the environmental shear, we consider the above model for $N_{data} - N_{parameter} = 12 - 8 = 4$ degrees of freedom, leaving either a fixed to zero (Model 1) or α fixed to 1 (Model 2). A simple χ^2 optimization finds the two models the models both have $\chi^2_{\rm best} = 3.2$, indicating both models can fit the data well. Which model is then more favorable? Without a prior preference to a particular model, there is no means with which to determine which model is better from the χ^2 alone. Instead, we aim to calculate the posterior distribution and the Bayesian Evidence using the Nested Sampling algorithm described above. We find that Model 1, with a fixed core radius and varying profile index, has log(Evidence) $= -17.1$. Model 2, fixed to an isothermal ($\alpha = 1$) profile but with a varying core radius, has log(Evidence) $= -18.1$. Computing the Bayes Factor, we find Model 1 is 10 times more likely than Model 2. Apparently, the data considered for PG 1115 strongly favors a model which has a running power law index. Without the evidence, this conclusion would not have been

[3]http://www.cfa.harvard.edu/castles/

possible. In Chapter 5, we apply the Nested Sampling technique in analyzing HE 0435-1223, finding it an invaluable tool in understanding small scale structure in the lens galaxy.

Bibliography

Agol, E., Gogarten, S. M., Gorjian, V., & Kimball, A. 2009, ApJ, 697, 1010

Auger, M. W., Treu, T., Bolton, A. S., Gavazzi, R., Koopmans, L. V. E., Marshall, P. J., Bundy, K., & Moustakas, L. A. 2009, ApJ, 705, 1099

Baker, A. J., Lutz, D., Genzel, R., Tacconi, L. J., & Lehnert, M. D. 2001, A&A, 372, L37

Barnabè, M., Czoske, O., Koopmans, L. V. E., Treu, T., Bolton, A. S., & Gavazzi, R. 2009, MNRAS, 399, 21

Baum, S. A., Gallimore, J. F., O'Dea, C. P., Buchanan, C. L., Noel-Storr, J., Axon, D. J., Robinson, A., Elitzur, M., Dorn, M., & Staudaher, S. 2010, ApJ, 710, 289

Blackburne, J. A., & Kochanek, C. S. 2010, ApJ, 718, 1079

Bolton, A. S., Burles, S., Treu, T., Koopmans, L. V. E., & Moustakas, L. A. 2007, ApJ, 665, L105

Bolton, A. S., Treu, T., Koopmans, L. V. E., Gavazzi, R., Moustakas, L. A., Burles, S., Schlegel, D. J., & Wayth, R. 2008, ApJ, 684, 248

Bradač, M., Schneider, P., Steinmetz, M., Lombardi, M., King, L. J., & Porcas, R. 2002, A&A, 388, 373

Chen, J., Rozo, E., Dalal, N., & Taylor, J. E. 2007, ApJ, 659, 52

Chiba, M. 2002, ApJ, 565, 17

Chiba, M., Minezaki, T., Kashikawa, N., Kataza, H., & Inoue, K. T. 2005, ApJ, 627, 53

Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486

Dai, X., Kochanek, C. S., Chartas, G., Kozłowski, S., Morgan, C. W., Garmire, G., & Agol, E. 2010, ApJ, 709, 278

Dalal, N., & Kochanek, C. S. 2002, ApJ, 572, 25

Diemand, J., Kuhlen, M., Madau, P., Zemp, M., Moore, B., Potter, D., & Stadel, J. 2008, Nature, 454, 735

Dobler, G., & Keeton, C. R. 2006, MNRAS, 365, 1243

Gao, L., & Theuns, T. 2007, Science, 317, 1527

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. 2003, Bayesian Data Analysis, Second Edition (Chapman & Hall/CRC Texts in Statistical Science), 2nd edn. (Chapman & Hall)

Gerhard, O., Kronawitter, A., Saglia, R. P., & Bender, R. 2001, AJ, 121, 1936

Gnedin, N. Y. 2000, ApJ, 542, 535

Grillo, C., Eichner, T., Seitz, S., Bender, R., Lombardi, M., Gobat, R., & Bauer, A. 2010, ApJ, 710, 372

Grillo, C., & Gobat, R. 2010, MNRAS, 402, L67

Grillo, C., Gobat, R., Lombardi, M., & Rosati, P. 2009, A&A, 501, 461

Inada, N., Oguri, M., Becker, R. H., White, R. L., Kayo, I., Kochanek, C. S., Hall, P. B., Schneider, D. P., York, D. G., & Richards, G. T. 2007, AJ, 133, 206

Irwin, M. J., Webster, R. L., Hewett, P. C., Corrigan, R. T., & Jedrzejewski, R. I. 1989, AJ, 98, 1989

Kaspi, S., Maoz, D., Netzer, H., Peterson, B. M., Vestergaard, M., & Jannuzi, B. T. 2005, ApJ, 629, 61

Kayser, R., Surdej, J., Condon, J. J., Kellermann, K. I., Magain, P., Remy, M., & Smette, A. 1990, ApJ, 364, 15

Keeton, C. R. 2009, arXiv:astro-ph/0908.3001

Keeton, C. R., Kochanek, C. S., & Seljak, U. 1997, ApJ, 482, 604

Keeton, C. R., & Moustakas, L. A. 2009, ApJ, 699, 1720

Klypin, A., Gottlöber, S., Kravtsov, A. V., & Khokhlov, A. M. 1999a, ApJ, 516, 530

Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999b, ApJ, 522, 82

Kochanek, C. 2004, in Saas-Fee Advanced Courses, Vol. 33, Gravitational Lensing: Strong, Weak and Micro, ed. G. Meylan, P. Jetzer, & P. North (Berlin, Germany; New York, U.S.A.: Springer), 91–268

Kochanek, C. S. 1995, ApJ, 445, 559

Kochanek, C. S., Morgan, N. D., Falco, E. E., McLeod, B. A., Winn, J. N., Dembicky, J., & Ketzeback, B. 2006, ApJ, 640, 47

Koopmans, L. V. E. 2005, MNRAS, 363, 1136

Koopmans, L. V. E., Bolton, A., Treu, T., Czoske, O., Auger, M. W., Barnabè, M., Vegetti, S., Gavazzi, R., Moustakas, L. A., & Burles, S. 2009, ApJ, 703, L51

Kormann, R., Schneider, P., & Bartelmann, M. 1994, A&A, 286, 357

Kravtsov, A. 2010, Advances in Astronomy, 2010

Laurent, O., Mirabel, I. F., Charmandaris, V., Gallais, P., Madden, S. C., Sauvage, M., Vigroux, L., & Cesarsky, C. 2000, A&A, 359, 887

Lawrence, C. R., Elston, R., Januzzi, B. T., & Turner, E. L. 1995, AJ, 110, 2570

Macciò, A. V., Kang, X., Fontanot, F., Somerville, R. S., Koposov, S., & Monaco, P. 2010, MNRAS, 402, 1995

MacLeod, C. L., Kochanek, C. S., & Agol, E. 2009, ApJ, 699, 1578

Madau, P., Diemand, J., & Kuhlen, M. 2008, ApJ, 679, 1260

Mandelbaum, R., van de Ven, G., & Keeton, C. R. 2009, MNRAS, 398, 635

Mao, S., & Schneider, P. 1998, MNRAS, 295, 587

Maraston, C. 2005, MNRAS, 362, 799

Maraston, C., Strömbäck, G., Thomas, D., Wake, D. A., & Nichol, R. C. 2009, MNRAS, 394, L107

Marshall, P. J., Treu, T., Melbourne, J., Gavazzi, R., Bundy, K., Ammons, S. M., Bolton, A. S., Burles, S., Larkin, J. E., Le Mignant, D., Koo, D. C., Koopmans, L. V. E., Max, C. E., Moustakas, L. A., Steinbring, E., & Wright, S. A. 2007, ApJ, 671, 1196

Mashchenko, S., Wadsley, J., & Couchman, H. M. P. 2008, Science, 319, 174

Metcalf, R. B., & Madau, P. 2001, ApJ, 563, 9

Metcalf, R. B., & Zhao, H. 2002, ApJ, 567, L5

Minezaki, T., Chiba, M., Kashikawa, N., Inoue, K. T., & Kataza, H. 2009, ApJ, 697, 610

Minezaki, T., Yoshii, Y., Kobayashi, Y., Enya, K., Suganuma, M., Tomita, H., Aoki, T., & Peterson, B. A. 2004, ApJ, 600, L35

Moore, B., Ghigna, S., Governato, F., Lake, G., Quinn, T., Stadel, J., & Tozzi, P. 1999, ApJ, 524, L19

Moore, B., Governato, F., Quinn, T., Stadel, J., & Lake, G. 1998, ApJ, 499, L5+

More, A., McKean, J. P., More, S., Porcas, R. W., Koopmans, L. V. E., & Garrett, M. A. 2009, MNRAS, 394, 174

Morgan, C. W., Kochanek, C. S., Morgan, N. D., & Falco, E. E. 2010, ApJ, 712, 1129

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Penarrubia, J., Benson, A. J., Walker, M. G., Gilmore, G., McConnachie, A., & Mayer, L. 2010, MNRAS, 406, 1290

Poindexter, S., & Kochanek, C. S. 2010, ApJ, 712, 658

Pooley, D., Blackburne, J. A., Rappaport, S., & Schechter, P. L. 2007, ApJ, 661, 19

Reddy, N. A., & Steidel, C. C. 2004, ApJ, 603, L13

Reddy, N. A., Steidel, C. C., Erb, D. K., Shapley, A. E., & Pettini, M. 2006a, ApJ, 653, 1004

Reddy, N. A., Steidel, C. C., Fadda, D., Yan, L., Pettini, M., Shapley, A. E., Erb, D. K., & Adelberger, K. L. 2006b, ApJ, 644, 792

Riechers, D. A., Walter, F., Brewer, B. J., Carilli, C. L., Lewis, G. F., Bertoldi, F., & Cox, P. 2008, ApJ, 686, 851

Romano-Díaz, E., Shlosman, I., Heller, C., & Hoffman, Y. 2009, ApJ, 702, 1250

Rusin, D., & Kochanek, C. S. 2005, ApJ, 623, 666

Rusin, D., Kochanek, C. S., & Keeton, C. R. 2003, ApJ, 595, 29

Salucci, P., Lapi, A., Tonini, C., Gentile, G., Yegorova, I., & Klein, U. 2007, MNRAS, 378, 41

Sand, D. J., Treu, T., Ellis, R. S., Smith, G. P., & Kneib, J. 2008, ApJ, 674, 711

Scannapieco, E., Thacker, R. J., & Davis, M. 2001, ApJ, 557, 605

Schneider, P., Ehlers, J., & Falco, E. 1992, Gravitational Lenses, Astronomy and Astrophysics Library (Berlin; New York: Springer)

Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337

Siana, B., Smail, I., Swinbank, A. M., Richard, J., Teplitz, H. I., Coppin, K. E. K., Ellis, R. S., Stark, D. P., Kneib, J., & Edge, A. C. 2009, ApJ, 698, 1273

Siana, B., Teplitz, H. I., Chary, R., Colbert, J., & Frayer, D. T. 2008, ApJ, 689, 59

Skilling, J. 2004, in American Institute of Physics Conference Series, Vol. 735, American Institute of Physics Conference Series, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 395–405

Smail, I., Swinbank, A. M., Richard, J., Ebeling, H., Kneib, J., Edge, A. C., Stark, D., Ellis, R. S., Dye, S., Smith, G. P., & Mullis, C. 2007, ApJ, 654, L33

Springel, V., Wang, J., Vogelsberger, M., Ludlow, A., Jenkins, A., Helmi, A., Navarro, J. F., Frenk, C. S., & White, S. D. M. 2008, MNRAS, 391, 1685

Stark, D. P., Swinbank, A. M., Ellis, R. S., Dye, S., Smail, I. R., & Richard, J. 2008, Nature, 455, 775

Strigari, L. E., Bullock, J. S., Kaplinghat, M., Diemand, J., Kuhlen, M., & Madau, P. 2007, ApJ, 669, 676

Suganuma, M., Yoshii, Y., Kobayashi, Y., Minezaki, T., Enya, K., Tomita, H., Aoki, T., Koshida, S., & Peterson, B. A. 2006, ApJ, 639, 46

Teplitz, H. I., McLean, I. S., Becklin, E. E., Figer, D. F., Gilbert, A. M., Graham, J. R., Larkin, J. E., Levenson, N. A., & Wilcox, M. K. 2000, ApJ, 533, L65

Tollerud, E. J., Bullock, J. S., Strigari, L. E., & Willman, B. 2008, ApJ, 688, 277

Treu, T., Auger, M. W., Koopmans, L. V. E., Gavazzi, R., Marshall, P. J., & Bolton, A. S. 2010, ApJ, 709, 1195

Treu, T., & Koopmans, L. V. E. 2002, MNRAS, 337, L6

—. 2004, ApJ, 611, 739

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

van de Ven, G., Mandelbaum, R., & Keeton, C. R. 2009, MNRAS, 398, 607

Vanderriest, C., Schneider, J., Herpe, G., Chevreton, M., Moles, M., & Wlerick, G. 1989, A&A, 215, 1

Vegetti, S., Czoske, O., & Koopmans., L. V. E. 2010, arXiv:astro-ph/1002.4708

Vegetti, S., & Koopmans, L. V. E. 2009, MNRAS, 400, 1583

Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T., & Gavazzi, R. 2009, arXiv:astro-ph/0910.0760

Wadepuhl, M., & Springel, V. 2010, arXiv:astro-ph/1004.3217

Wall, J. V., & Jenkins, C. R. 2003, Practical statistics for astronomers; electronic version, Cambridge Observing Handbooks for Research Astronomers (Leiden: Cambridge Univ. Press)

Williams, L. L. R., Foley, P., Farnsworth, D., & Belter, J. 2008, ApJ, 685, 725

Willman, B., Dalcanton, J. J., Martinez-Delgado, D., West, A. A., Blanton, M. R., Hogg, D. W., Barentine, J. C., Brewington, H. J., Harvanek, M., Kleinman, S. J., Krzesinski, J., Long, D., Neilsen, Jr., E. H., Nitta, A., & Snedden, S. A. 2005, ApJ, 626, L85

Woźniak, P. R., Udalski, A., Szymański, M., Kubiak, M., Pietrzyński, G., Soszyński, I., & Żebruń, K. 2000, ApJ, 540, L65

Wu, Y., Charmandaris, V., Hao, L., Brandl, B. R., Bernard-Salas, J., Spoon, H. W. W., & Houck, J. R. 2006, ApJ, 639, 157

Yee, H. K. C., Ellingson, E., Bechtold, J., Carlberg, R. G., & Cuillandre, J. 1996, AJ, 111, 1783

Yoshida, N., Sokasian, A., Hernquist, L., & Springel, V. 2003, ApJ, 591, L1

Chapter 2

Improved Constraints on the Gravitational Lens Q0957+561

This chapter is based on the published work in the journal article:

Fadely, R., Keeton, C. R., Nakajima, R., & Bernstein, G. B., 2010, ApJ, 711, 246

We present a detailed strong lensing analysis of an HST/ACS legacy dataset for the first gravitational lens, Q0957+561. With deep imaging we identify 24 new strongly lensed features, which we use to constrain mass models. We model the stellar component of the lens galaxy using the observed luminosity distribution, and the dark matter halo using several different density profiles. We draw on the weak lensing analysis by Nakajima et al. (2009) to constrain the mass sheet and environmental terms in the lens potential. Adopting the well-measured time delay, we find $H_0 = 85^{+14}_{-13}$ km s^{-1} Mpc^{-1} (68% CL) using lensing constraints alone. The principal uncertainties in H_0 are tied to the stellar mass-to-light ratio (a variant of the radial profile degeneracy in lens models). Adding constraints from stellar population synthesis models, we obtain $H_0 = 79.3^{+6.7}_{-8.5}$ km s^{-1} Mpc^{-1} (68% CL). We infer that the lens galaxy has a rising rotation curve and a dark matter distribution with an inner core. Intriguingly, we find the quasar flux ratios predicted by our models to be inconsistent with existing radio measurements, suggesting the presence of substructure in the lens.

2.1 Background

Since its inception, strong gravitational lensing has been used to probe galaxy mass distributions in a way that complements photometric and dynamical studies. In systems with numerous lensed features, especially partial or full Einstein rings, lensing can provide a fairly detailed description of the lens galaxy mass distribution (e.g., Kochanek et al. 1989, 2001; Kochanek 1995; Lehár et al. 1996, 1997; Trotter et al. 2000; Koopmans et al. 2006; Gavazzi et al. 2008; Suyu et al. 2008, 2009). In systems with only a few lensed images, however, there may be significant systematic uncertainties in the lens mass distribution. For example, four-image lenses typically constrain the angular structure of the lens potential reasonably well (e.g., Keeton et al. 1997), but often cannot determine the radial profile because the images lie at similar distances from the center of the galaxy (e.g., Keeton & Kochanek 1997; Kochanek 2002). Two-image lenses are better able to probe the radial profile, but provide only poor constraints on the angular structure of the lens potential (Rusin et al. 2003). In addition, further complications may arise if the lens galaxy lies in a group or cluster of galaxies that significantly affect the lens potential (see Keeton & Zabludoff 2004).

These issues explain the challenges associated with modeling the first gravitational lens discovered, Q0957+561 (Walsh et al. 1979). The two lensed images of the background quasar provide limited constraints on a lens potential that is complicated by the presence of a cluster surrounding the main lens galaxy (e.g., Kochanek 1991). To move beyond simple two-image reconstructions, Gorenstein et al. (1988b) and Garrett et al. (1994) used VLBI observations to resolve the radio structure of the quasars; but Grogin & Narayan (1996) and Barkana et al. (1999) found that the new radio constraints still could not strongly constrain the lens potential. Seeking yet more constraints, Bernstein et al. (1997) and Keeton et al. (2000) used the Hubble Space Telescope to discover lensed images of the quasar host galaxy, but even they found that a range of models could reproduce all the strong lensing data (also see Bernstein & Fischer 1999). There

were parallel efforts to obtain other kinds of data to provide independent constraints on the lens galaxy and cluster. Tonry & Franx (1997) measured the central velocity dispersion of the lens galaxy, which Romanowsky & Kochanek (1999) used in stellar dynamical models. Fischer et al. (1997) studied weak lensing by the cluster. Chartas et al. (2002) used the Chandra X-ray Observatory to detect the hot intracluster gas and estimate the cluster mass. Despite the growing amount of data for this system, a definitive description of the mass distribution has remained elusive.

Uncertainties in the mass distribution propagate into attempts to use the lens time delay to determine the Hubble constant, H_0. In principle, lens time delays offer a simple measurement of the Hubble constant at cosmological distances that bypasses the distance ladder. In practice, however, H_0 is degenerate with certain aspects of the mass model (Falco et al. 1985; Gorenstein et al. 1988a; Williams & Saha 2000; Kochanek 2002; Saha & Williams 2006). Due to these degeneracies, measurements of H_0 in individual lens systems have yielded quite varied results. Figure 2.1 shows the lens systems for which individual modeling of the mass distribution has been done to determine H_0. Examining the figure, it is unclear what value of H_0 such studies prefer. Roughly half of the previous studies are consistent with $H_0 < 60$ km s^{-1} Mpc^{-1}, with most preferring $H_0 = 65$–80 km s^{-1} Mpc^{-1}. Strikingly, only two studies, of Q0957+561 (Bernstein & Fischer 1999; Keeton et al. 2000) and PKS 1830-211 (Lidman et al. 1999), found H_0 values > 85 km s^{-1} Mpc^{-1}.

In non-lensing studies, measurements of H_0 have reached better agreement. Using Cepheid variable stars as distance indicators, Freedman et al. (2001) obtained $H_0 = 72 \pm 2$ (stat.) ± 7 (syst.) km s^{-1} Mpc^{-1}. Riess et al. (2009) recently analyzed a higher quality and more homogeneous sample of Cepheids and supernovae and found $H_0 = 74.2 \pm 3.6$ km s^{-1} Mpc^{-1}. Results from measurements of the Cosmic Microwave Background (CMB) have given similar results. Most recently, Dunkley et al. (2008) analyzed the 5th year data release of the Wilkinson Microwave Anisotropy

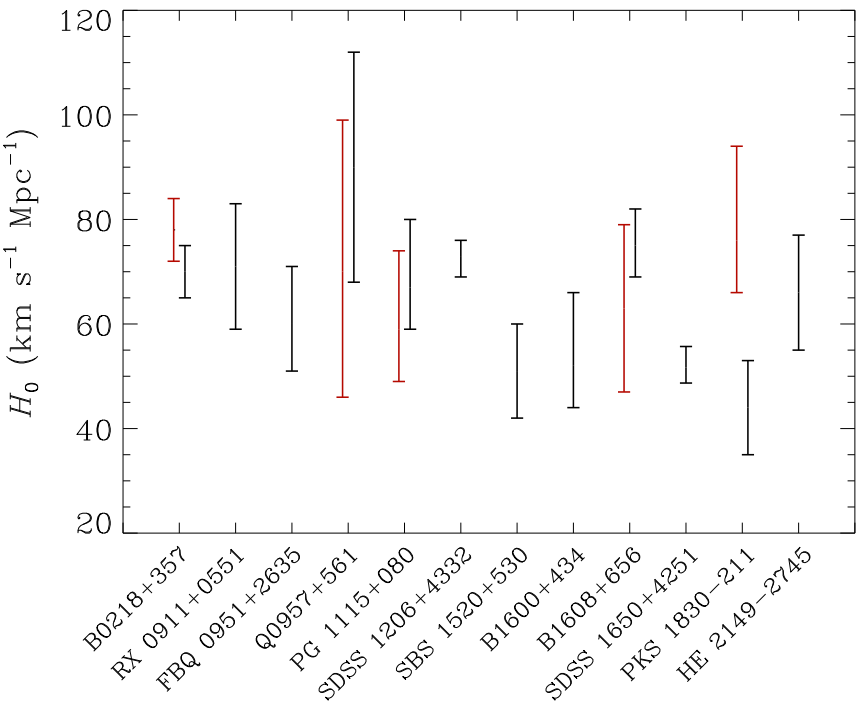

Figure 2.1 Current H_0 measurements in individual lens studies. Black errorbars indicate the most recent measurement for a given lens. When more than one such effort has been made, we present the second most recent effort in red. The measurements span a range of median values and uncertainties; a simple χ^2 test indicates the scatter is not purely statistical at $> 99\%$ confidence. *References*— B0218+357: Wucknitz et al. (2004); York et al. (2005), RX J0911+0551: Hjorth et al. (2002), FBQ0951+2635: Jakobsson et al. (2005), Q0957+561: Bernstein & Fischer (1999); Keeton et al. (2000), PG 1115+080: Treu & Koopmans (2002); Chartas et al. (2004), SDSS 1206+4332: Paraficz et al. (2009), SBS 1520+530: Burud et al. (2002b), B1600+434: Koopmans et al. (2000); Burud et al. (2000), B1608+656: Fassnacht et al. (2002); Koopmans et al. (2003), SDSS 1650+4251: Vuissoz et al. (2007), PKS 1830-211: Lidman et al. (1999); Winn et al. (2002), HE 2149-2745: Burud et al. (2002a).

Probe (WMAP) to find $H_0 = 71.9^{+2.6}_{-2.7}$ km s^{-1} Mpc^{-1}, assuming a universe with a flat geometry and a cosmological constant. With joint analyses of the various constraints, measurements of H_0 can now approach the percent level. For instance, Komatsu et al. (2008) combined H_0 measurements from the CMB, SNIa, and Baryon Acoustic Oscillations (BAO) to obtain $H_0 = 70.5 \pm 1.3$ km s^{-1} Mpc^{-1} or a 2.3% determination of H_0.

Given the apparent consensus in non-lensing studies of H_0, the scatter in results from lensing is puzzling. Are conventional lens models failing to account for variations in key

components of lens models, such as the logarithmic density slope, angular structure, and/or substructure? Are the (often large) uncertainties in lensing measurements of H_0 fully understood? How can the uncertainties be reduced?

One way to confront these concerns is to model a statistical ensemble of lenses when measuring H_0. Saha & Williams (2006) used non-parametric lens modeling to generate free-form mass maps of 10 lenses and obtain $H_0 = 72.5^{+7.8}_{-11.3}$ km s^{-1} Mpc^{-1}. Adding one lens to this sample, Coles (2008) used similar techniques and found $H_0 = 71^{+6}_{-8}$ km s^{-1} Mpc^{-1}. Oguri (2007) took a parametric approach but attempted to incorporate a reasonable amount of complexity and scatter in the models; his analysis of 16 time delay lenses yielded $H_0 = 68 \pm 6$ (stat.) ± 8 (syst.) km s^{-1} Mpc^{-1}. While these results are enticing, the systematic uncertainties may still be underestimated and cannot be beaten down with sample size. In particular, lensing constraints on H_0 are known to be affected by the "mass-sheet degeneracy": a uniform sheet of mass can be added to a lensing potential in a way that leaves the positions and brightnesses of the images unchanged but rescales the inferred Hubble constant (Falco et al. 1985; Gorenstein et al. 1988a). Groups and clusters surrounding lens galaxies or along the line of sight can act as mass sheets that bias lensing measurements of H_0 (Keeton & Zabludoff 2004). While there is considerable effort to characterize the local and line-of-sight environments of strong lenses (e.g., Fassnacht & Lubin 2002; Fassnacht et al. 2006, 2008; Momcheva et al. 2006; Williams et al. 2006; Auger et al. 2007), and Oguri (2007) attempted to account for the mass sheet in several lenses for which it is expected to be significant, it is still not clear whether the systematic uncertainties are fully understood.

In order to identify and control systematics, it seems that we still need to model individual lens systems in detail. One advantage of this approach is the ability to use additional sources of information, such as weak lensing (e.g., Fischer et al. 1997; Bernstein & Fischer 1999) or stellar dynamics (e.g., Grogin & Narayan 1996; Tonry & Franx 1997; Treu & Koopmans 2002; Barnabè & Koopmans 2007), to reduce the

uncertainties inherent in strong lensing. With the four-image lens B1608+656 (Myers et al. 1995), for example, Suyu et al. (2008, 2009) demonstrate that lensing and velocity dispersion information can be combined to get a $\sim 7\%$ determination of H_0, in which the modeling uncertainties are comparable to the uncertainties in the measured time delays themselves (Fassnacht et al. 2002). With Q0957, we have a system whose complicated lens potential not only presents certain challenges for measuring H_0, but also bestows certain opportunities. In particular, the presence of the cluster around the lens galaxies offers a rare chance to do a weak lensing analysis of an *individual* strong lens system (as opposed to a stacked ensemble; e.g., Gavazzi et al. 2007; Lagattuta et al. 2009), which is key to breaking the mass sheet degeneracy and controlling that vital systematic uncertainty. Furthermore, there is now a rich variety of data available for Q0957, which enables a broad-based analysis. We therefore study Q0957 to investigate statistical and systematic uncertainties in lens models and H_0. The strong lensing analysis presented here complements and draws upon the weak lensing analysis presented by Nakajima et al. (2009); both are based on the same new HST/ACS data. In this system extensive observational effort (e.g., Schild & Cholfin 1986; Kundic et al. 1997; Oscoz et al. 2001) has yielded a very precise time delay with an uncertainty of just $\sim 0.02\%$ (Colley et al. 2003). While we do not expect to achieve that level of precision in lens models, our goal is to see just how well we can do. We combine our new data with a number of independent constraints from previous observations as well as stellar population synthesis models, in order to obtain the most precise measurement of H_0 to date for Q0957.

In this paper we assume a flat universe with matter density $\Omega_M = 0.274$ and cosmological constant $\Omega_\Lambda = 0.726$, which is consistent with the 5-year WMAP+SN+BAO constraints from Komatsu et al. (2008).

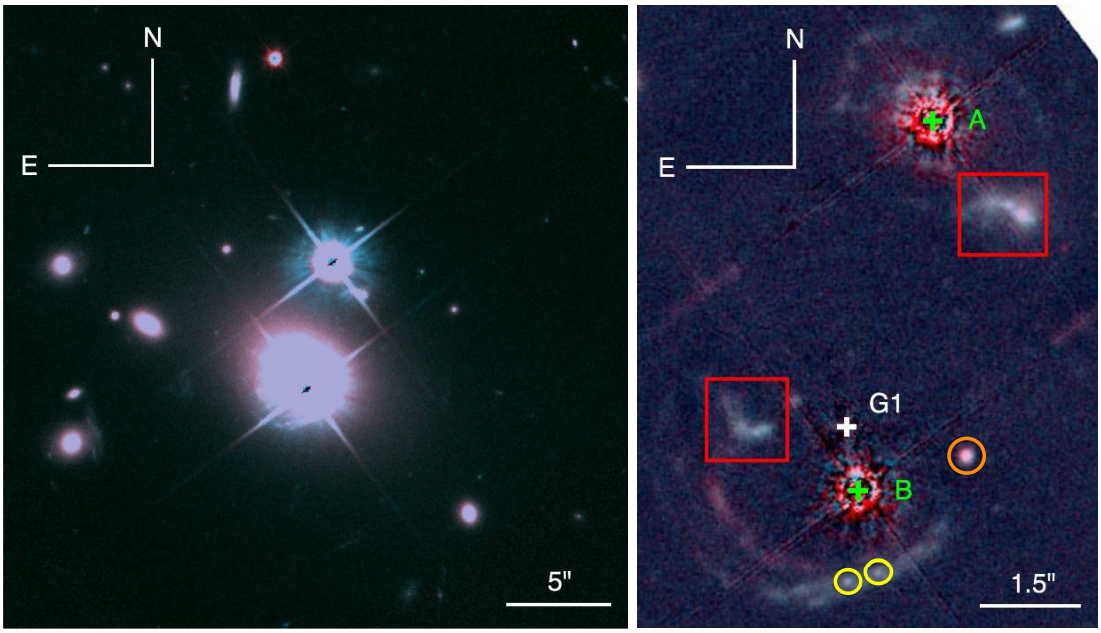

Figure 2.2 *(a, left)* A false color image of the central 30″ of our combined F606W and F814W images of Q0957+561. *(b, right)* Close-up of the strong lensing region, after the main lensing galaxy and quasar images have been subtracted. The white cross indicates the center of the lens galaxy G1, while the green crosses indicate the quasar positions A and B. The red boxes and yellow circles indicate the "blobs" and "knots" identified by Bernstein et al. (1997). Newly resolved faint features are seen south and east around quasar B and southwest of quasar A. The orange circle indicates an unknown object, not associated with any lensed features. Since the light profile of the object is consistent with the PSF, we surmise it is a faint halo star in the foreground of the lens.

2.2 Observations and Data Analysis

2.2.1 Observations

We conducted deep observations of Q0957+561 on 2005 October 10–11 with HST's Advanced Camera for Surveys as part of program GO-10569. Using four pointings of 7.7 ks in the F606W filter and 3.8 ks in the F814W filter, we created a $6' \times 6'$ mosaic of the field. We arranged the pointings to overlap in the central 30″ region, forming a 30 ks image in the F606W filter (15 ks in F814W) for our strong lensing analysis with a final pixel scale of 0.03″. The large number of exposures in this central region allows us to use a simple image-combining algorithm that avoids the undesirable PSF broadening and noise correlation of the common *Drizzle* algorithm (Fruchter & Hook 2002):

1. An astrometric solution is derived for each exposure by compounding the ACS/WFC coordinate map of Anderson (2002) with an additional affine transformation to account for pointing errors, stellar aberration, and slight plate-scale variations due to the HST "breathing mode." The coefficients of the affine transformation are derived by registering objects detected in individual exposures.

2. A grid of $0.03''$ pixels is created for the combined image. Each pixel in each exposure is mapped to a single destination pixel. Input pixels flagged as invalid due to detector defects, etc., are discarded.

3. For each destination pixel, we average all of the input pixels, using a sigma-clipping algorithm to eliminate pixels contaminated by cosmic rays.

The procedure is identical to the use of *Drizzle* with a "drop zone" of zero size. Since each input pixel contributes to only one output pixel, the output pixels have uncorrelated noise. The combining algorithm broadens the PSF only by an effective convolution with the output pixel square. The final pixel size is chosen such that it is small enough not to degrade the resolution, but coarse enough that there are enough input pixels for averaging and outlier rejection. We present a false color image of our combined F606W and F814W images in Figure 2.2a.

To look for new strong lensing constraints, we subtract the bright quasar images A and B using the PSF derived from observations of the star HD237859. Since the PSF varies with focal plane position as well as time, we observed the star as close as possible in time and chip position to each quasar image in each of the four pointings.

2.2.2 Lens Galaxy Properties

We model the main lensing galaxy using the IRAF ELLIPSE task, masking out regions where quasar subtraction takes place as well as any bright regions not associated with

Table 2.1. Lens Galaxy Photometry

Filter	Total Counts	m_{AB}	Zeropoint, AB[1]
F606W	3328.6	18.809 ± 0.061	27.614
F814W	5372.9	17.743 ± 0.065	27.068

[1]Quoted zeropoints differ from standard values as we must correct for a plate scale change from $0.05''$ to $0.03''$ pixels

the galaxy (e.g., other lensed features). Our resulting IRAF model provides a measurement of the galaxy's isophotes and total flux (see Table 2.1). As shown previously (Bernstein et al. 1997), the isophotes of the lens galaxy exhibit an ellipticity gradient and a position angle twist (Figure 2.3). These isophotal features may complicate the lensing potential, so we incorporate them directly into our lens models (§2.3.2). We also use the photometry of the lens galaxy to constrain stellar population synthesis models and estimate the stellar mass-to-light ratio (§2.4.3).

2.2.3 Faint Strong Lensing Features

We subtract the model galaxy from the quasar-subtracted image to produce the final image of the strong lensing region, which is shown in Figure 2.2b. This image reveals several new, previously unresolved or undetected strongly lensed features. Since the morphology is similar to the host galaxy arc from NICMOS (Keeton et al. 2000), we conjecture that the optical features are most likely images of star forming regions of the quasar host galaxy at $z = 1.41$.

The lensed "blobs" and "knots" indicated in Figure 2.2b were previously identified by Bernstein et al. (1997), and were used by Bernstein & Fischer (1999) and Keeton et al. (2000) as constraints on lens models. To derive new constraints from our new strongly lensed features, we use the models of Keeton et al. (2000) as a starting point.

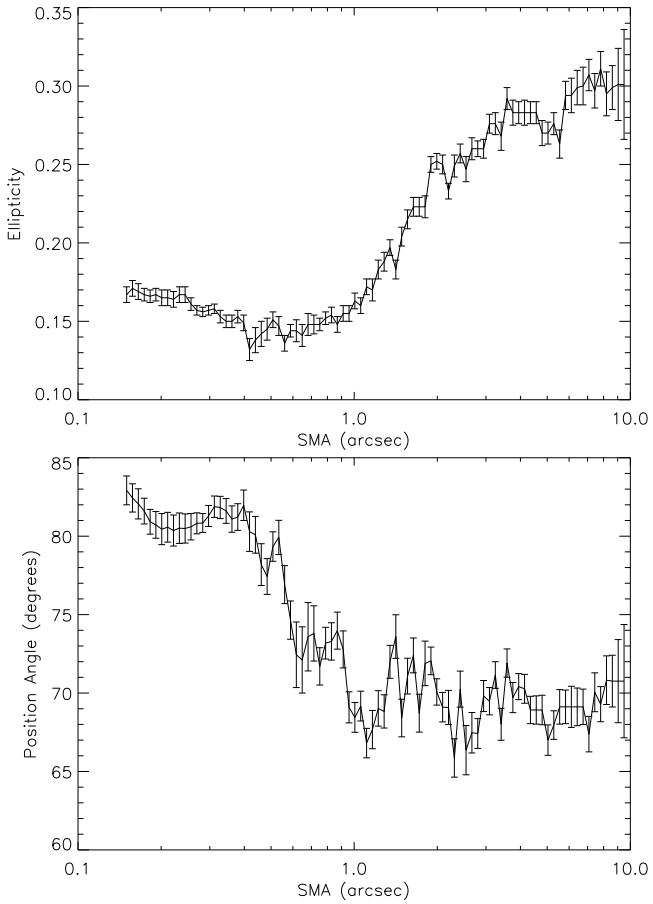

Figure 2.3 Ellipticity and position angle of galaxy isophotes, plotted as a function of semi-major axis. Note the increase in ellipticity beginning at $\sim 1''$ and the decrease in position angle beginning at $\sim 0.3''$.

Using the *lensmodel* software (Keeton 2001), we check to see how these older models would map new features in the image plane. Specifically, we take an observed image position, map it to the source plane, and then find all corresponding images using the old lens models. We then look for the predicted images in our new HST data. Unfortunately, we find the Keeton et al. models cannot sensibly reproduce the lensing we see in the HST data. These models fail most notably for the new features south and east of quasar B, mapping bright peaks in the data to blank regions of the sky.

In order to make sense of the new features, we examine the morphology of the images in the data. Specifically, in the area around the bright "knots" shown in Figure 2.2b we notice a distinct fork-like feature extending from either side of the knots. Using

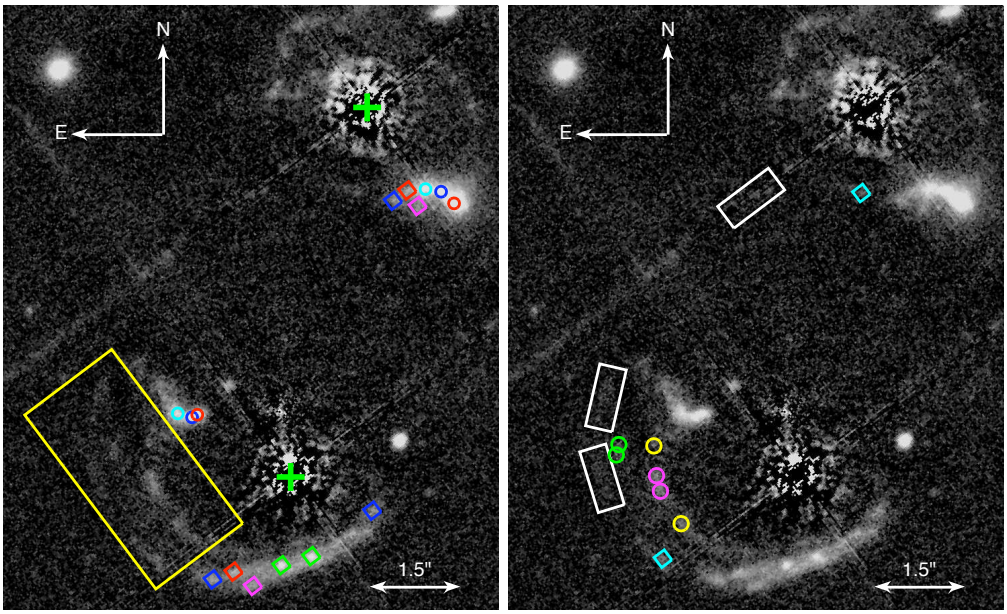

Figure 2.4 Our proposed mapping of the new lensed features (cf. Table 2.2). Corresponding shapes and colors indicate proposed images of a given source. The grid-like pattern that appears is due to the imperfect subtraction of the diffraction pattern of the quasar images. *(a, left)* We use the images indicated to constrain new lens models, and then use those models to test features in the yellow box. *(b, right)* We take one of the points in each shape/color pair, map it back to the source, and then find the corresponding images. The predicted positions match very well with features in the image, which gives us confidence both that the "primary" constraints are valid and that the additional lensed features are real. (Points in white rectangles indicate features that are too close to the noise level to provide confident peaks.) All of the colored points are used as constraints in our subsequent modeling.

peaks in this structure, we postulate a set of new constraints as shown in the left panel of Figure 2.4, which we call our "primary" set of new constraints. We use these primary features to constrain a singular isothermal ellipsoid lens model, and see how the resulting model maps other faint features found in the HST data. Specifically, we consider faint features to the east of quasar B, indicated by the large box in Figure 2.4a. To our surprise, the model derived from our primary features accurately describes the remaining faint features, mapping peaks of the images in reasonable ways (Figure 2.4b) and giving us confidence in the identification of new lensed features. We therefore add these features to our list of constraints, resulting in a final data set which is given in Table 2.2. We note that all presumed multiple images of each source are observed to

Table 2.2. Modeling Constraints

Feature	Position (″)[1]	Symbol in Figure 3
G1	$(0, 0) \pm 0.00001$	–
Quasar A	$(1.408, 5.034) \pm 0.03$	Green Plus
Quasar B	$(0.182, -1.018) \pm 0.03$	Green Plus
Jet A5	$(1.392, 5.080) \pm 0.03$	–
Jet B5	$(0.164, -0.962) \pm 0.03$	–
IA	$(2.878, 3.453) \pm 0.05$	Red Circle
IB	$(-1.362, -0.043) \pm 0.05$	Red Circle
IIA	$(2.666, 3.634) \pm 0.05$	Blue Circle
IIB	$(-1.457, -0.075) \pm 0.05$	Blue Circle
IIIA	$(2.395, 3.694) \pm 0.05$	Cyan Circle
IIIB	$(-1.682, -0.026) \pm 0.05$	Cyan Circle
IVA	$(0.021, -2.532) \pm 0.03$	Green Diamond
IVB	$(0.512, -2.386) \pm 0.03$	Green Diamond
VA	$(2.111, 3.664) \pm 0.05$	Red Diamond
VB	$(-0.768, -2.640) \pm 0.05$	Red Diamond
VIA	$(1.875, 3.488) \pm 0.12$	Blue Diamond
VIB	$(-1.128, -2.777) \pm 0.05$	Blue Diamond
VIC	$(1.523, -1.634) \pm 0.05$	Blue Diamond
VIIA	$(1.875, 3.488) \pm 0.12$	–
VIIB	$(-1.065, -2.786) \pm 0.05$	–
VIIC	$(1.489, -1.688) \pm 0.05$	–

have similar colors (within the photometric noise).

To obtain the position and uncertainty of each peak listed in Table 2.2, we examine the brightest pixel(s) within the peak. If the brightest pixel is more than $10\sigma_{noise}$ above any other pixel in the peak (as for the "knots" or source IV in Table 2.2), we set the position error to be ± 1 pixel or $\pm 0.03''$. If the brightest pixel is 3–$10\sigma_{noise}$ above the surrounding pixels, we conservatively set the error to ± 1.5 pixels or $\pm 0.05''$. If there are multiple pixels in the peak that are within $3\sigma_{noise}$ of the brightest, we take the average position of all such pixels and set the error to be the distance from this average to the farthest of the bright pixels, plus our nominal value of $0.05''$.

Table 2.2 (cont'd)

Feature	Position (″)[1]	Symbol in Figure 3
VIIIA	$(2.280, 3.391) \pm 0.09$	Magenta Diamond
VIIIB	$(-0.454, -2.873) \pm 0.09$	Magenta Diamond
IXA	$(-2.003, -2.435) \pm 0.05$	Cyan Diamond
IXB	$(1.293, 3.611) \pm 0.05$	Cyan Diamond
XA	$(-1.708, -1.878) \pm 0.08$	Yellow Circle
XB	$(-2.181, -0.551) \pm 0.08$	Yellow Circle
XIA	$(-2.070, -1.252) \pm 0.08$	Magenta Circle
XIB	$(-2.112, -1.053) \pm 0.08$	Magenta Circle
XIIA	$(-2.745, -0.516) \pm 0.08$	Green Circle
XIIB	$(-2.781, -0.699) \pm 0.08$	Green Circle

Note. — New features are indicated with Roman numerals

[1]Written as (x, y) where x is West and y is North.

2.3 Lens Modeling Methods

2.3.1 General Theory

In this section we present a brief review of the lensing theory that is particularly pertinent to our analysis of Q0957. For further discussion of strong lensing theory, please see Schneider et al. (1992) and Kochanek (2004).

As predicted by Einstein's General Relativity in the early 20th century, a mass concentration near the line of sight to a background object may significantly displace and distort a background image. The angular position $\boldsymbol{u}$ of the source and the angular position of $\boldsymbol{x}$ of an image are related by the lens equation,

$$\boldsymbol{u} = \boldsymbol{x} - \boldsymbol{\nabla}\phi\,, \tag{2.1}$$

where ϕ is the (scaled) gravitational potential due to mass at the lens redshift. The lens potential is given by

$$\nabla^2\phi(\boldsymbol{x}) = 2\kappa(\boldsymbol{x}) = 2\Sigma/\Sigma_{crit}, \tag{2.2}$$

where the convergence κ equals the surface mass density (Σ) scaled by the critical density for lensing (Σ_{crit}).

Since the deflected light travels along different ray paths, there is a difference in the light travel time for different images. This difference, known as the time delay, can be measured if the source is sufficiently variable. The time delay between images at positions $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ is given by

$$\begin{aligned} \Delta t_{ij} &= \frac{1+z_l}{c}\frac{D_{ol}D_{os}}{D_{ls}} \\ &\quad \times \left\{ \frac{1}{2}\left(|\boldsymbol{x}_i - \boldsymbol{u}|^2 - |\boldsymbol{x}_j - \boldsymbol{u}|^2\right) - \left[\phi(\boldsymbol{x}_i) - \phi(\boldsymbol{x}_j)\right] \right\}, \end{aligned} \tag{2.3}$$

where z_l is the lens redshift and D_{ol}, D_{os}, and D_{ls} are angular-diameter distances from the observer to the lens, the observer to the source, and the lens to the source respectively. Combining a measured time delay with a lens model (to infer the source position $\boldsymbol{u}$ and the lens potential ϕ) provides a measurement of the distance combination $D_{ol}D_{os}/D_{ls}$, which is inversely proportional to H_0. (The distance ratio also depends on cosmological parameters Ω_M and Ω_Λ, but that is typically a small effect compared with other uncertainties in the problem.)

A well-known problem in lensing constraints on H_0 is the "mass-sheet degeneracy" (Falco et al. 1985; Gorenstein et al. 1988a). For any potential ϕ that fits the data, one can construct another potential

$$\phi' = \frac{1}{2}\kappa'|\boldsymbol{x}|^2 + (1-\kappa')\phi \tag{2.4}$$

that yields exactly the same image positions and flux ratios. The addition of the mass sheet κ' does, however, rescale the time delays and hence the inferred Hubble constant, such that $H_0' = (1-\kappa')H_0$. The challenge for Q0957 is that the cluster around the lens contributes a term (κ_c in eq. 2.11 below) that acts like a mass sheet, which we cannot constrain by fitting the positions and fluxes of the strongly lensed images. We therefore set $\kappa_c = 0$ when doing the strong lens modeling, so we obtain some Hubble constant

estimate $H_{0,model}$. We then use weak lensing data to constrain κ_c, which yields our corrected Hubble constant estimate

$$H_0 = (1 - \kappa_c) H_{0,model}. \tag{2.5}$$

Additionally, since the mass sheet correction is a rescaling of the potential, we must multiply each term in the potential by the same factor of $1 - \kappa_c$ as in eq. (2.5). For our results in §2.4, we indicate all parameters to which this applies.

2.3.2 Mass Models

Stellar component

The lensing potential of Q0957+561 may be complicated by the ellipticity gradient and isophote twist seen in the luminous component of the lens galaxy (see Figure 2.3). To allow such features in lens models, Keeton et al. (2000) introduced "double pseudo-Jaffe" models featuring a superposition of two ellipsoidal mass distributions, centered on the galaxy position, with different ellipticities, orientations, and scale radii. We initially adopted similar models, but quickly judged them to be unsatisfactory. Constrained by the new lensed features, double pseudo-Jaffe models were adopting odd forms, such as one round and one very flattened component, that seemed unrealistic and unlike what is observed for giant elliptical galaxies.

A better approach is to incorporate the observed ellipticity gradient and isophote twist directly by using our isophotal model of the stellar component. We combine the model galaxy with an assumed stellar mass-to-light ratio to construct a convergence map. (We vary the mass-to-light ratio during the modeling, as discussed in §2.3.4.) To compute the corresponding lensing potential, we solve the Poisson equation using Fourier transforms. In Fourier space, the Poisson equation (2.2) has the form

$$-\boldsymbol{k}^2 F(\boldsymbol{k}) = 2K(\boldsymbol{k}) \tag{2.6}$$

where $K(\boldsymbol{k})$ is the Fourier transform of the convergence map and $F(\boldsymbol{k})$ is the Fourier transform of the lens potential. It is straightforward to construct the convergence on a two-dimensional grid, calculate $K(\boldsymbol{k})$ using FFTs, solve for $F(\boldsymbol{k})$, and then do an inverse Fourier transform to get $\phi(\boldsymbol{x})$. We can then obtain the lensing deflection and magnification by computing derivatives of $\phi(\boldsymbol{x})$ in Fourier space. This method is discussed in more detail by van de Ven et al. (2009).

Dark matter

The luminous galaxy is presumably embedded in its own dark matter halo and any halo associated with the surrounding cluster. The cluster in Q0957 is non-negligible: previous weak lensing studies constrained the mean convergence within $30''$ of the lens to be $\kappa_{30''} = 0.26 \pm 0.16$ and indicated a shear of $\gamma \sim 0.2$ (Fischer et al. 1997; Bernstein & Fischer 1999). (In §2.3.5 we report our own constraints on cluster mass models.) Also, X-ray observations of the intracluster gas indicated a convergence from the cluster at the quasar positions of $\kappa_A = 0.22^{+0.14}_{-0.07}$ and $\kappa_B = 0.21^{+0.12}_{-0.07}$ (Chartas et al. 2002).

Since the lens galaxy is the brightest and (presumably) most massive galaxy in the cluster, it seems natural to assume as a fiducial model that the galaxy lies at the center of the cluster. In this case the dark matter halo we insert in our models represents some combination of dark matter associated with the galaxy and dark matter associated with the cluster as a whole. We consider various profiles to encompass the range of possibilities. One model we use is the Navarro-Frenk-White profile (NFW, Navarro et al. 1997),

$$\rho = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2} \tag{2.7}$$

whose projected surface density has the form (Bartelmann 1996)

$$\kappa(r) = 2\kappa_s \frac{1 - \mathcal{F}(x)}{x^2 - 1} \tag{2.8}$$

where $x = r/r_s$, $\kappa_s = \rho_s r_s / \Sigma_{crit}$, and

$$\mathcal{F}(x) = \begin{cases} \frac{1}{\sqrt{x^2-1}} \tan^{-1}\sqrt{x^2-1} & (x > 1) \\ \frac{1}{\sqrt{1-x^2}} \tanh^{-1}\sqrt{1-x^2} & (x < 1) \\ 1 & (x = 1) \end{cases} \tag{2.9}$$

While these equations describe a spherically symmetric model, we can obtain an elliptical model by replacing r with the ellipse coordinate $\xi = (x^2 + y^2/q^2)^{1/2}$ where q is the projected axis ratio. The lensing potential and its derivatives can be computed for elliptical models using numerical integrals (Schramm 1990; Keeton 2001).

We also use three softened power law profiles with projected surface densities of the form

$$\kappa(\xi) = \frac{1}{2} \frac{b_d^{2-\alpha}}{(a^2 + \xi^2)^{1-\alpha/2}} \tag{2.10}$$

where ξ is the ellipse coordinate defined above, a is the core radius, b_d is a normalization parameter, and the power law index α is chosen such that $M(R) \propto R^\alpha$ for $R \gg a$. In this class of models we study one with an isothermal profile ($\alpha = 1$), one steeper than isothermal ($\alpha = 0.5$), and one shallower than isothermal ($\alpha = 1.5$). For $\alpha = 1$ the lensing potential and its derivatives can be computed analytically (Keeton & Kochanek 1998), but for $\alpha \neq 1$ they require numerical integrals.

We still need to consider the possibility that the cluster may not be centered on the lens galaxy, which is germane because the observed X-ray emission from the cluster is slightly offset from the lens: the X-ray centroid is 4.3 ± 1.3 arcsec East and $3.2^{+1.2}_{-0.6}$ arcsec North of Image B, or 4.7 ± 1.2 arcsec from the lens galaxy (Chartas et al. 2002). We carry out our full modeling analysis (as described in the remainder of the paper) treating the cluster as a distinct dark matter halo centered at the observed X-ray position. While this approach yields fits that are formally comparable to our fiducial results, we ultimately reject such models for two reasons. When we treat the offset cluster with a softened power law profile, the models require a large ($> 30''$)

core radius, which makes the "cluster" basically equivalent to a uniform mass sheet on the $\sim 6''$ scale of strong lensing and thus negates the effect of having an extra halo. NFW models cannot mimic that behavior, of course, since they do not have a flat core. Instead, we find that an offset NFW cluster must have an unreasonably large ellipticity of > 0.7. We conclude that the strong lens data do not favor and in fact disfavor an offset cluster, so we not include a separate cluster component in our main analysis.

Environmental terms

While our lens-centered mass profiles should account for the majority of the mass in the strong lensing region (apart from a mass sheet), they should not be expected to represent the full complexity of the cluster potential. The cluster halo may not be a simple ellipsoid on large scales, and it presumably has some lumpiness in the form of individual cluster galaxies. In general, we can write the lens potential due to structures outside the strong lensing region using a Taylor series expansion of the form

$$\begin{aligned} \phi_{env}(r,\theta) &= \frac{\kappa_c}{2}\, r^2 + \frac{\gamma}{2}\, r^2 \cos 2(\theta - \theta_\gamma) \\ &\quad + \frac{\sigma}{4}\, r^3 \cos(\theta - \theta_\sigma) + \frac{\delta}{6}\, r^3 \cos 3(\theta - \theta_\delta) + \ldots \end{aligned} \tag{2.11}$$

where κ_c is a mass sheet, γ is an external shear, and σ and δ represent higher, third-order terms. If the structures are "far" enough from the Einstein radius, the higher-order terms may be sufficiently small that the expansion can be truncated after the shear; indeed this approximation is used in many lens systems. However, as shown by Kochanek (1991) and many subsequent studies, the third-order terms cannot be neglected in Q0957. Bernstein & Fischer (1999) and Keeton et al. (2000) included the third-order terms but imposed the assumptions $\sigma = -2\delta/3$ and $\theta_\sigma = \theta_\delta$, which corresponds to the perturbation from a singular isothermal sphere. We find that this assumption is too restrictive, so we include general third-order terms in order to allow more freedom in the models to account for complex features we have not explicitly modeled.

Note that we have written eq. (2.11) in terms of the amplitudes (γ, σ, δ) and directions $(\theta_\gamma, \theta_\sigma, \theta_\delta)$ of the shear and third-order terms. We can think of these as "polar" coordinate versions of these parameters. We can equivalently define "Cartesian" coordinate versions of the same parameters:

$$\begin{aligned} \gamma_c &= \gamma \cos 2\theta_\gamma \,, \quad \gamma_s = \gamma \sin 2\theta_\gamma \,, \\ \sigma_c &= \sigma \cos \theta_\sigma \,, \quad \sigma_s = \sigma \sin \theta_\sigma \,, \\ \delta_c &= \delta \cos 3\theta_\delta \,, \quad \delta_s = \delta \sin 3\theta_\delta \,. \end{aligned} \tag{2.12}$$

While we quote results for the "polar" parameters, we actually carry out our modeling analysis using the "Cartesian" parameters. The translation is straightforward.

In summary, our composite models consist of a stellar component, an elliptical halo centered on the stellar component that accounts for dark matter in the lens galaxy and/or cluster halo, and an additional set of terms corresponding to a third-order Taylor series expansion of the potential from the lens environment. The 11 associated model parameters are listed in Table 2.3.

2.3.3 Model Constraints

Table 2.2 lists the positions of the lensed images used as constraints on our models. We supplement the faint images discussed in §2.2.3 with the quasar cores and radio jets resolved with Very Long Baseline Interferometry (VLBI; Gorenstein et al. 1988b; Garrett et al. 1994). With a resolution of ~ 0.1 mas, the VLBI observations reveal structure within the jets and provide astrometry with formal errors < 1 mas. Previous studies of Q0957 used this excellent astrometry as strong constraints on lens models (Barkana et al. 1999; Bernstein et al. 1997; Bernstein & Fischer 1999; Keeton et al. 2000). In this work, however, we reconsider the use of such stringent constraints. Many recent works have shown that lens galaxies often contain substructure in the form of CDM subhalos, which can perturb lens flux ratios by tens of percent or more and image

Table 2.3. Model Parameters

Parameter	Label
Stellar mass to light ratio	$\Upsilon_{\rm F606W}$
Halo ellipticity	e
Position angle	PA
Shear angle	θ_γ
Core radius (α models)	a
Scale radius (NFW models)	r_s
σ angle	θ_σ
δ angle	θ_δ
Halo mass normalization	b_d
External Shear	γ
$3^{\rm rd}$ order term	σ
$3^{\rm rd}$ order term	δ

Note. — Power law α models use a core radius a, while NFW models use a scale radius r_s.

positions at levels up to ~ 10 mas (e.g., Mao & Schneider 1998; Metcalf & Madau 2001; Dalal & Kochanek 2002; Chiba et al. 2005; Chen et al. 2007). Since our models do not contain any substructure,[1] they should not be expected to match the image positions to better than the ~ 10 mas residuals expected from substructure. We therefore broaden the errorbars and adopt conservative uncertainties on the VLBI quasar and jet positions of 30 mas. As a check, we verified that the VLBI derived positions of the quasars are in good agreement with the positions in our HST data.

We take the image position constraints together with constraints from our weak lensing analysis (§2.3.5) to comprise our main model constraints. Subsequently, we consider adding supplementary constraints in the form of lensed arcs detected with HST/NICMOS, the quasar radio flux ratio, and stellar population synthesis models

[1]It would certainly be interesting to add substructure to the models and use the precise radio positions and flux ratios to constrain the amount of substructure in the lens; but that is beyond the scope of the present work.

(see §2.4 for details). Throughout the analysis we use the time delay 417.09 ± 0.07 days from Colley et al. (2003) to infer the Hubble constant.

2.3.4 Strong Lensing Analysis

In Bayesian language, our ultimate goal is to determine the posterior probability distribution for our model parameters and H_0, given the observational constraints. We have $N_{data} = 60$ position constraints, compared with $N_{param} = 39$ free parameters (11 parameters for the mass model plus 28 source position parameters). Hence our analysis has $N_{data} - N_{param} = 21$ degrees of freedom. The Hubble constant analysis involves one additional constraint (the observed time delay) and one additional parameter (H_0), so it has the same number of degrees of freedom.

Formally, the posterior probability distribution has the form

$$P(\theta|d, M) = \frac{P(d|\theta, M)\, P(\theta|M)}{P(d|M)} \ , \tag{2.13}$$

where d denotes the data that provide constraints on the parameters θ associated with some model M. The likelihood of the data given the model is

$$\mathcal{L} \equiv P(d|\theta, M) \propto e^{-\chi^2/2} \ , \tag{2.14}$$

where χ^2 is the goodness of fit. The quantity $P(\theta|M)$ represents priors on the model parameters, which we take to be uniform.[2] The Bayesian Evidence $P(d|M)$ is discussed below.

To ease the search of our large parameter space, we treat the image positions using the approximate position χ^2 as defined by Keeton (2001):

$$\chi^2_p \ = \ \sum_i \delta \boldsymbol{x}_i^T \cdot S_i^{-1} \cdot \delta \boldsymbol{x}_i \ \approx \ \sum_i \delta \boldsymbol{u}_i^T \cdot \mu_i^T \cdot S_i^{-1} \cdot \mu_i \cdot \delta \boldsymbol{u}_i \, , \tag{2.15}$$

where the sum runs over all images, $\delta \boldsymbol{x}_i = \boldsymbol{x}_{obs,i} - \boldsymbol{x}_{mod,i}$ is the position residual for image i, $S_i = \mathrm{diag}(\sigma_i^2, \sigma_i^2)$ is the covariance matrix for the image positions, and μ_i is

[2]Note that our uniform priors apply to the "Cartesian" coordinate versions of the environmental parameters (eq. 2.12) and the ellipticity ($e_c = e \cos 2\theta_e$ and $e_s = e \sin 2\theta_e$).

the lensing magnification tensor. Using the lens equation, the position residual in the source plane is $\delta \boldsymbol{u}_i = \boldsymbol{x}_{obs,i} - \nabla\phi(\boldsymbol{x}_{obs,i}) - \boldsymbol{u}_{mod}$. The last step in eq. (2.15) is valid when the position residuals are small such that the image and source plane residuals are related by $\delta \boldsymbol{x}_i \approx \mu_i \cdot \delta \boldsymbol{u}_i$. The benefit of using this approach is that χ^2_p is quadratic in each source position $\boldsymbol{u}_{mod}$, so we can find the best-fit value analytically:

$$\boldsymbol{u}_{mod} = A^{-1}\boldsymbol{b}\,, \tag{2.16}$$

where

$$A = \sum_i \mu_i^T \cdot S_i^{-1} \cdot \mu_i\,, \tag{2.17}$$

$$\boldsymbol{b} = \sum_i \mu_i^T \cdot S_i^{-1} \cdot \mu_i \cdot \boldsymbol{u}_{obs,i}\,, \tag{2.18}$$

where these sums now run over the known images of a given source. The upshot is that we do not have to search explicitly through the 28 dimensions corresponding to the source parameters. An additional advantage is that we only have to compute the lens potential and its derivatives at the known positions of the images, which is useful for our dark matter models that require numerical integrals.

We must still search the 11-dimensional space of mass model parameters. We sample the posterior probability distribution in this space using an adaptive Metropolis-Hastings Monte Carlo Markov Chain (MCMC) algorithm (Roberts et al. 1997; Haario et al. 2001; Roberts & Rosenthal 2001). Each chain consists of a sequence of trial steps drawn from a multivariate Gaussian distribution. In 95% of the steps the Gaussian is based on an empirically updated covariance matrix to provide efficient sampling of a high-dimensional posterior distribution. In the remaining 5% of the steps the covariance matrix is diagonal so the algorithm takes independent and relatively large steps along the coordinate axes; this feature acts as a "fail-safe" to help the algorithm escape local minima in the χ^2 surface and potentially discover new features in the posterior distribution. The probability of accepting a trial step that modifies the likelihood from $\mathcal{L}_{old}$ to $\mathcal{L}_{new}$ is $\min(1, \mathcal{L}_{new}/\mathcal{L}_{old})$. In other words, if a trial step increases the likelihood

it is automatically accepted; while if it decrease the likelihood it is accepted with a probability given by the likelihood ratio.

We run 25 chains simultaneously and set them up to work "from the outside in." That is, we generate an initial sample of models by drawing $\sim 10^4$ random points in the parameter space and optimizing them; this provides an estimate of the allowed parameter ranges (although without the proper statistical sampling that MCMC provides). We then select models with maximal/minimal values of individual parameters to use as starting points for the MCMC chains. By starting with well spread chains, the MCMC algorithm can establish the covariance matrix more quickly, and spend more time sampling the tails of the distribution, than it would by starting with closely-spaced starting points (Gelman et al. 1995). The choice of initial models does not matter in detail, though, because for our final sampling we merge only the second halves of the chains in order to avoid sensitivity to the initial "burn-in" phase.

To assess whether an MCMC run has converged, we use the criterion presented by Gelman et al. (1995). For any given parameter θ we define

$$R(\theta) = \left[\frac{\mathrm{var}(\theta)}{\frac{1}{J}\sum_{j=1}^{J} \mathrm{var}_j(\theta)} \right]^{1/2}, \tag{2.19}$$

where $\mathrm{var}_j(\theta)$ is the variance of θ in the individual chain j, and $\mathrm{var}(\theta)$ is the variance of θ over the entire set of J chains. Heuristically, R is a comparison of the variance of the entire distribution (the numerator) and the variance within individual chains (the denominator). The ratio will be greater than 1 for disjoint chains, and it will decrease and asymptotically approach 1 as the chains converge. Gelman et al. (1995) find that stopping an MCMC run once R reaches values below 1.2 provides a sufficient description of the target distribution for most samplings. To be conservative we run until $R \leq 1.1$ for *every* parameter. We then repeat the entire MCMC analysis a total of three times to obtain robust sampling of the target distribution.

In general we let all 11 model parameters vary simultaneously. The only exception is

the scale radius in models with an NFW dark matter halo. There is a strong covariance between r_s and other parameters, which produces a narrow, curved ridge in the likelihood surface that is difficult for MCMC algorithms to sample efficiently. To deal with this challenge, we discretely sample r_s in logarithmic steps ranging from $0.1''$ to $1000''$. We checked that the median values of parameters in the posterior distributions do not vary more than 5% between steps. To combine results from individual r_s runs into the final posterior distribution, we need to normalize the individual results properly using eq. (2.13). In particular, we need to determine the normalization factor

$$P(d|M) = \int P(d|\theta, M)\ P(\theta|M)\ d\theta\,, \tag{2.20}$$

which is known as the marginal likelihood or Bayesian Evidence. Computing this integral usually requires techniques that are more computationally intensive than basic MCMC sampling, such as thermodynamic integration (e.g., Marshall et al. 2003; Lartillot & Phillipe 2006). However, it is possible to obtain a simple and useful surrogate for the evidence using the Bayesian Information Criterion,

$$\mathrm{BIC} = -2\ln\mathcal{L}_{max} + k\ln N\,, \tag{2.21}$$

where $\mathcal{L}_{max}$ is the maximum likelihood of the model, k is the number of free parameters, and N is the number points used in the fit. While the BIC does not provide a proper statistical treatment of the evidence, it has been widely used in statistics and astronomy (e.g., Schwarz 1978; Rapetti et al. 2007; Davis et al. 2007; Liddle 2007) and is sufficient for this study.

As discussed in §2.4, we first examine models constrained only by the positions of the strongly lensed images, and then add supplemental constraints from weak lensing and various other considerations. In practice, this means we run the MCMC analysis to sample the likelihood eq. (2.14) based on the goodness of fit χ^2_p from eq. (2.15). (Since we use uniform priors, the posterior probability distribution is proportional to the likelihood.) Now suppose we want to add some supplemental constraints characterized

by their own goodness of fit χ^2_s. The total χ^2 is just the sum $\chi^2_p + \chi^2_s$ (i.e., χ^2 values add and likelihoods multiply), so we now want to sample

$$P(\theta|d, M) \propto e^{-\chi^2_s/2} e^{-\chi^2_p/2}. \tag{2.22}$$

Since the MCMC analysis provides a set of points drawn from $e^{-\chi^2_p/2}$, all we need to do is take those points and re-weight them by a factor proportional to $e^{-\chi^2_s/2}$. This provides a simple way to apply additional constraints without having to repeat the full MCMC analysis.

2.3.5 Weak Lensing Analysis

In parallel with the strong lensing analysis, we have conducted a weak lensing analysis of the full $6' \times 6'$ map to constrain the mass sheet and other environmental terms in the lensing potential (see eq. 2.11). The techniques and results of our weak lensing analysis are presented by Nakajima et al. (2009). We find the cluster potential to be consistent with a core-softened isothermal sphere profile, $\kappa(r) = \kappa_0[1+(r/r_c)^2]^{-1/2}$, with a best-fit central convergence $\kappa_0 = 0.47 \pm 0.17$ for a core radius $r_c = 5.0''$, corresponding to a velocity dispersion of (420 ± 70) km s^{-1} for $h = 0.70$. Additionally, we find the cluster to be consistent with an NFW profile but are unable to provide useful constraints on the cluster concentration and scale radius.

One product of our weak lensing analysis is the average convergence within $30''$ of the lens: $\kappa_{w,30''} = 0.166 \pm 0.056$. This represents the net convergence including contributions from both the lens galaxy and the cluster. As discussed in §2.3.1, to determine the Hubble constant we need to know the correction from the cluster mass sheet,

$$1 - \kappa_c = \frac{1 - \kappa_{w,30''}}{1 - \kappa_{s,30''}} \tag{2.23}$$

where $\kappa_{w,30''}$ is the net convergence inside $30''$ from our weak lensing measurement while $\kappa_{s,30''}$ is the contribution from our strong lensing model. Since we do not impose

a priori limits on $\kappa_{s,30''}$ in our models, the possibility exists that $\kappa_{s,30''}$ could exceed $\kappa_{w,30''}$. In this sense, our measurement of $\kappa_{w,30''}$ provides an upper bound on strong lens models. We penalize models with $\kappa_{s,30''} > \kappa_{w,30''}$ by adding an additional χ^2 term of the form

$$\chi^2_\kappa = \begin{cases} \frac{(\kappa_{s,30''}-\kappa_{w,30''})^2}{\sigma^2_{\kappa_{w,30''}}}, & \kappa_{s,30''} > \kappa_{w,30''} \\ 0, & \kappa_{s,30''} < \kappa_{w,30''} \end{cases} \tag{2.24}$$

When we apply the mass sheet correction to H_0 and mass model parameters, we need to account for the measurement uncertainties in $\kappa_{w,30''}$. We do this using Monte Carlo techniques. Specifically, we take each model from our MCMC runs and generate a distribution of values for $1-\kappa_c$ by drawing from a Gaussian distribution for $\kappa_{w,30''}$ set by the measurement and uncertainty from our weak lensing analysis. (The factor of $\kappa_{s,30''}$ in eq. 2.23 comes from the model itself.) We then fold this distribution into our final results reported in §§2.4.1–2.4.4.

The weak lensing analysis also yields constraints on the average shear in an annulus centered on the lens galaxy extending from $20''$ to $40''$: the two "Cartesian" shear components are $\gamma_c = -0.009 \pm 0.045$ and $\gamma_s = 0.092 \pm 0.045$, or equivalently the two "polar" components are $\gamma = 0.093 \pm 0.045$ and $\theta_\gamma = 47.8 \pm 13.9$ deg.[3] To compare this measurement to our models, we calculate the mean shear in the same annulus. After multiplying our value of the mean shear by the $1-\kappa_c$ correction, we impose the weak lensing results as constraints on the lens models through additional χ^2 terms. These constraints penalize models with unusually small or large shears.

2.4 Results

We first present results from lens models based on our new HST/ACS data (§2.4.1). We then consider adding additional constraints from the quasar radio flux ratio (§2.4.2), stellar population synthesis models (§2.4.3), and physical properties of NFW halos

[3]Note that the uncertainties are likely to be non-Gaussian for the "polar" components.

(§2.4.4). All values we report for the parameters $(\Upsilon_{\rm F606W}, b_d, \gamma, \sigma, \delta)$, as well as the dimensionless Hubble constant $h = H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1})$, are corrected for the mass sheet through the factor $1 - \kappa_c$ (including the associated uncertainties; see §2.3.5).

2.4.1 Basic Results: Strong and Weak Lensing

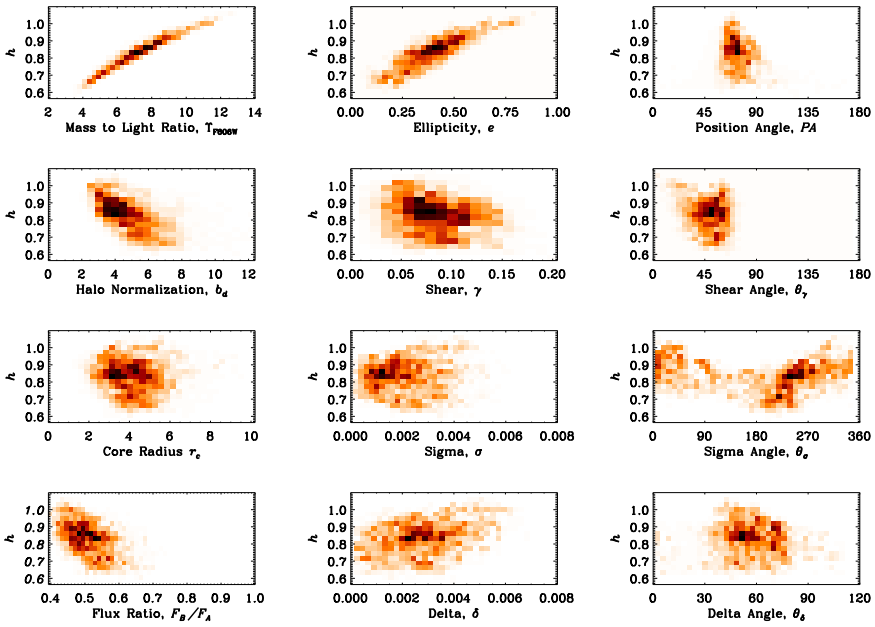

Figure 2.5 2D histograms depicting the marginalized joint probability distributions $P(\theta, h)$ for each model parameter θ and the dimensionless Hubble constant h. (For an explanation of parameters, see Table 2.3.) We also show the model flux ratio in the lower left panel. The colorscale is linear in the probability density, running from black at the peak to white at zero. These results are based on dark matter models with a softened isothermal profile ($\alpha = 1$).

For each of the four dark matter profiles we consider, we find a wide range of models that fit the HST strong lensing data well ($\chi^2_{reduced} < 1$).[4] For example, the stellar mass-to-light ratio can be anywhere in the range $\Upsilon_{\rm F606W} \sim 4$–12, the ellipticity $e \sim 0.1$–0.7, and shear $\gamma \sim 0.04$–0.12. Table 2.4 lists the median value and 68% and 95% confidence intervals for each model parameter (from the individual marginalized

[4] As shown in Table 2.4, the best-fit model in each model class has a reduced χ^2 somewhat less than unity. We used the χ^2 probability distribution to check that these values are statistically reasonable given the numbers of degrees of freedom.

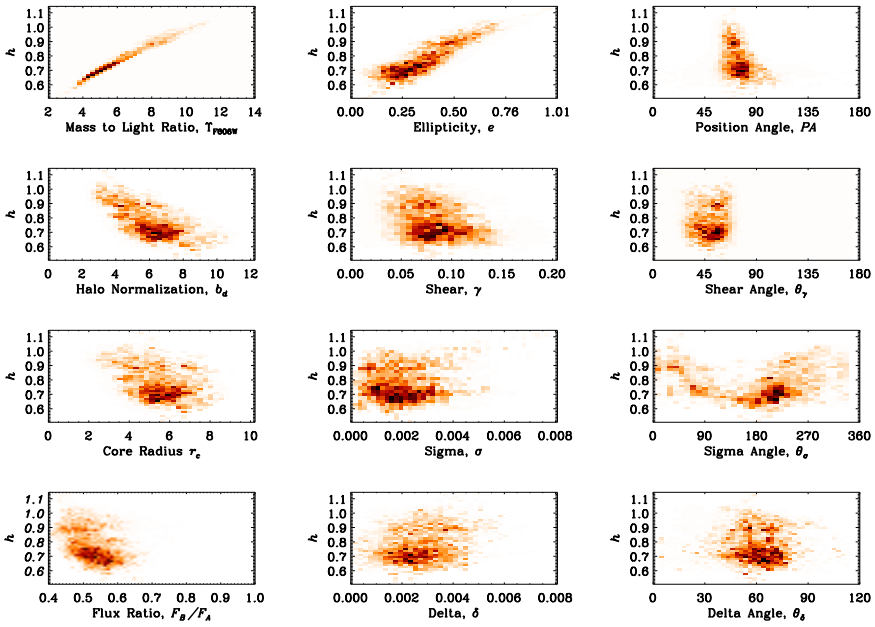

Figure 2.6 Similar to Fig. 2.5, but for models in which the dark matter halo has a power law profile that is steeper than isothermal ($\alpha = 0.5$).

posterior probability distributions). The Table also lists the relative probabilities of our four dark matter models obtained by integrating the posterior probability distributions over h, after weighting by the BIC. The range of allowed models is striking given that we now have so many strong lensing constraints. Examining the relative probabilities, we find an isothermal ($\alpha = 1$) dark matter profile is favored from lensing alone. It is interesting to note, however, that our isothermal models require a values of $\kappa_{s,30''}$ which are larger than, but still consistent with our weak lensing measurements.

There are various ways to examine the results, so let us begin with the Hubble constant. Figures 2.5–2.8 show the marginalized joint probability distributions $P(\theta, h)$ for each model parameter θ and the dimensionless Hubble constant h, for all four classes of dark matter models. Viewing the results this way helps reveal any important degeneracies or systematics that affect the inferred value of h. The most obvious feature is a strong degeneracy between h and the stellar mass-to-light ratio, $\Upsilon_{\rm F606W}$, which we discuss momentarily. Focusing on h itself, Figure 2.9 shows the marginalized cumulative

Table 2.4 Model Results: HST-ACS data

Median model values with 68% CL (95% CL) uncertainties

Parameter	NFW		$\alpha = 0.5$		$\alpha = 1.0$		$\alpha = 1.5$	
$\Upsilon_{\rm F606W}$	$7.7^{+3.3}_{-2.7}$	$(^{+4.9}_{-4.1})$	$5.4^{+2.5}_{-0.6}$	$(^{+3.9}_{-1.2})$	$7.4^{+1.5}_{-1.8}$	$(^{+3.0}_{-2.7})$	$8.6^{+1.1}_{-1.5}$	$(^{+2.2}_{-2.6})$
b_d ($''$)	$0.097^{+0.01}_{-0.022}$	$(^{+0.118}_{-0.034})$	$6.3^{+1.6}_{-1.4}$	$(^{+2.6}_{-2.3})$	$4.4^{+1.5}_{-0.8}$	$(^{+2.6}_{-1.5})$	$3.1^{+0.6}_{-0.5}$	$(^{+1.0}_{-0.8})$
e	$0.46^{+0.19}_{-0.19}$	$(^{+0.30}_{-0.28})$	$0.31^{+0.14}_{-0.13}$	$(^{+0.22}_{-0.18})$	$0.40^{+0.12}_{-0.14}$	$(^{+0.24}_{-0.22})$	$0.44^{+0.15}_{-0.11}$	$(^{+0.24}_{-0.17})$
θ_e ($^\circ$)	70^{+7}_{-6}	$(^{+12}_{-11})$	72^{+9}_{-10}	$(^{+14}_{-154})$	71^{+9}_{-7}	$(^{+14}_{-18})$	72^{+9}_{-6}	$(^{+14}_{-11})$
γ ($\times 10^2$)	$7.8^{+3.0}_{-2.7}$	$(^{+5.0}_{-4.3})$	$7.6^{+2.1}_{-2.8}$	$(^{+3.5}_{-4.0})$	$7.8^{+3.1}_{-2.5}$	$(^{+5.2}_{-4.0})$	$7.5^{+4.2}_{-3.4}$	$(^{+5.8}_{-5.2})$
θ_γ ($^\circ$)	52^{+11}_{-15}	$(^{+15}_{-25})$	51^{+11}_{-13}	$(^{+16}_{-20})$	51^{+10}_{-14}	$(^{+15}_{-25})$	42^{+16}_{-19}	$(^{+26}_{-33})$
a ($''$)	–	–	$5.7^{+1.2}_{-1.1}$	$(^{+2.0}_{-2.0})$	$4.0^{+1.0}_{-0.9}$	$(^{+1.7}_{-1.4})$	$1.8^{+0.5}_{-0.7}$	$(^{+1.0}_{-1.0})$
σ ($\times 10^3$)	$2.6^{+1.7}_{-1.3}$	$(^{+2.7}_{-1.9})$	$1.9^{+1.1}_{-0.9}$	$(^{+1.8}_{-1.3})$	$1.8^{+1.3}_{-0.9}$	$(^{+2.0}_{-1.3})$	$1.9^{+1.0}_{-1.4}$	$(^{+1.6}_{-1.5})$
θ_σ ($^\circ$)	335^{+5}_{-291}	$(^{+14}_{-322})$	210^{+46}_{-129}	$(^{+95}_{-180})$	233^{+61}_{-184}	$(^{+106}_{-218})$	266^{+51}_{-197}	$(^{+85}_{-242})$
δ ($\times 10^3$)	$2.8^{+1.8}_{-1.1}$	$(^{+3.0}_{-1.8})$	$2.6^{+1.0}_{-1.0}$	$(^{+1.8}_{-1.6})$	$2.4^{+1.5}_{-1.0}$	$(^{+2.6}_{-1.7})$	$3.3^{+0.9}_{-1.4}$	$(^{+1.9}_{-2.0})$
θ_δ ($^\circ$)	68^{+9}_{-17}	$(^{+17}_{-28})$	61^{+11}_{-13}	$(^{+18}_{-21})$	59^{+17}_{-13}	$(^{+21}_{-21})$	62^{+10}_{-5}	$(^{+15}_{-18})$
f_B/f_A	$0.55^{+0.10}_{-0.06}$	$(^{+0.22}_{-0.08})$	$0.53^{+0.05}_{-0.05}$	$(^{+0.09}_{-0.08})$	$0.50^{+0.06}_{-0.04}$	$(^{+0.09}_{-0.07})$	$0.51^{+0.07}_{-0.04}$	$(^{+0.12}_{-0.06})$
$\kappa_{s,30''}$	$0.12^{+0.07}_{-0.02}$	$(^{+0.10}_{-0.05})$	$0.15^{+0.07}_{-0.05}$	$(^{+0.13}_{-0.08})$	$0.21^{+0.07}_{-0.07}$	$(^{+0.11}_{-0.10})$	$0.27^{+0.04}_{-0.04}$	$(^{+0.07}_{-0.06})$
H_0	$89.4^{+16.8}_{-13.7}$	$(^{+27.3}_{-20.8})$	$75.3^{+13.8}_{-10.0}$	$(^{+23.4}_{-15.3})$	$84.4^{+11.1}_{-11.3}$	$(^{+18.3}_{-17.9})$	$89.9^{+10.1}_{-10.0}$	$(^{+17.0}_{-16.0})$
$\min(\chi^2_{reduced})$	0.38		0.91		0.90		0.58	
P_{rel}	0.014		0.575		1.000		0.114	

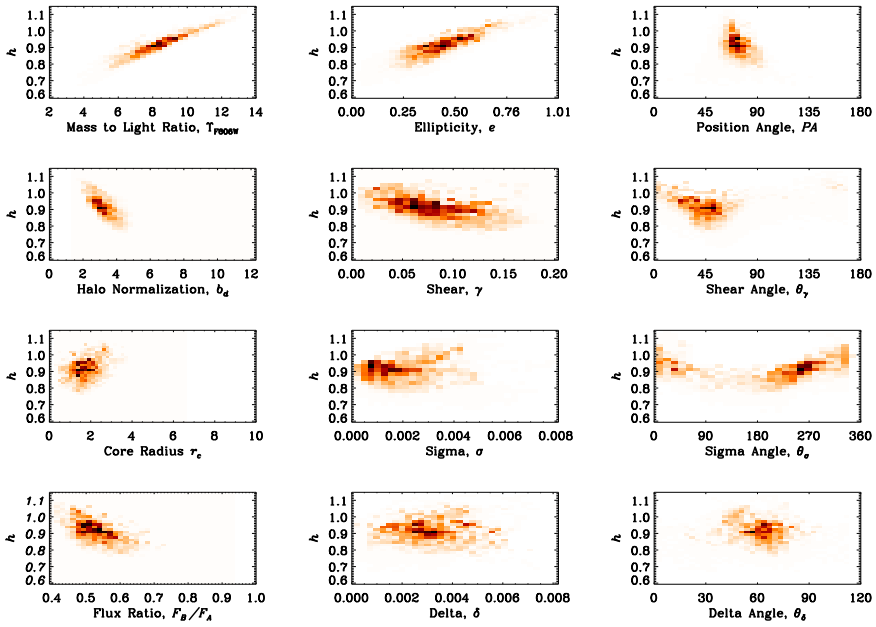

Figure 2.7 Similar to Fig. 2.5, but for models in which the dark matter halos has a power law profile that is shallower than isothermal ($\alpha = 1.5$).

posterior probability distribution for h from each of our dark matter models. Combining the four models (weighted by their relative probabilities), we find $H_0 = 85^{+14}_{-13}$ km s^{-1} Mpc^{-1} (68% CL). Our measurement of H_0 is somewhat higher than, but statistically consistent with, the recent determinations of $H_0 = 74.2 \pm 3.6$ km s^{-1} Mpc^{-1} from SNe (Riess et al. 2009) and $H_0 = 70.5 \pm 1.3$ km s^{-1} Mpc^{-1} from WMAP5+BAO+SNe (Komatsu et al. 2008). Compared with previous results from Q0957 (Bernstein & Fischer 1999; Keeton et al. 2000), our initial results have lowered the median[5] value from $\sim$ 90 to 85 and reduced the spread by $\sim$ 30%. The latter result is significant given the new complexity in our models, the relaxation of (previously tight) quasar and jet positions, and the elimination of the quasar flux ratio constraint.

The degeneracy between h and the stellar mass-to-light ratio $\Upsilon_{\rm F606W}$ arises because the total mass within the Einstein radius is fixed, so varying $\Upsilon_{\rm F606W}$ changes the balance between the concentrated stellar component and the more diffuse dark matter halo.

[5]The old median value for H_0 is not very well determined, because the previous analyses did not include the proper statistical sampling provided by our MCMC analysis.

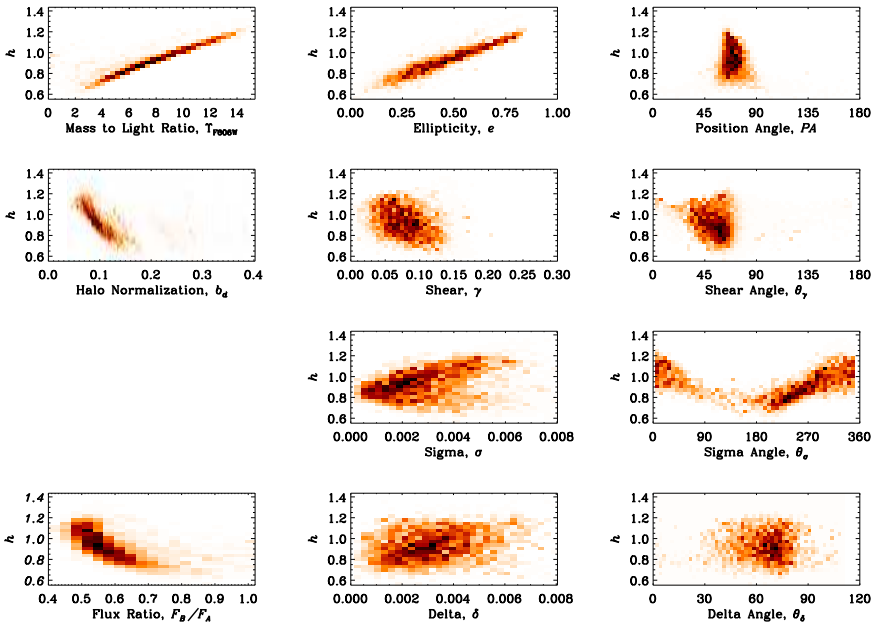

Figure 2.8 Similar to Fig. 2.5, but for models with an NFW dark matter halo. Here we do not show $P(r_s, h)$ since we discretely sample the scale radius (see §2.3.4).

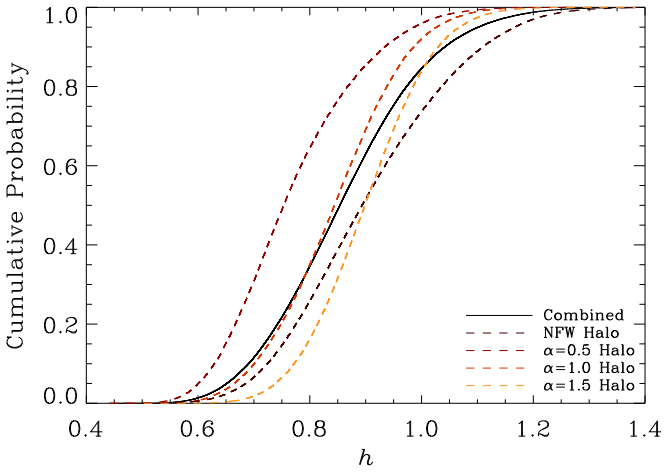

Figure 2.9 Cumulative posterior probability distribution for the dimensionless Hubble constant h for our four dark matter halo profiles. The solid black curve shows the result of combining these distributions, weighted by their relative probabilities.

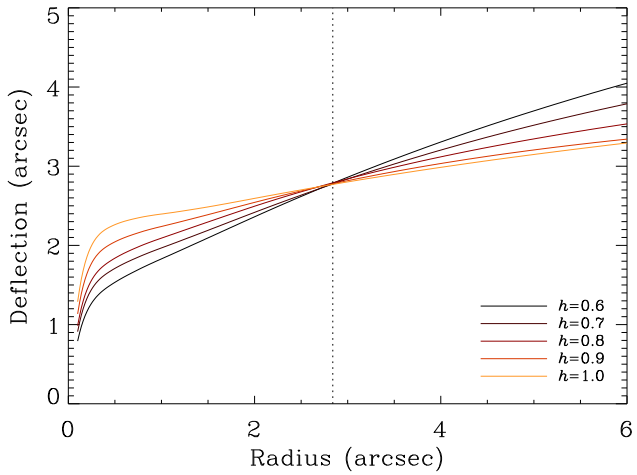

Figure 2.10 Monopole deflection profile for models with $h = 0.60$, 0.70, 0.80, 0.90 and 1.00. We actually plot the mean profile for all models within ± 0.005 of the nominal h value; the scatter among such models is $< 2\%$ across all radii and does not depend systematically on the particular dark matter profile. The vertical dotted line indicates the Einstein radius at $R_{ein} = 2.84''$. The scatter among models drops to $< 0.5\%$ in the vicinity of R_{ein}, indicating a robust and tight constraint on the Einstein radius.

That, in turn, modifies the slope of the total density profile, which is known to be the main factor that determines h (e.g., Williams & Saha 2000; Kochanek 2002). To illustrate these effects, we examine the monopole deflection curve, $\alpha(r) \propto M(r)/r$ where $M(r)$ is the projected mass within radius r (e.g., Blandford & Narayan 1986; Blandford & Kochanek 1987; Cohn et al. 2001; Kochanek et al. 2006; van de Ven et al. 2009). This is a 2D analog of the rotation curve. A flat deflection curve corresponds to an isothermal profile, while a rising (falling) curve corresponds to a profile shallower (steeper) than isothermal. Figure 2.10 shows the deflection curves for models with different values of h. There is a systematic change in the slope of the deflection curve with h, with very little ($< 2\%$) scatter among models at a fixed value of h. In other words, even though Q0957 has a complex lens potential, we still recover the familiar result that the slope of the total density profile is what principally determines h. In our models, the slope of the density profile is governed by the stellar mass-to-light ratio, which makes $\Upsilon_{\rm F606W}$ the key parameter responsible for the range of h values. We consider external

constraints on $\Upsilon_{\rm F606W}$ in §2.4.3 below.

One interesting aspect of Figure 2.10 is that Q0957 appears to have a rising deflection curve, corresponding to a total density profile that is shallower than isothermal, for any reasonable value of h. Many other lens galaxies have profiles that are closer to isothermal (e.g., Cohn et al. 2001; Rusin & Kochanek 2005; Koopmans et al. 2006). Q0957 is not, however, unique in this regard: Kochanek et al. (2006) found that HE 0435−1223 also has a rising deflection curve. They argue that different density profiles and deflection curves can arise as a consequence of how galaxies populate dark matter halos. In the halo model (see Cooray & Sheth 2002 for a review), a group or cluster of galaxies consists of a massive central galaxy surrounded by smaller satellite galaxies. Lying as it does at the center of the potential well, the central galaxy should have a higher dark matter surface density compared to its satellite neighbors, which would lead to a more diffuse mass distribution with a shallower profile and hence a rising deflection curve. In this context, Kochanek et al. (2006) argue that HE0435 may be the central galaxy in a group of galaxies. Q0957 seems to fit naturally into this picture because it lies at or near the center of a modest cluster of galaxies (see §2.3.2).

One useful way to characterize a strong lens system is with the lensing critical curves and caustics. The critical curves reveal highly magnified regions in the image plane, and the corresponding caustics separate regions in the source plane that lead to different numbers of lensed images. Figure 2.11 shows examples of the critical curves for our models of Q0957. It is clear that the newly identified HST/ACS features have tightly constrained where the critical curve lies, especially southeast of quasar B. Significant variations do still exist near the ends of the critical curves, suggesting that constraints from the faint, unused features indicated in Figure 2.4 could help to further constrain the critical curves and restrict the parameter space. Unfortunately, in the current data these features have a signal to noise ratio less then 3, making them not sufficiently reliable.

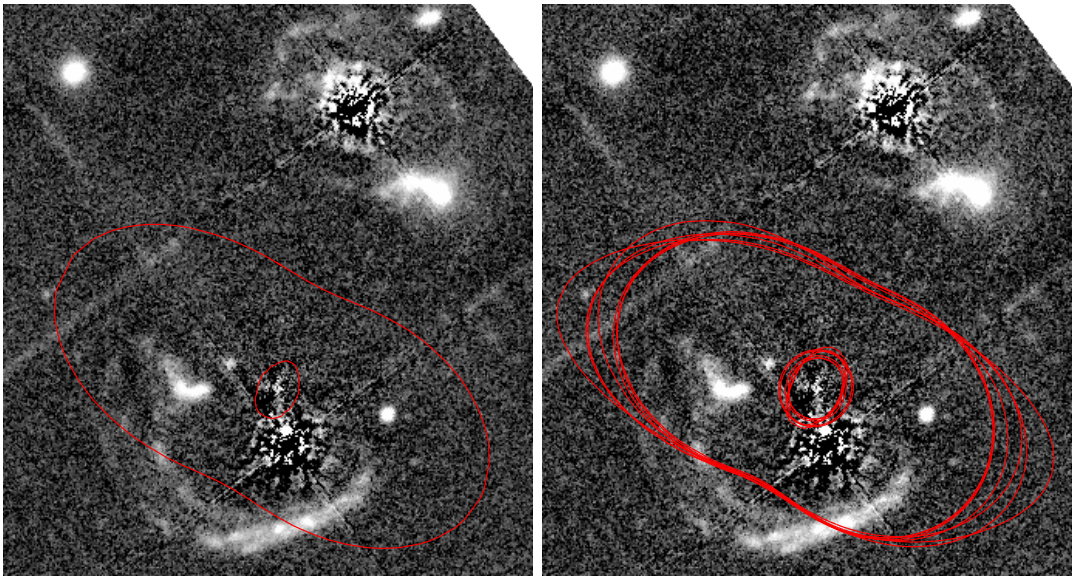

Figure 2.11 *(a, left)* A typical critical curve resulting from the new image constraints in Table 2. Notice the fold image pairs that the curve runs though South and East of quasar B. *(b, right)* Critical curves corresponding to models with minimal/maximal values of different parameters, for our models with an NFW dark matter halo (results are similar for other model classes). The critical curves show little variation along the semi-minor axis due to strong constraints from new images. Significant variation still exists along the semi-major axis of the curves.

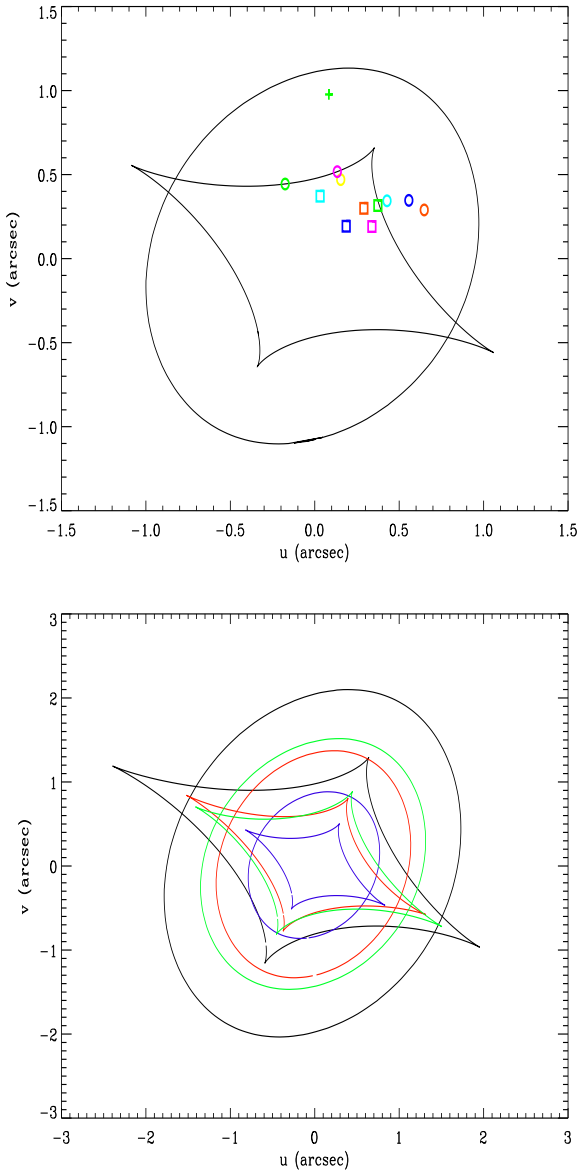

Figure 2.12 *(a, left)* Source plane caustic corresponding to the critical curve shown in Fig. 2.11a. The points show the positions of the sources corresponding to the observed images, with the same shape and color scheme as in Fig. 2.4. *(b, right)* Source plane caustics corresponding to the critical curves shown in Fig. 2.11b. The main variation is an overall rescaling due to different values of the convergence in the vicinity of the Einstein radius.

In addition to $\Upsilon_{\rm F606W}$, our models exhibit a second degeneracy between h and the ellipticity of the dark matter halo. While the exact origin of this is unclear, it is likely connected to the degeneracy in $\Upsilon_{\rm F606W}$. Since the stellar component has a fixed angular structure whose position angle ($\sim 67°$–$82°$, Fig. 2.3) does not quite match the necessary angular structure of the lens potential (with position angle closer to $63°$, see Fig. 2.11), the dark matter component needs to make up the difference. The amount of compensation increases as the mass of the stellar component increases, so the halo ellipticity rises with $\Upsilon_{\rm F606W}$.

Figure 2.12 shows the corresponding caustics in the source plane. Generically, the tangential caustic is elongated along a roughly NE-SW direction and extends beyond the radial caustic. There is some variation among the models, but the main effect is just an overall rescaling by $1 - \kappa_E$, where κ_E is the convergence at the Einstein radius (which is related to $\kappa_{s,30''}$). Many of our newly discovered sources are inferred to lie within the tangential caustic and should therefore have four images. Since we did not necessarily identify these as quad systems in our original detections (§2.2.3), it is interesting to examine the predicted counter-images. Figure 2.13 shows all the predicted images of the quadruply-imaged sources, for comparison with the detected images shown in Figure 2.4. (We show results for one particular model, but results for other models are similar.) We see that there are some predicted images that lie in relatively blank regions West of quasar B and South of quasar A. This is not a concern, however, because the undetected images have magnifications that are a factor of 10–100 smaller than the detected images lying to the East of quasar B, so their predicted fluxes lie well below the noise level of the HST image. Some of the predicted counter-images (among the ones indicated by diamonds) are not so far below the noise, but they still lack clear peak positions and thus cannot currently provide further robust constraints on lens models.

We consider one additional source of strong lensing constraints, namely extended

images of the quasar host galaxy observed with HST/NICMOS by Keeton et al. (2000). Following Keeton et al., we analyze the arcs by taking the resolved arc around quasar A, mapping it pixel-by-pixel to the source plane using the lens model of interest, then mapping the reconstructed source back to the image plane to predict the arc around quasar B. Generally, we find that all models generated from our ACS data reproduce the NICMOS arcs comparably well; the NICMOS arcs do not restrict the range of models significantly better than the ACS data. We infer that the ACS data have captured most of the information present in the NICMOS arcs, which is not surprising given that the ACS and NICMOS data span similar spatial regions and (presumably) both come from the lensed quasar host galaxy. Compared with the NICMOS arcs, the ACS data are somewhat cleaner to interpret because they avoid complicated interpolations to and from the image plane and offer a more straightforward counting of degrees of freedom. We take the compatibility of the NICMOS and ACS data as additional reassurance that the mapping of faint lensed features (§2.2.3) has been done correctly.

2.4.2 Quasar Flux Ratio

So far we have only considered image positions as lensing constraints. In order to include some information about the lensing magnification, we consider measurements of the quasar flux ratio. Bernstein & Fischer (1999) and Keeton et al. (2000) used the delay corrected VLA measurements at 4 cm and 6 cm by Haarsma et al. (1999) to constrain the quasar core flux ratio. While the VLA cannot resolve out the relative contributions of radio compontents, if the jets are invariant on decadal time scales, the ratio of the radio *fluctuations* gives a measurement of the core flux ratio. With this assumption, Haarsma et al. derive the core flux ratio to be 0.74 ± 0.02. Before applying this constraint to our lens models, we must consider whether our models should be expected to fit the observed flux ratio. The issue is whether dark matter substructure may perturb the flux ratio (e.g., Mao & Schneider 1998; Metcalf & Madau 2001; Dalal

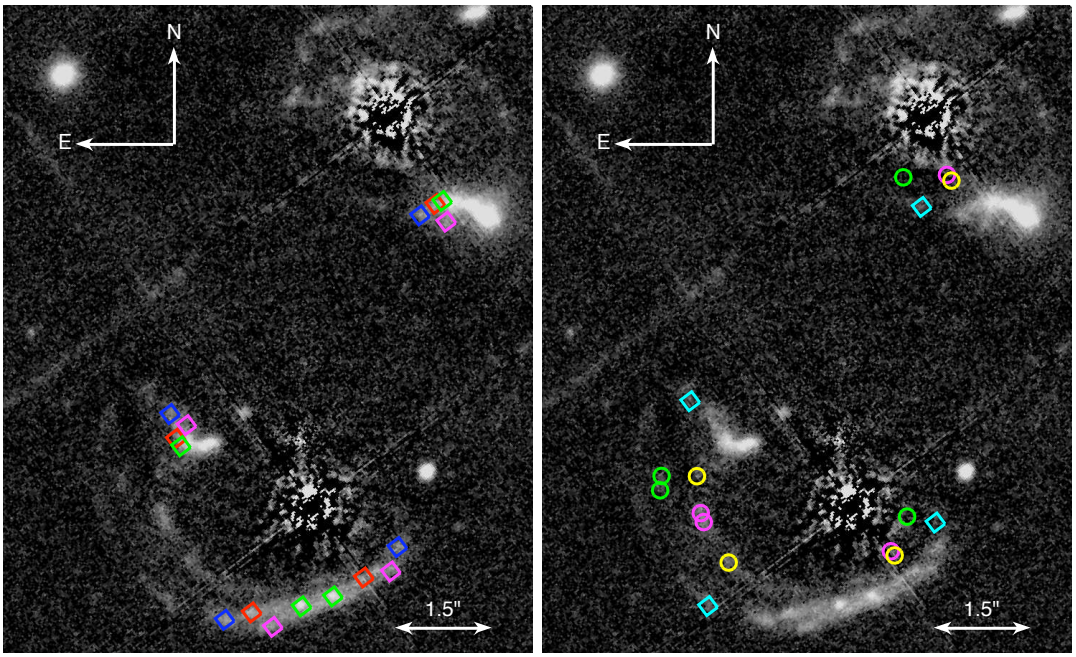

Figure 2.13 Predicted images of all sources that lie inside the tangential caustic for our lens models. Results are shown for the same model as Figure 2.12a but are similar for other models. The two panels correspond to different sets of sources, with the same arrangement as Figure 2.4. There are some predicted images shown here that have not been detected (i.e., they do not appear in Fig. 2.4); they are predicted to be below the noise in our HST data.

& Kochanek 2002; Chiba et al. 2005; MacLeod et al. 2009) in a way that our smooth models cannot reproduce.

High-resolution VLBI measurements show that the quasar images are $\lesssim 1.2$ mas in size (Gorenstein et al. 1988b), which corresponds to a size for the emission region in the source plane of $\lesssim 0.9$ mas or $\lesssim 5.4\,h^{-1}$ pc. We use the methods of Dobler & Keeton (2006) to estimate how a source of this size would be affected by an isothermal sphere clump placed near one of the images. We find that a clump of mass $\gtrsim 10^6\,M_\odot$ can easily change the lensing magnification by a factor of order unity, and N-body simulations predict such subhalos to be abundant ($\gtrsim 10^3$) in a galaxy with a mass of $\sim 10^{13}\,M_\odot$ (e.g., Springel et al. 2005; Angulo et al. 2008). Apparently we should not discount the possibility that substructure plays a significant role in the observed VLA flux ratio.

Our basic models generally predict a flux ratio in the vicinity of ~ 0.5 with at most a tail extended to the range of the VLA measurement (see Figs. 2.5–2.8 and Table 2.4). The discrepancy could be interpreted as evidence that the VLA flux ratio is indeed perturbed by substructure. Further support for this hypothesis comes from the fact that the magnification ratio inferred from the resolved radio jets is different from the ratio for the quasar cores, and closer to the smooth model prediction (Bonometti 1985; Gorenstein et al. 1988b; Conner et al. 1992). We should be careful, of course, not to think that a measurement that disagrees with our smooth models is "wrong" and one that agrees is "right" — or to assume that any discrepancy involving a flux ratio automatically implies dark matter substructure. Nevertheless, we conclude that existing evidence shows the flux ratio to be very intriguing and worthy of further study, both on its own and as possible evidence for substructure in Q0957.

With this in mind, it is not clear how strongly we should impose the VLA flux ratio as a constraint on our models. We consider using the measurement but inflating the errorbar by various factors to obtain a range of constraints from strong to weak. Figure 2.14 shows the marginalized cumulative posterior probability distribution for h for the

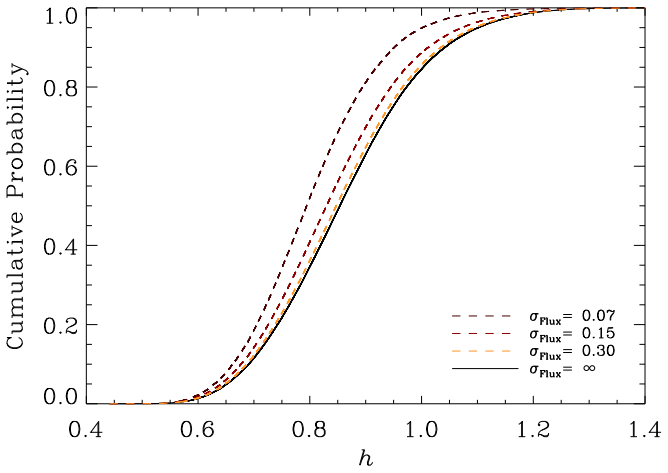

Figure 2.14 Cumulative posterior probability distribution for the dimensionless Hubble constant h with different assumptions about the constraint from the quasar radio flux ratio. The solid black line shows our fiducial results using the image positions alone (from Fig. 2.9). The dashed lines show how the results change when we impose constraints from the VLA flux ratio constraints (Haarsma et al. 1999) with the uncertainties increased by a factor of 3.5, 7.5, and 15, representing different levels at which effects from substructure might be understood. These values are consistent with a increase in the magnification of image B by a factor of 1.35, 1.20, and 1.0 due to a subhalo of mass $\gtrsim 10^6 M_\odot$. While adding flux constraints clearly tightens the h distribution, we do not include them in our final h distribution since we do not currently understand the extent to which substructure may be important in Q0957.

different cases. The flux ratio constraint tends to reduce the median value of h and tighten the distribution. While such a result seems enticing, we caution that it may be artificial if the flux ratio is really perturbed by substructure which is absent from our model. Given the concerns, we choose not to use the flux ratio as a constraint for our final results. If there were some way to determine the "macro" flux ratio, however, that might help improve constraints on h.

2.4.3 Stellar Mass to Light Ratio

As shown in Figures 2.5–2.8, our models demonstrate a strong correlation between h and $\Upsilon_{\rm F606W}$, the stellar mass-to-light ratio, as a consequence of the radial profile degeneracy in lensing. Any external constraints on $\Upsilon_{\rm F606W}$ could not only narrow the

overall parameter space but also tighten constraints on h. It is possible to obtain such constraints from stellar population synthesis (SPS) models thanks to the fact that we have excellent HST photometry of a relatively "simple," old stellar population in the lens galaxy.

We use two sets of SPS models. The first are the Flexible Stellar Population Synthesis (FSPS) models of Conroy et al. (2009). These models exhibit enormous flexibility and are aimed at addressing many of the uncertainties of SPS models. In particular, the user can not only consider the traditional effects of varying the star formation history, star formation epoch, metallicity, initial mass function (IMF), and dust, but also account for various treatments of the thermally pulsating asymptotic giant branch (TP-AGB), blue stragglers, and blue horizontal branch stars. See Conroy et al. (2009) for details of the FSPS models and the SPS uncertainties they address. Using the FSPS models, we follow the treatment of Conroy et al., adjusting 9 models parameters (see Table 2.5) that include the effects of varying the epoch of star formation, the star formation history, and dust. We consider six metallicities from 50% to 160% Solar using a Chabrier IMF (Chabrier 2003).

The second set of SPS models we use are from Maraston et al. (2009). These models are based on the same simple stellar populations (SSPs) as Maraston (2005), which treat the TP-AGB contribution to the SEDs, but also include a metal poor ($[Z/H] = -2.2$) population comprising 3% of the mass. As shown by Maraston et al. (2009), these models provide a good fit to the optical colors of galaxies in the Luminous Red Galaxy (LRG) sample from the Sloan Digital Sky Survey. This is encouraging because SPS models have historically had trouble fitting LRG colors (e.g., Eisenstein et al. 2001; Wake et al. 2006). Since the lens galaxy in Q0957 is luminous ($L \sim 6.5L_{\star}$) and red ($m_{F606W} - m_{F814W} = 1.057$), we postulate that these models should provide a good fit the galaxy's spectral energy distribution. To account for variation of the lens galaxy from the SSPs of Maraston et al., we allow for variation in the redshift at which star

Table 2.5. SPS Model Parameters

Parameter	Prior	FSPS	Maraston SPS
Formation redshift, z_f	$0.361 - 1089$	✓	✓
SFR e-folding time, τ	$0 - \infty$	✓	✓
Constant SFR, C	$0 - 1$	✓	✓
Dust around young stars, τ_1	$0 - \infty$	✓	✓
Diffuse dust,[1] τ_2	$P \propto e^{-1.086\tau_2}$	✓	✓
Fraction of blue HB stars,[2] f_{BHB}	$0 - 0.5$	✓	–
Specific frequency of blue stragglers,[2] $S_{\rm BS}$	$0 - 10$	✓	–
Shift in $log(L_{bol})$ along the TP-AGB,[2] Δ_L	$-0.4 - 0.4$	✓	–
Shift in $log(T_{eff})$ along the TP-AGB,[2] Δ_T	$-0.2 - 0.2$	✓	–

[1]We use a exponential prior on the diffuse dust content of the form $e^{-A_{F606W}/1.0\rm mag}$

[2]See Conroy et al. (2009) for details

formation begins, the star formation history, and dust. As in our FSPS models, we allow star formation to begin anywhere from the CMB redshift, $z = 1089$, to the redshift of the lens, $z = 0.361$. We adopt the star formation rate

$$\begin{aligned} \Psi(t) &= \frac{1-C}{\tau}\frac{e^{-t/\tau}}{e^{-T(z_{form})/\tau} - e^{-T(z_l)/\tau}} \\ &\quad + \frac{C}{T(z_l) - T(z_{form})}, \qquad T(z_{form}) \leq t \leq T(z_l), \end{aligned} \tag{2.25}$$

where C is the fraction of stars formed at a constant rate, τ is the e-folding time of the star formation rate, and $T(z)$ denotes the age of the universe at redshift z. This form of the star formation rate has the advantage of smoothly varying from a SSP to a constant star formation history. For dust, we consider the two-parameter model of Charlot & Fall (2000) which includes the effect of dust around young stars as well a diffuse dust component. Parameters for our Maraston SPS models are summarized in Table 2.5.

As constraints on the SPS models, we use our measurements of the F606W and F814W magnitudes of the lens galaxy (see Table 2.1) together with a reanalysis of the

NICMOS/F160W image obtained by Keeton et al. (2000). The revised AB magnitude $m_{\rm F160W,AB} = 16.84 \pm 0.15$ (Chien Peng, private communication) differs somewhat from the originally reported value because it is based on a more sophistocated deconvolution of the lens galaxy from the quasar images and host arcs. We correct for Galactic extinction using the methods of Schlegel et al. (1998), finding the value of $E(B-V) = 0.0095$.

Initial modeling found that large values of dust extinction, extending to $A_{\rm F606W} > 2.0$ mag, were allowed in the SPS models. This was unexpected since early type galaxies are known to have modest dust content (e.g., Schawinski et al. 2007). In lensed systems, differential extinction measurements have shown that lenses typically exhibit smaller values of dust extinction than found in nearby, late-type galaxies: Elíasdóttir et al. (2006) find a mean extinction of $A_V = 0.56 \pm 0.04$ in a sample of 10 lenses (also see Falco et al. 1999). For Q0957 we could in principle rely on previous attempts to measure the dust content of the lens galaxy, but the results are puzzling. Goicoechea et al. (2005) used HST/STIS observations to measure the flux ratio $F_B/F_A > 1$ at optical and ultraviolet wavelengths (also see the delay corrected ratios of Colley et al. 2003), which stands in stark contrast to the VLA measurement $F_B/F_A = 0.74 \pm 0.02$ and our models predictions. To explain this difference Goicoechea et al. invoke dust clouds in front of image A leading to extinction $A_V = 0.30$. It is counterintuitive to think that image A (at $18.6\,h^{-1}$ kpc from the center of the galaxy) is more heavily extinguished than image B (just $3.7\,h^{-1}$ kpc from the center). If extinction is indeed the cause of the wavelength dependence in the flux ratios, it remains unclear how the dust is distributed throughout the rest of the galaxy, whether it is clumpy and extends to large radii in other directions. For all these reasons, we choose not to constrain the dust in the SPS models to a particular value. Nevertheless, in order to avoid unreasonable large extinction values we impose a weak, exponential prior of the form $e^{-A_{F606W}/1.0{\rm mag}}$.

In order to derive constraints on $\Upsilon_{\rm F606W}$, we set up an MCMC analysis similar to

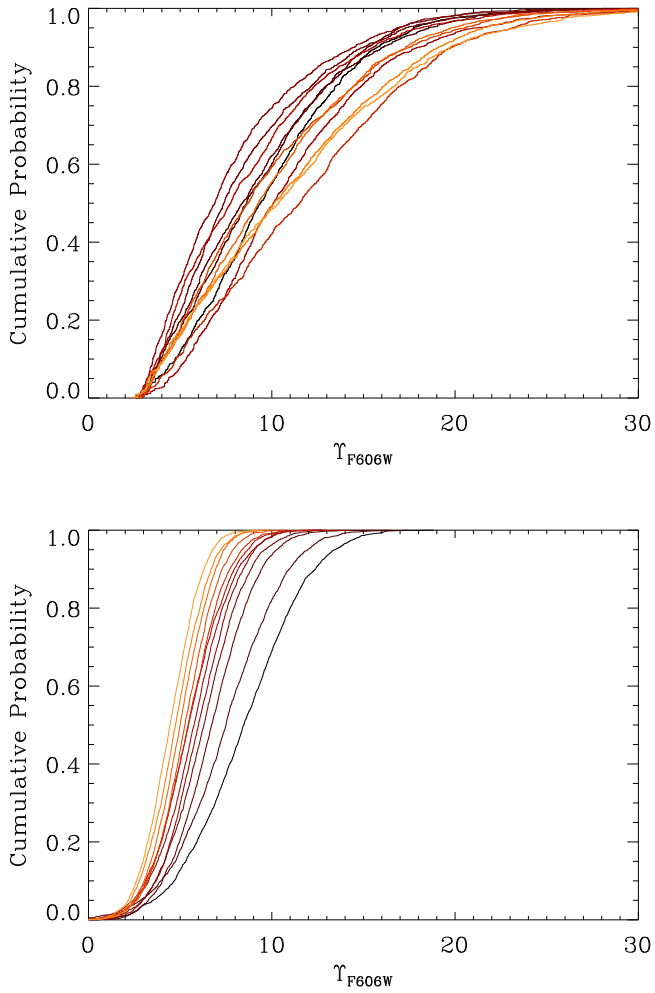

Figure 2.15 Cumulative posterior probability distributions for the stellar mass to light ratio, $\Upsilon_{\rm F606W}$. The different curves correspond to different values of h, varying from 0.55 (black) to 1.10 (light orange) in increments of 0.05. *(a, left)* Results from the FSPS models of Conroy et al. (2009). *(b, right)* Results from the SPS models of Maraston et al. (2009). The FSPS models tend to produce higher values of $\Upsilon_{\rm F606W}$ with more scatter than the Maraston models. Also, the FSPS models are not as *systematically* dependent on h, presumably because the large freedom in the models dominates the distribution of $\Upsilon_{\rm F606W}$.

what we use for lens models (see §2.3.4) to sample the posterior probability distribution. Since the value of $\Upsilon_{\rm F606W}$ depends on h (principally through the age of the universe as a function of redshift), we run the MCMC analysis for discrete values of h from 0.50 to 1.45 in steps of $\Delta h = 0.025$; the small steps ensure that the median and range of $\Upsilon_{\rm F606W}$ do not vary by more than 3% from one h value to the next, so we can interpolate accurately. Figure 2.15 shows the cumulative distributions for $\Upsilon_{\rm F606W}$ for different values of h, from both FSPS and Maraston models. In general, both models fit the observed F606W−F814W and F606W−F160W colors well ($\chi^2_{reduced} \leq 1$), but the FSPS models yield a larger range for $\Upsilon_{\rm F606W}$. This is not surprising given the amount of freedom available in the FSPS models. Previous studies of massive ellipiticals found values of Υ_B of $\sim 4-10$ (e.g., Gerhard et al. 2001; Grillo et al. 2009), in good agreement with the values found here.[6]

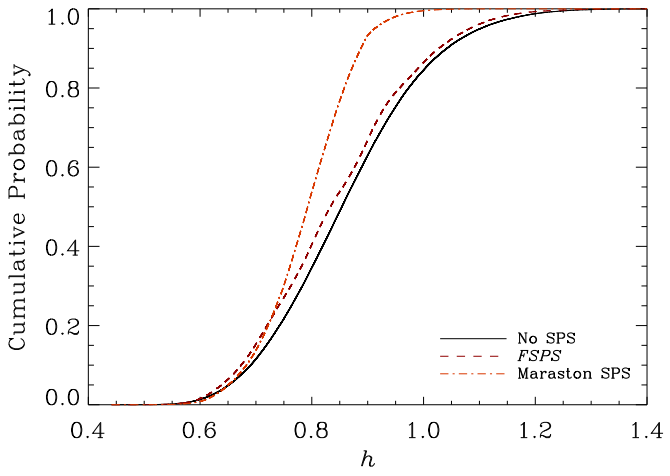

Figure 2.16 Cumulative posterior probability distribution for h with and without constraints from SPS models. Since the Maraston SPS models have been shown to fit LRG colors from the SDSS, we adopt them when quoting final values of h and model parameters. We find $H_0 = 79.3^{+6.7}_{-8.5}$ km s^{-1} Mpc^{-1} (68% CL).

Figure 2.16 shows what happens to the posterior probability distribution for the Hubble constant when we impose the SPS constraints on $\Upsilon_{\rm F606W}$. Using the FSPS

[6]At a redshift of $z = 0.36$, observed $\Upsilon_{\rm F606W}$ corresponds to rest frame Υ_B.

models has little effect on the h distribution, because these models allow for a large range of values for $\Upsilon_{\rm F606W}$. The Maraston models, by contrast, favor $\Upsilon_{\rm F606W} \sim 4.5$–$6.5$ and such values tend to reduce the median h and tighten the distribution. Clearly it is important to understand why the FSPS and Maraston models differ as to whether high values of $\Upsilon_{\rm F606W}$ are acceptable. Figure 2.17 shows that the high $\Upsilon_{\rm F606W}$ values attained in FSPS models correspond to large values of extinction — values that seem surprising for a luminous early-type galaxy in a modest cluster at redshift $z = 0.361$. We infer that the flexibility of FSPS models is allowing them to reproduce the observed colors of the galaxy even with models that do not make much sense astrophysically. One way to reconcile the FSPS and Maraston models is to constrain the amount of dust in our FSPS models. We find that adopting an extinction prior of $A_{F606W} = 0.45 \pm 0.2$ would bring the FSPS constraints on $\Upsilon_{\rm F606W}$ into agreement with those from the Maraston models. Imposing such a prior has little effect on the $\Upsilon_{\rm F606W}$ constraints from the Maraston models since those models show little or no correlation between dust extinction and $\Upsilon_{\rm F606W}$.

Ultimately we need to decide what to use for our final constraints on $\Upsilon_{\rm F606W}$. Since the Maraston models are constructed to match the SDSS LRG sample and require no ad hoc assumptions about the dust content of the lens galaxy, we elect to use them when reporting our final determination of h and model parameters (see Table 2.6). With these constraints on $\Upsilon_{\rm F606W}$ we find $H_0 = 79.3^{+6.7}_{-8.5}$ km s^{-1} Mpc^{-1} (68% CL).

2.4.4 Physical Properties of NFW Halos

When fitting models with an NFW dark matter halo, we previously took both r_s and $\kappa_s = \rho_s r_s / \Sigma_{crit}$ to be free parameters. However, in N-body simulations the two NFW parameters are actually related to one another, albeit with some scatter. We now consider whether our lens model parameters have reasonable values in general, and whether they are consistent with the correlation found in simulated halos.

Table 2.6 Model Results: HST-ACS data + Maraston SPS Models

	Median model values with 68% CL (95% CL) uncertainties							
Parameter	NFW		$\alpha = 0.5$		$\alpha = 1.0$		$\alpha = 1.5$	
$\Upsilon_{\rm F606W}$	$5.7^{+1.2}_{-1.2}$	$(^{+1.9}_{-1.9})$	$5.3^{+1.1}_{-0.9}$	$(^{+2.9}_{-1.5})$	$6.5^{+1.0}_{-1.8}$	$(^{+1.6}_{-2.0})$	$6.9^{+0.9}_{-1.0}$	$(^{+1.4}_{-1.4})$
b_d ($''$)	$0.118^{+0.028}_{-0.018}$	$(^{+0.107}_{-0.025})$	$6.4^{+1.5}_{-1.2}$	$(^{+2.5}_{-2.0})$	$5.0^{+1.3}_{-2.0}$	$(^{+2.3}_{-1.8})$	$3.5^{+0.5}_{-0.5}$	$(^{+0.9}_{-0.9})$
e	$0.34^{+0.09}_{-0.09}$	$(^{+0.14}_{-0.26})$	$0.30^{+0.6}_{-0.11}$	$(^{+0.16}_{-0.16})$	$0.32^{+0.10}_{-0.14}$	$(^{+0.115}_{-0.16})$	$0.34^{+0.08}_{-0.08}$	$(^{+0.11}_{-0.12})$
θ_e ($^\circ$)	70^{+8}_{-8}	$(^{+15}_{-13})$	73^{+8}_{-10}	$(^{+13}_{-154})$	72^{+9}_{-9}	$(^{+13}_{-15})$	76^{+8}_{-6}	$(^{+12}_{-11})$
γ ($\times 10^2$)	$8.6^{+3.0}_{-2.9}$	$(^{+4.5}_{-4.4})$	$7.7^{+2.8}_{-2.1}$	$(^{+4.1}_{-3.4})$	$8.2^{+2.8}_{-2.5}$	$(^{+5.1}_{-4.2})$	$9.8^{+2.9}_{-3.4}$	$(^{+6.1}_{-5.1})$
θ_γ ($^\circ$)	55^{+8}_{-12}	$(^{+13}_{-19})$	51^{+10}_{-14}	$(^{+16}_{-20})$	52^{+9}_{-12}	$(^{+15}_{-21})$	49^{+7}_{-9}	$(^{+15}_{-17})$
a ($''$)	$-$	$-$	$5.8^{+1.1}_{-1.1}$	$(^{+1.9}_{-1.9})$	$4.0^{+1.0}_{-0.9}$	$(^{+1.7}_{-1.5})$	$1.4^{+0.5}_{-0.5}$	$(^{+0.8}_{-0.8})$
σ ($\times 10^3$)	$2.0^{+1.5}_{-1.1}$	$(^{+2.5}_{-1.5})$	$1.9^{+1.1}_{-1.0}$	$(^{+1.7}_{-1.3})$	$1.8^{+1.4}_{-0.9}$	$(^{+2.0}_{-1.3})$	$1.9^{+1.0}_{-1.2}$	$(^{+1.6}_{-1.5})$
θ_σ ($^\circ$)	235^{+39}_{-95}	$(^{+80}_{-174})$	211^{+39}_{-119}	$(^{+81}_{-162})$	228^{+36}_{-96}	$(^{+82}_{-195})$	229^{+26}_{-70}	$(^{+62}_{-146})$
δ ($\times 10^3$)	$2.8^{+1.4}_{-1.4}$	$(^{+2.4}_{-1.9})$	$2.5^{+1.1}_{-0.9}$	$(^{+1.8}_{-1.5})$	$2.4^{+1.1}_{-1.2}$	$(^{+1.9}_{-1.9})$	$3.0^{+1.2}_{-1.3}$	$(^{+2.8}_{-2.0})$
θ_δ ($^\circ$)	68^{+10}_{-16}	$(^{+18}_{-28})$	61^{+11}_{-12}	$(^{+17}_{-21})$	60^{+16}_{-14}	$(^{+23}_{-23})$	67^{+10}_{-14}	$(^{+19}_{-23})$
f_B/f_A	$0.61^{+0.08}_{-0.06}$	$(^{+0.18}_{-0.09})$	$0.53^{+0.05}_{-0.05}$	$(^{+0.09}_{-0.07})$	$0.53^{+0.05}_{-0.06}$	$(^{+0.08}_{-0.07})$	$0.58^{+0.07}_{-0.06}$	$(^{+0.11}_{-0.09})$
$\kappa_{s,30''}$	$0.17^{+0.03}_{-0.05}$	$(^{+0.05}_{-0.09})$	$0.13^{+0.03}_{-0.02}$	$(^{+0.04}_{-0.04})$	$0.16^{+0.02}_{-0.03}$	$(^{+0.04}_{-0.05})$	$0.23^{+0.01}_{-0.02}$	$(^{+0.02}_{-0.02})$
H_0	$82.9^{+6.9}_{-7.3}$	$(^{+12.2}_{-12.6})$	$73.7^{+7.6}_{-7.6}$	$(^{+12.4}_{-11.9})$	$78.4^{+6.6}_{-7.6}$	$(^{+10.2}_{-12.6})$	$82.2^{+5.7}_{-6.2}$	$(^{+8.7}_{-10.4})$
$\min(\chi^2_{reduced})$	0.57		0.99		0.98		0.83	
P_{rel}	0.010		0.814		1.000		0.053	

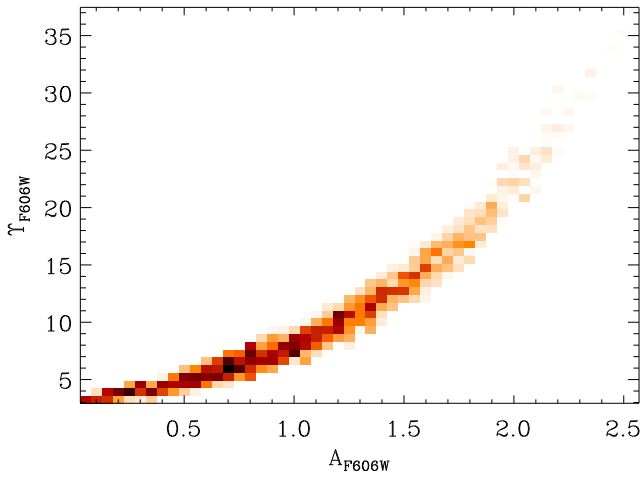

Figure 2.17 2D histogram showing the joint probability distribution for the stellar mass-to-light ratio, $\Upsilon_{\rm F606W}$, and the amount of extinction in the *F606W* band, from FSPS models. As the amount of extinction increases, the model magnitude in the F606W filter increases, leading to a larger values of $\Upsilon_{\rm F606W}$. Results are shown for models with $h = 0.7$ and Solar metallicity, but the distributions for other values of h and metallicity exhibit a similar behavior.

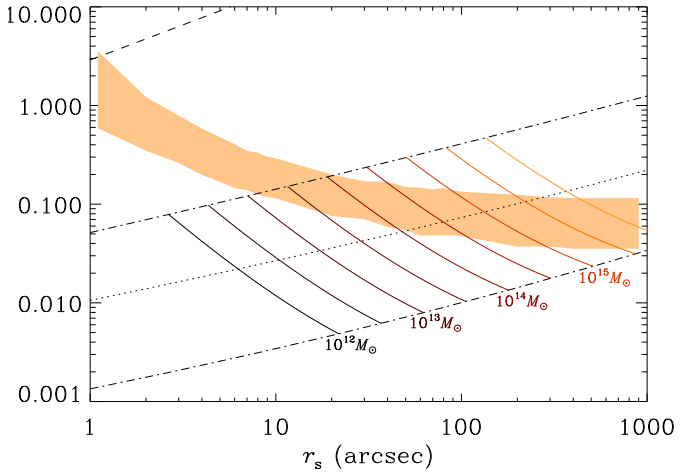

Figure 2.18 The space of scale radius, r_s, and halo normalization, κ_s, for NFW halos at the lens redshift $z = 0.361$. The shaded region indicates the parameter range occupied by our lens models with an NFW dark matter halo. The dotted curve shows the expected relation between κ_s and r_s based on the median concentration/mass relation found in N-body simulations by Macciò et al. (2007), while the dot-dash curves show the $\pm 3\sigma$ range, and the dashed curve in the upper left corner lies 6σ above the median. The colored solid curves represent the theoretical predictions at fixed virial mass, with the concentration varying $\pm 3\sigma$ around the median, ranging from $10^{12}\,M_\odot$ (black) to $10^{16}\,M_\odot$ (orange) in steps of 0.5 dex.

Parameterizing NFW halos with the virial mass M_{vir} and concentration c_v, Macciò et al. (2007) find that the parameters are related by $c_v(z) = 213^{+40}_{-34}\, M_{vir}^{-0.109\pm0.005}\,(1+z)^{-1}$. We can express our lens model parameters in terms of M_{vir} and c_v as follows:

$$\begin{aligned}\kappa_s &= \left(\frac{1+z}{c}\right)^2 \left[\frac{1}{2}\,(GM_{vir})^{1/2}\,H_0^2\Omega_m\Delta_{vir}\right]^{2/3} \\ &\quad \times \frac{D_{ol}D_{ls}}{D_{os}} c_v^2 \left[\ln(1+c_v) - \frac{c_v}{1+c_v}\right]^{-1} \qquad (2.26)\\ r_s &= \frac{1}{c_v(1+z)}\left[\frac{2GM_{vir}}{H_0^2\Omega_m\Delta_{vir}}\right]^{1/3} \qquad (2.27)\end{aligned}$$

where $\Delta_{vir} = 98$ is the virial overdensity (Mainini et al. 2003; Macciò et al. 2007). Figure 2.18 compares our recovered model values of r_s and κ_s with expectations based on the Macciò et al. (2007) relation for different values of the NFW halo mass. The first point is that the model parameters do indeed have reasonable values. Going into more detail, we see that our models with small scale radii ($r_s < 10''$) are consistent with relatively low cluster masses ($M_{vir} \lesssim 10^{13.5}\,M_\odot$) but concentrations that are 3–6σ above the median for that mass. As the scale radius increases, our models follow a track corresponding to increasing mass and decreasing concentration; indeed, lens models with $r_s \gtrsim 200''$ require an extraodinarily large mass of $M_{vir} \gtrsim 10^{15}\,M_\odot$. Such a large mass seems unreasonable for a fairly modest galaxy cluster, especially considering that X-ray measurements imply a mass within 1 h_{75}^{-1} Mpc of $9.9^{+1.9}_{-3.8}\times10^{13}\,M_\odot$ (Chartas et al. 2002). Therefore, we argue that our models with $r_s \gtrsim 200''$ are disfavored but models with $r_s \lesssim 200''$ have parameters that seem reasonable in comparison with simulated NFW halos. We note that recovering a concentration that is a few sigma above the median is not necessarily a concern, because there may be a selection bias such that a high concentration increases the lensing cross section (see, e.g., Mandelbaum et al. 2009).

2.5 Discussion and Conclusions

Since its discovery in 1979, Q0957 has presented many puzzles that we still cannot definitively solve. Nevertheless, by combining new HST/ACS data and stellar population synthesis models, we have presented a consistent picture of lensing in this system using a realistic treatment of both the stellar and dark matter components of the mass distribution. In §2.5.1 we discuss our results regarding measurements of H_0 with Q0957. Turning the tables, in §2.5.2 we adopt priors on H_0 from other measurements and examine the inferred properties of the lens mass distribution. Looking ahead, in §2.5.3 we discuss potential ways in which these measurements can be improved.

2.5.1 Hubble Constant

Motivated by our new ACS data, we conducted a joint strong+weak lensing analysis in the hope of obtaining the best constraints to date from Q0957. In §2.4.1 we found $H_0 = 85^{+14}_{-13}$ km s^{-1} Mpc^{-1} (68% CL) on the basis of lensing alone. This result is higher than, but still consistent with, measurements from other recent lensing (e.g., Jakobsson et al. 2005; Paraficz et al. 2009; Oguri 2007) and non-lensing (e.g., Freedman et al. 2001; Riess et al. 2009; Dunkley et al. 2008; Komatsu et al. 2008) studies. In spite of the extensive lensing data we have obtained, the uncertainty in H_0 from Q0957 is still larger than from most other lenses (see Fig. 2.1).

One source of uncertainty in Q0957 is the sheer complexity of the potential: with ellipticity, shear, and higher-order environmental terms to play with, models can find a wide range of combinations that fit the data well (see Figs. 2.5–2.8). The main systematic effect in lens models is a correlation between h and the stellar mass-to-light ratio of the lens galaxy. Varying $\Upsilon_{\rm F606W}$ changes the balance between stars and dark matter in the lens galaxy, which modifies the net density profile, which then affects h through the radial profile degeneracy (e.g., Kochanek 2002). We can actually turn this degeneracy to our advantage if we can place independent constraints on the stellar

mass-to-light ratio. In §2.4.3 we used the stellar population synthesis models from Maraston et al. (2009) to constrain $\Upsilon_{\rm F606W}$ and thereby reduce the uncertainties in our Hubble constant determination to $H_0 = 79.3^{+6.7}_{-8.5}$ km s^{-1} Mpc^{-1} (68% CL).

While this is a significant reduction in the uncertainty for H_0 from Q0957, we caution that SPS models are still improving and may ultimately be even more complicated than the Maraston models. When we used the FSPS models of Conroy et al. (2009), for example, we did not see much tightening of the H_0 constraints relative to lensing alone; and we traced the trouble to uncertainties in the amount of dust extinction in the lens galaxy. There is one additional source of uncertainty in SPS models that we did not explicitly address, namely the IMF. Variations in the IMF can alter the colors and mass-to-light ratios of SPS models (e.g., Conroy et al. 2009). In particular, a more bottom-heavy IMF (e.g., Salpeter) would raise the inferred value of the $\Upsilon_{\rm F606W}$ and hence our median value of H_0, while a more top-heavy IMF would have the opposite effect. For the Maraston SPS models, variations in the IMF must be relatively small or the models would cease to provide a good fit to the SDSS LRG sample. We attempt to compensate for IMF-related variations by allowing broad ranges for the star formation history, star formation epoch, and dust. Nevertheless, this remains an unknown, but presumably small, uncertainty in our models. Clearly there is a lot of room for improvement with a better understanding of the stellar population of the lens galaxy (see §2.5.3).

2.5.2 Mass Distribution

Instead of trying to measure H_0 ourselves, we can choose to place external priors on H_0 to see how well we can understand the mass distribution of the lens. We consider two determinations of H_0: the refurbished distance ladder measurement of $H_0 = 74.2 \pm 3.6$ km s^{-1} Mpc^{-1} by Riess et al. (2009), and the combined WMAP5+SNe+BAO value of $H_0 = 70.5 \pm 1.3$ from Komatsu et al. (2008). In Table 2.7 we show the model parameters recovered from this approach.

Table 2.7 Model Results: HST-ACS data + H_0 priors

	Median model values with 68% CL (95% CL) uncertainties							
Parameter	NFW		$\alpha = 0.5$		$\alpha = 1.0$		$\alpha = 1.5$	
$\Upsilon_{\rm F606W}$	$4.3^{+0.7}_{-1.2}$	$3.6^{+0.2}_{-0.6}$	$5.5^{+0.5}_{-0.5}$	$5.1^{+0.2}_{-0.3}$	$5.6^{+1.3}_{-0.6}$	$5.0^{+0.2}_{-0.3}$	$5.8^{+0.5}_{-0.5}$	$5.7^{+0.3}_{-0.4}$
b_d ($''$)	$0.135^{+0.028}_{-0.018}$	$0.144^{+0.037}_{-0.011}$	$6.3^{+1.3}_{-1.0}$	$6.7^{+1.2}_{-1.1}$	$5.5^{+1.1}_{-1.1}$	$5.6^{+1.2}_{-1.5}$	$4.1^{+0.5}_{-0.5}$	$4.1^{+0.5}_{-0.5}$
e	$0.25^{+0.06}_{-0.07}$	$0.19^{+0.05}_{-0.04}$	$0.31^{+0.07}_{-0.10}$	$0.26^{+0.07}_{-0.09}$	$0.28^{+0.08}_{-0.07}$	$0.21^{+0.08}_{-0.06}$	$0.25^{+0.06}_{-0.07}$	$25^{+0.07}_{-0.06}$
θ_e ($^\circ$)	69^{+12}_{-9}	69^{+11}_{-11}	74^{+9}_{-10}	74^{+9}_{-10}	72^{+9}_{-9}	72^{+8}_{-8}	77^{+9}_{-10}	77^{+7}_{-8}
γ ($\times 10^2$)	$9.6^{+3.0}_{-2.6}$	$10.0^{+2.0}_{-2.0}$	$7.6^{+2.7}_{-1.9}$	$7.8^{+2.4}_{-2.0}$	$8.2^{+2.5}_{-2.6}$	$8.6^{+2.2}_{-2.6}$	$11.0^{+2.3}_{-3.0}$	$11.0^{+3.0}_{-2.2}$
θ_γ ($^\circ$)	57^{+6}_{-11}	58^{+6}_{-8}	50^{+11}_{-14}	52^{+8}_{-14}	52^{+10}_{-12}	56^{+5}_{-10}	51^{+16}_{-8}	51^{+10}_{-7}
a ($''$)	–	–	$5.8^{+1.1}_{-1.0}$	$5.9^{+1.1}_{-1.0}$	$4.2^{+1.0}_{-0.7}$	$4.1^{+0.8}_{-0.6}$	$1.3^{+0.6}_{-0.7}$	$1.3^{+0.5}_{-0.5}$
σ ($\times 10^3$)	$2.5^{+1.6}_{-1.1}$	$3.3^{+1.2}_{-1.4}$	$1.9^{+1.1}_{-1.0}$	$2.1^{+0.9}_{-1.1}$	$1.9^{+1.4}_{-1.1}$	$2.0^{+1.5}_{-1.0}$	$2.2^{+0.9}_{-0.9}$	$2.2^{+0.9}_{-0.9}$
θ_σ ($^\circ$)	213^{+23}_{-42}	211^{+10}_{-36}	215^{+45}_{-127}	215^{+30}_{-114}	221^{+24}_{-74}	220^{+21}_{-26}	180^{+47}_{-66}	185^{+39}_{-60}
δ ($\times 10^3$)	$2.8^{+1.1}_{-1.5}$	$2.4^{+1.4}_{-1.2}$	$2.3^{+1.3}_{-0.7}$	$2.5^{+1.1}_{-1.6}$	$2.3^{+1.1}_{-1.5}$	$2.3^{+1.0}_{-1.3}$	$3.8^{+2.0}_{-1.3}$	$3.8^{+1.1}_{-1.0}$
θ_δ ($^\circ$)	68^{+10}_{-16}	70^{+19}_{-28}	61^{+11}_{-12}	62^{+11}_{-16}	63^{+15}_{-16}	63^{+17}_{-16}	69^{+4}_{-10}	68^{+4}_{-8}
f_B/f_A	$0.66^{+0.09}_{-0.06}$	$73^{+0.22}_{-0.07}$	$0.52^{+0.05}_{-0.04}$	$0.53^{+0.05}_{-0.04}$	$0.53^{+0.05}_{-0.04}$	$0.55^{+0.05}_{-0.03}$	$0.61^{+0.04}_{-0.05}$	$0.60^{+0.05}_{-0.04}$
$\kappa_{s,30''}$	$0.17^{+0.04}_{-0.05}$	$0.17^{+0.03}_{-0.04}$	$0.13^{+0.02}_{-0.02}$	$0.13^{+0.02}_{-0.02}$	$0.17^{+0.02}_{-0.02}$	$0.17^{+0.02}_{-0.02}$	$0.24^{+0.01}_{-0.01}$	$0.24^{+0.01}_{-0.01}$
P_{rel}	0.007	0.003	1.000	1.000	0.869	0.565	0.006	0.000

Examining the relative probabilities of the models, it is clear that the lensing data favor a power law profile of index $\alpha = 0.5$ or 1.0 over $\alpha = 1.5$ or an NFW profile. Both of the favored models exhibit relatively large core radii: $a = 5.8^{+1.1}_{-1.0}$ arcsec for $\alpha = 0.5$, or $a = 4.2^{+1.0}_{-0.7}$ arcsec for $\alpha = 1$. Given the reduced probability of our NFW models, we conclude that lensing provides strong evidence for a constant-density core (rather than a cusp) in the dark matter halo of Q0957.

In Figure 2.10 we showed that Q0957 exhibits a rising deflection profile, analogous to a rising rotation curve and indicative of a net density profile shallower than isothermal. While this is not the first case of a lens with a rising rotation curve (see Kochanek et al. 2006), the origin of phenomenon is unclear. One possible explanation involves the special location of the lens galaxy. As the central galaxy in a modest cluster, the lens is embedded in the most concentrated part of a massive dark matter halo. The higher than average surface density of dark matter leads to a shallow density profile and a rising deflection curve. To further explore this idea, we shown in Figure 2.19 the fraction of the deflection contributed by dark matter as a function of radius. This is equivalent to the projected enclosed dark matter fraction as a function of radius. At the effective radius, dark matter constitutes $(50 \pm 7)\%$ or $(57 \pm 7)\%$ of the enclosed mass (assuming the Riess or Komatsu priors on H_0, respectively). Such values are well above the minimum dark matter fraction found $(38 \pm 7)\%$ for galaxies in the SLACS survey (Bolton et al. 2008), and are greater than 16 of the 22 lens systems considered by (Jiang & Kochanek 2007). It is not very surprising, of course, to find a relatively large dark matter fraction in a brightest cluster galaxy.

We find the the dark matter halo in Q0957 must be well aligned with stellar mass distribution. With either the Riess or Komatsu H_0 priors, we find the position angle of the dark matter halo to be $\theta_e = 73^{+9}_{-10}$ deg, in good agreement with the measured position angle of the stellar component $\theta_\star = 71 \pm 5$ deg at large radii ($> 5''$). Perhaps more interesting is that the ellipticity of the dark matter appears to be in good agreement

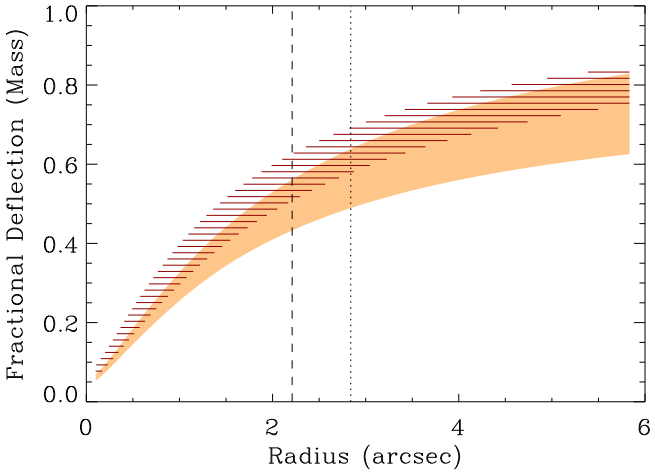

Figure 2.19 Fraction of the monopole deflection contributed by the dark matter halo, as a function of radius. Since $\alpha(r) \propto M(r)/r$, this plots also depicts the enclosed projected dark matter fraction as a function of radius. The vertical dotted line indicates the Einstein radius at $R_{ein} = 2.84''$ while the vertical dashed line marks the effective radius of the stellar light profile at $R_e = 2.21''$. The shaded regions show results from our models when we adopt different priors on the Hubble constant: the orange region corresponds to assuming $H_0 = 74.2 \pm 3.6$ km s^{-1} Mpc^{-1} from Riess et al. (2009), while the region indicated with red horizontal lines corresponds to $H_0 = 70.5 \pm 1.3$ km s^{-1} Mpc^{-1} from Komatsu et al. (2008).

with that of the stellar distribution. We find the ellipticity of the dark matter to be $e = 0.28^{+0.06}_{-0.07}$ or $e = 0.25^{+0.07}_{-0.09}$ for the Riess or Komatsu H_0 values, in good agreement with the measured value of the stellar ellipticity $e_\star \sim 0.3$ at large radii.

In general, we find that improving H_0 constraints from 5% to 2% has little impact on our understanding of the mass distribution, because most of our model parameters have little or no correlation with H_0. The only significant exception is in our determination of the stellar mass to light ratio, $\Upsilon_{\rm F606W}$. We find $\Upsilon_{\rm F606W} = 5.5^{+0.9}_{-0.5}$ using the Riess value for H_0, versus $\Upsilon_{\rm F606W} = 5.5^{+0.2}_{-0.3}$ using the Komatsu value. For comparison, the Maraston SPS models give $\Upsilon_{\rm F606W} = 5.9 \pm 1.9$. It is interesting to see that combining lensing with H_0 priors can provide excellent constraints on stellar populations, which may prove useful as SPS laboratories as multi-wavelength datasets for well-studied lenses grow.

We found the intriguing result that our models favor a quasar flux ratio around $F_B/F_A = 0.53 \pm 0.06$, which is substantially different from the radio measurement of $F_B/F_A = 0.74 \pm 0.02$. Since the radio emission is free from extinction by dust, and should be insensitive to microlensing by stars, we infer the quasar flux ratio in Q0957 seems to be "anomalous." The putative anomaly presumably indicates additional complexity in the lens potential. While the nature of that complexity is not yet clear, it is worth noting that dark matter substructure can easily produce flux ratio anomalies (e.g., Mao & Schneider 1998; Metcalf & Madau 2001; Dalal & Kochanek 2002; Chiba et al. 2005; MacLeod et al. 2009). Invoking substructure as an explanation seems especially enticing because the radio jets, a mere 80 mas away from the quasar images, have a flux ratio of 0.61 ± 0.04 (Bonometti 1985), in much better agreement with the macro models. While we cannot definitively identify substructure from the present analysis, the evidence is fascinating and warrants future study.

2.5.3 Future Prospects

Looking ahead, there are several ways in which we can hope to improve the measurement of H_0 in Q0957. Following this work, the best improvements are likely to come from a better understanding of the stellar population and $\Upsilon_{\rm F606W}$. In the near term, an extension of the photometric data to both bluer and redder wavelengths could reduce the uncertainties in existing SPS models and, therefore, the uncertainties in H_0. Over the longer term, extensive photometry spanning UV/optical/infrared wavelengths, coupled with a better understanding of SPS uncertainties (e.g., TP-AGB stars, blue stragglers, IMF), should boost the reliability and reduce the uncertainties in the SPS technique.

Improved lensing constraints may also help. One source of systematic uncertainty in our models is the uncertainty in the measurement of the total convergence from weak lensing. Currently, the total convergence is measured to $\sim 34\%$ precision. Deeper imaging would be observationally expensive but worthwhile, especially if coupled with

improved understanding of the point spread function and source redshift distribution for the weak lensing analysis. For strong lensing, we have noted that there are additional faint lensed features that we have not used, but that might further constrain the lensing critical curve (see the discussion accompanying Figs. 2.11 and 2.13). We note for the record that using the precise position and flux ratio constraints for the quasars is not likely to help us understand the (large-scale) mass distribution or H_0 any better; those data will ultimately be most useful for probing small-scale structure in the lens.

Finally, it is interesting to consider whether stellar dynamics data could further constrain our models. Previous studies have shown that joint lensing+dynamics analyses can provide more information than lensing alone about the mass distribution (e.g., Treu et al. 2006; Barnabè & Koopmans 2007). The joint approach has been used to good effect by the SLACS team to improve constraints on quantities like the total mass to light ratio and the slope of the inner density profile (Koopmans et al. 2006). For Q0957, Romanowsky & Kochanek (1999) have successfully combined information from stellar dynamics and lensing to tighten constraints on mass models. Adopting the measured central velocity dispersion from Tonry & Franx (1997), Romanowsky & Kochanek use orbit modeling techniques to constrain a spherical power law profile, measuring $H_0 = 61^{+13}_{-15}$ km s^{-1} Mpc^{-1} (2σ). While it is clear such an analysis would be useful to further constrain our models, it would require orbit modeling for each of our model classes, which is beyond the scope of this work. Furthermore, it would require a deprojection of the stellar component that we include in our models, which could be challenging (because of the varying ellipticity and orientation; cf. Figure 2.3) and would have uncertainties of its own.

Bibliography

Anderson, J. & King, I. R., 2002, Instrument Science Report ACS 2006-01, http://www.stsci.edu/hst/acs/documents/isrs/isr0601.ps

Angulo, R. E., Lacey, C. G., Baugh, C. M., & Frenk, C. S., 2009, MNRAS, 399, 983

Auger, M. W., Fassnacht, C. D., Abrahamse, A. L., Lubin, L. M., & Squires, G. K., 2007, AJ, 134, 668

Barkana, R., Lehár, J., Falco, E. E., Grogin, N. A., Keeton, C. R., & Shapiro, I. I., 1999, ApJ, 520, 479

Barnabè, M. & Koopmans, L. V. E., 2007, ApJ, 666, 726

Bartelmann, M., 1996, A&A, 313, 697

Bernstein, G., Fischer, P., Tyson, J. A., & Rhee, G., 1997, ApJ, 483, 79

Bernstein, G. & Fischer, P., 1999, ApJ, 118, 14

Blandford, R. D., & Narayan, R., 1986, ApJ, 310, 568

Dark Matter in the Universe, Proceedings of the 4th Jerusalem Winter School for Theoretical Physics, Jerusalem. World Scientific Publication, Singapore, 1987, p.133

Bolton, A. S., Treu, T., Koopmans, L. V. E., Gavazzi, R., Moustakas, L. A., Burles, S., Schlegel, D. J., & Wayth, R., 2008, ApJ, 684, 248

Bonometti, R. J., 1985, Ph.D. thesis, Massachusetts Institute of Technology

Burud, I., Hjorth, J., Jaunsen, A. O., Anderson, M. I., Korhonen, H., Clasen, J. W., Pelt, J., Pijpers, F. P., Magain, P., & Østensen, 2000, ApJ, 544, 117

Burud, I., Hjorth, J., Courbin, F., Cohen, J. G., Magain, P., Jaunsen, A. O., Kaas, A. A., Faure, C., & Letawe, G., 2002, A&A, 391,481

Burud, I., Courbin, F., Magain, P., Lidman, C., Hutsemékers, D., Kneib, J. P., Hjorth, J., Brewer, J., Pompei, E., Germany, L., Pritchard, J., Jaunsen, A. O., Letawe, G., & Meylan, G., 2002, A&A, 383, 71

Chabrier, G., 2003, PASP, 115, 763

Chartas, G., Gupta, V., Garmire, G., Jones, C., Falco, E. E., Shapiro, I. I., & Tavecchio, F., 2002, ApJ, 565, 96

Chartas, G., Dai, X., & Garmire, G. P., 2004, Carnegie Observatories Astrophysics Series, Volume 2, *Measuring and Modeling the Universe*, http://www.ociw.edu/ociw/symposia/series/symposium2/proceedings.html

Charlot, S. & Fall, S. M., ApJ, 539, 718

Chen, J., Rozo, E., Delal, N., Taylor, J. E., 2007, ApJ, 659, 52

Chiba, M., Minezaki, T., Nashikawa, N., Kataza, H., Inoue, K. T., 2005, ApJ, 627, 61

Cohn, J. D., Kochanek, C. S., McLeod, B. A., & Keeton, C. R., 2001, ApJ, 554, 1216

Coles, J., 2008, ApJ, 679, 17

Colley, W., et al., 2003, ApJ, 587, 71

Conner, S. R., Lehár, J., & Burke, B. F., 1992, ApJ, 387, L61

Conroy, C., Gunn, J. E., & White, M., 2009, ApJ, 699, 486

Cooray, A., & Sheth, R., 2002, Phys. Res., 372, 1

Dalal, N., & Kochanek, C. S., 2002, ApJ, 572, 25

Davis, T. M., et al., 2007, ApJ, 666, 715

Dobler, G. & Keeton, C. R., MNRAS, 365,1243

Dunkley, J., et al., 2008, ApJS, 180, 306

Eisenstein, D. J., 2001, AJ, 122, 2267

Elíasdóttir, Á., Hjorth, J., Toft, S., Burud, I., & Paraficz, D., ApJS, 166, 443

Falco, E. E., Gorenstein, M. V., & Shapiro, I. I., 1985, ApJ, 289, L1

Falco, E. E., Impey, C. D., Kochanek, C. S., Lehr, J., McLeod, B. A., Rix, H.-W., Keeton, C. R., Muñoz, J. A., & Peng, C. Y., 1999, ApJ, 523, 617

Fassnacht, C. D., & Lubin, L. M., 2002, AJ, 123, 627

Fassnacht, C. D., Xanthopoulos, E., Koopmans, L. V. E., & Rusin, D. 2002, ApJ, 581, 823

Fassnacht, C. D., Gal, R. R., Lubin, L. M., McKean, J. P., Squires, G. K., & Readhead, A. C. S., 2006, ApJ, 642, 30

Fassnacht, C. D., Kocevski, D. D., Auger, M. W., Lubin, L. M., Neureuther, J. L., Jeltema, T. E., Mulchaey, J. S., & McKean, J. P., 2008, ApJ, 681, 1017

Fischer, P., Bernstein, G., Rhee, G., & Tyson, J.A., 1997, AJ, 113, 521

Freedman, W. L., et al., 2001, ApJ, 553, 47

Fruchter, A. S., & Hook, R. N., 2002, PASP, 114, 144

Garrett, M. A., Calder, R. J., Porcas, R. W., King, L. J., Walsh, D., & Wilkinson, P. N., 1994, MNRAS, 270, 457

Gavazzi, R., Treu, T., Rhodes, J. D., Koopmans, L, V. E., Bolton, A. S., Burles, S., Massey, R. J., & Moustakas, L. A., 2007, ApJ, 667, 176

Gavazzi, R., Treu, T., Koopmans, L. V. E., Bolton, A. S., Moustakas, L. A., Burles, S., & Marshall, P. J., 2008, ApJ, 677, 1046

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B., 1995, Bayesian Data Analysis (New York: Chapman &Hall/CRC)

Gerhad, O., Kronawitter, A., Saglia, R. B., & Bender, R., 2001, AJ, 121, 1936

Goicoechea, Gil-Merino, R., & Ullán, A., 2005, MNRAS, 360, L60

Goicoechea, L. J., Shalyapin, V. N., Koptelova, E., Gil-Merino, R., Zheleznyak, A. P., & Ullán, A., 2008, New Astronomy, 13, 182

Gorenstein, M. V., Cohen, N. L., Shapiro, I. I., Rogers, A. E. E., Bonometti, R. J., Falco, E. E., Bartel, N., & Marcaide, J. M., 1988b, ApJ, 334, 42

Gorenstein, M. V., Shapiro, I. I., & Falco, E. E., 1988a, ApJ, 327, 693

Grillo, C., Gobat, R., Lombardi, M., & Rosati, P., 2009, A&A, 501, 461

Grogin, N. A., & Narayan, R., 1996, ApJ, 464, 92; erratum, 1996, ApJ, 473, 570

Haario, H., Saksman, E., & Tamminen, J., 2001, Bernoulli, 7, 223

Haarsma, D. B., Hewitt, J. N., Lehar, J., & Burke, B. F., 1999, ApJ, 510, 64

Hjorth, J., Burud, I., Jaunsen, A. O., Schechter, P. L., Kneib, J. P., Andersen, M. I., Korhonen, H., Clasen, J. W., Kaas, A. A., Østensen, R., Pelt, J., & Pijpers, F. P., 2002, ApJ, 572, L11

Jakobsson, P., Hjorth, J., Burud, I., Letawe, G., Lidman, C., & Courbin, F., 2005, A&A, 431, 103

Jiang, G. & Kochanek, C. S., 2007, ApJ, 671, 1568

Keeton, C. R. & Kochaneck C. S., 1997, ApJ, 487, 42

Keeton, C. R., Kochaneck C. S., & Seljak, U., 1997, ApJ, 482, 604

Keeton, C. R. & Kochaneck C. S., 1998, ApJ, 495, 157

Keeton, C. R., Falco, E. E., Impey, C. D., Kochanek, C. S., Lehr, J., McLeod, B. A., Rix, H.-W., Muñoz, J. A., & Peng, C. Y., 2000, ApJ, 542, 74

Keeton, C. R., 2001, preprint arXiv:astro-ph/0102340

Keeton, C. R., & Zabludoff, A. I., 2004, ApJ, 612, 600

Kelson, D. D., Zabludoff, A. I., Williams, K. A., Trager, S. C., Mulchaey, J. S., & Bolte, M., 2002, ApJ, 576, 720

Kochanek, C. S., Blandford, R. D., Lawrence, C. R., & Narayan, R., 1989, MNRAS, 238, 43

Kochanek, C. S., 1991, ApJ 382, 58

Kochanek, C. S., 1995, ApJ, 445, 559

Kochanek, C. S., Keeton, C. R., & McLeod, B. A., 2001, ApJ, 547, 50

Kochanek, C. S., 2002, ApJ, 578, 25

Kochanek, C. S., 2004, preprint arXiv:astro-ph/0407232

Kochanek, C. S., Morgan, N. D., Falco, E. E., McLeod, B. A., Winn, J. N., Dembicky, J., & Ketzeback, B., 2006, ApJ, 640,47

Komatsu, E., et al. 2008, ApJS, 180, 330

Koopmans, L. V. E., de Bruyn, A. G., Xanthopoulos, E., & Fassnacht, C. D., 2000, A&A, 356, 391

Koopmans, L. V. E., Treu, T., Fassnacht, C. D., Blandford, R. D., & Surpi, G., 2003, ApJ, 599, 70

Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S., & Moustakas, L. A., 2006, ApJ, 649, 599

Kundic, T., et al., 1997, ApJ, 482, 75

Lagattuta, D. J., Fassnacht, C. D., Auger, M. W., Bradac, M., Marshall, P. J., Treu, T., Gavazzi, R., & Schrabback, T., 2009, BAAS, 41, 445

Lartillot, N & Phillipe, H., 2006, Systematic Biology, 55, 195

Lehár, J., Cooke, A. J., Lawrence, C. R., Silber, A. D., & Langston, G. I., 1996, AJ, 111, 1812

Lehár, J., Burke, B. F., Conner, S. R., Falco, E. E., Fletcher, A. B., Irwin, M., McMahon, R. G., Muslow, T. W. B., & Schechter, P. L., 1997, AJ, 114, 48

Liddle, A. R., 2007, MNRAS, 377, L74

Lidman, C., Courbin, F., Meylan, G., Broadhurst, T., Frye, B., & Welch, W. J. W., 1999, ApJ, 514, L57

Lovell, J. E. J., Jauncey, D. L., Reynolds, J. E., Wieringa, M. H., King, E. A., 1998, ApJ, 508, L51

Macciò, A. V., Dutton, A. A., van den Bosch, F., Moore, B., Potter, D., & Stadel, J., MNRAS, 378, 55

MacLeod, C. L., Kochanek, C. S., & Agol, E., 2009, ApJ, 699, 1578

Macri, L. M., Stanek, K. Z., Bersier, D., Greenhill, L. J., & Reid, M. J., 2006, ApJ, 652, 1133

Mainini, R, Macciò, A. V., Bonometto, S. A., & Klypin, A., 2003, ApJ, 599, 24

Marshall, P. J., Hobson, M. P., & Slosar, A., 2003, MNRAS, 346, 489

Mandelbaum, R., van de Ven, G., & Keeton, C. R., 2009, MNRAS, 398, 635

Mao, S. & Schneider, P., 1998, MNRAS, 295, 587

Maraston, C., 2005, MNRAS, 362, 799

Maraston, C., Strömbäck G., Thomas, D. A., Wake, D. A., & Nichol, R. C., 2009, MNRAS, 394, L104

Metcalf, R. B. & Madau, P., 2001, ApJ, 563, 9

Myers, S. T., et al., 1995, ApJ, 447, L5

Momcheva, I., Williams, K., Keeton, C., & Zabludoff, A., 2006, ApJ, 641, 169

Nakajima, R., Bernstein, G., Fadely, R., Keeton, C.R., & Schrabback, T., 2009, ApJ, 697, 1793

Navarro, J. F., Frenk, C. S., & White, S. D. M., 1997, ApJ, 462,563

Oguri, M., 2007, ApJ, 660, 1

Oscoz, A., Alcalde, D., Serra-Ricart, M., Mediavilla, E., Abajas, C., Barrena, R., Licandro, J., Motta, V., & Muoz, J. A., 2001, ApJ, 552, 81

Paraficz, D., Hjorth, J., & Elíasdóttir, Á., 2009, A&A, 499, 395

Patnaik, A. R., Narasimha, D., 2001, MNRAS, 326, 1403

Poindexter, S., Morgan, N., Kochanek, C. S., & Falco, E. E., 2007, ApJ, 660, 146

Rapetti, D., Allen, S. W., Amin, M. A., & Blandford, R. D., 2007, MNRAS, 375, 1510

Reiss, A. G., Macri, L., Casertano, S., Sosey, M., Lampeitl, H., Ferguson, H. C., Filippenko, A. V., Jha, S. W., Li, W., Chornock, R., & Sarkar, D., 2009, ApJ, 699, 539

Roberts, G. O., Gelman, A., & Gilks, W. R., 1997, Annals of Applied Probability, 7, 110

Roberts, G. O. & Rosenthal, J. S., 2001, Statistical Science, 16, 351

Romanowsky, A. & Kochanek, C. S., 1999, ApJ, 516 , 18

Rusin, D., Kochanek, C. S., & Keeton, C. R., 2003, ApJ, 595, 29

Rusin, D., & Kochanek, C. S., 2005, ApJ, 623, 666

Saha, P., Coles, J., Macciò, A. V., & Williams, L. L. R., 2006, ApJ, 650, L17

Saha, P., & Williams, L. L. R., 2006, ApJ, 653, 936

Schawinski, K., Thomas, D., Sarzi, M., Maraston, C., Kaviraj, S., Joo, S.-J., Yi, S. K., & Silk, J., MNRAS, 2007, 382, 1415

Schild, R. E. & Cholfin, B., 1986, ApJ, 300, 209

Schlegel, D. J., Finkbeiner, D. P., & Davis, M., 1998, ApJ, 500, 525

Schneider, P., Ehlers, J., & Falco, E. E., 1992, Gravitational Lenses (New York: Springer)

Schramm, T., 1990, A&A, 231, 19

Schwarz, G., 1978, Ann. Stat., 6, 461

Springel, V., et al., 2005, Nature, 435, 629

Suyu, S. H., Marshall, P. J., Blandford, R. D., Fassnacht, C. D., Koopmans, L. V. E., McKean, J. P., & Treu, T., 2008, ApJ, 691, 277

Suyu, S. H., Marshall, P. J., Auger, M. W., Hilbert, S., Blandford, R. D., Koopmans, L. V. E., Fassnacht, C. D., & Treu, T., 2009, preprint arXiv:0910.2773

Tonry, J.L. & Franx, M., 1999, ApJ, 515, 512

Treu, T. & Koopmans, L. V. E., 2002, MNRAS, 337, L6

Treu, T., Koopmans, L. V., Bolton, A. S., Burles, S., & Moustakas, L. A., 2006, ApJ, 640, 662

Trotter, C. S., Winn, J. N., & Hewitt, J. N., 2000, ApJ, 535, 671

van de Ven, G., Mandelbaum, R., & Keeton, C. R., 2009, MNRAS, 398, 607

Vuissoz, C., Courbin, F., Sluse, D., Meylan, G., Ibrahimov, M., Asfandiyarov, I., Stoops, E., Eigenbrod, A., Le Guillou, L., van Winckel, & H., Magain, P., 2007, A&A, 464, 845

Wake, D. A., et al., 2006, MNRAS, 372, 537

Walsh, D., Carswell, R. F., & Weymann, R. J., 1979, Nature, 279, 381

Williams, K. A., Momcheva, I., Keeton, C. R., Zabludoff, A. I., & Lehár, J., 2006, ApJ, 646, 85

Williams, L. L. R. & Saha, P., 2000, ApJ, 119, 439

Winn, J. N., Kochanek, C. S., McLeod, B. A., Falco, E. E., Impey, C. D., & Rix, H.-W., 575, 103

Wucknitz, O., Biggs, A. D., & Browne, I. W. A., 2004, MNRAS, 349, 14

York, T., Jackson, N., Browne, I. W. A., Wucknitz, O., & Skelton, J. E., 2005, MNRAS, 357, 124

Chapter 3

Mid-Infrared Spectroscopy of Two Lensed Star-forming Galaxies

This chapter is based on work in the journal article:

Mid-Infrared Spectroscopy of Two Lensed Star-forming Galaxies

Fadely, R., Allam, S. S., Baker, A. J., Lin, H., Lutz, D., Shapley, A. E., Shin, M.-S., Smith, J. A., Stauss, M. A., & Tucker, D. L., 2010, ApJ submitted

We present low-resolution, rest-frame $\sim 5-12\,\mu$m *Spitzer*/IRS spectra of two lensed $z \sim 2$ UV-bright star-forming galaxies, SDSS J120602.09+514229.5 and SDSS J090122.37+181432.3. Using the magnification boost from lensing, we are able to study the physical properties of these objects in greater detail than is possible for unlensed systems. In both targets, we detect strong PAH emission at 6.2, 7.7, and 11.3 μm, indicating the presence of vigorous star formation. For J1206, we find a steeply rising continuum and significant [S IV] emission, suggesting that a moderately hard radiation field is powering continuum emission from small dust grains. The strength of the [S IV] emission also implies a sub-solar metallicity of $\sim 0.5\,Z_{\odot}$, confirming published rest-frame optical measurements. In J0901, the PAH lines have large rest-frame equivalent widths ($> 1\,\mu$m) and the continuum rises slowly with wavelength, suggesting that any AGN contribution to $L_{\rm IR}$ is insignificant, in contrast to the implications of optical emission-line diagnostics. Using [O III] line flux as a proxy for AGN strength, we estimate that the AGN in J0901 provides only a small fraction of its mid-infrared continuum flux. By combining the detection of [Ar II] with an upper limit on [Ar III] emission, we

infer a metallicity of $\gtrsim 1.3\,Z_{\odot}$. This work highlights the importance of combining rest-frame optical and mid-IR spectroscopy in order to understand the detailed properties of star-forming galaxies at high redshift.

3.1 Background

Rest-frame UV selection offers a prime view into populations of star-forming galaxies at $z > 1.5$. First used to identify $z \sim 3$ galaxies with sharp breaks in their spectral energy distributions due to absorption below the Lyman limit (Steidel et al. 1996), the Lyman break technique has now been modified and extended to both lower and higher redshift (e.g., Steidel et al. 1999; Lehnert & Bremer 2003; Adelberger et al. 2004). Since the development of the technique in the 1990s, thousands of star-forming galaxies have been spectroscopically confirmed at $z \sim 1.5 - 3.5$ (e.g., Steidel et al. 2003; Reddy et al. 2006a), resulting in a revolution in our understanding of galaxy formation and evolution. However, these high-redshift galaxies suffer from a fundamental problem: they are typically small and faint, with $R_{\rm AB} \geq 24\,{\rm mag}$, making it impossible to carry out detailed studies of individual objects unless they happen to be strongly lensed. In particular, this limitation applies to observations of UV-bright star-forming systems at the long wavelengths that can trace star formation even in highly obscured regions. While stacking analyses of large samples confirm that UV-selected galaxies at $z \sim 2$ have substantial fractions of their bolometric luminosities emerging in the far-infrared, with $\langle L_{\rm IR}/L_{\rm UV}\rangle \simeq 4 - 5$ (Reddy & Steidel 2004; Reddy et al. 2006b), understanding the parameters of obscured star formation in individual unlensed objects remains out of reach for current facilities.

The first bright lensed Lyman break galaxies, "cB58" (i.e., MS 1512-cB58) and the "Cosmic Eye" (i.e., LBG J213512.73−010143), were discovered serendipitously in the course of a cluster redshift survey (Yee et al. 1996) and a *Hubble Space Telescope* (*HST*) snapshot imaging survey of X-ray-bright clusters (Smail et al. 2007), respectively.

Since clusters are rich in strong caustics capable of producing fold arcs (like cB58) and in individual galaxies capable of producing nearly complete Einstein rings (like the Cosmic Eye), the circumstances of these discoveries were not surprising. More recently, however, several teams have begun to exploit the enormous footprint of the Sloan Digital Sky Survey (SDSS) to identify UV-bright high-redshift sources that are lensed by individual galaxies in field or group environments. This enterprise kicked off with the serendipitous discovery of the "8 O'Clock Arc" (Allam et al. 2007) and has now spawned a variety of systematic searches within the SDSS object catalog that rely on different selection criteria (e.g., Belokurov et al. 2007; Shin et al. 2008; Hennawi et al. 2008). By focusing on luminous red galaxies with multiple blue neighbors, and on close pairs with characteristic lens+arc morphologies, various authors of this paper have now contributed to the discovery of 11 new spectroscopically confirmed lenses at redshifts $0.4 \leq z \leq 2.4$ (Kubo et al. 2009; Lin et al. 2009; Diehl et al. 2009).

In this paper, we present *Spitzer*/IRS spectroscopy of two objects from this new SDSS sample. The first, SDSS J120602.09+514229.5 (a.k.a. the "Clone", hereafter J1206) is a $z = 2.00$ arc discovered by Lin et al. (2009), who determine a lensing magnification $\mathcal{M} = 27 \pm 1$. The second, SDSS J090122.37+181432.3 (hereafter J0901) is a $z = 2.26$ arc discovered by Diehl et al. (2009); preliminary lens modelling implies its magnification is also high. Both objects are 20–30 times brighter than galaxies at the knee of the $1.9 \leq z \leq 2.7$ rest-UV luminosity function (Reddy et al. 2008). Here we focus on what can be learned about the conditions in the dusty regions of these galaxies from their integrated mid-infrared spectra, based on comparisons with local galaxies, assuming that there is minimal differential lensing across the wavelength ranges of our IRS spectra. Analyses of *Spitzer*/IRAC and (for J1206) MIPS imaging of these targets in light of more refined *HST*-based lens models, together with detailed comparisons to the source-plane properties of similar lensed star-forming galaxies (Siana et al. 2008,

2009) and high-redshift systems selected through their dust emission rather than rest-UV colors (e.g., Lutz et al. 2005; Weedman et al. 2006; Valiante et al. 2007; Yan et al. 2007; Pope et al. 2008; Menéndez-Delmestre et al. 2009), are deferred to future papers.

3.2 Observations and Data Reduction

We used the Infrared Spectrograph (IRS: Houck et al. 2004) on board the *Spitzer Space Telescope* to obtain $14 - 38\,\mu$m spectra of both J1206 and J0901 in the instrument's "long low" mode ($R \sim 57 - 126$), for which the first (LL1) and second (LL2) orders cover wavelength ranges of $19 - 38\,\mu$m and $14 - 21.3\,\mu$m, respectively. In order to ensure optimal signal-to-noise ratios, we followed the recommendations of Teplitz et al. (2007) and mapped the targets at six positions across the slit. Observations of J1206 were taken during *Spitzer* Cycle 4 (PID 40430; PI S. Allam) on 2007 December 19–20, and consisted of 1×6 pointings in LL1 and LL2 for total integration times of 2.1 ks and 2.2 ks, respectively. J0901 was observed on 2009 May 15 during *Spitzer* Cycle 5 (PID 50086; PI S. Allam) using 2×6 pointings in LL1 totaling 2.2 ks and 2.0 ks, and 1×6 pointings in LL1 totaling 2.2 ks. Data were obtained under nominal operating conditions, with the exception of the second LL1 Astronomical Observation Request (AOR) for J0901. During this AOR, *Spitzer* began to warm up due to the depletion of its cryogen. The increased thermal background was marginal, raising data collection event (DCE) values by only 6%. The IRS support team deemed the data nominal, and we reduced them following the same procedures used for our other AORs.

Data reduction relied on standard analysis packages and followed the procedure described by Teplitz et al. (2007). Using IRAF, we removed latent charge in the IRS images row-by-row by fitting the linear background increase over time. Subsequently, we masked "rogue" pixels using the IDL routine IRSCLEAN. After cleaning, we constructed sky images for each target position using the five other pointings from the same mapping AOR. The resulting sky images were subtracted from the corresponding

science frames, and the differences were then co-added to produce a final 2D spectrum at each position. We extracted 1D spectra using the SPICE package provided by the Spitzer Science Center, using optimal extraction. We used an extraction aperture half the default size in order to avoid contamination from other sources (see below), but we corrected to full-aperture fluxes using observations of standard stars.

For J0901, extraction of the 1D spectra was complicated by an interloping source, SDSS J090125.59+181427.8 (see Figure 3.1a), lying $\sim 46''$ away from the lens. From its IRS spectra, this object is likely to be a quasar at $z \sim 1.3$, in agreement with the assessment of its optical colors by Richards et al. (2009). Its spectrum in one pointing often lay on or near the position of a J0901 spectrum in *another* pointing; combined with the fact that J0901 itself is bright and extended, this situation meant that a given "sky" frame included the true background, light from the interloping source, and residual flux from J0901 at other map positions, leading to oversubtraction in our final 2D spectra (e.g., Figure 3.1a). To correct for this effect, we measured the negative sky echoes at *different* locations in the 2D spectrum whose combination should have experienced the *same* oversubtractions (e.g., Figure 3.1b). By combining such measurements, we constructed a model for our oversubtraction of each of our target spectra; correcting led to a $\sim 10 - 25\%$ flux increase for the source.

For both J1206 and J0901, the error spectrum at each position was calculated using SPICE and standard deviation frames constructed in IRAF. The final error spectrum is that of all the positions added in quadrature. For J0901, we include an additional systematic uncertainty associated with our sky correction, raising the final error spectrum by $\sim \sqrt{2}$.

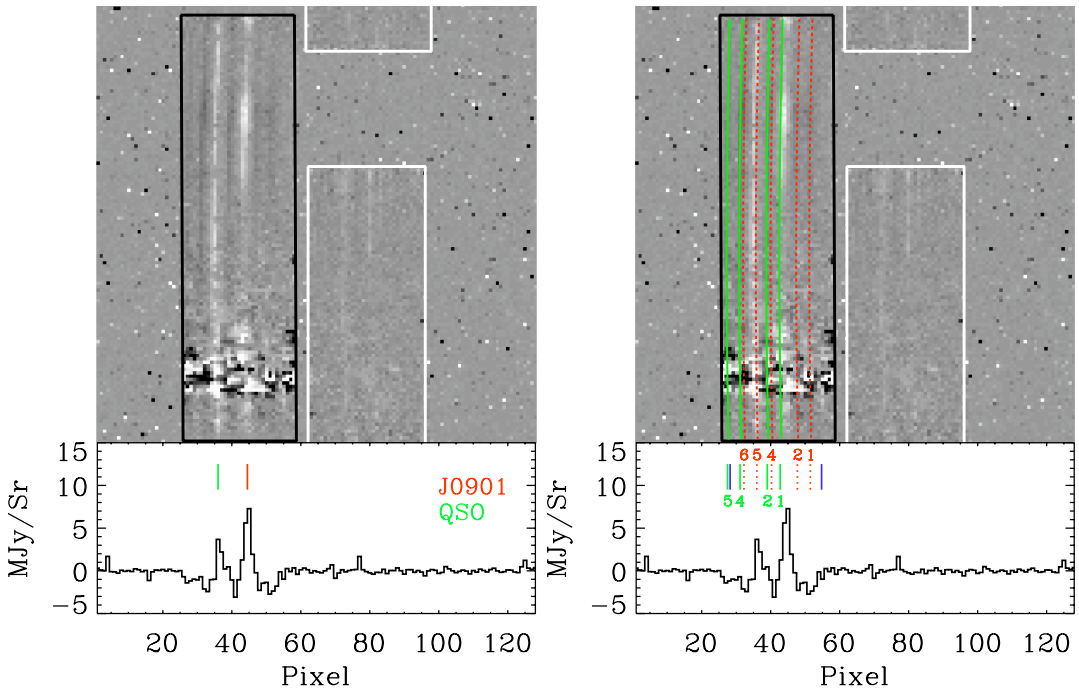

Figure 3.1 (a) Reduced 2D spectrum for J0901 at our third map position, after subtraction of a model for the sky. The black and white rectangles indicate the areas of the array exposed to the spectrometer for LL1 and LL2, respectively, with wavelength increasing downward. Two LL1 spectra are visible: the detection on the right is our target J0901, while on the left is the interloping source SDSS J090125.59+181427.8. From the pattern of the neighboring pixels, it is clear that the background is oversubtracted due to the relatively bright spectra of J0901 and the additional source. Plotted below the spectrum are the data for the row corresponding to the peak of the J0901 spectrum, where the oversubtraction is the worst. (b) The same data overplotted with the positions of the two spectra for the other map positions, indicated by the numbers in the bottom plot. The dotted red lines mark the positions of the J0901 spectra, while the solid green lines mark those of the additional source. For this (i.e., the third) map position, the spectrum for J0901 lies just next to the rightmost solid green line (the position of the interloping source for our first map position) and between two dashed red lines (the positions of J0901 for our second and fourth map positions). To correct for this effect, we measure the sky at appropriate positions, indicated by the blue line segments. The combination of these measurements gives an accurate model for the oversubtraction, which is then added to the data.

3.3 Results

3.3.1 SDSS J120602.09+514229.5

Figure 3.2 a shows the extracted IRS spectra for SDSS J1206. On top of a rising continuum, prominent PAH features are present at 6.2, 7.7, and 11.3 μm. In addition, a strong [S IV] feature is present at 10.5 μm. In order to compare the spectrum to those of local starbursting analogs, we fit template *Infrared Space Observatory* spectra of individual galaxies from Sturm et al. (2000) and the average *Spitzer* starburst template from Brandl et al. (2006). For the fit we allow a varying contribution from the templates (normalized to 6.2 μm) as well as an additional power law continuum: $F_{\nu,\mathrm{fit}} = C_1 \times [\mathrm{Template}/\mathrm{Template}_{6.2\,\mu\mathrm{m}}] + C_2 \times (\lambda/6.2\,\mu\mathrm{m})^{\alpha}$. The three parameters for the fit were sampled using a standard Metropolis-Hastings MCMC (Markov chain Monte Carlo) algorithm. Figure 3.2b shows two of the template spectra that provide good fits. Table 3.1 shows the inferred median, 68% confidence limits, and best-fit parameters. We find that the spectrum is well fit by the M82 template ($\chi^2_{\mathrm{red}} = 0.96$), and marginally fit by the NGC 253 and average starburst templates ($\chi^2_{\mathrm{red}} = 1.20, 1.14$). In contrast, the other templates from Sturm et al. (for 30 Doradus, Circinus, and NGC 1068) fit the spectrum poorly ($\chi^2_{\mathrm{red}} > 10$) due to the lack of strong PAH features. For each of the acceptable templates, the best fit includes an additional, steeply-rising continuum with a power law index of $\alpha \sim 3.3$.

To extract and interpret the features of the spectrum, we have also computed a fit using the Drude profiles defined in Draine & Li (2007) for PAH features, Gaussian line profiles for ionized species expected to be strong, and a simple power law continuum ($\propto \lambda^{\alpha}$). In addition, the fit includes extinction effects from 9.7 μm silicate absorption. The results of this fit are plotted in Figure 3.2c and tabulated in Table 5.5. To assess the strength of the PAH emission, we compare the rest-frame equivalent widths (EWs) of the 6.2, 7.7 ($\equiv 7.41 + 7.61 + 7.85$), and 11.3 ($\equiv 11.23 + 11.33$) μm PAH features to

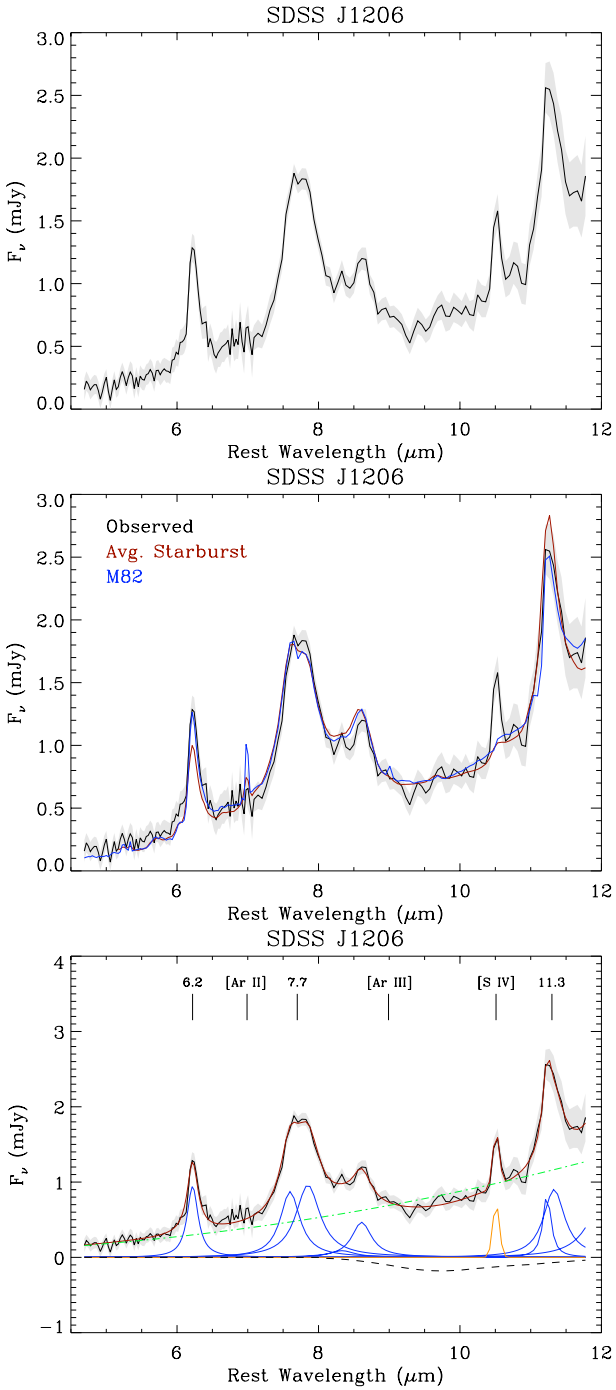

Figure 3.2 (a) IRS spectrum of J1206. The observed spectrum is plotted in black after redshifting to rest wavelength for $z = 2.00$. The light grey band corresponds to the 1σ uncertainty associated with the spectrum. (b) Plotted in blue and red, respectively, are the best fit template spectra for M82 and the average starburst of Brandl et al. (2006) for the parameters listed in Table 3.1. (c) A comprehensive fit (red) to the spectrum using Drude profiles for PAH features (blue), Gaussian profiles for ionic lines (orange), and a power law continuum (green dot dashed). The fit also includes 9.7 μm silicate extinction (black dashed).

those of local starbursting systems. Comparisons of EWs are known to be sensitive to the details of how authors define the underlying continuum. In particular, continuum levels are often defined by the values of the data seen on either side of emission features (see e.g., Brandl et al. 2006; Pope et al. 2008). Such definitions result in systematically higher continua and lower PAH EWs than found by comprehensive fits to the spectra (e.g., Siana et al. 2009). To facilitate interpretation of the EW values reported in Table 5.5, we analyze the average starburst spectrum of Brandl et al. (2006) using the same conventions as for J1206. We find EWs for the average starburst spectrum that are factors $\sim 1-8$ higher than derived by Brandl et al. (2006) for their own data. Nevertheless, this approach provides a consistent way of measuring and comparing EW values to our data. Relative to the average starburst spectrum, J1206 has EWs that are lower by factors of 1.2, 1.5, and 1.2 for the 6.2, 7.7, and 11.3 μm PAH features, respectively. This slight deficiency is not surprising, given the additional power-law continuum preferred by our template fitting above.

The two most striking features of the $\sim 4.5-12\,\mu$m spectrum of J1206 are its steep underlying continuum and prominent [S IV] emission. In high-resolution studies of local star-forming galaxies, the latter line is fairly common and appears weakly in starbursts (Bernard-Salas et al. 2009) and ultraluminous infrared galaxies (ULIRGs) (Farrah et al. 2007), but is much stronger in blue compact dwarfs (BCDs; Hao et al. 2009). At lower resolution this line is unresolved, and only the strongest emitters are detected (see, e.g., Brandl et al. 2006; Wu et al. 2006). Comparing the relative strengths of [S IV] and PAH emission of J1206 to those of local counterparts, we identify the two low-resolution mid-IR spectra of the starburst NGC 1222 (Brandl et al. 2006) and the BCD UGC 4274 (Wu et al. 2006) as close analogs.

Using the optical spectroscopy of Liu & Kennicutt (1995) for NGC 1222 and Ho et al. (1997) for UGC 4274, we calculate the oxygen abundance in each galaxy using the $N2$ and $O3N2$ indicators calibrated in Pettini & Pagel (2004). We find $12+\log(\mathrm{O/H})_{\mathrm{N2}} =$

$8.48, 8.43$ and $12+\log(\mathrm{O/H})_{\mathrm{O3N2}} = 8.38, 8.37$ for NGC 1222 and UGC 4274, respectively. For J1206, Hainline et al. (2009) use near-IR spectroscopy to find $12 + \log(\mathrm{O/H})_{\mathrm{N2}} = 8.50 \pm 0.18$ and $12 + \log(\mathrm{O/H})_{\mathrm{O3N2}} = 8.34 \pm 0.14$, in good agreement with the putative local counterparts. Given the consistency between the spectra and the oxygen abundances, we conclude that NGC 1222 and UGC 4274 have physical conditions similar to those of J1206.

In addition to oxygen abundances, we can consider the ratios of ionized sulfur and neon lines [S IV] $10.5\,\mu$m/[S III] $18.7\,\mu$m and [Ne III] $15.6\,\mu$m/[Ne II] $12.8\,\mu$m for NGC 1222 and UGC 4274. These quantities are well-known proxies for the hardness of the radiation field (see Figure 9 of Hao et al. 2009), which is also a function of metallicity (Wu et al. 2006). The galaxies have similar ratios of [Ne III]/[Ne II] ~ 1.30 and [S IV]/[S III] ~ 0.35, and lie between lower excitation starbursts and higher excitation BCDs on a [S IV]/[S III] versus [Ne III]/[Ne II] excitation diagram (Hao et al. 2009), implying a moderately hard radiation field. Using stellar models, Thornley et al. (2000) estimate the hardness of the radiation in starbursts by relating [Ne III]/[Ne II] to the ratio of the infrared and Lyman continuum luminosities ($L_{\mathrm{IR}}/L_{\mathrm{Lyc}}$). Since NGC 1222 and UGC 4274 have higher [Ne III]/[Ne II] ratios than the Thornley et al. sample, we extrapolate their results and find $3 \lesssim L_{\mathrm{IR}}/L_{\mathrm{Lyc}} \lesssim 20$ indicating a somewhat lower range than for their more typical starbursts ($4 \lesssim L_{\mathrm{IR}}/L_{\mathrm{Lyc}} \lesssim 30$). If present in J1206, such hard radiation would cause significant heating of small dust grains and naturally explain the steep continuum in our IRS spectrum. Unfortunately, [Ne III], [Ne II], and [S III] all lie outside of our spectral coverage, preventing definitive confirmation of this hypothesis.

In J1206, the apparent agreement between metallicity estimates in the optical and the infrared implies that we are not seeing discrepancies of the sort seen in the most violent local mergers. For ULIRGs, abundances derived from optical line diagnostics (even after careful extinction corrections) are lower than those derived from mid-infrared

spectra, likely because the former are depressed by inflows of metal-poor gas from the outskirts of progenitor disks (Rupke et al. 2008) while the latter reflect rapid local enrichment in the most deeply embedded star-forming regions (Veilleux et al. 2009). For J1206, metallicities appear internally consistent across both obscured and unobscured regions, suggesting a less traumatic recent history.

We can also consider the absence of other emission lines in our IRS spectrum, which shows no significant ionic or molecular features other than [S IV]. In Table 5.5, we report 3σ upper limits on [Ar II] at 6.99 μm, [Ar III] at 8.99 μm, and [Ne VI] at 7.65 μm. At first glance, it seems surprising we do not detect [Ar II], given that it is commonly detected in starburst systems (Brandl et al. 2006), including M82 and NGC 253. As discussed above, however, the presence of a rising continuum and [S IV] emission imply higher excitation in J1206 than in average starbursting systems. Thus, in J1206, [Ar II] is likely weak because most of the argon is more highly ionized. This effect is seen in BCD galaxies, which are known to exhibit high excitation states (Hao et al. 2009; Hunt et al. 2010). As a consistency check, we note that our upper limit for [Ar III] implies a ratio of [Ar III]/[S IV] that is consistent with values seen in local starbursts and BCDs (Brandl et al. 2006; Wu et al. 2006).

In principle, an alternative explanation for the steep continuum and [S IV] emission in J1206 is the presence of an AGN. Indeed, Seyfert 2 galaxies are known to show PAH and [S IV] features, and exhibit strong, rising continua due to heating of small dust grains. Considering the flux ratio of 6.2 μm PAH to $5.1 - 6.8$ μm continuum, however, we find that the strength of the 6.2 μm flux means J1206 more closely resembles PDR and HII type spectra, with $\lesssim 5\%$ of the emission contributed by an AGN (Laurent et al. 2000). Reinforcing this conclusion, Seyferts with similar [S IV] and [Ar II] strengths have 6.2 μm PAH EWs that are $> 2\sigma$ lower than for J1206 (Gallimore et al. 2010).[1]

[1]While such comparisons involve difficulties in continuum definitions (see above), we note that Laurent et al. and Gallimore et al. also use comprehensive fits of all relevant mid-IR features to derive their continua, and therefore should have results similar to ours.

Finally, J1206 has a ratio [Ne VI]/[S IV]< 0.38 (3σ), which is lower than that in any Seyfert for which [S IV] is detected (Sturm et al. 2002). We conclude that nuclear activity plays little role in the mid-IR spectrum of J1206.

3.3.2 SDSS J090122.37+181432.3

Using the reduction procedure outlined in Section 3.2, we have derived the final rest-frame spectrum for J0901 that is shown in Figure 3.3a. Following the same procedure above, we fit the same starburst templates to the spectrum as in Section 3.1 with the results shown in Table 3.1. As for J1206, we find that the spectrum is best fit by a scaled version of M82 ($\chi^2 = 0.95$), less well by NGC 253 or the average starburst template of Brandl et al. (2006) ($\chi^2 = 1.42, 1.26$, respectively), and poorly by the other Sturm et al. templates ($\chi^2_{\rm red} > 20$). For the NGC 253 and average starburst fits, the higher χ^2 values originate from the enhanced 9.7 μm silicate absorption in J0901, which is not seen in the templates (see Figure 3.3b). In contrast to J1206, J0901 exhibits only weak evidence for additional power-law emission, with fits favoring a component that is negligible or has highly uncertain parameters.

As for J1206, we simultaneously fit the spectrum with a combination of continuum and relevant PAH features and atomic emission lines. We find the degree of inferred silicate absorption is degenerate with the contributions of weak PAH emission features to the spectrum. In general, this degeneracy has little effect on the inferred PAH emission, except in the relative contributions of the 11.23 and 11.33 μm lines to the 11.3 μm PAH feature. In fact, the 6.2 μm and blended 7.7 μm and 11.3 μm PAH strengths are essentially unaffected. Ultimately, we have opted to use solutions with higher silicate extinction, in agreement with the template fits above. We present the results of this fit in Table 5.5 and plot the results in Figure 3.3c.

The PAH emission in J0901 is strong relative to local starbursts, with EWs of each feature that are a factor $\sim 1.6 - 1.9$ times larger than the Brandl et al. (2006) average

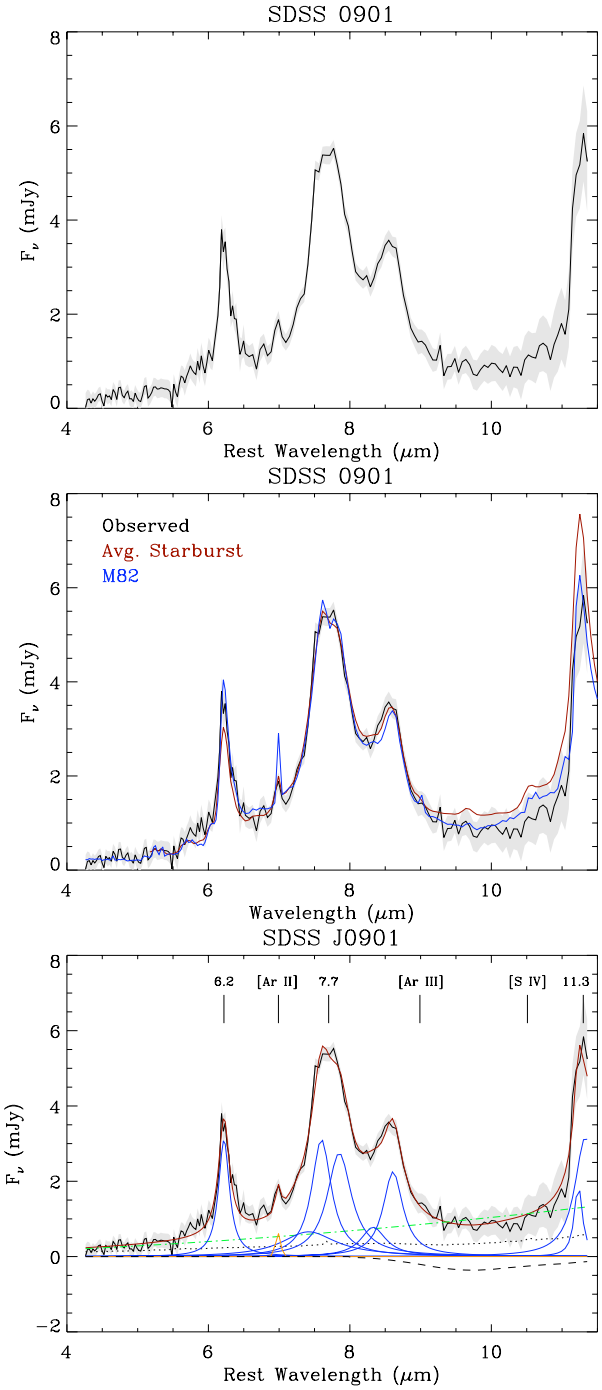

Figure 3.3 (a) Observed IRS spectrum of J0901 plotted in black after redshifting to rest wavelength for $z = 2.26$. (b) Template fits to the spectrum, with notation as in Figure 2b. (b) Comprehensive fit, with notation as in Figure 2c. The additional dotted line indicates the AGN contribution to the spectrum implied by rest-frame optical measurements (see Section 3.2).

values. The strengths of the 6.2, 7.7, and 11.3 μm PAH features serve as common diagnostics of the source(s) of infrared emission in active galaxies (e.g., Imanishi et al. 2007; Veilleux et al. 2009; Baum et al. 2010; Gallimore et al. 2010). In particular, suppressed PAH emission (especially at smaller wavelengths) is indicative of a significant AGN contribution to the bolometric infrared luminosity. Given the strength of the observed PAH emission, coupled with the shallow continuum, we conclude that accretion plays a small role in the mid-IR properties of J0901. However, Hainline et al. (2009) examined the optical emission line ratios and found values of [O III]/Hβ and [N II]/Hα indicative of an AGN, consistent with the significant [N V] and weak [Si IV] and [C IV] emission seen in the object's rest-UV spectrum (Diehl et al. 2009).

To assess this apparent contradiction, we consider the expected flux contribution to our mid-IR spectrum by an AGN whose optical properties resemble those of J0901. Specifically, we treat [O III] 5007 Å emission as a proxy for AGN strength and scale the *ISO* template of NGC 1068 by the factor required to reduce the integrated [O III] flux of NGC 1068 (Moustakas & Kennicutt 2006) to the Hainline et al. value. The dotted line in Figure 3.3c shows the corresponding contribution of the scaled NGC 1068 spectrum: roughly 57% and 35% of the continuum flux at 5 and 10 μm, respectively. This comparison demonstrates that reliance on rest-frame UV/optical measurements alone may provide a bolometrically unrepresentative picture of the physical properties of high-redshift systems.

Finally, we examine our IRS spectrum of J0901 for ionic and molecular emission. Here J0901 is quite unlike J1206, showing [Ar II] but no [S IV] emission. The EW of [Ar II] is similar to that in the average starburst spectrum of Brandl et al. (2006) (see Figure 3.3b). As discussed above, [Ar II] is weaker for systems with high-excitation interstellar media. Therefore, J0901 must be bathed in a softer radiation field than J1206, reinforcing the conclusion that AGN emission plays little role in the IR properties of the system. With our upper limit on [Ar III] we find [Ar III]/[Ar II]< 0.83, which

when combined with the argon excitation versus abundance relation from Verma et al. (2003) implies a super-solar metallicity of $Z \gtrsim 1.3\,Z_{\odot}$ for J0901.

3.4 Conclusions

We have obtained *Spitzer*/IRS spectra of two $z \sim 2$ UV-bright star-forming galaxies, that are magnified by strong gravitational lensing. At rest wavelengths of $\sim 5 - 12\,\mu$m, the spectra reveal strong PAH emission at 6.2, 7.7, and 11.3 μm, indicating that these objects are undergoing intense star formation. The strength of the PAH emission implies these objects have properties in line with those of local starbursting galaxies. We find this similarity to local starburst galaxies is confirmed by our empirical template fits, in which both galaxies are well fit by simple, rescaled versions of M82. In detail, however, analysis of PAH strengths and emission line and continuum diagnostics reveals disparate properties. We summarize our conclusions as follows:

1. In J1206, we find PAH EWs lower than those in the local starburst spectrum of Brandl et al. (2006), due to an enhanced power-law continuum. In contrast, J0901 exhibits PAH EWs that are factors 1.6–1.9 times larger than the local average.

2. We detect significant [S IV] emission in J1206. By analogy with two local galaxies with similar mid-IR spectra, NGC 1222 and UGC 4274, we infer a sub-solar metallicities of $\sim 0.5\,Z_{\odot}$, in agreement with the published optical measurement (Hainline et al. 2009). The consistency of the optical and infrared metallicity estimates suggests J1206 has not undergone a recent violent merger. Considering the [S IV]/[S III] and [Ne III]/[Ne II] ratios of the local objects, we argue that J1206 is characterized by a moderately hard radiation field, which naturally explains the steeply rising continuum and lack of [Ar II] emission.

3. In J0901, we detect strong PAH emission but no [S IV] or significant rising continuum. These results indicate that the mid-IR properties of J0901 are consistent

with purely starburst-driven energetics. This inference contrasts with the implications of optical spectroscopy, where emission line ratios show the presence of an AGN; however, scaling from the [O III] flux of a local AGN implies the AGN contributes $< 57\%$ of the mid-IR continuum. Thus, from its rest-frame UV thru IR properties, J0901 likely hosts a narrow line AGN whose IR emission is overwhelmed by that of its surrounding starburst. This analysis highlights the need for future IR studies of high-redshift objects if we are to determine their physical properties robustly.

4. With the detection of [Ar II], we are able to put an lower limit on the metallicity of J0901. Using the argon abundance and excitation relation of Verma et al. (2003), we find $Z \gtrsim 1.3\, Z_{\odot}$, similar to many local starbursts.

Table 3.1. Local starburst template fits.

Template	$\log C_1$	$\log C_2$	α	Comment
		J1206		
M82	$2.063^{+0.006}_{-0.004}$	$3.29^{+0.09}_{-0.15}$	$3.3^{+0.2}_{-0.2}$	MCMC results
	2.060	3.3	3.1	Best Fit: $\chi^2_{\rm red} = 0.96$
NGC 253	$1.938^{+0.128}_{-0.102}$	$3.13^{+0.10}_{-0.12}$	$3.2^{+0.2}_{-0.2}$	MCMC results
	2.030	3.24	3.0	Best Fit: $\chi^2_{\rm red} = 1.20$
Avg. Starburst	$1.960^{+0.051}_{-0.051}$	$2.98^{+0.11}_{-0.19}$	$3.3^{+0.2}_{-0.2}$	MCMC results
	2.012	3.18	3.1	Best Fit: $\chi^2_{\rm red} = 1.14$
		J0901		
M82	$2.567^{+0.054}_{-0.057}$	$3.11^{+0.189}_{-0.141}$	$3.1^{+2.2}_{-2.1}$	MCMC results
	2.604	3.30	2.6	Best Fit: $\chi^2_{\rm red} = 0.95$
NGC 253	$1.822^{+0.045}_{-0.058}$	$3.32^{+0.13}_{-0.18}$	$3.2^{+0.8}_{-0.6}$	MCMC results
	1.87	3.42	3.01	Best Fit: $\chi^2_{\rm red} = 1.42$
Avg. Starburst	$2.449^{+0.038}_{-0.035}$	$-0.24^{+0.18}_{-0.17}$	$0.6^{+0.1}_{-0.1}$	MCMC results
	2.478	-0.24	0.5	Best Fit: $\chi^2_{\rm red} = 1.26$

Note. — Columns on the left use $H_0 = 74.2 \pm 3.6$ km s^{-1} Mpc^{-1} from Riess et al. (2009) while columns on the right use $H_0 = 70.5 \pm 1.3$ km s^{-1} Mpc^{-1} from Komatsu et al. (2008).

Note. — $F_{\nu,\rm fit} = C_1 \times [\text{Template}/\text{Template}_{6.2\,\mu\rm m}] + C_2 \times (\lambda/6.2\,\mu\rm m)^{\alpha}$ mJy. MCMC results report the median and 68% confidence intervals.

Table 3.2. Derived feature strengths.

	J1206		J0901	
Wavelength μm	Observed Flux 10^{-15} erg s^{-1} cm^{-1}	Rest EW μm	Observed Flux 10^{-15} erg s^{-1} cm^{-1}	Rest EW μm
6.22 PAH	21.7 (1.2)	1.12	70.2 (3.1)	2.12
6.99 [Ar II]	< 1.9	< 0.12	3.5 (0.9)	0.11
7.42 PAH	–	–	52.0 (10.8)	1.59
7.60 PAH	24.2 (1.6)	1.21	84.7 (4.5)	2.61
7.65 [Ne VI]	< 0.8	< 0.10	< 2.6	< 0.08
7.85 PAH	30.9 (1.8)	1.55	87.0 (4.1)	2.70
8.33 PAH	3.1 (1.6)	0.20	22.8 (3.9)	0.71
8.61 PAH	10.2 (1.2)	0.51	47.9 (2.9)	1.51
8.99 [Ar III]	< 2.8	< 0.05	< 2.9	< 0.09
10.51 [S IV]	2.1 (0.1)	0.10	< 4.8	< 0.15
11.23 PAH	3.7 (1.3)	0.19	9.6 (6.6)	0.31
11.33 PAH	13.2 (3.4)	0.61	42.1 (15.4)	1.37
11.99 PAH	3.2 (1.4)	0.78	–	–

Note. — Equivalent width values are based on the green dot dashed lines (power-law continuum fits) in Figures 3.2c and 3.3c. Upper limits are based on 3σ uncertainties in the spectra.

Bibliography

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

Allam, S. S., Tucker, D. L., Lin, H., Diehl, H. T., Annis, J., Buckley-Geer, E. J., & Frieman, J. A. 2007, ApJ, 662, L51

Baum, S. A., et al. 2010, ApJ, 710, 289

Belokurov, V., et al. 2007, ApJ, 671, L9

Bernard-Salas, J., et al. 2009, ApJS, 184, 230

Brandl, B. R., et al. 2006, ApJ, 653, 1129

Diehl, H. T., et al. 2009, ApJ, 707, 686

Draine, B. T., & Li, A. 2007, ApJ, 657, 810

Farrah, D., et al. 2007, ApJ, 667, 149

Gallimore, J. F., et al. 2010, ApJS, 187, 172

Hainline, K. N., Shapley, A. E., Kornei, K. A., Pettini, M., Buckley-Geer, E., Allam, S. S., & Tucker, D. L. 2009, ApJ, 701, 52

Hao, L., Wu, Y., Charmandaris, V., Spoon, H. W. W., Bernard-Salas, J., Devost, D., Lebouteiller, V., & Houck, J. R. 2009, ApJ, 704, 1159

Hennawi, J. F., et al. 2008, AJ, 135, 664

Ho, L. C., Filippenko, A. V., Sargent, W. L. W., & Peng, C. Y. 1997, ApJS, 112, 391

Houck, J. R., et al. 2004, ApJS, 154, 18

Hunt, L. K., Thuan, T. X., Izotov, Y. I., & Sauvage, M. 2010, ApJ, 712, 164

Imanishi, M., Dudley, C. C., Maiolino, R., Maloney, P. R., Nakagawa, T., & Risaliti, G. 2007, ApJS, 171, 72

Kubo, J. M., Allam, S. S., Annis, J., Buckley-Geer, E. J., Diehl, H. T., Kubik, D., Lin, H., & Tucker, D. 2009, ApJ, 696, L61

Laurent, O., Mirabel, I. F., Charmandaris, V., Gallais, P., Madden, S. C., Sauvage, M., Vigroux, L., & Cesarsky, C. 2000, A&A, 359, 887

Lehnert, M. D., & Bremer, M. 2003, ApJ, 593, 630

Lin, H., et al. 2009, ApJ, 699, 1242

Liu, C. T., & Kennicutt, Jr., R. C. 1995, ApJ, 450, 547

Lutz, D., Valiante, E., Sturm, E., Genzel, R., Tacconi, L. J., Lehnert, M. D., Sternberg, A., & Baker, A. J. 2005, ApJ, 625, L83

Menéndez-Delmestre, K., et al. 2009, ApJ, 699, 667

Moustakas, J., & Kennicutt, Jr., R. C. 2006, ApJS, 164, 81

Pettini, M., & Pagel, B. E. J. 2004, MNRAS, 348, L59

Pope, A., et al. 2008, ApJ, 675, 1171

Reddy, N. A., & Steidel, C. C. 2004, ApJ, 603, L13

Reddy, N. A., Steidel, C. C., Erb, D. K., Shapley, A. E., & Pettini, M. 2006a, ApJ, 653, 1004

Reddy, N. A., Steidel, C. C., Fadda, D., Yan, L., Pettini, M., Shapley, A. E., Erb, D. K., & Adelberger, K. L. 2006b, ApJ, 644, 792

Reddy, N. A., Steidel, C. C., Pettini, M., Adelberger, K. L., Shapley, A. E., Erb, D. K., & Dickinson, M. 2008, ApJS, 175, 48

Richards, G. T., et al. 2009, ApJS, 180, 67

Rupke, D. S. N., Veilleux, S., & Baker, A. J. 2008, ApJ, 674, 172

Shin, M., Strauss, M. A., Oguri, M., Inada, N., Falco, E. E., Broadhurst, T., & Gunn, J. E. 2008, AJ, 136, 44

Siana, B., Teplitz, H. I., Chary, R., Colbert, J., & Frayer, D. T. 2008, ApJ, 689, 59

Siana, B., et al. 2009, ApJ, 698, 1273

Smail, I., et al. 2007, ApJ, 654, L33

Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, ApJ, 519, 1

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

Steidel, C. C., Giavalisco, M., Dickinson, M., & Adelberger, K. L. 1996, AJ, 112, 352

Sturm, E., Lutz, D., Tran, D., Feuchtgruber, H., Genzel, R., Kunze, D., Moorwood, A. F. M., & Thornley, M. D. 2000, A&A, 358, 481

Sturm, E., Lutz, D., Verma, A., Netzer, H., Sternberg, A., Moorwood, A. F. M., Oliva, E., & Genzel, R. 2002, A&A, 393, 821

Teplitz, H. I., et al. 2007, ApJ, 659, 941

Thornley, M. D., Schreiber, N. M. F., Lutz, D., Genzel, R., Spoon, H. W. W., Kunze, D., & Sternberg, A. 2000, ApJ, 539, 641

Valiante, E., Lutz, D., Sturm, E., Genzel, R., Tacconi, L. J., Lehnert, M. D., & Baker, A. J. 2007, ApJ, 660, 1060

Veilleux, S., et al. 2009, ApJS, 182, 628

Verma, A., Lutz, D., Sturm, E., Sternberg, A., Genzel, R., & Vacca, W. 2003, A&A, 403, 829

Weedman, D. W., Le Floc'h, E., Higdon, S. J. U., Higdon, J. L., & Houck, J. R. 2006, ApJ, 638, 613

Wu, Y., Charmandaris, V., Hao, L., Brandl, B. R., Bernard-Salas, J., Spoon, H. W. W., & Houck, J. R. 2006, ApJ, 639, 157

Yan, L., et al. 2007, ApJ, 658, 778

Yee, H. K. C., Ellingson, E., Bechtold, J., Carlberg, R. G., & Cuillandre, J. 1996, AJ, 111, 1783

Chapter 4

Multi-wavelength Studies of Lens Flux Ratios

This chapter is based on observations from the approved NOAO proposal:
A Gemini Search for Dark Matter Substructure,
GN-2008B-Q-49, R. Fadely (PI) & C. R. Keeton

Using K and L' images obtained at the Gemini North 8m telescope, we examine the wavelength dependence of flux ratios for six gravitational lenses. Selecting lenses with source redshifts $z_s < 2.8$, our K band images probe rest frame optical emission from accretion disks, while L' band images originate from the rest frame near-infrared, emitted (in part) from the more extended surrounding torus. Since the observations correspond to different source sizes, our K and L' flux ratios are sensitive to substructure on different scales and may be useful for identifying and studying small structure in the lenses. We identify two lenses, HE 0435-1223 and SDSS 0806+2006, whose flux ratios show significant multi-wavelength variation. Additionally, we discuss the utility and limitations of the current data for studying dark matter substructure.

4.1 Background

While the cold dark matter (CDM) paradigm for structure formation successfully describes cosmological observations on large (CMB and cluster) scales, there is notable disagreement with small-scale observations. Among other issues, N-body simulations (e.g., Via Lactea, Aquarius) predict the existence of many CDM subhalos, with masses

$M \sim 10^4$–10^9 $M_\odot$ on galaxy scales and $M \sim 10^8$–10^{12} $M_\odot$ on galaxy cluster scales. This has proved troubling observationally, because there are many fewer dwarf galaxies in our own Milky Way than predicted by CDM. Since the discrepancy may be due to baryon stripping from subhalos (e.g., Madau et al. 2008; Macciò et al. 2010), we need ways to probe dark matter substructure directly, regardless of the presence of baryonic material.

Gravitational lensing provides a unique way to detect CDM substructure in distant galaxies and clusters (e.g., Metcalf & Madau 2001; Chiba 2002; Dalal & Kochanek 2002). Stars and CDM substructure perturb the lens potential on micro- to milli-arcsecond scales, which can have dramatic effects on the properties of lensed images. Most notably, lensing from stars and dark matter substructure can alter the flux ratios from those of smoothly distributed mass profiles. As shown by Dobler & Keeton (2006), lens flux ratios depend on the size of the source compared to the size of the perturber. When the source is very small, it is effectively a point source for the lens substructure, resulting an additional magnification boost ($\mu_{\rm sub}$) on top of the macroscopic lens properties. As the source increases in size, $\mu_{\rm sub}$ may increase or decrease depending on the source's location relative to the substructure and the parity of the "macro" magnification. For very large source sizes, the substructure is too small to affect the image and $\mu_{\rm sub} \to 1$. Figure 4.1 provides an illustration of the behavior of $\mu_{\rm sub}$ as a function of source size. This phenomenon implies that by measuring flux ratios at different wavelengths, corresponding to different source sizes, substructure may be mapped out on a variety of scales.

Heuristically, a quasar emitting region of size R_S is significantly affected by a subhalo with Einstein radius R_E only if $R_S \lesssim R_E$. For typical lens and source redshifts ($z_l = 0.5, z_s = 2.0$), the Einstein radius of a subhalo of mass M is $R_E \sim 10^{16}$ cm $(M/M_\odot)^{1/2}$. Since the optically emitting regions of QSOs have $R_S \sim 10^{15}$–10^{16} cm (Wyithe et al. 2000; Pooley et al. 2007; Morgan et al. 2010), optical lens flux ratios are sensitive to both

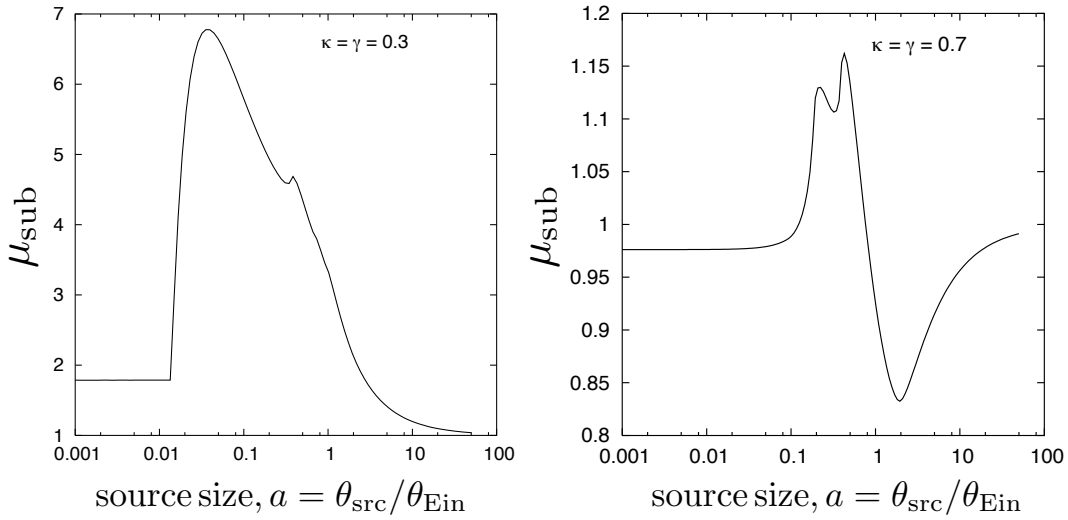

Figure 4.1 Two panels showing the additional magnification factor $\mu_{\rm sub}$ provided by a small singular isothermal clump near an image with positive parity (*left*) and negative parity (*right*), as a function of the relative size of a uniform, circular source a. For sources which are small compared to the Einstein radius of the clump ($\theta_{\rm Ein}$), $\mu_{\rm sub}$ is constant and not equal to unity. As a increases, $\mu_{\rm sub}$ changes due to lensing caustics provided by the clump. Finally, for very large sources, the clump is unable to change the magnification of the image and asymptotes to unity. This figure is adopted from Figures 5 and 6 of Dobler & Keeton (2006).

microlensing by stars and millilensing by CDM substructure. By contrast, the more extended infrared emitting regions with $R_S \gtrsim 1$ pc (Chiba et al. 2005; Minezaki et al. 2009; Agol et al. 2009) can only be affected by relatively massive subhalos. Comparing lens flux ratios at different wavelengths therefore makes it possible to constrain the amount of micro- and milli-lensing present in the system, as well as the sizes of the perturbers.

Infrared studies of lens flux ratios have provided interesting investigations of substructure properties. With Suburu 11 μm imaging of four image lenses, substructure with masses $\gtrsim 10^5 M_\odot$ has been found in the lenses B1422+231 and MG0414+0534 (Chiba et al. 2005; Minezaki et al. 2009, respectively). Additionally, the use of infrared flux ratios has been broadly tested, with recent studies finding both luminous satellites (More et al. 2009; MacLeod et al. 2009) and null detections (Chiba et al. 2005; Minezaki et al. 2009; Agol et al. 2009). Expanding this work, we obtained K (2.2 μm) and L'

(3.8 μm) band images of six lenses using Gemini North during the 2008B semester.[1] At rest frame wavelengths upwards of 1μm, an increasing percentage (from $\sim$ 20% to 100%) of AGN luminosity comes from thermal emission by the inner dusty torus ($\sim$ 1 pc) (Minezaki et al. 2004; Agol et al. 2009). Limiting our source redshifts to $z < 2.8$, L' band observations correspond to emission whose luminosity originates from large emitting regions. In contrast, K band images originate mostly from the small accretion disk. Thus, comparison between K and L' flux ratios may provide the source size baseline necessary to identify substructure. Below we discuss the observations and their interpretation in detail.

4.2 Observations in the Near Infrared

To explore the multi-wavelength behavior of flux ratios, we obtained the first K (2.2 μm) and L' (3.8 μm) images for six gravitational lenses using the Near Infrared Imager (NIRI) located at the Gemini North facility. Our lens sample was selected from a set of well known lenses which can be found in the CASTLES database[2]. The lenses (Table 4.2) consist of a subset of the CASTLES sources, which have relatively bright ($m_H < 18.5$) and well separated ($> 1''$) images. The lenses were also selected to have source redshifts $z < 2.8$, corresponding to rest frame emission $> 1.0\,\mu$m in our L' band observations. Information on the observations is also given in Table 4.2.

All of the lenses in the sample have been previously studied using ground and space based photometry. As part of the CASTLES survey, all of the lenses have single epoch *Hubble Space Telescope* (*HST*) V, I, and H band photometry which have been used in earlier studies to derive lens models. Many other ground based studies have been conducted on our sample. Of these, optical monitoring efforts of Kochanek et al. (2006),

[1] In 2009B, two nights of classical time were granted at Gemini North. Unfortunately, poor weather conditions prevented successful observations.

[2] http://www.cfa.harvard.edu/castles/

Koptelova et al. (2008), and Goicoechea et al. (2008) are particularly useful in examining the multi-wavelength behavior of the lenses, since they help quantify intrinsic variations in the central engine, as well as degree of microlensing present (e.g., Kochanek et al. 2006).

Observing strategies in the near-infrared are quite different than for typical optical studies. At wavelengths greater than $2\,\mu$m, thermal emission from the Earth's atmosphere becomes an increasingly significant contributor to the flux recorded at the telescope. In order to remove the (often dominant) emission from the foreground sky, observations in the infrared are nodded between positions which are on and off source. Fortunately, the targets listed in Table 4.2 are all small in size (image separations $\lesssim 3''$) and require small nods. Since our L' observations have a field of view of $22''$, we decided to "nod on chip" whereby a small 2×2, $6''$ dither pattern was used. With this dither pattern, for any single dither position, three other positions are available to construct sky model images. This dither strategy, therefore, effectively doubles our observing efficiency by eliminating the need for off-source. In addition to providing sky subtraction images, dithering reduces the sensitivity of the observations to hot or bad pixels in the NIRI array.

For K band images, single exposures of $20 - 40$s were selected so that the bright quasar images (extrapolated from 1.6 to $2.2\,\mu$m assuming $f_\nu \propto \nu^{-0.5}$, Stern et al. (2005)) would not exceed more than 50% of the detector well. In the L' band, the foreground emission is much brighter than at shorter wavelengths and exposure times of 1s were selected to prevent saturation. To accumulate the necessary integration time, these short exposures were co-added in 30s blocks at each dither position.

Data reduction of K and L' images from Gemini was conducted with standard IRAF routines provided by the Gemini Observatory. The reduction proceeded as follows. First bad pixel masks and sigma clipping algorithms were applied to the data, accounting for defective and hot pixels in the detector. A flat fielding was then applied to the

Table 4.1 Targets and Observational Information

Target	z_{lens}	z_{source}	Image sep. ($''$)	Night(s) (HST)†	Total Integration K (min)	Total Integration L' (min)
Q0142-100	0.491	2.719	2.24	8.21.2008	3	32
SDSS 0246-0825	0.723	1.680	1.04	8.18.2008, 8.20.2008	4.8	76
HE 0435-1223	0.455	1.689	2.42	8.31.2008,12.22.2008	4.8	64
SDSS 0806+2006	0.573	1.540	1.40	12.22.2008	2.4	20
SBS 0909+523	0.830	1.376	1.17	12.22.2008	1.6	2
HE 2149-2745	0.603	2.033	1.70	12.22.2008	3	32

† Hawaii Standard Time

data. For K band images, sky brightness is subdominant and separate calibration flats were taken the same night as the observations. Since the sky dominates L' band observations, the images themselves are essentially sky flats and are used to construct flat field images. After flat fielding, the data are sky-subtracted using observations from other positions in the same dither. L' sky images were limited to the positions immediately before and after the observation, since sky brightness can vary on the scale of minutes at 3.8 μm (Glass 1999). Once each dither position has been flat fielded, sky subtracted, and sigma-clipped, the final image is produced by combining the various dither positions. Figure 4.2 illustrates various steps of the reduction procedure for the lens SBS 0909+523. In Figures 4.3 and 4.4 we present the reduced K and L' images of the targets listed in Table 4.2.

4.3 Flux Ratio Extraction

Using the `GALFIT` image modeling package (Peng et al. 2002), we extract the K and L' fluxes of the lensed images and (when detected) the lensing galaxy. `GALFIT` provides not only a means of deriving accurate photometric measurements of our images, but also provides a method to separate the flux of lensed images from that of the nearby lens galaxy. The latter feature is particularly important in our K band images, in which we detect the lens galaxy in every target except SBS 0909+523.

We produce our photometry using the procedure below, which is similar to that of Sluse et al. (2008). Initially, we select the lensed quasar image with the largest separation from the lens galaxy as our initial model for the point spread function (PSF)[3]. We then perform a simultaneous fit to all the lens images (modeled as a scaled PSF) and any galaxies present (modeled as a PSF-convolved Sérsic profile). With this fit in hand, we subtract the galaxy (or galaxies) from the original image and then re-exact a new PSF model from the same lens image, which now contains less contamination from

[3]If image separations are comparable, we select the brightest image.

the lens galaxy. Using this cleaner PSF, a simultaneous fit is again performed and the reduced χ^2_r value of the residuals is examined. χ^2_r values are calculated using the regions defined by the PSF size for the quasar images, and in regions of 1, 2, & 3 $\times$ the PSF size for the lens galaxy. If the $\chi^2_r = 1$ within these regions, we conclude the procedure is converged. If not, additional iterations of PSF extraction are performed with an improved model for the lens(es) in each step, until convergence is achieved. Our K band images typically require 3 or 4 iterations, while our L' images require only 1 or 2 since the lens galaxy is rarely detected in L'. An example of our `GALFIT` modeling procedure is shown in Figure 4.5. In order to calibrate of the flux measurements provided by `GALFIT`, we measure the flux of standard stars for each target using aperture photometry. Lastly, we note that the quasar image which serves as the PSF model is fit perfectly. For these images, uncertainty is derived by computing the range in flux where $\Delta\chi^2 = 1$ in the PSF region. When images are clearly distinct, derived flux values and uncertainties are verified using aperture photometry. We report the flux ratios for the targets in Table 4.2.

Table 4.2. K and L' Flux Ratios

Target	Image	K	L'
Q0142-100	A	$\equiv 1.0$	$\equiv 1.0$
	B	0.128 ± 0.002	0.132 ± 0.006
SDSS 0246-0825	A	$\equiv 1.0$	$\equiv 1.0$
	B	0.258 ± 0.015	0.331 ± 0.016
HE 0435-1223	A	1.837 ± 0.130	1.706 ± 0.129
	B	1.271 ± 0.073	0.991 ± 0.074
	C	$\equiv 1.0$	$\equiv 1.0$
	D	0.851 ± 0.049	0.809 ± 0.102
SDSS 0806+2006	A	$\equiv 1.0$	$\equiv 1.0$
	B	0.406 ± 0.030	$< 0.164\,(3\sigma)$
SBS 0909+523	A	$\equiv 1.0$	$\equiv 1.0$
	B	0.973 ± 0.028	0.973 ± 0.033
HE 2149-2745	A	$\equiv 1.0$	$\equiv 1.0$
	B	0.280 ± 0.006	0.240 ± 0.009

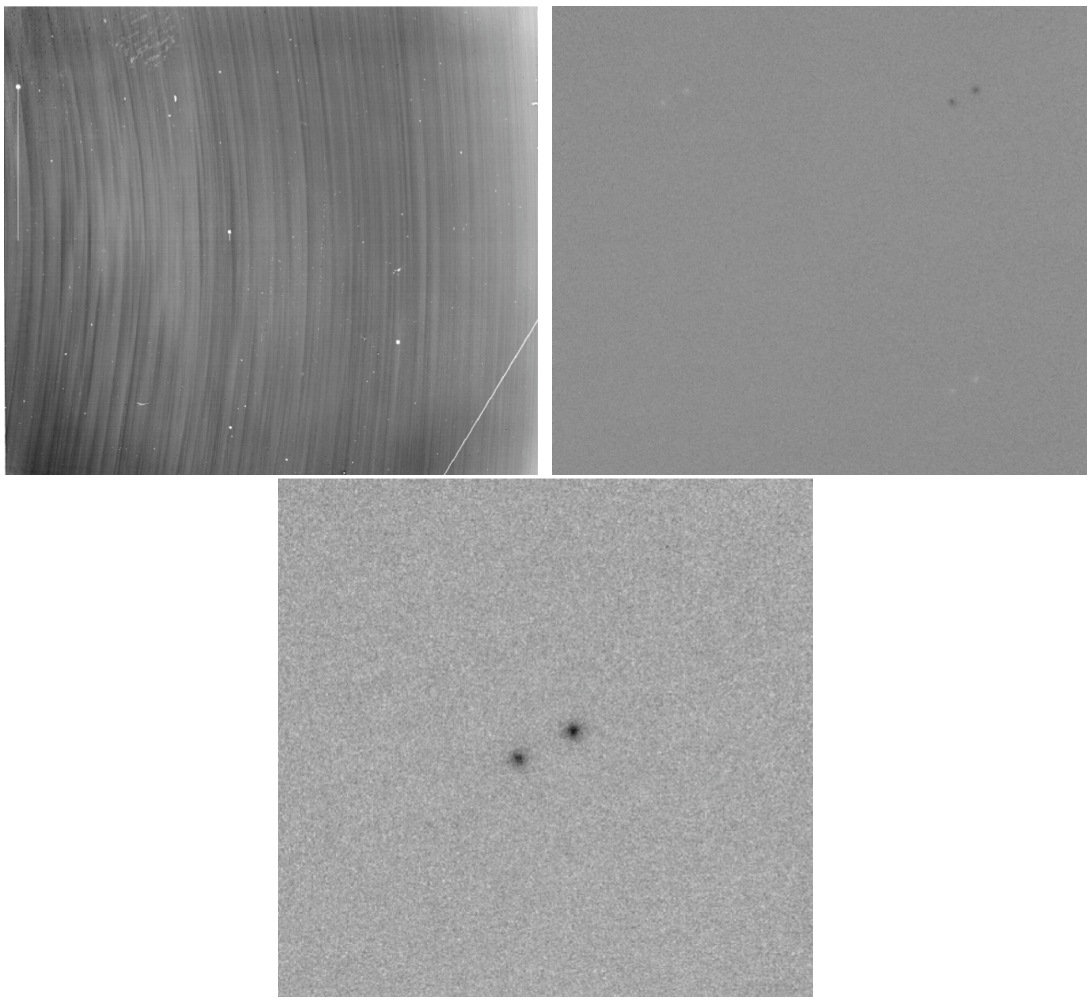

Figure 4.2 The L' band reduction sequence of SBS 0909+523. The top left panel shows the raw data from 30s of coadded exposures (pixel counts $\sim 10^5$ ADU). Detector defects, like the prominent crack in the right hand corner, are clearly visible. After flat fielding, dither positions immediately before and after are used to subtract a model for the bright foreground sky, resulting in the image on the top right. The two lensed images (black) and the residuals from the sky subtraction (white) are just visible in this single 30s coadded image. The bottom panel shows the final image (noise counts at ~ 50 ADU, S/N > 10), after combining all the data from the four dither positions.

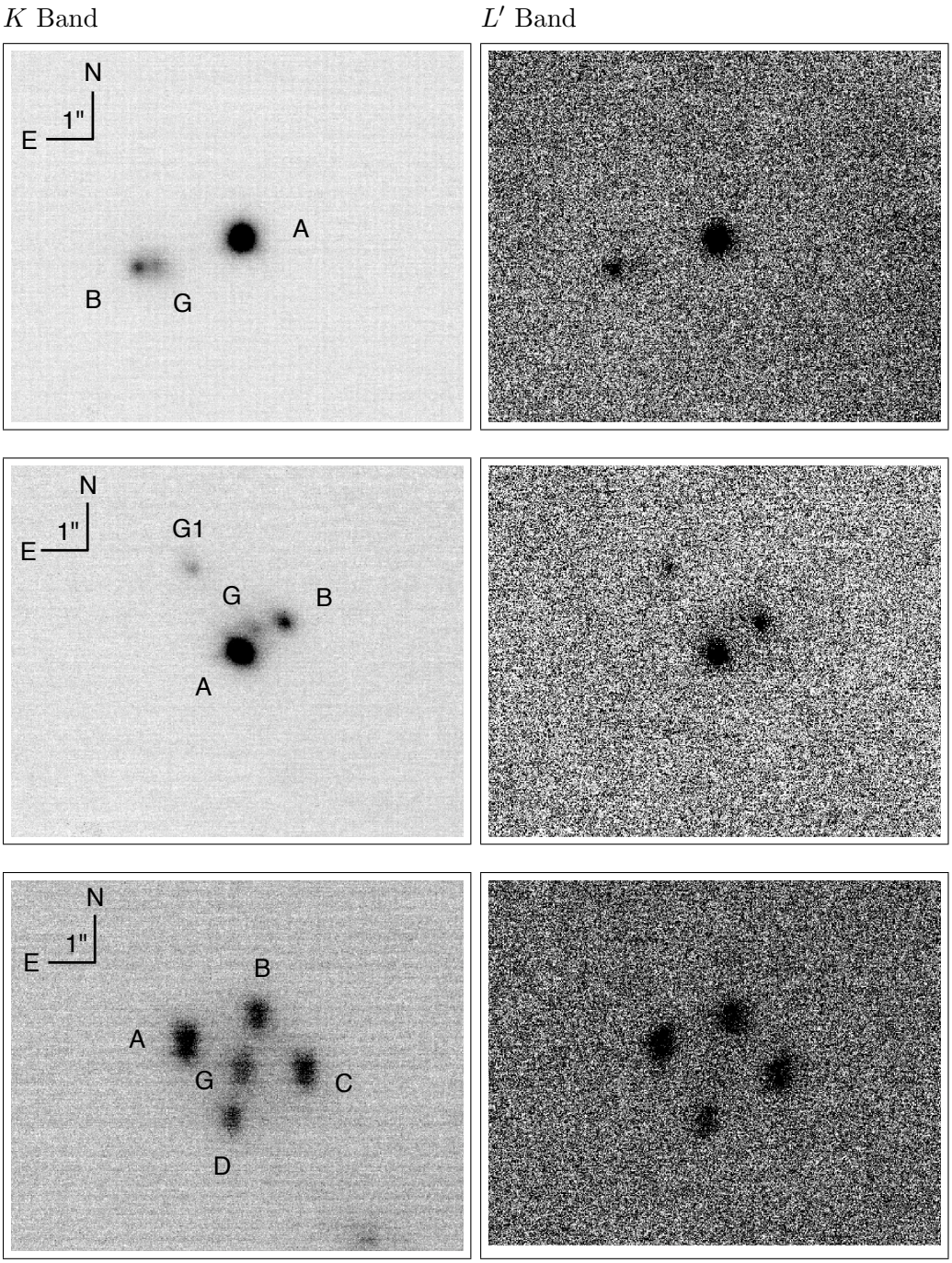

Figure 4.3 From top to bottom, the K (left) and L' (right) images of Q0142-100, SDSS 0246-0825, and HE 0435-1223 are presented. Lensed images are labeled as A − D with the lens galaxies labeled as G or G1. Lens galaxies are clearly visible in the K band images. With the exception of G1 in SDSS 0246-0825, lens galaxies are not seen the L' images.

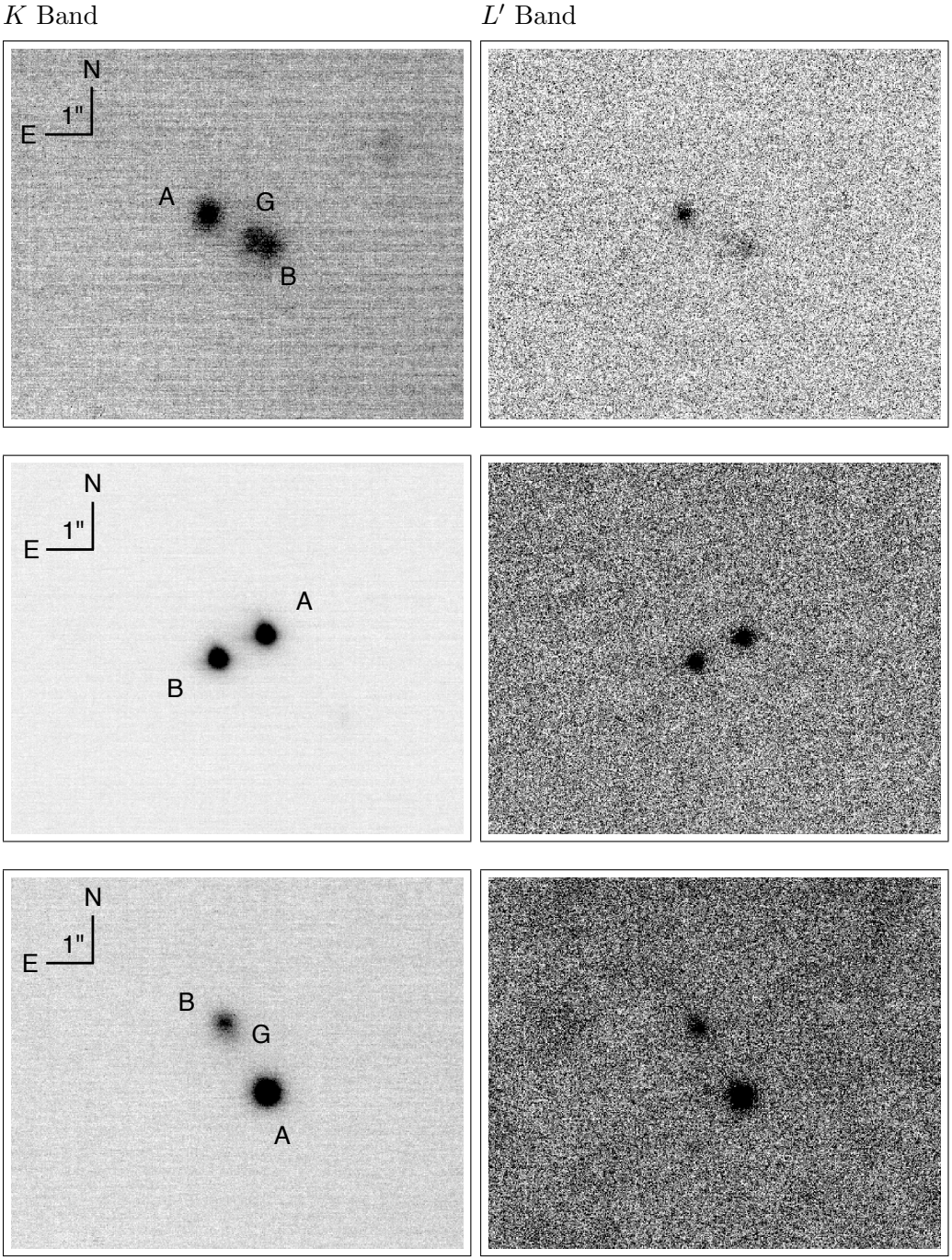

Figure 4.4 The same as for Figure 4.3, but for the lenses SDSS 0806+2006, SBS 0909+523, and HE2149-2745.

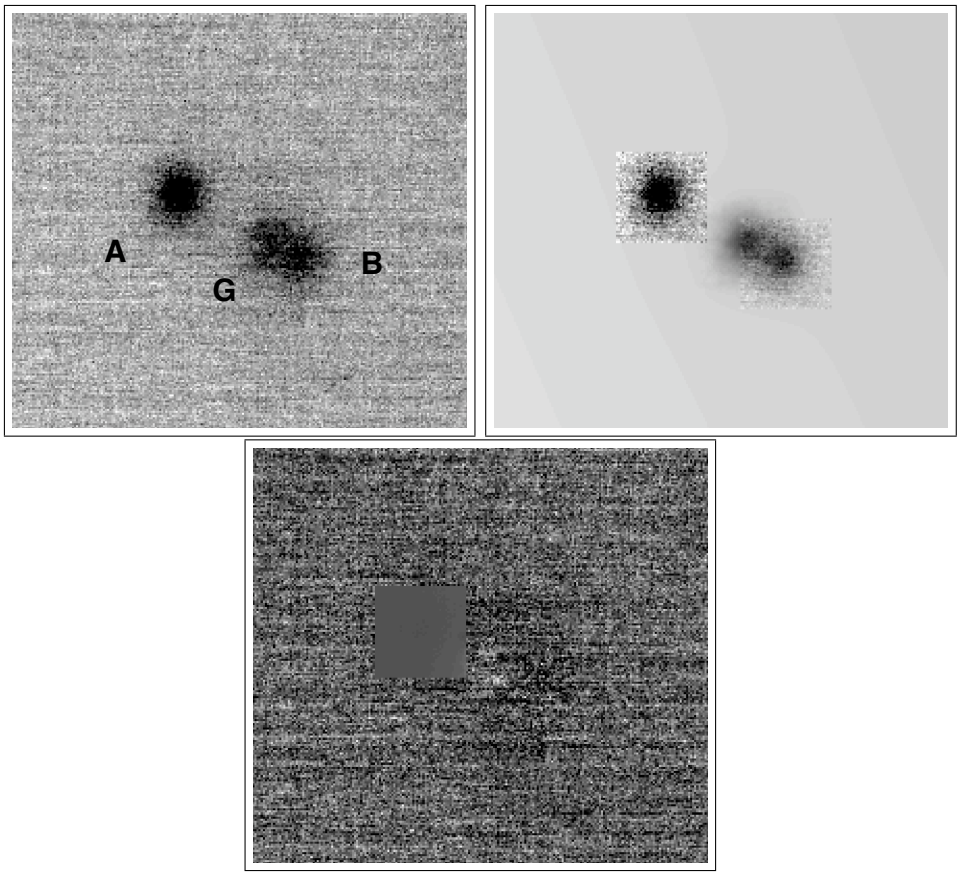

Figure 4.5 A photometric model for the lens SDSS 0806+2006 produced by our `GALFIT` fitting of the K band data in Figure 4.4. The top left panel shows the original data with the lens galaxy and quasar images labeled as G, A, and B, respectively. Image A serves as the model PSF for the lens, with the model for the various components presented in the top right panel. The bottom panel shows the residuals after model subtraction, which follow a Gaussian distribution with mean zero and $\chi_r^2 = 1$.

Table 4.3. Existing Flux Ratio Measurements

Target	Filter	B/A	Ref.
Q0142-100	$HST\,F555W$	0.127 ± 0.009	1
	$HST\,F675W$	0.121 ± 0.004	1
	Johnson $R^{\dagger}$	0.146 ± 0.009	2
	$HST\,F814W$	0.146 ± 0.016	1
	$HST\,F160W$	0.121 ± 0.004	1
SDSS 0246-0825	Sloan u	0.319 ± 0.033	3
	Sloan g	0.310 ± 0.010	3
	$HST\,F555W$	0.483 ± 0.010	4
	Sloan r	0.328 ± 0.011	3
	Sloan i	0.340 ± 0.016	3
	$HST\,F160W$	0.291 ± 0.010	4
	Johnson H	0.302 ± 0.062	3
	Johnson K'	0.290	3
SDSS 0806+2006	Johnson V	0.581 ± 0.012	5
	Johnson R	0.673 ± 0.014	5
	Johnson I	0.581 ± 0.012	5
	$0.39 - 1.10\,\mu$m	0.7	5
	$0.45 - 0.87\,\mu$m	0.77	6
	C III & Mg II BEL	0.454	6
	Johnson H	0.649 ± 0.013	5
	Johnson H	0.474 ± 0.043	6
	Johnson K'	0.689	5
SBS 0909+523	$0.3 - 8$ keV	3.125 ± 0.030	7
	$HST\,F555W$	0.444 ± 0.077	1
	Sloan $r^{\dagger}$	0.647 ± 0.003	8
	$HST\,F814W$	0.705 ± 0.073	1
	$HST\,F160W$	0.887 ± 0.029	1

Table 4.3 (cont'd)

Target	Filter	B/A			Ref.
HE 2149-2745	Johnson B	0.236 ± 0.009			9
	$HST\,F555W$	0.209 ± 0.015			1
	Johnson V	0.236 ± 0.018			9
	Johnson R	0.231 ± 0.006			9
	$HST\,F814W$	0.238 ± 0.005			1
	$HST\,F160W$	0.238 ± 0.009			1
Target	Filter	A/C	B/C	D/C	Source
HE 0435-1223	$HST\,F555W$	1.84 ± 0.11	1.08 ± 0.09	0.95 ± 0.06	10
	KPNO $R^{\dagger}$	1.75 ± 0.10	1.00 ± 0.04	0.85 ± 0.05	10
	$HST\,F814W$	1.69 ± 0.04	1.01 ± 0.04	0.82 ± 0.03	10
	$HST\,F160W$	1.57 ± 0.05	1.00 ± 0.03	0.79 ± 0.03	10

†Reported flux ratio is based on ground-based monitoring. Uncertainty includes scatter due to intrinsic and/or microlensing variability.

Note. — The above is a non-comprehensive list of existing flux ratio measurements for our targets. References: 1− Lehár et al. (2000), 2− Koptelova et al. (2010), 3− Inada et al. (2005), 4− CASTLES (http://www.cfa.harvard.edu/castles/), 5− Inada et al. (2006), 6− Sluse et al. (2008), 7− Dai & Kochanek (2009), 8− Goicoechea et al. (2008), 9− Lopez et al. (1998) 10− Kochanek et al. (2006)

4.4 Results

In this section, we compare the K and L' flux ratio measurements for each lens in Table 4.2. Since L' emission comes (in part) from the extended torus surrounding the AGN, difference between the two near-infrared flux ratios may indicate the presence of substructure. In Section 4.5, we discuss and interpret the flux ratios in conjunction with existing measurements at different wavelengths.

- *Q0142-100* In K, the flux ratio $F_B/F_A = 0.128 \pm 0.002$ while in L', $F_B/F_A = 0.132 \pm 0.006$. Consistent within one sigma, we do not detect any significant difference amongst the two measurements.

- *SDSS 0246-0825* We detect a significant discrepancy between our near-infrared flux ratios. We find $F_B/F_A = 0.258 \pm 0.015, 0.331 \pm 0.016$ in the K and L' bands, respectively. This difference is inconsistent with simple measurement error at $> 99.9\%$ confidence.

- *HE 0435-1223* From our measurements, we consider the flux ratios of the four images with respect to the flux of image C. Though measurement errors are $< 10\%$, we find the flux ratios F_A/F_C and F_D/F_C are consistent between our NIR wavebands. For the flux ratio F_B/F_C, however, we find a large discrepancy of 1.271 ± 0.073 and 0.991 ± 0.074 between K and L'. These flux ratios are discrepant at $> 99\%$ confidence, indicating significant multi-wavelength variation in the lens.

- *SDSS 0806+2006* In the L' band, we do not significantly detect image B, with an upper limit of $F_B/F_A < 0.164$ (3σ). Interestingly, the K band value of the flux ratio is $F_B/F_A = 0.406 \pm 0.030$, far above the L' measurement. The difference between the two measurements is the strongest amongst our sample, indicating particularly intriguing wavelength behavior. In Section 4.5 we address this behavior in detail.

- *SBS 0909+523* We measure identical flux ratios of $F_B/F_A = 0.973$ in K and L' images of SBS0909+523. Nevertheless, this measurement is intriguing as a null detection, especially since the flux ratio is known to exhibit strong multi-wavelength variation (see Section 4.5).

- *HE 2149-2745* We find flux ratio of image B and image A to be $F_B/F_A = 0.280 \pm 0.006$ in K and $F_B/F_A = 0.240 \pm 0.009$ in L', indicating a significant difference between the two wavebands ($> 99.99\%$ CL).

4.5 Discussion

Interpretation of the image flux ratios of our targets is complicated by a number of factors. At optical wavelengths, lensed emission is probing the rest frame UV emission from the accretion disk of the AGN. Therefore, the size associated with the optical flux ratios is on micro-arcsecond scales and may significantly altered by lensing from the stellar distribution in the lens (e.g., Chartas et al. 2009). Furthermore, optical flux ratios can also be altered by differential extinction from dust in the lens galaxy (e.g., Elíasdóttir et al. 2006).

At near-infrared (NIR) wavelengths, the size of accretion disk is larger and less sensitive to stellar microlensing. Depending on the source redshift, L' flux ratios can correspond to rest frame NIR wavelengths around $1-2\,\mu$m, where the luminosity of the accretion disk is falling rapidly and the emission from the innermost regions of the dusty torus is quickly rising (Rowan-Robinson 1995; Nenkova et al. 2008). Thus, the source size of our L' observations may be small and point-like ($0.01-0.1\,$pc, $10^{-6\prime\prime}$–$10^{-5\prime\prime}$) or more extended ($0.5-5\,$pc, $10^{-4\prime\prime}$–$10^{-3\prime\prime}$) depending on the relative contribution of of the accretion disk and dusty torus. If dark matter substructure affecting the lensed images, flux ratios can vary strongly as a function of source size (Dobler & Keeton 2006). As a result, L' flux ratios *may* exhibit the signature of substructure if the source is indeed much more extended than at lower wavelengths.

One final complication which affects flux ratios across all wavelengths is the intrinsic variation of the source. Since lensed light rays travel different paths, variations in source flux arrive at different times. Thus for single epoch measurements, the images reflect a snapshot of the source plane, seen at different times. Photometric monitoring can be used quantify the variations, but has only been done for a limited number of lenses (Eigenbrod et al. 2006; Kochanek et al. 2006; Vuissoz et al. 2007, 2008; Koptelova et al. 2008; Goicoechea et al. 2008). For our sample, this includes Q0142-100, HE 0435-1223,

SBS 0909+523, and HE 2149-2745.

Clearly, disentangling the variety of effects which can alter flux ratios across wavelengths is difficult. Nevertheless, differences amongst multi-wavelength flux ratios provide tantalizing evidence for interesting effects which warrant further study. Below, we discuss the multi-wavelength properties of each lens by comparing published flux ratios in the literature (Table 4.3) and our new NIR data (Table 4.2).

- *Q0142-100* We detect no significant difference between our K and L' flux ratios. With a source redshift of $z = 2.719$, the L' flux ratios correspond to rest frame $1.0\,\mu$m emission where the accretion disk luminosity is still quite strong relative to extended torus emission, and is likely the dominant source of flux (e.g., Hönig et al. 2008). Thus it is little surprise that we see no difference between our K and L' flux ratios since, even if substructure were present, the change in the size of the emitting region is minimal (on subhalo scales). At optical wavelengths, monitoring by Koptelova et al. (2010) gives a R band flux ratio of 0.146, quite different from our NIR value of 0.128. The amplitude of the flux ratio variation during their season is found to be 0.009. However, single epoch HST measurements have found $F_B/F_A = 0.127, 0.121, 0.121$ in $F555W$, $F657W$, $F160W$, indicating that variability is likely much larger for the system. We therefore attribute the difference between R and K/L' flux ratios to instrinsic/microlensing variation.

- *SDSS 0246-0825* Comparing observations in similar filters taken at various epochs, it is clear intrinsic and/or microlensing variability plays a role in the lens flux ratios (see Table 4.3, and references therein). For instance, SDSS i and $HST\,F814W$ measurements taken 20 months apart found $F_B/F_A = 0.340 \pm 0.016$ and 0.247 ± 0.017 respectively. Taken at the same observing periods, SDSS g and $HST\,F555W$ gave $F_B/F_A = 0.310 \pm 0.010$ and 0.483 ± 0.010. In the NIR, we find the K and L' flux ratios disagree at the $\sim 3\sigma$ level. It is unclear whether the same variation

responsible for discrepancies at lower wavelengths is the origin of NIR differences, or whether millilensing by substructure might be at play. Future observations in K and L', or in the mid-IR may help to discriminate between the two possibilities.

- *HE 0435-1223* We detect a significant difference between the K and L' flux ratio of images B and C. In Chapter 5, we discuss this interesting detection in detail.

- *SDSS 0806+2006* In Section 4.4 we find a strong difference between the image flux ratios in K and L' wavebands. The origin of this discrepancy is unclear. The VLT optical spectra of Sluse et al. (2008) show that chromatic microlensing and dust extinction are not significant in the lens. One possibility is that intrinsic variation of the AGN is responsible. At a redshift of $z = 1.54$, however, the L' emission comes from rest frame $1.5\,\mu$m. Therefore, if intrinsic variation is responsible, an explanation of how the source brightness can vary by a factor ~ 2 must be made. As an alternative, if the L' band source is indeed much larger than at the K band, millilensing may account for the difference. Either way the detection is exciting, and should be investigated further with modeling and follow up observations.

- *SBS 0909+523* For SBS0909+523, strong differential dust extinction is known to affect the flux ratio (Motta et al. 2002). Even though we see no difference between our K and L' flux ratios, our observations may be useful in constraining dust models in the system since they represent the longest wavelengths yet measured. In fact, the flux ratio at $F160W$ is $F_B/F_A = 0.887 \pm 0.02$ and at K is $F_B/F_A = 0.973 \pm 0.028$, indicating differential dust extinction is still significant in the rest frame near-infrared. In terms of dark matter substructure, the agreement of our K and L' flux ratios is interesting as a null detection. With rest frame wavelengths of $0.9\,\mu$m (in K) and 1.6μm (in L'), there should be a significant difference in the size of the two emitting regions. If the source sizes are indeed very different, this implies there is no substructure near the images with the right sized Einstein

radius to produce a chromatic effect.

- *HE 2149-2745* Across all wavelengths in Tables 4.3 and 4.2 ($\sim 0.5 - 3.8\,\mu$m), the flux ratio measurements for HE 2149-2745 are in good agreement. The exception to this is our K band flux ratio of 0.280. However, the system is known to vary with recorded flux ratio fluctuations as large as 0.03 (Burud et al. 2002). Given the source redshift ($z = 2.033$) and the known variability, the anomalously high value of our K band measurement is likely due to intrinsic and/or microlensing variability.

4.6 Conclusions

Multi-wavelength observations of lensed flux ratios are a powerful probe of quasar variability, microlensing, dust, and *potentially* dark matter substructure. For the latter, chromatic effects are present if the size of the emitting region varies strongly as a function of wavelength. We have attempted to search for these effects with new K and L' band images of six gravitational lenses. By selecting targets with source redshifts < 2.78, our L' images measure emission from rest frame wavelengths $> 1.0\mu$m. Some of this flux should originate from the extended torus of gas surrounding the central accretion disk, possibly providing the conditions for chromatic millilensing.

Of the six lenses, we detect strong differences between our K and L' images in HE 0435-1223 and SDSS 0806+2006. In Chapter 5, we discuss HE 0435-1223 in detail. For SDSS 0806+2006, our upper limit on the L' flux ratio is interesting and will be useful for understanding the origin of the long-wavelength suppression of image B. In particular, future lens modeling should be able to test whether substructure is capable of producing the anomaly, or whether variability is at play.

For the remainder of the sample, variability in accretion or microlensing are likely responsible for any differences in our NIR observations. Multi-epoch observations of

the lenses SDSS 0246-0825 and HE 2149-2745 have shown fluctuations of $\sim 40\%$ and $\sim 15\%$ in the flux ratios, respectively. The possibility exists, however, that chromatic millilensing may still be the source of the NIR discrepancies. Future observations are needed to quantify the variability in fluxes beyond $1\,\mu$m (rest frame). Alternatively, mid-IR flux ratios would be useful in deconstructing chromatic effects, since emission would definitively originate from the extended torus. Unfortunately, even the brightest lenses currently require almost a full night on $8-10$m telescopes. The *James Webb Space Telescope*, therefore, will likely serve as a valuable tool in future millilensing studies.

Finally, for the lenses Q0142-100 and SBS0909+523, we detect no difference in the flux ratios between K and L' measurements. These results are interesting null detections, indicating no substructure is close enough to the images for chromatic millilensing to take place. Our observations of the lenses, therefore, may serve as useful data for statistical studies of substructure in the future.

Bibliography

Agol, E., Gogarten, S. M., Gorjian, V., & Kimball, A. 2009, ApJ, 697, 1010

Burud, I., Courbin, F., Magain, P., Lidman, C., Hutsemékers, D., Kneib, J., Hjorth, J., Brewer, J., Pompei, E., Germany, L., Pritchard, J., Jaunsen, A. O., Letawe, G., & Meylan, G. 2002, A&A, 383, 71

Chartas, G., Kochanek, C. S., Dai, X., Poindexter, S., & Garmire, G. 2009, ApJ, 693, 174

Chiba, M. 2002, ApJ, 565, 17

Chiba, M., Minezaki, T., Kashikawa, N., Kataza, H., & Inoue, K. T. 2005, ApJ, 627, 53

Dai, X., & Kochanek, C. S. 2009, ApJ, 692, 677

Dalal, N., & Kochanek, C. S. 2002, ApJ, 572, 25

Dobler, G., & Keeton, C. R. 2006, MNRAS, 365, 1243

Eigenbrod, A., Courbin, F., Dye, S., Meylan, G., Sluse, D., Vuissoz, C., & Magain, P. 2006, A&A, 451, 747

Elíasdóttir, Á., Hjorth, J., Toft, S., Burud, I., & Paraficz, D. 2006, ApJS, 166, 443

Glass, I. S. 1999, Handbook of Infrared Astronomy, First Edition, 1st edn. (Cambridge University Press)

Goicoechea, L. J., Shalyapin, V. N., Koptelova, E., Gil-Merino, R., Zheleznyak, A. P., & Ullán, A. 2008, New Astronomy, 13, 182

Hönig, S. F., Prieto, M. A., & Beckert, T. 2008, A&A, 485, 33

Inada, N., Burles, S., Gregg, M. D., Becker, R. H., Schechter, P. L., Eisenstein, D. J., Oguri, M., Castander, F. J., Hall, P. B., Johnston, D. E., Pindor, B., Richards, G. T., Schneider, D. P., White, R. L., Brinkmann, J., Szalay, A. S., & York, D. G. 2005, AJ, 130, 1967

Inada, N., Oguri, M., Becker, R. H., White, R. L., Gregg, M. D., Schechter, P. L., Kawano, Y., Kochanek, C. S., Richards, G. T., Schneider, D. P., Barentine, J. C., Brewington, H. J., Brinkmann, J., Harvanek, M., Kleinman, S. J., Krzesinski, J., Long, D., Neilsen, Jr., E. H., Nitta, A., Snedden, S. A., & York, D. G. 2006, AJ, 131, 1934

Kochanek, C. S., Morgan, N. D., Falco, E. E., McLeod, B. A., Winn, J. N., Dembicky, J., & Ketzeback, B. 2006, ApJ, 640, 47

Koptelova, E., Oknyanskij, V. L., Artamonov, B. P., & Burkhonov, O. 2010, MNRAS, 401, 2805

Koptelova, E. A., Artamonov, B. P., Bruevich, V. V., Burkhonov, O. A., & Sergeev, A. V. 2008, Astronomy Reports, 52, 270

Lehár, J., Falco, E. E., Kochanek, C. S., McLeod, B. A., Muñoz, J. A., Impey, C. D., Rix, H., Keeton, C. R., & Peng, C. Y. 2000, ApJ, 536, 584

Lopez, S., Wucknitz, O., & Wisotzki, L. 1998, A&A, 339, L13

Macciò, A. V., Kang, X., Fontanot, F., Somerville, R. S., Koposov, S., & Monaco, P. 2010, MNRAS, 402, 1995

MacLeod, C. L., Kochanek, C. S., & Agol, E. 2009, ApJ, 699, 1578

Madau, P., Diemand, J., & Kuhlen, M. 2008, ApJ, 679, 1260

Metcalf, R. B., & Madau, P. 2001, ApJ, 563, 9

Minezaki, T., Chiba, M., Kashikawa, N., Inoue, K. T., & Kataza, H. 2009, ApJ, 697, 610

Minezaki, T., Yoshii, Y., Kobayashi, Y., Enya, K., Suganuma, M., Tomita, H., Aoki, T., & Peterson, B. A. 2004, ApJ, 600, L35

More, A., McKean, J. P., More, S., Porcas, R. W., Koopmans, L. V. E., & Garrett, M. A. 2009, MNRAS, 394, 174

Morgan, C. W., Kochanek, C. S., Morgan, N. D., & Falco, E. E. 2010, ApJ, 712, 1129

Motta, V., Mediavilla, E., Muñoz, J. A., Falco, E., Kochanek, C. S., Arribas, S., García-Lorenzo, B., Oscoz, A., & Serra-Ricart, M. 2002, ApJ, 574, 719

Nenkova, M., Sirocky, M. M., Nikutta, R., Ivezić, Ž., & Elitzur, M. 2008, ApJ, 685, 160

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H. 2002, AJ, 124, 266

Pooley, D., Blackburne, J. A., Rappaport, S., & Schechter, P. L. 2007, ApJ, 661, 19

Rowan-Robinson, M. 1995, MNRAS, 272, 737

Sluse, D., Courbin, F., Eigenbrod, A., & Meylan, G. 2008, A&A, 492, L39

Stern, D., Eisenhardt, P., Gorjian, V., Kochanek, C. S., Caldwell, N., Eisenstein, D., Brodwin, M., Brown, M. J. I., Cool, R., Dey, A., Green, P., Jannuzi, B. T., Murray, S. S., Pahre, M. A., & Willner, S. P. 2005, ApJ, 631, 163

Vuissoz, C., Courbin, F., Sluse, D., Meylan, G., Chantry, V., Eulaers, E., Morgan, C., Eyler, M. E., Kochanek, C. S., Coles, J., Saha, P., Magain, P., & Falco, E. E. 2008, A&A, 488, 481

Vuissoz, C., Courbin, F., Sluse, D., Meylan, G., Ibrahimov, M., Asfandiyarov, I., Stoops, E., Eigenbrod, A., Le Guillou, L., van Winckel, H., & Magain, P. 2007, A&A, 464, 845

Wyithe, J. S. B., Webster, R. L., Turner, E. L., & Mortlock, D. J. 2000, MNRAS, 315, 62

Chapter 5

Substructure in the lens HE 0435-1223

This chapter is based on work in preparation for
submission to the Astrophysical Journal:
Fadely, R. & Keeton, C. R., 2010, ApJ in prep.

We investigate the properties of dark matter substructure in the gravitational lens HE 0435-1223 ($z_l = 0.455$) via its effects on the positions and flux ratios of the multiply imaged background quasar ($z_s = 1.689$). We produce a series of lens models which add individual, truncated isothermal clumps near the lensed images and calculate the Bayesian evidence associated with each model. Through the image flux ratios, we decisively detect the presence of a clump near image A with a mass of $\log(\frac{M_A(<R_{Ein})}{M_\odot\, h_{70}^{-1}}) = 7.68^{+0.92}_{-0.85}$. In addition, our models indicate support for a second clump near image B, with $\log(\frac{M_B(<R_{Ein})}{M_\odot\, h_{70}^{-1}}) = 6.6^{+1.02}_{-1.52}$, although evidence for this clump is not decisive. Using Monte Carlo simulations of substructure populations, we find the mass fraction of substructure at the Einstein radius to be $f_{sub} \gtrsim 0.00092$, assuming a substructure mass function with slope $\alpha = -1.9$ with upper and lower mass thresholds $m_{\rm high} = 10^{10} M_\odot$ and $m_{\rm low} = 10^7 M_\odot$, respectively.

5.1 Background

A tension has arisen between cold dark matter (CDM) theories and modern astronomical observations. On mass scales below $10^{10} M_\odot$, N-body simulations of dark matter

have produced a consensus on the small scale structure of galaxies. In particular, CDM simulations firmly predict that small "subhalos" should should follow a power-law mass function, $dN/dm \propto m^{\alpha}$. Probing down to $\sim 10^4\,M_{\odot}$, recent simulations like *Via Lactea* and *Aquarius* find the slope of this function to be $\alpha \approx -1.9$ (Diemand et al. 2007; Springel et al. 2008). In total, the fractional amount of substructure, M_{sub}/M_{host}, is found to be $\sim 8\%$ in the *Via Lactea* simulation and around $\sim 11\%$ in the *Aquarius* simulation, due to the lower mass resolution of the latter.

Observationally, however, this picture of small scale structure in galaxies has yet to be confirmed. In the Local Group, observational censuses have sought to characterize the abundance, masses, and distribution of luminous satellites surround large galaxies like the Milky Way and M31 (e.g., Simon & Geha 2007; Kalirai et al. 2010). Before 2005, objects known to surround the Milky Way totaled only 11 in number, as only the most massive and luminous dwarfs galaxies had been found (Mateo 1998). After 2005, the Sloan Digital Sky Survey (York et al. 2000) opened up the possibility to detect extremely faint satellites (e.g., Willman et al. 2005; Irwin et al. 2007; Liu et al. 2008; Belokurov et al. 2009, 2010). These "ultra-faint" dwarfs, with absolute magnitudes as low as $M_V \sim -2$, have since more than tripled the number of Milky Way satellites to 35 (for a current list, see Wadepuhl & Springel 2010). Yet, in spite of this dramatic leap forward, this number falls severely short of the hundreds predicted by N-body simulations (Klypin et al. 1999; Moore et al. 1999).

A clear contributor to this disparity is the lack of a complete and thorough survey of the local volume. Indeed, while a huge improvement over previous observational data, the SDSS survey is limited in both sky coverage ($\sim 1/5$ of the sky) and in depth ($m_g < 22.2$). Accounting for the limitations of Sloan, a volumetrically complete survey will likely find many more missing satellites, possibly eliminating the problem altogether (Tollerud et al. 2008). However, depending on assumptions based on the currently known population, even such a survey may fail to find all the missing satellites. If

so, ant remaining discrepancy between theoretical predictions and observations may be attributed to the intrinsic luminosity of small mass dwarfs. Recently, it has been shown that satellites with total mass $< 10^7 M_\odot$ are likely to experience extremely suppressed, even quenched, star formation (e.g., Strigari et al. 2007; Macciò et al. 2010). In particular, effects from cosmic reionization, UV photo-evaporation, ram pressure or tidal stripping, supernovae, or cosmic rays may all play a role in hampering the conditions for star formation (Gnedin 2000; Scannapieco et al. 2001; Strigari et al. 2007; Madau et al. 2008; Mashchenko et al. 2008; Macciò et al. 2010; Penarrubia et al. 2010; Wadepuhl & Springel 2010). While the precise mechanisms are still debated, the plausibility of such arguments points to a large population of "dark dwarfs", whose luminosities are so low that they will elude traditional observation techniques.

Intriguingly, while local measurements of satellite galaxies seem to fall short of CDM predictions, observations in more distant galaxies exhibit the opposite conflict. Sensitive to mass alone, strong gravitational lensing provides a unique tool to study the population of satellites in cosmologically distant galaxies, regardless of their luminosities (e.g., Dalal & Kochanek 2002; Vegetti et al. 2009). On large angular scales ($\sim 1''$), the bulk of the lens properties of multiply imaged quasars are determined from the macroscopic mass distribution of the lens galaxy and its surrounding environment. Upon detailed inspection, however, perturbations to these large scale effects may be caused by substructure in the mass distribution (Mao & Schneider 1998; Metcalf & Madau 2001; Chiba 2002; Dalal & Kochanek 2002; Metcalf & Zhao 2002; Bradač et al. 2002; Koopmans et al. 2002; Chen et al. 2007; Keeton & Moustakas 2009; Keeton 2009). Thus, from the flux ratios, positions, and time delays of images, properties of the small scale structure of lens galaxies may be inferred.

Currently, some of the best constraints on dark matter substructure (outside of the Local Group) come from the analysis of "anomalous" flux ratios, detected in four-image gravitational lenses. In the radio, seven lenses have been show to violate "cusp"

and "fold" relations predicted for smooth mass models (Keeton et al. 2003, 2005). Performing a combined statistical analysis of the sample, Dalal & Kochanek (2002) found the mass fraction in substructure to be $0.006 < f_{sub} < 0.07$ (90% confidence) at the Einstein radii of the lenses. This lies in contradiction to CDM predictions, which predict $f_{sub} \sim 0.001 - 0.003$ at similar projected radii (Mao et al. 2004; Amara et al. 2006; Macciò et al. 2006; Macciò & Miranda 2006). In fact, Xu et al. (2010) recently showed that N-body simulations predict $f_{sub} \sim 0.002$ around typical Einstein radii, even when considering other sources of small scale structure (e.g., globular clusters, stellar streams).

Observational constraints, therefore, seem at odds. Around the Milky Way, tallies of satellites seem to indicate a dearth of substructure, while lensing points to a surplus. Confronting this on the lensing side, an expansion of sample size and methods (e.g., use of time delays, galaxy-galaxy lenses) is currently being sought. For quasar lenses, investigations in the mid-infrared (see Chapter 4) have begun to increase the number of lenses available for flux ratio studies (e.g., Chiba et al. 2005; MacLeod et al. 2009). Additionally, it has been recently shown that the brightness distribution of lensed galaxies can provide excellent constraints on substructure properties, detailing the mass and even the abundance of subhalos (Vegetti et al. 2009; Vegetti & Koopmans 2009b; Vegetti et al. 2010). Using Bayesian inversion techniques, Vegetti et al. (2009) measured $f_{sub} = 0.0215^{+0.0201}_{-0.0125}$ (assuming $\alpha = -1.9 \pm 0.1$, $m_{\rm high} = 10^{9.6} M_{\odot}$ and $m_{\rm low} = 10^{6.6} M_{\odot}$), consistent with, but higher than, predictions from simulations.

Here, we investigate the properties of the four image gravitational lens HE 0435-1223 (hereafter HE 0435), selected for its relatively bright ($F160W < 18.1$) and well separated ($2.4''$) images. Since its discovery (Wisotzki et al. 2002), HE 0435 has been extensively studied using ground- and space-based observations. From the ground, optical spectroscopy provided early evidence for stellar microlensing and for little to no differential dust extinction in the lens (Wisotzki et al. 2003). More recently, Kochanek

et al. (2006) have quantified the intrinsic and microlensing variability, providing an initial estimate of the time delays between images. Using *Hubble Space Telescope* (*HST*) imaging, Morgan et al. (2005) studied the environment of HE 0435, finding photometric evidence that the lens lies in a group of galaxies. Also identified by Morgan et al. was the nearby companion galaxy, G22. Assumed to be at the lens redshift ($z = 0.455$), G22 was shown to be important in reproducing the lensed images (Kochanek et al. 2006). More recently, Wong et al. (2010) studied the lens environment in detail, providing the first spectroscopic confirmation of the group members. Using the collective set of data available, as well as new near-infrared photometry, we examine the mass distribution of HE 0435, focusing on substructure in the lens. New in our analysis is the inclusion of both individual- and population-based simulations of substructure, which allow us to constrain both the masses of the clumps and connect them to the properties of the broader substructure distribution. In addition, we examine near infrared data from Gemini North, illuminating the multi-wavelength behavior of the lens. In this chapter we assume a flat cosmology with $\Omega_m = 0.27$ and $H_0 = 70.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, which is similar to the mean *WMAP*+BAO + H_0 values presented in Komatsu et al. (2010).

5.2 Infrared Observations and Data

We have obtained the first K (2.2 μm) and L' (3.78 μm) band images of the lens HE 0435-1223 using the NIRI instrument on the Gemini North 8m telescope. The data were taken in queue mode over two nights during the 2008B observing semester. During the first, August 31, 2008, the majority (84%) of the L' data were collected. The remaining L' exposures were taken on December 12, 2008, the same night as our K band observations. After data reduction was completed, L' data from the two nights were compared, and show no signs of variability. Seeing was better than average during the observation nights, with mean values of $0.53''$ and $0.44''$ in the K and L' images, respectively.

As discussed in Chapter 4, the observing strategy for our near-IR data was to dither the exposures over 4 positions on the sky, using 6″ shifts in right ascension and declination. Advantages of this strategy were twofold. First, dithering reduced the impact of hot or bad pixels and cosmic rays. Second, since HE 0435 has a relatively small angular size ($\sim 3''$), we were able to construct sky frames for each dithering position by combining observations at other pointings. This effectively doubled our observation efficiency by eliminating the need for nodding off source. The K band data consist of a single 24s exposure at each dither, with a total integration time of 288s. For the L' data, short exposures of 1.0s were used to prevent saturation of the detector by the bright foreground sky. At each dither, L' band observations were coadded in 30s blocks and total 3870s on-source. In addition, photometric standard stars were observed immediately adjacent to the observations on both nights.

We reduced the data using the standard NIRI IRAF packages provided by Gemini Observatory. Special care was taken in accounting for hot and bad pixels in our data. Specifically, we began with template bad pixel masks and used iterative sigma clipping routines in IRAF until the noise in the final reduced images reflected a Gaussian distribution with a mean of zero. One additional concern in our K band data was the presence of fixed pattern noise, common with low read noise observations using the NIRI instrument. We used the Python routine `nirinoise.py` (provided by Gemini) to correct the effect, with the corrected residual pattern comprising a variation of $< 1\%$ of the median sky value.

We present the reduced K and L' band images in Figure 5.1. To extract the observed flux of lensed images in our data, we use the `GALFIT` image fitting routine (Peng et al. 2002) to construct a simultaneous model for the lens galaxy and images (method discussed previously in Chapter 4). Due to the lack of bright stars in our small (22.5″) field of view, we opt to construct our PSF model using the lensed quasar images themselves. Specifically, we take the brightest quasar image (image A) and construct a model PSF

using a 80×80 pixel^2 or $1.8'' \times 1.8''$ region around the image. This PSF is then used to construct a model for the remaining three quasar images. For the lens galaxy we use a PSF-convolved Sérsic profile of varying index. For simplicity we fix half-light radius to $R_e = 0.86''$, corresponding to the value measured in F160W by Kochanek et al. (2006). We conduct our fitting procedure is as follows. First, we fit a model for the lensed images and galaxy. Subsequently, the lens galaxy model is subtracted from the data and a new PSF is measured from image A. Using this new PSF model (now with less contaminating flux), the original data is then re-fit. After subtraction, the image residuals and reduced χ^2_r values are inspected. We calculate χ^2_r values within a region defined by the PSF size for the quasar images, and within three regions which are 1, 2 , and 3 times the PSF size for the lens galaxy. If the model is unsatisfactory (due to persistent lens galaxy contamination of the PSF model), the PSF is then remeasured with the new lens galaxy model subtracted. This procedure is then iteratively repeated until there are little to no discernible residuals and $\chi^2_r \sim 1$ in the vicinity of the quasars and lens. As a check, our entire modeling procedure is repeated using the second brightest quasar image as the PSF source. We find our `GALFIT` modeling is important in the K band, where the lens galaxy provides a small but significant contribution to the flux near the quasar images. In the L' band, we do not detect the lens galaxy and our `GALFIT` modeling agrees with simple aperture photometry. We report the results of our fitting in Table 5.1.

5.3 Constraints

In addition to our new K and L' images, we consider a suite of existing data to constrain lens models of HE 0435. Astrometrically, we consider the HST derived centroids of the lensed images, the lens galaxy, and the companion galaxy G22 (Morgan et al. 2005). For the image flux ratios, we focus on the optical R band monitoring of Kochanek et al. (2006). As shown by Kochanek et al., HE 0435 exhibits significant intrinsic quasar

Table 5.1. Gemini Data

Filter	Image A	Image B	Image C	Image D	Lens Galaxy, G1
K	16.70 ± 0.01	17.10 ± 0.02	17.36 ± 0.05	17.68 ± 0.05	15.6 ± 0.6
L'	15.00 ± 0.02	15.59 ± 0.05	15.58 ± 0.05	15.81 ± 0.11	–

Note. — Apparent magnitudes in the Vega system. Not include above are the uncertainties associated with the zeropoint derived from standard stars. These uncertainties are ± 0.01 and ± 0.03 for K and L' magnitudes, respectively. For our purposes, uncertainties associated with the zeropoint are not important, since we only deal with relative photometry.

variability ($\sim 0.5 - 0.8$ mag.) as well as magnification variations due to microlensing by stars, both of which may affect image flux ratios. For simplicity, we assume the data of Kochanek et al. to be a representative sample of the variability of HE 0435. We then compute the mean and standard deviation of their data, treating the result as a measurement of the true R band flux ratios with uncertainties which reflect the variability due to intrinsic and microlensing variations. While the possibility exists that the R band measurements may not fairly represent variablity in HE 0435, recent work indicates the typical timescale for optical quasar variability is ~ 200 days (Kelly et al. 2009; Kozłowski et al. 2010; MacLeod et al. 2010). Since the Kochanek et al. observations comprise almost 2 years of monitoring, it is likely the variability is well sampled. In Figure 5.2, we plot the values of the R, K, and L' flux ratios as a function of wavelength and we report these data, along with other relevant measurements, in Table 5.2.

Given the growing amount of observational data for HE 0435, we must carefully consider which data to use as constraints for lens models. First, we adopt the positions of the lens, its companion G22, and the lensed images as in Table 5.2. For the image flux ratios, which constraints to impose is less clear. Amongst the options, there is good

Table 5.2. HE 0435 Constraints

Images					
	Position		R band flux	K band flux	L band flux
Image A	-1.165 ± 0.003	0.573 ± 0.003	1.751 ± 0.098	1.837 ± 0.086	1.706 ± 0.085
Image B	0.311 ± 0.004	1.126 ± 0.004	0.998 ± 0.037	1.271 ± 0.063	0.991 ± 0.065
Image C	1.302 ± 0.005	0.030 ± 0.005	$\equiv 1.0$	$\equiv 1.0$	$\equiv 1.0$
Image D	-0.226 ± 0.003	-1.041 ± 0.003	0.851 ± 0.049	0.745 ± 0.049	0.809 ± 0.090
Lens Galaxies					
	Position		$F555W$ (mag)	$F814W$ (mag)	$F160W$ (mag)
G	$\equiv 0.0 \pm 0.002$	$\equiv 0.0 \pm 0.002$	21.55 ± 0.13	18.85 ± 0.13	16.86 ± 0.04
G22	2.585 ± 0.005	3.637 ± 0.005	22.25 ± 0.04	21.26 ± 0.01	~ 18.8

Note. — R band flux data is reflects the average and standard deviation from the published data of Kochanek et al. (2006), and includes scatter from intrinsic and microlensing variability. The astrometry of the images and lens, and well as the magnitudes of the lens are from Kochanek et al.. The data for G22 is found in Morgan et al. (2005), with the exception of the $F160W$ magnitude which is estimated in Kochanek et al.

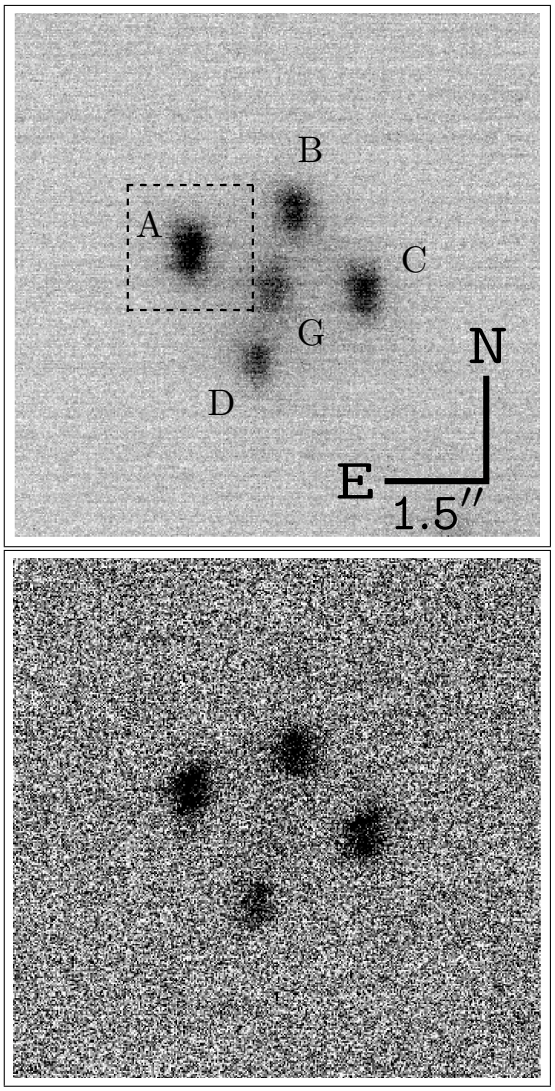

Figure 5.1 Near infrared images of the lens HE 0435-1223, obtained with the NIRI instrument at Gemini North. A K band image is presented in the top panel, with the lens galaxy clearly visible. On the bottom, we present the L' band image of the lens. The lensed images are labeled A through D, with the lens galaxy labeled G. The dashed box represents the PSF extraction region, used for our `GALFIT` photometric modeling.

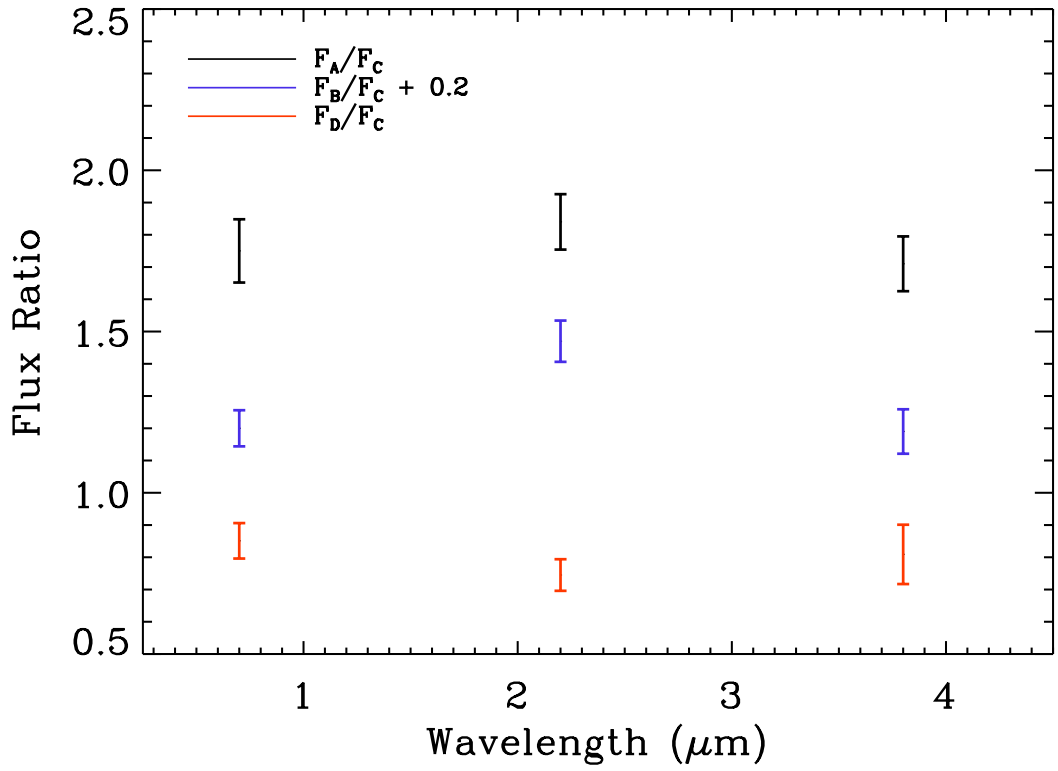

Figure 5.2 Flux ratios in R, K, and L' bands for HE 0435-1223. Note the flux ratio B/C has been shifted up by 0.2 for visual clarity. We find the R and L' flux ratios in good agreement. In K, the flux ratio of B/C is a factor 1.3 higher than corresponding R and L' measurements.

agreement between the optical R band flux ratios and our Gemini data. One exception is the value of the B/C flux ratio in our K band image, which is a factor ~ 1.3 higher than in R or L'. Potentially, all three wavebands are suitable to constrain our models.

At a source redshift of $z_s = 1.689$, R and K band images represent rest-frame UV, optical wavelengths which are dominated by thermal emission from the hot quasar accretion disk. With sizes of order 10^{15-17} cm (Morgan et al. 2010), such emission is effectively a point source for lensing masses $\gtrsim$ few $\times\, 100 M_{\odot}$, since the source size is much smaller than the corresponding Einstein radii ($R_{src} \ll R_{Ein}$). Therefore, our R and K band measurements should result in the same image magnifications. For image B, this is clearly not the case. This discrepancy might be attributed to the effects of dust extinction, variability, and/or microlensing. For the early type lens galaxy in HE 0435, dust is not a concern. Using optical integral field spectroscopy, Wisotzki et al. (2003) found no evidence of dust extinction both from the slopes of the rest-frame UV slopes and from the *identical* broad emission line profiles found in the spectra of the four images. We postulate that our K band flux ratios are likely affected by microlensing, which we further discuss in Section 5.7.3.

For emission in the L' band, the story is more complicated. At a rest-frame wavelength of 1.4μm, the quasar flux represents a combination of emission from the central accretion disk and the surrounding dust torus (Rowan-Robinson 1995; Nenkova et al. 2008). The exact contribution of each of these radiation sources is unclear, with the accretion disk likely contributing between $20 - 80\%$ of the luminosity (e.g., Wittkowski et al. 2004; Hönig et al. 2008). In the case of substructure lensing, extra light bending on small scales is significant when the source is much smaller than the Einstein radius of the clump $R_{src} < R_{Ein}$. If the source is larger, with a size comparable to the Einstein radius ($R_{src} \sim R_{Ein}$), magnification can be quite different from the small source case (see e.g., Dobler & Keeton 2006, also Chapter 4.1 here). For HE 0435, some of the L' emission comes from the hot accretion disk ($R_{src} < R_{Ein}$), while the remainder

originates from the more extended torus ($R_{src} \sim R_{Ein}$). In turn, the two components of the L' band source may experience very different magnifications due to the finite source effects of millilensing.

Given the complexity and uncertainty associated with the source of L' band flux ratios, and the possible microlensing present in K band flux ratios, we choose to constrain our lens models using the optical R band measurements. This has the advantage of knowing more precisely the effects of quasar and microlensing variability, and is simple to model as a point source. In Section 5.7.3 we revisit our new IR data, after analyzing the existing R band data in detail.

Finally, we note that we do not use the measured time delays of Kochanek et al. (2006) to constrain our lens models. With only 2 seasons of monitoring data, it is difficult to accurately disentangle the complicated variations from quasar and microlensing variability, and the uncertainties of the published values are quite likely underestimated. Indeed, Blackburne & Kochanek (2010) have recently reported that new time delay estimates are discrepant by $2-5\sigma$ from the Kochanek et al. values. Since there is currently no accurate time delay published for HE 0435, we do not constrain the delays in our models.

5.4 Analysis of HE 0435: Motivation, Strategy, and Goals

Using the collection of data available for HE 0435, we conduct initial modeling to examine the properties of the lens. Using a singular isothermal ellipsoid (SIE) mass distribution for the main lens and its companion (G22), a "flux ratio anomaly" between images A and C was identified. Seen in many previous works (e.g., Mao & Schneider 1998; Bradač et al. 2002), flux ratio anomalies arise in models which fail to reproduce the observed flux ratio, and allude to small scale structure in a lens. In Section 5.7.1 we discuss the identified flux ratio anomaly and show it is a general phenomenon for HE 0435, even for more complex lens models.

Motivated by this anomaly, we aim to study the properties of substructure in the lens HE 0435-1223. First, we seek to constrain the mass and positions of any clumps affecting the images and quantify the related effects on the macroscopic lens model. To accomplish this goal, we construct a highly flexible macroscopic lens model and successively add individual clumps of mass near the lensed images. For each model (Section 5.5), we conduct a full search of the parameter space and compute the Bayesian evidence, using the approach discussed in Section 5.6. As a final result, for each model we have an estimate of the clump masses (from parameter marginalization), as well as a ranking of models with different numbers of clumps (from the evidence).

For the anomaly, we find the observed A/C flux ratio is higher than predicted from smooth models. In order to resolve this discrepancy, therefore, we first consider substructure which might affect images A and C. In HE 0435, images A and C are both positive parity images. Under the influence of substructure, positive parity images experience an increase in magnification, while negative parity images (usually) experience a decrease in magnification (Schechter & Wambsganss 2002; Keeton 2003). Therefore, to explain the A/C ratio, we adopt a single clump near image A. We do not consider clumps near image C, since increasing its magnification further exacerbates the anomaly.

In addition to a clump near A, we examine the presence of substructure near the images B and D. While neither of these exhibit striking anomalies, the possibility exists that perturbations to these images are more subtle, affecting either the flux ratios or the positions. In summary, we consider four models with individual clumps near images A, A and B, A and D, and A, B, and D. For simplicity, we label these models A, AB, AD, and ABD, respectively. In Section 5.7.1 we discuss the results of these models in detail.

Following this analysis, we conduct Monte Carlo simulations of substructure populations in the lens (Section 5.7.2). Doing so, we randomly generate realizations of

the population by drawing from a mass function motivated by N-body simulations of dark matter (e.g., Diemand et al. 2007; Springel et al. 2008). For each simulation, we conduct many realizations, holding the fraction of mass in substructure, f_{sub}, fixed. We then vary the value of f_{sub} between simulations, and compute the Bayesian evidence. Our final result is an estimate of f_{sub}, located at the Einstein radius of HE 0435.

5.5 Model Selection

Using the publicly available `lensmodel` code (Keeton 2001), we seek to construct realistic lens models which reproduce the observed quasar image positions and flux ratios. Following the work of Kochanek et al. (2006), we assume a "minimal" lens model which consists of a single power law mass model for the main lens (G1) of the form

$$\kappa(\xi) = \frac{1}{2}\frac{b_{\rm G1}^{2-\beta}}{(s^2+\xi^2)^{1-\beta/2}} \tag{5.1}$$

where s is the core radius, $\xi = \sqrt{x^2 + y^2/q^2}$ is the ellipse coordinate (in the major axis frame), and q is the projected axis ratio. For flexibility, we allow all model parameters of the main lens to vary. One particularly important parameter we vary is the index of the power law, β. For HE 0435, Kochanek et al. found a projected density profile whose inner slope is rising ($\beta > 1.0$ here), and inconsistent with steeper profiles. Thus, varying β in our models is important if we are to account for all possible distributions of the lens mass.

In addition, our minimal model includes the effects of tidal shear. Here, we explicitly model the local effects of the lens companion galaxy, G22, as a SIE profile. Using an SIS profile for G22, Kochanek et al. found a best-fit Einstein radius of $0.22''$, which should provide negligible surface density at a projected distance of $R_{G22} \sim 4.4''$. Since G22 manifests itself mostly in local shear and higher order terms, its radial profile should not effect significantly affect our lens models. We therefore assume our SIE model is sufficient. Lastly, we also consider the effects of an additional external shear component,

which may be provided by the surrounding group of galaxies (Morgan et al. 2005; Wong et al. 2010).

When moving beyond our minimal lens model, we consider possible effects due to small scale substructure in the lens. We do so by adding single clumps near the lensed images, modeling the clumps as a spherical pseudo-Jaffe profile:

$$\kappa(r) = \frac{b_{\rm clump}}{2}\left[\frac{1}{r} - \frac{1}{\sqrt{a^2+r^2}}\right]. \tag{5.2}$$

We select this clump profile since it includes the effects of tidal truncation and is computationally easy to model. Furthermore, pseudo-Jaffe profiles have been used in previous studies (e.g., Dalal & Kochanek 2002; Vegetti et al. 2010), and will therefore facilitate interpretation of our results. Following Dalal & Kochanek (2002), we set the truncation radius $a = \sqrt{\langle b_{\rm G1}\rangle\, b_{\rm clump,max}}$, where $\langle b_{\rm G1}\rangle$ is the average mass normalization of G1 and $b_{\rm clump,max}$ is the maximum Einstein radius of the clump. For HE 0435, we consider a fixed truncation of $a = 0.367''$.

In summary, we use a minimal, smoothly distributed "macro" model as the basis of our lens models. To this model, we add sources of additional complexity in the form of small scale structure (clumps), and consider each addition (one at a time) in detail. In Table 5.3 we summarize our various model parameters.

5.6 Methodology

For our array of lens models, we aim to compute the posterior probability distribution

$$P(\theta|d,M) = \frac{P(d|\theta,M)P(\theta|M)}{P(d|M)} \tag{5.3}$$

where d is the data which constrain the parameters θ for model M. We calculate the likelihood, $\mathcal{L} = P(d|\theta,M)$, by using the χ^2 goodness-of-fit: $\mathcal{L} \propto e^{-\chi^2/2}$. Since we are only concerned with the *relative* posterior probablity, we ignore the proportionality constant in our analysis and set $\mathcal{L} = e^{-\chi^2/2}$. We take the prior distribution, $P(\theta|M)$, to be uniform for the parameters listed in Table 5.3.

Table 5.3. Model parameters and priors

Parameter	MCMC prior	Nested Sampling prior
Minimal, smooth model		
$\log(b_{\rm G1}/'')$	$-\infty$: ∞	0.04 : 0.13
$x_{\rm G1}$	$-\infty$: ∞	$-0.003''$: $0.003''$
$y_{\rm G1}$	$-\infty$: ∞	$-0.003''$: $0.003''$
$e_{\rm c,G1}$	-1.0 : 1.0	-0.50 : 0.50
$e_{\rm s,G1}$	-1.0 : 1.0	-0.50 : 0.50
$\gamma_{\rm c}$	-1.0 : 1.0	-0.04 : 0.06
$\gamma_{\rm s}$	-1.0 : 1.0	-0.03 : 0.03
$s_{\rm G1}$	$0.00''$: ∞	$0.00''$: $0.02''$
$\beta_{\rm G1}$	$-\infty$: ∞	0.95 : 1.60
$\log(b_{\rm G22}/'')$	-1.7 : ∞	-1.70 : -0.22
$x_{\rm G22}$	$-\infty$: ∞	$2.572''$: $2.597''$
$y_{\rm G22}$	$-\infty$: ∞	$3.625''$: $3.650''$
$e_{\rm c,G22}$	-1.0 : 1.0	-0.70 : 0.70
$e_{\rm s,G22}$	-1.0 : 1.0	-0.70 : 0.70
Clump models		
$\log(b_{\rm A}/'')$	$-\infty$: ∞	-4.00 : -1.00
$x_{\rm A}$	$-\infty$: ∞	$-1.40''$: $-0.70''$
$y_{\rm A}$	$-\infty$: ∞	$0.40''$: $0.80''$
$\log(b_{\rm B}/'')$	$-\infty$: ∞	-4.00 : -1.00
$x_{\rm B}$	$-\infty$: ∞	$0.20''$: $0.80''$
$y_{\rm B}$	$-\infty$: ∞	$1.00''$: $1.50''$
$\log(b_{\rm D}/'')$	$-\infty$: ∞	-4.00 : -1.00
$x_{\rm D}$	$-\infty$: ∞	$-0.40''$: $0.50''$
$y_{\rm D}$	$-\infty$: ∞	$1.00''$: $1.50''$
Multipole models		
A_3	-1.0 : 1.0	-0.05 : 0.05
A_4	-1.0 : 1.0	-0.05 : 0.05

Note. — Parameter priors are uniform within the intervals listed above.

The normalization of the posterior, $P(d|M)$, is the marginal likelihood of the model, usually referred to as the Bayesian Evidence. In astrophysical studies, it often the case that only one model is being tested. In such cases the evidence may simply be ignored, since the normalization of the posterior is not needed to calculate confidence intervals of marginalized parameters. Ignoring the evidence, the posterior is simply the combination of the likelihood and prior distributions, and can easily be calculated with Monte Carlo Markov Chains (MCMC).

However, this approximation of the posterior is no longer adequate once comparisons between models are to be made. Here, the Bayesian Evidence is the key quantity which allows one to distinguish between various models, since it quantifies the overall probability of a particular model. Since we wish to compare models of varying complexity, we use the Nested Sampling algorithm (Skilling 2004) to calculate both marginalized parameter ranges and the Bayesian Evidence. As a computational tool, Nested Sampling has been used in a variety of astrophysical studies (e.g., Mukherjee et al. 2006; Humphrey et al. 2009), including gravitational lensing (Vegetti & Koopmans 2009a; Barnabè et al. 2009).

The heart of the (basic) Nested Sampling algorithm is to execute many random draws from the parameter space defined by the prior volume, searching for the next highest point in likelihood. In practice, this means that oddly shaped likelihood surfaces (especially in high parameter dimensions) can be quite computationally demanding to sample. This problem is further exacerbated when very large parameter volumes are searched. We alleviate much of the computational demand by adopting a two step approach to our sampling. First, we execute a MCMC sampling of the posterior using uniform priors defined in Table 5.3. Our MCMC methods, including the algorithm, techniques, and convergence criteria are described in Section 3.4 of Fadely et al. (2010). Using the posterior derived from our MCMC runs, we then construct a reduced set of priors which encompass the 99.999% CL parameter ranges (Table 5.3). The reduced

Table 5.4. Jeffreys' Scale

Bayes Factor	Significance
$1 - 10^{1/2}$	Barely worth mentioning
$10^{1/2} - 10$	Substantial
$10 - 10^{3/2}$	Strong
$10^{3/2} - 100$	Very strong
> 100	Decisive

set of priors is then adopted for subsequent Nested Sampling runs, thus reducing the time spent sampling regions of extremely low likelihood ($\chi^2 > 10^6$).

Once computed, we use the Bayesian Evidence to perform model comparison by computing the ratio:

$$\frac{P(M|d)}{P(M'|d)} = \frac{P(M)}{P(M')}\frac{P(d|M)}{P(d|M')} \tag{5.4}$$

which provides a means of comparing two competing models given the data (MacKay 2003; Gelman et al. 2003). Since we assume equal prior probabilities between models ($\frac{P(M)}{P(M')} = 1$), this ratio is simply the ratio of the evidences for models M and M'. This ratio, called the Bayes Factor, provides a meaningful mechanism for model comparison. Nevertheless, the significance of Bayes factors are not clear cut and various scales are often employed to facilitate their interpretation. The most common scale is the Jeffreys' scale (Jeffreys 1961), which grades the Bayes Factors as shown in Table 5.4. In this work, we use the Jeffreys' scale as a guideline for judging between our models.

5.7 Results and Discussion

5.7.1 Individual Clump Models

As a null hypothesis, we test the ability of our minimal lens model to reproduce the observed *HST* image positions and flux ratios from ground-based optical monitoring.

The model, described in Section 5.5 and Table 5.3, accounts for the mass distribution of the large, "macro" scale structures and does not include small scale perturbations due to substructure. Following the procedure of Section 5.6 we have derived the parameter ranges and model evidence, and present the results in Table 5.5. We find the model does poorly in reconstructing the observations, with a best fit $\chi^2 = 24.6$ for $N_{dof} = -1$. Given that the model is formally underconstrained, it is surprising that $\chi^2 \neq 0$, indicating the model lacks some key freedom. We find the primary failure of the model is in accounting for the flux ratio of A/C. In Figure 5.3 we show the model distributions for the three flux ratios in HE 0435. We find the smooth model is unable to account for the flux ratio of A/C at high confidence.

The strong discrepancy between the observed A/C flux ratio and that inferred for our minimal model motivates us to consider additional complexity in the lens in the form of small scale structure. Doing so, we add a single clump (whose profile is described in Section 5.5) one at a time near various images and calculate the evidence for the new model.

We first add a single clump near image A to our macro model. As shown in Figure 5.4 and Table 5.5, we find the addition of the clump does much to alleviate the discrepancy between the flux ratio predicted by our smooth model and that of the observations, resulting in a model value of $A/C = 1.72^{+0.22}_{-0.29}$. With the addition of a small scale clump near image A, the model is now able to reproduce the data perfectly with a corresponding $\chi^2 = 0$. It is no surprise the data can be perfectly reconstructed, given the model is underconstrained with $N_{dof} = -4$.

In Figure 5.5, we plot the position of the clump near image A for three different upper limits on the clump masses of $M_A(< R_{Ein}) < 10^6, 10^7, 10^8\, M_\odot$. We find that the position of the clump is degenerate with its mass, similar to previous studies (e.g., Dalal & Kochanek 2002; Keeton 2009). This degeneracy is no surprise since, heuristically, a less massive clump may be placed closer to the image and still produce the same

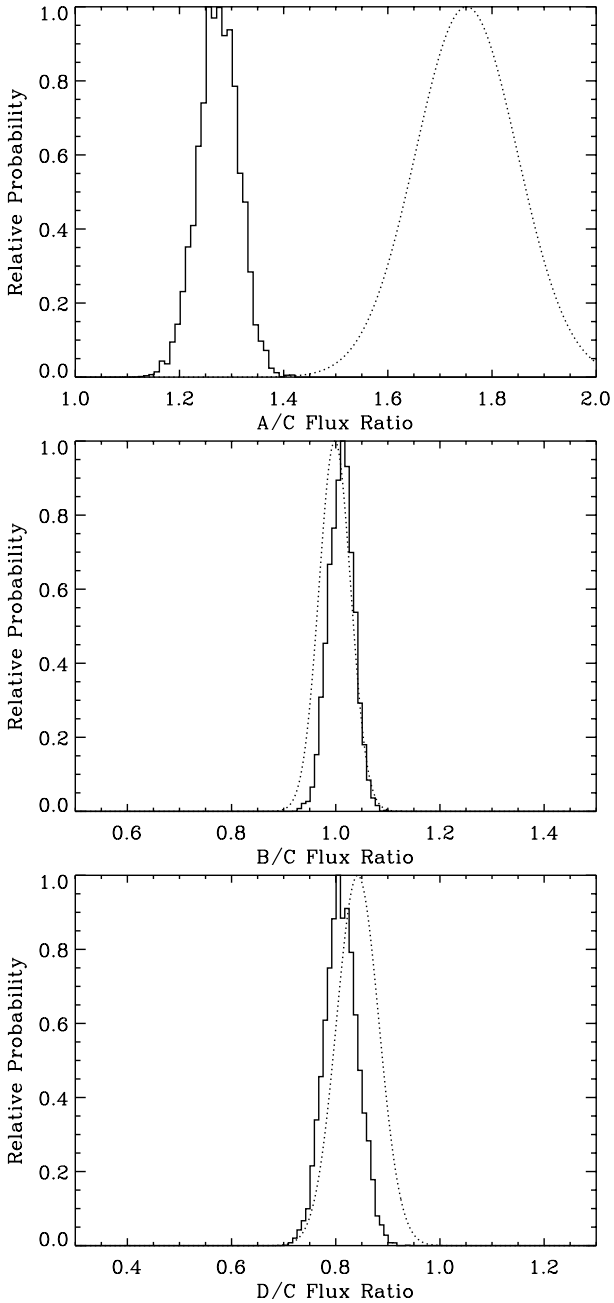

Figure 5.3 Plotted are the marginalized probability distributions for the three image flux ratios (solid lines) inferred from our minimal (smooth) mass model. Also plotted are the values of the observed distributions (dotted lines). Interestingly, our minimal model fails to explain the flux ratio A/C. The flux ratios for B/C and D/C, are clearly in good agreement with the observed R band ratios.

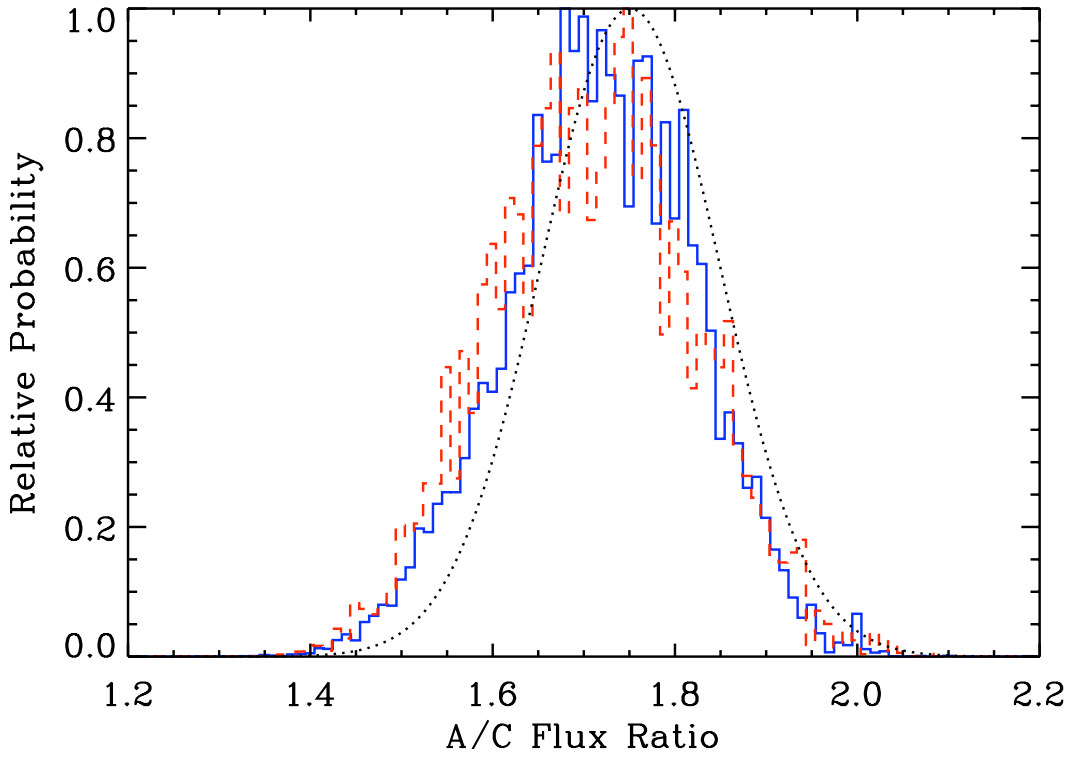

Figure 5.4 Same as the top left panel of Figure 5.3, but for our model with an additional clump near image A (solid, blue) and for our model with three clumps near images A, B, and D (dashed, red). Adding a clump near image A clearly brings the models and data into agreement for the flux ratio A/C. The distribution of flux values is negligibly changed when clumps near other images are added, indicating constraints on clump A are fairly independent from the presence of clumps B and D.

magnification. From flux constraints alone the degeneracy is, in principle, unbounded. Formally, the magnification perturbation due to a nearby clump scales as $\delta\mu \propto R_{Ein}^2/d^2$, where d is the distance from the clump to the image. Thus a star, placed very close to the image, can produce the same magnification as a massive galaxy placed further away, so long as R_{Ein}^2/d^2 is constant. Adding position constraints can break this degeneracy. If the clump causes astrometric perturbations to the image, the deflection scales with $\delta\alpha \propto R_{Ein}^2/d$. Thus, if both the position and flux are affected by the clump, the two different scalings allow constraints on the mass to be obtained. On the low mass end, a very small mass clump is simply unable to affect the image position, regardless of location. On the high mass side, the clump will disturb the image position too much, and potentially affect the other lensed images. In HE 0435 we are able to find bounds on the mass of clump A, with $\log(\frac{M_A(<R_{Ein})}{M_\odot\, h_{70}^{-1}}) = 7.68^{+0.92}_{-0.85}$. We conclude, therefore, that the clump is constrained by the combination of both flux and astrometric data.

We consider the possibility that more than one clump might be affecting the images in HE 0435. Specifically, we consider models with clumps near images AB, AD, and ABD. Figure 5.4 shows the effect of the higher number of clumps on the A/C flux ratio. We find that adding additional substructure near images B and D does not significantly alter the inferred A/C value. Since the flux ratio of A/C provides the main constraint on clump A, the mass associated with the clump is also relatively unaffected (Table 5.5).

To assess the relative probabilities of the various models, we have computed the Bayesian evidence and report the values in Table 5.5. We find our models with clumps near images A, AB, AD, and ABD all have evidence values that are each at least three orders of magnitude greater than our minimal smooth model. Clearly, the data strongly prefer models with at least one clump near the lensed images. Examining the evidence in detail, we see the model with clumps near images A and B has the highest value, and is 0.63 dex larger than our single clump model. The model with clumps near images

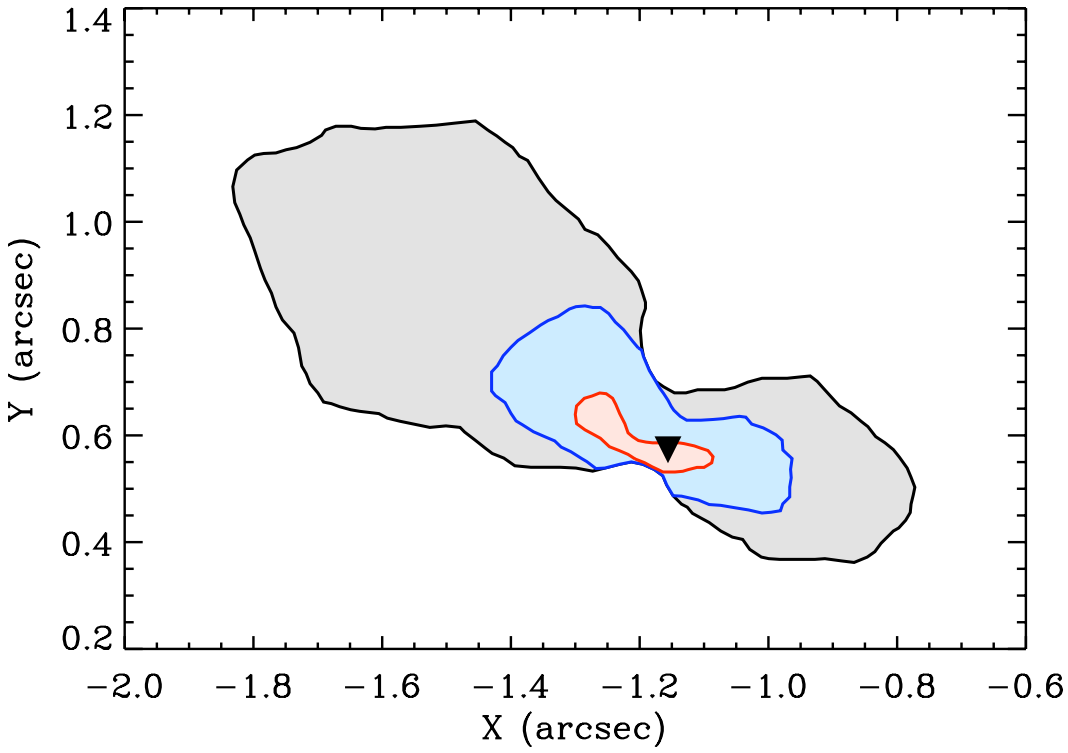

Figure 5.5 Plotted are the positions of the clump near image A, inferred after marginalizing over all of our model parameters. The black triangle indicates the observed position of image A. Red, blue, and black shaded regions indicate the positions of all clumps with masses less than $10^6, 10^7$, and $10^8\ M_\odot$, respectively. We find larger/smaller clumps lie farther/closer to the position of image A.

A and D exhibits a lower evidence than model AB, with an evidence consistent with our single clump model. Lastly, our three clump model has evidence that is equivalent (within 1σ) to our best model, indicating firmly that models with a clump near image D are not strongly favored. In comparison to model AB, the addition of the clump near image D must add parameters (clump D mass and position) which do not improve model ABD's ability to reproduce the global constraints. Thus, the evidence seen for model ABD is an example of Occam's Razor in action.

Robustly, the data prefer a model with a single clump near image A. This clump has a mass within its Einstein radius of $M_A(< R_{\rm Ein}) = 10^{7.7^{+0.92}_{-0.85}}\ M_\odot$. According to Jeffreys' Scale (Table 5.4), there is "substantial" evidence pointing to a clump near image B,

with a mass $M_B(< R_{\rm Ein}) = 10^{6.6^{+1.02}_{-1.52}}\,M_\odot$. We note, however, that this evidence is only 0.63 dex higher than for model A, and that both models (A and AB) carry evidence uncertainties of 0.12 dex. Therefore, while the evidence for clump B is intriguing, it is far from decisive.

With the addition of substructure to our minimal model, it is interesting to study how the macroscopic distribution changes. In Figure 5.6, we plot the joint posterior probability distribution of a few key parameters, before and after adding a single clump near image A. In general, we find the inferred distributions are broadened by the addition of clump A (e.g., Figure 5.6, top). This result is not surprising, given the increased flexibility afforded by the clump. For some parameters, the posterior distribution is not simply broadened but develops significant structure, indicative of degeneracies in the 14 dimensional space. For instance, the bottom panel of shows the distribution of the mass normalizations for the main lens and the companion G22. Before the addition of clump A, the distribution of the parameters is relatively compact, resembling that of a 2D elliptical gaussian. After clump A is added, the distribution is clearly much more distorted, with a clear degeneracy developing between the parameters. Nevertheless, even under such distortions the median values of the distributions are not significantly altered, typically shifting within the 68% confidence interval of the no-clump model. One particularly interesting parameter of our macro model is the slope of the density profile of the main lens, β. Using the surface brightness of the quasar host galaxy, the (estimated) time delays, and the image positions, Kochanek et al. (2006) found the slope to be shallower than isothermal, corresponding to $\beta > 1.0$ in our models. We find $\beta = 1.19^{+0.13}_{-0.13}$ and $1.19^{+0.17}_{-0.15}$ for our models with and without a clump near image A. Thus, we conclude that the density slope in HE 0435 corresponds to a rising rotation curve, regardless of the presence of substructure.

The mass constraints we find on substructure in HE 0435 are a first for quasar lenses. Previous work in the radio and mid-infrared have found many lens systems

Table 5.5 Model Results: Clump Parameters, and Evidence

Model	$\log(M_A)$	x_A	y_A	$\log(M_B)$	x_B	y_B	$\log(M_D)$	x_D	y_D	log(Evidence)
Minimal	–	–	–	–	–	–	–	–	–	-17.66 ± 0.11
+ A	$7.48^{+0.61}_{-0.70}$	$-1.05^{+0.13}_{-0.17}$	$0.51^{+0.07}_{-0.06}$	–	–	–	–	–	–	-13.83 ± 0.12
+ AB	$7.68^{+0.92}_{-0.85}$	$-1.13^{+0.08}_{-0.13}$	$0.56^{+0.05}_{-0.06}$	$6.56^{+1.02}_{-1.52}$	$0.15^{+0.24}_{-0.29}$	$1.04^{+0.12}_{-0.12}$	–	–	–	-13.20 ± 0.12
+ AD	$7.33^{+0.59}_{-0.63}$	$-1.13^{+0.08}_{-0.13}$	$0.56^{+0.05}_{-0.06}$	–	–	–	$5.80^{+1.56}_{-1.72}$	$0.15^{+0.19}_{-0.30}$	$-1.04^{+0.12}_{-0.13}$	-13.76 ± 0.13
+ ABD	$7.50^{+0.69}_{-0.69}$	$-1.07^{+0.09}_{-0.11}$	$0.55^{+0.05}_{-0.05}$	$6.15^{+1.33}_{-1.88}$	$0.45^{+0.14}_{-0.13}$	$1.31^{+0.10}_{-0.11}$	$5.87^{+1.52}_{-1.80}$	$0.13^{+0.11}_{-0.22}$	$-1.05^{+0.10}_{-0.11}$	-13.31 ± 0.13

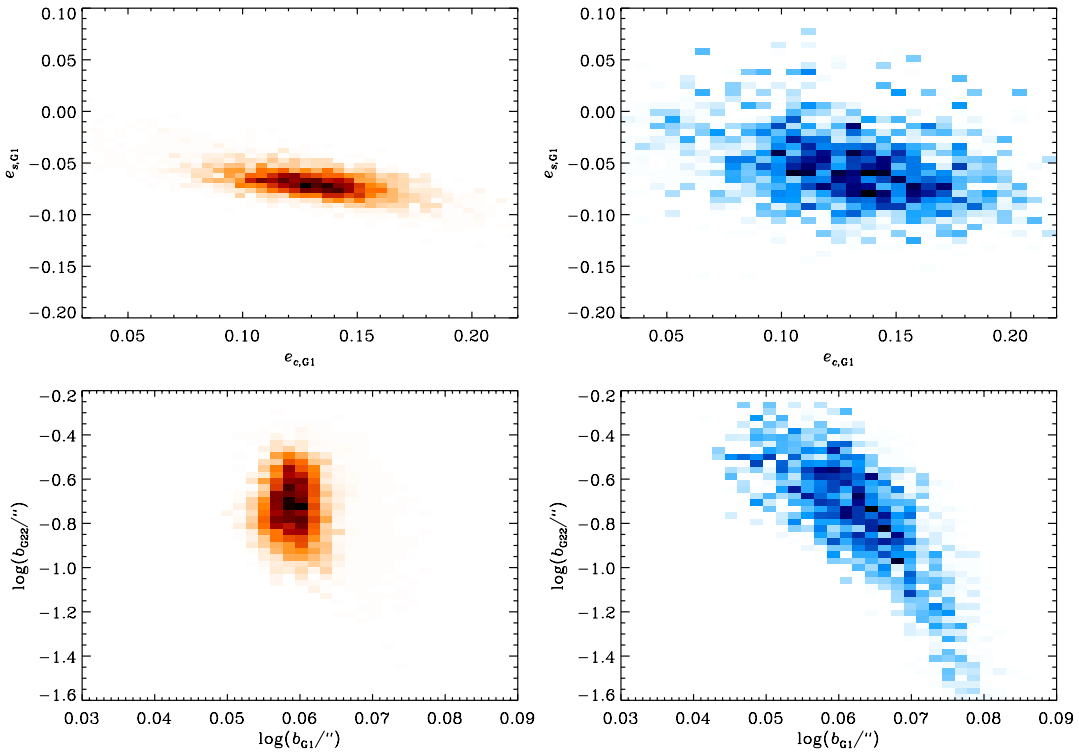

Figure 5.6 Joint posterior probability distributions for four of our macro model parameters. The left column corresponds to parameters from our model with no substructure, while the right column corresponds to the same parameters for our single clump model. Probability of parameter values increases from white ($P(\theta_1, \theta_2|d, M) \approx 0$) to black ($P(\theta_1, \theta_2|d, M) = P_{max}$). The top row presents the distributions for the two components (sine and cosine) of the ellipticity of the lens G1. The bottom row presents the distributions for the mass normalizations of the lens G1 and its companion G22. Generally, adding substructure (right column) broadens the parameter distributions of our models. In some cases, the additional freedom from clumps distorts the shape of the distributions and introduces parameter degeneracies (e.g., bottom row).

with evidence of substructure. However, the vast majority have been unable to place upper and lower bounds on the substructure mass (e.g., Chiba et al. 2005; Minezaki et al. 2009), presumably since constraints from the combination of flux ratios and image positions are not strong enough in the lenses. More recently, studies of the lens systems B2045+265 (McKean et al. 2007), MG 2016+112 (More et al. 2009), and H1413+117 (MacLeod et al. 2009), have been able to find proper constraints on clump masses. In all of these cases the substructure has been linked to a luminous satellite, whose position was subsequently constrained in lens models. Fixing the positions of these satellites,

in turn, breaks the position-mass degeneracy associated with flux perturbations, therefore providing excellent constraints on the substructure masses ($\sigma_M \sim 0.1 - 0.3$ dex). Contrary to these cases, for HE 0435 we vary clump positions in our lens models and allow the data to constrain them. Additionally, we note the masses of clumps A and B are the smallest found in any lens system to date, and there is no known luminous counterpart for the clumps.

5.7.2 Substructure Population Models

So far we have considered the effects of a small number of individual clumps, each lying near one of the lensed images of HE 0435. In reality, however, these clumps are unlikely to be isolated examples of substructure but rather the rare, fortunate few of a larger population, which happen to lie near the lens images (Diemand et al. 2007; Springel et al. 2008). To study the properties of such a population, we conduct Monte Carlo simulations in which clumps are randomly drawn from a mass function of the form $dN/dm \propto m^{\alpha}$ and added to our minimal lens model. We conduct our analysis using a power law slope of $\alpha = -1.9$ (Diemand et al. 2007; Springel et al. 2008) with a fixed upper and lower threshold on the clumps masses of $m_{\rm high} = 10^{10} M_{\odot}$ and $m_{\rm low} = 10^{7} M_{\odot}$, respectively. Clumps are drawn from a uniform spatial distribution, following the work of Dalal & Kochanek (2002). While realistic spatial distributions of substructure are unlikely to be uniform (e.g., Springel et al. 2008), such a distribution is sufficient for HE 0435 since we only constrain the flux ratios and positions of images, which are primarily sensitive to clumps near the images (Keeton 2009).

Before undertaking extensive simulations, we would like to see if we can use the clumps we have inferred so far to estimate the properties of the larger population. While such an estimate should not be taken too seriously, it may help guide our exploration of population models. In Appendix A we present a simple analysis based on the idea that a clump population should have one and only one clump in a position to produce a

strong flux perturbation in image A. Assuming the substructure has a uniform spatial distribution, we derive an expression for the likelihood of the mean convergence in substructure,

$$\mathcal{L}_{\rm clump}(\kappa_{\rm s}) = \frac{A_A}{m_A}\kappa_{\rm s}{\rm exp}\left[-\frac{A_A}{m_A}\kappa_{\rm s}\right]. \tag{5.5}$$

where $\kappa_{\rm s}$ is the substructure convergence, m_A is the mass of clump A, and $A_A = \pi d_A^2$ is the area defined by the distance d_A from the clump to image. Simple in form, this distribution has a peak at $\kappa_{\rm s} = 4\frac{m_A}{A_A}$, a mean of $\kappa_{\rm s} = 3\frac{m_A}{A_A}$, and a standard deviation of $\sigma_{\kappa_{\rm s}} = 2\frac{m_A}{A_A}$. With this in hand, we can estimate the appropriate value(s) of $\kappa_{\rm s}$ to consider by examining the range in $\frac{m_A}{A_A}$ from our model with a single clump near image A (Section 5.7.1). We find the range of most likely values of $\kappa_{\rm s} = 0.025^{+0.074}_{-0.022}$ (95% CL). Therefore, we consider the values of $\kappa_{\rm s} = [0.00022, 0.00046, 0.001, 0.0022, 0.0046, 0.01, 0.022, 0.046, 0.10]$ as the amount of convergence in substructure for our simulations.

For each value of $\kappa_{\rm s}$ we conduct 5000 random realizations of the corresponding substructure population, drawing from the mass function stated above. The ultimate goal of these realizations is to determine which values of $\kappa_{\rm s}$ are favored given the lens data. In the Bayesian approach, this question is answered by comparing the evidence for each value of $\kappa_{\rm s}$. Here, the posterior for $\kappa_{\rm s}$ is written as:

$$P(\kappa_{\rm s}|d, M) = Z^{-1}\int dc\, \mathcal{L}(d|c, \kappa_{\rm s}, \theta_{SM})P(c|\kappa_{\rm s})P(\theta_{SM}|SM) \tag{5.6}$$

where Z is the evidence, $\mathcal{L}(d|c, \kappa_{\rm s}, \theta_{SM})$ is the likelihood of the substructure + smooth model (SM), and $P(\theta_{SM}|SM)$ represents the priors on the smooth macro model. $P(c|\kappa_{\rm s})$ is the probability of the clump positions and masses (denoted by c), given the amount of substructure $\kappa_{\rm s}$. In our models, $P(c|\kappa_{\rm s})$ is uniform for clump positions and defined by the mass function (above) for clump masses. In practical terms, evaluating Equation 5.6 amounts to drawing many realizations from the clump prior such that

$$\int dc\, P(c|\theta, SM) \rightarrow \sum_{c_j} \tag{5.7}$$

which gives

$$P(\theta|d, M) = Z^{-1} \sum_{j} \mathcal{L}(d|c_j, \kappa_{\rm s}, \theta_{SM}) P(c_j|\kappa_{\rm s}) P(\theta_{SM}|SM). \tag{5.8}$$

As a result, the average of the evidence values over all the realizations is then the evidence for that particular value of $\kappa_{\rm s}$. Unfortunately, calculating evidence values for 9×5000 models is unrealistic presently, since our current approach takes from ~ 3 hours, up to ~ 2 days, for a single realization, depending on the value of $\kappa_{\rm s}$.

Instead, we explore the possibility of using the minimum χ^2 value for each realization (which is much easier to determine) as a proxy for the evidence. The minimum χ^2 provides a measure of the peak likelihood, so it may be more or less indicative of the evidence depending on how much scatter there is in the width of the likelihood distribution. As we optimized the 14 lens (Table 5.3) and 2 source parameters for our smooth macro model, we initially found that some realizations led to extremely large ellipticities (~ 0.9 for the companion G22. In order to prevent this, we adopted a mild Gaussian prior of 0.0 ± 0.2 on both (quasi-Cartesian) ellipticity components. While ad hoc, tests indicate the prior is tight enough to prevent unrealistic values for the ellipticity and to stabilize the χ^2 values, yet broad enough to allow large range of ellipticity.

After optimizing all the realizations, we calculate the full Bayesian evidence for a subset of them and investigate the relationship between the peak likelihood (best χ^2) and the evidence. Figure 5.7 shows the relationship for three different values of $\kappa_{\rm s}$. Examining the top panel, we find a distinct pattern in the points. At values of $\chi^2 \lesssim 6.5$, we find χ^2 is an unreliable tracer of the evidence, with normalized values which range from $\sim 10^8$ down to 10^1. Presumably, this arises from certain realizations

being able to produce good peak likelihoods (small χ^2) only for a very highly tuned set of macro parameters, while other realizations are generally better and can have a large range of macro parameters which produce reasonable likelihood. Above $\chi^2 \sim 6.5$, we find the χ^2 values are much more reliable tracers of the evidence. As χ^2 values increase, the evidence values decrease. This trend continues down to $\chi^2 \sim 21$, at which point the values are consistent with that of the smooth, minimal model. For values of $\kappa_{\rm s} \lesssim 0.01$, the relationship between the best χ^2 and the evidence resembles closely the pattern seen in the top panel of Figure 5.7. As $\kappa_{\rm s}$ increases, however, we find the scatter amongst evidence values increases (Figure 5.7 middle and bottom panels). We attribute this increased scatter to the larger number of clumps associated with higher values of $\kappa_{\rm s}$. Such larger clump number densities may over-perturb image A, or affect other lensed images, resulting in a greater range of evidence values at fixed χ^2.

In Figure 5.7, we also plot the binned values of our subset of models. We use the binned values to connect the χ^2 of a particular realization to its associated evidence by interpolating between bins to find the average and uncertainty of the evidence at that χ^2 value. We then scatter this average value by a random amount corresponding to the (interpolated) uncertainty. Using this "lookup" scheme, we are able to calculate the evidence values for *all* of our substructure realizations and compute the posterior as in Equation 5.8.

Figure 5.8 shows the cumulative probability distribution of χ^2 values for each value of $\kappa_{\rm s}$. As the value of $\kappa_{\rm s}$ is increased in our simulations, the distribution of χ^2 values gets broader. In other words, when more substructure is present, there is a higher chance the images will be perturbed and the model will move away from the smooth case. In general, models with modest values of substructure ($\kappa_{\rm s} \lesssim 0.046$) have χ^2 values scattered moderately from our smooth model. A notable exception is our simulation with $\kappa_{\rm s} = 0.10$, for which $\sim 40\%$ of the models have $\chi^2 > 30$

Figure 5.9 depicts our results for the model evidence as a function of the substructure

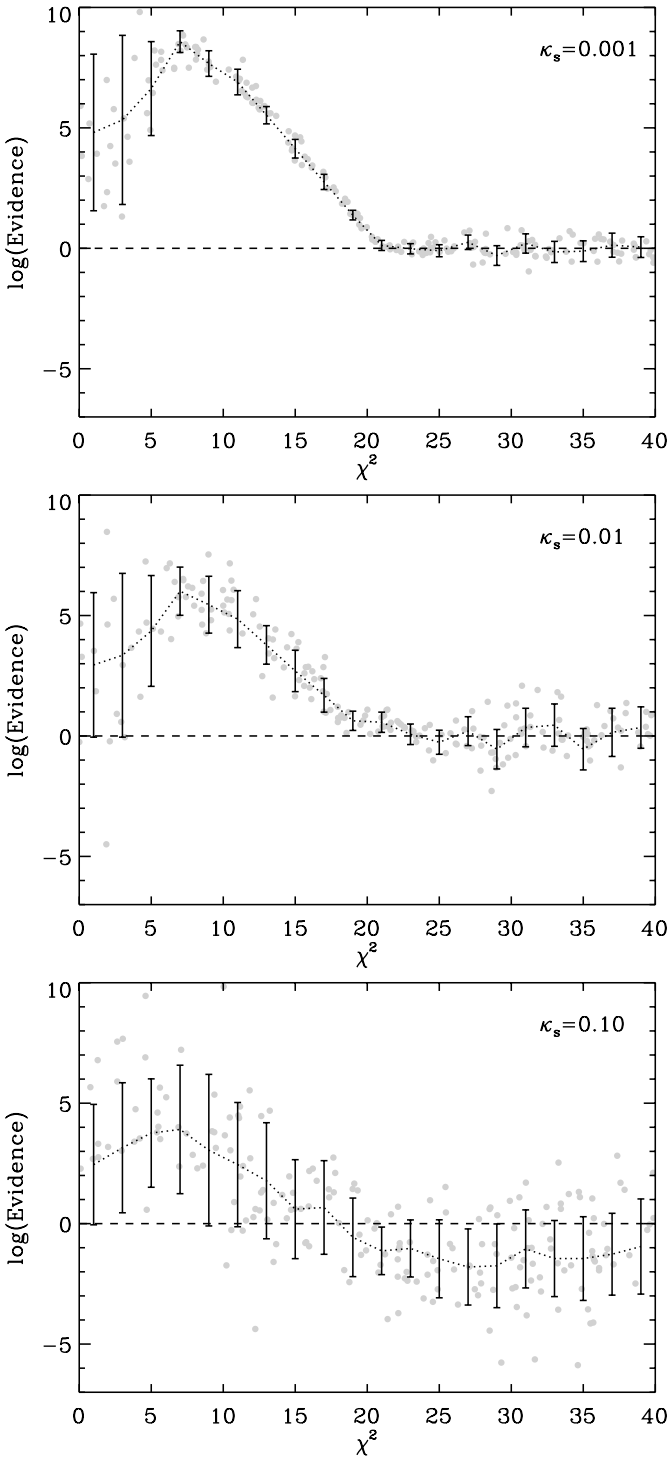

Figure 5.7 χ^2 versus the Bayesian evidence for a subset of our simulations. The top, middle, and bottom panels show the values of evidence (grey points) for κ_s =0.001, 0.01, and 0.10, respectively. In all cases, χ^2 poorly traces the evidence below $\chi^2 \sim 6.5$. For values of $\kappa_s \lesssim 0.0022$, the evidence is tightly correlated with χ^2 values above 6.5 (e.g., top panel). As κ_s increases, so does the scatter of the evidence values. Overplotted are the binned values of the evidence, which are used to connect the evidence to a particular χ^2 of a given realization.

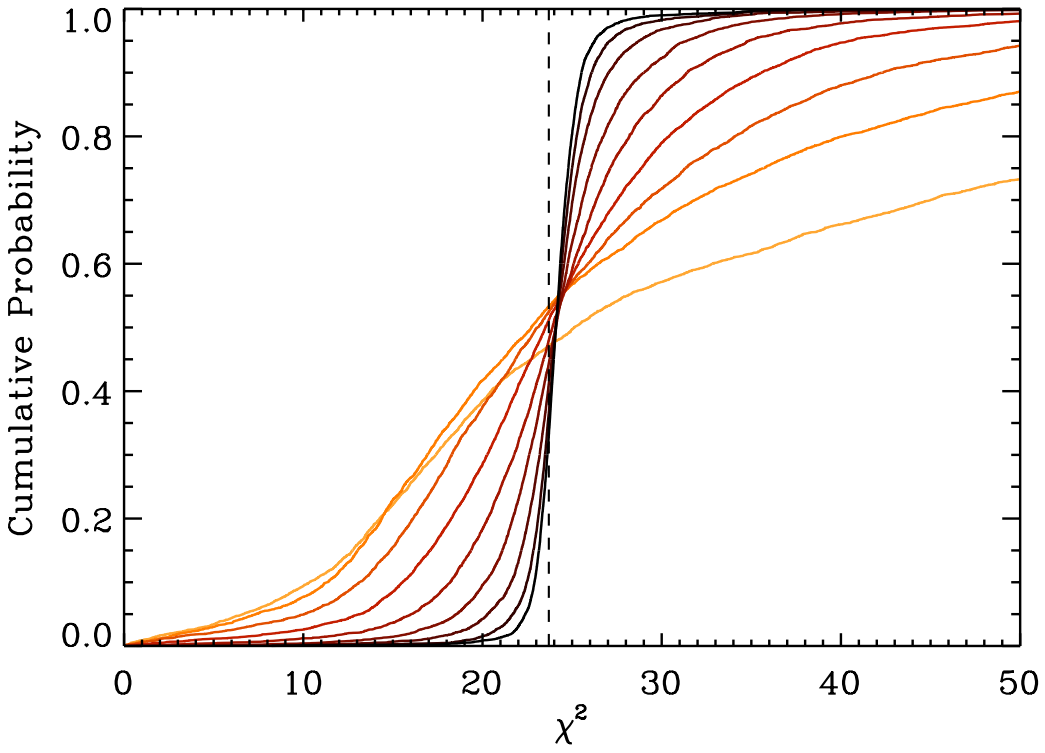

Figure 5.8 The cumulative probability distribution of our models as a function of χ^2. Colors of the lines correspond to different values of κ_s, ranging from 0.00022 (black) to 0.10 (light orange). The vertical dashed line corresponds to the optimized χ^2 of our smooth, minimal model. As κ_s is increased, the distribution is broaden since the presence of substructure allows for better/worse fits to the data.

convergence. We find models with $\kappa_{\rm s} \geq 0.001$ exhibit significantly higher values of evidence than models without substructure. For a uniform spatial distribution, we can write the fraction of mass in substructure as $f_{sub} \simeq 2\kappa_{\rm s}$ since the surface density at the Einstein radius is approximately $\Sigma_{\rm crit}/2$ for an isothermal macro model. Translating our results to mass fraction of substructure, we find $f_{sub} > 0.00092$ in HE 0435. Examining Figure 5.9, it is interesting that the evidence for f_{sub} is still large at high values, even when $f_{sub} = 0.20$. We speculate that such high f_{sub} values are possible thanks to the flexibility and freedom available to our macro model.

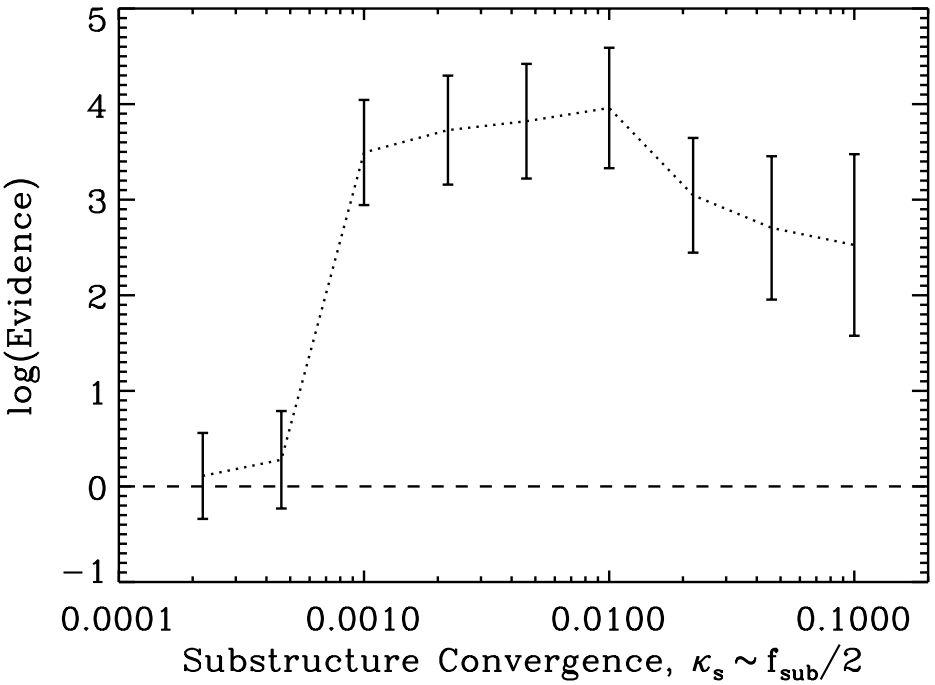

Figure 5.9 Values of the log of the evidence from our substructure realizations, as a function of $\kappa_{\rm s}$. Evidence values are normalized to the value associated with our smooth macro model, represented by the dashed horizontal line. We find models with $\kappa_{\rm s} \leq 0.00046$ exhibit evidence values similar to those without substructure. Decisively, models with $\kappa_{\rm s} \geq 0.001$ are strongly favored at $> 3\sigma$.

Our lower bound of $f_{sub} > 0.00092$ is consistent with other lensing based measurements. Using a sample of seven radio quads, Dalal & Kochanek (2002) found $0.006 < f_{sub} < 0.07$ at 90% confidence. In the case of SDSS 0946+1006, Vegetti et al.

(2009) found $f_{sub} = 0.0215^{+0.0205}_{-0.0125}$ (68% confidence), assuming $\alpha = -1.9 \pm 0.1$ for the slope of the substructure mass function and mass thresholds $m_{\rm high} = 10^{9.6} M_{\odot}$ and $m_{\rm low} = 10^{6.6} M_{\odot}$. These results, however, all point to values of f_{sub} which are significantly higher than those found in N-body simulations $\sim 0.002 - 0.003$ (e.g., Diemand et al. 2007; Springel et al. 2008; Xu et al. 2010). Contrary to previous results, in HE 0435 we find lens-based values of f_{sub} which are fully consistent with N-body simulations, with values of $f_{sub} \sim 0.001 - 0.003$ capable of reproducing the lensing.

Our models also permit values of f_{sub} which are much higher than CDM simulations predict, similar to the previous results of Dalal & Kochanek (2002) and Vegetti et al. (2009). Therefore, it is worth considering the origin of high f_{sub} values, and whether they are truly discrepant from CDM models. One possible explanation for high f_{sub} values is that the number of surviving subhalos in N-body simulations is underestimated, due to the lack of baryons in the simulations. In the presence of baryons, dark matter halos are expected to contract adiabatically and become more concentrated. Thus, baryonic material might create subhalos which are more resilient to tidal disruption (e.g., Dolag et al. 2009). However, in the presence of baryons the primary halo of the galaxy will also become more concentrated, providing a higher density contrast and increasing the rate of tidal disruption (e.g., Romano-Díaz et al. 2010). Other possible explanations exist. Currently, the number of lenses available for studying f_{sub} is small in number (~ 10). Expansion of the current sample, therefore, will greatly improve the statistics. Additionally, lenses are known to exhibit significant observational biases (e.g., Mandelbaum et al. 2009), favoring objects which lie in the more massive and concentrated halos, which may have preferential orientations along the line of sight. Furthermore, lenses are known to inhabit denser environments, possibly enhancing the value of f_{sub} (Oguri 2005). It is reasonable, therefore, that there may be a bias towards higher values of f_{sub} in lenses, if f_{sub} depends significantly on such factors.

In addition to our results on f_{sub}, there are a couple aspects of our analysis worth

highlighting. First, we have sought to connect our results from individual clump models to those for full substructure populations. We have done so by taking our individual model, using it to infer $\kappa_{\rm s}$, and subsequently using the *same* macro model to study the population. This is a first for quasar lenses. For galaxy-galaxy lenses, Vegetti et al. (2009) have detected a clump via gravitational imaging, and used that to infer f_{sub} by constructing a likelihood similar (in concept) to the likelihood we present in Appendix A.

Another key feature of our analysis is our method used in our population simulations. Here, the only comparable previous result is that of Dalal & Kochanek (2002). Due to the complexity and computational demand of the study, Dalal & Kochanek chose to linearly optimize their lens macro models. In our analysis, we not only do a full marginalization of the macro model, but we estimate the Bayesian evidence associated with each realization. Shown in Figure 5.7, we find the best χ^2 value can be an unreliable tracer of the evidence. Thus, the Dalal & Kochanek result may not reflect the true value of f_{sub}, if such behavior is present in their models. Additionally, Dalal & Kochanek assumed a constant mass for their substructure population. Here, we consider a more realistic population which is drawn from a mass function (with $\alpha = -1.9$), with a range of substructure masses ($10^7 - 10^{10}\,M_\odot$). Due to the computational demand of our simulations, we have so far only considered one value of α and one range of clump masses. Future work will explore the dependence of our f_{sub} results on these parameters.

5.7.3 K band flux ratios

Our substructure models presented in Sections 5.7.1 and 5.7.2 are able to account for the optical R band flux ratios and, due to their similarity, the L' flux ratios listed in Table 5.2. As noted in Section 5.3, however, we find some of our measured K band flux ratios discrepant from those in the R and L' bands. Specifically, we find the magnification of image B is a factor 1.27 higher in the K band data than in R and L' bands. This

anomaly is particularly perplexing since, at a rest-frame wavelength of 0.818 μm, the K band emission presumably originates from the same (effectively) point-like accretion disk as in the R band.

Since differential dust extinction is not likely in HE 0435 (Wisotzki et al. 2003), we hypothesize the K band anomaly is due to microlensing by stars. We test the reasonability of this hypothesis by simulating the expected microlensing magnification distributions near image B, using the ray-shooting code of Wambsganss (1999).

Conceptually these simulations are simple. First a small box L is created for the simulation. The size of L is selected such that it is much larger than the average Einstein radius of the stellar distribution. Subsequently, a uniform convergence and shear is applied to the box. This convergence is broken down into two components, one associated with the stars and one with any smoothly distributed matter. The stars are then randomly distributed by drawing from a user-specified mass function, such that the amount of stars matches the stellar convergence. Once the mass distribution is set up, rays are shot through the box, producing a magnification map across the box. From here, the magnification map is convolved with the specified source, producing a probability distribution function (PDF) for the microlensing magnification. This entire process is then repeated many times to insure fair sampling of the PDF.

For each simulation, we set the total convergence and shear to the values predicted by our best-fit two clump (AB) lens model: $\kappa_{\rm local} = 0.694$, $\gamma_{\rm local} = 0.486$. In order to calculate microlensing magnification distributions, we first need an estimate of how much convergence is granular in form (i.e., stars) and how much is smoothly distributed (i.e., dark matter). To do so, we split up the local convergence into a stellar and dark components using the models of Kochanek et al. (2006). Using separate stellar and dark matter distributions in their lens model, Kochanek et al. relate the stellar convergence to their dark matter profile as $\langle \kappa_\star \rangle = 0.05 \log(h r_c \, {\rm kpc})$, where r_c is the NFW scale radius in their models. We estimate the minimal and maximal stellar convergence using the

smallest and largest of Kochanek et al. well-fitting values for r_c, 2.5″ and 20.0″. For these radii, the relation yields $\langle\kappa_\star\rangle = 0.05$ and 0.10, respectively.

For both values of $\langle\kappa_\star\rangle$, we conduct 100 random realizations of the stellar distribution, drawn from a mass function $dN/dm \propto m^{-1.3}$ with $m_{\rm low} = 0.01M_\odot$ and $m_{\rm high} = 1.5M_\odot$. Such a mass function is selected to be in agreement with measurements from the Galactic bulge (Gould 2000), and has been used previously in microlensing studies (Morgan et al. 2010; Poindexter & Kochanek 2010).

For each realization a magnification map is produced using a box size of L = $15R_E(M_\odot)$, at a resolution of L/1024. Subsequently, we convolve each map by a uniform circular profile whose size corresponds to the expected range in the K band source size. To estimate the size of the source, we use the empirically derived result for the I band source of Morgan et al. (2010). We then scale the I band source (rest-frame 0.260μm) to our K band source using the Shakura & Sunyaev (1973) $R \propto \lambda^{4/3}$ relation. We find a K band source size of $\log(\frac{R_{\rm src}}{1''}) = -6.1^{+0.5}_{-0.7}$, a factor $0.2-3.1$ (68% CL) times the Einstein radius of the mean stellar mass ($\log(\frac{R_{\langle m_\star\rangle}}{1''}) = -6.1$).

We present the results of our microlensing simulations in Figure 5.10. In order to explain the observed K band B/C flux ratios, microlensing must provide a factor ~ 1.27 increase to the magnification of our best fit model. In general, we find microlensing in HE 0435 can account for our K band data for all but the largest source sizes in our simulations. To assess which source sizes can reasonably account for the data, we consider the range in sizes which produce magnifications > 1.27, more than 16% of the time (i.e., a two-sided 68% confidence limit). Under our assumed stellar density fractions and stellar mass functions, we find the needed K band magnification of image B occurs for source sizes $\lesssim 0.48 \times R_{\langle m_\star\rangle}$ for $\langle\kappa_\star\rangle = 0.05$, and $\lesssim 0.73 \times R_{\langle m_\star\rangle}$ for $\langle\kappa_\star\rangle = 0.10$. These source sizes are well within our inferred K band source sizes ($0.2-3.1 \times R_{\langle m_\star\rangle}$). Given the significant uncertainties in the source sizes, stellar density fraction, stellar mass function and flux ratio measurements, we conclude microlensing

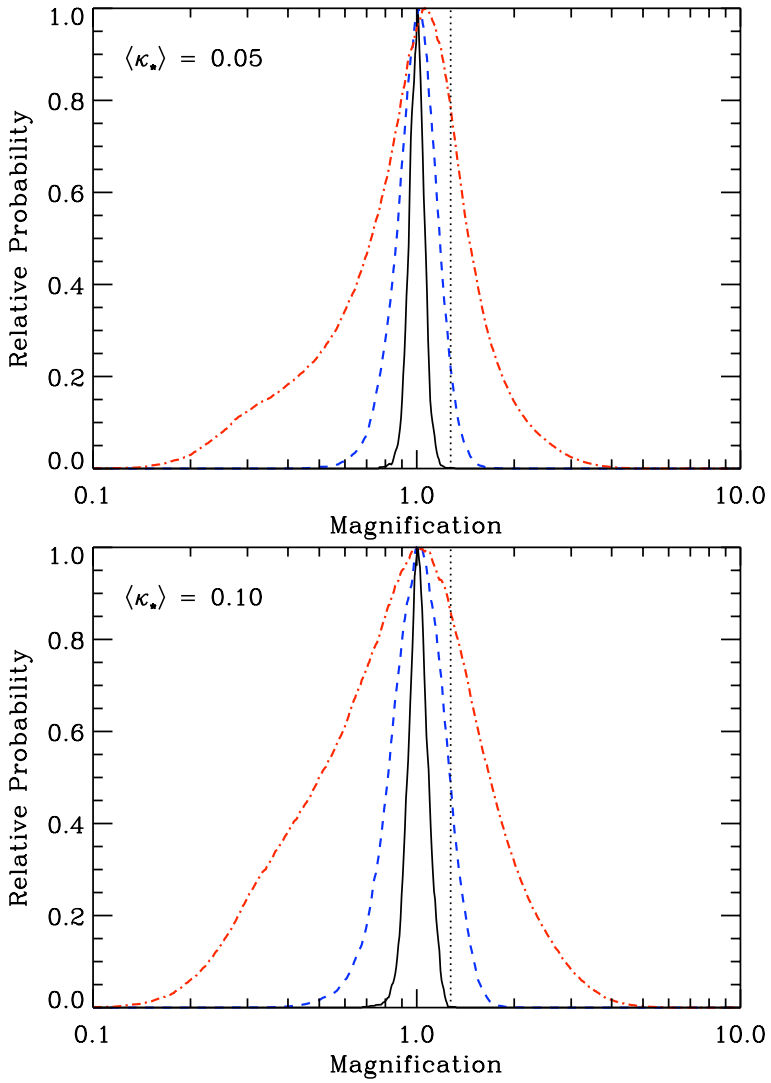

Figure 5.10 Above we plot the probability distributions of the magnification due to microlensing by stars for various source sizes. The magnification is normalized to that for Image B from our best fit model. The three distributions plotted are for source sizes corresponding to the median (dashed, blue line) and 68% CL (black, solid and dot-dashed, red lines) sizes inferred from Morgan et al. (2010). The vertical dotted line represents the mean value needed to reproduce our K band observations. We find small source sizes, $\lesssim 0.48, 0.73 \times R_{\langle m_\star \rangle}$, reproduce the observations well, for $\langle \kappa_\star \rangle = 0.05, 0.10$, respectively.

is a plausible explain for the image B flux in our K band data. Confirmation of this explanation will be possible with future K band imaging of the system, which will quantify the variability of the microlensing magnification.

5.8 Conclusions

We have conducted a fully Bayesian analysis of the lens HE 0435-1223. With precise astrometry from HST and R band flux ratios from ground-based monitoring, we examine the mass distribution down to milli-arcsecond scales, testing various models of substructure through their Bayesian Evidence. Additionally, we have obtained new near-infrared images, providing insight into the multi-wavelength properties of the lens. We summarize our conclusions as follows:

- The flux ratio of images A and C cannot be reproduced by macroscopic, smoothly distributed lens models. The source of this failure cannot be due to microlensing or intrinsic variability of the background quasar, since optical monitoring has quantified such variations. Instead, we show that a single clump with mass $\log(\frac{M_A(<R_{Ein})}{M_\odot\, h_{70}^{-1}}) = 7.68^{+0.92}_{-0.85}$ can account for the flux ratio, once added to the smooth mass distribution.

- We analyze other sources of small scale structure by including additional clumps near the lens images. Using the Bayesian evidence to discriminate between the various possibilities, we find a model with clumps near images A and B is most favored. This model an associated evidence which is 0.63 dex greater than our single clump model, and implies a mass for clump B of $\log(\frac{M_B(<R_{Ein})}{M_\odot\, h_{70}^{-1}}) = 6.6^{+1.02}_{-1.52}$. While interesting, the evidence uncertainties of models A and AB are 0.12 dex, making the case for clump B much weaker than for clump A.

- Connecting these clumps to a larger population of subhalos, we conduct Monte Carlo simulations of the full ensemble of substructures. We set the abundance ($\kappa_{\rm s}$)

of the clumps from the mass and location of our single clump model, and draw realizations from a mass function with slope $\alpha = -1.9$. Calculating the evidence associated with each value of $\kappa_{\rm s}$, we infer the mass fraction of substructure to be $f_{sub} > 0.00092$ near the Einstein radius. Our measurement of f_{sub}, unlike other lensing based measurements, is fully consistent with that predicted by CDM simulations ($f_{sub} \approx 0.003$).

- Near infrared flux ratio measurements in the K and L' wavebands generally agree with those from optical monitoring. The lone exception is the K band value of the flux ratio B/C. Using a series of microlensing simulations, we generate magnification maps from the stellar distribution of the lens. We find that microlensing may indeed be responsible for the B/C ratio, if the K band source is $\lesssim 10^{6.24''}$ in size. Estimations based on results from Morgan et al. (2010) show that such a source size is reasonable.

Appendix A

Connecting individual clumps to the population

A.1 Likelihood for a single clump

In Section 5.7.2 we have described the form of the mass function and the spatial distribution from which our subhalo population is drawn. A remaining key quantity we must consider is the amount of substructure appropriate for our models, which we express as $\kappa_{\rm s} = \Sigma_{\rm s}/\Sigma_{\rm crit}$, otherwise known as the convergence due to substructure in the lens.

To estimate $\kappa_{\rm s}$, we seek to construct a likelihood that one and only one clump drawn from a population of subhalos is close enough to perturb a given image. Let $m = M/\Sigma_{\rm crit} = \pi R_E^2$ be the scaled mass of a point mass clump that has units of angular area. For a population with a number density of clumps per unit mass of dn/dm, the mean surface mass density in substructure is then

$$\kappa_{\rm s} = \int m \frac{dn}{dm} dm. \tag{A.1}$$

For clumps near an image, let $A(m)$ be the area which defines the"region of influence" for a given clump of mass m. Examining only perturbations to the magnification of the image, we have $A(m) \propto R_E^2 \propto m$ (since $\delta\mu$ scales with R_E^2). If we define the clump near image i to have a scaled mass m_i at a distance d_i from the image, we can then use our individual clump models (Section 5.7.1) to calculate the proportionality

$$A(m) = \frac{A_i}{m_i} m, \tag{A.2}$$

where A_i is simply the geometric area πd_i^2.

Under these assumptions, the mean number of clumps near an image for a given population is simply

$$\left\langle \hat{N}_i \right\rangle = \int A(m) \frac{dn}{dm} dm = \frac{A_i}{m_i} \kappa_{\mathrm{s},i} \tag{A.3}$$

where $\kappa_{\mathrm{s},i}$ is the mean surface mass density in substructure in the vicinity of image i. The probability that there are exactly N_i clumps affecting the image is a Poisson distribution of the form

$$P(N_i | \left\langle \hat{N}_i \right\rangle) = \frac{\left\langle \hat{N}_i \right\rangle^{N_i} e^{-\langle \hat{N}_i \rangle}}{N_i!}. \tag{A.4}$$

Finally, we can assume the clumps are independent since the clump near image i is small enough such that it does not affect the magnification of image $j \neq i$. The final likelihood for a single clump near a given number of images is then the product

$$\mathcal{L}_{\mathrm{clump}} = \prod_i P(1 | \left\langle \hat{N}_i \right\rangle) = \prod_i \left\langle \hat{N}_i \right\rangle e^{-\langle \hat{N}_i \rangle}. \tag{A.5}$$

In Section 5.7.2, we apply Equation to the case of image A alone.

Bibliography

Amara, A., Metcalf, R. B., Cox, T. J., & Ostriker, J. P. 2006, MNRAS, 367, 1367

Barnabè, M., Czoske, O., Koopmans, L. V. E., Treu, T., Bolton, A. S., & Gavazzi, R. 2009, MNRAS, 399, 21

Belokurov, V., Walker, M. G., Evans, N. W., Gilmore, G., Irwin, M. J., Just, D., Koposov, S., Mateo, M., Olszewski, E., Watkins, L., & Wyrzykowski, L. 2010, ApJ, 712, L103

Belokurov, V., Walker, M. G., Evans, N. W., Gilmore, G., Irwin, M. J., Mateo, M., Mayer, L., Olszewski, E., Bechtold, J., & Pickering, T. 2009, MNRAS, 397, 1748

Blackburne, J. A., & Kochanek, C. S. 2010, ApJ, 718, 1079

Bradač, M., Schneider, P., Steinmetz, M., Lombardi, M., King, L. J., & Porcas, R. 2002, A&A, 388, 373

Chen, J., Rozo, E., Dalal, N., & Taylor, J. E. 2007, ApJ, 659, 52

Chiba, M. 2002, ApJ, 565, 17

Chiba, M., Minezaki, T., Kashikawa, N., Kataza, H., & Inoue, K. T. 2005, ApJ, 627, 53

Dalal, N., & Kochanek, C. S. 2002, ApJ, 572, 25

Diemand, J., Kuhlen, M., & Madau, P. 2007, ApJ, 657, 262

Dobler, G., & Keeton, C. R. 2006, MNRAS, 365, 1243

Dolag, K., Borgani, S., Murante, G., & Springel, V. 2009, MNRAS, 399, 497

Fadely, R., Keeton, C. R., Nakajima, R., & Bernstein, G. M. 2010, ApJ, 711, 246

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. 2003, Bayesian Data Analysis, Second Edition (Chapman & Hall/CRC Texts in Statistical Science), 2nd edn. (Chapman & Hall)

Gnedin, N. Y. 2000, ApJ, 542, 535

Gould, A. 2000, ApJ, 535, 928

Hönig, S. F., Prieto, M. A., & Beckert, T. 2008, A&A, 485, 33

Humphrey, P. J., Buote, D. A., Brighenti, F., Gebhardt, K., & Mathews, W. G. 2009, ApJ, 703, 1257

Irwin, M. J., Belokurov, V., Evans, N. W., Ryan-Weber, E. V., de Jong, J. T. A., Koposov, S., Zucker, D. B., Hodgkin, S. T., Gilmore, G., Prema, P., Hebb, L., Begum, A., Fellhauer, M., Hewett, P. C., Kennicutt, Jr., R. C., Wilkinson, M. I., Bramich, D. M., Vidrih, S., Rix, H., Beers, T. C., Barentine, J. C., Brewington, H., Harvanek, M., Krzesinski, J., Long, D., Nitta, A., & Snedden, S. A. 2007, ApJ, 656, L13

Jeffreys, H. 1961, Theory of probability, 3rd edn. (Oxford University Press)

Kalirai, J. S., Beaton, R. L., Geha, M. C., Gilbert, K. M., Guhathakurta, P., Kirby, E. N., Majewski, S. R., Ostheimer, J. C., Patterson, R. J., & Wolf, J. 2010, ApJ, 711, 671

Keeton, C. R., 2001, arXiv:astro-ph/0102340

—. 2003, ApJ, 584, 664

—. 2009, arXiv:astro-ph/0908.3001

Keeton, C. R., Gaudi, B. S., & Petters, A. O. 2003, ApJ, 598, 138

—. 2005, ApJ, 635, 35

Keeton, C. R., & Moustakas, L. A. 2009, ApJ, 699, 1720

Kelly, B. C., Bechtold, J., & Siemiginowska, A. 2009, ApJ, 698, 895

Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82

Kochanek, C. S., Morgan, N. D., Falco, E. E., McLeod, B. A., Winn, J. N., Dembicky, J., & Ketzeback, B. 2006, ApJ, 640, 47

Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Nolta, M. R., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2010, arXiv:astro-ph/1001.4538

Koopmans, L. V. E., Garrett, M. A., Blandford, R. D., Lawrence, C. R., Patnaik, A. R., & Porcas, R. W. 2002, MNRAS, 334, 39

Kozłowski, S., Kochanek, C. S., Udalski, A., Wyrzykowski, Ł., Soszyński, I., Szymański, M. K., Kubiak, M., Pietrzyński, G., Szewczyk, O., Ulaczyk, K., Poleski, R., & The OGLE Collaboration. 2010, ApJ, 708, 927

Liu, C., Hu, J., Newberg, H., & Zhao, Y. 2008, A&A, 477, 139

Macciò, A. V., Kang, X., Fontanot, F., Somerville, R. S., Koposov, S., & Monaco, P. 2010, MNRAS, 402, 1995

Macciò, A. V., & Miranda, M. 2006, MNRAS, 368, 599

Macciò, A. V., Moore, B., Stadel, J., & Diemand, J. 2006, MNRAS, 366, 1529

MacKay, D. J. C. 2003, Information Theory, Inference, and Learning Algorithms (Cambridge University Press), available from http://www.inference.phy.cam.ac.uk/mackay/itila/

MacLeod, C. L., Ivezić, Ž., Kochanek, C. S., Kozłowski, S., Kelly, B., Bullock, E., Kimball, A., Sesar, B., Westman, D., Brooks, K., Gibson, R., Becker, A., & de Vries, W. H. 2010, arXiv:astro-ph/1004.0276

MacLeod, C. L., Kochanek, C. S., & Agol, E. 2009, ApJ, 699, 1578

Madau, P., Diemand, J., & Kuhlen, M. 2008, ApJ, 679, 1260

Mandelbaum, R., van de Ven, G., & Keeton, C. R. 2009, MNRAS, 398, 635

Mao, S., Jing, Y., Ostriker, J. P., & Weller, J. 2004, ApJ, 604, L5

Mao, S., & Schneider, P. 1998, MNRAS, 295, 587

Mashchenko, S., Wadsley, J., & Couchman, H. M. P. 2008, Science, 319, 174

Mateo, M. L. 1998, ARA&A, 36, 435

McKean, J. P., Koopmans, L. V. E., Flack, C. E., Fassnacht, C. D., Thompson, D., Matthews, K., Blandford, R. D., Readhead, A. C. S., & Soifer, B. T. 2007, MNRAS, 378, 109

Metcalf, R. B., & Madau, P. 2001, ApJ, 563, 9

Metcalf, R. B., & Zhao, H. 2002, ApJ, 567, L5

Minezaki, T., Chiba, M., Kashikawa, N., Inoue, K. T., & Kataza, H. 2009, ApJ, 697, 610

Moore, B., Ghigna, S., Governato, F., Lake, G., Quinn, T., Stadel, J., & Tozzi, P. 1999, ApJ, 524, L19

More, A., McKean, J. P., More, S., Porcas, R. W., Koopmans, L. V. E., & Garrett, M. A. 2009, MNRAS, 394, 174

Morgan, C. W., Kochanek, C. S., Morgan, N. D., & Falco, E. E. 2010, ApJ, 712, 1129

Morgan, N. D., Kochanek, C. S., Pevunova, O., & Schechter, P. L. 2005, AJ, 129, 2531

Mukherjee, P., Parkinson, D., & Liddle, A. R. 2006, ApJ, 638, L51

Nenkova, M., Sirocky, M. M., Nikutta, R., Ivezić, Ž., & Elitzur, M. 2008, ApJ, 685, 160

Oguri, M. 2005, MNRAS, 361, L38

Penarrubia, J., Benson, A. J., Walker, M. G., Gilmore, G., McConnachie, A., & Mayer, L. 2010, MNRAS, 406, 1290

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H. 2002, AJ, 124, 266

Poindexter, S., & Kochanek, C. S. 2010, ApJ, 712, 658

Romano-Díaz, E., Shlosman, I., Heller, C., & Hoffman, Y. 2010, ApJ, 716, 1095

Rowan-Robinson, M. 1995, MNRAS, 272, 737

Scannapieco, E., Thacker, R. J., & Davis, M. 2001, ApJ, 557, 605

Schechter, P. L., & Wambsganss, J. 2002, ApJ, 580, 685

Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337

Simon, J. D., & Geha, M. 2007, ApJ, 670, 313

Skilling, J. 2004, in American Institute of Physics Conference Series, Vol. 735, American Institute of Physics Conference Series, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 395–405

Springel, V., Wang, J., Vogelsberger, M., Ludlow, A., Jenkins, A., Helmi, A., Navarro, J. F., Frenk, C. S., & White, S. D. M. 2008, MNRAS, 391, 1685

Strigari, L. E., Bullock, J. S., Kaplinghat, M., Diemand, J., Kuhlen, M., & Madau, P. 2007, ApJ, 669, 676

Tollerud, E. J., Bullock, J. S., Strigari, L. E., & Willman, B. 2008, ApJ, 688, 277

Vegetti, S., Czoske, O., & Koopmans., L. V. E. 2010, arXiv:astro-ph/1002.4708

Vegetti, S., & Koopmans, L. V. E. 2009a, MNRAS, 392, 945

—. 2009b, MNRAS, 400, 1583

Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T., & Gavazzi, R. 2009, arXiv:astro-ph/0910.0760

Wadepuhl, M., & Springel, V. 2010, arXiv:astro-ph/1004.3217

Wambsganss, J. 1999, Journal of Computational and Applied Mathematics, 109, 353

Willman, B., Blanton, M. R., West, A. A., Dalcanton, J. J., Hogg, D. W., Schneider, D. P., Wherry, N., Yanny, B., & Brinkmann, J. 2005, AJ, 129, 2692

Wisotzki, L., Becker, T., Christensen, L., Helms, A., Jahnke, K., Kelz, A., Roth, M. M., & Sanchez, S. F. 2003, A&A, 408, 455

Wisotzki, L., Schechter, P. L., Bradt, H. V., Heinmüller, J., & Reimers, D. 2002, A&A, 395, 17

Wittkowski, M., Kervella, P., Arsenault, R., Paresce, F., Beckert, T., & Weigelt, G. 2004, A&A, 418, L39

Wong, K. C., Keeton, C. R., Williams, K. A., Momcheva, I. G., & Zabludoff, A. I. 2010, ApJ, submitted

Xu, D., Mao, S., Cooper, A., Wang, J., Gao, L., Frenk, C., & Springel, V. 2010, arXiv:astro-ph/1004.3094

York, D. G., et al., 2000, AJ, 120, 1579

CURRICULUM VITAE

ROSS FADELY

EDUCATION

- **Rutgers University**—Department of Physics and Astronomy

  PhD in Astrophysics - October 2010

  Advisor: Charles R. Keeton

- **Rutgers University**—Department of Physics and Astronomy

  B.S. in Physics - May 2004

  High Honors

POSITIONS HELD

- **7/2008-Present** Graduate Assistant in Astrophysics, Rutgers
  - reference: Prof. Charles R. Keeton
- **7/2007-6/2008** Teaching Assistant, Rutgers
  - references: Prof. Mohan Kalelkar, Prof. Andrew Baker
- **7/2006-6/2007** Graduate Assistant in Astrophysics, Rutgers
  - reference: Prof. Charles R. Keeton
- **9/2004-6/2006** Teaching Assistant, Rutgers
  - reference: Prof. Mohan Kalelkar

REFEREED PUBLICATIONS

- R. Nakajima, G. M. Bernstein, **R. Fadely**, C. R. Keeton, & T. Schrabback, 2009, ApJ, 697, 1793, "Improved Constraints on the Gravitational Lens Q0957+561. I. Weak Lensing"

- **R. Fadely**, C. R. Keeton, R. Nakajima, & G. M. Bernstein, 2010, ApJ, 711, 246, "Improved Constraints on the Gravitational Lens Q0957+561. II. Strong Lensing"

SUBMITTED PUBLICATIONS

- **R. Fadely**, S. S. Allam, A. J. Baker, H. Lin, D. Lutz, A. E. Shapley, M.-S. Shin, J. A. Smith, M. A. Stauss, & D. L. Tucker, 2010, ApJ submitted, "Mid-Infrared Spectroscopy of Two Lensed Star-forming Galaxies"

PUBLICATIONS IN PREPARATION

- **R. Fadely** & C. R. Keeton, "Dark Matter Substructure in the Lens HE 0435-1223"
- **R. Fadely** & C. R. Keeton, "Identifying Substructure with Near-Infrared Flux Ratios"