Markov Chain Monte Carlo Modeling of High-Redshift Quasar Host Galaxies in
Hubble Space Telescope Imaging

by

Matt R. Mechtley

A Dissertation Presented in Partial Fulfillment
of the Requirement for the Degree
Doctor of Philosophy

Approved December 2014 by the
Graduate Supervisory Committee:

Rogier A. Windhorst, Chair
Nathaniel Butler
Rolf A. Jansen
James Rhoads
Paul Scowen

ARIZONA STATE UNIVERSITY

May 2014

ABSTRACT

Quasars, the visible phenomena associated with the active accretion phase of super-massive black holes found in the centers of galaxies, represent one of the most energetic processes in the Universe. As matter falls into the central black hole, it is accelerated and collisionally heated, and the radiation emitted can outshine the combined light of all the stars in the host galaxy. Studies of quasar host galaxies at ultraviolet to near-infrared wavelengths are fundamentally limited by the precision with which the light from the central quasar accretion can be disentangled from the light of stars in the surrounding host galaxy.

In this Dissertation, I discuss direct imaging of quasar host galaxies at redshifts $z \simeq 2$ and $z \simeq 6$ using new data obtained with the Hubble Space Telescope. I describe a new method for removing the point source flux using Markov Chain Monte Carlo parameter estimation and simultaneous modeling of the point source and host galaxy. I then discuss applications of this method to understanding the physical properties of high-redshift quasar host galaxies including their structures, luminosities, sizes, and colors, and inferred stellar population properties such as age, mass, and dust content.

"Everything we see hides another thing, we always want to see what is hidden by what we see. There is an interest in that which is hidden and which the visible does not show us." — René Magritte

To my grandmother Maureen Dodson, who watched every Space Shuttle launch, provided me with National Geographic magazines, and makes the best pies.

To my brothers Adam and Brandon Mechtley, whose interests and expertise continue to complement my own, over a quarter century later.

To my parents Alan Mechtley and Patricia Dodson, who encouraged us to learn and explore, even if they didn't understand the curious white words on the blue screen:
`"DO...LOOP...END"`

To Cayley, meow meow, meow meow meow.

To my partner Alisha Rossi, a true intellectual romantic.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my research advisor, Dr. Rogier Windhorst. He has provided unique advice, encouragement, financial support, and insights that come only with decades of experience. He has consistently put his students' needs before his own, helped provide myriad opportunities for collaboration, and ensured I left well-prepared to write proposals, grants, and papers on my own. I would also like to thank Drs. Seth Cohen and Rolf Jansen, who have provided immeasurable advice and support on a daily basis. I would like to thank my entire committee for their contributions and critiques that made this dissertation possible: Prof. Nathaniel Butler, Dr. Rolf Jansen, Prof. James Rhoads, Prof. Paul Scowen, and Prof. Rogier Windhorst.

I have been fortunate to collaborate with many talented researchers across the globe during the course of my graduate work. No research is accomplished alone, and none of the work herein would have been possible without support and advice from the following collaborators and experts: Prof. Xiaohui Fan, Dr. Nimish P. Hathi, Dr. Knud Jahnke, Dr. Linhua Jiang, Prof. William C. Keel, Dr. Anton M. Koekemoer, Dr. Samuel Lawrence, Dr. Knox Long, Dr. John MacKenty, Prof. Phillip Mauskopf, Dr. Norbert Pirzkal, Prof. Mark Robinson, Tony Roman, Prof. Huub Röttgering, Dr. Russell E. Ryan, Jr., Prof. Evan Scannapieco, Prof. Donald P. Schneider, Dr. Glenn Schneider, Prof. Michael A. Strauss, Dr. Lisa Will, and Prof. Haojing Yan. Several professors have also contributed significantly to my graduate education. I'd like to thank Profs. Sangeeta Malhotra, James Rhoads, Evan Scannapieco, Paul Scowen, and Patrick Young for the significant effort that goes into teaching graduate courses. I'd also like to thank the academic support staff that have worked tirelessly to keep my ducks in a row and shield me from the iniquities of bureaucracy: Scott Smas, Sunny Thompson, Pattie Dodson, Becca Dial, and Becky Polley.

Sometimes a good friend, a good beer, and an hour of ranting can solve a problem faster than reading a dozen papers. I'm grateful to the following brilliant people for their friendship: Teresa Ashcraft, Danny Baranowsky, Sarah Braden, Ben Braman, Alex Burley, Chrissy Chubala, Seth Cohen, Cynthia d'Angelo, Carola Ellinger, Kim Emig, Alex Fink, Steve Finkelstein, Bhavin Joshi, Katie Kaleida, Boom Kittiwisit, Dave Koontz, Paul Hegel, Pascale Hibon, Natalie Hinkel, Josh Kessler, Duho Kim, Hwi Kim, Mike Kime, Karen Knierman, Mike Lesniak, Emily McLinden, Michael Mein, Jackie Monkiewicz, Andy Nealen, Corey Nolan, Peter Nguyen, Lillian Ostrach, Kyle Pulver, Tommy Refenes, Mark Richardson, Erin Robinson, Ben Ruiz, Mike Rutkowski, Russell Ryan, Tae-hyeon Shin, Brent Smith, Keely Snider-Finkelstein, Amber Straughn, Steve Swink, Kaz Tamura, Todd Veach, Kim Ward-Duong, Matthew Wegner, Rebecca Wershba-Kessler, Caleb Wheeler, Shawn White, and Haojing Yan.

I also want to thank my many friends in the indie video games, maker, and media arts and sciences communities. It makes me extremely happy to describe my research to folks. I hope that my enthusiastic blathering about quasars and the wonders of the

---

Universe has provided you as much inspiration as your wonderful creations provide me. May our narratives continue to interweave.

Finally, I could not have completed the work herein without the love and encouragement of my family, especially my partner Alisha.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure   Page

Chapter 1

INTRODUCTION

1.1   Quasars and Active Galactic Nuclei

Quasars were first recognized as optically-luminous extragalactic objects by Maarten Schmidt (1963) from the optical spectrum of a variable star-like object coincident with the radio source 3C 273. It was found to have a blue non-thermal power-law continuum spectrum $F_\nu \propto \nu^{0.28}$ (Oke 1963), where F_ν is the flux density per unit frequency and ν is the frequency. The power-law spectrum was punctuated by extremely broad emission lines (equivalent widths of $\simeq$50 Å), indicating emission from ionized gas with extremely high velocities ($\simeq$10,000 km s^{-1}). Schmidt identified emission lines as the Balmer series of Hydrogen, with a systematic redshift of $z = 0.158$, confirming its nature as an extremely luminous extragalactic object.

It was found that the dominant optical emission in quasars comes from an extremely compact region located in the nucleus of a host galaxy, with this nuclear emission outshining the starlight of the entire host galaxy, such that the system appears point-like in optical images (Kristian 1973). This, combined with the similarity to spectral lines observed in Seyfert galaxies (Seyfert 1943), identified quasars as the most luminous class of active galactic nuclei (AGN). The other AGN classes include Seyfert galaxies — spiral galaxies with luminous nuclear sources that do not exceed the host stellar luminosity; radio galaxies — giant elliptical galaxies with strong radio emission (e.g., Miley 1980); blazars (or BL Lac Objects) — galaxies with strongly variable nuclear sources that show no emission or absorption features (Stein et al. 1976; Angel & Stockman 1980); and LINERS (low-ionization nuclear emission

1

Table 1.1: Classification of Active Galactic Nuclei

AGN Class	UV-Optical Continuum	UV-Optical Emission Lines
Radio-loud		
Blazar	Power-law	None
Radio-loud Quasar[a]	Power-law	Very Broad
Broad-line Radio Galaxy	Power-law	Broad
Narrow-line Radio Galaxy	None/Stellar	Narrow
Radio-quiet		
Radio-quiet Quasar[a]	Power-law	Very Broad
Seyfert 1	Power-law	Broad
Seyfert 2	None/Stellar	Narrow
LINER	None/Stellar	Narrow; Low-ionization

[a] Historically, the term quasar was limited to radio-loud sources, with the term quasi-stellar object (QSO) used for those that were radio-quiet. A strong distinction between these terms no longer exists, so the term quasar will be used interchangeably for radio-loud and radio-quiet sources, with radio fluxes noted when relevant.

regions) — galaxies with nuclear low-ionization emission lines from star formation, but with emission-line ratios indicating a non-thermal contribution (Heckman 1980). These classes are summarized in Table 1.1.

A unification scheme for AGN began to emerge as similarities and differences among the various classes were described (e.g., Scheuer & Readhead 1979; Orr & Browne 1982; Heckman et al. 1984b; Antonucci & Ulvestad 1985; Barthel 1989). Summarized by Antonucci (1993), this unification model has AGN activity driven by a central super-massive black hole (SMBH) surrounded by a torus of dusty material, and an accretion disk of infalling material captured from the inner edge of the torus. AGN are broadly classified as radio-loud or radio-quiet, depending upon the level of detected flux at radio wavelengths. Other differences are then attributed primarily to intrinsic luminosity differences and orientation effects, with various emission regions

or mechanisms obscured or enhanced when the SMBH-torus system is observed from different angles. The major components of the AGN unification model are:

- Relativistic jets, which may or may not be detected, and are responsible for radio emission in radio-loud AGN. Sometimes the jets can also be seen at optical and x-ray wavelengths. They are beamed in a narrow opening angle and are also responsible for the strongly variable and dominant optical continuum in blazars.

- The accretion disk, producing the strong UV-optical continuum in quasars, Seyfert 1 galaxies, and broad-line radio galaxies. Although non-thermal in shape (not a black body spectrum), the consensus model has the emission produced thermally, but at a range of temperatures as a function of radius from the SMBH (i.e., a "sum of black bodies" model, Shields 1978; Malkan & Sargent 1982; Kishimoto et al. 2008)

- The broad-line region, high-velocity clouds of ionized gas close to the SMBH. These clouds are responsible for the broad emission lines in quasars, Seyfert 1 galaxies, and broad-line radio galaxies

- The narrow-line region, lower-velocity clouds of ionized gas farther from the SMBH. These clouds are responsible for the prominent narrow emission lines in Seyfert galaxies and narrow-line radio galaxies.

- An optically-thick dusty torus, which obscures the accretion disk and broad-line region when the torus is viewed from significantly off-axis, as occurs for Seyfert 2 galaxies and narrow-line radio galaxies. The inner surface of the torus also reflects some emission from the broad-line region and disk, causing the polarized spectra of Seyfert 2 nuclei to resemble Seyfert 1 nuclei.

Relativistic jets in this scheme are the only *optional* component, with all the other components assumed to be present in all AGN, and dominant or obscured based on the viewing angle and luminosity. Quasars are then intrinsically luminous AGN that are viewed nearly face-on (unobscured), marked by strong, very broad emission lines and a nuclear point source continuum that is more luminous than the integrated starlight of the host galaxy.

1.2 Black Hole Growth and Mass Estimates

An outstanding question in AGN research is the mechanism by which the SMBH gains its mass. In the local Universe, there is a strong correlation between SMBH mass and the stellar mass of the kinematically hot spheroid component of the host (the central bulge in spiral galaxies, or the entire galaxy in elliptical galaxies): the so-called $M_{BH} - M_{bulge}$ relation (Kormendy & Richstone 1995; Magorrian et al. 1998; Marconi & Hunt 2003; Häring & Rix 2004; Peterson et al. 2004). This suggests a physical relation between bulge star formation and black hole mass buildup, i.e., the black hole and bulge may be coeval and grow in lockstep.

The physical process causing this relation remains uncertain, with several mechanisms suggested that assume a direct physical link. The buildup of stellar mass in the bulge requires gas for star formation, so it has been suggested that the same gas source may also feed the SMBH (e.g., Sanders et al. 1988). Other possible sources of black hole growth include cold mode accretion, where cold gas from the intergalactic medium is accreted directly (Dekel et al. 2009), and black hole mergers (Volonteri & Rees 2006; Li et al. 2007). It has even been recently suggested that the local relationship may arise from an initially uncoupled population (Peng 2007; Jahnke & Macciò 2011). It is likely that all of the above processes play some role in black hole mass buildup, whether or not they are directly responsible for the $M_{BH} - M_{bulge}$ relation.

The physical accretion happens via the thin accretion disk, which is limited by the Eddington rate, where radiation pressure from the accreting gas balances gravitational infall. The maximum black hole growth rate is then proportional to the black hole mass, with the total mass growing exponentially with time (Springel et al. 2005a). For seed black holes with masses $10^2 - 10^4$ $M_\odot$, the growth time for a 10^9 $M_\odot$ black hole ranges from $0.45 - 2.0$ Gyr depending upon the radiative efficiency (Volonteri & Rees 2006). Schwarzschild black holes have lower efficiencies and shorter growth times, while disk-accreting black holes with angular momentum have higher efficiencies and longer growth times.

Black hole masses for local AGN are measured using a technique known as reverberation mapping, currently the most broadly-applicable technique for accurately measuring SMBH masses (Blandford & McKee 1982; Peterson et al. 2004). Reverberation mapping uses time delays between continuum and emission-line variability to measure the spatial extent of the line-emitting ionization regions, from which rotation curves can be calculated. These masses have been further used to calibrate scaling relationships between AGN emission-line widths and SMBH masses, allowing estimation of so-called "virial masses" from single-epoch spectra (Wandel 1999; Vestergaard 2002; McLure & Jarvis 2002; McLure & Dunlop 2004; Greene & Ho 2005; Vestergaard & Peterson 2006; McGill et al. 2008; Vestergaard & Osmer 2009; Wang et al. 2009).

Since the bulge represents the old stellar component of nearby galaxies, observations at high redshift have the unique ability to distinguish between theoretical models for the $M_{BH} - M_{bulge}$ relation. Studies of quasar spectra at high redshift (e.g., Barth et al. 2003; Iwamuro et al. 2004; Shen et al. 2011; Mortlock et al. 2011) have inferred black hole masses of $\gtrsim 10^9$ $M_\odot$ for the most luminous objects. In the current concordance ΛCDM cosmology (e.g., Komatsu et al. 2011; Hinshaw et al. 2013), this means that these billion solar mass black holes have been built up in a few hundred

million years following the formation of the first stars (Bromm et al. 2002). Hierarchical formation models (Volonteri & Rees 2006; Li et al. 2007) are just barely able to reproduce these black hole masses within the required time, with the stellar masses of their host galaxies exceeding 10^{12} $M_\odot$, roughly matching the local $M_{BH} - M_{bulge}$ relation. If the host galaxies truly have as much stellar mass as predicted, they should be among the brightest galaxies at $z = 6$.

1.3 The Quasar-Merger Connection

There is reason to suspect gas-rich galaxy mergers in particular as the trigger for quasar activity. Quasar hosts undergoing mergers are well-documented in the literature (e.g., Brotherton et al. 1999; Stockton et al. 1999; Canalizo et al. 2000; Canalizo & Stockton 2000a,b; Bennert et al. 2008). The collision of two galaxies significantly disturbs their gravitational potentials, which provides a mechanism for large amounts of gas to lose angular momentum, and fall toward the nucleus where it can subsequently accrete onto the SMBH (e.g., Toomre & Toomre 1972; Heckman et al. 1984a; Di Matteo et al. 2005; Springel et al. 2005b). Models for AGN accretion generally have their absolute luminosity be a function of accretion rate (Springel et al. 2005b), so quasars, as the most luminous AGN class, require the highest accretion rates and thus largest quantities of infalling gas. Triggered star formation is also well-documented in gas-rich mergers (e.g., Larson & Tinsley 1978; Soifer et al. 1984; Keel et al. 1985; Lawrence et al. 1989; Duc et al. 1997), providing a mechanism for the buildup of stellar mass in lockstep with the black hole.

Similarities have particularly been noted between quasar host galaxies and ultra-luminous infrared galaxies (ULIRGs), those galaxies with integrated far-infrared luminosities $> 10^{12}$ $L_\odot$ (rest-frame $8 - 1000$ μm, Sanders et al. 1988). All ten of the ULIRGs studied by Sanders et al. (1988) were found to be strongly interacting merger

6

systems with non-thermal ionization and near-infrared colors characteristic of AGNs, suggesting that they may represent obscured or early-stage AGNs. More recent studies, such as the Quasar and ULIRG Evolution Study (QUEST, Schweitzer et al. 2006; Netzer et al. 2007; Schweitzer et al. 2008; Veilleux et al. 2009), have strengthened this connection, finding significant indications of star formation in local quasar hosts, especially those with the largest FIR luminosities.

1.4   Direct Imaging of Quasar Host Galaxies

Ground-based studies have successfully detected the host galaxies of the nearest quasars (e.g., Boroson & Oke 1982; McLeod & Rieke 1994; Taylor et al. 1996; Jahnke et al. 2004a, 2007), but are limited by atmospheric seeing and show that observing the underlying stellar populations of high-redshift quasars is not trivial. The highest redshift for which underlying stellar populations have been unambiguously identified in data obtained from ground-based telescopes is $z \simeq 4$ (McLeod & Bechtold 2009; Targett et al. 2012). Targett et al. (2012) found that these galaxies are extremely compact (half-light radii of $\simeq 1.8$ kpc), much smaller than quasar hosts at lower redshift (e.g., Dunlop et al. 2003). While providing important data regarding the integrated host stellar populations and luminosities, the hosts in high-redshift ground-based studies are generally marginally resolved and do not provide detailed structural information.

The Hubble Space Telescope (HST), with its high on-orbit angular resolution and stable point spread function, has been instrumental in examining the detailed spatial distributions of quasar host stellar populations. These studies have led to important constraints on the morphology, luminosity, and stellar populations of quasar hosts (e.g., Bahcall et al. 1994, 1995b,a; Disney et al. 1995; McLeod & Rieke 1995; Bahcall et al. 1997; McLure et al. 1999; Keeton et al. 2000; Kukula et al. 2001; McLeod

& McLeod 2001; Percival et al. 2001; Ridgway et al. 2001; Hutchings et al. 2002; Dunlop et al. 2003; Marble et al. 2003; Floyd et al. 2004; Jahnke et al. 2004a,b; Kuhlbrodt et al. 2004; Peng et al. 2006; Zakamska et al. 2006; Urrutia et al. 2008; Cales et al. 2011). The high sensitivity of HST's newest imaging camera, the Wide Field Camera 3 (WFC3), and low on-orbit near-infrared sky background make it the ideal instrument for extending detailed structural studies of quasar host galaxies to high redshifts.

There are several techniques that may be employed for high-contrast imaging, the general term for attempting to detect flux from a relatively faint source that is spatially close to a relatively bright source. The classic method for doing this is with a coronagraph, where a Lyot stop physically blocks incoming light from the bright source. This technique has seen limited use for quasar host galaxy studies (Martel et al. 2003), since the stop size must be extremely small or the quasar nearby to avoid blocking a significant fraction of the host galaxy light as well. For recent studies of stellar companion objects such as exoplanets, the LOCI algorithm developed by Lafrenière et al. (2007) has proven effective at finding faint companions. However, this algorithm relies upon multiple matched PSFs (such as from multiple roll angles) and thus is generally too costly in observing time for use in high-redshift quasar studies. The preferred method for most quasar host studies is then direct subtraction, where a reference star is observed to characterize the instrument point spread function (PSF), then used to subtract a model of the quasar point source. There are many details that must be considered when choosing the best PSF star; these will be discussed briefly in Chapter 2 and in-depth in Chapter 3.

1.5 Overview of the Studies Presented Herein

The main goals of the new studies described in this Dissertation are twofold. First: to search for signatures of stellar emission from the host galaxies of $z \simeq 6$ quasars, to better understand mechanisms for SMBH growth and how star formation relates to AGN activity at the cosmic dawn, from the epoch of first light to the end of reionization. Second: to examine quasar feeding mechanisms by measuring the merger fraction for quasar host galaxies at $z \simeq 2$, currently the oldest cosmic epoch for which rest-frame optical emission can be imaged at HST resolution.

Chapter 2 presents a first attempt to image a $z \simeq 6$ quasar host with WFC3IR, and constraints placed on the host galaxy SED, confirming high UV extinction like in local ULIRGs. Chapter 3 introduces a new Markov Chain Monte Carlo approach to modeling AGN+host galaxy light distributions, unique in its careful handling of uncertainty in the supplied PSF. Chapter 4 introduces a sample of nineteen $z \simeq 2$ quasars imaged with WFC3IR, and a study of the host galaxy merger fraction, finding a merger fraction almost double that of galaxies at $z \simeq 2$ not hosting quasars. Chapter 5 introduces a continuing study of quasar hosts at $z \simeq 6$, including detection of a candidate major-merger host system for the quasar NDWFS J1425+3254. Finally, Chapter 6 presents concluding remarks and a discussion of avenues for future work.

Throughout this Dissertation, a concordance ΛCDM cosmology is adopted with $H_0 = 70 \; \mathrm{km \; s^{-1} \; Mpc^{-1}}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$ (Komatsu et al. 2011; Hinshaw et al. 2013). Unless otherwise stated, all magnitudes use the AB system (Oke 1974) and have been corrected for Galactic extinction using the map of Schlegel et al. (1998).

The work presented in Chapter 2 has been previously published in the Astrophysical Journal Letters. It appeared as Mechtley, M., et al. 2012, ApJ, 756, L38, © and published by the American Astronomical Society.

Chapter 2

DIRECT IMAGING OF THE HOST GALAXY OF SDSS J1148+5251

2.1   Background and Introduction

This chapter describes near-infrared imaging and analysis of the $z = 6.42$ quasar SDSS J114816.64+525150.3 (hereafter J1148+5251) with the Hubble Space Telescope (HST) Wide Field Camera 3 (WFC3) infrared channel, and methods for characterizing and subtracting the instrument and telescope Point Spread Function (PSF). This pilot program represents the first attempt to apply PSF subtraction methods to quasar host galaxies in WFC3IR images.

J1148+5251 is the best-studied member of the $z \simeq 6$ quasar population, having been extensively observed at multiple wavelengths since its discovery by Fan et al. (2003). Near-infrared spectroscopy by Willott et al. (2003) and Iwamuro et al. (2004) measured the Mg II and Fe II features, estimating a mass of 3×10^9 $M_\odot$ for the SMBH, an accretion rate near the Eddington limit, and an Fe II/Mg II ratio consistent with quasars at lower redshifts. Radio observations of CO lines (e.g., Walter et al. 2003; Riechers et al. 2009) indicate the presence of $2.2 - 2.4 \times 10^{10}$ $M_\odot$ of high-excitation molecular gas extending to $\simeq 2.5$ kpc ($r = 0\rlap{.}''45$). Studies of the [C II] line at $158\mu m$ (Maiolino et al. 2005; Walter et al. 2009) provide evidence that the quasar host galaxy is undergoing a vigorous starburst, with an estimated star formation rate density of $\simeq 1000$ $M_\odot$ yr^{-1} kpc^{-2} extending over kiloparsec scales. Studies of the continuum emission at far-infrared (FIR) wavelengths indicate a warm dust component with an AGN-corrected FIR luminosity of 9.2×10^{12} $L_\odot$ (Wang et al. 2010) and corresponding dust mass of $4.2 - 7.0 \times 10^8$ $M_\odot$ (Bertoldi et al. 2003; Robson et al. 2004; Beelen et al.

2006). Locally, this is in the range of ultra-luminous infrared galaxies (ULIRGs, galaxies with $L_{FIR} > 10^{12} L_\odot$). Near- and mid-infrared Spitzer Space Telescope observations by Jiang et al. (2006) show clear evidence for prominent hot dust within the galaxy. All of these observations argue for a host galaxy with a significant stellar mass component undergoing an extreme episode of star formation.

§2.2 describes new near-infrared WFC3 imaging of J1148+5251 and a nearby star, used for PSF characterization. §2.3 describes a method for subtracting the central point source from the quasar images. Simulations for assessing the reliability of the subtraction method are described in §2.4. Finally, in §2.5, the implications of these results and plans for future investigation are discussed.

2.2 Observations and Data Reduction

The HST observations of J1148+5251 were performed on 2011 January 31 (HST Program ID 12332, PI: R. Windhorst) using the WFC3 IR channel with the F125W (Wide J) and F160W (WFC3 H) filters. Previous programs (e.g., Hutchings et al. 2002) found that the quality of empirical quasar point source subtractions was significantly affected by PSF time variability. This variability is mostly due to so-called "spacecraft breathing" effects, i.e., thermally induced defocus of the HST secondary mirror due to movement of the Optical Telescope Assembly as the telescope goes into and out of Earth shadow (e.g., Hershey 1998). There are two primary sources of thermal variation that were expected to affect these observations — the spacecraft attitude and the orbital day-night cycle.

To minimize thermal variations due to spacecraft attitude, the PSF star was constrained to be within 5° of the target quasar, and we required that the quasar and PSF star observations be collected in contiguous orbits. It was further requested (and granted) that the observations be scheduled immediately following another observa-

11

tion near the same celestial coordinates, thus minimizing thermal equilibration time in the first orbit. The PSF star was selected to match the quasar near-infrared colors $((J - H) = 0.55$ mag, $(H - K') = 0.72$ mag) as closely as possible, to minimize differences in wavelength-dependent PSF features. Many stars with colors similar to the quasar also have SDSS spectra, since they were targeted by SDSS as high-redshift quasar candidates. These spectra were examined where available, to reject obvious spectroscopic binaries or other contaminants. After checking the resulting candidates for HST guide stars, we selected the star 2MASS J11552259+4937342 (spectral type K7, $m_J = 16.272 \pm 0.079$ Vega mag, $(J - H) = 0.488 \pm 0.158$ mag, $(H - Ks) = 0.651 \pm 0.176$ mag, Skrutskie et al. 2006) as the final PSF target.

Unfortunately, there is no way to eliminate thermal variation due to the orbital day/night cycle. Thus, the orbits were constructed to ensure the cycle was fully sampled, and that equal fractions of the final quasar and PSF star images would come from a given location in orbital phase. Each quasar exposure was matched by several shorter PSF star exposures, taken at the same sub-pixel dither point and phase within the orbit. This phase-space sampling is summarized in Figure 2.1. No detector location was exposed beyond half-well depth in a single orbit, thus avoiding saturation or detector persistence. The quasar was observed in this pattern for three orbits, and the PSF star was observed for a single orbit. [1] Total exposure times for each combination of object and filter are summarized in Table 2.2. Analysis was performed on `Multidrizzle`-combined images (Koekemoer et al. 2002, 2011) with an output pixel scale of 0.''065, to achieve Nyquist sampling of the PSF in both filters, enabling accurate spatial shifting. The cosmic ray rejection step of `Multidrizzle` was disabled, since the `MULTIACCUM` readout mode provided sufficient cosmic ray rejection.

---

[1]An additional PSF star, 2MASS J11403198+5620582, was also observed for a single orbit to allow for inter-orbit interpolation of the PSF measurement. This observation suffered from a poor guide star acquisition and was unusable.
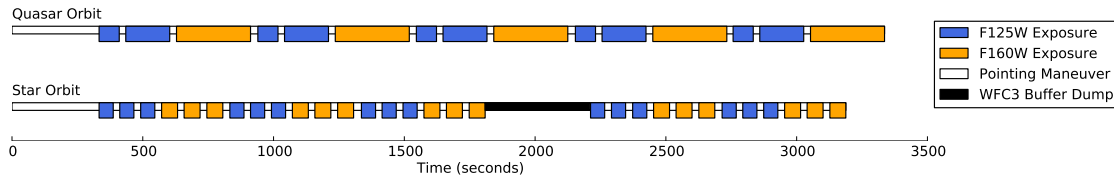
Figure 2.1: PSF star and one quasar orbit, highlighting the relative phasing of corresponding dither points to compensate for spacecraft breathing. Exact phase matches are not possible due to buffer dumps and specific readout sequences. Corresponding dither points were centered at similar positions on the detector to account for field-dependent PSF variability.

Table 2.1: Exposure summary for J1148+5251

Target	Filter	Exp. Time (s)	S/N
SDSS J1148+5251	F125W	2478	2400
SDSS J1148+5251	F160W	3646	3760
2MASS J11552259+4937342	F125W	208	1730
2MASS J11552259+4937342	F160W	335	2200

2.3 Point Source Subtraction

The software `GalFit` (Peng et al. 2002, 2010) was used to fit a PSF single-component model to the quasar image. The `Multidrizzle`-generated weight maps were transformed into uncertainty maps as in Dickinson et al. (2004), including the effects of correlated noise and shot noise from the quasar, and these were supplied to `GalFit` as the pixel-to-pixel uncertainty ("sigma") image. The best-fit model was then subtracted from the original image, and the residual inspected.

This subtraction was first attempted using the image of the PSF star as the model PSF. The results are shown in Figure 2.2. The residuals were measured using a $0\!\!''\!5$ radius aperture, obtaining upper limits of $m_J > 22.8$ mag, $m_H > 23.0$ mag ($2\,\sigma$). This includes the total noise contribution from both the quasar and the empirical PSF, measured by scaling the PSF uncertainty map by the same factor as in the fit, and adding it in quadrature to the quasar uncertainty map. The noise contribution

13

Figure 2.2: Empirical PSF subtraction. Left panels: `Multidrizzle`-combined F125W (J, top) and F160W (H, bottom) WFC3 images of J1148+5251 after sky subtraction. Pixels are $0\rlap{.}''065$. Middle Panels: Scaled and shifted PSF star, as fit by `GalFit`. Right panels: Fit residuals, showing a net positive residual flux, but high noise. We measured the integrated residual flux using a $0\rlap{.}''5$ radius aperture (gray circle), obtaining upper limits of $m_J > 22.8$ mag, $m_H > 23.0$ mag ($2\,\sigma$). All images are displayed with the same logarithmic stretch.

from the subtracted PSF is comparable to that of the quasar since the two images have comparable S/N (see Table 2.2), which leads to cosmetic defects (holes) in the subtraction, despite the net positive residual.

We also generated a `TinyTim` [2] (Krist et al. 2011) model of the PSF, which was calibrated to the PSF star observations. A $5\times$ spatially oversampled `TinyTim` model was generated for each WFC3 exposure to allow for subpixel shifting. The spectrum of J1148+5251 obtained by Iwamuro et al. (2004) was used as the model spectrum. The observatory pointing accuracy data (jitter files) for each exposure were included

---

[2] http://www.stsci.edu/hst/observatory/focus/TinyTim

in the models. The HST Focus Model [3] (Cox & Niemi 2011) was used to estimate the secondary mirror despace for each exposure, and the field-dependent coma and astigmatism measurements built into `TinyTim` were included. The individual HST detectors have different mean focus offsets in the Focus Model, but the offset for the WFC3 IR channel has not been characterized. The mean $Z4$ Zernike coefficient in `TinyTim` (R_2^0 in the original formulation of Zernike 1934) was therefore allowed to float as a free parameter in the optimization, which was then added to the Focus Model estimate for each exposure. These models for individual exposures were then combined, weighted by exposure time, to produce a composite PSF for each `Multidrizzle`-combined science image. `GalFit` also accepts a pixel response convolution kernel for oversampled PSFs. This kernel was generated by drizzling copies of the empirical WFC3 pixel response convolution kernel (modeling inter-pixel capacitance and jitter, see Hartig 2008) using the same shifts applied to the real images.

The result of the `TinyTim` PSF subtraction, with a significantly reduced noise floor compared to the direct subtraction, is shown in Figure 2.3. No host galaxy is detected, to a limiting surface brightness from $r = 0\farcs3$ to $0\farcs5$ radius of $\mu_J > 23.5$, $\mu_H > 23.7$ mag arcsec^{-2} ($2\,\sigma$). The inner $0\farcs3$ was excluded from the fit, as the best-fit `TinyTim` models produce PSF cores that are consistently narrower than those observed, despite the inclusion of the observatory pointing stability data. Visible residual structures (diffraction spikes and spots) are also seen when subtracting this model from the PSF star observations.

2.4   Host Galaxy Simulations

Having established no host galaxy detection using the `TinyTim` PSF, we sought to quantify this subtraction method's ability to recover the host galaxy flux as a function
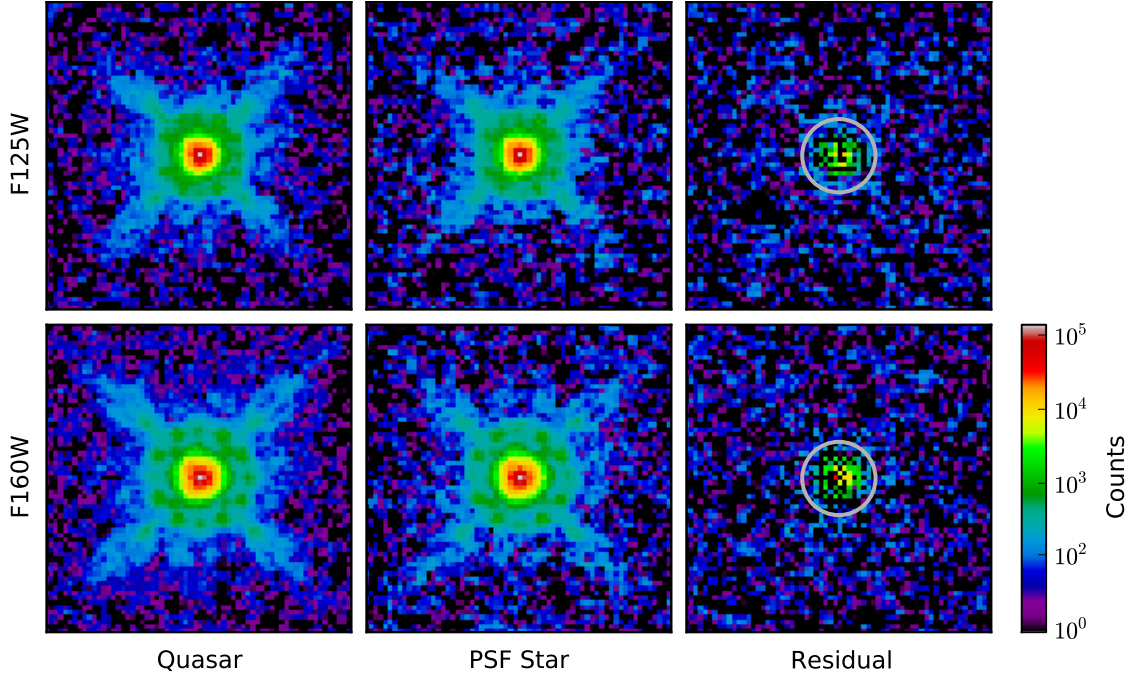
---

[3]http://www.stsci.edu/hst/observatory/focus/FocusModel
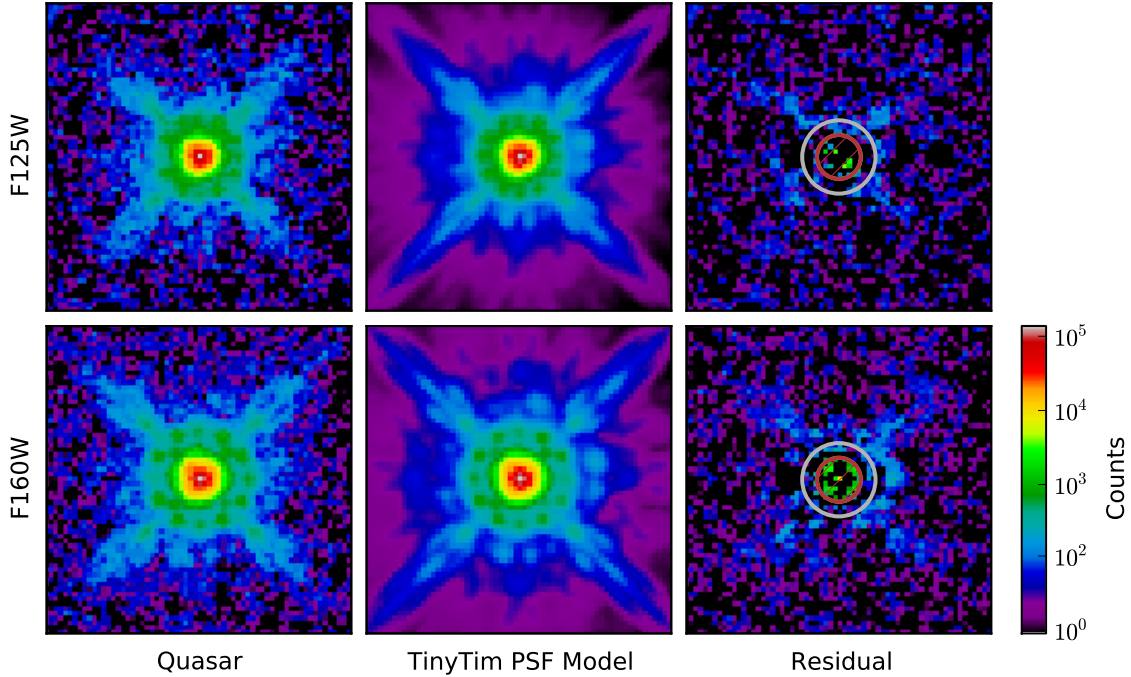
Figure 2.3: Model PSF subtraction. Left panel: `Multidrizzle`-combined F125W (J, top) and F160W (H, bottom) WFC3 images of J1148+5251. Middle Panels: `TinyTim` models of the quasar point source, constructed by optimizing parameters for the PSF star observations, then scaled and shifted by `GalFit`. Right panel: Fit residuals, showing no significant detection of the underlying galaxy beyond $0\farcs3$ radius. The over-subtracted flux in the central $0\farcs3$ (inner circle) occurs because the best-fit model PSFs have more power in the central peak than the observations, and is also seen in residuals when modeling the PSF star. This region was excluded from the fit. The noise floor in the residual panel is 40% that of the residual panel in Figure 2.2. From $r = 0\farcs3 - 0\farcs5$ (between inner and outer circles) we measure a limiting surface brightness of $\mu_J > 23.5$, $\mu_H > 23.7$ mag arcsec^{-2} ($2\,\sigma$). The noise in the quasar image and uncertainties in the PSF model contribute roughly equally within this region. Visible residual structures (diffraction spikes and spots) are also seen when subtracting this model from the PSF star image.

of host galaxy parameters. To do so, `GalFit` was used to simulate a point source along with a Sérsic profile host galaxy, with total flux adding up to $m_J = 19.1$ mag, the measured flux in F125W. This simulated image contained no noise, so both shot noise from the object and a Gaussian noise field drizzled in the same manner as the real quasar image were added, to match its correlated noise properties. The same analysis that was used on the real quasar image was then performed, using `GalFit` to subtract a `TinyTim`-generated point source and measuring the surface brightness from $r = 0\!''\!3$ to $0\!''\!5$ in the residual image.

A grid of 256 models was run using this technique, varying the total integrated flux of the host galaxy from $m_J = 20 - 26$ mag, the effective radius from $r_e = 0\!''\!1 - 0\!''\!9$, and Sérsic indexes $n = 1.0$ and 4.0. The magnitude range represents host galaxies with luminosities from $\simeq 1/2$ to $1/500$ of the total quasar luminosity. Fainter host galaxies than this are undetectable due to shot noise from the point source. The range in effective radius corresponds to $r_e \simeq 0.6 - 5.0$ kpc at $z = 6.42$.

Figure 2.4 summarizes the results of these simulations, plotting the measured surface brightness from $r = 0\!''\!3 - 0\!''\!5$ and contours representing the 1, 2, and $5\,\sigma$ detection limits. Inspecting the residuals of these model subtractions, it was found that bright ($m_J < 22.5$ mag), compact ($r_e < 0\!''\!3$) host galaxies cause the method to significantly over-subtract the PSF. This would show negative residuals from the diffraction spikes, which are not seen in Figure 2.3.

The model surface brightness from $r = 0\!''\!3 - 0\!''\!5$ reaches the $2\,\sigma$ upper limit of $\mu_J > 23.5$ mag arcsec^{-2} for a host galaxy of $m_J > 22 - 23$ mag, depending upon Sérsic index and effective radius.
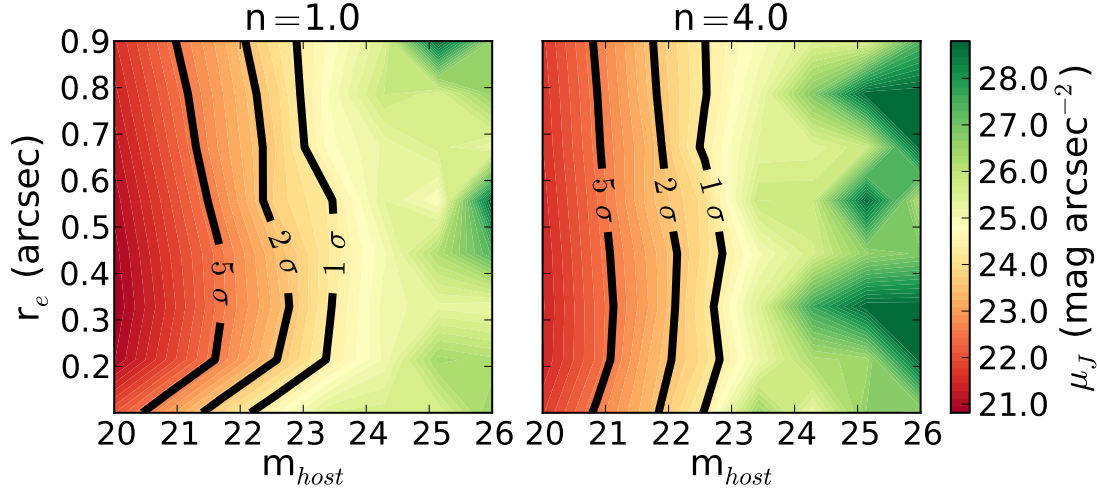
Figure 2.4: PSF subtraction on simulated host galaxies. Surface brightness predicted from $r = 0\overset{\prime\prime}{.}3 - 0\overset{\prime\prime}{.}5$ as a function of simulated host galaxy parameters. The simulated observation is approximated as a PSF component along with a Sérsic profile, with total flux adding up to the observed m_J from the WFC3 image. The integrated magnitude (m_{host}), effective radius (r_e), and Sérsic index ($n = 1.0$, left panels and $n = 4.0$, right panels) of the Sérsic profile are varied with each model. The 1, 2, and 5σ detection significance levels are plotted as black lines. The measured surface brightness reaches the 2σ limit ($\mu_J > 23.5$ mag arcsec^{-2}) for a host galaxy of $m_J > 22 - 23$ mag, depending upon Sérsic index and effective radius.

2.5   Discussion

Point source subtraction was performed on the $z = 6.42$ quasar J1148+5251, with both empirical and modeled PSFs. Using direct subtraction, an upper limit of $m_J > 22.8$ mag, $m_H > 23.0$ mag (2σ) was measured. With the modeled PSF subtraction, a limiting surface brightness was measured from $0\overset{\prime\prime}{.}3 - 0\overset{\prime\prime}{.}5$ of $\mu_J > 23.5$ mag arcsec^{-2}, $\mu_H > 23.7$ mag arcsec^{-2} (2σ). Performing the same subtraction method on simulated quasars, this surface brightness limit was found to correspond to a host galaxy of $m_J > 22 - 23$ mag, consistent with the direct subtraction limit.

Using the direct subtraction limits, the upper limits on the rest-frame monochromatic luminosity (λL_λ) at 1700 Å and 2200 Å are $L_{1700} < 8.4 \times 10^{11} \ L_\odot$ and $L_{2200} < 5.4 \times 10^{11} \ L_\odot$, assuming a flat spectrum within each band when applying

the K-correction (Oke & Sandage 1968). This is comparable to the most luminous Lyman break galaxies at $z \simeq 2 - 3$ (Hoopes et al. 2007).

Using the upper limits for the host galaxy flux and Equation 1 from Kennicutt (1998), which relates L_ν to star formation rate, a star formation rate of SFR < $210 - 250\ M_\odot\ \mathrm{yr}^{-1}$ is estimated. This estimate ignores dust attenuation and assumes a continuous star formation rate over 10^8 years or longer. A younger population would decrease this upper limit, while dust would allow for a higher (absorption-corrected) rate. The star formation rate estimated from the AGN-corrected FIR luminosity by Wang et al. (2010) is 2380 $M_\odot\ \mathrm{yr}^{-1}$. Since J1148+5251 would be classified as a ULIRG locally, this discrepancy is likely due to significant UV absorption by dust.

The infrared excess (IRX) of the host galaxy can also be constrained, defined as the infrared to far-ultraviolet (FUV) luminosity ratio L_{IR}/L_{FUV} (e.g., Howell et al. 2010), usually expressed in logarithmic units. Using the upper limit for L_{1700} and an AGN-corrected infrared luminosity $L_{IR} = 9.2 \times 10^{12}\ L_\odot$ (Wang et al. 2010) implies $\log(IRX) > 1.0$, consistent with local luminous infrared galaxies (LIRGs) and ULIRGs (Howell et al. 2010), but greater than local starburst galaxies and high-redshift Lyman break galaxies (Overzier et al. 2011).

Figure 2.5 plots broad-band measurements for J1148+5251 taken from Fan et al. (2003); Iwamuro et al. (2004); Jiang et al. (2006); Beelen et al. (2006); Robson et al. (2004), and Bertoldi et al. (2003). Also plotted are the upper limits for the host galaxy flux at 1700 Å and 2200 Å from this work, and the AGN-corrected FIR measurements of Wang et al. (2010). Also plotted are the spectral energy distributions (SEDs) of four local galaxy systems — three LIRGs (Arp 220, IRAS 22491-1808, and IC 1623), representing the range in IRX from Howell et al. (2010), and the star-forming spiral NGC 4631, representing a galaxy with $\log(IRX) < 1.0$, which is thus excluded as a

potential host. The photometric points for the local galaxies were taken from NED [4] and SEDs have been normalized to match the AGN-corrected emission of J1148+5251 between 40 and 200 microns.

Using the relation between the IRX flux ratio and A_{FUV} (e.g., Overzier et al. 2011, Equation 1) provides an estimate of $A_{FUV} > 2.1$ mag of UV absorption. Using the empirical relation $IRX_{M99,inner}$ ($A_{FUV} = 4.54 + 2.07\beta \pm 0.4$) from Overzier et al. (2011) gives a limit of $\beta > -1.2 \pm 0.2$. This matches local (U)LIRGs (Howell et al. 2010), but is redder than almost all local star-forming galaxies and $z \simeq 6$ Lyman break galaxies (Overzier et al. 2011; Bouwens et al. 2012).

The `TinyTim`-based subtraction may be improved in the future with more accurate WFC3 IR PSFs. Since uncertainties introduced by the PSF model scale with PSF brightness, our further WFC3 observations target quasars where the contrast ratio between point source and host galaxy is expected to be smaller, such as optically faint $z \simeq 6$ quasars with large FIR luminosities. While J1148+5251 is too far north to be observed with the Atacama Large Millimeter Array (ALMA), future observations with the Combined Array for Research in Millimeter-wave Astronomy (CARMA) or the upgraded Plateau de Bure interferometer may be able to provide additional morphological constraints. The James Webb Space Telescope will enable use of the PSF subtraction method at rest-frame ultraviolet and optical wavelengths with better-sampled empirical PSFs in a more stable thermal environment.

---

[4]The NASA/IPAC Extragalactic Database (NED) is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.
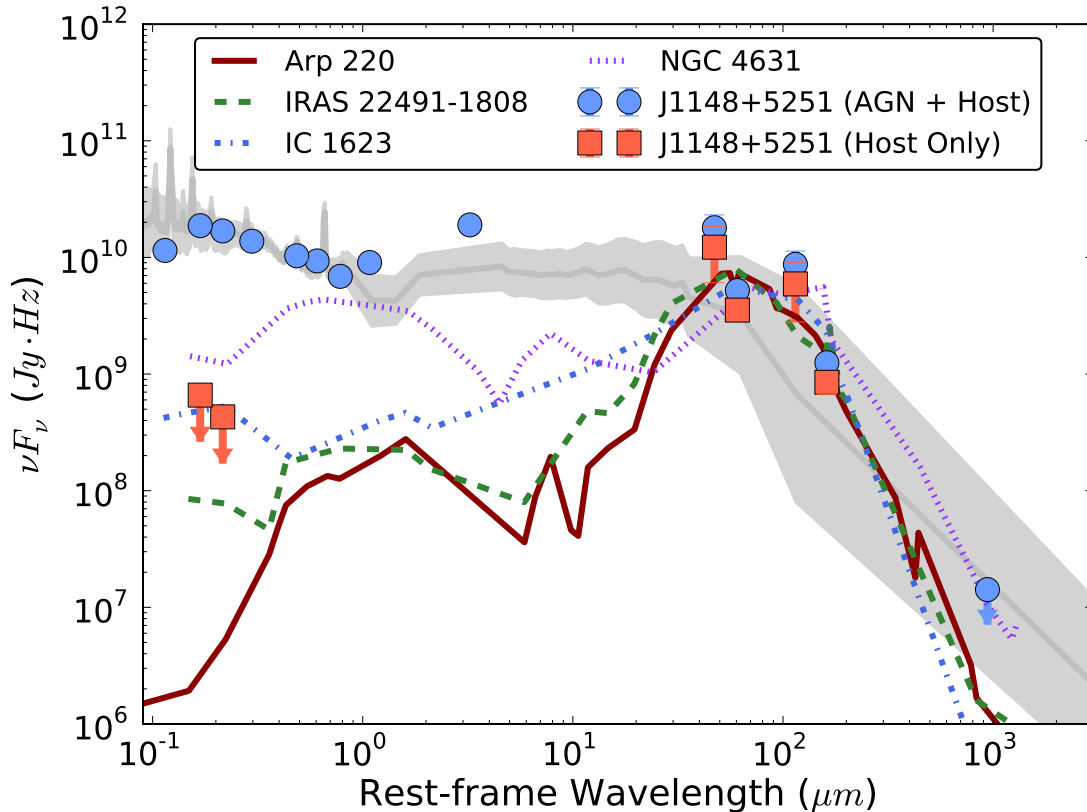
Figure 2.5: Comparison of broad-band photometry for J1148+5251 to local galaxies. Blue circles are broad-band (quasar and host) photometry of J1148+5251 taken from Fan et al. (2003); Iwamuro et al. (2004); Jiang et al. (2006); Beelen et al. (2006); Robson et al. (2004), and Bertoldi et al. (2003). Red squares show upper limits for the host galaxy flux at 1700 Å and 2200 Å from this work, and the AGN-corrected FIR measurements from Wang et al. (2010). The light gray spectrum is the average radio-quiet quasar spectrum of Shang et al. (2011), normalized to J1148+5251 from $0.1 - 1\mu m$. The dotted purple SED is NGC 4631, a local spiral with $\log(IRX) < 1.0$. Other SEDs are those of the local LIRGs Arp 220 (solid red, $\log(IRX) = 3.423$), IRAS 22491-1808 (dashed green, $\log(IRX) = 2.198$), and IC 1623 (dot-dashed blue, $\log(IRX) = 1.379$), representing high, average, and low IRX LIRGs, respectively (Howell et al. 2010). The local galaxy SEDs have been normalized to match the AGN-corrected emission of J1148+5251 between rest-frame 40 and 200 microns. The constraint of $\log(IRX) > 1.0$ (and the 2200 Å flux limit) matches most local LIRGs, but is greater than almost all local star-forming galaxies and high-redshift Lyman break galaxies (Howell et al. 2010; Overzier et al. 2011).

Chapter 3

TWO-DIMENSIONAL SURFACE BRIGHTNESS MODELING WITH THE

MARKOV CHAIN MONTE CARLO METHOD

3.1 Introduction and Motivation

The experience with the J1148+5251 pilot program, as described in Chapter 2, was particularly instructive in two respects. First, for further studies of $z \simeq 6$ quasars with the WFC3 IR channel, the likelihood of a host galaxy detection could be improved by selecting targets with intrinsically fainter AGN. This new quasar sample and preliminary results will be detailed in Chapter 5. Second, there were fundamental limitations to currently-existing fitting algorithms for two-dimensional surface brightness distributions, such as `GalFit`, when applied to the point source subtraction problem. To address the latter problem, I developed a new two-dimensional fitting algorithm, `psfMC`, that is purpose-built for the quasar point source subtraction problem. This chapter details the two-dimensional fitting algorithm.

`GalFit` (Peng et al. 2002, 2010) represents the current state-of-the-art for two-dimensional fitting of galaxy surface brightness distributions. Although other software exist for two-dimensional image modeling (e.g., Shaw & Gilmore 1989; Byun & Freeman 1995; de Jong 1996; Simard 1998; Wadadekar et al. 1999; Pignatelli et al. 2006), these generally use similar techniques with similar strengths and limitations. A notable exception is `GALPHAT` (Yoon et al. 2011), which uses a Markov Chain Monte Carlo approach to galaxy modeling. It is optimized for fitting individual galaxies, rather than quasar+host models, and still assumes error-free PSFs.
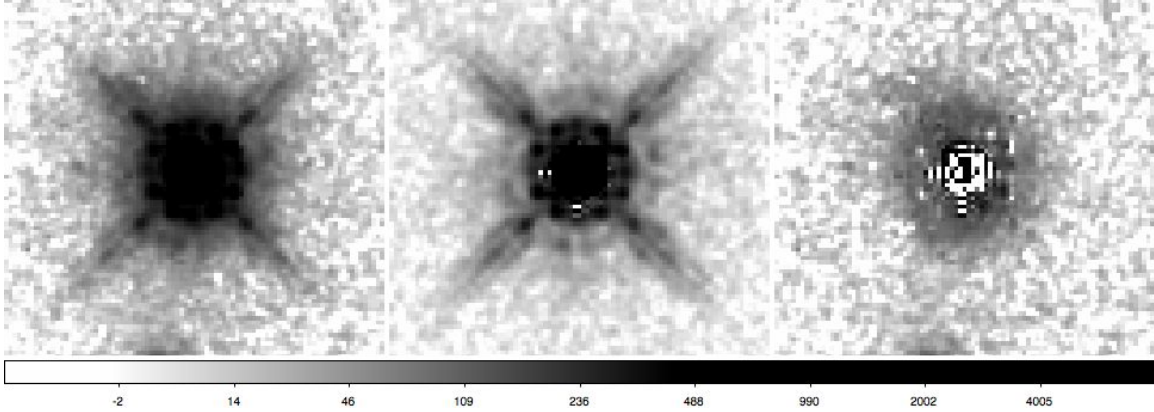
Figure 3.1: Example of point source over-subtraction in a single-component model. Left: WFC3 infrared image of a $z = 2$ quasar in the F160W filter. Middle: Single-component PSF model, generated using `GalFit` from a focus- and color-matched star. Right: residual after subtracting the point source model, showing significant over-subtraction in the center. This is a general problem of quasar point source subtraction, and necessitates simultaneously modeling the host galaxy flux with the point source.

`GalFit` has become the software of choice among these for several reasons. It provides many different surface brightness distribution options (e.g., various radial profiles, generalized ellipses, Fourier modes, various spiral parameterizations, among many others), and allows for convolution with an arbitrary user-supplied point spread function. It implements a standard, computationally efficient algorithm (Levenberg-Marquardt gradient descent) for parameter optimization, which allows rapid fitting of many objects. It is also publicly available and relatively easy to use. However, there are several reasons that `GalFit` is not optimal for quasar point source subtraction.

The simulations performed for J1148+5251 (§2.4) show that single-component point source subtraction methods tend to over-subtract the point source when the host galaxy flux is detectable. This over-subtraction is proportional to the luminosity of the underlying galaxy, and is particularly pronounced for hosts that are comparable in size to the point spread function, such as high-redshift galaxies. An example of such an over-subtraction is shown in Figure 3.1.

23

A solution is to simultaneously model the point source and the underlying host galaxy. Such two-component models (e.g., a point source and a Sérsic profile modeling the coeval stellar population of the galaxy bulge) have fundamentally covariant parameters. For example, the fluxes of the two components are necessarily correlated since the total flux of the two-component model must match the measured flux from the image being fit. `GalFit` does not provide a way to quantify these covariances — error estimates are provided for each fit parameter, but these are based on an ellipsoidal approximation of the surface of constant $\Delta\chi^2$, and the largest projected vector sums of the principal axes of that ellipsoid (see §2.6 in Peng et al. 2002).

`Galfit` achieves some computational efficiency by using the Levenberg-Marquardt method of performing least-squares minimization (e.g., Press et al. 2007). This reduces the total number of samples required in parameter space, which means models converge quickly. However, this particular method (and similar gradient descent methods) find local minima of the objective function (χ^2) in parameter space, rather than the global minimum, which means they can be sensitive to the initial guess of parameter values. [1]   The method also involves calculating a gradient image during each iteration to determine the parameter values to use for the subsequent iteration. Extremely compact models — e.g., a Sérsic profile with small effective radius (r_e) and large index (n) — have all their gradient information contained within a single pixel, and so the Sérsic degrees of freedom are used to fit aberrant pixels due to PSF mismatch, rather than the true host galaxy distribution. This essentially creates a false minimum in parameter space from which the gradient descent cannot escape.

---

[1]The GIM2D two-dimensional modeling software (Simard 1998) uses the Metropolis algorithm (Metropolis et al. 1953) rather than gradient descent, so can escape local minima and is less susceptible to the initial parameter guess. In other respects described in the text, such as lack of parameter covariance information and assuming a PSF with infinite S/N, it is similar to `GalFit`.

A related problem is that `GalFit` assumes that the supplied PSF is without error, and has infinite signal-to-noise ratio. Even without systematic PSF uncertainties (i.e., a PSF exactly matching the telescope focal history, spectral energy distribution of the quasar point source, etc.), the photon or shot noise from the supplied PSF can be large enough to become significant. It is often the case that the most desirable star has S/N comparable to the quasar, such as in the analysis of J1148+5251 presented in Chapter 2. For a star whose S/N exceeds the quasar by a factor of $1.5-10$, this means that when performing the point source subtraction, the PSF contributes $1-20\%$ of the per-pixel RMS error. This is significant when attempting to detect host galaxies that may be $50-100$ times fainter than the point source.

3.2   Bayesian Parameter Estimation

An alternative method of modeling and parameter estimation that addresses some of these problems is Markov Chain Monte Carlo (MCMC, e.g., Gelman et al. 2011, Chapter 11). MCMC is one of many Bayesian simulation methods, meaning that it is motivated by probability theory. Given a set of observed data y, and a model described by a set of parameters θ, the goal is to make inferences about the probability distribution of the model parameters. Bayes' Theorem is given by Equation 3.1:

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)} \tag{3.1}$$

In the above equation, $P(\theta|y)$ is the posterior probability distribution, the probability distribution of model parameters given the observed data (the eventual goal of the simulation). $P(\theta)$ is the *prior* probability distribution of the model parameters, determined e.g., from previous observations or physical first principles. $P(y|\theta)$ is the model *likelihood function*, the probability of the observed data for a given set of model parameter values. $P(y)$ is the prior probability distribution of the observed data, a

25

measure of the overall probability of measuring the particular values for y. This can be taken as a proportionality constant since the observed data do not change as different parameter values are tried (see discussion in Gelman et al. 2011, Chapter 1).

As a Markov chain method, MCMC works by drawing successive samples from the model's parameter space, where the parameter values of the next sample, θ_{t+1}, are based only on the parameter values of the current sample, θ_t, i.e., the Markov property. The selection of samples in the chain is constructed via a *step method* such that with each successive sample, the distribution of sampled points becomes closer to the true posterior distribution $P(\theta|y)$. The most common general-purpose step method is the Metropolis-Hastings algorithm (Metropolis et al. 1953; Hastings 1970). In the basic Metropolis algorithm (a particular case of Metropolis-Hastings), a proposed sample θ_* is drawn from a proposal distribution, such as a multivariate normal distribution centered at the current sample θ_t. The ratio of the posterior probability densities, $r = P(\theta_*|y)/P(\theta_t|y)$, is then calculated and the proposed sample is accepted or rejected based on the Metropolis criterion. If $r \geq 1$, the sample is accepted. Otherwise, the sample is accepted with probability r. Probabilistic acceptance is implemented by generating a uniform random number u between 0 and 1. If $u \leq r$, the sample is accepted, and if $u > r$, the sample is rejected. If a sample is accepted, it is added to the chain as θ_{t+1}. If a sample is rejected, it is discarded entirely and a new proposed sample is drawn, starting again from θ_t.

When the algorithm has finished (usually, when the chain reaches some pre-determined size, or some statistical criteria are met), the result is a series of parameter samples that approximates $P(\theta|y)$, the posterior probability distribution of parameters given the observed data (see Gelman et al. 2011, Chapter 11 for a proof). Samples can then be analyzed statistically to provide insights about the inferred posterior distribution, such as moments, covariance, or multi-modality.

3.3 Description of the `psfMC` Software

`psfMC` is built upon `pyMC` (Patil et al. 2010), a module for performing Bayesian stochastic modeling with the Python programming language (van Rossum & Drake 1990). The computational book-keeping tasks — such as drawing samples, proposing new values via the Markov Chain step method, and saving sample traces — are handled by `pyMC`. The `psfMC` software then allows the user to build models that simultaneously model an arbitrary number of components. At this time, point sources and Sérsic profiles are provided, though additional surface brightness distributions may easily be added. The free parameters for each component (e.g., position, total magnitude, Sérsic index, etc.) can either be supplied as a fixed numeric value or as an arbitrary prior probability distribution. `pyMC` provides many common built-in distributions, and additional distributions can be easily added (e.g., a Schechter luminosity function with a faint-end cutoff as the prior for a galaxy's integrated magnitude). An arbitrary number of PSF images can also be supplied, with the most likely PSF selected based on the data.

The software requires Python version ≥ 2.6, and depends upon only four additional Python modules: The standard scientific packages `numpy` (version ≥ 1.6) and `scipy` (version ≥ 0.10), the `pyMC` module for Bayesian stochastic modeling (version ≥ 2.0), and the `pyfits` module (version ≥ 3.0) for manipulating astronomical FITS format images. Two additional modules are optional: `pyregion` (version ≥ 1.0) to use SAOImage `ds9` region files for masking, and `numexpr` (version ≥ 2.0) to parallelize the computation of Sérsic profiles on multi-core systems.

The user specifies the model components using a simple Python file (see Appendix A for an example), which may include numeric expressions, additional variables, etc. The user then calls the fitting function, supplying values for the following

additional parameters. All images are supplied in the astronomical FITS format (Wells et al. 1981; Pence et al. 2010), either as strings representing a relative path on disk, or as pre-opened `pyfits` HDU objects. Multi-extension FITS (MEF) files will use the first image extension, so it is recommended that the individual extensions be first opened as `pyfits` HDU objects if the data are in MEF format.

- `obs_file` The filename of the observed image

- `obsIVM_file` The filename of the weight (inverse variance) map for the observed image

- `psf_files` The filename of the PSF image or a list of multiple filenames, in which case the fitting process will select the most likely PSF. [2]

- `psfIVM_files` The filename(s) of the weight (inverse variance) map for the PSF image

- `model_file` The filename of the model definition file (see Appendix A)

- `mag_zeropoint` The instrumental magnitude zeropoint used to convert instrumental magnitudes into observed magnitudes. Magnitudes supplied as parameters of the model components should be relative to the same zeropoint.

- `mask_file` (optional) An image, with non-zero pixels denoting exclusion from the fit. Alternatively, an SAOImage `ds9` region file describing which pixels should be included in the fit. Regions can have the exclude flag set to exclude pixels from the fit. Any galaxies not intended to be modeled should be excluded from the fit, see discussion below.

---

[2]Multiple supplied PSFs are currently experimental. The feature will work best when the PSFs are sorted in some logical sequence, such as by estimated focus or measured full width at half maximum. Using this feature may also affect measurements of host galaxy Sérsic parameters, since there is some degeneracy between Sérsic parameters and PSF shape when fitting the two simultaneously.

Weight maps must be provided for both the observed quasar and the PSF. These should be provided in inverse variance ($1/\sigma^2$) form, and should include all noise and uncertainty contributions, including detector read noise, sky noise, and Poisson or shot noise from the object photon counts. For `Multidrizzle`-combined HST images, maps produced by the "ERR" weighting scheme include all these contributions, with the caveat that correlated noise may cause these maps to underestimate the true per-pixel noise (see the discussion of correlated noise with respect to the model variance map below). It should be noted that shot noise from the objects requires careful consideration when working with non-destructive multiple-read CMOS detectors such as the HST WFC3 infrared channel or NICMOS. These generally use an up-the-ramp fitting routine to fit a count rate for each pixel, in order to achieve high dynamic range. This process makes the understanding of shot noise non-trivial. One cannot simply multiply the count rate by the exposure time to get the original counts, because for pixels that saturated (such as the central pixels of the point source), or that were affected by cosmic rays, count rates will be based on fewer non-destructive reads than low count rates. The error data extensions produced by the HST pipeline take this into account, so should be treated as the true per-pixel errors unless the images are re-calibrated manually.

The number of samples to be drawn from the posterior distribution can be specified by the user. An arbitrary number of samples may also be discarded as a *burn-in period*, allowing the Markov chain to converge to a stable region of parameter space before samples are retained for analysis. A *thinning interval* can also be specified, where only every nth sample is retained, to account for the fact that the Markov process produces correlated samples. Since the correlation factor is rarely known *a priori*, and in principal is different for each model parameter, it is recommended that the number of effective samples instead be estimated after fitting (see discussion in

Chapter 4). For more detailed discussions of burn-in and sample thinning in general, see Gelman et al. (2011, Chapter 11). Samples are drawn from parameter space using a Metropolis-Hastings algorithm that updates one parameter value at a time. For most parameters, the basic Metropolis algorithm is used with the user-specified prior as the proposal distribution. For (x, y) positions, the Adaptive Metropolis step method is used (Haario et al. 2001), which quantifies covariance between the individual vector components during the burn-in period and uses a rotated multivariate normal distribution for proposals.

The model is constructed with a simple, flat hierarchy, with the individual parameters acting as children only of the final model images, which in turn determine the likelihood function. Each time a proposed sample is drawn from the parameter space, psfMC generates a model image of the intrinsic surface brightness distribution described by the parameters (hereafter, "raw model"), without the telescope and instrument PSF. The raw model is composed of whichever components were specified by the user in the model definition file (usually a point source and one or more Sérsic components for quasars). This raw model is then used to generate two further images — one convolved with the PSF ("convolved model"), and a model variance map that includes the uncertainty in the supplied PSF.

The model variance map is simply the square of the model image convolved with the PSF variance map. Consider the intensity, $I_{CM}(\boldsymbol{p})$, of a pixel $\boldsymbol{p}$ in the convolved model, which is given by:

$$I_{CM}(\boldsymbol{p}) = (I_{RM} * W_{PSF})(\boldsymbol{p}) = \sum_{\boldsymbol{q}} I_{RM}(\boldsymbol{q}) \cdot W_{PSF}(\boldsymbol{p} - \boldsymbol{q}) \qquad (3.2)$$

The bold-faced variables denote two-dimensional vector-valued image coordinates, i.e., (i, j). The $*$ symbol denotes convolution, and the summation is over all pixels $\boldsymbol{q}$. $I_{RM}(\boldsymbol{q})$ is the raw model intensity, and $W_{PSF}(\boldsymbol{p} - \boldsymbol{q})$ is the PSF weight for pixel

30

$\boldsymbol{p}$ with the kernel centered at pixel $\boldsymbol{q}$. If the kernel weights have associated variances $\sigma_{PSF}^2(\boldsymbol{p}-\boldsymbol{q})$ and are independent, then the uncertainties sum in quadrature, and the variance of pixel $\boldsymbol{p}$ in the convolved model is given by:

$$\sigma_{CM}^2(\boldsymbol{p}) = \left(I_{RM}^2 * \sigma_{PSF}^2\right)(\boldsymbol{p}) = \sum_{\boldsymbol{q}} I_{RM}(\boldsymbol{q})^2 \cdot \sigma_{PSF}^2(\boldsymbol{p}-\boldsymbol{q}) \qquad (3.3)$$

This definition assumes that the individual pixel variances in the PSF are independent (there are no covariance terms when adding the pixel variances). For real data there is some correlated noise between adjacent pixels, whether from charge diffusion in CCD detectors, inter-pixel capacitance in CMOS detectors, or the resampling process in drizzled images (such as all undistorted HST data). For charge diffusion and inter-pixel capacitance, the degree of the correlation is generally small (a few percent of the incident flux, e.g., Hartig 2008). For `Multidrizzle`-combined images, the correlation factor will depend upon the parameters used, particularly the output linear pixel scale and the pixfrac parameter (Koekemoer et al. 2002, 2011). Techniques exist for quantifying correlated noise in drizzled images and propagating this estimate back into the variance map (e.g., measuring the sky autocorrelation function as in Guo et al. 2013). These should be used when working with drizzled data, to ensure that correlated noise is at least understood, if not mitigated.

The conditional probability $P(y(\boldsymbol{p})|\theta)$ of each observed pixel value $y(\boldsymbol{p})$ given the model is then calculated as:

$$P(y(\boldsymbol{p})|\theta) = \frac{1}{\sqrt{2\pi\sigma^2(\boldsymbol{p})}} \exp\left(\frac{(y(\boldsymbol{p}) - I_{CM}(\boldsymbol{p}))^2}{2\sigma^2(\boldsymbol{p})}\right); \quad \sigma^2(\boldsymbol{p}) = \sigma_y^2(\boldsymbol{p}) + \sigma_{CM}^2(\boldsymbol{p}) \quad (3.4)$$

That is, it is calculated using a normal distribution with mean equal to the pixel value in the PSF-convolved model, and variance equal to the sum of the pixel variances in the observed variance map and the model variance map. The likelihood function $P(y|\theta)$ is then the joint probability (product) of the individual pixel probabilities.

Models may also be built where the pixel conditional probabilities use other distributions (e.g., Poisson, Skellam, or Binomial for discrete CCD data). Since the software has been developed for use with WFC3 infrared channel data processed by the HST pipeline where a Poisson likelihood is not an obvious choice (see discussion of weight maps above), Normal is the default likelihood.

The user also has full control over the prior distributions of individual parameters. These quantify prior knowledge about a given parameter's distribution, e.g., the distribution of Sérsic indexes or the galaxy luminosity function for a particular population. For a given sample, the (joint) prior probability of all the parameters $P(\theta)$ is calculated and multiplied by the likelihood function to obtain the posterior probability $P(\theta|y)$ for that sample. Although the choice of priors is to some degree subjective, they codify assumptions or prior knowledge explicitly, rather than the rash and often implicit assumptions (e.g., normality) of other methods. As with all assumptions, they should be considered carefully and stated explicitly. If there is no reason to prefer any parameter value over another, then uniform priors spanning a range of reasonable values can be used.

When building models, the user is encouraged to make generous use of the masking functionality in psfMC. Any object or structure that is not explicitly treated in the model — such as tidal tails or unrelated field galaxies — should be masked out. The reason is two-fold: first, goodness-of-fit diagnostics will be more meaningful if unmodeled objects are masked, and second, the MCMC process may try to fit unmasked flux of these objects, thus affecting the best-fit parameters. Since masking is important, I made masks simple to construct and edit, by letting the user supply them in the SAOImage ds9 region format. Regions are interpreted as the section of image to include; to exclude a section of pixels, the region's "exclude" property must be set. An example showing masked nearby galaxies is shown in Figure 3.2.

It is also recommended that cutout or "stamp" images be used for both the quasar and the PSF. Sérsic profiles and Fast Fourier Transform (FFT) PSF convolution in particular are computationally expensive and can quickly become slow if the image is large. Since FFT is used for convolution, rounding to a power-of-two image size may also increase speed. For HST images of high-z quasars, 128×128 or 256×256 is generally optimal. Array calculations and FFT are performed using the `NumPy` module (Oliphant 2006), which uses a combination of compiled C and Fortran routines for fast numeric computations. Generating 100,000 accepted samples of a PSF+Sérsic model with 256×256 stamp images currently takes $60 - 90$ minutes on a desktop workstation (dual-core 3.6 GHz Intel i5 processor), and there is potential for further significant optimization in the future.

Several tests are also provided with `psfMC` to ensure that correct output is produced at various steps in the modeling process. In particular, the surface brightness distributions are tested by comparing the raw model to a model with the same parameters produced by `GalFit`. `GalFit` uses a more sophisticated integration scheme, with the result that it produces more accurate results for the central pixels of very concentrated (e.g., high Sérsic index) profiles. For well-sampled images (e.g., the HST images described in Chapters 4 and 5), this results in a fractional error of less than 3% per pixel. Future versions of the software will use adaptive integration step sizes to reduce this source of error.

3.4   Analyzing `psfMC` Output

The primary output of the fitting function is the MCMC trace database (saved in the standard Python "Pickle" format), which allows detailed analysis of the full posterior probability distribution. Several FITS format images are also output by the fitting function. By default, these include the raw model, convolved model, compos-
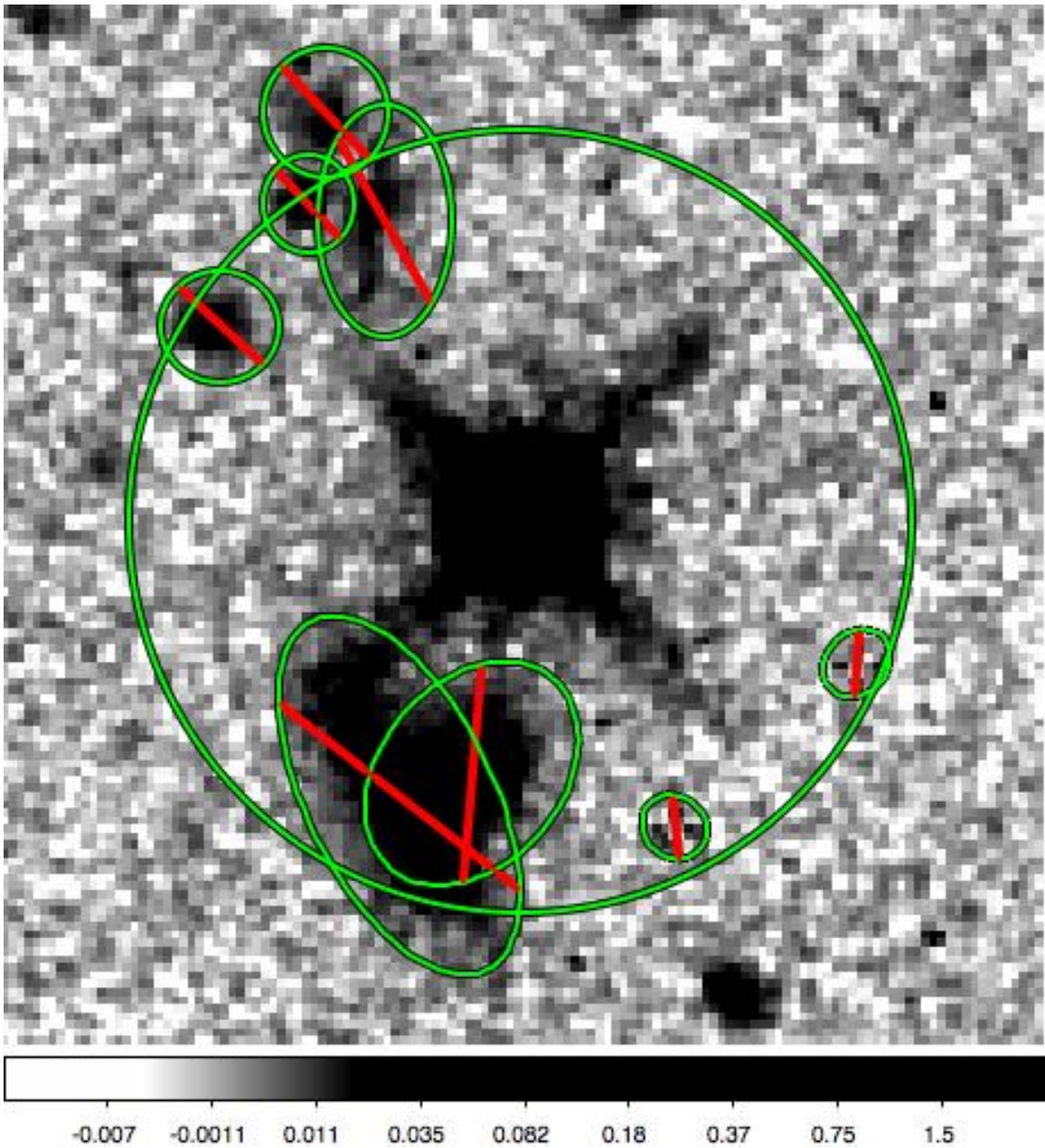
Figure 3.2: Example of using SAOImage ds9 region files to mask unmodeled background galaxies and bad pixels. The large outer circle is the region to include in fitting. The smaller ellipses with a slash through the center are regions to exclude, selected by setting their exclude flag in ds9.

ite inverse variance map (combined model and observation variance), an image with all point sources subtracted, and a residual image with all modeled components subtracted. A list of desired output products can be specified when the sampler is run. Examples of these images are shown in Figure 3.3. The images output by the fitting routine should be considered diagnostic, and should not be used without analyzing the MCMC chain. The parameter values used to create these images are simply the values corresponding to the sample that maximized the posterior probability. If the model has not converged, or if the posterior distribution has multiple distinct modes, the maximum a posteriori estimate may be an incorrect or incomplete description of the likely parameter values. The best way to find and estimate these modes depends upon the nature of the posterior distribution, and many techniques for analyzing high-dimensional posteriors exist, including kernel density estimation methods, clustering methods such as k-means, and the Patient Rule Induction Method (PRIM, Friedman & Fisher 1999). An example of using kernel density estimation to examine parameter covariance in a two-component point source+Sérsic model is shown in Figure 3.4.

Since the Markov chain can be run for as many samples as desired, it is important to know when it is appropriate to stop simulating. The goal of the MCMC process is to estimate, via sampling, the posterior parameter distribution $P(\theta|y)$. The number of samples required to estimate this will vary as a function of the particular problem and the accuracy required. For most quasar+Sérsic host models, this will be several tens of thousands of samples, since these models have high-dimensional (> 10) parameter spaces.

To assess whether samples are mixing well (providing an accurate estimate of the posterior distribution), two techniques may be particularly useful. The first is to run multiple chains, and see whether they have converged to the same region of parameter space, by comparing the variance within a single chain to the variance
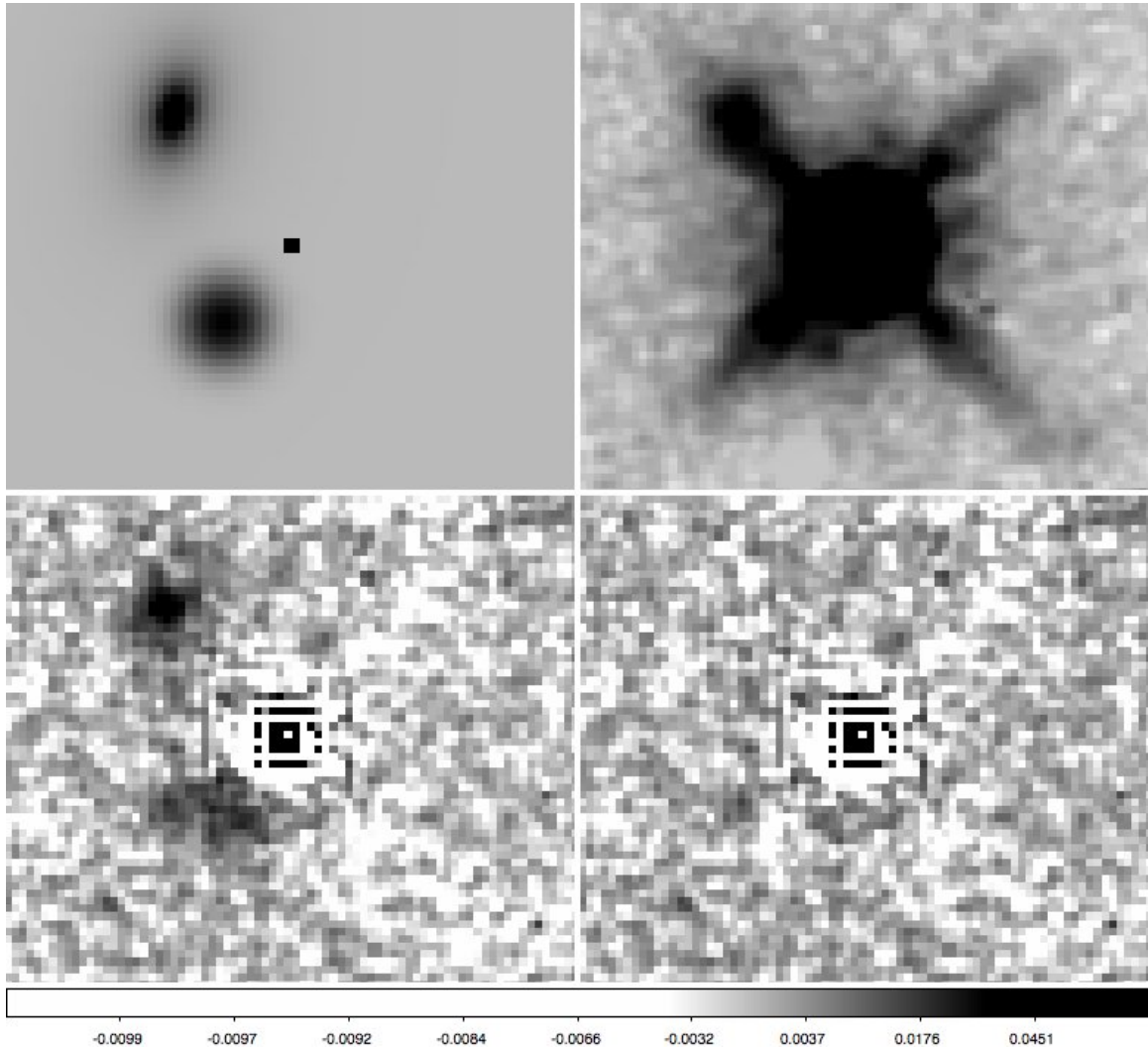
Figure 3.3: Output images produced by `psfMC`. All images are displayed with the same logarithmic stretch. Upper left: Raw model, the intrinsic surface brightness distribution of the maximum posterior model. Upper right: Convolved model, the raw model after convolution with the supplied point spread function. Lower left: Point source subtracted, the original data with all point sources from the maximum posterior model subtracted. Lower right: Residual, the original data with the maximum posterior model subtracted. Not pictured: Composite inverse variance map, including variance contributions from both the original data and the convolved model.
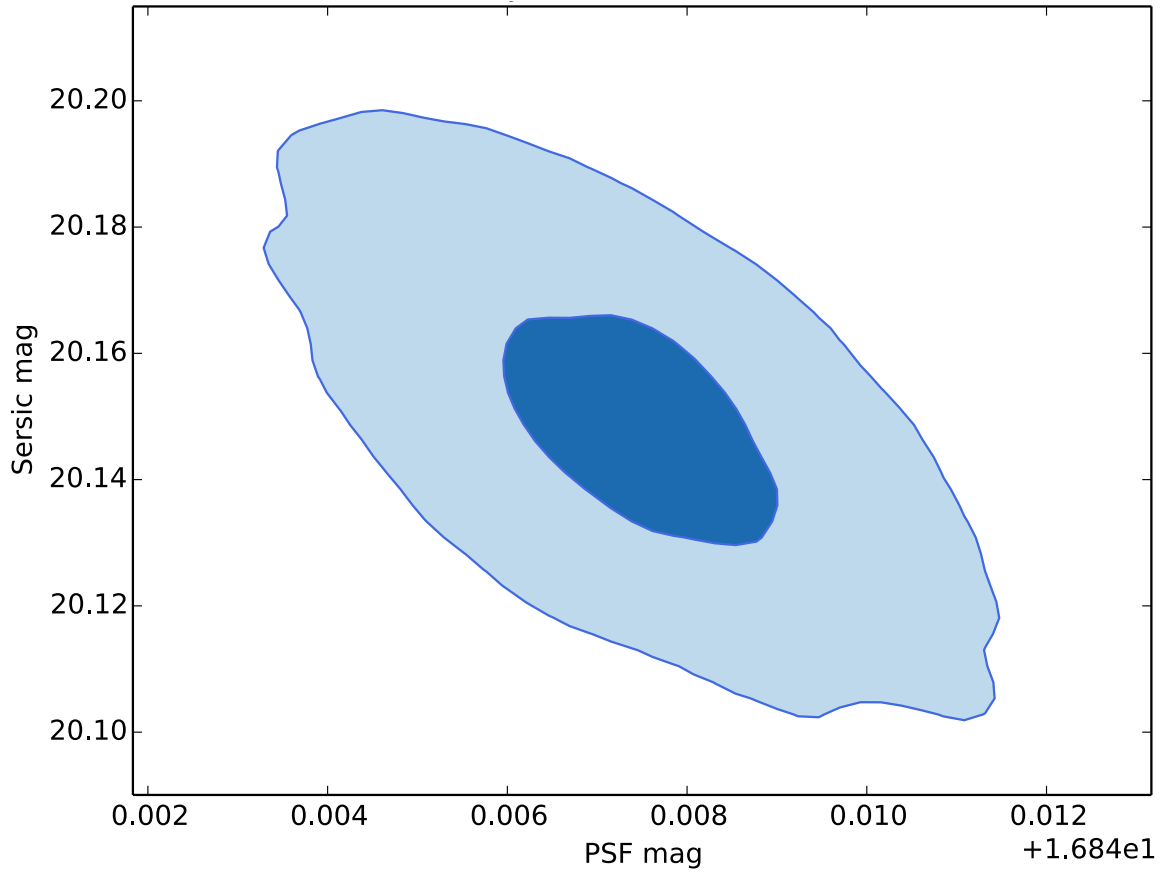
Figure 3.4: Example of parameter covariance analysis using kernel density estimation, for a two-component point source+Sérsic profile model of the quasar SDSS J215954.45−002150.1. Shown is the joint two-dimensional marginalized distribution of PSF component magnitude and Sérsic component magnitude, showing significant parameter covariance. Contours show 68% and 95% confidence levels, based on 80,000 retained MCMC samples. The x-axis zero point is 16.84 mag, as indicated in the lower right corner. This covariance is as expected, since the total flux of the point source+Sérsic model must equal the observed flux of the object.

between chains (Gelman & Rubin 1992; Gelman et al. 2011). Since this requires running multiple chains, it may be computationally expensive if fitting many objects. Ideally, the starting points for the individual chains should be over-dispersed, such that they begin in significantly different regions of parameter space.

When multiple chains cannot be run, an alternative is sample time-series analysis, such as that of Geweke (1992) which compares mean and variance of a parameter at the beginning and end of a chain. In particular, time-series analysis can be performed on the sample *deviance*, defined as $D(\theta) = -2(\ln(P(y|\theta)) - C)$ (i.e., -2 times the log-likelihood of the given sample), where C represents the log-likelihood of some model being compared to, for instance the maximum posterior model from the MCMC trace. This quantity gives some indication of the stability of sampling for the model as a whole, rather than having to analyze traces for each individual parameter. Examples showing an un-converged sampler and a converged sampler are shown in Figures 3.5 and 3.6. Note that if the un-converged sampler had been stopped after 100,000 samples, it would be indistinguishable from a converged sampler. This underscores the benefit of instead using the multiple-chain method whenever possible. An example of using multi-chain convergence diagnostics is discussed in Chapter 4.

Model selection and comparison can be performed using standard least-squares methods on the maximum posterior model (such as reduced χ^2), or using Bayesian methods that incorporate information from the entire posterior distribution, such as the Deviance Information Criterion (Gelman et al. 2011, Chapter 6) and Bayesian Predictive Information Criterion (Ando 2007, 2011). These can be used, for instance, to decide whether a PSF+Sérsic model over-fits the data with respect to a PSF-only model. Even without performing formal model comparison analyses, it may sometimes be obvious that a two-component model does not call for the Sérsic component. If the PSF is well-matched (see below for a description of PSF matching criteria) and
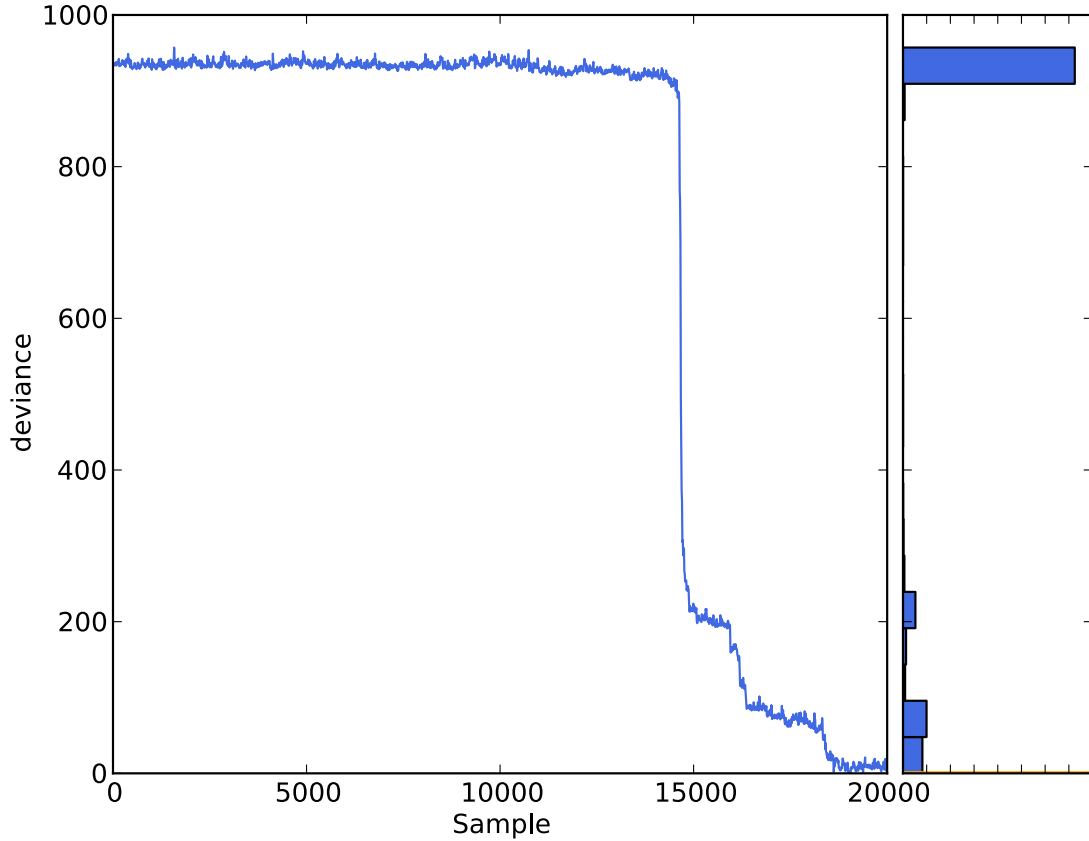
Figure 3.5: Example of a sampler that has not yet converged, where the distribution of deviance values at the beginning of the trace is significantly different from the distribution at the end of the trace. The sampler rapidly finds a significantly higher-probability region of parameter space near sample 15,000. In such a case the sampler can be continued to ensure the desired number of posterior samples are available for analysis, or re-run from the start with a longer burn-in period.

the observation and PSF weight maps are accurate, an effective host non-detection will manifest as Sérsic parameters that generate profiles with low surface brightness, such that individual pixels fall within the observed point source+PSF noise budget. If the weight maps are inaccurate, or there are significant mismatches between the observation and model PSFs, then host non-detections will manifest as models with small radii and high Sérsic index, which concentrate their flux within a small number of pixels.
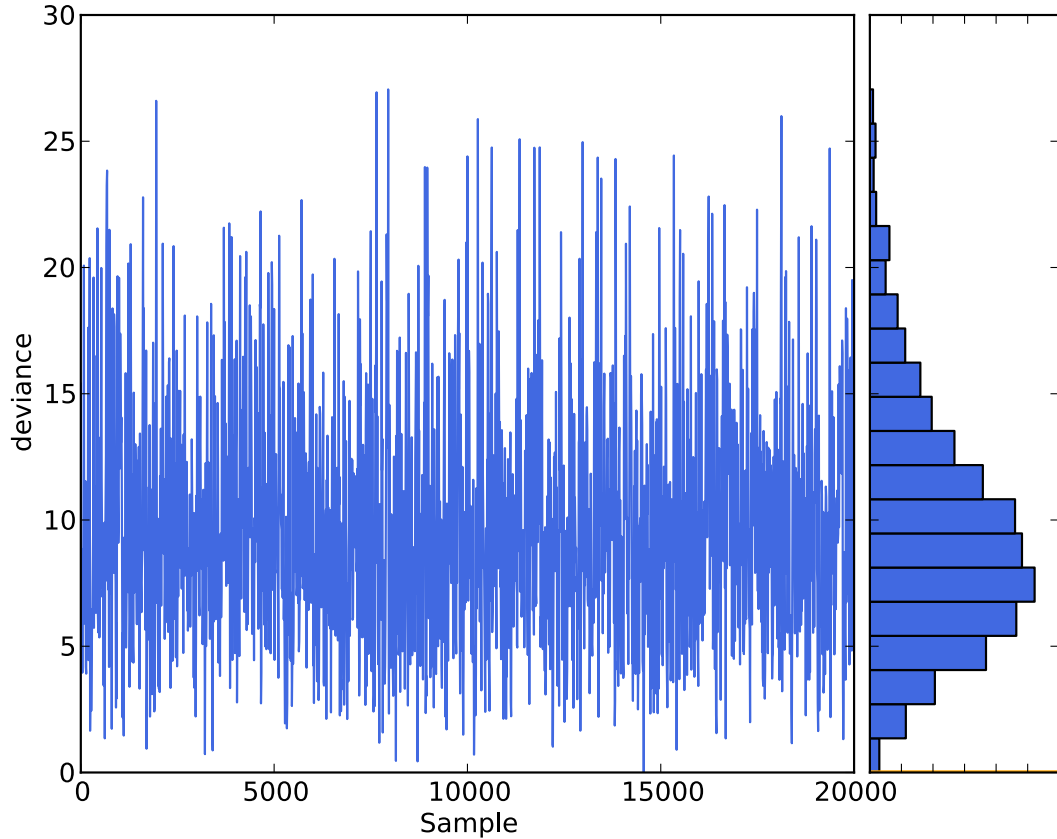
Figure 3.6: Example of a converged sampler, where the distributions of deviance values at the beginning and end of the trace is identical. Note the difference in vertical scale compared to Figure 3.5. These samples are ready to be analyzed, and are representative of the true posterior distribution.

3.5 Selecting the Best Model PSF

Selecting the proper PSF is key to enabling the high-contrast imaging necessary for high-z quasar studies. A PSF that significantly deviates from that in the quasar image will affect measurements of the underlying host galaxy parameters, or, in the worst cases, make morphological assessments impossible. Selecting the best PSF means understanding the sources of variation in the PSF and matching the conditions between the quasar and PSF observations as much as possible. This is true whether using a space-based instrument such as WFC3, where the variation is due primarily

to instrumental effects, or ground-based imaging, where the variation is primarily due to atmospheric effects. The important point is that simply stacking many stars to obtain an average PSF is unlikely to produce the best results, especially if there is significant PSF variation across the instrument field of view, or the observations are short compared to the typical timescale of PSF variations. I describe the primary sources of variability for WFC3 infrared imaging below. Imaging with other space-based instruments will have many of the same sources, and ground-based imaging will additionally have atmospheric effects.

As summarized in Chapter 2, a major source of variability in the HST PSF comes from the spacecraft breathing effect, where variations in the telescope thermal environment cause de-space of the secondary mirror, leading to variations in focus. There are two primary sources of thermal variability: the angle of the telescope with respect to the sun (mainly a function of season and the object being observed), and the orbital day-night cycle as the telescope orbits the Earth every 90 minutes (Hershey 1998). The focal variations can either be measured directly, through carefully coordinated observations of a PSF star (such as in Chapter 2), or by estimating the secondary mirror de-space for an exposure after the fact, using the HST Focus Model (Cox & Niemi 2011).

For observations through broad-band filters, the spectrum of the observed star or quasar may also significantly affect the PSF. The monochromatic instrument PSF may change significantly across the filter bandpass, including both size (due to the diffraction limit's λ/D dependence), and structure (due to wavelength dependence of optical aberrations, diffraction spikes, etc.). Thus a star with a blue spectral energy distribution may have an integrated PSF significantly different from a star with a red spectral energy distribution. This effect was explored by Bahcall et al. (1997), who found that using a blue $((B - V) = -0.07$ mag) versus a red $((B - V) =$

41

1.11 mag) PSF star significantly affected the cosmetics of the residual images, but not the photometry of their bright ($m_{F606W} \simeq 17.5$), low-redshift ($z \simeq 0.15$) quasars. For high-redshift quasars whose hosts are fainter and smaller compared to the PSF, large color mismatches will significantly affect *both* photometry and morphological assessments.

The point spread function of the HST WFC3 infrared channel also has significant optical aberration variations across the field of view. In particular, field-dependent focus, astigmatism, and coma have been measured (Hook & Stoehr 2008; Biretta 2012). Figure 3.7 shows the locations of two stars near the field of view center, separated by 37 arcsec in HST dataset IBR213020. Since these two stars come from the same set of exposures, their thermal focus history is identical. Their NIR ($F110W - F160W$) colors are 0.19 (psf27) and 0.23 (psf28). Thus the only significant possible source of PSF mismatch is field-dependent PSF aberration. Figure 3.8 shows the result of fitting one star with the other as the PSF. The spots of negative and positive residuals to the left and right of the center are tell-tale signs of coma mismatch, and could easily be mistaken for high-z galaxy features if the subtraction image had been a quasar. Three example subtractions for stars that were selected only from within $0\rlap{.}''25$ of the field center are shown in Figure 3.9. The PSF model for each star was focus-matched to the observation by estimating each exposure's focus using the HST Focus Model. The region of very poor matching is confined within the central $0\rlap{.}''3$ radius.

3.6 Discussion

The software presented in this chapter currently represents the most precise treatment of two-dimensional surface brightness modeling for quasar host galaxies. It is unique in its explicit treatment of uncertainties in the supplied PSF, which are not treated by any other current modeling software. By adopting a Markov Chain Monte

42
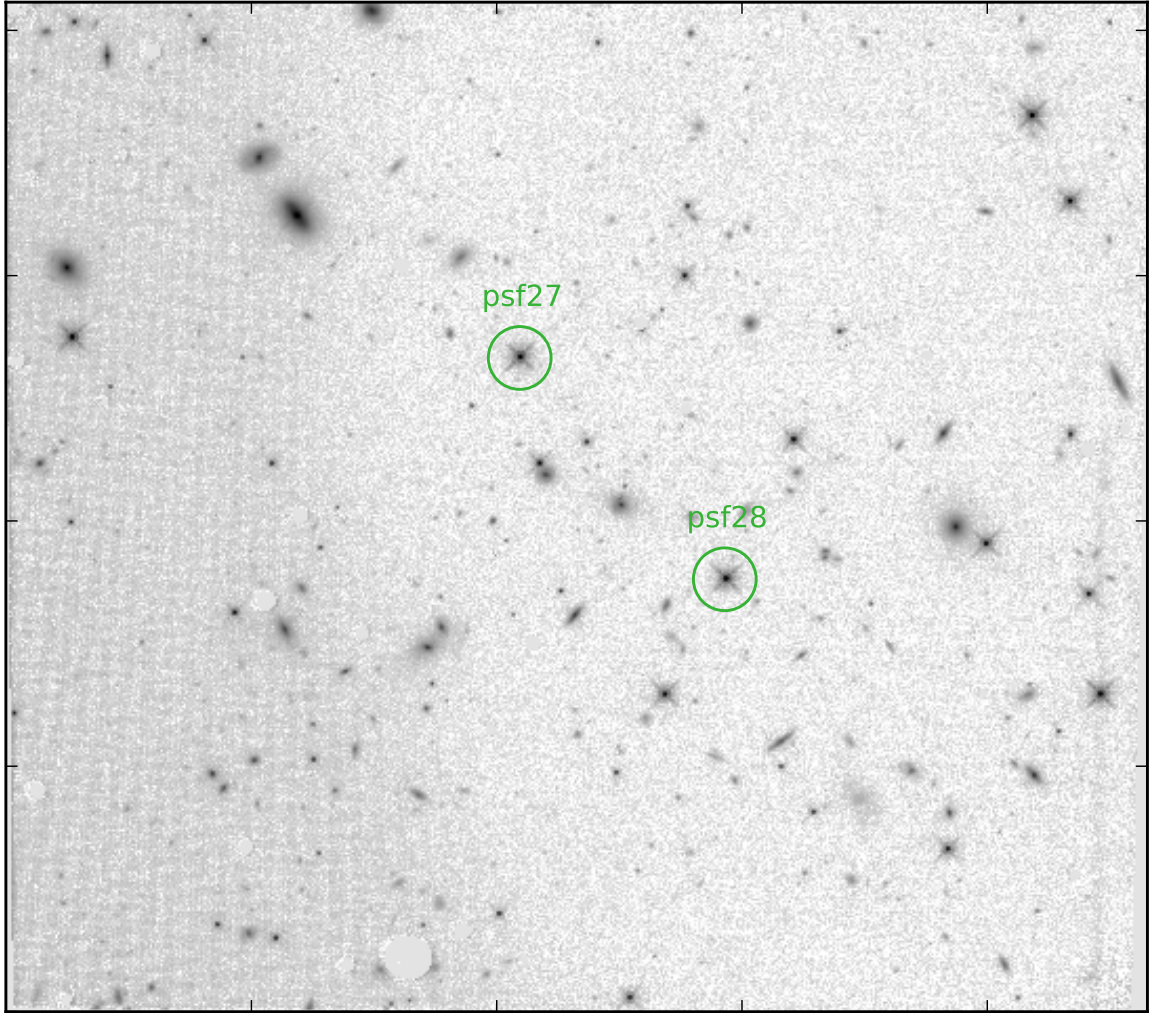
Figure 3.7: Locations of two stars used for characterizing the effect of field-dependent PSF aberrations. The stars are separated by 37 arcsec and are marked by the labeled circles. The image shows the entire WFC3 infrared field of view, roughly $2\rlap{.}'3 \times 2\rlap{.}'1$.
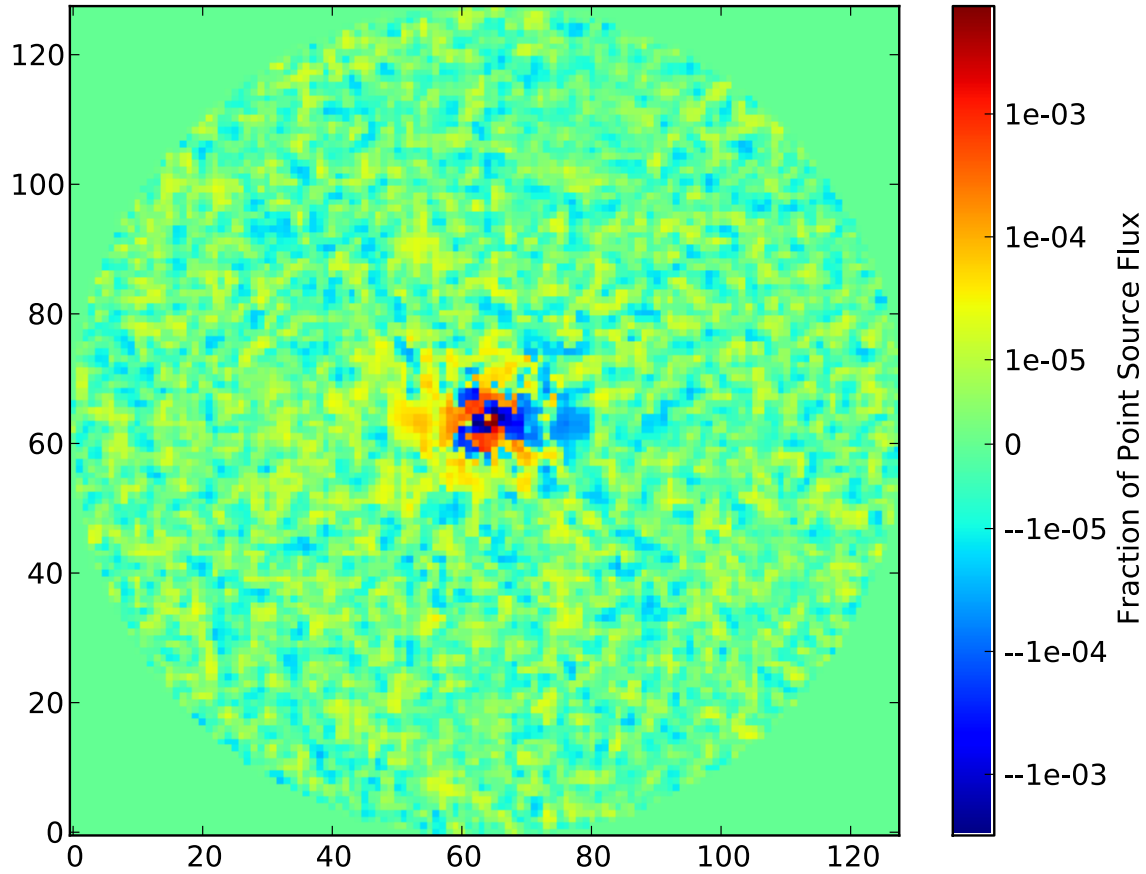
Figure 3.8: Fitting residual after subtracting one of the stars in Figure 3.7 from the other. X and Y coordinates are in pixels, with color representing the significance of the PSF mismatch structures, using the `psfMC` composite variance map to measure the noise. The over- and under-subtracted patches to the left and right of the center could be mistaken for host galaxy structures in a real quasar subtraction. Units are in fractional flux of the original point source, before subtraction.

Carlo approach to characterizing the posterior parameter distribution, it enables the detailed analyses necessary to understand the complex parameter covariances and non-gaussian parameter probability distributions inherent in quasar+host models.
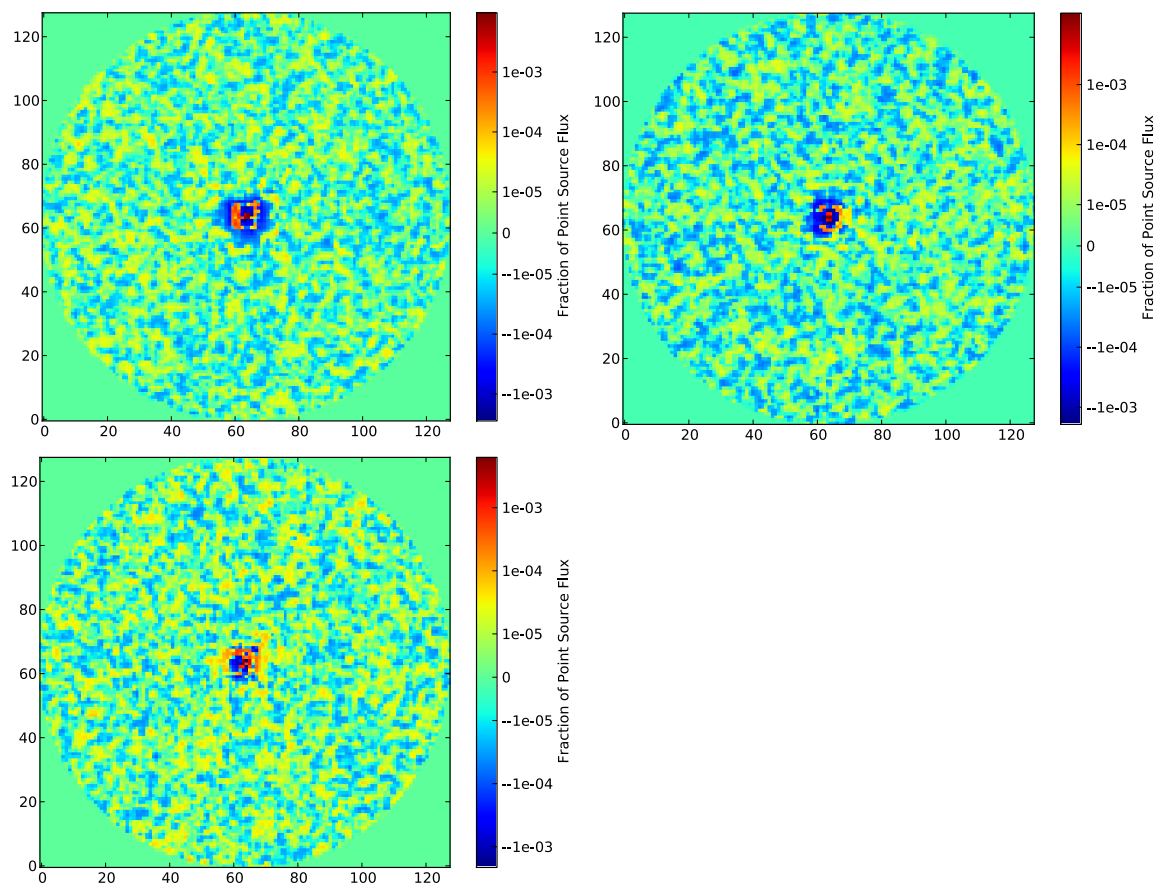
Figure 3.9: Three examples of fitting residuals after subtracting one star from another, where the stars have been selected only from within $0\overset{\prime\prime}{.}25$ of the WFC3 field of view center. X and Y coordinates are in pixels, with color representing the significance of the PSF mismatch structures. Units are in fractional flux of the original point source, before subtraction.

Chapter 4

QUASAR HOST GALAXY MORPHOLOGIES AT $z \simeq 2$

4.1   Introduction

An outstanding question in the study of AGN is the importance of major mergers as triggering mechanisms for the AGN phase. As discussed in Chapter 1, there is reason to suspect gas-rich major mergers as a triggering mechanism, and indeed many previous studies have found signatures of such mergers in quasar hosts. However, in a recent study of x-ray selected AGN in the Cosmic Evolution Survey (COSMOS) field, Cisternas et al. (2011) found no evidence for an enhancement of the merger fraction in AGN hosts at $z \simeq 0.3 - 1.0$. That is, AGN hosts were no more likely to be undergoing a major merger than galaxies without AGN. Thus, at least at $z \lesssim 1.0$, other triggering mechanisms may dominate. However, this may not be true at higher redshifts, or specifically for quasars, as the most luminous AGN with the highest accretion rates. The HST WFC3 now allows such detailed structural studies to be extended to $z \simeq 2$, since rest-frame optical wavelengths can now be imaged at HST resolution using the WFC3 infrared channel.

Several previous studies have examined the structural properties of galaxies not hosting quasars at $z \simeq 2$, especially after the installation of WFC3IR. In particular, multi-wavelength data from large HST programs such as the WFC3 ERS2 field (Early Release Science 2, Windhorst et al. 2011), the HUDF (Hubble Ultra Deep Field, Beckwith et al. 2006; Oesch et al. 2007; Illingworth et al. 2013; Koekemoer et al. 2013), and CANDELS (Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey, Grogin et al. 2011; Koekemoer et al. 2011) have enabled large-scale

NIR structural studies for galaxies beyond $z \simeq 1.4$ with secure redshifts. Detailed studies of the $z = 1.893$ galaxy FW4871 by van Dokkum & Brammer (2010) and Ferreras et al. (2012) provided evidence for merger-induced star formation and associated young stellar populations in $z \simeq 2$ early-type galaxies. Ryan et al. (2012) and Cassata et al. (2013) examined the properties of passively-evolving red galaxies, finding that massive galaxies grow in size more rapidly than lower-mass galaxies at $z < 3$. Targett et al. (2013) studied the structures of (sub-)mm-selected galaxies at $z \simeq 1 - 3$, finding them to be disk-dominated systems with high star-formation rates. However, detailed structural studies of quasar hosts at $z \simeq 2$ using WFC3IR have not yet been attempted.

In this chapter, I describe the first steps toward searching for an enhancement to the quasar host merger fraction at $z \simeq 2$: near-infrared imaging of 19 $z \simeq 2$ quasars with the HST WFC3IR, methods for characterizing and subtracting the AGN emission, and the rest-frame optical properties of the host galaxies. §4.2 describes the quasar sample selection and existing data. §4.3 describes the WFC3IR imaging and data reduction. §4.4 details how stars were selected from the HST archive to act as PSF models for each quasar. The simultaneous two-component image modeling with `psfMC` is detailed in §4.5, and the results of the fitting process in §4.6. Finally, in §4.7, I discuss the fitting results and the inferred quasar host merger fraction.

4.2 Sample Definition and Existing Data

A sample of 115 quasars with systematic redshift $z = 2.0 \pm 0.1$ was selected from the SDSS 5th Data Release Quasar Catalog (Schneider et al. 2007) for an HST Cycle 19 SNAPshot study (Program ID 12613, PI: Jahnke). Between October 2011 and September 2012, 19 of these quasars were observed with WFC3IR in the F160W filter (H-band). The locations of these 19 quasars are plotted in Figure 4.1.
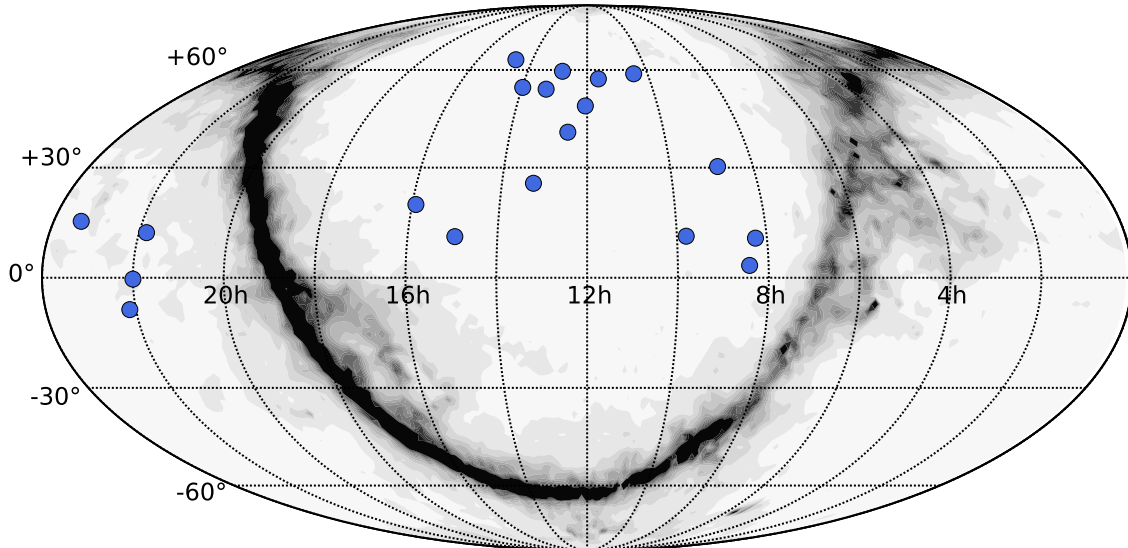
Figure 4.1: All-sky map (Mollweide projection) showing the locations of the 19 $z \simeq 2$ quasars. The grayscale background is the Galactic A_V extinction map of Schlegel et al. (1998).

Shen et al. (2011) classified the spectral properties of all the SDSS Data Release 7 quasars, including the 19 in this sample. The luminosities of the 19 $z \simeq 2$ quasars at rest-frame 3000 Å range from $\log(L_{3000}) = 12.54 - 13.33$ $L_\odot$. The virial black hole masses, as estimated from the C IV line, range from $\log(M_{BH}) = 9.3 - 9.9$ $M_\odot$. The continuum power-law index α at a rest-frame wavelength of 2500 Å $(f_\nu = C\nu^{-\alpha})$ ranges from $\alpha = 0.11 - 1.1$. [1] One object was flagged as a C IV broad absorption line (BAL) quasar; examining the SDSS spectra confirms that it is the only BAL quasar among the sample. Eighteen of the quasars are covered by the FIRST radio survey (White et al. 1997), with one quasar having a strong 1.4 GHz radio flux $(S_{1.4} = 878$ mJy) indicating it is radio-loud, with lobe-dominant morphology (following the classifications in Jiang et al. 2007). These properties are summarized in Table 4.2.

---

[1]Shen et al. (2011) measure this value when fitting the Mg II line, using the parameterization $f_\lambda = A\lambda^{\alpha_\lambda}$. These values have been converted to the frequency-dependent formalism more common in the literature.

Table 4.1: Summary of Spectral Properties of $z \simeq 2$ Quasars

Quasar (SDSS J)	Redshift z	L_{3000} $\log(L_\odot)$	Continuum Index α	M_{BH} $\log(M_\odot)$	$S_{1.4}$ mJy
081518.99+103711.5	2.021	12.72	0.41	9.6	(0.98)
082510.09+031801.4	2.035	12.91	0.81	9.7	(1.80)
085117.41+301838.7	1.917	12.54	0.39	9.6	(0.92)
094737.70+110843.3[a]	1.905	12.66	0.66	9.8	(1.00)
102719.13+584114.3	2.020	12.67	0.26	9.7	(1.01)
113820.35+565652.8	1.917	12.76	0.74	9.8	(0.99)
120305.42+481313.1	1.988	12.79	1.10	9.3	(1.01)
123011.84+401442.9	2.049	13.17	0.82	9.6	(0.91)
124949.65+593216.9	2.052	12.88	0.97	9.5	1.69
131501.14+533314.1	1.921	12.70	0.25	9.3	(0.94)
131535.42+253643.9	1.926	12.57	0.11	9.5	(0.92)
135851.73+540805.3	2.066	12.80	0.29	9.5	(0.93)
143645.80+633637.9	2.066	13.22	0.76	9.4	877.84
145645.53+110142.6	2.017	12.79	0.64	9.4	(0.98)
155447.85+194502.7	2.091	12.71	0.95	9.9	(1.77)
215006.72+120620.6	1.993	12.66	0.25	9.8	(1.01)
215954.45−002150.1	1.963	13.33	0.54	9.5	2.85
220811.62−083235.1	1.923	12.77	0.15	9.9	(1.03)
232300.06+151002.4	1.989	12.81	0.94	9.5	—

[a] C IV BAL quasar

Notes: Column 1: Name, giving the full SDSS celestial coordinates; Column 2: Redshift (Schneider et al. 2010); Column 3: Continuum luminosity at rest-frame 3000 Å (Shen et al. 2011); Column 4: Power-law index of the continuum emission at rest-frame 2500 Å (Shen et al. 2011, re-parameterized as $f_\nu = C\nu^{-\alpha}$, see note in text); Column 5: Virial black hole mass estimated from the C IV line by Shen et al. (2011), using the formalism of Vestergaard & Peterson (2006); Column 6: 1.4 GHz flux density from the FIRST catalog (White et al. 1997). Values in parentheses are upper limits, given by the FIRST catalog 5σ detection limit calculated at the quasar position, including CLEAN bias.

4.3 Hubble Space Telescope Data

All 19 quasars were observed with WFC3IR using the F160W filter (H-band, rest-frame 5200 Å). None of the quasars had existing HST imaging using the Wide Field and Planetary Camera 2, Advanced Camera for Surveys, or WFC3 UVIS channel. As a SNAP program, these observations by necessity have short exposure times of 1596 seconds per target (less than one orbit). Despite the short exposure time, the excellent sensitivity of WFC3IR and low on-orbit near-infrared sky allow the data to reach a 1σ limiting surface brightness of 24.47 mag arcsec^{-2}, sufficient to identify tidal disturbances at $z \simeq 2$, such as those described by van Dokkum & Brammer (2010) and Ferreras et al. (2012), or the simulations in Kaviraj et al. (2012). Targets were observed using the standard four point sub-pixel dithering pattern to improve PSF sampling and to assist in the rejection of bad pixels and cosmic rays.

Data processing began with individual flat-fielded, flux-calibrated exposures delivered by the HST archive. All four exposures for each pointing were combined using the `Multidrizzle` software package (Koekemoer et al. 2002, 2011) with an output plate scale of $0\rlap{.}''060$ per pixel and a `pixfrac` parameter of 0.8. For these observations, this provides Nyquist sampling of the PSF and relatively uniform weighting of the individual pixels. Inverse variance weighting was used for the final image combination step. Variance maps that include all sources of noise (including shot noise from the object flux) were generated by copying the WFC3 "ERR" arrays into the standard image arrays, and re-running the drizzle process with the same parameters and weighting (as suggested in Gozanga et al. 2012). Accurate variance maps including shot noise are necessary since the photon count rate from the quasar point source is significantly higher than that of the sky, and under-estimated errors can lead to problems with multi-component fitting, as described in Chapter 3.

4.4 Point Spread Function Models

As a SNAP program, the HST data did not have dedicated observations of stars to measure the instrument and telescope PSF. Few of the exposures contained high S/N stars, and simulated PSF models are currently poor matches to observations (e.g., Mechtley et al. 2012; Biretta 2012), so empirical PSF models were built from HST WFC3 archival data.

The HST focus is affected by the telescope's thermal environment, with changes in solar illumination resulting in de-space of the secondary mirror, the so-called "spacecraft breathing" effect (Bély et al. 1993; Hershey 1998; Cox & Niemi 2011). In addition, the WFC3IR PSF has other aberrations (e.g., coma and astigmatism) that vary with position within the field of view. To find similar exposures from which to extract PSF models, the HST archive was searched for all single-orbit F160W observations. Only single-orbit data were used because these provide the highest degree of pointing stability ($2-5$ milliarcsecond RMS, Gozanga et al. 2012), which matches that of the quasar observations. Observations with fewer than four dither points were thrown out, to ensure sub-pixel sampling equal or better than the quasar observations.

All point sources falling within $0\rlap{.}''25$ of the WFC3 field of view center were then selected, and those which were identified as optical quasars or radio sources in the NASA/IPAC Extragalactic Database were excluded. The remaining PSF stars were inspected both visually and with the IRAF task `psfmeasure` (part of the STSDAS package), and those that were contaminated by background galaxies or that were significantly elliptical (probable binary stars) were excluded. This left a library of nine star images whose S/N exceeded the quasars, but which still had accurate count rates in their cores (i.e., were not effectively saturated). Drizzled images of these stars were created with the same plate scale and weighting scheme described in §4.3.

The HST Focus Model (Cox & Niemi 2011) was then used to estimate the effective spacecraft focus during each quasar and star observation. A cubic spline was fit to the focus estimates for each observation (typically available at five-minute intervals), and used to calculate the average focus value over the entire WFC3 exposure. [2] Each quasar was then paired with the star having the closest average focus. No attempt was made to match the spectral slope of the star within the F160W filter to that of the quasar point source. However, most stars at high galactic latitudes are low-mass dwarf stars subject to little reddening. Their $(J - H)$ colors should range from roughly 0.5 to 0.7 Vega mag (Skrutskie et al. 2006), which matches the $(J-H)$ colors of high-z quasars (e.g., simulation tracks in Hewett et al. 2006).

4.5   Point Source Subtraction and Host Galaxy Modeling

A two-component Sérsic+PSF fit was first attempted using the software `GalFit` (Peng et al. 2002, 2010). As discussed in Chapter 3, and noted in §6.2 of Peng et al. (2010), `GalFit` employs several design decisions that make it computationally efficient, but not optimal for quasar modeling and point source subtraction. To address these problems, I developed the Markov Chain Monte Carlo simultaneous fitting software `psfMC`, described in Chapter 3.

Each quasar in the sample was modeled using `psfMC` to simultaneously fit two components — a point source and a single Sérsic profile. The prior distributions adopted for the model parameters are summarized in Table 4.5. At $z = 2$, the drizzled $0\rlap{.}''060$ linear pixel scale corresponds to a physical size of 0.52 kpc. The position priors correspond to a maximum separation of 6.1 kpc between the point source and the center of the underlying galaxy. A Weibull distribution was used as

---

[2]I have made the software to query the focus database and perform the spline fit freely available at http://github.com/mmechtley/HSTFocusModel

Table 4.2: Adopted Prior Distributions of Fitting Parameters

Parameter	Distribution	Value Range
PSF Component		
X,Y Position	Uniform	Centroid ± 4 pixels
Magnitude	Uniform	$m_H\,{}^{-0.2}_{+1.5}$ mag
Sérsic Component		
X,Y Position	Uniform	Centroid ± 8 pixels
Magnitude	Uniform	Upper: m_H, Lower: 26 mag
Major/Minor Axes	Uniform	$3.0 - 15.0$ pixels $(0\rlap{.}''18 - 0\rlap{.}''90)$
Index n	Weibull[a]	$\lambda = 4$, $k = 1.5$
Angle	Uniform	$0 - 180$ degrees

[a] The Weibull distribution is selected to model the empirical distribution of Sérsic indexes (e.g., Ryan et al. 2012).

Notes: Ranges are for intrinsic quantities, before convolution with the PSF. "Centroid" refers to the flux centroid of the (point-source dominated) quasar+host galaxy in the HST image. m_H is the total measured H-band magnitude of the quasar+host system.

the prior distribution of Sérsic indexes, modeling the empirical distribution of indexes from Ryan et al. (2012). The ranges on the other parameter priors were selected to model the entire range of values that are both physically reasonable and detectable. Nearby galaxies that were not coincident with the quasar point source were masked out.

Four separate MCMC chains were run in parallel for each model, each with an initial length of 100,000 samples, with the first 80,000 discarded as a burn-in period and the final 20,000 retained for analysis. This resulted in a total of 80,000 retained samples for posterior analysis (4 chains with 20,000 samples each). Convergence was then assessed in two ways. First, each individual chain was analyzed using the method of Geweke (1992), examining the Z-score of the difference of means between the first 2,000 retained samples and the last 10,000, requiring $|Z| < 2$. This ensured that each chain represented a stable distribution without a strong long-term trend.

Second, the Potential Scale Reduction diagnostic ($\hat{R}$, Gelman & Rubin 1992; Gelman et al. 2011) was calculated for each parameter to compare the within-chain variance to the between-chain variance, ensuring that all four chains had converged to the same stationary distribution. All fitting parameters were required to have $\hat{R} < 1.05$. Five quasars did not initially meet these criteria (i.e., had not converged after 100,000 samples), so were run for an additional 100,000 samples, again with the final 20,000 from each chain retained. Two still had not converged, although the point source parameters had. In both cases the host galaxy was effectively too faint to constrain its parameters (see discussion below). Figure 4.2 shows an example of diagnostic plots used to assess convergence.

Markov Chains produce correlated samples, and the degree of correlation can vary significantly, both for individual parameters and depending upon the particular path the random walk takes through parameter space. This correlation is often quantified as the effective number of samples for a given parameter, $n_{eff} = n/\tau$, i.e., the total number of retained samples n divided by the characteristic correlation length τ. This length can be estimated by summing the initial positive sequence of autocorrelation function values (Geyer 1992; Thompson 2010), i.e., the values for lags 0 to k, where k is the first lag with negative autocorrelation. An example is shown in Figure 4.3, estimating n_{eff} for the chains plotted in Figure 4.2. The effective number of samples is additionally estimated using the method described by Gelman et al. (2011, Chapter 11), which compares the within-chain variances to the between-chain variance. The latter method is expected to provide a cruder estimate, since only four chains were run, so is used for comparison only.
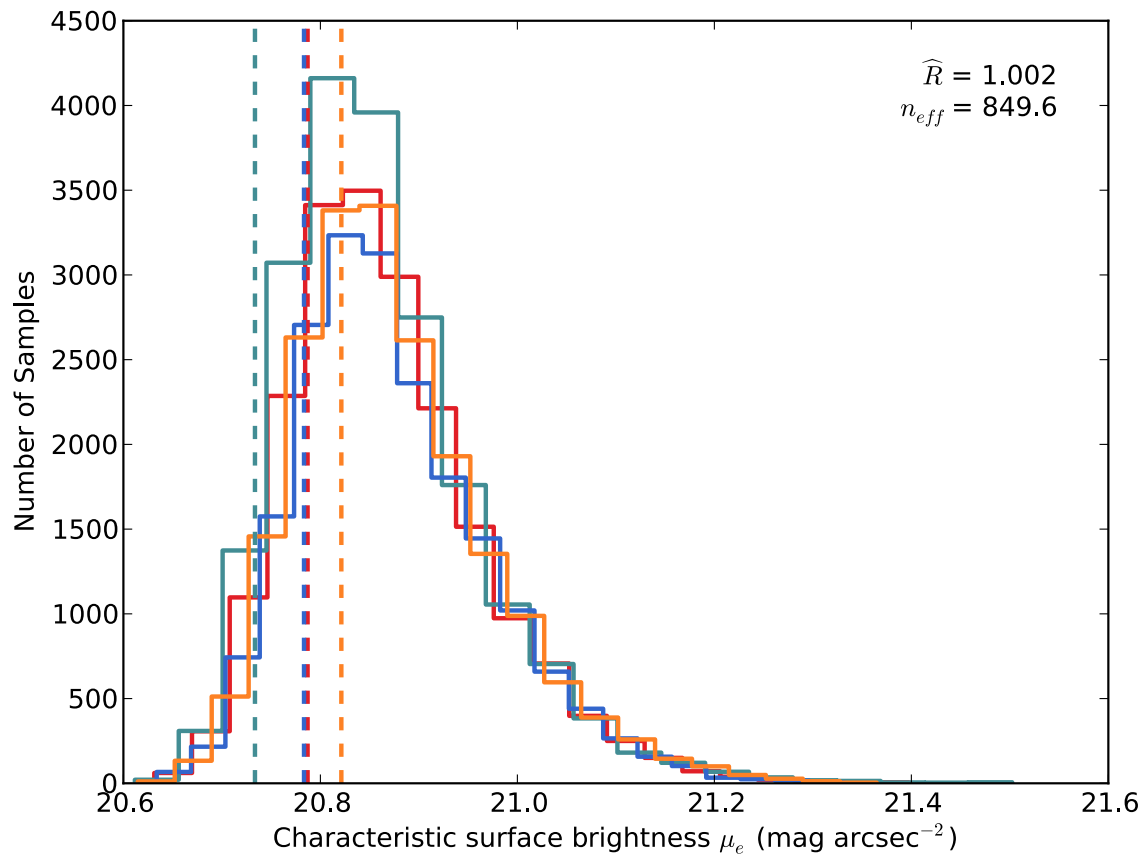
Figure 4.2: Posterior distributions of Sérsic profile characteristic surface brightness (μ_e) for four MCMC chains run for the same model (SDSS J081518.99+103711.5). This is not an explicit model parameter, but rather an example of a scalar estimand that can have its probability distribution analyzed after fitting, since thousands of MCMC samples have been retained. Each of the four chains is represented by a different color, with dashed vertical lines showing the values from the maximum posterior model for each chain. $\hat{R}$ is the Potential Scale Reduction diagnostic, and n_{eff} is the effective total number of samples after accounting for Markov Chain autocorrelation, estimated using the method of Gelman et al. (2011).
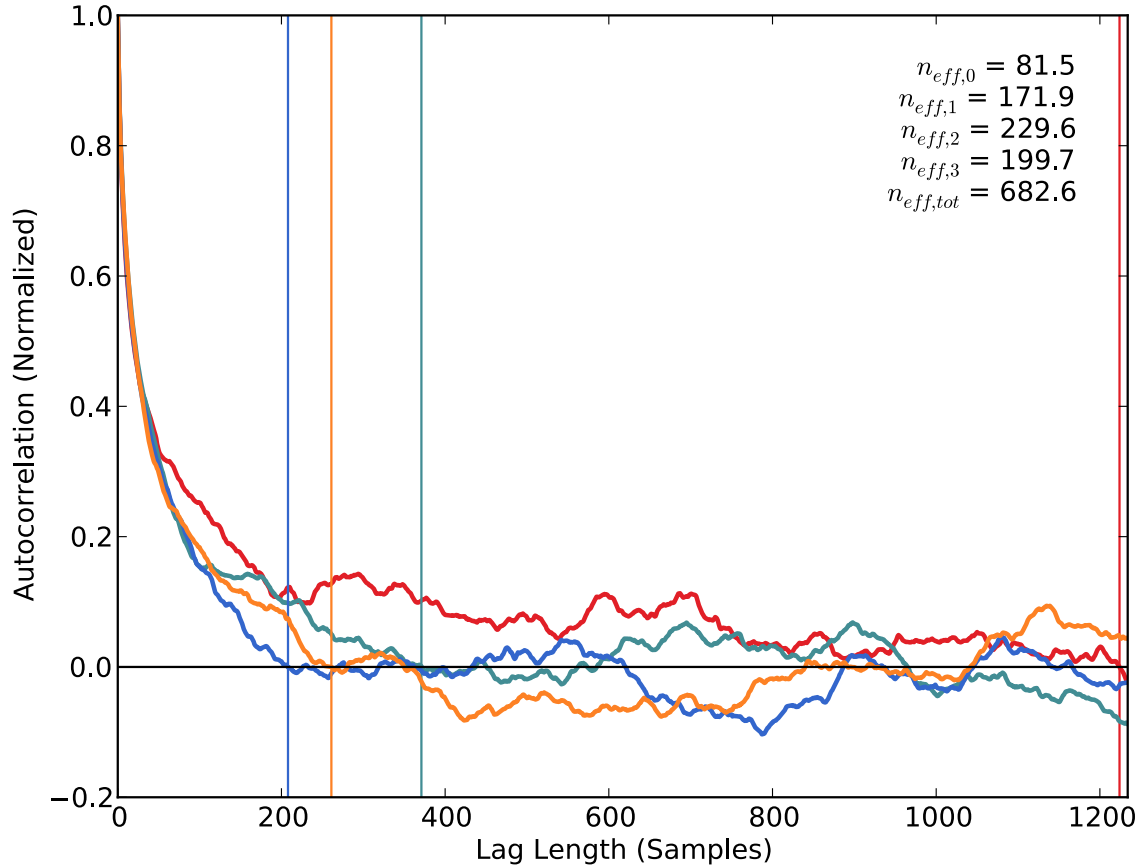
Figure 4.3: Autocorrelation functions of the chains from Figure 4.2. Vertical lines plot the first zero-crossing, i.e., the cutoff sample for the initial positive sequence. This then gives an estimate for the characteristic correlation length τ in each chain, which is used to calculate n_{eff} for each chain, and subsequently for all four chains as a whole.

4.6 Results of MCMC Fitting

The maximum posterior models from the MCMC fitting process for each quasar are shown in Figures 4.4−4.22. Table 4.6 summarizes the range of posterior values for each fitting parameter. Notes on individual quasars follow.

SDSS J081518.99+103711.5: The model is well-converged, with individual parameters having > 250 effective samples each. The Sérsic profile is not significantly offset from the point source, and the point source-subtracted residual is not significantly asymmetric. One other galaxy falls within a $2\rlap{.}''0$ radius.

Table 4.3: Summary of Posterior Parameter Values From MCMC Fitting

Quasar (SDSS J)	m_{AGN}	m_{host}	Index n	R_e (arcsec)	b/a	Offset (arcsec)
081518.99+103711.5	18.708 ± 0.006	21.3 ± 0.05	2.30 ± 0.26	0.21 ± 0.01	0.89 ± 0.03	0.029 ± 0.006
082510.09+031801.4	17.744 ± 0.001	—	—	—	—	—
085117.41+301838.7	18.987 ± 0.002	22.8 ± 0.05	0.20 ± 0.04	0.49 ± 0.04	0.92 ± 0.05	0.12 ± 0.017
094737.70+110843.3	19.226 ± 0.009	21.0 ± 0.09	1.45 ± 0.45	0.39 ± 0.01	0.77 ± 0.01	0.088 ± 0.007
102719.13+584114.3	18.800 ± 0.006	—	—	—	—	—
113820.35+565652.8	18.154 ± 0.002	21.9 ± 0.03	0.15 ± 0.01	0.58 ± 0.07	0.88 ± 0.04	0.06 ± 0.016
120305.42+481313.1	18.297 ± 0.002	20.9 ± 0.02	2.04 ± 0.12	0.55 ± 0.01	0.67 ± 0.02	0.231 ± 0.003
123011.84+401442.9	17.442 ± 0.009	20.4 ± 0.03	4.78 ± 0.44	0.89 ± 0.01	0.93 ± 0.04	0.173 ± 0.002
124949.65+593216.9	18.252 ± 0.002	21.3 ± 0.03	0.81 ± 0.06	0.64 ± 0.02	0.78 ± 0.02	0.152 ± 0.009
131501.14+533314.1	18.611 ± 0.002	22.1 ± 0.04	0.52 ± 0.06	0.67 ± 0.03	0.85 ± 0.04	0.12 ± 0.017
131535.42+253643.9	18.717 ± 0.006	20.2 ± 0.14	15.3 ± 4.19	0.32 ± 0.07	0.74 ± 0.02	0.059 ± 0.005
135851.73+540805.3	18.360 ± 0.001	—	—	—	—	—
143645.80+633637.9	17.155 ± 0.001	20.4 ± 0.03	0.82 ± 0.08	0.61 ± 0.01	0.82 ± 0.02	0.067 ± 0.007
145645.53+110142.6	18.295 ± 0.007	20.7 ± 0.14	13.0 ± 3.82	0.42 ± 0.09	0.54 ± 0.03	0.085 ± 0.008
155447.85+194502.7	19.109 ± 0.007	21.1 ± 0.05	5.26 ± 0.58	0.87 ± 0.03	0.70 ± 0.03	0.073 ± 0.001
215006.72+120620.6	18.847 ± 0.004	20.7 ± 0.06	1.09 ± 0.16	0.85 ± 0.03	0.76 ± 0.01	0.062 ± 0.008
215954.45−002150.1	16.776 ± 0.002	20.2 ± 0.03	0.46 ± 0.05	0.55 ± 0.01	0.81 ± 0.03	0.15 ± 0.011
220811.62−083235.1	18.668 ± 0.005	21.3 ± 0.06	0.96 ± 0.19	0.43 ± 0.02	0.88 ± 0.03	0.100 ± 0.009
232300.06+151002.4	18.345 ± 0.005	20.3 ± 0.09	6.16 ± 1.32	0.22 ± 0.02	0.90 ± 0.02	0.062 ± 0.004

Notes: Column 1: Quasar name (SDSS J); Column 2: AB magnitude of point source component; Column 3: Total AB magnitude of host Sérsic component; Column 4: Sérsic index n; Column 5: Host effective radius (R_e) in arcsec; Column 6: Host axis ratio b/a (minor/major); Column 7: Offset between point source and Sérsic component center in arcsec. Errors are the standard deviations of the 1D-marginalized MCMC samples. The three quasars whose hosts were not detected (SDSS J082510.09+031801.4, SDSS J102719.13+584114.3, and SDSS J135851.73+540805.3) have no values for parameters that depend upon the Sérsic component.

SDSS J082510.09+031801.4: The model is well-converged, with individual parameters having > 600 effective samples each. The PSF has lower S/N than the quasar, so the limiting surface brightness is significantly brighter; the host galaxy is effectively not detected. No additional galaxies are visible within a 2″.0 radius.

SDSS J085117.41+301838.7: The model is well-converged, with individual parameters having > 600 effective samples each. The point source-subtracted residual is not significantly asymmetric. No additional galaxies are visible within a 2″.0 radius.

SDSS J094737.70+110843.3: This quasar is the c IV BAL quasar. The model is well-converged, with individual parameters having > 250 effective samples each. The Sérsic profile is significantly offset from the point source (0″.088), likely due to asymmetry in the galaxy's flux distribution, which has faint emission extending to the North. In addition to this extended emission, there are two additional objects within a 2″.0 radius. The residual image also shows what appears to be a weak spiral or bar-like residual pattern.

SDSS J102719.13+584114.3: The model is formally converged by the criteria defined above (all parameters have $\hat{R} < 1.05$, and there are no long-term trends in individual chains). The point source parameters are well-determined with > 600 effective samples. However, the Sérsic profile parameters are poorly constrained (preferring models which minimize the total flux). The host galaxy is effectively not detected. There is one additional galaxy within a 2″.0 radius.

SDSS J113820.35+565652.8: The model is well-converged, with individual parameters having > 1000 effective samples each, with the exception of the Sérsic index and effective radius. Models with large R_e and low n are preferred, possibly due to a worse than average PSF mismatch.

SDSS J120305.42+481313.1: The model is well-converged, with individual parameters having > 400 effective samples each. The center of the Sérsic profile is offset significantly from the point source (0″231). Faint emission extends both East and West of the host. There is one additional galaxy within a 2″0 radius.

SDSS J123011.84+401442.9: After 200,000 samples the model still has not formally converged, although the PSF parameters are well-converged. The quasar is much brighter than the others, and none of the PSF stars has sufficient S/N to model it, and the Sérsic profile may vary without constraint due to the high noise. There appears to be asymmetric emission in the point source-subtracted image, but without a higher S/N star it remains uncertain.

SDSS J124949.65+593216.9: The model is well-converged, with individual parameters having > 1300 effective samples each. The Sérsic profile is significantly offset from the point source, likely due to the model attempting to fit the low surface brightness emission extending asymmetrically Northeast of the point source. There are two additional galaxies within a 2″0 radius, both of which are blended with the host galaxy.

SDSS J131501.14+533314.1: The model is well-converged, with individual parameters having > 2100 effective samples each. There is one additional galaxy within a 2″0 radius.

SDSS J131535.42+253643.9: The model is formally converged by the criteria defined above (all parameters have $\hat{R} < 1.05$, and there are no long-term trends in individual chains). However, the Sérsic profile parameters are highly uncertain. The host galaxy has a highly asymmetric tail visible in the point source-subtracted image (which was masked out). Besides this tail-like structure, there is one additional galaxy within a 2″0 radius, which also exhibits a tail-like structure, extending to the North.

SDSS J135851.73+540805.3: The model is well-converged, with individual parameters have > 380 effective samples each. The models are compact and faint, so the host galaxy is effectively not detected. There are no galaxies within a $2\rlap{.}{''}0$ radius.

SDSS J143645.80+633637.9: This quasar is a radio-loud quasar. The model is well-converged, with individual parameters having > 1000 effective samples each. The quasar is particularly bright, so the lower S/N PSF limits the surface brightness detection limit. There are no additional galaxies within a $2\rlap{.}{''}0$ radius.

SDSS J145645.53+110142.6: The model is formally converged by the criteria defined above (all parameters have $\hat{R} < 1.05$, and there are no long-term trends in individual chains). However, the Sérsic profile parameters are highly uncertain. There are three additional galaxies within a $2\rlap{.}{''}0$ radius.

SDSS J155447.85+194502.7: After 200,000 samples the model still has not formally converged, although the PSF parameters are well-converged. The host galaxy appears significantly asymmetric in the point source-subtracted image, and is blended with two nearby galaxies that fall within a $2\rlap{.}{''}0$ radius.

SDSS J215006.72+120620.6: The model is well-converged, with individual parameters having > 150 effective samples each. The host galaxy emission is extended over a larger area than the other hosts ($0\rlap{.}{''}85 \times 0\rlap{.}{''}65$ ellipse), and is disk-like ($n \simeq 1$). The emission is slightly asymmetric, as visible in the residual image (the northern half is slightly under-subtracted, while the southern half is slightly over-subtracted). There are no additional galaxies within a $2\rlap{.}{''}0$ radius.

SDSS J215954.45−002150.1: The model is well-converged, with individual parameters having > 1800 effective samples each. The quasar is brighter than the others and consequently none of the PSF stars have sufficient S/N to model it well. The detection limit is thus brighter than for the other quasars. No other galaxies are visible within a $2\rlap{.}{''}0$ radius.
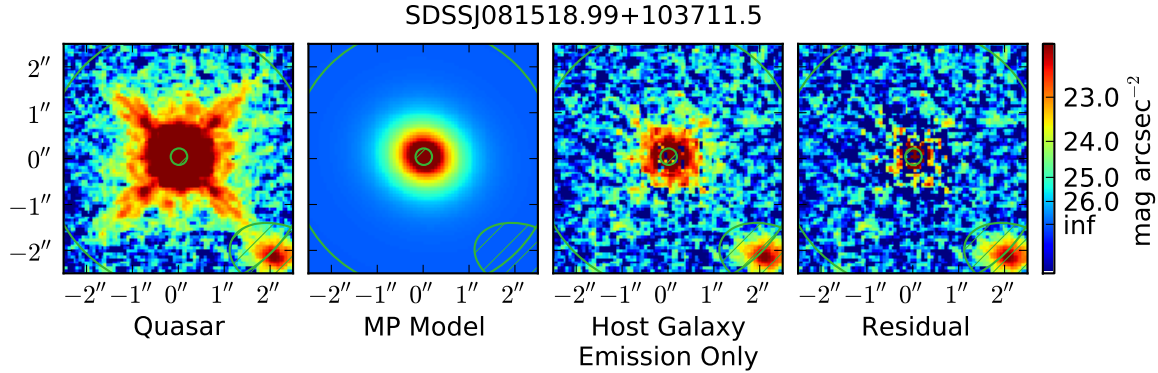
Figure 4.4: Maximum posterior (MP) model for SDSS J081518.99+103711.5. All images are displayed with the same arcsinh color stretch. Far left: Drizzled, undistorted F160W image with $0.''060$ pixels. Middle left: MP model from MCMC fitting process, before convolution with the focus-matched PSF. Middle right: Residual after subtracting only the MP model's point source from the original image. Far right: Residual after subtracting the MP model's point source and host galaxy from the original image. Hatched green regions show areas that were excluded from the fitting, due to saturated PSF cores, nearby galaxies, or other structures not included in the model (e.g., spiral arms, tidal tails, etc.).

SDSS J220811.62−083235.1: The model is well-converged, with individual parameters having > 400 effective samples each. The host emission is disk-like ($n \simeq 1$) and has a bright knot on one side. Other than this knot, there are no additional galaxies within a $2.''0$ radius.

SDSS J232300.06+151002.4: The model is well-converged, with individual parameters having > 250 effective samples each, with the exception of the Sérsic index n, which has only 77 effective samples. There is significant asymmetric extended emission in the form of a bright trail North of the host galaxy. Other than this trail, there are no other galaxies visible within a $2.''0$ radius.
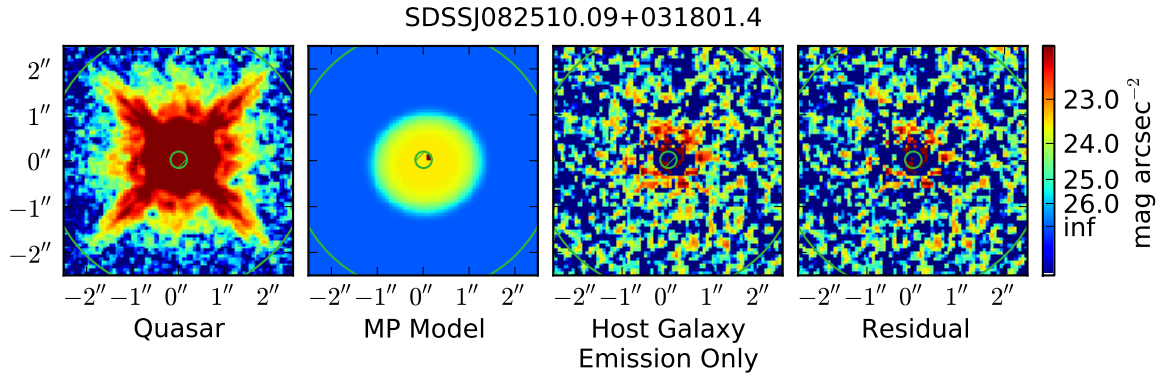
Figure 4.5: Maximum posterior model for SDSS J082510.09+031801.4. For a description of individual panels, see Figure 4.4.
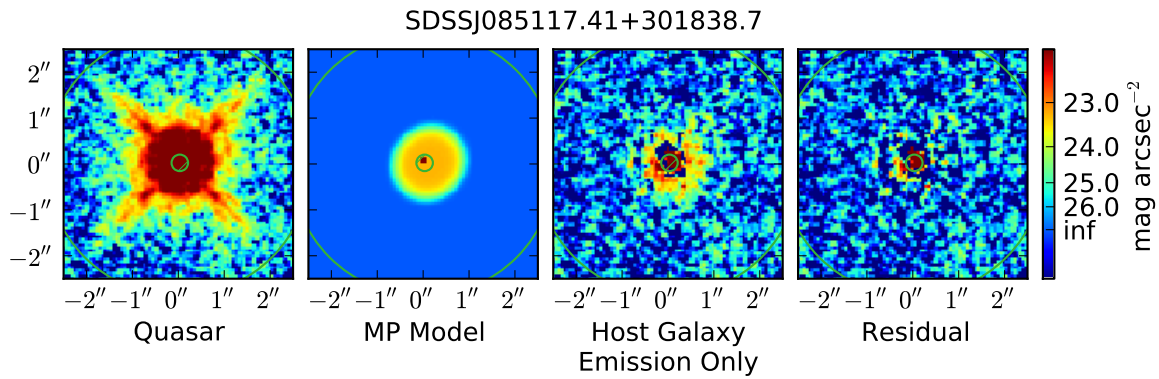

Figure 4.6: Maximum posterior model for SDSS J085117.41+301838.7. For a description of individual panels, see Figure 4.4.

Figure 4.7: Maximum posterior model for SDSS J094737.70+110843.3. For a description of individual panels, see Figure 4.4.

Figure 4.8: Maximum posterior model for SDSS J102719.13+584114.3. For a description of individual panels, see Figure 4.4.

Figure 4.9: Maximum posterior model for SDSS J113820.35+565652.8. For a description of individual panels, see Figure 4.4.

Figure 4.10: Maximum posterior model for SDSS J120305.42+481313.1. For a description of individual panels, see Figure 4.4.

Figure 4.11: Maximum posterior model for SDSS J123011.84+401442.9. For a description of individual panels, see Figure 4.4.

Figure 4.12: Maximum posterior model for SDSS J124949.65+593216.9. For a description of individual panels, see Figure 4.4.
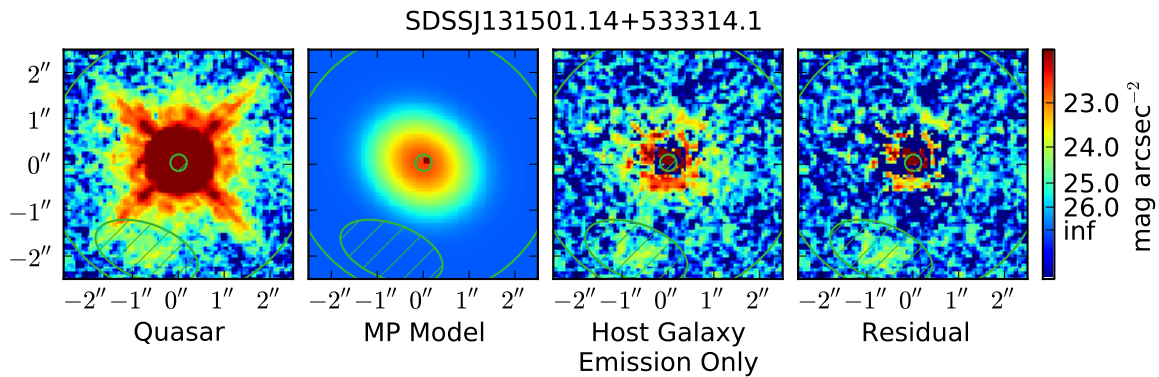
Figure 4.13: Maximum posterior model for SDSS J131501.14+533314.1. For a description of individual panels, see Figure 4.4.
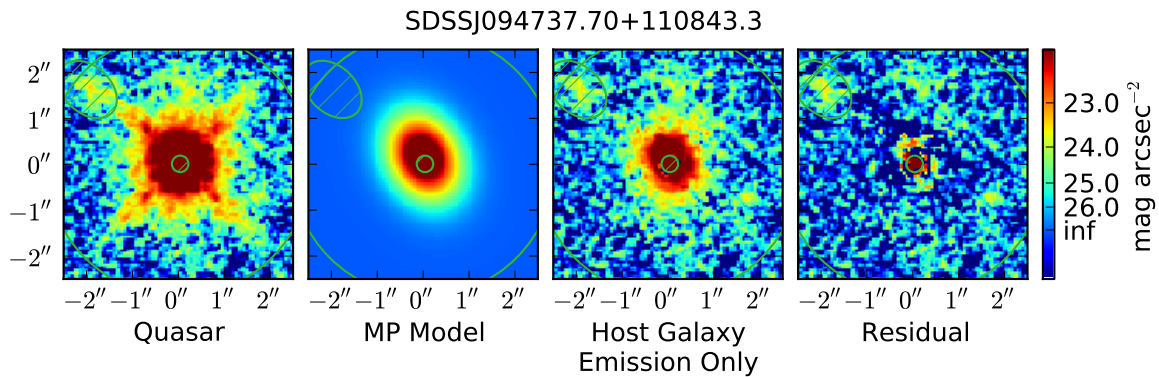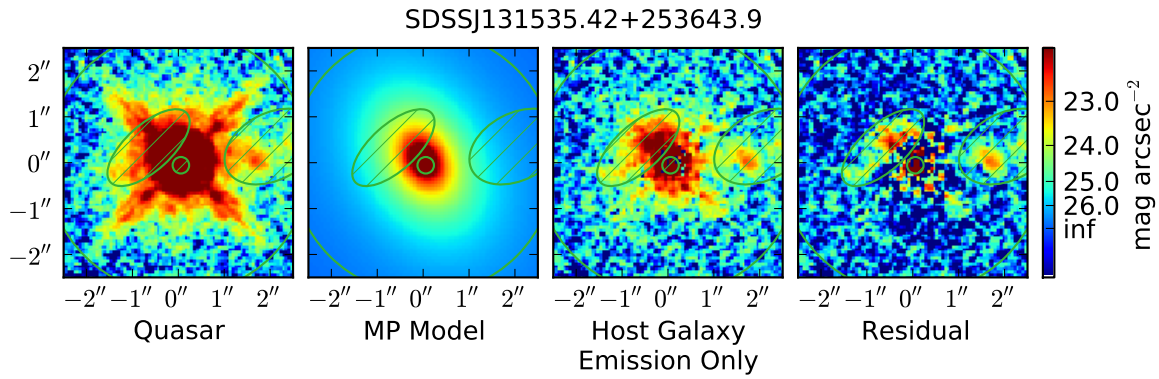
Figure 4.14: Maximum posterior model for SDSS J131535.42+253643.9. For a description of individual panels, see Figure 4.4.
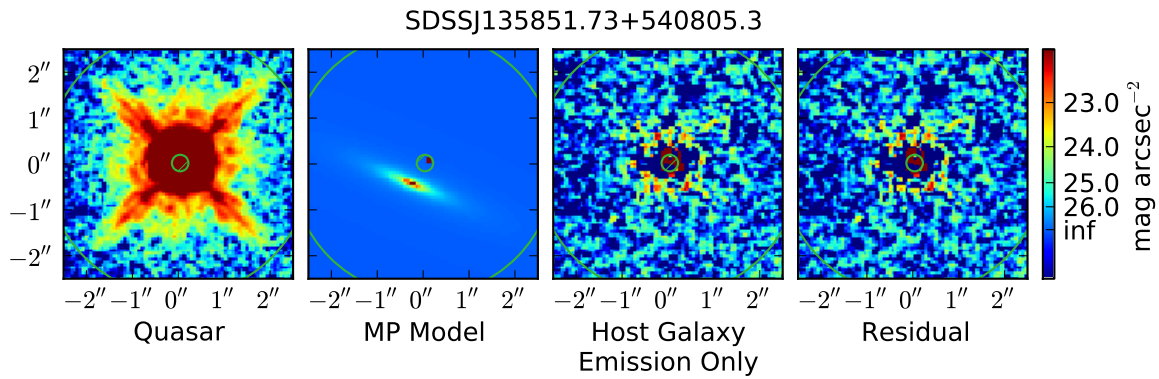
Figure 4.15: Maximum posterior model for SDSS J135851.73+540805.3. For a description of individual panels, see Figure 4.4.
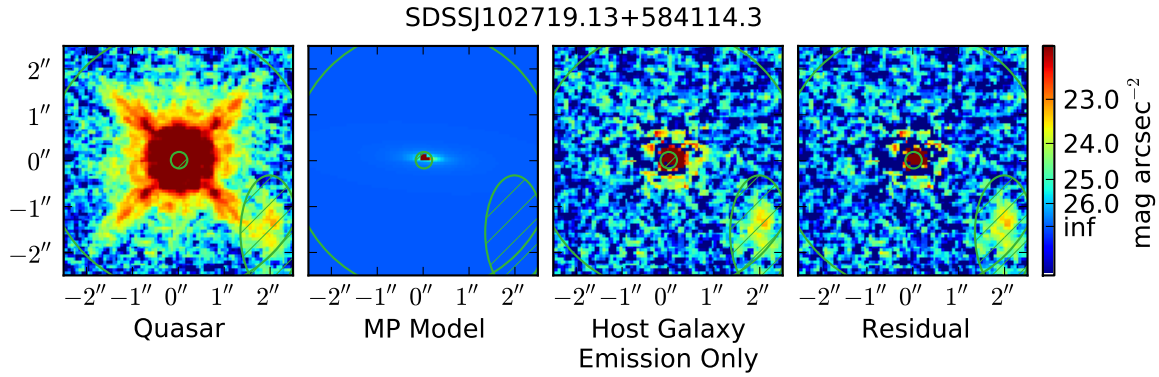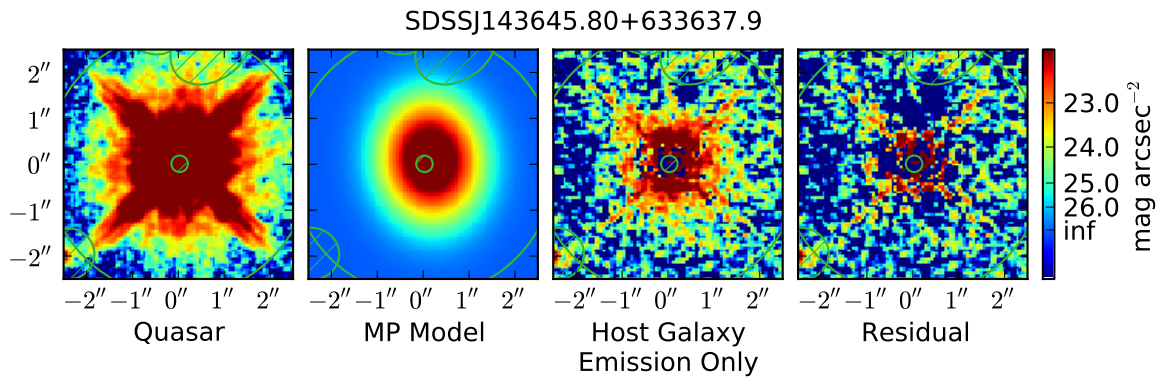

Figure 4.16: Maximum posterior model for SDSS J143645.80+633637.9. For a description of individual panels, see Figure 4.4.
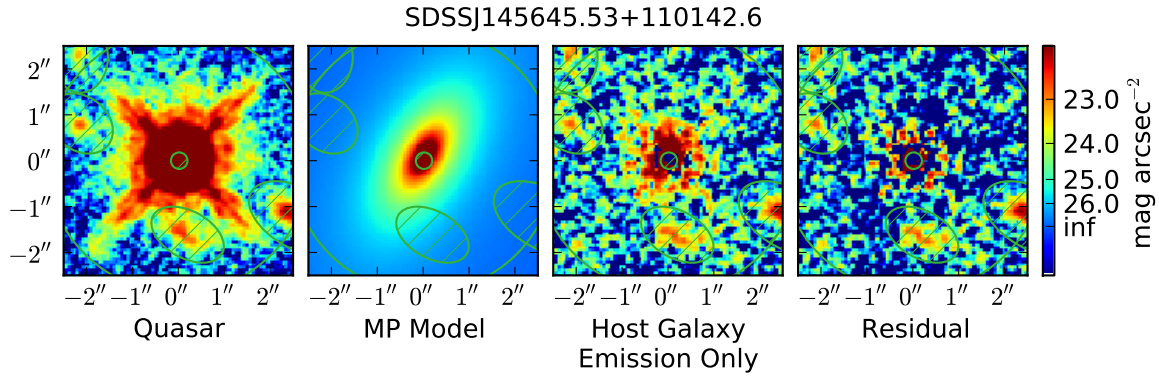
Figure 4.17: Maximum posterior model for SDSS J145645.53+110142.6. For a description of individual panels, see Figure 4.4.
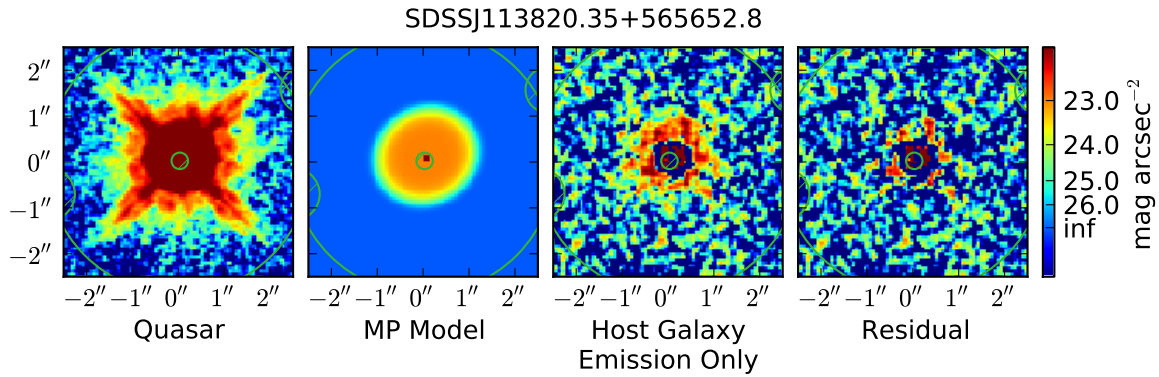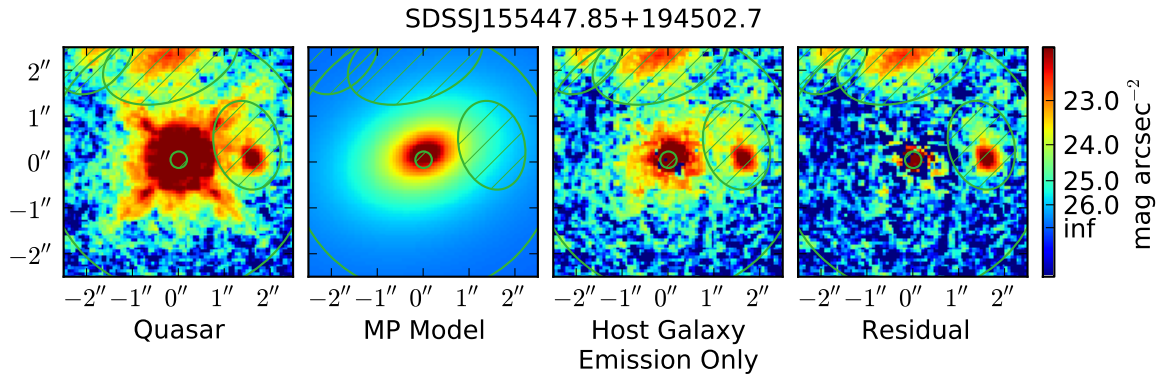

Figure 4.18: Maximum posterior model for SDSS J155447.85+194502.7. For a description of individual panels, see Figure 4.4.

Figure 4.19: Maximum posterior model for SDSS J215006.72+120620.6. For a description of individual panels, see Figure 4.4.
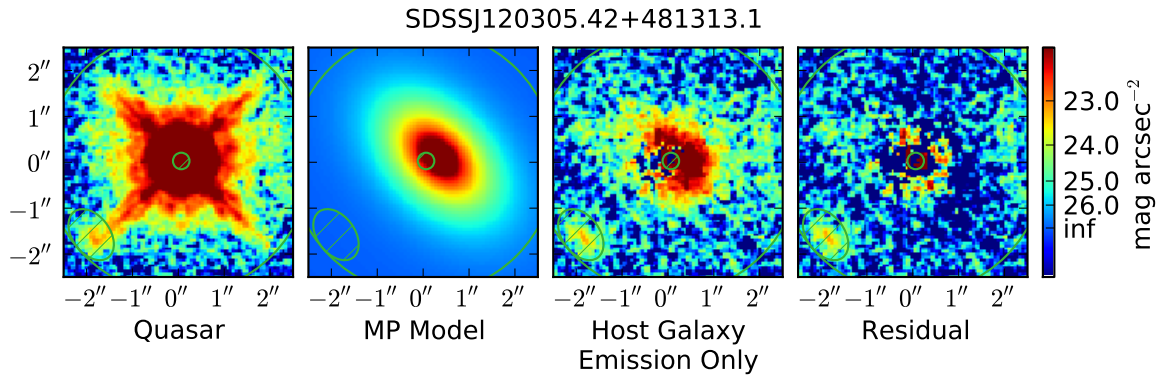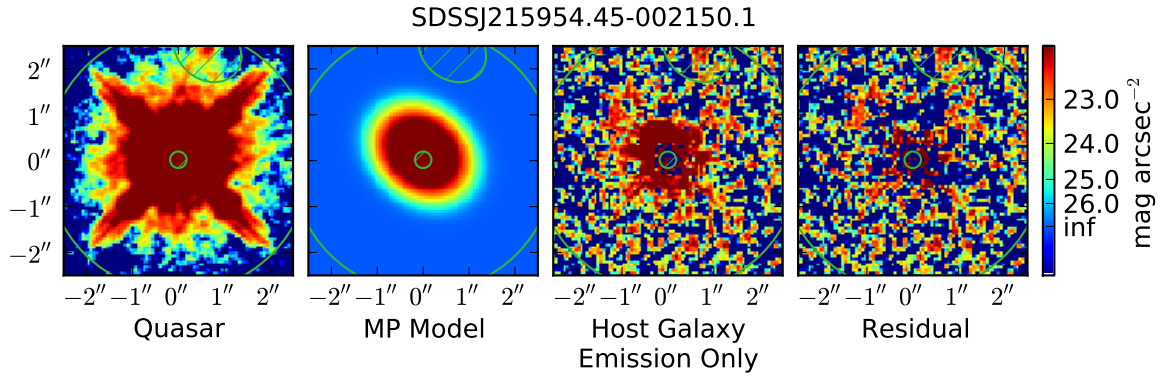
Figure 4.20: Maximum posterior model for SDSS J215954.45−002150.1. For a description of individual panels, see Figure 4.4.
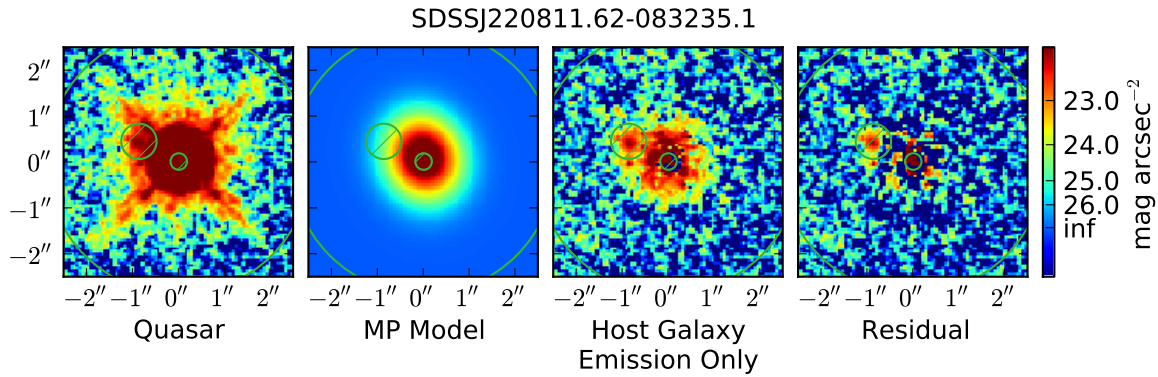
Figure 4.21: Maximum posterior model for SDSS J220811.62−083235.1. For a description of individual panels, see Figure 4.4.
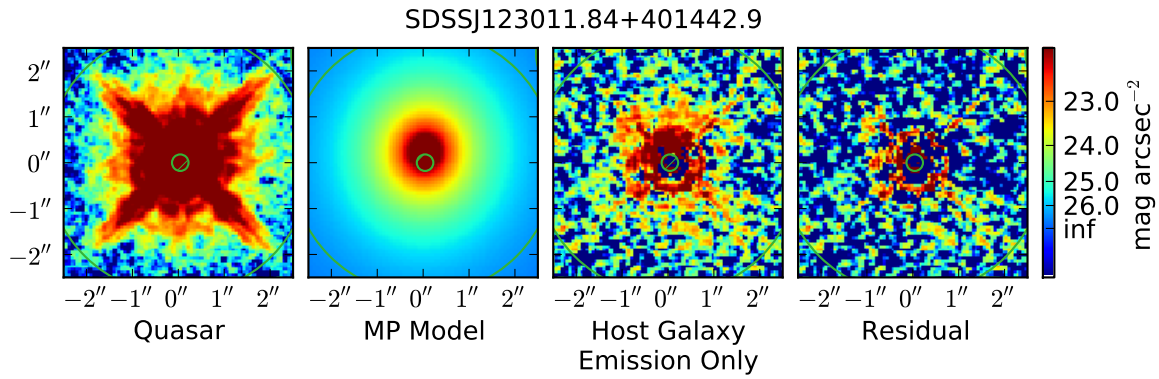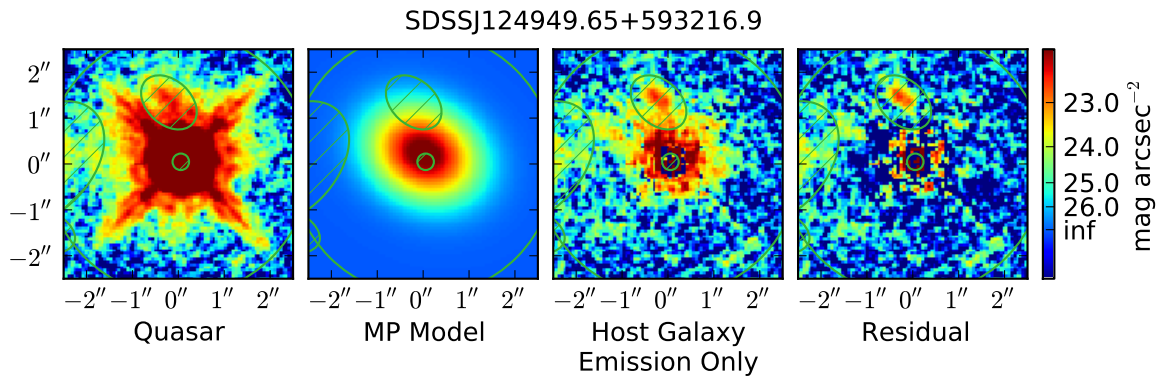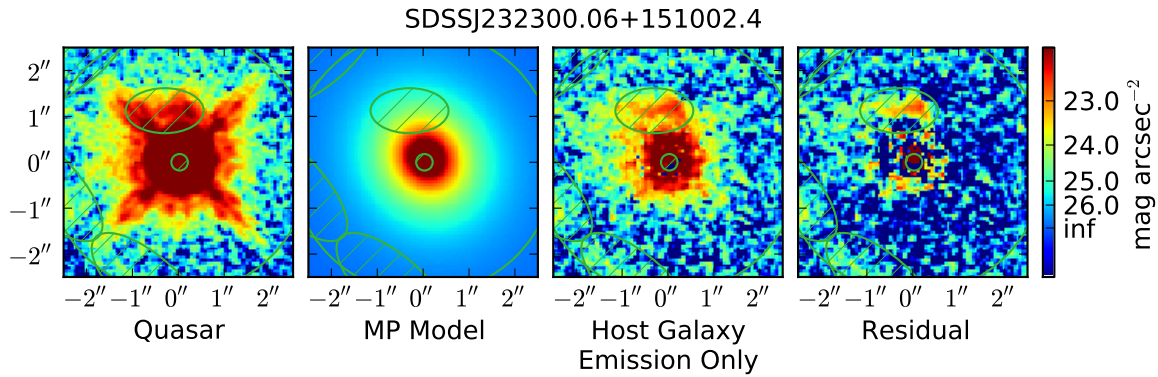

Figure 4.22: Maximum posterior model for SDSS J232300.06+151002.4. For a description of individual panels, see Figure 4.4.
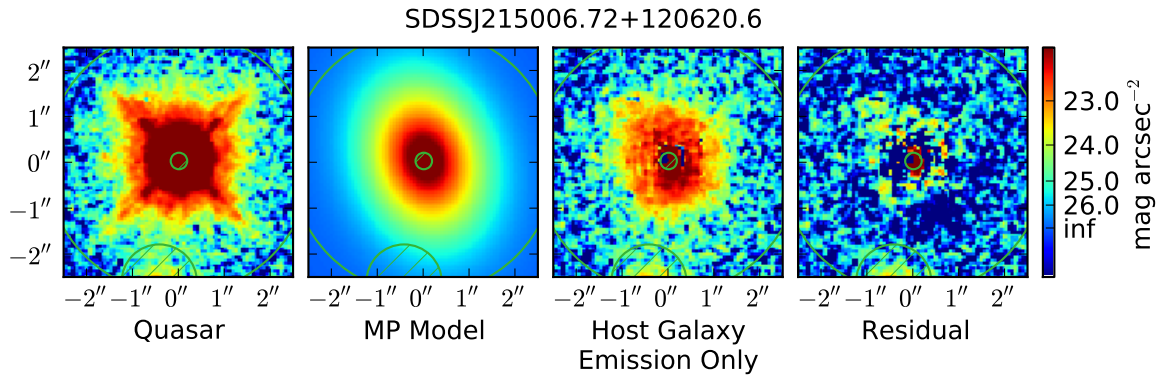
4.7 Discussion

The presence of fits with very high contrast ratios $(m_{host} - m_{AGN})$ means that three host galaxies are too faint to detect, and reduces the number of hosts that can have their morphologies and structures analyzed to sixteen. Two additional hosts do not have converged Sérsic fits, however their point source fits have converged so their structures may still be analyzed, with the caveat that the numeric values of the Sérsic component cannot be used for these two objects.

The offset between point source component and the center of the Sérsic component can be used as a quantitative diagnostic for gravitationally disturbed morphology. In an undisturbed host galaxy, the SMBH should fall within the center of the host galaxy mass (and thus flux) distribution, whether the galaxy is bulge-dominated or disk-dominated. The point source may be significantly offset from the Sérsic profile center for one of two reasons, either of which indicate a disturbed morphology. Either the black hole is physically not at the center of the host gravitational potential (so the system is not relaxed), or the host galaxy flux distribution is so significantly asymmetric or lopsided, so that the best-fit Sérsic profile is offset to better model the total flux distribution. For the purpose of analysis, the offset is considered significant if it is at least one resolution element (i.e., 2 pixels or $0\rlap{.}''12$, corresponding to 1.0 kpc projected distance). Six out of sixteen detected host galaxies meet this criterion, indicating a significantly disturbed morphology. Four additional host galaxies have neighboring companions or knots that fall within the Sérsic profile effective radius, but were masked out prior to fitting. These result in offsets $>0\rlap{.}''12$ if they are not masked, bringing the total number to 10/16, for a merger fraction of $f_M = 0.625$.

Thus, 62.5% of the detected host galaxies have some indication of an ongoing or recent strong interaction. Comparing this to the major merger fraction for inactive

galaxies determined by other studies is not straightforward, since the selection and systematics of other samples are significantly different, especially due to the point source subtraction in the quasars. Careful construction of a comparison sample of inactive galaxies from the CANDELS survey (Grogin et al. 2011) will be the subject of a future paper. For now, I will compare to several other studies of the $z \simeq 2$ merger fraction for inactive galaxies.

Conselice et al. (2003) used a method that measured galaxy asymmetry to identify mergers at $z \leq 3$ in the Hubble Deep Field North. At $z = 1.73$, they found a merger fraction range $f_M = 0.09 - 0.17$, depending upon the rest-frame B-band galaxy luminosities. The quasar host merger fraction as calculated above is significantly higher than this.

Stott et al. (2013) used a morphology-based approach that identified mergers in the HiZELS survey (Sobral et al. 2012) based on the M_{20} coefficient, which quantifies the second-order moment of the brightest 20% of the pixels within a galaxy. They find a merger fraction of $f_M = 0.32$ at $z = 2.23$. Again, the $z \simeq 2$ quasar host merger fraction is roughly double this value.

López-Sanjuan et al. (2013) measured a major merger fraction of $f_{MM} = 0.220^{+0.137}_{-0.073}$ from 12 star-forming galaxies at $1.5 \leq z \leq 1.8$ in the MASSIV sample (Contini et al. 2012). Extrapolating their fit for $f_{MM}(z)$ to $z = 2$ gives an estimate of $f_{MM}(z = 2) = 0.484$. Again, the quasar hosts have a significantly enhanced merger rate compared to these star-forming galaxies. This represents one of the few merger fraction estimates at $z \gtrsim 1.5$ based entirely on galaxies with spectroscopic redshifts. However, the merger criteria are not based on morphologies, but rather the identification of two distinct objects with i-band magnitude difference $\Delta m_i \leq 1.5$, a projected distance $< 20h^{-1}$ kpc, and systematic velocity offset $\Delta v \leq 500$ km s^{-1}.

Again, it is stressed that since a different diagnostic has been used in the quasar study described in this chapter, and the point source subtractions are not perfect, a true comparison sample based on inactive $z \simeq 2$ galaxies and using the same methods is the best way to search for an enhancement to the quasar host merger fraction. However, it is cautiously noted that the quasar host merger fraction found above is significantly higher (by a factor of $\gtrsim 2$) than all of the other studies described above. This suggests that the most luminous AGN at $z \simeq 2$ are indeed preferentially hosted in systems that are merger-driven.

Chapter 5

PRELIMINARY RESULTS OF A HUBBLE SPACE TELESCOPE STUDY OF THE HOST GALAXIES OF UV-FAINT QUASARS AT $z \simeq 6$

5.1 Introduction

Recent deep surveys have begun to find UV-faint quasars at $z \simeq 6$, with observed rest-frame UV emission $m_{1450} > 20.2$ mag (Cool et al. 2006; Jiang et al. 2008, 2009; Willott et al. 2007, 2009, 2010). Most of the known $z \simeq 6$ quasars have properties consistent with low-redshift ULIRGs, including high CO luminosities, copious host dust, high FIR luminosities, and high infrared excess (e.g., Jiang et al. 2006; Wang et al. 2010, 2011; Mechtley et al. 2012; Wang et al. 2013). A study of the rest-frame far-infrared (FIR) continuum emission of these systems by Wang et al. (2011) confirmed five with FIR luminosities $> 10^{12}$ $L_\odot$. The faint UV emission suggests the AGN is not the dominant source of the FIR emission (25% or less using the radio-quiet quasar template of Elvis et al. 1994), but rather dust heated by star formation, with inferred rates of > 500 $M_\odot$ yr^{-1} (Wang et al. 2011). Locally, such galaxies are predominately systems undergoing major mergers (Howell et al. 2010).

Since their AGN emission is fainter than the UV-bright quasar described in Chapter 2, and the host galaxy detection limit is dominated by the shot noise from the point source, these quasars represent the best candidates for host galaxy detections. That is, given their known FIR luminosities, the fainter point sources enable probing fainter rest-frame UV fluxes, to levels consistent with the L_{FIR}/L_{FUV} ratios found in local ULIRGs. My collaborators and I were awarded 25 orbits of HST time to observe these five known UV-faint, FIR-bright $z \simeq 6$ quasars (GO 12974, PI: Mechtley).

5.2   HST Data and Point Source Subtraction

Currently, WFC3IR data have been taken for two of the five targets in the HST program: NDWFS J142516.3+325409 ($z = 5.85$, Cool et al. 2006, hereafter ND-WFS J1425+3254) and SDSS J205406.49000514.8 ($z = 6.06$, Jiang et al. 2008, hereafter SDSS J2054−0005). The observing strategy for this program is similar to that described for SDSS J1148+5251 in Mechtley et al. (2012) and Chapter 2. For each quasar, a PSF star is selected that 1) is within 5°of the quasar, to best match the spacecraft attitude-dependent thermal environment, 2) matches the quasar ($J - H$) color as closely as possible, and 3) has been pre-selected as a non-binary by ground-based adaptive optics imaging. The PSF star is then observed in the HST orbit directly following the quasar observations, with a matched dither pattern and exposures that fully sample spacecraft breathing effects from the the orbital day-night cycle. A four-point half-pixel stepped dither pattern is used, with inter-point spacing of $\simeq 2\rlap{.}''0$ (larger than the standard pattern) to avoid self-persistence from the bright central pixels of the point source. Both PSF star and quasar are imaged in the F125W (J) and F160W (H) bands.

The data were processed similarly to those in Chapters 2 and 4. Flux-calibrated, flat-fielded exposures were obtained from the HST archive, then projected to a common reference grid with $0\rlap{.}''060$ pixels, using the `Multidrizzle` software with inverse variance map weighting and a `pixfrac` parameter of 0.8 (Koekemoer et al. 2002, 2011). The `Multidrizzle` cosmic ray rejection step was turned off, since the WFC3IR up-the-ramp count rate fitting algorithm provided sufficient cosmic ray rejection. The ERR extensions from the HST exposures were used to generate per-pixel RMS error maps that include all sources of error, including shot noise from the objects, and account for correlated noise introduced in the drizzling process.

Once calibrated images were produced, the quasar point sources were subtracted using psfMC, first with a single-component point source model. The point source only model used 4 chains of 20,000 samples each, with the first 10,000 of each discarded as a burn-in period. Convergence was assessed as in Chapter 4; all models had converged. SDSS J2054−0005 had no host galaxy detection, to a limit of $m_J > 23.62$ mag, $m_H > 23.30$ mag ($2\,\sigma$, $0\rlap{.}''5$ radius aperture).

5.3   Detection of a Candidate Host System for NDWFS J1425+3254

The point source subtraction for NDWFS J1425+3254 resulted in a successful detection of a candidate merger system, with two components separated from the host galaxy by $0\rlap{.}''8$ and $1\rlap{.}''4$. The psfMC MCMC fitting was re-run with a point source plus two Sérsic profiles, to better model the blended flux from all three components and avoid oversubtracting the point source. It should be stressed however that if the galaxies are associated with the quasar and undergoing a merger, their rest-frame UV emission need not be distributed in anything like a Sérsic profile. Rather, this is simply used as a way to model their flux to avoid oversubtraction. Models for the detected components in J and H bands, as well as their residuals after point source subtraction, are shown in Figures 5.1 and 5.2.

The results of the fitting process for both components are compiled in Table 5.3. The northeastern component has magnitude $m_H = 24.6$ mag with $(J-H) = 0.1 \pm 0.1$ and is separated from the quasar point source by $1\rlap{.}''4$, or 8 kpc projected distance at $z = 5.85$. The southeastern component has magnitude $m_H = 25.2$ mag with $(J - H) = 0.1 \pm 0.1$ and a separation of $0\rlap{.}''8$ (4.8 kpc). Unfortunately, existing observations at sub-mm to radio wavelengths do not resolve the individual sources, so morphological comparisons were not possible. The photometry along with existing data for this object are plotted in Figure 5.3. The overall SED, including the rest-
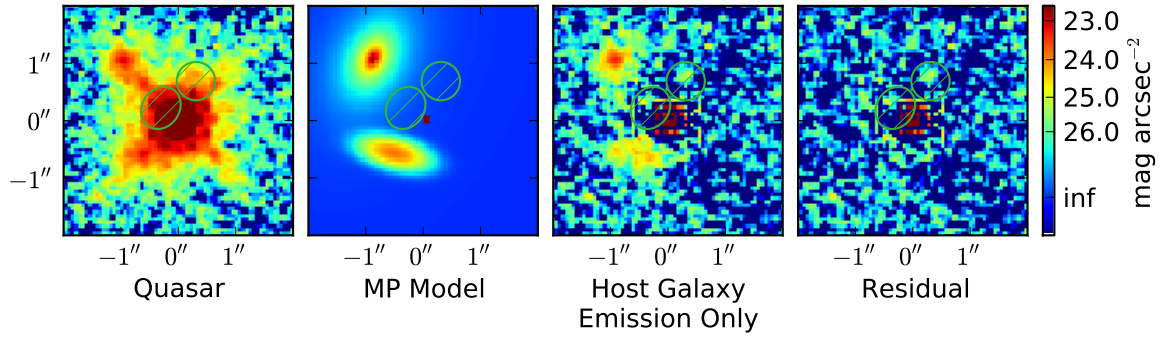
Figure 5.1: HST WFC3 IR channel F125W (J-band) modeling of the combined surface brightness distribution. North is up, East is left, and linear pixel scale is 0″.060 in all images. Far left: HST observation of J1452+3254. Center left: Best-fit (maximum posterior, MP) model from Markov Chain Monte Carlo fitting, before convolution with the PSF. Center right: Best-fit (ML) model after PSF convolution. Far right: Residual after subtracting the point source component only. The hatched green ellipses are regions excluded from the fitting process, due to residual flux not included in the 3-component model.
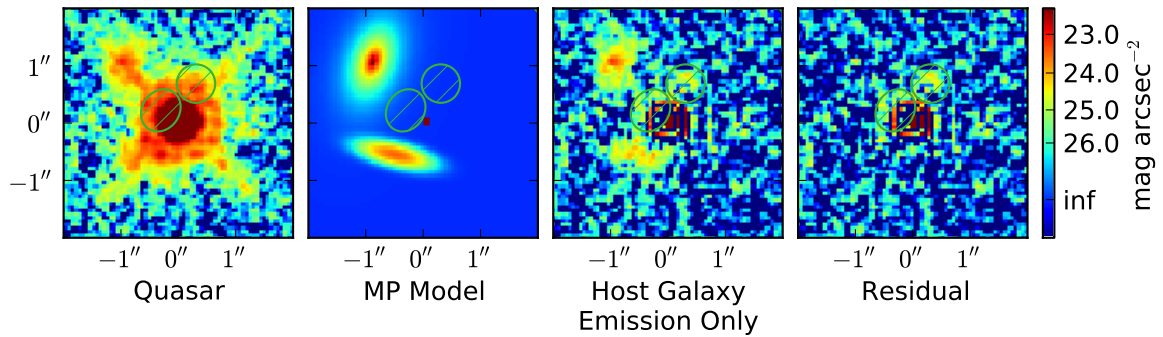
Figure 5.2: Same as Figure 5.1, but for HST WFC3 IR channel F160W (H-band).

frame FIR detection from Wang et al. (2011), is consistent with local ULIRGs, with infrared excess log(IRX) = 1.3 and morphology consistent with a major merger in progress. A third component, coincident with the point source and with higher UV extinction, cannot be ruled out by these data. The dust extinction of local ULIRGs is very dependent upon the line of sight (Scoville et al. 2000). If the AGN is hosted within an extremely dusty region with very high IRX ($\gtrsim$ 2.0), then its UV flux is below the detection limit for these data.
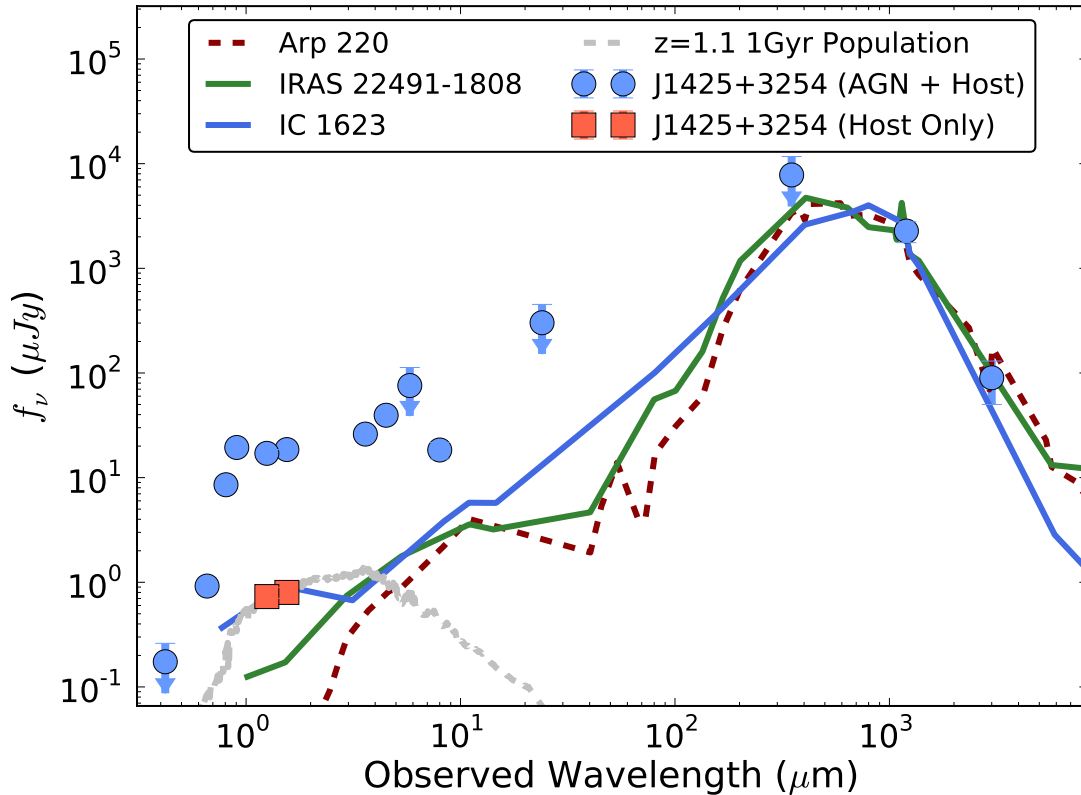
Figure 5.3: Comparison of broadband photometry for NDWFS J1425+3254 with local LIRGs. Blue circles are broad-band (quasar and host system) photometry of J1425+3254 taken from Cool et al. (2006) and this work. Red squares show the measurements for the combined host system components at rest-frame 1800 Å and 2300 Å. The blue, green, and red lines are the SEDs of the local LIRGs IC 1623, IRAS 22491-1808, and Arp 220, respectively, representing LIRGs with increasingly red UV slopes (Howell et al. 2010). Extremely red SEDs like Arp 220 are ruled out, but the galaxy $(J-H)$ colors are consistent with a range of internal reddening values. The dashed grey curve, a 1 Gyr simple stellar population redshifted to $z = 1.1$, serves as a fiducial model for a foreground interloper. The local LIRG SEDs have been normalized to the rest-frame FIR emission of NDWF J1425+3254, and the $z = 1.1$ model has been normalized to match the HST photometry of the merger candidate in J band.

Table 5.1: Summary of Candidate Host System Sérsic Parameter Values

Component	m_{tot}	Index n	R_e (arcsec)	b/a
Northeast (F125W)	24.70 ± 0.10	3.60 ± 0.90	0.40 ± 0.06	0.7 ± 0.2
Southeast (F125W)	25.26 ± 0.08	0.56 ± 0.06	0.39 ± 0.03	0.4 ± 0.1
Northeast (F160W)	24.60 ± 0.10	2.87 ± 1.00	0.38 ± 0.06	0.6 ± 0.2
Southeast (F160W)	25.20 ± 0.08	0.63 ± 0.15	0.36 ± 0.04	0.4 ± 0.1

Notes: Column 1: Component and filter; Column 2: Total (integrated) magnitude; Column 3: Sérsic index n; Column 4: Sérsic effective radius; Column 5: Axis ratio (minor axis/major axis)

The probability that these objects are foreground interlopers can be calculated from wide-field multi-band surveys. Using the data from the WFC3 Early Release Science 2 (ERS2) field (Windhorst et al. 2011), the probability of finding a single galaxy with $(J - H) = 0.1 \pm 0.1$ within $0\rlap{.}''8$ of a uniform random position is 2.5%. The probability of finding two such objects within $1\rlap{.}''4$ of a random position is $< 1\%$.

If they were foreground objects, constraints can also be made from gravitational lensing calculations. Assuming a fiducial foreground redshift of $z = 1.1$ (since they are not detected in the NDWFS B_W band), the fact that no quasar counter-image is seen places a constraint on the stellar mass-to-light ratio of < 0.39 for the northeastern component, and < 3.68 for the southeastern component (R. Barone-Nugent 2013, priv. comm.). This provides a weaker constraint than the probabilistic argument above, but does rule out foreground objects with high mass-to-light ratios.

The discovery spectrum (Cool et al. 2006, Figure 2) also provides evidence for the major merger interpretation. This spectrum has a significant absorption feature at roughly 8350 Å, 20 Å redder than Lyα at 8329.6 Å. A probable source of this feature is foreground H I absorption from a companion galaxy, slightly redshifted because the galaxy is infalling with respect to the quasar. This gives $v/c \simeq 0.002$, or an infall velocity of 720 km s^{-1}. Assuming a virialized system, and using the projected

distance of the southeastern host system component (4.8 kpc), this corresponds to a dynamical mass of $\simeq 5.8 \times 10^{11}\ M_\odot$. For a fiducial dynamical-to-stellar mass ratio of 10 and an estimated star formation rate of $\simeq 500\ M_\odot\ \mathrm{yr}^{-1}$ (Wang et al. 2011), this gives an approximate timescale for stellar mass build-up of 10^8 yr, within the available time budget of $0.5 - 0.7$ Gyr between the epoch of first light at $z \simeq 12 - 15$ and $z = 5.85$.

5.4   Discussion

The five $z \simeq 6$ quasars for this HST program were selected as those optically-faint quasars with confirmed sub-mm detections, and thus the greatest rest-frame L_{FIR}/L_{UV} ratios. This selects for host systems with the greatest non-AGN contribution to the FIR flux, with inferred ULIRG-class FIR luminosities ($> 10^{12}\ L_\odot$). Locally, such systems are largely mergers with high inferred star formation rates (Howell et al. 2010).

After analyzing the first of the HST WFC3IR data using the point-source subtraction method detailed above, we detected a candidate merger system around the quasar NDWFS J1425+3254. While the HST J and H band data provide high-resolution structural constraints, the single $(J - H)$ color provides only limited information about the young stellar population and internal reddening of the candidate merger components, and does not definitively rule out low-redshift interlopers. However, the quasar discovery spectrum does provide independent evidence for a merging system. Follow-up observations will focus on better constraining the spectral energy distributions of both components, including deep r-band imaging to rule out low-redshift interlopers, and adaptive optics-corrected K-band imaging to attempt to better constrain the rest-frame UV SED.

Chapter 6

CONCLUSIONS

6.1   Black Hole Growth and the Quasar Host Merger Fraction

The work presented in Chapters 2 and 5 addresses mechanisms for black hole growth and the coevolution of SMBHs and their host galaxies at the earliest cosmic epoch where AGN are currently known. The non-detection of host galaxy flux coincident with the quasar point sources provides additional evidence that the hosts are starburst ULIRG-like galaxies with extreme reddening and UV extinction. Meanwhile, the merger candidates found surrounding NDWFS J1425+3254 provide a direct example of a star-forming system undergoing mass build-up at high redshift.

The work presented in Chapter 4 is a direct measurement of the quasar host galaxy merger fraction at $z \simeq 2$, currently the earliest epoch for which detailed structural studies of galaxies can be made in rest-frame optical light. More than 60% of the sixteen detected hosts show some indication of ongoing strong gravitational interactions (mergers or tidal disruption). While a carefully selected control sample of galaxies without quasars will be the subject of future work, comparisons with published merger fractions from the literature indicate that this is significantly higher (roughly double) the $z \simeq 2$ merger fraction for field galaxies. This indicates that for the most luminous AGN at $z \simeq 2$, strong galaxy-galaxy interactions are the dominant triggering mechanism for the AGN activity.

6.2 The Quasar-Starburst-Merger Connection

It is clear that, at least for the currently known $z \simeq 6$ quasars, the often-proposed model of quasar hosts as "post-starburst" systems appears too simplistic. In total there are 12 quasars at $z \simeq 6$ with secure sub-mm measurements, as summarized by Wang et al. (2010, 2011, 2013). These represent $> 25\%$ of the known $z \simeq 6$ quasar population, and their FIR luminosities are much too high to be driven by AGN activity alone, especially for the UV-faint quasars described in Chapter 5. The implied ULIRG-like reddening and FUV extinction in the three that have already been observed with HST support this conclusion as well. So, at least at the highest redshifts, it appears that the luminous AGN phase may coincide with, rather than follow, the starburst phase. The true chronology of these stages is likely dependent upon the details of the given system being studied, particularly merger geometry for systems that are merger-driven.

The work presented in Chapter 4 seems to indicate a significant enhancement to the merger fraction among $z \simeq 2$ quasars. Even those host galaxies from the $z \simeq 2$ sample that most resemble giant ellipticals — large, luminous hosts with Sérsic index $\gtrsim 4$, such as SDSS J123011.84+401442.9, SDSS J155447.85+194502.7, and SDSS J232300.06+151002.4 — do not resemble simple passively-evolving elliptical galaxies, but show companions and/or significant disturbances. Perhaps then, quasars, as the most luminous AGN with the highest-mass black holes, require mergers with large gas-rich galaxies as the feeding mechanism for their gluttonous appetites, while lower-luminosity AGN snack on smaller galaxies or the remains of star-formation episodes.

6.3   Future Uses for `psfMC`

The `psfMC` software presented in Chapter 3 represents the most precise treatment of quasar PSF modeling and subtraction to date. Thus, it may potentially improve PSF subtractions for any quasars observed with HST — both from previous and future observations. It has also shown promise as a method for subtracting point sources in ground-based imaging. Preliminary tests subtracting the NDWFS 1425+3254 point source from the ground-based data used by Cool et al. (2006) have resulted in surprisingly clean subtractions, which might be used put an upper limit on the host galaxy flux in the R-, I-, and z-bands, albeit much less stringent than the limit imposed by the higher-resolution HST data. The `psfMC` software also potentially has utility subtracting point sources for stars within our own galaxy, to search for companion objects such as binary stars, exoplanets, or circum-stellar disks. This may especially be useful when multiple-visit methods such as those described in Chapter 1 are deemed too time-intensive, or if subtractions are done on existing data.

6.4   Future Improvements to Synthetic PSFs

The current `TinyTim` WFC3 model PSFs are still based on data from the ground-based thermal vacuum tests, rather than on-orbit data. Now, four years after the camera's installation, there is a plethora of archival data that might be used to improve these models. An improved `TinyTim` that includes a spacecraft breathing model as well as better determinations of the WFC3 field-dependent aberration coefficients could significantly improve quasar point source subtractions. This is especially true of the $z \simeq 2$ quasars in Chapter 4, since they did not have dedicated PSF observations, and since the observed bandpass includes a strong emission line (redshifted Hα) that is certainly not present in any PSF star.

6.5 Future Prospects for High-Redshift Quasar Studies

Within the next decade, high-z quasar studies should see a renaissance if not a revolution. Two observatories in particular stand to significantly improve our understanding of high-z quasar hosts. The Atacama Large Millimeter Array provides unprecedented sensitivity and spatial resolution of the gas and dust within high-z quasars, and early science programs are already pursuing these studies (e.g., Wang et al. 2013). The James Webb Space Telescope (JWST) will be able to explore new frontiers where HST's Wide Field Camera 3 leaves off. With unprecedented sensitivity and spatial resolution at near- and mid-infrared wavelengths, JWST will be able to observe AGN at $z \simeq 6$ and beyond at rest-frame optical wavelengths, such as HST did for the $z \simeq 2$ quasars described in Chapter 4. JWST will also provide the much-needed capability to take NIR spectra of the fainter objects (both quasars and galaxies), without being limited by Earth's atmosphere as ground-based observatories are.

REFERENCES

Abraham, R. G., Tanvir, N. R., Santiago, B. X., et al. 1996, MNRAS, 279, L47

Ando, T. 2007, Biometrika, 94, 443

—. 2011, Am. J. of Math. and Mgmt. Sciences, 31, 13

Angel, J. R. P., & Stockman, H. S. 1980, ARA&A, 18, 321

Antonucci, R. 1993, ARA&A, 31, 473

Antonucci, R. R. J., & Ulvestad, J. S. 1985, ApJ, 294, 158

Bahcall, J. N., Kirhakos, S., Saxe, D. H., & Schneider, D. P. 1997, ApJ, 479, 642

Bahcall, J. N., Kirhakos, S., & Schneider, D. P. 1994, ApJ, 435, L11

—. 1995a, ApJ, 450, 486

—. 1995b, ApJ, 447, L1

Barth, A. J., Martini, P., Nelson, C. H., & Ho, L. C. 2003, ApJL, 594, L95

Barthel, P. D. 1989, ApJ, 336, 606

Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., et al. 2006, AJ, 132, 1729

Beelen, A., Cox, P., Benford, D. J., et al. 2006, ApJ, 642, 694

Bély, P. Y., Hasan, H., & Miebach, M. 1993, Orbital Focus Variations in the Hubble
    Space Telescope, Instrument Science Report SESD-1993-16, STScI, Baltimore

Bennert, N., Canalizo, G., Jungwiert, B., et al. 2008, ApJ, 677, 846

Bertoldi, F., Carilli, C. L., Cox, P., et al. 2003, A&A, 406, L55

Biretta, J. 2012, TinyTIM Modeling of WFC3/IR Images, Instrument Science Report
    WFC3 2012-13, STScI, Baltimore

Blandford, R. D., & McKee, C. F. 1982, ApJ, 255, 419

Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476

Boroson, T. A., & Oke, J. B. 1982, Nature, 296, 397

Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., & Franx, M. 2006, ApJ, 653, 53

Bouwens, R. J., Illingworth, G. D., Oesch, P., et al. 2012, ApJ, 754, 83

Bromm, V., Coppi, P. S., & Larson, R. B. 2002, ApJ, 564, 23

Brotherton, M. S., van Breugel, W., Stanford, S. A., et al. 1999, ApJ, 520, L87

Byun, Y. I., & Freeman, K. C. 1995, ApJ, 448, 563

Cales, S. L., Brotherton, M. S., Shang, Z., et al. 2011, ApJ, 741, 106

Canalizo, G., & Stockton, A. 2000a, ApJ, 528, 201

—. 2000b, AJ, 120, 1750

Canalizo, G., Stockton, A., Brotherton, M. S., & van Breugel, W. 2000, AJ, 119, 59

Cassata, P., Giavalisco, M., Williams, C. C., et al. 2013, ApJ, 775, 106

Cisternas, M., Jahnke, K., Inskip, K. J., et al. 2011, ApJ, 726, 57

Cohen, S. H., Ryan, Jr., R. E., Straughn, A. N., et al. 2006, ApJ, 639, 731

Conselice, C. J., Bershady, M. A., Dickinson, M., & Papovich, C. 2003, AJ, 126, 1183

Contini, T., Garilli, B., Le Fvre, O., et al. 2012, A&A, 539, A91

Cool, R. J., Kochanek, C. S., Eisenstein, D. J., et al. 2006, AJ, 132, 823

Cox, C., & Niemi, S.-M. 2011, Evaluation of a temperature-based HST focus model, Instrument Science Report TEL 2011-01, STScI, Baltimore

de Jong, R. S. 1996, A&AS, 118, 557

Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Nature, 457, 451

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

Dickinson, M., Stern, D., Giavalisco, M., et al. 2004, ApJL, 600, L99

Disney, M. J., Boyce, P. J., Blades, J. C., et al. 1995, Nature, 376, 150

Driver, S. P., Fernández-Soto, A., Couch, W. J., et al. 1998, ApJL, 496, L93

Driver, S. P., Windhorst, R. A., Ostrander, E. J., et al. 1995, ApJL, 449, L23

Duc, P.-A., Brinks, E., Wink, J. E., & Mirabel, I. F. 1997, A&A, 326, 537

Dunlop, J. S., McLure, R. J., Kukula, M. J., et al. 2003, MNRAS, 340, 1095–1135

Elvis, M., Wilkes, B. J., McDowell, J. C., et al. 1994, ApJS, 95, 1

Fan, X., Strauss, M. A., Schneider, D. P., et al. 2003, AJ, 125, 1649

Ferreras, I., Pasquali, A., Khochfar, S., et al. 2012, AJ, 144, 47

Floyd, D. J. E., Kukula, M. J., Dunlop, J. S., et al. 2004, MNRAS, 355, 196

Friedman, J. H., & Fisher, N. I. 1999, Statistics and Computing, 9, 123–143

Gelman, A., Carlin, J. B., Stern, H. S., et al. 2011, Bayesian data analysis, 3rd edn. (Boca Raton, Fla.; London: Chapman & Hall/CRC)

Gelman, A., & Rubin, D. B. 1992, J. Stat. Sci., 7, pp. 457

Geweke, J. 1992, in Bayesian Statistics 4, ed. J. Bernardo, J. Berger, A. Dawid, & A. Smith (Valencia, Spain: Oxford University Press), 169–193

Geyer, C. J. 1992, Stat. Sci., 7, 473

Glazebrook, K., Ellis, R. S., Santiago, B., & Griffiths, R. 1995, MNRAS, 275, 19L

Gozanga, S., Hack, W., Fruchter, A. S., & Mack, J. 2012, The DrizzlePac Handbook, 1st edn. (Baltimore, Maryland: STScI)

Greene, J. E., & Ho, L. C. 2005, ApJ, 630, 122

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35

Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, ApJS, 207, 24

Haario, H., Saksman, E., & Tamminen, J. 2001, Bernoulli, 7, 223242

Häring, N., & Rix, H.-W. 2004, ApJL, 604, L89

Hartig, G. F. 2008, WFC3 IR PSF Evaluation in Thermal-Vacuum Test #3, Instrument Science Report WFC3 2008-41, STScI, Baltimore

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

Heckman, T. M. 1980, A&A, 87, 152164

Heckman, T. M., Bothun, G. D., Balick, B., & Smith, E. P. 1984a, AJ, 89, 958

Heckman, T. M., Miley, G. K., & Green, R. F. 1984b, ApJ, 281, 525

Hershey, J. L. 1998, MODELLING HST FOCAL-LENGTH VARIATIONS, Instrument Science Report SESD-1997-01, STScI, Baltimore

Hewett, P. C., Warren, S. J., Leggett, S. K., & Hodgkin, S. T. 2006, MNRAS, 367, 454

Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19

Hook, R., & Stoehr, F. 2008, WFC3 Support in Tiny Tim, Instrument Science Report WFC3 2008-14, STScI, Baltimore

Hoopes, C. G., Heckman, T. M., Salim, S., et al. 2007, ApJS, 173, 441

Howell, J. H., Armus, L., Mazzarella, J. M., et al. 2010, ApJ, 715, 572

Hutchings, J. B., Frenette, D., Hanisch, R., et al. 2002, AJ, 123, 2936

Illingworth, G. D., Magee, D., Oesch, P. A., et al. 2013, ApJS, 209, 6

Iwamuro, F., Kimura, M., Eto, S., et al. 2004, ApJ, 614, 69

Jahnke, K., Kuhlbrodt, B., & Wisotzki, L. 2004a, MNRAS, 352, 399

Jahnke, K., & Macciò, A. V. 2011, ApJ, 734, 92

Jahnke, K., Wisotzki, L., Courbin, F., & Letawe, G. 2007, MNRAS, 378, 23

Jahnke, K., Sanchez, S. F., Wisotzki, L., et al. 2004b, ApJ, 614, 568

Jiang, L., Fan, X., Ivezi, ., et al. 2007, ApJ, 656, 680

Jiang, L., Fan, X., Hines, D. C., et al. 2006, AJ, 132, 2127

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

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

Kaviraj, S., Cohen, S., Windhorst, R. A., et al. 2012, MNRAS, 429, L40

Keel, W. C., Kennicutt, R. C., J., Hummel, E., & van der Hulst, J. M. 1985, AJ, 90, 708

Keel, W. C., & Wehrle, A. E. 1993, AJ, 106, 236

Keeton, C. R., Falco, E. E., Impey, C. D., et al. 2000, ApJ, 542, 74

Kennicutt, R. C. 1998, ARA&A, 36, 189

Kishimoto, M., Antonucci, R., Blaes, O., et al. 2008, Nature, 454, 492

Koekemoer, A. M., Fruchter, A. S., Hook, R. N., & Hack, W. 2002, in The 2002 HST Calibration Workshop: Hubble after the Installation of the ACS and the NICMOS Cooling System (Baltimore: STScI), 337

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36

Koekemoer, A. M., Ellis, R. S., McLure, R. J., et al. 2013, ApJS, 209, 3

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18

Kormendy, J., & Richstone, D. 1995, ARA&A, 33, 581

Krist, J. E., Hook, R. N., & Stoehr, F. 2011, in Optical Modeling and Performance Predictions V., ed. M. A. Kahan, Vol. 8127, San Diego, CA, 81270J–16

Kristian, J. 1973, ApJ, 179, L61

Kuhlbrodt, B., Wisotzki, L., & Jahnke, K. 2004, MNRAS, 349, 1027

Kukula, M. J., Dunlop, J. S., McLure, R. J., et al. 2001, MNRAS, 326, 1533

Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007, ApJ, 660, 770

Larson, R. B., & Tinsley, B. M. 1978, ApJ, 219, 46

Lawrence, A., Rowan-Robinson, M., Leech, K., Jones, D. H. P., & Wall, J. V. 1989, MNRAS, 240, 329

Li, Y., Hernquist, L., Robertson, B., et al. 2007, ApJ, 665, 187

Longair, M. S. 1998, Galaxy formation (Berlin: Springer)

López-Sanjuan, C., Fvre, O. L., Tasca, L. A. M., et al. 2013, A&A, doi:10.1051/0004-6361/201220286

Lupton, R., Blanton, M. R., Fekete, G., et al. 2004, PASP, 116, 133

Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285

Maiolino, R., Cox, P., Caselli, P., et al. 2005, A&A, 440, L51

Malhotra, S., Rhoads, J. E., Pirzkal, N., et al. 2005, ApJ, 626, 666

Malkan, M. A., & Sargent, W. L. W. 1982, ApJ, 254, 22

Marble, A. R., Hines, D. C., Schmidt, G. D., et al. 2003, ApJ, 590, 707

Marconi, A., & Hunt, L. K. 2003, ApJ, 589, L21

Martel, A. R., Ford, H. C., Tran, H. D., et al. 2003, AJ, 125, 2964

McGill, K. L., Woo, J., Treu, T., & Malkan, M. A. 2008, ApJ, 673, 703

McLeod, K. K., & Bechtold, J. 2009, ApJ, 704, 415

McLeod, K. K., & McLeod, B. A. 2001, ApJ, 546, 782

McLeod, K. K., & Rieke, G. H. 1994, ApJ, 420, 58

—. 1995, ApJ, 454, L77

McLure, R. J., & Dunlop, J. S. 2004, MNRAS, 352, 1390

McLure, R. J., & Jarvis, M. J. 2002, MNRAS, 337, 109

McLure, R. J., Kukula, M. J., Dunlop, J. S., et al. 1999, MNRAS, 308, 377–404

Mechtley, M., Windhorst, R. A., Ryan, R. E., et al. 2012, ApJ, 756, L38

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, J. Chem. Phys., 21, 1087

Miley, G. 1980, ARA&A, 18, 165

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

Netzer, H., Lutz, D., Schweitzer, M., et al. 2007, ApJ, 666, 806

Oesch, P. A., Stiavelli, M., Carollo, C. M., et al. 2007, ApJ, 671, 1212

Oke, J. B. 1963, Nature, 197, 1040

—. 1974, ApJS, 27, 21

Oke, J. B., & Sandage, A. 1968, ApJ, 154, 21

Oliphant, T. E. 2006, A Guide to NumPy, Vol. 1 (Provo, UT: Trelgol Publishing USA)

Orr, M. J. L., & Browne, I. W. A. 1982, MNRAS, 200, 1067

Overzier, R. A., Heckman, T. M., Wang, J., et al. 2011, ApJ, 726, L7

Patil, A., Huard, D., & Fonnesbeck, C. J. 2010, J. Stat. Softw., 35, 1

Pence, W. D., Chiappetti, L., Page, C. G., Shaw, R. A., & Stobie, E. 2010, A&A, 524, A42

Peng, C. Y. 2007, ApJ, 671, 1098

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

—. 2010, AJ, 139, 2097

Peng, C. Y., Impey, C. D., Rix, H., et al. 2006, ApJ, 649, 616

Percival, W. J., Miller, L., McLure, R. J., & Dunlop, J. S. 2001, MNRAS, 322, 843

Peterson, B. M., Ferrarese, L., Gilbert, K. M., et al. 2004, ApJ, 613, 682

Pignatelli, E., Fasano, G., & Cassata, P. 2006, A&A, 446, 373

Pirzkal, N., Xu, C., Malhotra, S., et al. 2004, ApJS, 154, 501

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. T. 2007, Numerical recipes: the art of scientific computing, 3rd edn. (Cambridge, UK ; New York: Cambridge University Press)

Ridgway, S. E., Heckman, T. M., Calzetti, D., & Lehnert, M. 2001, ApJ, 550, 122

Riechers, D. A., Walter, F., Bertoldi, F., et al. 2009, ApJ, 703, 1338

Robson, I., Priddey, R. S., Isaak, K. G., & McMahon, R. G. 2004, MNRAS, 351, L29

Ryan, R. E., McCarthy, P. J., Cohen, S. H., et al. 2012, ApJ, 749, 53

Ryan, Jr., R. E., Hathi, N. P., Cohen, S. H., et al. 2007, ApJ, 668, 839

Ryden, B. S. 2003, Introduction to cosmology (San Francisco: Addison-Wesley)

Sanders, D. B., Soifer, B. T., Elias, J. H., et al. 1988, ApJ, 325, 74

Scheuer, P. A. G., & Readhead, A. C. S. 1979, Nature, 277, 182

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

Schmidt, M. 1963, Nature, 197, 1040

Schneider, D. P., Hall, P. B., Richards, G. T., et al. 2007, AJ, 134, 102

Schneider, D. P., Richards, G. T., Hall, P. B., et al. 2010, AJ, 139, 2360

Schweitzer, M., Lutz, D., Sturm, E., et al. 2006, ApJ, 649, 79

Schweitzer, M., Groves, B., Netzer, H., et al. 2008, ApJ, 679, 101

Scoville, N. Z., Evans, A. S., Thompson, R., et al. 2000, AJ, 119, 991

Seyfert, C. K. 1943, ApJ, 97, 28

Shang, Z., Brotherton, M. S., Wills, B. J., et al. 2011, ApJS, 196, 2

Shaw, M. A., & Gilmore, G. 1989, MNRAS, 237, 903

Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, ApJS, 194, 45

Shields, G. A. 1978, Nature, 272, 706

Simard, L. 1998, in A.S.P. Conference Series, Vol. 145, Astronomical Data Analysis
    Software and Systems VII, ed. R. Albrecht, R. N. Hook, & H. Bushouse, 108

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

Sobral, D., Smail, I., Best, P. N., et al. 2012, MNRAS, 428, 1128

Soifer, B. T., Rowan-Robinson, M., Houck, J. R., et al. 1984, ApJ, 278, L71

Spergel, D. N., Bean, R., Dore, O., et al. 2007, ApJS, 170, 377

Springel, V., Di Matteo, T., & Hernquist, L. 2005a, MNRAS, 361, 776

Springel, V., White, S. D. M., Jenkins, A., et al. 2005b, Nature, 435, 629

Stein, W. A., O'Dell, S. L., & Strittmatter, P. A. 1976, ARA&A, 14, 173

Stockton, A., Canalizo, G., & Ridgway, S. E. 1999, ApJ, 519, L131

Stott, J. P., Sobral, D., Smail, I., et al. 2013, MNRAS, 430, 1158

Targett, T. A., Dunlop, J. S., & McLure, R. J. 2012, MNRAS, 420, 3621

Targett, T. A., Dunlop, J. S., Cirasuolo, M., et al. 2013, MNRAS, 432, 2012

Taylor, G. L., Dunlop, J. S., Hughes, D. H., & Robson, E. I. 1996, MNRAS, 283, 930

Thompson, M. B. 2010, A Comparison of Methods for Computing Autocorrelation Time, Statistics Department Technical Report 1007, University of Toronto, Toronto

Thompson, R. I., Illingworth, G., Bouwens, R., et al. 2005, AJ, 130, 1

Toomre, A., & Toomre, J. 1972, ApJ, 178, 623

Urrutia, T., Lacy, M., & Becker, R. H. 2008, ApJ, 674, 80

van Dokkum, P. G., & Brammer, G. 2010, ApJL, 718, L73

van Rossum, G., & Drake, F. L. 1990, The Python Language Reference

Veilleux, S., Rupke, D. S. N., Kim, D.-C., et al. 2009, ApJS, 182, 628

Vestergaard, M. 2002, ApJ, 571, 733

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

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

Volonteri, M., & Rees, M. J. 2006, ApJ, 650, 669

Wadadekar, Y., Robbason, B., & Kembhavi, A. 1999, AJ, 117, 1219

Walter, F., Riechers, D., Cox, P., et al. 2009, Nature, 457, 699

Walter, F., Bertoldi, F., Carilli, C., et al. 2003, Nature, 424, 406

Wandel, A. 1999, ApJ, 527, 649

Wang, J.-G., Dong, X.-B., Wang, T.-G., et al. 2009, ApJ, 707, 1334

Wang, R., Carilli, C. L., Neri, R., et al. 2010, ApJ, 714, 699

Wang, R., Wagg, J., Carilli, C. L., et al. 2011, AJ, 142, 101

—. 2013, ApJ, 773, 44

Wells, D. C., Greisen, E. W., & Harten, R. H. 1981, A&AS, 44, 363

White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D. 1997, ApJ, 475, 479

Willott, C. J., McLure, R. J., & Jarvis, M. J. 2003, ApJ, 587, L15

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

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

—. 2010, AJ, 139, 906

Windhorst, R., Hathi, N., Cohen, S., et al. 2008, Adv. in Space Research, 41, 1965

Windhorst, R. A., Cohen, S. H., Hathi, N. P., et al. 2011, ApJS, 193, 27

Wright, E. L. 2006, PASP, 118, 1711

Yan, H., & Windhorst, R. A. 2004, ApJ, 600, L1

Yoon, I., Weinberg, M. D., & Katz, N. 2011, MNRAS, 414, 1625

Zakamska, N. L., Strauss, M. A., Krolik, J. H., et al. 2006, AJ, 132, 1496

Zernike, F. 1934, MNRAS, 94, 377

APPENDIX A

EXAMPLE PSFMC MODEL DEFINITION FILE

Listing A.1 is the `psfMC` model definition file used for the host system of the $z = 5.85$ quasar NDWFS J1425+3254 in Chapter 5. This demonstrates the inclusion of multiple model components (one PSF component, two Sérsic components, and the background sky as a free parameter), and the use of various probability distributions as parameter priors (Uniform, Normal, TruncatedNormal). It also shows how the user can create custom variables that can be used in calculations (zp, totalflux, and center). Lines beginning with the # symbol are comments.

Listing A.1: Example `psfMC` Model File

```python
from numpy import array
from psfMC.ModelComponents import Sky, PSF, Sersic
from psfMC.distributions import Normal, Uniform, TruncatedNormal

# Since the model is a Python file, we can define variables
zp = 26.2303
totalflux = -5.5 + zp

# We can treat the sky as unknown if the subtraction is uncertain
Sky(adu=Normal(mu=0, tau=100))

# Point source component
center = array((64, 64))
PSF(xy=Normal(mu=center, tau=(1.0, 1.0)),
    mag=Uniform(lower=totalflux - 0.2, upper=totalflux + 1.5))

# First Sersic profile, modeling the first host component
center = array((49, 83))
Sersic(xy=Uniform(lower=center-(4, 4), upper=center+(4, 4)),
       mag=TruncatedNormal(mu=-1.14+zp, tau=1.0, a=19, b=28),
       reff=Uniform(lower=2.0, upper=8),
       reff_b=Uniform(lower=2.0, upper=8),
       index=Uniform(lower=0.5, upper=8),
       angle=Uniform(lower=0, upper=180), angle_degrees=True)

# Second Sersic profile, modeling the second host component
center = array((57, 54))
Sersic(xy=Uniform(lower=center-(4, 4), upper=center+(4, 4)),
       mag=Uniform(lower=22, upper=26),
       reff=Uniform(lower=2.0, upper=8),
       reff_b=Uniform(lower=2.0, upper=8),
       index=Uniform(lower=0.5, upper=8),
       angle=Uniform(lower=155, upper=175), angle_degrees=True)
```

APPENDIX B

APPRECIATING HUBBLE AT HYPERSPEED: AN INTERACTIVE
COSMOLOGY VISUALIZATION TOOL

We have developed a Java-based teaching tool, "Appreciating Hubble at Hyperspeed" ("AHaH"), intended for use by students and instructors in beginning astronomy and cosmology courses, which we have distributed via the World Wide Web. [1] This tool lets the user hypothetically traverse the Hubble Ultra Deep Field in three dimensions at over $\sim 500 \times 10^{12}$ times the speed of light, from redshifts $z = 0 - 6$. Users may also view the Universe in various cosmology configurations and two different geometry modes — standard geometry that includes expansion of the Universe, and a static pseudo-Euclidean geometry for comparison. In this paper we detail the mathematical formulae underlying the functions of this Java application, and provide justification for the use of these particular formulae. These include the manner in which angular sizes of objects are calculated in various cosmologies, as well as how the application's coordinate system is defined. We also briefly discuss the methods used to select and prepare the images in the application, the data used to measure the redshifts of the galaxies, and the qualitative implications of the visualization — that is, what exactly users see when they "move" the virtual telescope through the simulation.

B.1   Introduction

In beginning astronomy courses, many non-science majors appear to have a significant lack of understanding — even after taking the introductory courses — of basic concepts such as wavelength, electro-magnetic spectrum, the speed of light, lookback time, redshift, and expansion of the Universe. We believe this lack of concept acquisition or retention represents a significant shortcoming of the currently available teaching tools. While pictures, figures, and other static media are certainly effective at communicating many concepts, they tend to be poor at showing effects in three

---

[1]http://www.asu.edu/clas/hst/www/ahah/download.html

94

dimensions or that evolve over time. Since virtually all cosmological effects require very large time or distance scales to become apparent, a different teaching medium is preferable.

"Appreciating Hubble at Hyper-speed" is an educational tool that aims to address these issues of concept acquisition and retention by providing a visual and interactive learning medium. The project uses data from the HST Cycle 12 Project "GRAPES" (Grism-ACS Program for Extragalactic Science; Pirzkal et al. 2004) to build a redshift-sorted database of over 5000 galaxies within the Hubble Ultra Deep Field (HUDF). These galaxies range from redshift $z \simeq 0.05$ to $z \simeq 6$, spanning nearly 90% of the history of the Universe (Yan & Windhorst 2004; Malhotra et al. 2005; Bouwens et al. 2006). Since these data represent the deepest optical image of the Universe ever obtained, they are uniquely suited to helping students understand the effects of the expanding Universe and the evolution of galaxies across cosmic time.

B.2 Data Selection and Preparation

We first created a custom-balanced RGB version of the HUDF image. While the image provided in the original press releases would have been adequate, it has the undesirable characteristic that very bright areas, such as bulges in large spirals, appear burnt-out and lack fine detail. The raw HUDF data consist of 56 HST orbits each in the B-, and V-bands, and 144 orbits each in the i'-, and z'-bands (Beckwith et al. 2006), so we created a three-channel color image by first combining the B- and V-bands, applying weights based on the sky SNR (0.235 weight in B and 0.765 weight in V). We then used the algorithm developed by Lupton et al. (2004) to create the combined RGB image, with the combined $B + V$-bands as the blue channel, the

i'-band as the green channel, and the z'-band as the red channel. [2] Besides showing more detail in bright areas, this method has the added benefit that an object with a specified astronomical color has a unique color in the composite RGB image. A comparison of the original STScI color images and our prepared images is shown in Figure B.1. The full HUDF image using this color preparation technique is also available as an interactive map on the World Wide Web. [3]

The galaxies represented in the AHaH application were i'-band selected using `SExtractor` with a detection threshold of $\sigma = 1.8$ above sky. The i'-band dropouts of Yan & Windhorst (2004) were later added by hand. We then created color JPEG "stamp" images for each individual object, using the `SExtractor`-generated segmentation map to mask as black any pixels outside the detected source. These "stamps" were then converted pixel-for-pixel to PNG images, which employ a lossless compression algorithm — no image quality was thus lost. We then developed a transparency map based on each pixel's brightness, which was saved into the PNG alpha channel. The resulting images can thus be displayed as semi-transparent, allowing objects in the distance to show through the dim regions of objects in the foreground. This is to illustrate that the outskirts of galaxies — beyond a few effective radii — are in fact semi-transparent (see e.g. Keel & Wehrle (1993)).

Photometric redshifts for the galaxies were measured with `HyperZ` (Bolzonella et al. 2000), using a combination of the original HST-ACS four-band $(BVi'z')$ data from the HUDF, along with J- and H-band data from HST-NICMOS (Thompson et al. 2005). We have supplemented the photometric redshifts with spectrophotometric redshifts measured by Ryan et al. (2007), which incorporate the aforementioned $BVi'z'JH$ data as well as grism spectra from GRAPES (Pirzkal et al.

---

[2] The channels were first scaled as follows, proportional to the data zero points — Red: 716.474, Green: 345.462, Blue: 254.449

[3] http://www.asu.edu/clas/hst/www/jwst/clickonHUDF/

2004), U-band observations from CTIO-MOSAIC II, and K_s-band data from VLT-ISAAC. For a summary of all these data and the quality of the spectro-photometric redshifts, see Ryan et al. (2007). When available, we have chosen to use the more reliable spectro-photometric redshifts.

B.3   Development of Formulae

As previously discussed by Wright (2006), there are a number of different methods for calculating distances in cosmology. For our purposes, the most meaningful of these is the comoving radial distance, D_R, representing the spatial separation of an object and an observer with zero peculiar velocity at a common time. This distance takes into account the expansion of the Universe, and so is more useful when dealing with distances on very large scales (and thus very large look-back times), as is the case with galaxies in the HUDF. Henceforth, we shall adopt the convention of referring to the comoving radial distance from Earth to a galaxy as D_R, and the comoving coordinate distance between two arbitrary points in the coordinate system as r_{ij}.

We also wish to calculate the angular sizes of objects as they would be observed from redshifts other than zero. To do so, we need a formula for the angular size distance, D_A. That is, the distance which satisfies the equation $d = \theta D_A$ for an object with transverse diameter d subtending an angle θ in the field of view. In a simple Euclidean space, this is the same as the radial distance, but again we must take into account the expansion (and possible curvature) of the Universe, so we must use a separate equation in the AHaH tool.

Additionally, we need to consider how we wish to define the coordinate system for the HUDF objects within the Java tool. Although we have very deep HST imaging data that allow us to show how the Universe has changed over time, all of these

97

data were collected at a common time. [4]   Moreover, the principal distance measure
that we have available, D_R, also assumes a common time. Thus the most sensible
coordinate system is one with three spatial dimensions that makes all calculations
for a common time, viz. when the data were collected. We can then contract the
distances in this "comoving coordinate system" as necessary to simulate observations
from redshifts other than zero. The question remains of how we should derive such
coordinates from the data that we have in such a way that they will be useful to us
— this is discussed in §B.3.3 below, prior to deriving the equations.

B.3.1   Comoving Radial Distance

To begin, we need the comoving radial distance, D_R, from the Earth to an object
at redshift z, derived from the Robertson-Walker metric, as discussed previously by
e.g. Longair (1998, Ch. 7), Ryden (2003, Eq. 6.8), and Wright (2006, Eq. 6). We
express this as the integral:

$$D_R(z) = \int_t^{t_0} \frac{c \cdot dt}{a} = \int_{\frac{1}{1+z}}^1 \frac{c \cdot da}{a\dot{a}} = \frac{c}{H_0} \int_0^z \frac{dz}{(1+z)\dot{a}} \qquad \text{(B.1)}$$

Here the scale factor is $a = 1/(1+z)$. The derivative of a with respect to time, $\dot{a}$, is
given by the expression:

$$\dot{a} = (\Omega_M/a + \Omega_R/a^2 + \Omega_\Lambda \cdot a^2 + \Omega_K)^{1/2} \qquad \text{(B.2)}$$

Ω_M, Ω_R, and Ω_Λ are energy density parameters, corresponding to the fractions of
the Universe's average energy density that are attributable to matter, radiation, and
dark energy, respectively. Ω_K is the term arising from the curvature of the Universe's
spatial geometry, and is defined as $1 - \Omega_M - \Omega_\Lambda - \Omega_R$, thus assuming that these are
the only meaningful contributions to the total energy density.

---

[4]Sep. 2003–Jan. 2004 For a summary of the HUDF epochs, see Cohen et al. (2006)

We evaluate this integral in steps of 0.05 in z from $z = 0$ to $z = 20$ to create a look-up table, interpolating linearly to find the value for any arbitrary redshift. This is because we must make the calculation frequently and for many objects, so computing the integral manually every time would be computationally prohibitive. The resultant error in this method is generally small enough that it translates to less than one pixel's difference even on high-resolution displays, so it can safely be ignored for the purposes of the application. We evaluate the integral using the simple midpoint method, which is not the most accurate solution, but was simple to implement and adequately efficient. As with the linear interpolation, higher-accuracy numerical integration would result in less than one pixel's difference when displayed.

B.3.2   Angular Size Distance

To develop the angular size distance, D_A, we first need to develop a generalized form of D_R, to express the distance measure to an object at redshift z_j as measured by an observer at redshift z_i. This distance is given by the formula:

$$D_R(z_i, z_j) = \begin{cases} \Re_i \sin(r_i/\Re_i) & \text{if } \Omega_K < 0 \\ r_i & \text{if } \Omega_K = 0 \\ \Re_i \sinh(r_i/\Re_i) & \text{if } \Omega_K > 0 \end{cases} \tag{B.3}$$

where $\Re_i$ is the radius of curvature of the spatial geometry at redshift z_i and r_i is the value of the comoving coordinate distance at the same redshift (Longair 1998; Wright 2006). These correspond to the cases where the spatial geometry is positively curved, flat, and negatively curved, respectively. Recalling that both r_i and $\Re_i$ scale as $1/(1 + z_i)$, and that $\Re_0 = (c/H_0)/\sqrt{|\Omega_K|}$, we next define an intermediate quantity U, representing the argument of sin and sinh in Equation B.3 above:

$$U = r_i/\Re_i = r_0/\Re_0 = (H_0/c)\sqrt{|\Omega_K|}r_0 \tag{B.4}$$

We note that since U now depends only upon the cosmology selected by the user and the object's redshift, we may calculate U once per object and re-use it, thus saving CPU time. Using this quantity, we may now rewrite D_R as:

$$D_R(z_i, z_j) = \frac{\delta(U)}{1 + z_i} r_0 \tag{B.5}$$

Here, $\delta(U)$ is simply some function of U. By substituting U into Equation B.3 above, we get the following expression for $\delta(U)$:

$$\delta(U) = \begin{cases} \frac{\sin(U)}{U} & \text{if } \Omega_K < 0 \\ 1 & \text{if } \Omega_K = 0 \\ \frac{\sinh(U)}{U} & \text{if } \Omega_K > 0 \end{cases} \tag{B.6}$$

Note that $\delta(U)$ expressly depends upon r_0. The case where $\Omega_K = 0$ comes from the limit of both $\sin(U)/U$ and $\sinh(U)/U$ as $\Omega_K \to 0$ — one may observe that in this case Equation B.5 simplifies to $r_0/(1 + z_i)$, which is precisely r_i as in Equation B.3.

Thus, using our Equation B.5 and the equation relating the angular size distance and the distance measure as developed by Longair (1998, Eq. 7.50), the angular size distance from redshift z_i to z_j is given by:

$$D_A(z_i, z_j) = D_R(z_i, z_j)\frac{1 + z_i}{1 + z_j} = \frac{\delta(U)}{1 + z_j} r_0 \tag{B.7}$$

B.3.3 Comoving Coordinate System

Now that we have developed formulae for any D_R and D_A, we can consider the best way to create a coordinate system for the Java application. The data we start with are the redshift of an object (with which we can calculate D_R) and four angular measurements: the object's angular size (from the height and width of its image) and the angular separation between the object and the x and y axes, which we define as lines going through the center of the original image. These angles are calculated by

taking the corresponding size in pixels and multiplying by the scale in arcsec/pixel of the original HST image. [5]

We would like to use this information to create a coordinate system with the original telescope position at the origin. In a Euclidean space this would present no problem, but we have already remarked that the *observed* angular sizes are not the same in an expanding Universe as they would be in a Euclidean space. Further, it would be desirable for the Euclidean coordinate distance to correspond to the comoving radial distance, as this would make calculations significantly simpler. We can accomplish this, but when we create coordinates for each object as such, we need to "correct" the angles. That is, we want a "Euclidean angular size" associated with the angular size of an object at redshift z_j as viewed by an observer at redshift z_i. We will call this θ_E. Using the small-angle approximation and Equation B.7, an object's angular size is related to its physical transverse diameter, d, by the following equation:

$$d = \theta D_A = \theta \frac{\delta(U)}{1 + z_j} r_0 = \theta_E \frac{1}{1 + z_i} r_0 \tag{B.8}$$

Note that in the Euclidean case we must contract r_0 by a factor of $1/(1 + z_i)$ to get the comoving distance from z_i to z_j as measured from z_i (r_i in Equation B.3 above). This is because the spatial separation between any two points in the current epoch is stretched with respect to the spatial separation at redshift z_i, so when this separation is observed from redshift z_i, it must be scaled appropriately. Hence, the equivalent Euclidean distance between any two points is $D_E = r_0/(1 + z_i)$, in which case we get the Euclidean small-angle approximation, $d = \theta_E D_E$.

---

[5]The drizzled HUDF image scale is $0\rlap{.}''03$ per pixel

Now, canceling r_0, we get the following expression for θ_E:

$$\theta_E = \theta \delta(U) \frac{1 + z_i}{1 + z_j} \tag{B.9}$$

In our initial HUDF data z_i is simply zero, so we create coordinates (X, Y, Z) for an object at redshift z like:

$$X = \sin\left(\frac{\delta(U)\theta_X}{1 + z}\right) \cos\left(\frac{\delta(U)\theta_Y}{1 + z}\right) D_R(0, z), \tag{B.10}$$

and similarly for Y and Z. We have thus developed a coordinate system of X, Y, and Z in comoving Mpc with the original telescope position at the origin.

B.3.4 Simulating Observations From Vantage Points Other Than $z = 0$

Now, when we "move" the camera, we do so by some X_c, Y_c, and Z_c in the coordinate space. By construction, the distance measure here is just the Euclidean coordinate distance:

$$D_E = ((X - X_c)^2 + (Y - Y_c)^2 + (Z - Z_c)^2)^{1/2} \tag{B.11}$$

Now, to determine where to display an object with redshift z_o after we have moved the camera back in time to some redshift z_c $(z_c < z_o)$, we simply contract all the position coordinates by $1 + z_c$.

For an object's angular size, we know that in the Euclidean case:

$$d = \theta_0 D_R = \theta_E D_E \tag{B.12}$$

where θ_0 is the Euclidean angular size from redshift zero, and D_E is the radial coordinate distance from the camera to the object. We then solve for θ_E and substitute into Equation B.9 to obtain an expression for the desired angular size θ as observed from z_c:

$$\theta = \theta_0 \left(\frac{D_R}{D_E}\right) \frac{1 + z_o}{\delta(U)(1 + z_c)} \tag{B.13}$$

B.4 Standard Display Mode

We believe it is very instructive to consider the qualitative implications of these equations' use — that is, a description of what exactly we see when we hypothetically "move" the camera in the Java application. For the sake of completeness, we will also detail a number of cosmological effects that have been omitted from the application due to technical limitations. An example of the standard AHaH display mode is shown in Figure B.3.

When we move the camera to a certain redshift position in the HUDF data cube, we are in general viewing the Universe as it would appear from that point and at that cosmic epoch. We must qualify this statement by noting that the simulation accounts *only* for geometrical cosmological effects of changing the camera position — no dynamical, lensing, evolutionary, or other effects are simulated. In this sense, AHaH thus truly, though hypothetically, allows the user to travel through the Universe at "hyper-speed," but with the cosmic clock ticking normally, so that far-away galaxies are only seen as they appeared long ago.

The somewhat counterintuitive relationship between an object's angular size and its redshift is readily apparent in the standard display mode. If a user slowly increases the redshift of the camera, high redshift objects will begin to decrease in angular size, eventually reach a minimum angular size around $z = 1.65$ in WMAP year 1 cosmology (Spergel et al. 2007), and then increase in angular size. Also visible are the effects of galaxy evolution and merging over time. For example, when viewing the Universe from redshift $z = 0.5$ as in Figure B.3, there are many large spiral and elliptical galaxies visible. However, when viewing the Universe from redshift $z = 1.5$ as in Figure B.4, the Universe is dominated by small and compact blue galaxies. This is a well-known phenomenon established by HST, that the population of elliptical and

spiral galaxies seen today transitions into the population of actively star-forming and merging galaxies seen at $z \geq 1 - 2$ (Driver et al. 1995, 1998; Abraham et al. 1996; Glazebrook et al. 1995).

It should be noted that the application does not make calculations for cosmological surface brightness dimming or changes in color due to redshift or spectral evolution. While certainly feasible to simulate, performing such image manipulation techniques on large numbers of galaxies in real-time is currently too difficult for consumer computers. Moreover, we must also recall that the HUDF data are limited in both apparent magnitude and effective horizon by what could be observed from low Earth orbit. When we view the data from redshifts other than zero, we would expect to see more galaxies overall — including fainter galaxies — than are represented in the current HUDF data. We could choose to simulate these objects as extensions of our data set if we desired, but we felt this would not be particularly instructive, and could lead to potential confusion. Moreover, such simulations have a high degree of uncertainty and, by significantly increasing the size of the data set, would add prohibitively to the computation times. Likewise, we have chosen not to simulate galaxies outside of the original field, which would of course enter the camera's field of view as the user pans around.

B.5   Static Geometry Mode

When a user presses the "G" key in the Java tool, they are told that they are viewing the simulation with "static geometry" turned on. What this means is that angular sizes as derived above are no longer affected by the scale factor or curvature of the Universe — after we develop our original coordinates, as in Equation B.10, all calculations for angles are simply done with $\theta = \theta_E$. This has the visual effect of all galaxies appearing smaller, since all initial angles have been contracted by a factor of $1 + z$ (when Ω_K is zero). In this static case, galaxies will also simply increase in angular size as we approach them, as opposed to the angular sizes of objects in the real Universe, which decrease, reach a minimum around $z = 1.65$, and then increase.

This static mode of viewing the simulation has no physical analog — it is simply meant to convey to the user that there are non-Euclidean aspects of the Universe's geometry, and that the angular sizes we observe in the present have been made larger due to the Universe's expansion. One should note that this display mode only considers expansion as it relates to angular size — the comoving radial distance is still calculated using the redshift and curvature factors that would not be present in a strictly Euclidean Universe. That is, in the static display mode, we assume that the Hubble Law distance, $D = v/H_0 = (c/H_0)z$, is simply a Euclidean distance unrelated to expansion. This is primarily because our method of calculating the comoving radial distance relies upon redshift, which is a phenomenon specific only to an expanding Universe, and is therefore the only way we could calculate distances for all the galaxies, even in the Euclidean case.

B.6 Conclusion

We believe that this software provides students and instructors with an unprecedented ability to interactively visualize many of the effects of an expanding Universe, among its other capabilities. The application should help clarify these concepts, and allow students to develop a deeper intuitive understanding of the material. Certain cosmological effects — such as bandpass shifting, k-correction, surface brightness dimming, evolutionary effects, gravitational lensing, and the effects of the magnitude limit and object sizes on the sample completeness limit — have been omitted due to computational or dataset limitations, but we believe these to be inessential for the understanding of the included effects. For a discussion of most of these effects, see e.g. Windhorst et al. (2008).

For the convenience of those who wish to see or modify the particular implementation of the above formulae within the Java software, we have provided source
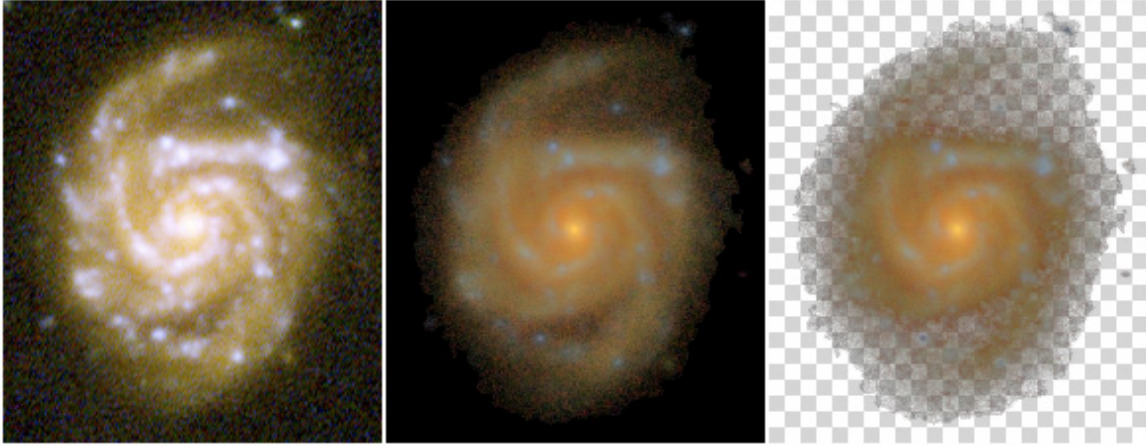
Figure B.1: A comparison of three images of HUDF galaxy 7556. The left image is that from the original STScI release, clearly showing the bright, burnt-out knots characteristic of the standard logarithmic image stretch. The center image is our prepared image using the arcsinh stretch described by Lupton et al. (2004), as it appears in the AHaH application. The right image is our prepared image against an artificially imposed chessboard pattern, showing the included transparency. Note that pixels outside the source are all completely transparent, since they have been removed entirely using the `SExtractor` segmentation map. The outskirts of the galaxy — beyond a few effective radii — are semi-transparent, as in real galaxies.

code with the standard distribution of the tool. It is included in the src/ directory of ahah.jar and may be extracted using the java `jar` utility or any zlib-compatible de-compressor such as `unzip`. The release version of the tool may be downloaded for Windows, Macintosh, or Linux computers from the AHaH website. [6]
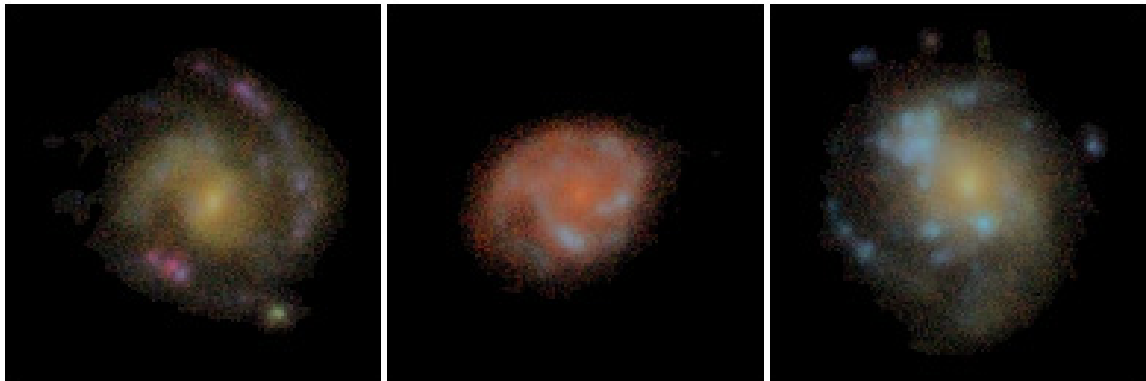
---

[6]http://www.asu.edu/clas/hst/www/ahah/

Figure B.2: Our prepared images of three galaxies from the HUDF, using the arcsinh stretch described by Lupton et al. (2004). Shown are galaxy 3180 (left), galaxy 5805 (center), and galaxy 6974 (right).

Figure B.3: The HUDF data as viewed from redshift $z = 0.5$ in the AHaH application, using standard geometry mode, which properly calculates angular sizes. Note how the image is dominated by luminous red early-type galaxies.

Figure B.4: The HUDF data as viewed from redshift $z = 1.5$ in the AHaH application, using standard geometry mode. Note how this image is dominated by blue irregular and merging star-forming galaxies.