IMPROVED CONSTRAINTS ON THE GRAVITATIONAL LENS Q0957+561. I. WEAK LENSING

R. NAKAJIMA\textsuperscript{1,2}, G. M. BERNSTEIN\textsuperscript{1}, R. FADELY\textsuperscript{3}, C. R. KEETON\textsuperscript{3}, and T. SCHRABBACK\textsuperscript{4,5}

\textsuperscript{1}Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
\textsuperscript{2}Space Sciences Laboratory, University of California, Berkeley, CA 94720, USA
\textsuperscript{3}Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854, USA
\textsuperscript{4}Argelander-Institut für Astronomie, Universität Bonn, D-53121 Bonn, Germany
\textsuperscript{5}Leiden Observatory, Universiteit Leiden, NL-2333 CA Leiden, Netherlands

Received 2008 December 18; accepted 2009 March 25; published 2009 May 14

Abstract

Attempts to constrain the Hubble constant using the strong gravitational lens system Q0957+561 are limited by systematic uncertainties in the mass model, since the time delay is known very precisely. One important systematic effect is the mass-sheet degeneracy, which arises because strong lens modeling cannot constrain the presence or absence of a uniform mass sheet $\kappa$, which rescales $H_0$ by the factor $(1 - \kappa)$. In this paper, we present new constraints on the mass sheet derived from a weak-lensing analysis of the Hubble Space Telescope imaging of a 6 arcmin square region surrounding the lensed quasar. The average mass sheet within a circular aperture (the strong lens model region) is constrained by integrating the tangential weak gravitational shear over the surrounding area. We find the average convergence within a 30 arcsec radius around the lens galaxy to be $\kappa(<30') = 0.166 \pm 0.056$ (1$\sigma$ confidence level), normalized to the quasar redshift. This includes contributions from both the lens galaxy and the surrounding cluster. We also constrain a few other low-order terms in the lens potential by applying a multipole aperture mass formalism to the gravitational shear in an annulus around the strong-lensing region. Implications for strong lens models and the Hubble constant are discussed in an accompanying paper.

Key words: galaxies: clusters: individual (Q0957+561) – galaxies: halos – gravitational lensing

Online-only material: color figures

1. INTRODUCTION

Gravitational lensing allows for a measurement of the Hubble constant $H_0$ that is independent of the “distance ladder” and is not susceptible to the peculiar velocities of the local universe (Refsdal 1964). This unique opportunity is available for special lenses strong enough to generate multiple images from a single source, and when this source has intrinsic variability, such that a differential time delay can be observed between the multiple images.

The doubly imaged quasar Q0957+561 is the first confirmed example of strong gravitational lensing (SL; Walsh et al. 1979). The quasar is variable; with the time delay accurately measured $<1\%$ between the two images (Kundic et al. 1997; Colley et al. 2003; Shalyapin et al. 2008), the lack of precision in obtaining $H_0$ from this system in past measurements ($\pm 35\%$, 2$\sigma$) lies in determining the mass distribution of the lens (Bernstein & Fischer 1999; Keeton et al. 2000).

One well-known problem in uniquely determining the mass distribution (and hence $H_0$) is the “mass-sheet degeneracy” (Falco et al. 1985; Gorenstein et al. 1988). The lens mass is modeled given the multiple images as constraints; the lensed light is traced back from the multiple image according to a given lens model, and checked for a consistent, single source. Unfortunately, the true angular position of the source is not observable and hence not constrained; the mass-sheet degeneracy occurs because the angular position of the source object is degenerate with the constant surface mass density, or the “mass sheet” (Bernstein & Fischer 1999). Thus, the mass sheet manifests itself as a uniform magnification of the (unobservable) source plane.

This degeneracy in the SL model can be lifted if the average mass overdensity in the SL model region is known. This is done by using the aperture mass weak gravitational-lensing (WL) technique (Kaiser & Squires 1993). WL generates a subtle but coherent distortion of the background galaxies, which can be used to infer the mass distribution. In particular, the integrated tangential shear around a circular aperture nonparametrically determines the two-dimensional average mass overdensity within the aperture.

The WL measurement suffers from its own mass-sheet degeneracy, because the WL signal must formally be integrated over the surrounding region out to infinity to obtain the mass distribution. While this is not practical, a realistic scheme to estimate the average mass sheet in an SL region is to integrate the WL distortions over an annular region surrounding the SL model. WL then measures the differential mass sheet $\Delta \kappa$ of the disk region defined by the inner and outer circles of the annulus, where the average convergence within the large circle is small, and can be estimated from an assumed mass profile. The mass sheet for Q0957+561 has previously been measured in this way (Fischer et al. 1997).

There are other sources of degeneracy in obtaining $H_0$ from SL. The Q0957+561 lens is a galaxy which sits within a modest cluster, so the lens mass model must include both sharply and smoothly varying terms, corresponding to that of the galaxy and cluster, respectively. The cluster potential can be expanded in Taylor series; the second-order terms (the lowest order contributing to gravitational lensing) correspond to the mass sheet and constant shear within the SL region. Past studies of the lens model have shown that the uncertainty in $H_0$ measured from Q0957+561 is dominated by the degeneracy between the galaxy ellipticity and cluster shear (Keeton et al. 2000, and references therein). For many models, the parametric fits tend to converge to a large cluster shear ($\gamma_c \sim 0.1–0.3$), while the quasar host galaxy lensed images seem to imply a small shear
be constrained by (and hence not included in) the SL models. Specifically, for a lens potential \( \psi(\theta) \) and source position \( \beta \) that satisfies the SL constraints,

\[
\psi'(\theta) = \frac{1}{2} \kappa_0 |\theta|^2 + (1 - \kappa_0) \psi(\theta),
\]

also satisfies the constraints, where \( \kappa_0 \) is the constant convergence (see Equation (6)) from the additional uniform mass sheet. Under this degenerate transformation, the time delay is

\[
(\Delta t)' = (1 - \kappa_0) \Delta t,
\]

requiring the Hubble parameter to rescale by \( H_0' = (1 - \kappa_0)H_0 \).

The lensing convergence \( \kappa(\theta) \) is a normalized surface density, and is defined as

\[
\nabla^2 \psi(\theta) = 2\kappa(\theta) = \frac{\Sigma(\theta)}{\Sigma_{crit}},
\]

where \( \Sigma(\theta) \) is the surface mass density, and the critical surface density is determined from the lens and source redshifts, and the assumed cosmology:

\[
\Sigma_{crit} = \frac{c^2}{4\pi G} \frac{D_{OS}}{D_{OL}D_{LS}}.
\]

An SL region which can generate multiple images has \( \Sigma \gtrsim \Sigma_{crit} \), or \( \kappa \gtrsim 1 \), while WL techniques are generally valid only in a region where \( \kappa \ll 1 \).

To remove the \( (1 - \kappa_0) \) degeneracy in \( H_0 \), one must obtain the uniform mass sheet \( \kappa_0 \) not included in the SL model. The overall average mass sheet \( \langle \kappa \rangle \) can be constrained via WL, which constrains the sum of \( \kappa_{SL} \) and \( \kappa_0 \), where \( \kappa_{SL} \) is the average convergence of the SL model lens mass distribution. The WL information is obtained over a field wider than the area where \( \psi(\theta) \) is modeled, and its procedure is described in Section 4.6.

2.3. Cluster Model

The Q0957+561 lens system consists of the primary lens galaxy G1 at redshift \( z_{lens} = 0.355 \), the relatively weak cluster which contains G1, and the quasar itself which lies at \( z_{src} = 1.41 \). The lens model in the SL analysis consists of a mass concentration with near elliptical symmetry representing the galaxy, with additional cluster potential components that are Taylor expanded in position to the third order (Bernstein & Nakajima 2009; Fadely et al. 2009). The expansion is defined within \( r \lesssim R_{SL} \), centered at the central peak of G1, and is fully general to order \( (r/R_{SL})^3 \), where \( R_{SL} \) is the radius of the circular boundary where the SL modeling takes place:

\[
\psi(r) = (1 - \kappa_c) \left[ \psi_c(r) + \text{Re} \left( \frac{\gamma_c}{2} r^2 e^{-2i\phi} + \frac{\sigma_c}{4} r^3 e^{-i\phi} \right) + \frac{\delta_c}{6} r^3 e^{-3i\phi} \right] + \frac{k_c}{2} r^{2.5} \quad (r \lesssim R_{SL}),
\]

where \( r \equiv (x, y) \equiv (r, \phi) \) is the transverse distance in the lens plane (and is equivalent to \( \theta \) in the previous sections), \( \psi_c(r) \) is the galaxy potential, and \( \kappa_c, \gamma_c, \sigma_c, \text{ and } \delta_c \) are the (complex) constants corresponding to the cluster mass sheet, constant shear, internal mass dipole moment, and \( m = 3 \) shear
moment, respectively. The complex notation corresponds to the two-dimensional description of the gravitational potential such that, e.g., the derivative is

$$\partial \equiv \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} = e^{i\phi} \left( \frac{\partial}{\partial r} + \frac{i}{r} \frac{\partial}{\partial \phi} \right).$$

The potential and the mass sheet $\kappa_c$ are always real. The first four terms in the square bracket are those modeled in SL, and are multiplied by $(1 - \kappa_c)$ due to the mass-sheet degeneracy associated with the analysis; $\kappa_c$ is the mass sheet that is not part of the SL model, and cannot be constrained from SL model within $r < R_{\text{SL}}$ alone, due to degeneracy (Section 2.2). The four quantities $\kappa_c$, $\gamma_c$, $\sigma_c$, and $\delta_c$ are constrained by WL measurements from data in $R_{\text{SL}} < r < R_{\text{max}}$ and the multipole formalism (Section 4.6). The inner and outer annular radii were chosen to be $R_{\text{SL}} = 30''$ and $R_{\text{max}} = 186''$ for our analysis, such that we consider a region sufficiently removed from the SL region, but with reasonable WL signal from the cluster.

3. OBSERVATIONS AND DATA

3.1. HST/ACS Imaging

The images for the cluster around the double quasar Q0957+561 were taken as a $2 \times 2$ mosaic of HST ACS pointings of the Wide Field Channel (WFC) detector on 2005 October 10–13 (Figure 1). Each ACS WFC pointing has $202'' \times 202''$ field of view (FOV), resulting in a $6' \times 6'$ mosaic image. Each pointing has four orbits ($4 \times 1920$ s) in the F606W (Broad V) filter and two orbits ($2 \times 1880$ s) in the F814W (Broad I) filter; the mosaic has the central 30'' region overlapped such that the region around the quasar images have four times the depth (30 ks and 15 ks for F606W and F814W, respectively) for studies of SL features (Fadely et al. 2009).

Figure 1. 6' × 6' combined image of the 0957+561 field in ACS F606W. Each quadrant is a single pointing (with pixel scale dithering) which has been imaged to a depth of 7.5 ks, while all four quadrants overlap in the central 30'' region to yield a depth of 30 ks. The chip gap is apparent in each of the quadrants. The orientation and the image scale are as labeled. The apparent shear in the FOV is from correcting for the image distortion due to the ACS focal plane located off-center from the telescope axis.

Figure 2. Size–magnitude diagram of all detected objects from native GLFit. The boxed region shows a portion of the stellar locus with no saturation and no galaxy contamination. The stellar size $\sigma$ is approximately the size of the ACS WFC pixel, $0.054''$, where $\sigma$ is the characteristic width scale for the best-fit Gaussian (for objects with elongated shapes, $\sigma$ is the geometric mean of the two characteristic widths). The crosses are all detected objects, while the circles indicate objects which satisfy the size, magnitude, flags, and signal to noise (S/N) >25 cut for use in WL analysis. The size and magnitude cuts are indicated by a dotted horizontal and a dotted vertical line, respectively.

(A color version of this figure is available in the online journal.)

The HST WFC is a mosaic of two 4096 × 2048 pixel charge-coupled devices (CCDs) of approximately $0.05$ pixel$^{-1}$. Each orbit is split into four dithered exposures to sample the diffraction-limited point-spread function (PSF) and to remove defective pixels and cosmic rays in the combined image.

3.2. Object Catalog

We created an initial stacked mosaic image using MultiDrizzle (Koekemoer et al. 2002), where we refine relative shifts between individual exposures by cross-correlating the positions of compact sources in overlapping regions. Objects were detected in this image using SExtractor (Bertin & Arnouts 1996). We took the initial SExtractor detection parameters and refined the centroid, size, and shape of each object using a native GLFit (Section 4.1; Nakajima & Bernstein 2007). The refined parameters determined here are still those which are convolved with the PSF, therefore the size–magnitude diagram allows us to determine the stellar locus (Figure 2). The stellar objects have an average full width half-maximum (FWHM) of $0.13$ in F606W, or a corresponding $\sigma = 0.054''$, where $\sigma$ is the characteristic width of the best-fit Gaussian.

Once the stars and galaxies were separated, the true (pre-PSF) galaxy shapes were obtained by performing a deconvolution GLFit, based on an interpolated PSF (Section 4.2). Both the native and deconvolution GLFit are done as a “multifit”—an individual object is fitted simultaneously over pixel data from each exposure rather than from the combined image—hence the measured shapes are not affected by distortion or aliasing induced by the image-combining process (Section 4.3).

4. WEAK-LENSING ANALYSIS

4.1. EGL Method

For our WL analysis, we employ the Elliptical Gauss Laguerre (EGL) method of shape measurement and shear estimation (Bernstein & Jarvis 2002; Nakajima & Bernstein 2007), implemented as GLFit; we provide here a brief description. In
characterizing their shapes, star and galaxy images are decomposed into two-dimensional Gauss–Laguerre (GL) basis functions (Bernstein & Jarvis 2002). In the EGL method, the GL basis functions are described upon a “basis ellipse,” i.e., a sheared and stretched coordinate system such that a unit circle appears as an ellipse of chosen size, shape or orientation; and so that the lowest order GL function is an elliptical two-dimensional Gaussian. The combination of the basis ellipse and coefficients to the GL functions then fully describes the image.

The choice of basis ellipse varies depending on its purpose. In describing the PSF, we choose a common basis ellipse: a circle with radius of the average PSF size. The PSF variation across the FOV can then be traced through the variation in the GL coefficients, and allows for a simple description of the necessary interpolation.

In contrast, when measuring galaxy shapes, the basis ellipse is chosen such that in the simplified case of no PSF smearing) the lowest order GL function is the best-matched two-dimensional Gaussian ellipse to the galaxy image. This basis ellipse then describes the shape (ellipticity) of the galaxy. In the observed image, the true galaxy shape has been convolved with the PSF. So our fitting does a simple deconvolution by modeling the true galaxy shape with a GL expansion, and adjusting the coefficients so that convolution of this model with the PSF best matches the data (Nakajima & Bernstein 2007).

4.2. PSF Interpolation

In order to obtain the true galaxy shapes, the PSF must be known at the position of the galaxy image. This is done by selecting the stars in the image and interpolating the shape at the galaxy position. There are two problems that make this difficult for our images. First, the PSF spatial variation is not stable over time: HST is known to “breathe” going in and out of the Earth’s shadow, causing its focus to vary on timescales of a single orbit (≈1.5 hr). Hence, it is unlikely that a single PSF spatial variation model would be valid for all exposures (Schrabback et al. 2007, and references therein). Second, there are not enough stars (≈20 per chip) in our cluster image to create a reliable model of the PSF variation across every exposure.

To circumvent the aforementioned difficulties, we use publicly available ACS WFC stellar field (SF) images to create the models of the PSF variation for every exposure. A given PSF spatial pattern is highly reproducible, since the spatial variation of PSF patterns in different exposures is caused by the thermal breathing of the HST focus. If the few PSF available on a given exposure can constrain a portion of the pattern, the rest of the PSF pattern can be predicted (Rhodes et al. 2005; Jee et al. 2007).

The noise in the interpolation is minimized by utilizing principal component analysis (PCA). Our procedure is as follows.

1. Obtain the SF images available from the HST archive. There were 184 archival F606W exposures of dense SFs; the criteria for selection are described in Schrabback et al. (2007).

2. The PSF anisotropy is measured from each star in terms of the anisotropy kernel $q$ (Schrabback et al. 2007), and its variation across the FOV is characterized by a third-order polynomial fit.

3. We then choose, for each of our Q0957 exposure, the SF image that best matches the anisotropy pattern of the stars.

The “best-fit” SF is identified by that which yields the minimal

$$\chi^2_{SF} = \sum_{i=1}^{N_{stars}} \left[ q_i - q_{model}(x_i, y_i) \right]^2$$

for a given exposure, where $q_i$ is the kernel for the $i$th star in our exposure (which have very few stars) and $q_{model}(x_i, y_i)$ is the interpolated kernel at the position of the $i$th star for the matching SF.

4. For the PSF interpolation, obtain the PSF GL coefficients by (1) dividing each SF exposure into a $8 \times 8$ grid, and then (2) fitting all stars within the grid to a constant PSF model for that cell, resulting in 64 PSFs per exposure (32 PSFs per chip). We choose to average the PSF in this way because the ACS WFC undersamples the PSF; by using multiple PSF images within the grid, the PSF is effectively “dithered” to obtain better sampling for the fit. We then have, for a given SF $f$ and within every grid cell $g$, a GL coefficient vector $b_{fg} = \{b_{fg}^i\}$ describing the PSF, where $i$ runs over the GL indices. All PSF coefficients from $f, g$ are described over a common basis ellipse, which we take to be a circle with radius equal to the average size of the PSF.

5. Extract the principal components $b^e = \{b^e_i\}$ of the collection of PSF GL coefficients $\{b_{fg}^i\}$ (Jarvis & Jain 2004). Express the PSF in terms of the PCA vectors, $b_{fg} = \sum \alpha \cdot b^e_{fg}$, where $\alpha$ runs over the PCA vectors. The PSF is now described in terms of PCA coefficients $\beta = \{\beta^e_i\}$ instead of the GL coefficients $b = \{b_i\}$.

6. Truncate the PCA coefficients by using only the major principal components which correspond to the PSF (spatial and temporal) variation, but not to the noise. These components are identified by ranking the variance of each, and identifying a gap in the variance (see Figure 3). Most of the variation in PSF is described in the first several terms in $\alpha$.

7. Obtain a third-order, two-dimensional polynomial fit across the image for the relevant PCA components $\alpha$ for each field $f$ such that the interpolated PCA coefficients are described as

$$\hat{\beta}^e_{a} = \sum_{0 \leq i,n \leq 3} c_{a,nn}^{f} x_{g}^{n} y_{g}^{n},$$

where $\chi^2 = \sum_{g} (\beta^e_{a} - \hat{\beta}^e_{a(g)})^2$ is minimized, $c_{a,nn}^{f}$ is the polynomial coefficient for $\alpha$ in field $f$, and $(x_g, y_g)$ are the grid center coordinates. Each PCA description of the PSF on every grid ($\hat{\beta}^e_{a}$) was visually inspected; obvious failures in the PSF descriptions (due to the lack of sufficient stars within the grid) were eliminated before being used in this fit.

8. The PSF at any location $(x, y)$ over any field $f$ can then be modeled using the known principal components and its two-dimensional polynomial fit coefficients across the given exposure.

Figure 3 shows the variance of each of the PSF principal components, for each WFC ACS CCD. The zeroth component is the average PSF shape over all grid cells and exposures. We find that the first seven components constitute the primary variance and are sufficient to describe the PSF variation; from visual inspection, the residual is consistent with ≈1% noise with respect to the peak. These seven components are used to model the PSF across the exposure.
4.3. Galaxy Multifit

Once the PSF is known at a given galaxy location, we perform the deconvolution GLFit to obtain the true galaxy shape. We utilize the multiple exposures taken of each galaxy without combining them. The “multifit” technique performs simultaneous fitting over individual exposures, each with distinct PSFs, assuming a single “true” galaxy model, and the fitting done with the convolved GL basis functions of each exposure to their respective PSFs. The simultaneous fitting is iterated over the basis ellipses to obtain the best-fitting true galaxy shape.

The pixel image is distorted with respect to the true image, so distortion correction is necessary within the multifit procedure. The galaxy model, described in sky coordinates, is fitted to the pixel flux information via a pixel-to-sky coordinate map, which corrects for optical distortions in the pixel image. The pixel map is based on a known, stable solution of fourth-order polynomial in combination with a supplementary look-up table (Anderson 2002, 2006), to which a linear correction (rotation, translation, shear, and uniform scale) is applied to fit to the USNO-B catalog (Monet et al. 2003) for alignment and absolute astrometry. Bad pixels, such as those affected by cosmic rays or saturation, are not used in the fitting procedure. The bad pixel masks for each exposure were generated using the multidrizzle package (Koekemoer et al. 2002).

4.4. Galaxy Selection for Weak Lensing

Once the object catalog has been generated, we select the galaxies for WL analysis by imposing the following cuts:

1. **Size.** The stellar contamination is removed by choosing objects which have characteristic size $\sigma$ which are $\gtrsim 1.5$ times that of the PSF. Our cut is conservative to ensure PSF deconvolution to be in the range where it performs well (Nakajima & Bernstein 2007).

2. **Significance (S/N).** From our performance analysis of GLFit (Nakajima & Bernstein 2007), we know that the deconvolution starts to produce biases below S/N of 20–40, depending on the galaxy size with respect to the PSF. Here, we choose the cut to be S/N $> 25$.

3. **Magnitude.** Since we only want the background galaxies to the cluster for the lensing analysis, foreground cluster member contamination must be minimal. Lacking redshift information, and unable to utilize color information for this purpose for reasons listed below, we cut the brightest objects ($m_{F606W} < 24$) as foreground. By plotting the galaxy number density as a function of the radius from the central brightest galaxy of the cluster (Figure 4), we estimate that there is at most 10% cluster member contamination for $r < 75''$ from the cluster center for the $m_{F606W} > 24$ objects.

4. **Flags.** Any object whose GLFit flags indicate that the shape has not been measured (an unsuccessful fit to the basis ellipse) is rejected from the catalog.

After these cuts, 1866 galaxies remain for our WL analysis, or galaxy number density of 50 arcmin$^{-2}$ (Figure 4).

We have not included a color cut (for a rough exclusion of cluster members) for the following reason: although several of the red-sequence cluster members of known redshifts have the predicted $F606W–F814W$ color of 1.0 (Bruzual & Charlot 1993) in our color–magnitude plot, no peak was found in the galaxy count at this color. This is because (1) majority of the cluster members are blue and (2) the expected member galaxy count in this halo is small. The first point is verified from > 50% of cluster members of known redshifts which are bluer than the red sequence, with $F606W–F814W$ colors in the range 0.3–0.8.

![Figure 3](image-url) (A color version of this figure is available in the online journal.)

![Figure 4](image-url) (A color version of this figure is available in the online journal.)
Many of these galaxies were found to have spiral morphology (Angonin-Willaime et al. 1994). The Q0957 cluster mass (and hence the number of its luminous members) is expected to be low based on the cluster richness (Johnston et al. 2007). The richness $N_{200}$ is defined by the number of member galaxies which are consistent with the red-sequence color, is within a certain radius of the brightest cluster galaxy, and have luminosity above $0.4L_*$ (see Koester et al. 2007, for details). Although their definition for $N_{200}$ is given only to $z = 0.3$, we extend their red-sequence color cut to $z = 0.35$ based on Table 1 of Eisenstein et al. (2001), and determine the equivalent richness to be $N_{200} = 2$ from their Sloan Digital Sky Survey (SDSS) magnitudes (Adelman-McCarthy et al. 2008). This is lower than the lowest richness bin available in Johnston et al. (2007), an indication that the Q0957 cluster is more of a group than a cluster, and hence that the member count is low.

4.5. CTE Correction

The galaxy shape is affected by the degradation of the charge transfer efficiency (CTE) in the ACS WFC CCD. When a charge is read out from a CCD pixel, some of the charge is trapped and released at a later time, causing the galaxy image to elongate along the charge readout direction, adding an excess shear $\gamma_{\text{CTE}} \propto <0$ along the pixel axis, where $\gamma_+ > 0$ or $<0$ indicates elongation in the horizontal or vertical direction, respectively (see Section 4.6.2 for the definition of the orthogonal shears $\gamma_+$ and $\gamma_-$). The effect of charge trailing on the object shape is worse for objects that are smaller, have low flux, require larger number of transfers to the readout register (i.e., objects located close to the chip gap in the ACS WFC), and exposures that have been taken at a later date.

We correct for CTE effects using a parametric model given the flux, size, location on CCD and observation date to estimate the spurious elongation, and correct for it by removing this shear from each object in the shape catalog. The model was derived from galaxies in the HST/COSMOS Survey similarly to that used by Rhodes et al. (2007), but additionally takes sky background variations into account (see T. Schrabback et al. 2009, in preparation for further details).

The mean correction averaged over all galaxies in our sample is $\langle \gamma_{\text{CTE}} \rangle = -0.008 \pm 0.010$, with the worst case correction for small, dim objects near the chip gap being up to $\gamma_{\text{CTE}} = -0.04$. Given the $2 \times 2$ mosaic configuration, we expect the CTE effect on the measured shear multipole moments to vanish to first order for the even multipoles. Overall, we find the correction to the mass distribution to be negligible, of the order $<3\%$ (well within the error bar, see Section 5) for the mass sheet, and within $<10\%$ of 1σ error bar for the multipole moments.

4.6. Mass Sheet and Multipole Moments

The two-dimensional multipole mass distribution within an aperture can be obtained in a nonparametric fashion from the WL shear information surrounding the aperture. The overall mass sheet (i.e., the combined average mass sheet from the SL model and the unconstrained cluster term $\kappa_c$) corresponds to the monopole. Here, we summarize the results from Bernstein & Nakajima (2009).

4.6.1. Mass Sheet

For any mass distribution, the expected azimuthally averaged tangential shear at radius $r$ from the mass center is (Miralda-Escudé 1991)

$$\bar{\gamma}_t(r) = \frac{\bar{\Sigma}(r) - \bar{\Sigma}(0)}{\Sigma_{\text{crit}}} \equiv \frac{\Delta \Sigma(r)}{\Sigma_{\text{crit}}} = \bar{\kappa}(r) - \bar{\kappa}(0),$$

where $r$ is the angular distance from the cluster center, $\bar{\gamma}_t(r)$ is the tangential shear averaged at $r$, $\bar{\Sigma}(r)$ is the azimuthally averaged surface density at radius $r$, $\Sigma_{\text{crit}}$ is the critical surface density, and $\Delta \Sigma(r) \equiv \Sigma_{\text{crit}} - \bar{\Sigma}(r)$ is the value often quoted in the WL literature. The ratio of the surface density to $\Sigma_{\text{crit}}$ is the convergence $\kappa$. We can directly estimate the mass sheet within an aperture using the aperture mass method (Fahlman et al. 1994): the average convergence within a circular aperture of radius $R$ is

$$\bar{\kappa}(r) \equiv \frac{1}{\pi R^2} \int_{r < R} d^2r \kappa(r) = \frac{1}{\pi} \int_{r < R} d^2r \frac{\bar{\gamma}_t(r)}{r^2},$$

where the tangential shear is integrated over $R < r < \infty$ (Kaiser & Squires 1993). The average convergence gives the average surface density if the critical surface density $\Sigma_{\text{crit}}$ is known. We discuss $\Sigma_{\text{crit}}$ estimation in Section 4.7.

The practical estimator for Equation (10) is a summation over galaxy shapes

$$\bar{\kappa}(r) \equiv \frac{1}{nR} \sum_{i(r > R)} \frac{\gamma_{t,i}}{r^2},$$

where $n$ is the surface number density of source galaxies and $\gamma_{t,i}$ is the tangential component of the $i$th galaxy shape. Here, we have assumed that the locally averaged galaxy shape (tangential component) is an estimator of the shear, $\gamma_t \approx \langle \gamma_{t,i} \rangle / R$, where $R$ is the responsivity, the multiplicative correction factor for the estimation of shear from galaxy shapes (Bernstein & Jarvis 2002).

In practice, the summation over all galaxies at $r > R$ is truncated at a maximum radius $R_{\text{max}}$. Equation (11) then becomes

$$\bar{\kappa}(r) - \bar{\kappa}(R_{\text{max}}) = \frac{1}{nR} \sum_{i(R < r < R_{\text{max})}} \frac{\gamma_{t,i}}{r^2}.$$
The truncation at $R_{\text{max}}$ is a source of uncertainty; that is, only the difference in the mass sheet at different aperture radii can now be determined (this is the WL mass-sheet degeneracy). We then assume an appropriate model to derive the mass distribution within $R_{\text{max}}$, and hence $\kappa(<R_{\text{max}})$. We feel safe in doing so, since this correction is reasonably smaller than $\kappa(<R)$.

If the WL measurement yields an average convergence of $\kappa$ within $r = R$ (normalized to the quasar redshift) and the SL modeling (assuming $\kappa_c = 0$ in Equation (8)) yields an average convergence of $\kappa_{\text{SL}}$, then the relation of the degenerate mass sheet $\kappa_c$ to the measured mass sheet $\bar{\kappa}$ is $\bar{\kappa} = \kappa_c + (1 - \kappa_c)\kappa_{\text{SL}}$, or

$$1 - \kappa_c = \frac{\bar{\kappa}}{1 - \kappa_{\text{SL}}}.$$  \hfill (13)

### 4.6.2. Multipole Moments

The method of aperture mass (Fahlman et al. 1994; Schneider 1996; Schneider & Bartelmann 1997), which integrates tangential shear to obtain the mass sheet within an aperture, can be generalized to obtain the “interior” and “exterior” multipoles defined relative to a circle of radius $R$ (Bernstein & Nakajima 2009). The “interior monopole” corresponds to the mass sheet, and is constrained from the shear information exterior to $r = R$. The other multipoles are constrained in a similar manner, where the shear information exterior to $r = R$ constrains the mass sheet moment interior to $r = R$, and vice versa.

In order to utilize the shear information to extract mass multipoles, first we define the complex shear relative to the circle, and is not orthogonal to which elongates in the radial direction is simply a negative tangential shear, and is not orthogonal to $\gamma_t$. The shear orthogonal to the tangential/radial directions are aligned $\pm 45^\circ$ from tangential/radial; hence we follow past conventions and rename this the “skew” shear, $\gamma_s$.

The interior and exterior mass multipole moments relative to a circle of radius $R$ are defined as

$$Q_n^{(m)}(r) \equiv \int_{r < R} d^2 r r^m e^{-im\phi} \kappa(r),$$  \hfill (15)

$$Q_n^{(m)}(r) \equiv \int_{r > R} d^2 r r^m e^{im\phi} \kappa(r),$$  \hfill (16)

where these definitions are normalized to agree with Equations (B5) of Schneider & Bartelmann (1997), but with an alteration in the phase convention. If the cluster potential at $r < R$ is described as Equation (8), then, assuming no contribution from the galaxy potential $\psi_g$, the interior and exterior multipoles $Q_n^{(m)}$ are related to the cluster constants as

$$Q_n^{(0)}(R) = \pi R^2 \kappa_c,$$  \hfill (17)

$$Q_n^{(2)}(R) = \frac{\pi}{4} R^2 \kappa_c^2 (1 - \kappa_c),$$  \hfill (18)

$$Q_n^{(3)}(R) = -\pi \kappa_c (1 - \kappa_c),$$  \hfill (19)

$$Q_n^{(3)}(R) = -\frac{\pi}{2} \delta_c (1 - \kappa_c),$$  \hfill (20)

independent of the cluster mass distribution at $r > R$, where $\kappa$ and $\sigma$ are the monopole and dipole mass and $\gamma$ and $\delta$ are the constant and $m$ = 3 shear terms within $r = R$, respectively, and $\sigma^*$ indicates complex conjugation. The terms $Q_n^{(2)}$ and $Q_n^{(3)}$ vanish (i.e., there are no mass distribution within $R$ of these multipoles), since $\kappa \equiv \frac{1}{2} \nabla^2 \psi = 0$ if $\psi \propto r^m e^{im\phi}$, and the $Q_n^{(3)}$ term produces no shear internal to $r < R$. These terms have no effect upon the lens model or the time delay, and therefore can be ignored. The $m = -2$ and $m = -3$ shear patterns within $R$ (corresponding to $\gamma_s$ and $\delta_c$ terms in Equation (8)) are generated by a quadrupole and sextupole mass distribution, respectively, external to $R$ (see Figure 1 of Bernstein & Nakajima 2009), hence these terms correspond to the external mass distributions $Q_n^{(m)}(r)$ or $Q_n^{(m)}(\infty)$ for the galaxy potential enter each of the multipole terms, and the WL measurement constrains the sum of the SL galaxy model and cluster terms.

The multipoles of the convergence $\kappa$ are related to the integrals over the complex shear $\Gamma$, Equation (14), as

$$\int_{r > R} d^2 r \Gamma(r) e^{-m^2 \phi} = R^{-2m^2} Q_n^{(m)}(R) (m \geq 0),$$  \hfill (21)

$$\int_{r < R} d^2 r \Gamma(r) e^{m^2 \phi} = R^{2m^2} Q_n^{(m)}(R) (m \geq 1),$$  \hfill (22)

where the integral can be converted to a summation over source galaxy shapes as an estimator to the integral. The practical estimators avoid summation over $R \rightarrow \infty$ or in the $R < R_1$ region (where the weak-lensing shear estimator breaks down) by integrating over an annular region $R_1 < r < R_2$:

$$\frac{1}{nR} \sum_{R_1 < r_j < R_2} r_j^{-3} \gamma_t,j = R_1^{-3} Q_1^{(0)}(R_1) - R_2^{-3} Q_1^{(0)}(R_2),$$  \hfill (23)

$$\frac{1}{nR} \sum_{R_1 < r_j < R_2} \gamma_t,j - i \gamma_s,j = R_1^{-2} Q_2^{(0)}(R_1) - R_2^{-2} Q_2^{(0)}(R_2),$$  \hfill (24)

$$\frac{1}{nR} \sum_{R_1 < r_j < R_2} (\gamma_t,j + i \gamma_s,j) e^{-i\phi_j} = R_1^{-2} Q_2^{(2)}(R_1) - R_2^{-2} Q_2^{(2)}(R_2),$$  \hfill (25)

where the summation is over the $j$th galaxies at radial distance $r_j$ and azimuth angles $\phi_j$. In Equation (23), the summation over $\gamma_t$ is referred to as the $E$-mode aperture mass, which corresponds to the physical quantity which is the surface mass density, while a summation over $\gamma_s$ is the $B$-mode aperture mass which is expected to vanish, and hence provides some measure of systematic present in the morphological (similar measures are not available for the higher order multipole terms, as discussed in Bernstein & Nakajima 2009). The multipoles on the right-hand sides are that of the combination of the cluster and galaxy potential; hence the summation constrains the combination of the cluster and galaxy multipoles.
4.7. Redshift Estimation

Our mass estimators from tangential shear yields estimates of the convergence, but not the mass sheet density $\Sigma$ itself. It is important to know the critical surface mass density $\Sigma_{\text{crit}}$, which normalizes the convergence $\kappa \equiv \Sigma/\Sigma_{\text{crit}}$ and is redshift dependent, since the convergence obtained from WL (at a given source redshift distribution) must be converted to an appropriate value for SL modeling, where the source is at $z_{\text{src}} = z_{\text{quasar}} = 1.41$.

In order to obtain the critical surface density $\Sigma_{\text{crit}}$ for WL, the source galaxy redshift distribution must be estimated. We adopt the magnitude-dependent parameterization (Baugh & Efstathiou 1993)

$$\frac{dN}{dz}(z, m) \propto z^2 \exp \left[-\left(\frac{z}{z_{\text{med}}(m)}\right)^{3/2}\right],$$

(27)

where $\frac{dN}{dz}(z, m)$ is the magnitude-dependent redshift distribution and $z_{\text{med}}(m)$ is the median redshift as a function of magnitude $m$. From the COSMOS$^7$ ACS data, Leauthaud et al. (2007) obtain the median redshift as a function of F814W magnitude as

$$z_{\text{med}} = (0.18 \pm 0.01) \times m_{\text{F814W}} - (3.3 \pm 0.2).$$

(28)

This relation is valid over the magnitude range $20 < m_{\text{F814W}} < 24$; however, their data suggest that the relation can be extended out to $m_{\text{F814W}} < 26$, based on the UDF$^8$ data (Coe et al. 2006), which show agreement in the $20 < m_{\text{F814W}} < 24$ region with COSMOS data. Hence, we use Equation (28) to estimate our median redshift, since our objects are within this magnitude range (Figure 5).

The final redshift distribution was estimated in the following manner:

1. Divide the galaxies into $\Delta m_{\text{F814W}} = 0.25$ bins.
2. Determine the redshift distribution for each bin using Equation (27) and $z_{\text{med}}$ from Equation (28).
3. Sum over each distribution with each bin properly weighted.

The weight of each bin is determined from the weight each galaxy shape gets in estimating the shear (Bernstein & Jarvis 2002). From the estimated source redshift distribution, we calculate the mean lensing strength, which is proportional to $\Sigma_{\text{crit}}^{-1}$, and hence to

$$\left\{\frac{D_{\text{LS}}}{D_{\text{OS}}}\right\} = \frac{\int dz \int dm \frac{dN}{dz}(z, m) w(m)}{\int dm w(m)} \left(\frac{D_{\text{LS}}}{D_{\text{OS}}}\right),$$

(29)

where $\frac{dN}{dz}(z, m)$ is normalized to unit integral and $w(m)$ is the weight per magnitude bin. We find $\left(\frac{D_{\text{LS}}}{D_{\text{OS}}}\right) = 0.572$, or the mean weak-lensing critical density to be

$$\Sigma_{\text{crit}} = \frac{c^2}{4 \pi G} \frac{1}{D_{\text{LS}}} \left\{\frac{D_{\text{LS}}}{D_{\text{OS}}}\right\}^{-1} = 3800 h M_{\odot} \text{pc}^{-2},$$

(30)

$^7$ Cosmic Evolution Survey (http://cosmos.astro.caltech.edu/). COSMOS is an ACS survey over 1.67 deg$^2$, where a single orbit (~2000 s) exposure is tilled over this relatively wide field in F814W. It has 50% completion for sources 0.5' in diameter at F814W $I_{\text{AB}} = 26.0$ (Scoville et al. 2007).

$^8$ Hubble Ultra Deep Field (http://www.stsci.edu/hst/udf/). The UDF data are a multicolor, deep image over a single ACS FOV (11.97 arcmin$^2$) and have a 10σ limiting magnitude (for a 0.5' diameter aperture) of 28.4 in the F850LP filter, with 144 orbit exposures. The F850LP filter is narrower and less efficient than the F814W filter.

Figure 5. Distribution of F606W (solid line) and F814W (dotted line) AB magnitudes of galaxies used in the WL analysis. The objects were detected in the F606W/F814W combined image, and the magnitudes in each were determined using the SExtractor double-image mode (Bertin & Arnouts 1996). The majority (92%) of our objects have F814W magnitude less than 26. (A color version of this figure is available in the online journal.)

for a flat $\Lambda$CDM cosmology with $\Omega_m = 0.24$, and where $\sigma_{v,0} = 0.355$.

Since the geometry dependence of the gravitational lens system is contained in $\Sigma_{\text{crit}}$, its value differs for the same lens if the source object is at different redshift. We compare the critical density $\Sigma_{\text{crit}}$ obtained above to $\Sigma_{\text{crit}}^0 = 3350 h M_{\odot} \text{pc}^{-2}$ for the lensed quasar source at $z = 1.41$, and find that the lensing signal, which is proportional to $\Sigma_{\text{crit}}^{-1}$, is enhanced by 13% relative to the WL galaxy population. In other words,

$$\kappa^0(r) = \frac{\Sigma_{\text{crit}}}{\Sigma_{\text{crit}}} \kappa(r),$$

(31)

where $\kappa^0(r)$ is the convergence at quasar redshift, $\Sigma_{\text{crit}}$ and $\Sigma_{\text{crit}}^0$ are the critical density for sources at the WL galaxy population and quasar, respectively, and $\kappa(r)$ is the convergence with respect to the WL galaxy population.

5. RESULTS

5.1. Mass Sheet

The average tangential shears at different radii are plotted in Figure 6. The data points, although noisy, are independent of each other, so it is appropriate to fit a model to these data. We assume a core-softened isothermal sphere (CIS) mass distribution

$$\kappa(r) = \frac{\Sigma_0}{\Sigma_{\text{crit}}} [1 + (r/r_c)^2]^{-1/2},$$

(32)

and use Equation (9) to compute the corresponding tangential shear. Fixing the core radius at $r_c = 5', we find $\Sigma_0/\Sigma_{\text{crit}} = 0.47 \pm 0.17$, or $\Sigma_0 = (1800 \pm 600) h M_{\odot} \text{pc}^{-2}$ (1σ error estimates), with reduced $\chi^2$ of 0.93. This corresponds to a cluster velocity dispersion of $\sigma_v = 420 \pm 70 \text{ km s}^{-1} (h = 0.7)$. A fit to a Navarro–Frenk–White (NFW) model (Navarro et al. 1996) was attempted, but did not provide any useful constraint on the concentration.

Figure 7 shows the aperture mass statistic as defined in Equation (12), with respect to a varying inner annulus. The
points in this figure are correlated, since each point is an integral over galaxies in the annulus from $R_{\text{max}} = 186''$ to the inner radii $R$.

The points show the azimuthally averaged radial profile of the mass concentration around Q0957. Because of the WL mass-sheet degeneracy, it is not the true radial profile, but a relative value with respect to a mass sheet averaged within a circular aperture of $r < R_{\text{max}}$.

From the aperture mass, we find the average convergence overdensity to be

$$\Delta \kappa = \bar{\kappa} (< 30'') - \bar{\kappa} (< 186'') = 0.122 \pm 0.048,$$

where the radii $R = 30''$ and $R_{\text{max}} = 186''$ are with respect to G1. The average convergence within $R_{\text{max}} = 186''$ can be estimated from the CIS model, whose fit value yields $\bar{\kappa} (< 186'') = 0.024 \pm 0.012$. The error here is conservative (i.e., not based on the error in the fit) to account for various possible mass distributions. The two results can then be combined to yield

$$\bar{\kappa} (< 30'') = 0.146 \pm 0.049.$$

From the redshift distribution of the WL source galaxies, this convergence corresponds to a mean mass sheet density of

$$\bar{\Sigma} (< 30'') = (550 \pm 190) \, h M_\odot \, \text{pc}^{-2}.$$

For sources at the quasar redshift $z = 1.41$, this mass sheet then corresponds to a convergence of

$$\bar{\kappa} (< 30'') = 0.166 \pm 0.056,$$

for the strong-lensing analysis (see Equation (31)). As explicitly stated above, this “mean mass sheet” includes the G1 mass averaged over the $r < 30''$ disk; hence to obtain the cluster mass sheet $\kappa_c$, the galaxy mass (modeled from SL) must be properly removed from the above quantity (Equation (13); Fadely et al. 2009).

5.1.1. Comparison with Fischer et al.

Our result is consistent with the WL results based on Canada–France–Hawaii Telescope (CFHT) images from Fischer et al.
They obtain \( \bar{\kappa}(\sim 30') = 0.16 \pm 0.05 \) (from their Figure 6) compared with our \( \bar{\kappa}(\sim 30') = 0.146 \pm 0.049 \), where we have compared the shears as obtained from the galaxy shapes, without normalizing to the quasar redshift. While the two results and error magnitudes are nearly identical, they are obtained in a completely different fashion. We obtain the shear calibration factor (our responsivity \( R \)) and propagate the shape noise and shape measurement errors from individual galaxies into the estimated shear in a deterministic manner based on the formalisms of Bernstein & Jarvis (2002), while Fischer et al. calibrate the shear and obtain the error based on Monte Carlo simulations.

The similarity in the error magnitude could possibly be attributed to the similar number of galaxies we use (1651 and 1866 galaxies for Fischer et al. (1997) and this study, respectively), since shear estimate error is typically dominated by the statistical shape noise. Both studies use source galaxies in a similar area and magnitude range. However, the resolution of the images used in the two studies are vastly different: the CFHT images have 0.6 FWHM seeing with 0.207 pixel\(^{-1}\), compared with our 0.13 FWHM with 0.05 pixel\(^{-1}\). Since the images with the smaller seeing and pixel scale should reveal the galaxy shapes better, the similarity in the results indicates that the error is indeed dominated by the statistical error. Therefore, in order to improve upon the mass sheet precision, we would need to increase the number count of the source galaxies, which would be possible if source galaxies of lower S/N in the HST images can be utilized. Currently, the accuracy in our galaxy shape measurement and shear estimation method is limited to high S/N objects (greater than 25). Hence, a shape measurement method which can be shown to be accurate at low S/N would improve precision in the mass sheet measurement, given the same HST images.

Although the two observed convergence are consistent with each other, Bernstein & Fischer (1999) report a convergence of \( \bar{\kappa}(\sim 30') = 0.26 \pm 0.08 \) based on the results of Fischer et al. (1997), when normalized to the quasar redshift. This discrepancy comes from the difference in the assumed \( \Sigma_{\text{crit}} \) values: although the redshift of the quasar has not changed, both the assumed WL source galaxy distribution and the cosmology are different between the two analyses. While we use a magnitude-based empirical redshift distribution (Section 4.7) with a flat ΛCDM with \( \Omega_m = 0.24 \), Fischer et al. (1997) estimate the redshift distribution based on Monte Carlo simulation of an assumed redshift evolution of galaxy size, along with an open CDM cosmology with \( \Omega_m = 0.1 \).

### 5.1.2. Comparison with Other Cluster Mass Estimates

Chartas et al. (2002) use \( 2' \times 2' \) Chandra observation to determine the mass distribution of the cluster. From the spatial distribution of the X-ray luminosity and the X-ray temperature, they find the cluster mass to be \( 5.0^{+1.3}_{-1.0} \times 10^{15} M_\odot \) within a radius of 0.5 \( h_{75} \) Mpc (\( \sim 2' \)) of the cluster center. Our corresponding WL mass estimate within the same radius is \( 3.3 \pm 1.1 \times 10^{15} M_\odot \) (for \( h = 0.7 \)), an agreement within \( \sim 1 \sigma \). We note that since the mass here is an integral over a radial density profile, the cluster mass estimate agreement between WL and X-ray, which assumes a different radial profile (CIS and β-model, respectively), will differ depending on the outermost radius chosen.

On the other hand, Garrett et al. (1992) and Angonin-Willaime et al. (1994) obtain a cluster member velocity dispersion of \( \sigma_v = 715 \pm 130 \) km s\(^{-1}\) from redshifts for 21 probable cluster members, compared with our \( \sigma_v \) equivalent of \( 420 \pm 70 \) km s\(^{-1}\) (\( h = 0.7 \)). The dispersion data correspond to cluster mass of \( (17.8 \pm 6.5) \times 10^{13} M_\odot \) (\( h = 0.7 \)) within a 0.5 \( h_{75}^{1/2} \) Mpc radius. The 2\( \sigma \) discrepancy between the WL and velocity-dispersion mass estimates suggests that the velocity dispersion is possibly highly anisotropic, with an enhanced peculiar velocity along the line of sight. Angonin-Willaime et al. (1994) suggest that cluster member selection could also be the cause for the large dispersion, where they find removing a single galaxy at the edge of the velocity distribution reduces the velocity dispersion to 660 km s\(^{-1}\).

### 5.2. Multipole Moments

The rest of the (differential) multipole moments as obtained from Equations (24)–(26) are listed in Table 1, where the differential is between the values at \( R = 30' \) and \( R_{\text{max}} = 186' \). Figure 8 plots the multipoles with respect to the value at \( R_{\text{max}} \) for various inner radii \( R \), which highlights the problem with constraining the \( Q_{\text{in}}^{(m)} \) terms. Due to the powers of \( R \) involved in the summation, the shear signal, as well as the shape noise, is weighted by \( R^{-2m-2} \) for \( Q_{\text{in}}^{(m)} \), and by \( R^{2m-2} \) for \( Q_{\text{out}}^{(m)} \). For \( Q_{\text{in}}^{(m)} \), the outer region (dominated by shape noise) is weighted less, and the signal at smaller radii can be seen (Figure 8 left).

At \( Q_{\text{out}}^{(m)} \), \( m > 1 \), (Figure 8 center and right), whatever signal existing at the inner radii is overwhelmed by the shape noise in the outer radii, and hence the differential signal as \( R \) decreases to 30' is merely noise which dominates at larger radii. This effect is more severe for \( Q_{\text{out}}^{(3)} \) (Figure 8 right), where higher powers of \( R \) are involved, that the shear from \( R < 100' \) hardly contributes to the integrated shear.

If we are to constrain the multipoles at \( R = 30' \), it is also necessary to have a handle on their values at \( R_{\text{max}} \). This value is estimated to be up to 20% of the values at \( R = 30' \) for the \( Q_{\text{in}}^{(2)} \) and \( Q_{\text{out}}^{(3)} \) terms, and around 1% for the \( Q_{\text{in}}^{(1)} \) term, for a CIS cluster whose center is displaced from the G1 center by 5'–30'. Since our multipole signals are already significantly smaller than the measurement errors, we deduce that the correction from the \( R_{\text{max}} \) term is significantly smaller than the errors if it arises from a cluster that resembles a CIS profile beyond \( R_{\text{max}} \).

Ignoring the \( R_{\text{max}} \) terms from the cluster and assuming the lensing potential as stated in Equation (8), the cluster expansion coefficients (normalized to the quasar redshift) are constrained to

\[
(1 - \kappa_c) \sigma_c + \frac{4}{\pi} (1 - \kappa_c) \left[ R_1^{-4} Q_{\text{in},g}^{(1)} (R_1) - R_2^{-4} Q_{\text{in},g}^{(1)} (R_2) \right] = (-0.0018 \pm 0.0042) + i(-0.0031 \pm 0.0052) \text{arcsec}^{-1},
\]

\[
(1 - \kappa_c) \gamma_c + \frac{1}{\pi} (1 - \kappa_c) \left[ Q_{\text{out},g}^{(2)} (R_1) - \left( \frac{R_2}{R_1} \right)^2 Q_{\text{out},g}^{(2)} (R_2) \right] = (0.00 \pm 0.34) + i(0.09 \pm 0.33),
\]

\[
(1 - \kappa_c) \delta_c + \frac{2}{\pi} (1 - \kappa_c) \left[ Q_{\text{out},g}^{(3)} (R_1) - \left( \frac{R_2}{R_1} \right)^4 Q_{\text{out},g}^{(3)} (R_2) \right] = (+0.115 \pm 0.098) + i(-0.047 \pm 0.097) \text{arcsec}^{-1}.
\]

As explicitly stated above, the galaxy model contribution \( Q_{g}^{(m)} \) to each moment must be subtracted from the results to obtain...
the cluster moments $\gamma_c, \sigma_c$, and $\delta_c$; the measurements constrain only the sum of the galaxy and cluster contributions.

As seen in Figure 8, the error in the $\gamma_c$ and $\delta_c$ terms do not converge to any sufficient degree. However, we attempt to constrain $\gamma_c$ by the following means: since this term is simply the constant shear across $r < 30''$, we measure the average of the (WL) shear within this region. However, since the central portion of this region is strongly lensed, we average the shear signal over an annular region of $20'' < r < 40''$ to estimate the constant shear within $r < 30''$. We find the average shear, normalized to the quasar redshift, to be

$$ (1 - \kappa_c) \gamma_c + \frac{1}{\pi} (1 - \kappa_c) \left[ \frac{Q_{\text{out},g}^{(2)}(R_1)}{R_1} - \frac{Q_{\text{out},g}^{(2)}(R_2)}{R_2} \right] $$

$$ = (-0.009 \pm 0.045) + i(+0.092 \pm 0.045). $$

This allows us to shrink the error bar considerably, and by doing so we have assumed that $\gamma_c$ is generated exclusively by mass exterior to 40''. The directionality of the constant shear term is consistent with an exterior mass quadrupole distribution with positive weight along the northeast/southwest direction with respect to G1.

6. CONCLUSION

The Q0957+561 gravitational lens system allows for a determination of the Hubble constant $H_0$, based on its firm time delay between the double images of the quasar; the remaining uncertainty originates from our knowledge of the lens mass distribution. The lens consists of the galaxy G1 and the Q0957+561 cluster to which the galaxy belongs. WL offers complementary constraints to the lens mass provided by the SL modeling analysis, by providing the mass sheet and multipole moments of the underlying cluster. We have developed and utilized the formalism for an exact solution for aperture mass multipoles from WL shear data in the thin-lens approximation.

The mean convergence in the SL region $r < 30''$ is estimated to be $\kappa = 0.166 \pm 0.036 (1\sigma)$ normalized to the quasar redshift, based on shear data from the annular region $30'' < r < 186''$, where the radii are centered on the lens galaxy G1. Although our shear measurements are consistent with Fischer et al. (1997), the new quasar-normalized $\kappa$ is more reliable since they are based on better data, a more accurate galaxy redshift distribution, and a standard cosmology. Our results give a 7% precision in the mass-sheet degeneracy term $(1 - \pi)$.

The uncertainty in the external multipole terms $\gamma_c$ and $\delta_c$ is too large to provide useful constraints to the lens potential $\psi$ within $r < 30''$. However, the constant shear term $(1 - \kappa)\gamma_c$ within $r < 30''$ can be estimated by straightforward averaging of the shear in the WL region at $20'' < r < 40''$. The internal dipole $(1 - \kappa)\sigma_c$ (whose galaxy contribution is expected to be negligible) has a value of $0.006 \pm 0.006$, compared with the Monte Carlo parametric fit values $-0.011 \pm 0.007$ obtained by Keeton et al. (2000). The implications for $H_0$ with the full SL lens modeling are discussed in an accompanying paper (Fadely et al. 2009).

This work has been supported by grant HST-GO-10569 from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. R.N. and G.M.B. acknowledge additional support from NASA grant BEFS-04-0014-0018. T.S. acknowledges financial support from the Netherlands Organization for Scientific Research (NWO) and the German Ministry for Education and Research (BMBF) through TR33.

REFERENCES

Adelman-McCarthy, J. K., et al. (SDSS DR6) 2008, ApJS, 175, 297
Anderson, J. 2002, in The 2002 HST Calibration Workshop, Space Telescope Science Institute, ed. S. Arribas, A. Koekemoer, & B. Whimber (Baltimore, MD: STScI), 13
Anderson, J. 2006, in The 2005 HST Calibration Workshop, Space Telescope Science Institute, ed. A. M. Koekemoer, P. Goudfrooij, & L. L. Dressel (Baltimore, MD: STScI), 1
Angonin-Willaine, M.-C., Soucail, G., & Vanderriest, C. 1994, A&A, 291, 411
Baugh, C. M., & Efstathiou, G. 1993, MNRAS, 265, 145
Bernstein, G. M., & Fischer, P. 1999, AJ, 118, 14
Bernstein, G. M., & Jarvis, M. 2002, AJ, 123, 583 (BJ02)
Bernstein, G. M., & Nakajima, R. 2009, ApJ, 693, 1508
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bruzual, A. G., & Charlot, S. 1993, ApJ, 405, 538
Chartas, G., Gupta, V., Garmire, G., Jones, C., Falco, E. E., Shapiro, I. I., & Tavecchio, F. 2002, ApJ, 565, 96
Coe, D., Benítez, N., Sánchez, S. E., Jee, M., Bouwens, R., & Ford, H. 2006, AJ, 132, 926
Colley, W., et al. 2003, ApJ, 598, 71
Eisenstein, D. J., et al. 2001, AJ, 122, 2267
Fadely, R., Keeton, C. R., Nakajima, R., & Bernstein, G. B. 2009, ApJ, submitted
Fahlman, G., Kaiser, N., Squires, G., & Woods, D. 1994, ApJ, 437, 56
Falco, E. E., Goreinstein, M. V., & Shapiro, I. I. 1985, ApJ, 289, L1
Fischer, P., Bernstein, G., Rhee, G., & Tyson, J. A. 1997, AJ, 113, 521
Garrett, M. A., Walsh, D., & Carswell, R. F. 1992, MNRAS, 254, 27
Gorenstein, M. V., Cohen, N. L., Shapiro, I. I., Rogers, A. E. E., Bonometti, R. J., Falco, E. E., Bartel, N., & Marcaide, J. M. 1988, ApJ, 327, 693
Jarvis, M., & Jain, B. 2004, arXiv:astro-ph/04112234
Jee, M. J., Blakeslee, J. P., Sirianni, M., Martel, A. R., White, R. L., & Ford, H. C. 2007, PASP, 119, 1403
Johnston, D. E., et al. 2007, arXiv:0709.1159
Kaiser, N., & Squires, G. 1993, ApJ, 404, 441
Keeton, C. R., et al. 2001, ApJ, 542, 74

Figure 8. Left: differential $R^{-1}Q_{\text{out},g}^{(2)}(R)$ with respect to the value at $R_{\text{max}}$. Center: differential $R^2Q_{\text{out},g}^{(2)}(R)$. Right: differential $R^4Q_{\text{out},g}^{(2)}(R)$. In all three figures, the solid squares indicate the real components, while the open squares indicate the imaginary components. The complex numbers encode the multipole orientation, where the axes are aligned with west ("x") and north ("y"). The internal mass dipole term has a directionality that points approximately to the northeast with respect to G1, consistent with previous studies (Fischer et al. 1997; Chartas et al. 2002).

(A color version of this figure is available in the online journal.)
Kochanek, C. S. 1991, ApJ, 382, 58
Koekemoer, A. M., Fruchter, A. S., Hook, R. N., & Hack, W. 2002, in The 2002 HST Calibration Workshop, Space Telescope Science Institute, ed. S. Arribas, A. Koekemoer, & B. Whitmore (Baltimore, MD: STScI), 337
Koester, B. P., et al. 2007, ApJ, 660, 239
Kundic, T., et al. 1997, ApJ, 482, 75
Leauthaud, A., et al. 2007, ApJS, 172, 219
Miralda Escudé, J. 1991, ApJ, 380, 1
Monet, D. G., et al. 2003, AJ, 125, 984
Nakajima, R., & Bernstein, G. M. 2007, AJ, 133, 1763
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 463, 563
Refsdal, S. 1964, MNRAS, 128, 307
Rhodes, J. D., et al. 2005, in The 2005 HST Calibration Workshop Space Telescope Science Institute, ed. A. M. Koekemoer, P. Goudfrooij, & L. L. Dressel (Baltimore, MD: STScI), 21
Rhodes, J. D., et al. 2007, ApJS, 172, 203
Schneider, P. 1996, MNRAS, 283, 837
Schneider, P., & Bartelmann, M. 1997, MNRAS, 286, 696
Schrabback, T., et al. 2007, A&A, 468, 823
Scoville, N., et al. 2007, ApJS, 172, 38
Shalyapin, V. N., et al. 2008, A&A, 492, 401
Walsh, D., Carswell, R. R., & Weymann, R. J. 1979, Nature, 279, 381