



A Likely Supermassive Black Hole Revealed by Its Einstein Radius in Hubble Frontier Fields Images

Mandy C. Chen¹ , Tom Broadhurst^{2,3,4}, Jeremy Lim¹, Jose M. Diego⁵ ,

Youichi Ohyama⁶ , Holland Ford⁷ , and Narciso Benítez⁸

¹ Department of Physics, The University of Hong Kong, Pokfulam Road, Hong Kong; chenc@connect.hku.hk

² Department of Theoretical Physics, University of the Basque Country UPV/EHU, Bilbao, Spain

³ IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

⁴ Donostia International Physics Center (DIPC), 20018 Donostia-San Sebastian (Gipuzkoa), Spain

⁵ ICA, Instituto de Física de Cantabria (UC-CSIC), Av. de Los Castros s/n, E-39005 Santander, Spain

⁶ Institute of Astronomy and Astrophysics, Academia Sinica, Taipei 10617, Taiwan

⁷ Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA

⁸ Instituto de Astrofísica de Andalucía (IAA-CSIC), Glorieta de la Astronomía, E-18008 Granada, Spain

Received 2018 May 7; revised 2018 June 20; accepted 2018 July 2; published 2018 August 17

Abstract

At cosmological distances, gravitational lensing can in principle provide direct mass measurements of supermassive black holes (SMBHs). Here, we directly estimate the mass of a SMBH in the brightest cluster galaxy (BCG) of MACS J1149.5+2223 at $z = 0.54$ using one of the multiply lensed images of a background spiral galaxy at $z = 1.49$ projected close to the BCG. A lensed arc is curved toward the BCG center, corresponding to an intrinsically compact region in one of the spiral arms. This arc has a radius of curvature of only $\sim 0''.6$, betraying the presence of a local compact deflector. Its curvature is most simply reproduced by a point-like object with a mass of $8.4_{-1.8}^{+4.3} \times 10^9 M_{\odot}$, similar to SMBH masses in local elliptical galaxies having comparable luminosities. The SMBH is noticeably offset by 4.4 ± 0.3 kpc from the BCG light center, which is plausibly the result of a kick imparted $\sim 2.0 \times 10^7$ years ago during the merger of two SMBHs, placing it just beyond the stellar core. A similar curvature can be produced by replacing the offset SMBH with a compact galaxy having a mass of $\sim 2 \times 10^{10} M_{\odot}$ within a cutoff radius of < 4 kpc, and an unusually large $M/L > 50(M/L)_{\odot}$ to make it undetectable in the deep Hubble Frontier Fields image, at or close to the cluster redshift. However, such a lensing galaxy perturbs the adjacent lensed images in an undesirable way.

Key words: galaxies: clusters: individual (MACS J1149.5+2223) – galaxies: elliptical and lenticular, cD – galaxies: evolution – galaxies: nuclei – gravitational lensing: strong

1. Introduction

While the ubiquity of supermassive black holes (SMBHs) at the centers of relatively massive galaxies is widely accepted, the origin and growth of these enigmatic objects remain poorly understood (if at all known). The now familiar $M_{\text{BH}}-\sigma$ relation suggests a co-evolution in mass between the SMBH and its host galaxy (e.g., Kormendy & Ho 2013), providing support for hierarchical mergers in structure formation theories. However, at very early times ($z > 6.0$), surprisingly large SMBH masses of $\sim 10^9 M_{\odot}$ have been inferred to be associated with host galaxies having relatively low masses (e.g., Trakhtenbrot et al. 2015; Wu et al. 2015). Similarly, in the local Universe, SMBHs that lie well above the established $M_{\text{BH}}-\sigma$ relationship have been found, comprising perhaps “fossil” cases with little past merging (van den Bosch et al. 2012; Emsellem 2013). As the accuracy of our measurements of local black hole and host galaxy masses improves, the traditional $M-\sigma$ relations (Ferrarese & Merritt 2000) have become more complex (Graham 2016). To understand the growth of SMBH masses over cosmic time, what is clearly needed is the ability to accurately measure their masses over a broad range of epochs.

To date, three widely used methods have been employed to measure SMBHs: modeling of stellar or gas kinematics, reverberation mapping, and scaling relations developed in large part from the results of reverberation mapping. For local galaxies where high spatial resolutions are possible, SMBH

masses are determined through measurements of stellar or gas kinematics within a region where the gravitational force of the SMBH is dominant (e.g., Ferrarese & Ford 1999; Kormendy 2004; Meyer et al. 2012). For more distant galaxies that have optically bright active galactic nuclei (AGNs), reverberation mapping can be used to infer the masses of their SMBHs. This approach requires the measurement of two parameters during a change (increase) in the brightness of an AGN. One of these parameters is the width of the emission lines from the broad-line region (BLR) of the AGN. The width of a given emission line is attributed to the orbital motion of that line-emitting gas surrounding the accretion disk of the SMBH. The second parameter is the time delay between a change in the continuum (from the accretion disk) and the emission-line (from the BLR) fluxes. This time delay corresponds to the light travel time between the accretion disk (which is assumed to be very small) and the region in the BLR at which a given emission line arises and, therefore, the radius of the line-emitting region from the SMBH. The mass of the SMBH can be derived from these two parameters. The major source of uncertainty in this method is the uncertainty in the BLR geometry. Because derivations of SMBH masses from reverberation mapping requires measurements with high signal-to-noise, this method has so far been restricted primarily to AGNs in nearby galaxies ($z < 0.3$) (Kaspi et al. 2005, 2007). The results from reverberation mapping have been used to develop empirical scaling relations between the radius of BLR and the AGN optical luminosities based on different emission lines (e.g., $H\beta$ and Mg II), whereby

the radius of the BLR have been found to scale with the AGN luminosity. This relationship implies a long time delay between changes in the continuum and emission-line fluxes at high AGN luminosities, making reverberation mapping impractical for these objects. Instead, the scaling relations that have been developed from reverberation mapping have been used to infer SMBH masses in luminous AGNs and quasi-stellar objects (QSOs) at intermediate or high redshifts (see review by Bentz et al. 2009).

Gravitational lensing provides a promising new approach to directly measure SMBH masses—one that does not depend on whether the SMBH is active or not and which is almost irrespective of distance. The Einstein radius, θ_e , of a point mass depends simply on the distances involved and scales slowly

with mass: $\theta_e = 0''.3 \left(\frac{M}{10^{10} M_\odot} \right)^{\frac{1}{2}} \left(\frac{D}{\text{Gpc}} \right)^{-\frac{1}{2}}$ for a typical lens

redshift of $z \simeq 0.5$ ($D = \frac{D_l D_s}{D_l}$ where D_l , D_s and D_{ls} are the angular diameter distances to the lens, the source, and between the lens and the source, respectively). At a limiting angular resolution of $\simeq 0''.1$, this Einstein radius is resolvable for a wide range in point masses of $M > 10^8 M_\odot$. In the situation where the background source and the foreground lensing object are closely aligned in the sky, a central de-magnified image is generic to this lensing geometry; such that the larger the SMBH mass, the more this central image is attracted toward the SMBH and is de-magnified (Mao et al. 2001; Hezaveh et al. 2015; Rusin et al. 2005).

One caveat of this approach, however, arises from the fact that when producing a certain magnification factor on the central lensed image, the SMBH mass is degenerate with the slope of the lensing galaxy’s central mass profile; as is clearly illustrated in Hezaveh et al. (2015). Consequently, the presence (or absence) of a central image can only provide the upper (or lower) limit of the SMBH mass. The first example was the multiply lensed quasar PMN J1632-0033. Observations with the Very Long Baseline Array (VLBA) and Very Large Array (VLA) revealed a central image, thus constraining the mass of the central SMBH in the foreground lensing galaxy to be $M_{\text{BH}} < 2 \times 10^8 M_\odot$ (Winn et al. 2003, 2004). More recently, radio observations with the Atacama Large Millimeter/submillimeter Array (ALMA) of the lensed system SDP81 have failed to detect a central image within the Einstein ring of the lensed background galaxy, thus placing a lower limit of $M_{\text{BH}} \sim 3 \times 10^8 M_\odot$ on the mass of a central SMBH in the foreground lensing elliptical galaxy at $z = 0.3$ (Tamura et al. 2015; Wong et al. 2015). The SMBH masses thus derived for the foreground lensing galaxies in both PMN J1632-0033 and SDP81 are in agreement with the local $M_{\text{BH}}-\sigma$ relation.

Quinn et al. (2016) discussed the implications for the non-detection of a central image in the lens system CLASS B1030+074 with the data from VLA and the extended Multi-Element Remote-Linked Interferometer (e-MERLIN), and argued in favor of a central SMBH with a mass slightly greater than that implied by the local $M_{\text{BH}}-\sigma$ relation; yet again, whether the SMBH is required in the lens model is dependent on the mass profile chosen for the lensing galaxy. A relatively large SMBH mass of $\sim 1.2 \times 10^{10} M_\odot$, lying well above the $M_{\text{BH}}-\sigma$ relation, has been inferred for the brightest cluster galaxy (BCG; central giant elliptical galaxy) in the cluster Abell 1201 at $z = 0.17$ based on the detection of a faint central image based on observations with the *Hubble Space Telescope* (HST) (Smith et al. 2017). In this case, however, the authors find that

the observed parameters of the central image can be equally well explained by a cuspy stellar M/L ratio for the BCG.

In this paper, we report a direct measurement of the mass of an SMBH through gravitational lensing. This SMBH is hosted by the BCG in the galaxy cluster MACS J1149.5+2223 (hereafter MACS 1149) at a redshift of $z = 0.543$ (Ebeling et al. 2007). Unlike the previous examples, in this case the background lensed galaxy and BCG are not closely aligned in the sky. Instead, one of the multiply lensed images of this background galaxy happens to be projected close to the BCG, which locally perturbs this image. One of the numerous compact H II regions in the background spiral galaxy is lensed into a curved arc, pointing to and betraying the presence of an SMBH. MACS 1149 is one of the six clusters from the Hubble Frontier Fields (HFF) program (PI: J. Lotz). Gravitational lensing by this cluster has been intensively studied (e.g., Smith et al. 2009; Zitrin & Broadhurst 2009; Zheng et al. 2012; Rau et al. 2014; Sharon & Johnson 2015; Grillo et al. 2016). The first (and so far only) multiply lensed supernova was detected in this cluster (Kelly et al. 2015, 2016; Oguri 2015; Treu et al. 2016). The discovery of a transient in a lensed image in this cluster has been attributed to microlensing by intracluster stars of a single star in a background lensed galaxy (Diego et al. 2018; Kelly et al. 2018).

Understanding and reproducing the curved lensed arc in the image of the aforementioned background spiral galaxy involves the following separate steps: (1) deducing a robust cluster lens model for MACS 1149 and refining the mass model for the BCG to produce all of the multiply lensed images found toward the cluster, particularly those projected close to the BCG, as described in Sections 2 and 3; (2) inferring the need for a local deflector to reproduce the curvature of L1, as described in Section 4; and (3) determining the nature of this local deflector and its physical parameters, as described in Sections 5 and 6. Our results are discussed in Section 7, and a concluding summary is presented in Section 8.

2. Data

2.1. HFF and CLASH

We retrieved the already reduced HFF public imaging data (epoch2) for MACS 1149 taken with the Hubble ACS and WFC3 (PI: J. Lotz) from the Mikulski Archive for Space Telescopes (MAST).⁹ From the HFF, a total of 140 orbits in the optical and infrared bands (filters: F435W, F606W, F814W, F105W, F125W, F140W and F160W) are devoted to MACS 1149. We used this data to construct the cluster lens model. MACS 1149 is also one of the clusters observed in the Cluster Lensing And Supernovae Survey with Hubble (CLASH) program (PI: M. Postman), for a total of 18 orbits in the UV, optical, and infrared bands spanning 17 filters.¹⁰ For the purpose of this study, the F435W data from CLASH (1.5 orbits) was also added to the F435W data from HFF (18 orbits), albeit leading to only a small improvement of the signal-to-noise ratio due to the significantly shorter exposure time of CLASH.

Figure 1 shows a $0''.8 \times 0''.8$ multi-band image of MACS 1149 constructed from the F125W (red), F814W (green) and F435W (blue) filters of the HFF. The most prominent lensed image is that of a multiply lensed spiral galaxy at $z = 1.4888$.

⁹ <https://archive.stsci.edu/pub/hlsp/frontier/mac1149/images/hst/v1.0-epoch2/>

¹⁰ <https://archive.stsci.edu/prepds/clash/>

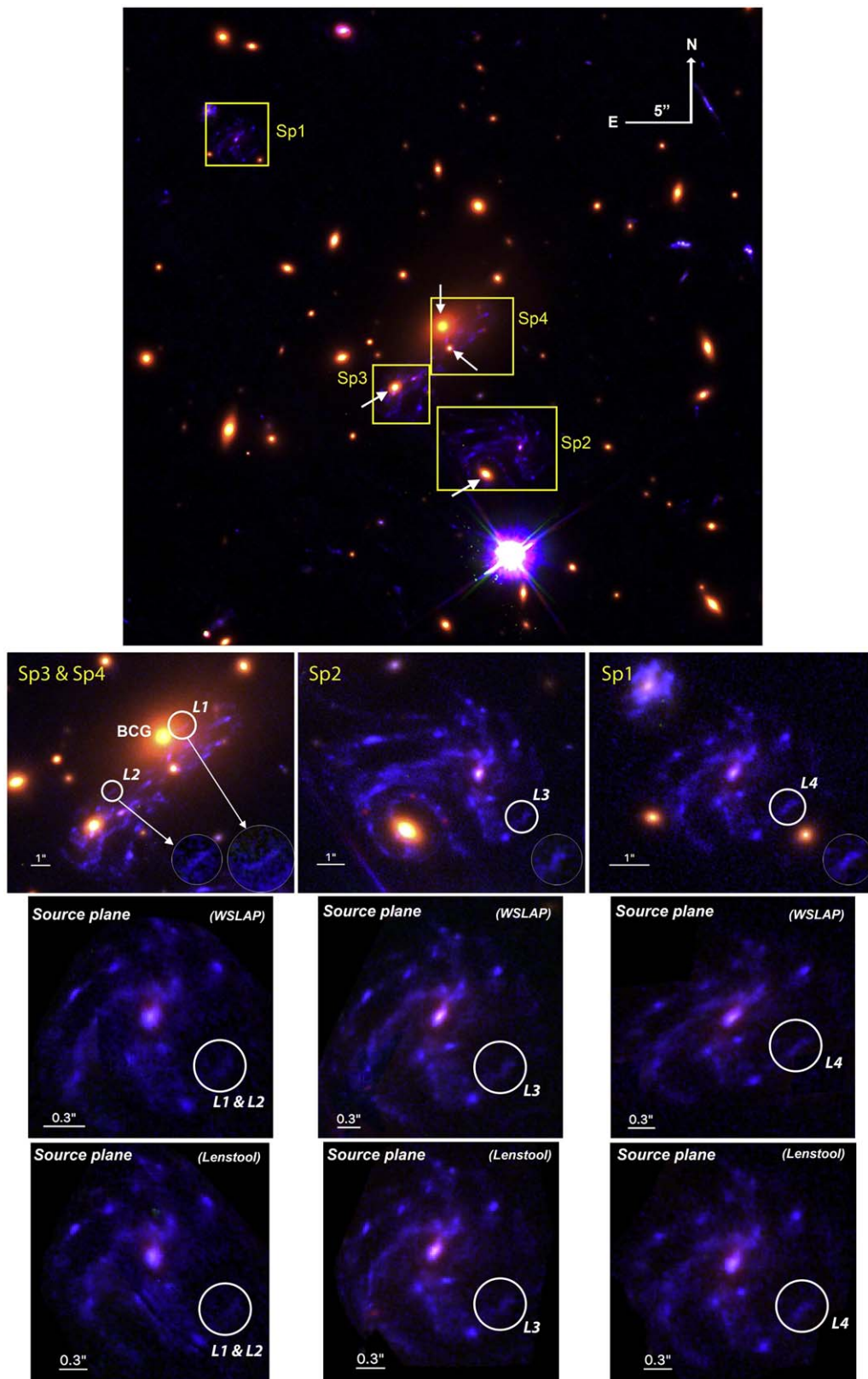


Figure 1. Row 1: HFF data of MACS 1149 ($0'.8 \times 0'.8$) in the F125W (red), F814W (green) and F435W (blue) bands. Row 2: close-up of the individual lensed images Sp1-4. Sp4 is projected close to the BCG, with an H II region in its spiral arm labeled L1 projected close to the BCG center. This H II region appears four times, labeled L2-4 in the other counter images, and is shown in greater detail by a magnified insert at the bottom right-hand corner of each panel. Row 3 and 4: the delensed images of the spiral galaxy obtained by applying a free-form WSLAP+ lens model and a parametric *Lenstool* model to each of the multiple images, respectively. Sp3 and Sp4 are combined to form one delensed image because Sp4 is lensed from only a small segment of the entire spiral galaxy. The three independent source plane images of this spiral galaxy are in good agreement with each other, even though Sp1 is relatively far from the cluster center where the lens model is less well constrained.

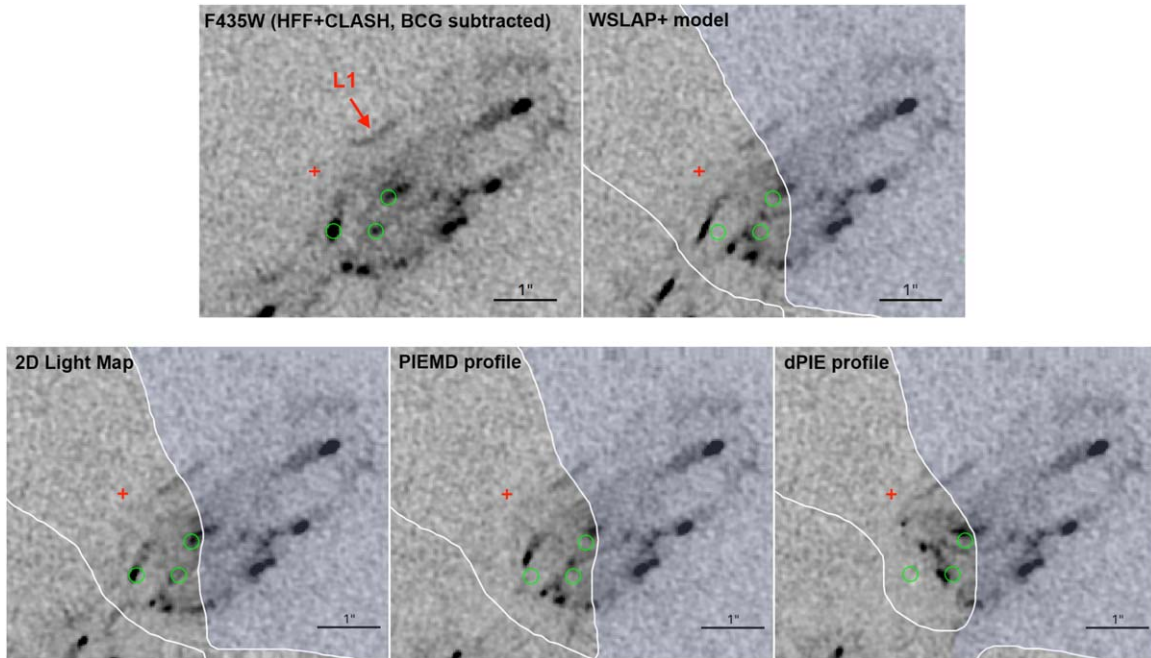


Figure 2. Red cross indicates the light centroid of the BCG. Upper left: the data in the region of interest close to the center of the BCG (after the BCG light is subtracted). Upper right: the relensed image from our free-form lens model (WSLAP+) where a NFW mass profile is assumed for the BCG. The blue-highlighted region on the right-hand side of a critical curve (white line) is the input data that is used for relensing, and that on the left-hand side the corresponding relensed images. Note that image L1 is predicted to be straighter than is actually observed. Lower row: same as upper right, but differs in the way the BCG mass distribution is modeled to seek closer agreement between the predicted and observed relensed images. By using the 2D light distribution for the BCG, we obtain a better agreement in terms of the reproduction of the relensed images (lower left-hand) but again with the obvious exception of the curved appearance of L1. In the lower middle and right-hand panels, we model the BCG with a dPIE and a PIEMD mass profile, as adopted by Grillo et al. (2016). The green circles point out three major features close to the center of the BCG in the relensed region, and they are plotted in all the model panels to indicate the offset of the model predicted positions. The mean rms offset of these three relensed images from their positions in the data ($\langle \text{rms}_i \rangle$) in the four models are: $0''.22$, $0''.13$, $0''.21$, and $0''.49$.

With the lensed images magnified by up to ~ 200 times, this object is one of the most highly magnified lensed galaxies yet discovered (Zitrin & Broadhurst 2009). We label the individual multiply lensed images as Sp1-4, as shown in Figure 1 (row 2). Sp1-3 corresponds to contain complete images of this spiral galaxy, whereas Sp4 corresponds to only a partial image; note that certain features in both Sp3 and Sp4 appear more than once owing to additional lensing by the BCG and the bright elliptical galaxy near Sp3 (indicated by the white arrow closest to this lensed image). One of the numerous H II regions in Sp4, which is labeled “L1” and which lies closest to the BCG, has a curved (banana-like) shape. As we will show in Section 3, L1 is itself a doubly lensed image with a critical curve passing through it, as predicted by different constructions of lens models for this cluster. The radius of curvature of L1 is only $\sim 0''.6$, which as we will show is much smaller than the Einstein radius of either the BCG or the cluster and, therefore, requires a compact lensing mass that is close to L1 on the sky.

To better reveal L1 and also nearby lensed features at the vicinity of the BCG center, we also subtracted the BCG light from the F435W data, as shown in Figure 2, 1st panel. The subtraction is not model dependent given the uniformity of the color of early type galaxies because we can simply scale the very bright image of the BCG at F160W band and subtract it from the F435W band. The scaling factor is decided by trial and error so as to best remove the BCG light without over-subtraction. This removal is straightforward in the F435W band because the BCG is barely detectable and lensed images from the background spiral galaxy at $z = 1.4888$ have the highest

contrast against the BCG light in this passband. As can be seen in Figure 2, 1st panel, only noise remains after subtraction.

2.2. MUSE/VLT

We analyzed the data of MACS 1149 taken with Multi-Unit Spectroscopic Explorer (MUSE) at the Very Large Telescope (VLT). Sp1-4 was observed for the purpose of studying the kinematics of the background spiral galaxy, which assists the strong lensing analysis. We retrieved the “phase-3” processed data of program ID 294.A-5032 (PI: C. Grillo) from the ESO Science Archive Facility,¹¹ and used the “MUSE-DEEP” product, where all individual exposures taken on different days under the same program ID are fully calibrated (with MUSE pipeline muse-1.6.1) and combined. This dataset was originally presented in Grillo et al. (2016), and all details of the observations are described there.

The MUSE instrument was configured for its wide field of view (FOV) ($1' \times 1'$) mode without the AO system. The standard MUSE IFU and spectrograph configuration provides a spectral cube covering a wide wavelength range (4750–9350 Å) with a spectral resolution of $R \sim 3000$ at a pixel scale of $0''.2 \times 0''.2$. The total exposure time was 17280 s, and the seeing size was $1''.0$ based on the measurement of a bright star within the FOV on the combined frame at a wavelength near the redshifted [O II] emission (~ 9300 Å; see Section 4.1). We refined the astrometry by aligning the stellar continuum map generated from our spectral fitting analysis with the *HST* images to an accuracy of $< 0''.1$.

¹¹ <http://archive.eso.org/cms/data-portal.html>

3. Lens Modeling

3.1. WSLAP+ Model of MACS 1149

The global lens model of the cluster is derived using our code WSLAP+ (Weak and Strong Lensing Analysis Package +) (Diego et al. 2005, 2007; Ponente & Diego 2011; Sendra et al. 2014). This method adopts a free-form philosophy where the lens plane is divided into a pixelated grid. Each pixel is represented as a Gaussian mass profile, where the FWHM may be varied to generate a multi-resolution grid or is held constant to provide a uniform grid (Diego et al. 2005). The division of the lens plane into grid points allows us to divide the deflection field, α , into the individual contributions to the deflection field from the pixel grid. For MACS 1149, we used a multi-resolution grid with 280 cells in total. The resolution of the grid in the cluster center is $2''.16$ and the resolution in the edges is $6''.56$. A further improvement was implemented by including member galaxies of the cluster with NFW mass profiles scaled by their measured luminosities (Sendra et al. 2014), and for which the only free parameter is the scaling of the M/L ratio (size scale), for all bright member galaxies included in the model. This M/L ratio and the Gaussian masses in the grid points are derived by minimizing a quadratic function. The minimum of this quadratic function is also the solution of a system of linear equations that describe the observed data (see Equation (3) below). Our method is described below.

Given the standard lens equation,

$$\beta = \theta - \alpha(\theta, \Sigma(\theta)), \quad (1)$$

where θ is the observed position of the source, α is the deflection angle, $\Sigma(\theta)$ is the projected surface mass density of the cluster at the position θ , and β is the position of the background source. Both the strong lensing and weak lensing observables can be expressed in terms of derivatives of the lensing potential,

$$\psi(\theta) = \frac{4GD_b D_{ls}}{c^2 D_s} \int d^2\theta' \Sigma(\theta') \ln(|\theta - \theta'|), \quad (2)$$

where D_b , D_{ls} and D_s are the angular diameter distances to the lens, from the lens to the source and from the observer to the source, respectively. The unknowns of the lensing problem are in general the surface mass density (or masses in our grid points) and the positions of the background sources. The weak and strong lensing problem can be expressed as a system of linear equations that can be represented in a compact form (Diego et al. 2007),

$$\Theta = \Gamma X, \quad (3)$$

where the measured strong and weak lensing observables are contained in the array Θ of dimension $N_\Theta = 2N_{SL} + 2N_{WL}$, the unknown surface mass density and source positions are in the array X of dimension $N_X = N_c + N_g + 2N_s$, and the matrix Γ is known (for a given grid configuration and fiducial galaxy deflection field, see below) and has dimension $N_\Theta \times N_X$. N_{SL} is the number of strong lensing observables (each one contributing with two constraints, x , and y), N_{WL} is the number of weak lensing observables (each one contributing with two constraints, γ_1 , and γ_2), and N_c is the number of grid points (or cells) that we use to divide the field of view. N_g is the number of deflection fields (from cluster members) that

we consider. N_s is the number of background sources (each contributes with two unknowns (Sendra et al. 2014), β_x , and β_y). The solution is found after minimizing a quadratic function that estimates the solution of the system of Equation (3). For this minimization, we use a quadratic algorithm that is optimized for solutions with the constraint that the solution, X , must be positive (Diego et al. 2005). This is particularly important because by imposing this constraint we avoid the unphysical situation where the masses associated to the galaxies are negative; which could, from the formal mathematical point of view, otherwise provide a reasonable solution to the system of linear Equation (3). Imposing the constrain $X > 0$ also helps in regularizing the solution because it avoids large negative and positive contiguous fluctuations. We did not use weak lensing constraints to model MACS 1149. Strong lensing constraints come from 16 multiply lensed galaxies, some of which contain individual resolved features that add up to the total number of constraints. In the Appendix, we mark the locations of multiply lensed galaxies in Figure 21 and we list the coordinates of multiple images used as constraints in Table 1.

Previous work has shown how the addition of the small deflection fields from member galaxies can help to improve the mass determination when enough constraints are available (Kassiola et al. 1992; Kneib et al. 1996; Sendra et al. 2014). In a previous study, we quantified via simulations how the addition of deflections from all the main member galaxies helps improve the mass reconstruction with respect to our previous standard non-parametric method (Sendra et al. 2014). Strongly-lensed galaxies are often locally affected by member galaxies. However, these perturbations cannot be recovered in grid based reconstructions because the lensing information is too sparse to resolve member galaxies.

For this study, we select elliptical galaxies in the cluster and assign a mass according to luminosity. For the fiducial deflection field, we assume that the mass of the member galaxies scales by a fixed M/L ratio. The optimization procedure determines the proportionality constant that allows for the best reproduction of the data. As mentioned previously, for the mass profiles, we assume a NFW profile (Navarro et al. 1996) and adopt a self-similarity so that the scale radius is proportional to mass. Note that the choice of the particular profile for these perturbing galaxies is not very relevant in terms of reproducing the multiply lensed images of a given background galaxy because the deflection angle is small compared to that produced by the smooth cluster component. We use two deflection fields for the cluster members (i.e $N_g = 2$, see definition of N_g above), thus allowing for different M/L ratios for the separate deflection fields. The first is associated to the BCG and three other galaxies that are the most close to the lensed images of the background spiral galaxy of interest (pointed with the white arrows in Figure 1) and the second contains the deflection field from the remaining dominant galaxies in the cluster. Each deflection field contributes in our model as one free parameter (its amplitude with respect to the fiducial amplitude). More details of the global lens modeling of this cluster can be found in our previous study: Diego et al. (2016). This model successfully predicted the reappearance of the SN Refsdal (Treu et al. 2016).

Table 1
Coordinates of Multiply Lensed Images Used as Constraints for Lens Modeling

Image ID	R.A.	Decl.
1.1.1	11:49:35.283	+22:23:45.64
1.1.2	11:49:35.213	+22:23:43.35
1.1.3	11:49:35.574	+22:23:44.27
1.1.3	11:49:35.453	+22:23:44.82
1.1.3	11:49:35.370	+22:23:43.94
1.1.3	11:49:35.474	+22:23:42.68
1.1.4	11:49:35.158	+22:23:44.16
1.1.5	11:49:35.558	+22:23:46.86
1.1.6	11:49:35.384	+22:23:47.09
1.1.7	11:49:35.307	+22:23:48.19
1.1.8	11:49:35.187	+22:23:46.70
1.1.9	11:49:35.414	+22:23:45.99
1.1.10	11:49:35.479	+22:23:47.63
1.1.11	11:49:35.639	+22:23:45.96
1.1.12	11:49:35.144	+22:23:46.50
1.1.13	11:49:35.349	+22:23:46.37
1.1.14	11:49:35.319	+22:23:42.76
1.1.15	11:49:35.250	+22:23:46.37
1.1.16	11:49:35.298	+22:23:44.58
1.1.17	11:49:35.254	+22:23:44.74
1.1.18	11:49:35.265	+22:23:43.42
1.1.19	11:49:35.272	+22:23:47.90
1.1.20	11:49:35.323	+22:23:47.42
1.1.21	11:49:35.177	+22:23:45.36
1.1.22	11:49:35.111	+22:23:44.51
1.1.23	11:49:35.214	+22:23:47.66
1.1.24	11:49:35.120	+22:23:45.71
1.1.25	11:49:35.498	+22:23:45.73
1.1.26	11:49:35.541	+22:23:45.11
1.1.27	11:49:35.277	+22:23:47.21
1.2.1	11:49:35.858	+22:23:50.81
1.2.2	11:49:35.942	+22:23:51.02
1.5.3	11:49:36.026	+22:23:48.10
1.2.4	11:49:35.883	+22:23:51.95
1.2.5	11:49:35.824	+22:23:48.62
1.2.6	11:49:35.798	+22:23:49.99
1.2.7	11:49:35.705	+22:23:51.48
1.2.8	11:49:35.754	+22:23:52.35
1.2.9	11:49:35.896	+22:23:49.78
1.2.11	11:49:35.872	+22:23:47.73
1.2.12	11:49:35.754	+22:23:52.92
1.2.13	11:49:35.840	+22:23:50.26
1.5.14	11:49:36.076	+22:23:48.93
1.2.15	11:49:35.801	+22:23:51.35
1.2.16	11:49:35.913	+22:23:50.50
1.2.17	11:49:35.889	+22:23:50.94
1.2.19	11:49:35.718	+22:23:51.78
1.2.20	11:49:35.752	+22:23:51.00
1.2.22	11:49:35.842	+22:23:52.54
1.2.23	11:49:35.709	+22:23:52.45
1.2.27	11:49:35.757	+22:23:51.46
1.3.1	11:49:36.820	+22:24:08.77
1.3.2	11:49:36.778	+22:24:07.23
1.3.3	11:49:36.906	+22:24:07.37
1.3.4	11:49:36.711	+22:24:08.03
1.3.5	11:49:36.921	+22:24:09.24
1.3.6	11:49:36.862	+22:24:09.54
1.3.7	11:49:36.809	+22:24:10.34
1.3.8	11:49:36.724	+22:24:09.66
1.3.9	11:49:36.888	+22:24:08.68
1.3.10	11:49:36.899	+22:24:09.84
1.3.11	11:49:36.944	+22:24:08.69
1.3.12	11:49:36.686	+22:24:09.51
1.3.13	11:49:36.850	+22:24:09.14

Table 1
(Continued)

Image ID	R.A.	Decl.
1.3.14	11:49:36.862	+22:24:06.77
1.3.15	11:49:36.784	+22:24:09.33
1.3.16	11:49:36.846	+22:24:08.08
1.3.17	11:49:36.807	+22:24:08.25
1.3.18	11:49:36.823	+22:24:07.30
1.3.19	11:49:36.789	+22:24:10.21
1.3.20	11:49:36.824	+22:24:09.92
1.3.21	11:49:36.730	+22:24:08.84
1.3.22	11:49:36.652	+22:24:08.25
1.3.23	11:49:36.749	+22:24:10.17
1.3.24	11:49:36.668	+22:24:09.04
1.3.25	11:49:36.911	+22:24:08.30
1.3.26	11:49:36.911	+22:24:07.88
1.3.27	11:49:36.796	+22:24:09.81
1.4.4	11:49:35.617	+22:23:55.28
1.4.7	11:49:35.542	+22:23:53.69
1.4.8	11:49:35.464	+22:23:55.65
1.4.8	11:49:35.681	+22:23:53.62
1.4.12	11:49:35.446	+22:23:56.25
1.4.19	11:49:35.501	+22:23:54.33
1.4.22	11:49:35.549	+22:23:56.15
1.4.23	11:49:35.437	+22:23:55.20
1.4.24	11:49:35.494	+22:23:56.28
1.4.27	11:49:35.563	+22:23:54.19
1.4.27	11:49:35.630	+22:23:53.69
1.4.28	11:49:35.597	+22:23:54.32
1.4.28	11:49:35.621	+22:23:54.11
1.5.1	11:49:35.967	+22:23:49.69
1.5.2	11:49:36.031	+22:23:49.93
1.5.9	11:49:35.936	+22:23:48.98
1.5.16	11:49:35.999	+22:23:49.11
1.5.17	11:49:35.990	+22:23:49.56
1.5.18	11:49:36.044	+22:23:49.35
1.5.25	11:49:35.954	+22:23:48.31
1.5.26	11:49:35.991	+22:23:48.17
2.1.1	11:49:36.581	+22:23:23.10
2.2.1	11:49:37.450	+22:23:32.92
2.3.1	11:49:37.579	+22:23:34.39
3.1.1	11:49:33.772	+22:23:59.36
3.1.2	11:49:33.784	+22:23:59.45
3.1.3	11:49:33.825	+22:23:59.50
3.1.4	11:49:33.738	+22:23:59.04
3.1.5	11:49:33.795	+22:23:59.67
3.2.1	11:49:34.282	+22:24:11.73
3.2.2	11:49:34.252	+22:24:11.10
3.2.3	11:49:34.180	+22:24:09.19
3.2.4	11:49:34.326	+22:24:12.73
3.2.5	11:49:34.212	+22:24:10.34
3.3.1	11:49:36.279	+22:24:25.88
3.3.2	11:49:36.311	+22:24:25.86
3.3.3	11:49:36.394	+22:24:25.74
3.3.4	11:49:36.206	+22:24:25.86
3.3.5	11:49:36.339	+22:24:25.92
4.1.1	11:49:34.320	+22:23:48.57
4.2.1	11:49:34.651	+22:24:02.65
4.3.1	11:49:37.001	+22:24:22.06
5.1.1	11:49:35.940	+22:23:35.02
5.2.1	11:49:36.259	+22:23:37.77
6.1.1	11:49:35.930	+22:23:33.16
6.2.1	11:49:36.439	+22:23:37.89
7.1.1	11:49:35.750	+22:23:28.82
7.2.1	11:49:36.821	+22:23:39.37
7.3.1	11:49:37.819	+22:24:04.47
8.1.1	11:49:35.640	+22:23:39.66

Table 1
(Continued)

Image ID	R.A.	Decl.
8.2.1	11:49:35.950	+22:23:42.16
8.3.1	11:49:37.702	+22:24:17.00
9.1.1	11:49:36.890	+22:23:52.03
9.2.1	11:49:36.679	+22:23:47.96
9.3.1	11:49:36.010	+22:23:37.89
10.1.1	11:49:34.001	+22:24:12.56
10.2.1	11:49:33.799	+22:24:09.53
11.1.1	11:49:36.034	+22:23:24.58
11.2.1	11:49:36.965	+22:23:34.42
12.1.1	11:49:37.082	+22:23:12.13
12.2.1	11:49:37.920	+22:23:20.61
12.3.1	11:49:38.177	+22:23:25.47
13.1.1	11:49:38.484	+22:23:19.50
13.2.1	11:49:38.213	+22:23:15.71
13.3.1	11:49:37.495	+22:23:07.33
14.1.1	11:49:38.330	+22:23:15.59
14.2.1	11:49:38.371	+22:23:16.21
15.1.1	11:49:38.388	+22:23:14.08
15.2.1	11:49:38.695	+22:23:18.46
16.1.1	11:49:38.306	+22:23:11.98
16.2.1	11:49:38.899	+22:23:20.60

Notes. The locations of the multiply lensed image systems are shown in Figure 21.

^a ID1.ID2.ID3: ID1 is the image system; ID2 is a multiple image belonging to the image system ID1; and ID3 is an individual feature belonging to the multiple image ID2.

^b Images 1.X.Xs belong to the background spiral galaxy at $z = 1.4888$. Image 1.1.1.3 corresponds to the multiply lensed SN Refdal.

3.2. Fine-tuning the BCG Mass Model

In Figure 2 (upper row), we show the critical curves in the region around the BCG as predicted by lens model produced using WSLAP+. Lensed images within ~ 6.4 kpc radius from the BCG center are triply lensed because of a local perturbation to the lensing associated with the BCG. We can, therefore, delens (i.e., trace the lensed feature back to the source plane with the deflections predicted by the lens model) one set of the multiple images back to the source plane and we can then relens (i.e., re-trace the source to the image plane with the deflections predicted by the lens model) them back to the image plane to reproduce the other two sets of images in this region. The detailed features of the images that are thus reproduced serve as local constraints to the BCG mass distribution. Figure 2 shows that although the matching between the actual images and those reproduced through delensing and then relensing is reasonably good, there is clearly room for improvement by locally adjusting the BCG mass distribution. Another point to notice is that L1 is predicted to be straight in the WSLAP+ solution. The same straight appearance for L1 is found in the cluster lens model produced by Grillo et al. (2016) using the parametric algorithm GLEE (Suyu et al. 2012), which will be discussed in more detail Section 4.2. In Section 6, we show that the cluster lens model that we independently produce using *Lenstool* (Jullo et al. 2007) also predicts L1 to be straight.

To improve the mass model of the BCG, we tried three different profiles for its projected two-dimensional mass distribution (while keeping the rest of the WSLAP+ solution fixed, which is necessary to preserve the good global agreement obtained for all the multiply lensed images with this solution).

The first profile simply assumes that the projected two-dimensional mass distribution of the BCG follows its projected two-dimensional light distribution, with the only parameter being the normalization factor (i.e., M/L ratio). The second profile assumes a dual pseudoisothermal elliptical mass distribution (dPIE) (Elfasdóttir et al. 2007) with the convergence given by

$$\kappa = \frac{\theta_E}{2} \left(\frac{1}{\sqrt{R_\epsilon^2 + r_c^2}} - \frac{1}{\sqrt{R_\epsilon^2 + r_t^2}} \right), \quad (4)$$

where r_c is the core radius and r_t is the truncation radius. The third profile assumes a pseudoisothermal elliptical mass distributions (PIEMD) (Kassiola & Kovner 1993), with the convergence given by

$$\kappa = \frac{\theta_E}{2\sqrt{R_\epsilon^2 + r_c^2}}, \quad (5)$$

where r_c is the core radius. In both dPIE and PIEMD profiles, R_ϵ is defined as

$$R_\epsilon = \frac{x^2}{(1 + \epsilon)^2} + \frac{y^2}{(1 - \epsilon)^2}, \quad (6)$$

where ϵ is the ellipticity of the profile, and the position angle is fixed to be at the observed position angle of the BCG's stellar component.

In Figure 2 (lower row), we show the relensed reproductions from three different models for the BCG mass distribution as described above. In both the dPIE and PIEMD profiles, we began with an initial zero ϵ . For the dPIE profile, we used $r_c = 0''.24$, which corresponds to the core radius in the light profile of the BCG, and $r_t = 1''.5$, which is close to the effective radius of the BCG. For the PIEMD profile, we also used $r_c = 0''.24$. We found that by changing the parameters (ϵ , r_c and r_t) in the dPIE and PIEMD mass profiles to improve the agreement between the predicted and observed relensed images, the mass distribution approached the two-dimensional light distribution of the BCG. Hence, in all of the following steps, we use the lens model where the BCG mass distribution is represented by its two-dimensional light distribution. With the best-fit normalization factor, the mass contributed from the BCG scaled with the two-dimensional light distribution is $6.3 \times 10^{11} M_\odot$ within a cylinder of $r < 30$ kpc, while the total projected mass within the same region is $7.1 \times 10^{12} M_\odot$.

At this point, we would like to emphasize that L1 is predicted to be straight, both prior to and after improving the mass distribution of the BCG. This result reflects the fact that, to curve L1 on a radius of curvature of $\sim 0''.6$, a compact local deflector is needed; as described in Sections 5 and 6.

4. Detailed Investigation of L1

4.1. Spectroscopic Analysis Using MUSE Data

To ensure that L1 is a single contiguous feature and that its lensed counterparts have been correctly identified, we extracted the [O II] $\lambda\lambda$ 3726, 3729 doublet emission of the background lensed spiral galaxy using simple spectral fitting. This doublet is the only emission line feature expected for star-forming galaxies at $z = 1.4888$ (Grillo et al. 2016) within the MUSE spectral coverage. We modeled the [O II] doublet as two Gaussian functions having the same line width at the same

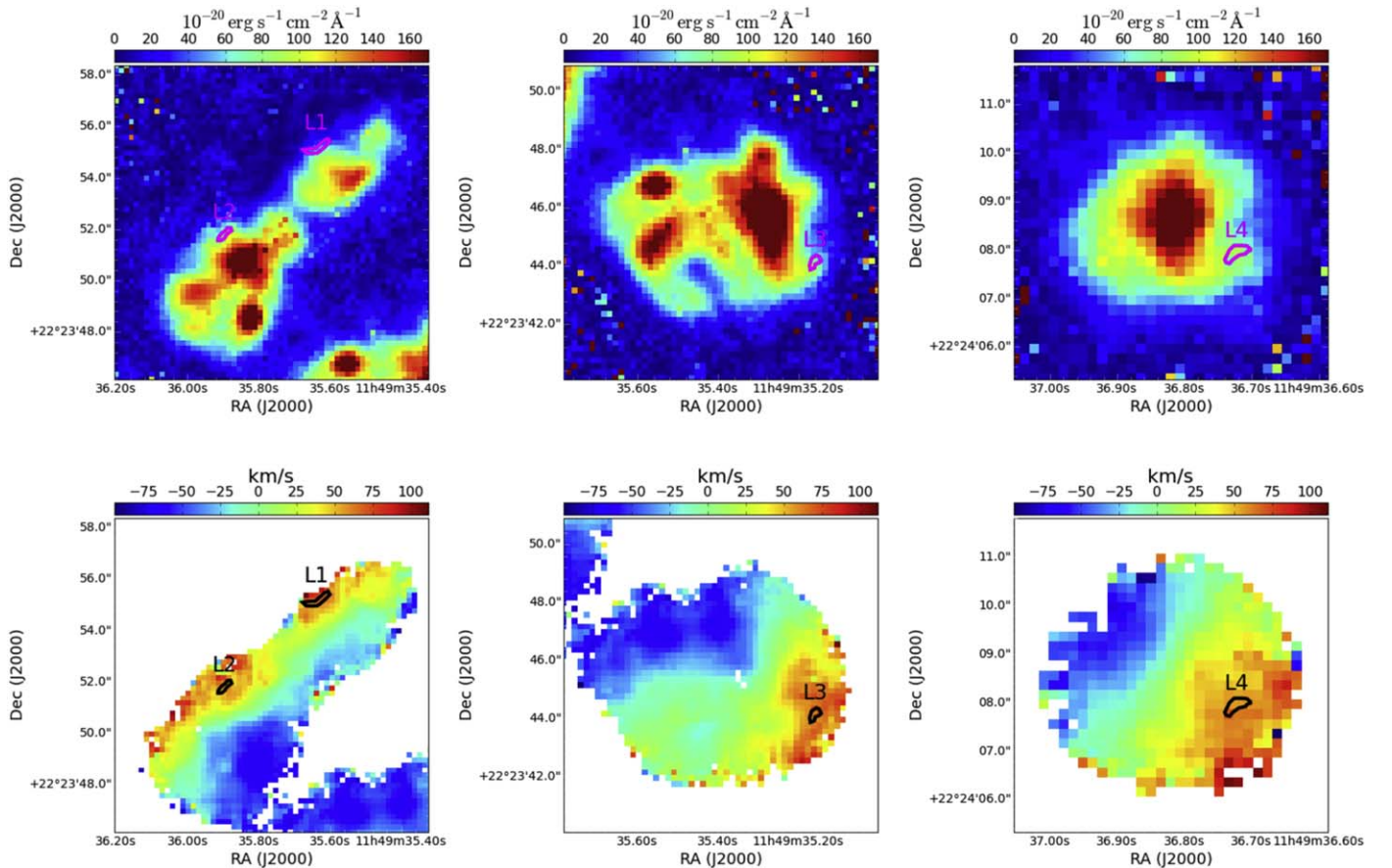


Figure 3. [O II] flux map and velocity map at the rest-frame of the spiral galaxy derived from the [O II] line. Upper row: the isophotal contours indicating the positions of L1-4 are plotted on the [O II] flux map. Lower row: L1-4 belong to the same kinematic region of the background spiral galaxy, confirming the multiple image identifications.

redshift but with different amplitudes, and we fitted these functions to the spectra at $9200\text{--}9300\text{ \AA}$ (where sky emission lines are weak) spanning the [O II] doublet, where sky emission lines are weak. Stellar absorption lines are seen around L1, where stellar light from the BCG makes the spectrum relatively bright. We extracted the spectrum of the BCG at its northeastern region where no contamination by [O II] from the background spiral galaxy is evident and we then used it as a template for the BCG stellar spectrum in the fit. In the fitting, we summed spectra over 3×3 pixels ($0''.6 \times 0''.6$) to improve the S/N without losing the seeing-limited ($1''.0$ at FWHM) spatial resolution. The free parameters of the fit are redshift, line width, flux ratio of the doublet, and scaling factor of the BCG stellar light. We utilized a nonlinear least squares curve fitting library MPFIT (Markwardt 2009) for the fit. From the results of the fit, we generated the velocity field map for Sp1-4 as shown in Figure 3, the detailed local spectral profile for L1 as shown in Figure 4, and the one-dimensional spectral profiles for L1-4 as shown in Figure 5. L1-4 show similar spectral profiles and correspond to the same kinematic region in the lensed background spiral galaxy, which is in agreement with our multiple image identification. The consistent spectral profiles of L1 taken from the three apertures shown in Figure 4 (left-hand panel) confirm that the entire arc of L1 belongs to the same kinematic region in the lensed background spiral galaxy, as our model predicts.

4.2. Examination of the Curvature of L1 Using All Multiply Lensed Images

In this section we will examine the need for a local deflector to produce the curvature of L1 by delensing each of the four spiral galaxy images (Sp1-4) back to the source plane and then making relensed projections to the image plane (see Figure 1 for the locations of Sp1-4 in the cluster). This also allows us to determine the level of agreement between these independently produced images. This agreement check is useful for establishing confidence in the overall accuracy of our cluster lens model. We are particularly interested in the predicted appearance of L1 when delensed and then relensed from its different counterparts (L2-L4) and, therefore, we investigate the level of agreement for this feature between the different lensed images.

The lens model on the large scale is generally very accurate at the level of 3% in terms of the deflection angle, which can be appreciated by delensing any lensed image and relensing it to form the counter images that are well separated in angle. In Figure 6, we show the agreement between our WSLAP+ free-form model and the parametric model of Grillo et al. (2016), both of which successfully predicted the position and time of the reappearance of SN Refsdal in one of the lensed images of the background spiral galaxy. In Figure 6 (lower row), we predict the large-scale distribution of the lensed images of the spiral galaxy by delensing and then relensing, alternatively, the lensed images labeled “Input image (Sp1)” and “Input

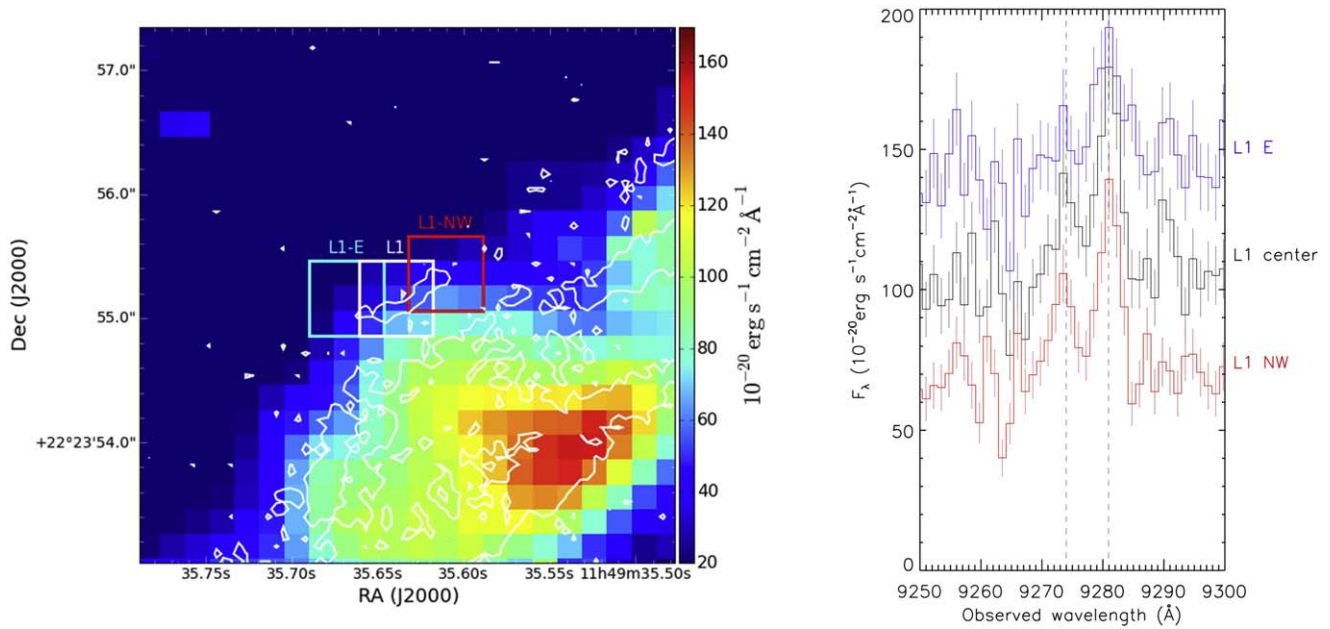


Figure 4. Spectral profile of L1 derived from the [O II] line. Left: the spectral aperture on top of the [O II] flux map. White contours are the 0.0015 electron/s isophotal contours from the F435W HFF data image. Right: the one-dimensional spectral profiles taken from the three apertures indicated in the left-hand panel. The vertical dashed lines are the rest-frame [O II] doublet lines. The agreement of the spectral profiles from the three apertures spanning the entire L1 indicates that the whole arc of L1 belongs to the background spiral galaxy.

image (Sp2).” The level of agreement that we find is comparable with the general level of relensing accuracy that has been determined for this cluster by other independent work (Zitrin & Broadhurst 2009; Diego et al. 2016; Treu et al. 2016), and also for other complex clusters with deep Hubble data (Broadhurst et al. 2005; Halkola et al. 2008). Figure 6 deliberately follows the published image format from Grillo et al. (2016) to make this comparison as clear as possible. Very good agreement is seen between our free-form method and their parametric approach to lens modeling.

When delensing and then relensing Sp1-4, the agreement between counter images is better for more closely separated images, as expected. A small systematic shift of $1''.1$ and $0''.3$ were applied to the more distant counter-images Sp1 and Sp2, and less than $0''.1$ to Sp3 so that they become better aligned. These shifts show an accuracy limitation at the large scale. The corresponding adjustments corrects the systematic offset between the predictions based on different relensed images and the data in the central region of the lensed field, thus allowing us to make a detailed comparison of the internal features of the spiral galaxy between each relensed images and the region of interest around the BCG. We emphasize at this point that these small systematic errors reflect the accuracy of the cluster-scale lens model given the constraints available from all the multiply lensed background galaxies, and that no modification to this cluster model can produce the curved appearance of L1 with a radius of curvature of $\sim 0''.6$. The large and smooth deflection angles that the cluster-scale lensing induces are $\sim 30''$ in size, and any modification of the cluster deflection field can, therefore, only shift a small image uniformly without bending it.

Figure 7 shows delensed and then relensed images in the region around the BCG by delensing and then relensing, alternatively, the lensed images Sp1-4. Note that when relensing Sp4, we only include the part of Sp4 that is indicated by the blue-highlighted region (lower right-hand panel) and,

therefore, only the portion of L1 to the right of the critical curve, to compare its delensed and relensed image to its counter image on the opposite side of the critical curve. We use the best-fit solution as described earlier comprising the grid to represent the large-scale distribution of dark matter, the 2D light map of the BCG to account for its mass, and a 2D NFW mass distribution associated with each identified member galaxy. As can be seen, all the bright features seen in the data appear in each of the four delensed and then relensed images. To form a deeper delensed and then relensed image, we added together all of the individual relensed images, as shown in Figure 7 (lower left-hand panel). This deeper image is now noticeably blurred with respect to the individual lensed images and with respect to the data, reflecting the small residual level of difference between the lens mapping and the four independent source images. The purpose of this additional image is to demonstrate that these effects are small, being not much larger than the angular resolution of the data. This does not affect significantly our subsequent conclusions regarding the shape of L1 and the neighboring features. A clear conclusion from this comparison is that L1 is straight as predicted by delensing and then relensing all four counter images, in contrast with the actual appearance of L1. Also note that the neighboring internal details are all well-reproduced by the delensed and then relensed images with little evidence of any systematic difference in either shape or orientation. As can be seen in Figure 6 (upper row), the straight appearance predicted for L1 when delensing and then relensing its counter images is also apparent in the independent lens model presented by Grillo et al. (2016).

4.3. Shear Effect at the Position of L1

We now demonstrate the shear effect at the position of L1, and contrast it with the shear effect at the positions of L3 and L4. As we will elaborate, this shear difference implies that the

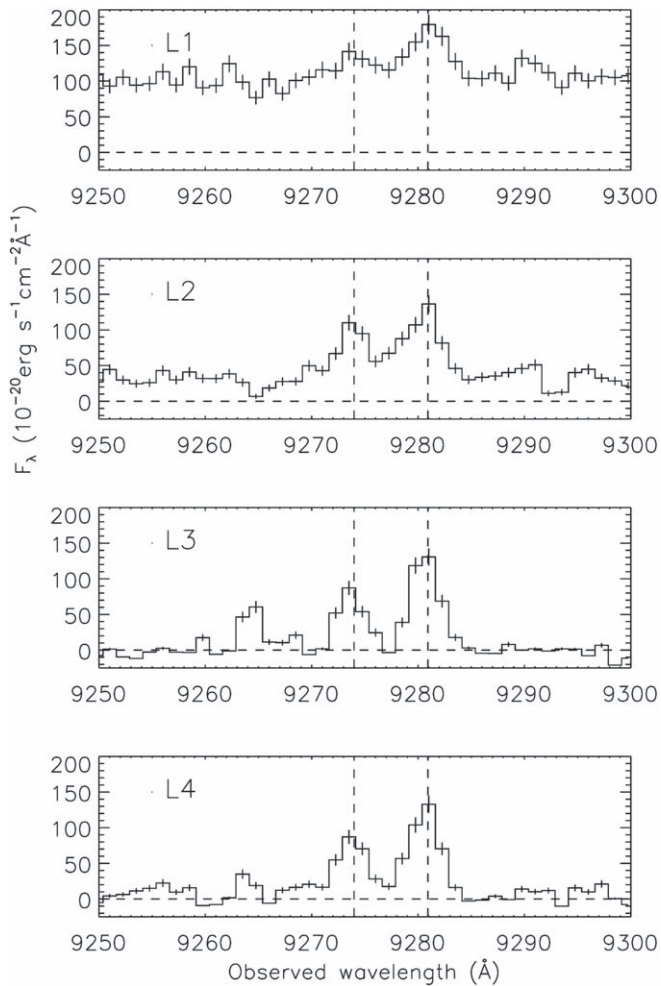


Figure 5. Spectral profile of L1-4 derived from the [O II] line. The vertical dashed lines are the rest-frame [O II] doublet lines. The line widths at the position of L1 and L2 are bigger than that at the position of L3 and L4 due to the large shear induced by lensing.

bending of L1 cannot be accounted for by intrinsic substructures in the source image.

As shown in Figure 8, we first delensed L3 (contained in Sp1) and L4 (contained in Sp2) to the source plane and we then put a color-coded circle at the position of the delensed L3 and L4. We relensed this circle back to the image plane, hence visualizing the shear magnitude and orientation of L3 and L4 at the locations of their counter images. As is clearly demonstrated in Figure 8, the magnitude of shear at position of L1 is much bigger than the shear at the positions of L3 and L4, making L1 more stretched than L3 and L4. This shear difference can be understood from the fact that the critical curve passes through L1, indicating the gravitational potential gradient is much higher in the vicinity of L1 compared to positions further away from critical curves such as L3 and L4. Therefore, although L1 has intrinsic substructures as observed in L3 and L4, the image formed at the position of L1 is guaranteed to possess a linear shape owing to the high magnitude shear at that particular position.

An important point to note here is that the small radius of curvature of L1 is only $\sim 0''.6$. The bending on such a small scale cannot be produced by BCG lenses (which have typical

Einstein radii of $\sim 5''$) and cluster lenses (which have Einstein radii of $\sim 30''$ and are responsible for bending the giant arcs). Although an accurate large-scale cluster lens model is essential to account for the lensed images of the spiral galaxy as a whole, the bending of L1 is a local effect that is beyond the influence of the combined lensing from the BCG and the cluster. This fact makes the exploration of the bending imposed on L1 independent of the large-scale lens model (given sufficient accuracy of the cluster lens model as we have demonstrated in the previous sections), as confirmed with a different parametric lens model constructed by *Lenstool* that we will describe in Section 6. Furthermore, whatever is responsible for this image curvature should have a deflection field that drops quickly with distance, otherwise the neighboring images will be noticeably deflected while failing to bend L1 on the small scale observed, as we will show in Section 5.3. In the next section, we will show that a point mass—an SMBH—uniquely satisfies these constraints and satisfactorily reproduces the degree of bending seen for L1.

5. Adding a Local Deflector

5.1. A Point Mass

As we will now show, the curvature of L1 can be reproduced by adding a point mass near the center of the BCG to the lens model.

We start by adding a point mass *at the center* of the BCG light distribution, then delensing and relensing each of the Sp1-4 images to obtain the model prediction for L1 as well as its nearby images, similar to the process adopted in Figure 7. As before, when relensing Sp4, we only include the part of Sp4 that is indicated by the blue-highlighted region (lower right-hand panel of Figure 7) and, therefore, only the portion of L1 to the right of the critical curve. The effect of adding this central point mass is shown in Figure 9. As can be seen, L1 is predicted to stretch toward the point mass and, therefore, to the center of the BCG, which is unlike that actually seen in the data. Thus, a point mass at the center of the BCG light cannot reproduce the curvature of L1 at the location observed.

Next, we consider an offset point mass. The position and mass of this point mass have been explored as three free parameters to obtain the best fit; details of the parameter optimization will be presented in the next section. Note that after adding a point mass either at the light centroid of the BCG or at an offset position close in projection to L1, we decrease the BCG mass by a certain amount to counteract the systematic effect of this extra point mass on lensed images close to the center of BCG. Further away from the center, we find that these modifications to the BCG mass distribution do not noticeably influence the relensed images. The relensed images from different multiple images for this case are shown in Figure 10.

Finally, we investigate what constraints we can place on a point mass at the BCG light centroid in addition to the offset point mass responsible for producing the curvature of L1. As shown in Figure 11, the presence of a central point mass with a mass equal to the offset point mass does not obviously bring about any improvement between the actual lensed images and their delensed-relensed counterparts for either L1 or other neighboring lensed features. We conclude that the present lensing data neither favors nor excludes the possibility of having a dual point mass system.

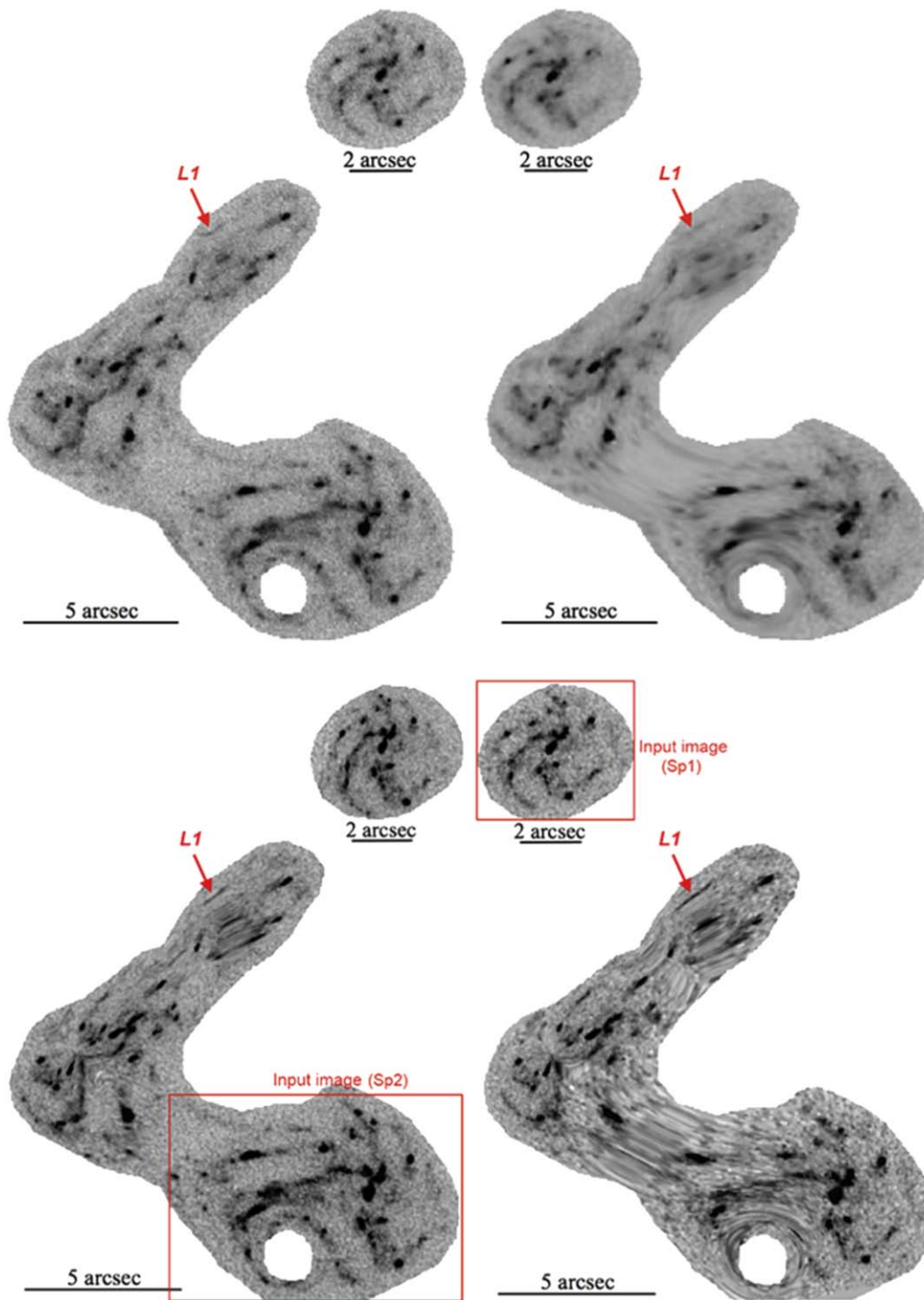


Figure 6. Upper row: Figure 7 of Grillo et al. (2016). The left-hand side is the original data but with cluster members subtracted to better reveal the multiply lensed spiral galaxy. Notice that L1 appears curved in the subtracted data, similar to its appearance in our subtracted image. The image on the right-hand side is the model-generated image from Grillo et al. (2016), showing good general agreement with the data on the left. Notice that L1 is predicted to be straight in their model, as in our work when we do not impose an extra deflector in the BCG. Lower row: predicted released images of the spiral galaxy obtained by relensing Sp2 in the left-hand panel and Sp1 in the right-hand panel, utilizing our lens model from WSLAP+ solution and the 2D light distribution of the BCG for its mass profile. In each case, there is good general agreement with the data, except that L1 is straight rather than being curved as is observed.

In Figures 9–11, all the bright lensed features seen in the data appear in each of the four delensed and then relensed images, and are similar with each other in each model. This good agreement clearly demonstrates that an offset point mass can plausibly reproduce the bending of L1, while a point mass centered at the BCG centroid can

only elongate the image L1 while not bending it enough. A double point mass solution makes no obvious difference to the relensed images, hence no conclusion can be made concerning a second point mass at the BCG centroid. We present the different models described above in Figure 12.

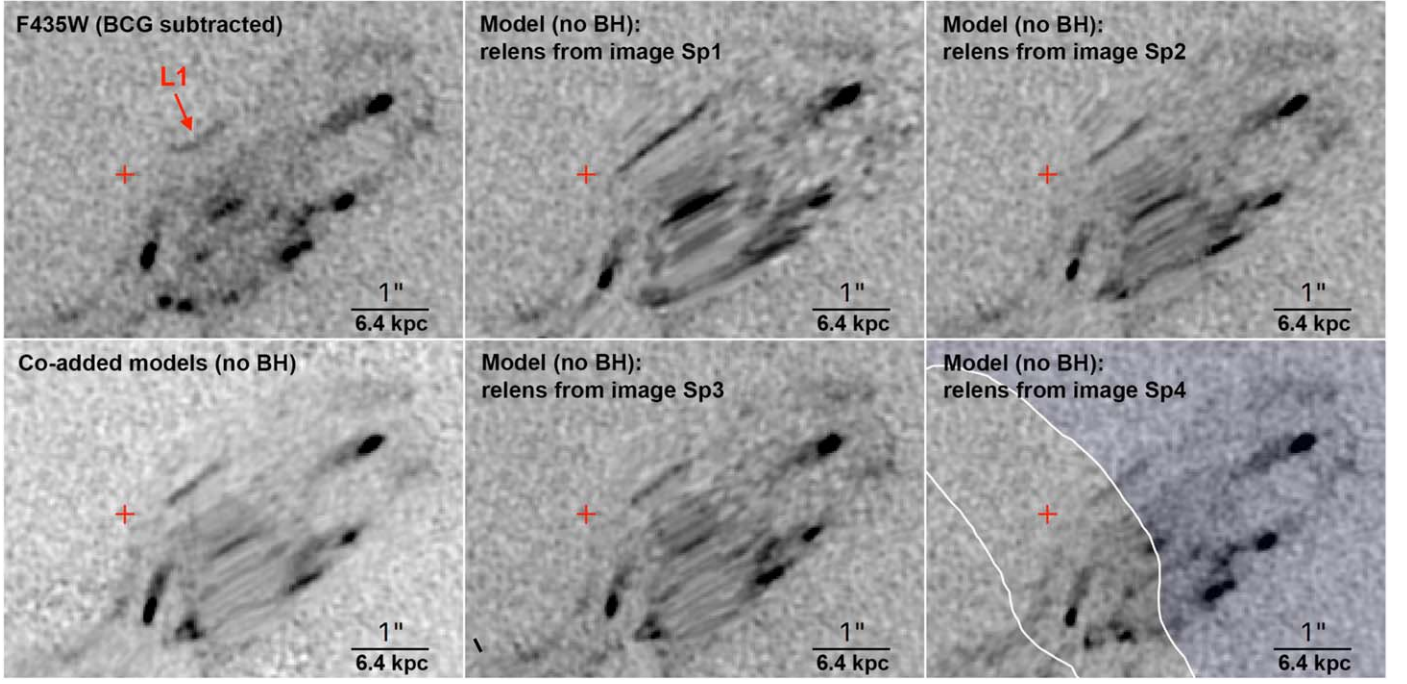


Figure 7. Detailed comparison of lensed features around the BCG based on re-lensing of four independent lensed images of the background spiral galaxy (Sp1-4, see Figure 1 for the locations of these images in the cluster). Upper left: image in the F435W band. The re-lensed images in the remaining panels all assume the same BCG mass distribution, based on its 2D light map as described in Section 3.2. The light centroid is marked with a red cross for reference in all the panels. Note that in the lower right-hand panel, only the blue-highlighted region on the right-hand side is used as the input for re-lensing. The consistency between the re-lensed images and their actual appearance is good considering the differences in location and magnifications of each of the input data images for re-lensing. The obvious exception is the curvature observed in L1, which was not predicted in this model. This underscores the need for additional deflection local to L1.

5.2. Statistical Uncertainties of SMBH Parameters

We will now describe the statistical constraints that are placed on the parameters of the additional point mass, which is presumably an SMBH. Figure 13 demonstrates that the shape of the re-lensed L1 image is sensitive to both the position and mass of this SMBH. The models with no SMBHs predict a relatively straight L1 (position (1,1) in Figure 13). Placing the SMBH at the light center (first row with $0''$ offset) predicts images that are too long and straight. Therefore, we can constrain the parameters of the SMBH by comparing the actual image of L1 with its delensed and then re-lensed counterparts as predicted by the lens model. Its position and mass (X_{BH} , Y_{BH} and M_{BH}), as well as the total mass of the BCG (M_{BCG}), are constrained simultaneously through Markov Chain Monte Carlo (MCMC) sampling. The posterior probability distribution of the parameters ζ is sampled given the observed data d :

$$p(\zeta|d) \propto \mathcal{L}_{\text{pos}}(d|\zeta) \mathcal{L}_{L1_angle}(d|\zeta) \mathcal{L}_{L1_flux}(d|\zeta). \quad (7)$$

\mathcal{L}_{pos} is the likelihood function for the positions predicted for three bright and compact features close to the BCG center enclosed in the green circles of Figure 2, which is defined as

$$\mathcal{L}_{\text{pos}}(d|\zeta) \propto \exp\left[-\frac{1}{2} \sum_{i=1}^3 \frac{(x_{\text{pred}}^i - x_{\text{obs}}^i)^2}{0.5^2}\right], \quad (8)$$

where we take the uncertainty of position measurements in the data as 0.5 pixel ($0''.015$). The position prediction of these three major features strongly constrains the total mass of the BCG. \mathcal{L}_{L1_angle} is the likelihood function for the predicted angle of L1 as defined in the manner described below. In each model with certain parameter values, we predict the shape of L1 through

the delensing and re-lensing process as demonstrated in Figure 12. According to where the critical curve passes through L1 in each model, we divide the predicted L1 into two parts and we then fit a straight line to each part to obtain the angle between the two lines, θ_{L1_pred} . We use the same cutting point to divide L1 in the actual image and we obtain θ_{L1_obs} accordingly. Then, we define \mathcal{L}_{L1_angle} as

$$\mathcal{L}_{L1_angle} \propto \exp\left[-\frac{1}{2} \frac{(\theta_{L1_pred} - \theta_{L1_obs})^2}{\sigma_{\theta_L1_obs}^2}\right], \quad (9)$$

where the uncertainty $\sigma_{\theta_L1_obs}$ is propagated from the slope uncertainties of the two fitted lines to L1 data. In addition to the bending angle of L1 thus defined, we also measure the average flux density in the predicted region of L1 to ensure that the surface brightness is correctly reproduced in the best-fit models. Thus, we added a third likelihood function \mathcal{L}_{L1_flux} which is defined as

$$\mathcal{L}_{L1_flux} \propto \exp\left[-\frac{1}{2} \frac{(\bar{J}_{\text{pred}} - \bar{J}_{\text{obs}})^2}{\sigma_{\bar{J}_{\text{obs}}}^2}\right], \quad (10)$$

by taking all the pixels that are re-lensed into the calculation.

The result of the MCMC sampling is shown in Figure 14. We obtained a mass of $M_{\text{BH}} = 8.4_{-1.8}^{+4.3} \times 10^9 M_{\odot}$ with an offset from the BCG center of $\Delta_{\text{BH}} = 4.4 \pm 0.3$ kpc. Note that this offset position is beyond the BCG stellar light core that we derive from a standard Nuker fitting using GALFIT (Peng et al. 2010), for which we obtain a break radius of 1.5 kpc, as shown in Figure 15. There is an observed degeneracy between the SMBH mass and its offset with respect to the BCG center.

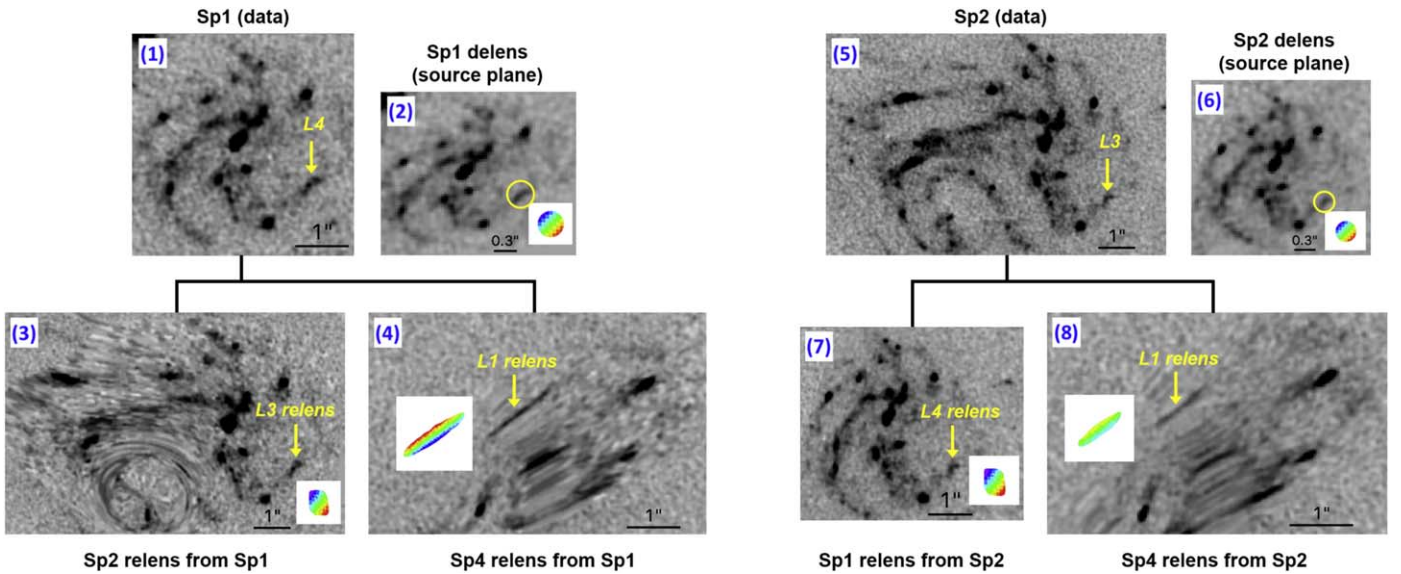


Figure 8. Illustration of shear effect on positions of L1, L3 and L4. (1): data zooming into the Sp1 region. (2): delensed Sp1 in the source plane. A color-coded circle as shown in the insert is placed at the position of delensed L4 in the source plane (yellow circle). (3) and (4): the relensed images of Sp1 at Sp2 and Sp4 region, respectively. The projection of the color-coded circle from source plane to the image plane at position of L3 and L1 are shown in the inserts. (5)–(8) are same as (1)–(4), but with Sp2 as the input data for delensing and relensing. This figure demonstrates that the magnitude of shear at position of L1 is much bigger than at the positions of L3 and L4, making L1 more stretched than L3 and L4.

This degeneracy is due to the fact that within a small region enclosing the optimal position for the offset SMBH, when the offset SMBH is closer to the BCG, it is further away from L1; hence, it requires a bigger mass to cause the same degree of change on the local shear of L1.

5.3. Adding an Extended Deflector

Instead of an offset SMBH, we will now examine whether a spatially extended substructure can account for the bending of L1; and if so, then we determine its required parameters. For this purpose, we replace the offset SMBH with a singular isothermal halo that is centered at the same position. The general effect of this can be seen in Figure 16 (upper row). Note that after adding an extended halo, the BCG’s mass needs to be reduced accordingly to retain a relative satisfactory fit for all the nearby lensed features close to the BCG (which are also visible in Figure 16). We find that although an extended halo has the effect of bending L1 in the manner desired, it also significantly influences the neighboring details of the lensed spiral galaxy in a noticeably undesirable way. Specifically, the more massive this extended halo and, therefore, the greater the reduction required for the BCG mass, the worse the match between model predictions and data for the other neighboring lenses images becomes. This effect can be appreciated through the relative error in model residuals obtained by $(\text{model}-\text{data})/\text{data}$. In the relensed region enclosed by the critical curve shown in Figure 16 (upper middle panel), the $(\text{model}-\text{data})/\text{data}$ has a standard deviation of 12.3, and a median of -0.18 . In comparison, in the relensed region that is enclosed by the critical curve shown in Figure 12 (central middle panel), the $(\text{model}-\text{data})/\text{data}$ has a standard deviation of 7.4 and a median of -0.11 . Furthermore, it can be observed that three major features indicated by green circles in Figure 16 are poorly reproduced after an extended halo is added to the model.

We also tried halos that have different mass profiles. A NFW halo has a shallower central profile than a singular isothermal

halo so that the deflection angle increases with the projected angle out to several Einstein radii, making matters even worse; as shown in Figure 16 (lower row). Later on in Section 6, we will show that a simultaneous fitting of an extended halo along with the BCG cannot produce a solution more plausible than (i.e., inferior to that of) a point-mass solution, as described above.

In this context, we can also ask whether the luminous stars expected to be associated with any extended deflector should be visible in contrast against the BCG light. We adopt a M/L of $10(M/L)_{\odot}$, and convert the total luminosity L to luminosity density in I band L_I according to the bolometric correction estimated in Buzzoni (2005). For the extended halo, we assume an elliptical galaxy having an age of 4.0 Gyr (note that the look back time to the cluster is 5.4 Gyr) and we scale the luminosity density as $L_I \simeq 0.15L$ (Buzzoni 2005). Requiring the light distribution to follow the two-dimensional mass distribution of the extended deflector, we then convert L_I to signals in *HST* F140W data with respect to the CCD inverse sensitivity (i.e., the PHOTFLAM keyword in data header). We convolve this light map with a point-spread function (PSF) that is adopted from an isolated star (at R.A. = 11:49:32.697, decl. = +22:24:08.61) in the HFF F140W data. Finally, we add this converted light map to the F140W data, as shown in Figure 17 (upper row). We expect to clearly see associated stars. Increasing M/L reduces the associated starlight contrast but must reach a large value of $M/L > 50(M/L)_{\odot}$ to be lost in the contrast against the BCG, as shown in Figure 17 (lower row). Although ultra-faint dwarfs can have M/L as large as $\sim 1000 (M/L)_{\odot}$, these objects have very low masses (with equivalent velocity dispersion $< 10 \text{ km s}^{-1}$) (Simon & Geha 2007). We know of no objects having the required equivalent velocity dispersion of $> 100 \text{ km s}^{-1}$ and such a large M/L . The arguments presented above argue against an extended halo rather than a black hole to explain the appearance of L1.

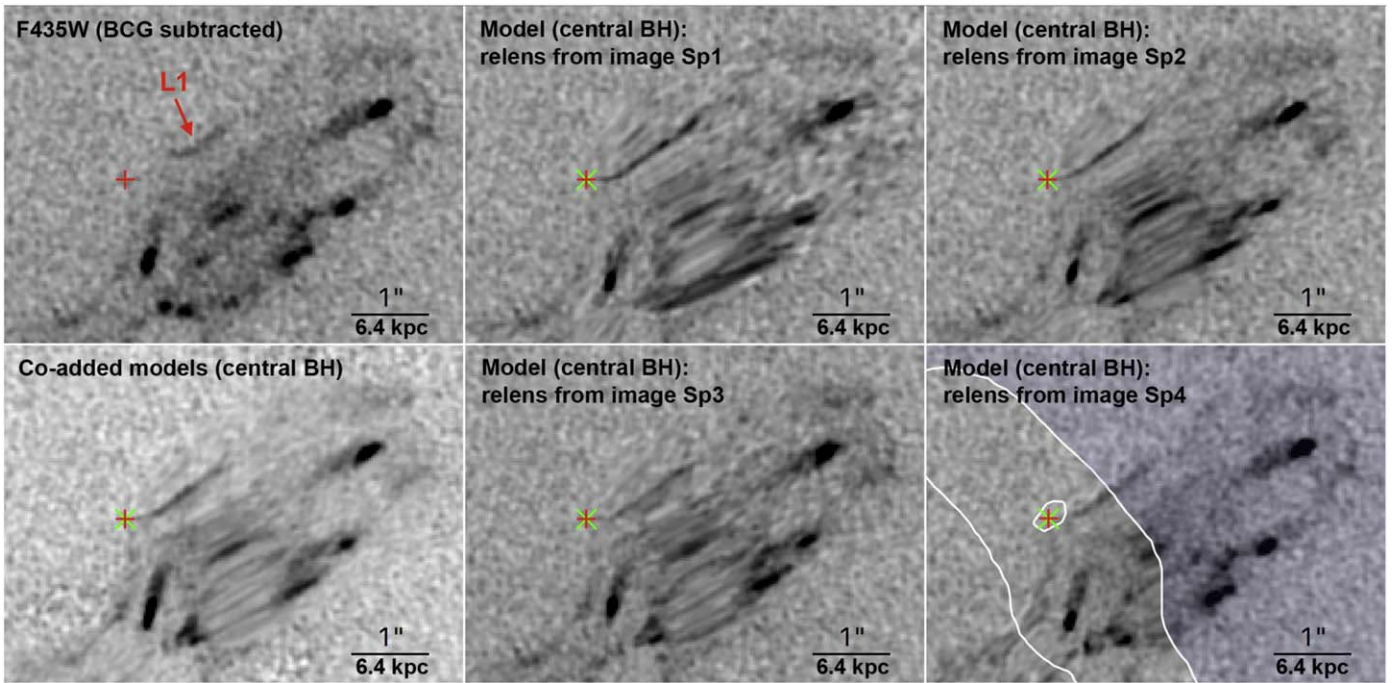


Figure 9. Same as Figure 7 but with the addition of a massive central black hole (green cross) at the location of the light centroid (red cross). The black hole is modeled as a point mass of $8 \times 10^9 M_{\odot}$. The independent predictions and their co-addition clearly shows that adding such a point mass significantly elongates the predicted length of L1 toward the light center, quite unlike in the actual data where L1 is bent away from the center.

6. *Lenstool* Model of MACS 1149

As an independent check of the cluster-scale lens model derived using our non-parametric, grid based method (WSLAP+), we employ the parametric lens modeling package *Lenstool* (Jullo et al. 2007) to construct an independent lens model for MACS 1149. As we will show, this package also has the advantage of permitting simultaneous fitting of the BCG and a local deflector required to bend L1. We begin by allowing *Lenstool* to freely describe all three main sources of deflection—namely the cluster, the BCG and the member galaxies—to find a best fit to the full set of detected multiple images, including the internal substructures of the well resolved spiral galaxy at $z = 1.49$. Three PIEMD halos (see Equation (5)), which are found through trial and error, are used to represent the cluster-scale mass distribution. The BCG together with three other galaxies that are close to the lensed spiral galaxy images, as pointed out in Figure 1, are each described by a single PIEMD halo. The rest of the cluster member galaxies are modeled by PIEMD halos that are scaled with respect to their luminosities, as described in Jullo et al. (2007). In doing so, we only restrict the BCG to be centered at its observed light centroid, so that *Lenstool* is free to obtain the BCG density profile, core radius, ellipticity and position angle. The model cluster halo has the additional freedom to be centered without restriction and with any additional substructures according to the *Lenstool* prescription. We used the same set of multiple image constraints from the WSLAP+ modeling, as listed in Table 1. The lens model that we obtained from *Lenstool* is very similar to the WSLAP+ model that we use for all the previous analyzes, boosting the confidence in our results (as presented above).

The best-fit model from *Lenstool* has a mean rms dispersion between the predicted and observed centroids for each set of multiply lensed images (which are used to constrain the lens

model) as measured in the image plane of $\langle \text{rms}_i \rangle = 0''.25$. Compared with WSLAP+ model, this *Lenstool* model has a higher large-scale accuracy toward the outer region of the cluster. We show this *Lenstool* solution in Figure 1 (4th row) and also Figure 18. The corresponding de-lensed source plane images of Sp1-4 from *Lenstool* (Figure 1, 4th row) show a very high degree of agreement in shape and position, exceeding the accuracy of our WSLAP+ solution. Most importantly for our conclusions regarding the curved image, L1, the *Lenstool* critical curves around the BCG are very similar to our previous WSLAP+ solution, passing through the center of the curved L1 image. The similar critical curves at the position of L1 found using either WSLAP+ or *Lenstool* indicates that our conclusion does not depend strongly on the assumed mass profile of the BCG. As shown in Figure 18, the released image of L1 is quite straight just like our previous solution, and is highly sheared because it straddles the critical curve; as discussed in detail in Section 4.3. This critical curve also passes through the center of the other nearby highly sheared images and close pairs that evidently straddle this critical curve, which are indicated by the encircled features in Figure 18. These pairs provide a convenient measure of the accuracy of the lens model in the region around the BCG. Without an addition of a local deflector to bend L1, the mean rms dispersion between the predicted and observed centroids for each set of these multiply lensed images (i.e., the circles in Figure 18) as measured in the image plane is $\langle \text{rms}_i \rangle = 0''.2$.

6.1. Adding a Local Point Mass

Next, we require *Lenstool* to find a best-fit solution with a point mass added to produce the observed curvature of L1. Because the goal here is to reproduce L1 based on a good cluster lens model that we have already obtained, we restrict the model constraints to the images in the vicinity of

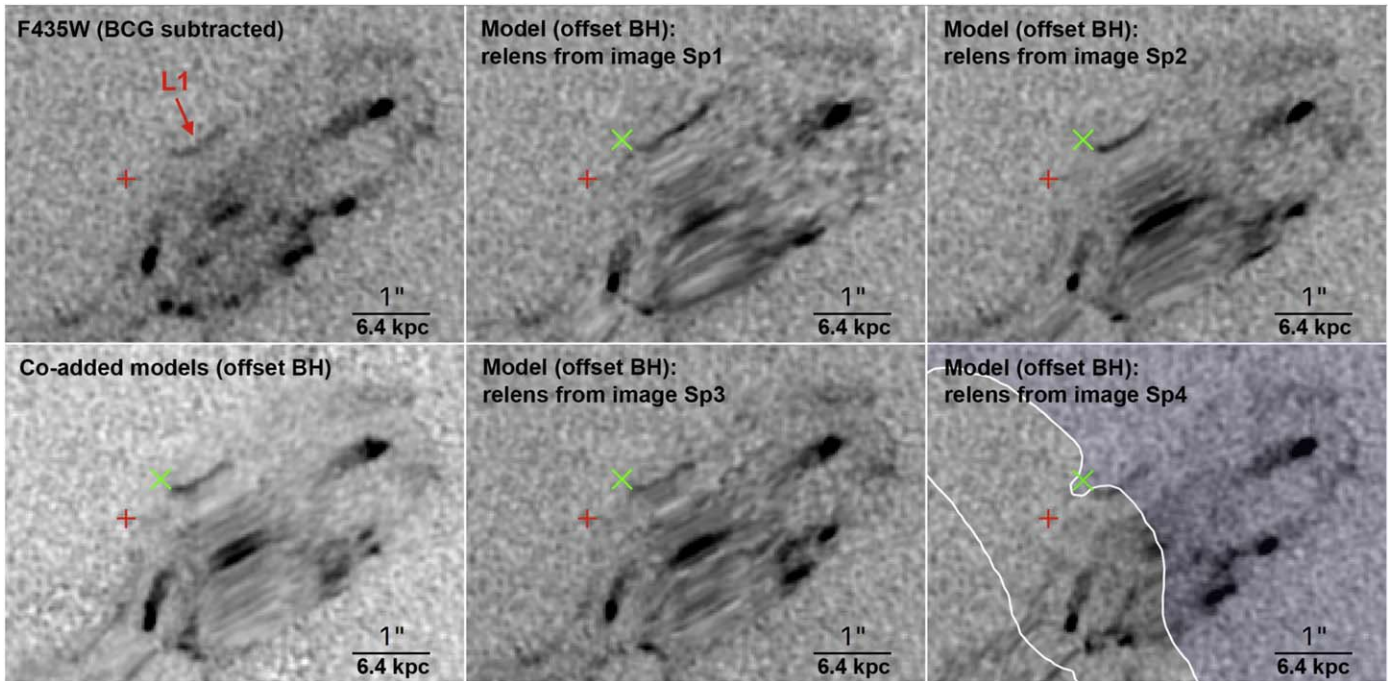


Figure 10. Same as Figure 9 but with the addition of a point mass of $8 \times 10^9 M_{\odot}$ at the position of the green cross. Such a point mass introduces a significant curvature in L1, consistent with that seen in the data, without affecting the good fit already obtained for the other images. Note that the BCG mass is scaled down by $\sim 6\%$ in this combined best fit, as described in the Supplementary Methods.

the BCG, including L1, while excluding images from different background galaxies and images far away from the BCG center. In turn, we fix the three cluster PIEMD halos because they are not well constrained by the restricted set of image constraints. The BCG’s PIEMD halo position is fixed to be at its observed light centroid with all other parameters free to vary. We also fix the PIEMD parameters for the three other galaxy halos except for their velocity dispersions (i.e., total mass). The point mass added is allowed to lie within a $1'' \times 1''$ square centered on the centroid of the BCG, with a wide uniform prior on its mass, relevant for SMBHs. We use “forme = -3” in the *Lenstool* parameter file so that the minor and major axes of L1 are included as constraints. The optimization of the aforementioned free parameters is performed in the image plane.

The best-fit *Lenstool* solution found this way is centered near the radius of curvature of L1, with a mass whose Einstein radius is in correspondence with the approximate curvature radius of L1. The effect on the shape of the critical curve is the same as the one that we obtained previously with WSLAP+, forming an indentation near L1; as shown in Figure 18. *Lenstool* constrains this point mass offset from the BCG as $\Delta_{\text{BH}} = 4.2^{+0.7}_{-0.5}$ kpc with a higher best-fit mass of $M_{\text{BH}} = 1.25^{+0.14}_{-0.27} \times 10^{10} M_{\odot}$. These parameters are in good agreement with those derived in Section 5 based on WSLAP+ lens model within the uncertainties. The posterior probability distributions of the SMBH parameters and BCG mass obtained from *Lenstool* model are shown in Figure 19. The mean rms dispersion between the predicted and observed centroids for each set of the multiply lensed images circled in Figure 18 as measured in the image plane remains unchanged at $\langle \text{rms}_i \rangle = 0''.2$, indicating that the addition of a point mass has not appreciably perturbed the neighboring lensed images.

6.2. Adding a Local Extended Halo

Lenstool can also readily illuminate the question of an extended deflector, allowing us to obtain constraints on its mass, profile slope, core and truncation radii. We add an extended PIEMD halo to the *Lenstool* input (without a point mass) and then first place this extended galaxy within the lens plane ($z = 0.543$). The best-fit result requires the extra halo to be cuspy and highly truncated with a core radius of 0.8 kpc and truncation radius of 3.7 kpc. This best-fit halo has a mass of $2.3 \times 10^{10} M_{\odot}$, producing a high surface mass density giving its small truncation radius. In this case, the mean rms dispersion between the predicted and observed centroids for each set of the multiply lensed images circled in Figure 18 as measured in the image plane is $\langle \text{rms}_i \rangle = 1''.3$, which is much poorer than in the case of adding a point mass. Thus, although a massive and sharply truncated halo can bend L1 to the degree required, it also perturbs neighboring lensed images in an undesirable manner, as we found earlier by adding a halo in the WSLAP+ model. With the Bayesian likelihood values output by *Lenstool*, we also compute the Bayesian Information Criterion (BIC, Schwarz 1978) for three models: (1) model without an extra deflector local to L1, (2) model with an extended halo as a local deflector, and (3) model with a point mass as a local deflector. Their BIC values are 17.97, 27.25, and -0.80 , respectively. The model with a point mass is preferred because it has the lowest BIC value, while the model with an extended halo has the highest BIC value and is disfavored.

We then explore the addition of an extended deflector at a different redshift ranging over the redshift interval 0.3 to 1.2. The best-fit mass and position for this extra halo at different redshifts are shown in Figure 20. We conduct this analysis by manually adding a second lens plane containing only the extended deflector to the *Lenstool* modeling. This is not a full

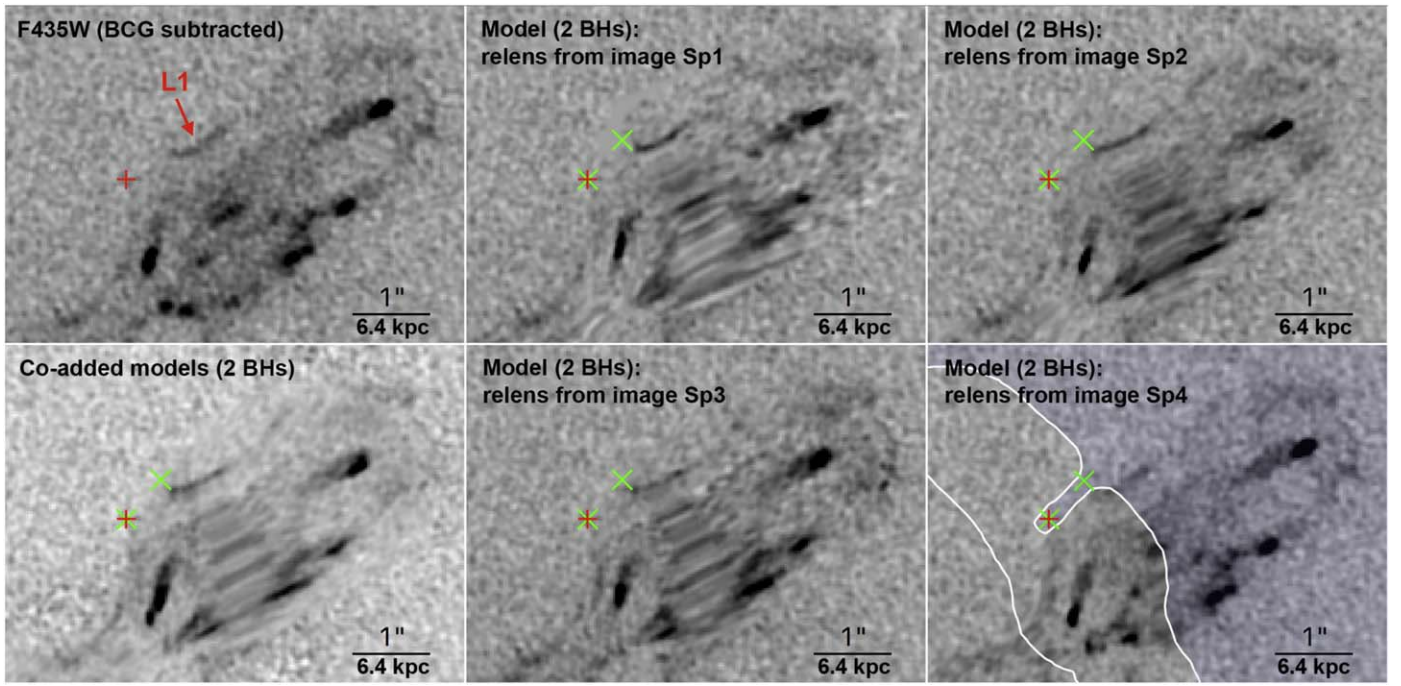


Figure 11. Same as Figure 10 but but now with the addition of two point masses, one at the center as in Figure 9 and the other offset as in Figure 10. This figure shows that there is no need for an additional central black hole, although our model does not exclude its presence.

multi-plane analysis, which is currently not supported by *Lenstool*. As can be seen, this halo is required to lie along the line of sight between the projected source position and the position of L1, increasing in mass and moving closer to the source position with increasing redshift. All of the solutions prefer a small truncation radius (~ 3 kpc) comparable to the radius of curvature of the lensed image, with an unconstrainably small core radius. As can be seen, as the halo moves significantly away from the cluster redshift (either toward lower or higher redshifts), it cannot alter the shape of the critical curve in the cluster lens plane sufficiently to produce the observed curvature of L1.

Similar to Figure 17, we also explore the visibility of this extended mass in contrast to the BCG light. We convert the mass distribution to light distribution with a $M/L = 10(M/L)_\odot$, $M/L = 50(M/L)_\odot$, and $M/L = 100(M/L)_\odot$ for the three halos at redshift 0.3, 0.543, and 1.0, respectively, following the method described in Section 5.3. Note that when the extended halo is located beyond the cluster redshift of 0.543, the halo image is lensed by the cluster, which will impose a magnification of $\mu > 2$. In the case of a halo at $z = 1.0$, this halo is lensed into three multiple images with a total magnification of $\mu \simeq 8.5$. We add the converted light maps for the extended halo at $z = 0.3$, $z = 0.543$, and $z = 1.0$ to F140W data, as shown in Figure 20 (lower row). If at $z = 0.3$ and having $M/L = 10(M/L)_\odot$, this halo should be clearly visible in the data. A much larger M/L is required to make any such halo undetectable. For halos at $z > 0.543$, a large M/L of $> 100(M/L)_\odot$ is required for the extended halo to be invisible in the data. In combination with its even more unfavorable ability to bend L1 to the degree required, both these factors make an intervening extended halo untenable.

7. Discussion

7.1. Rocket Effect

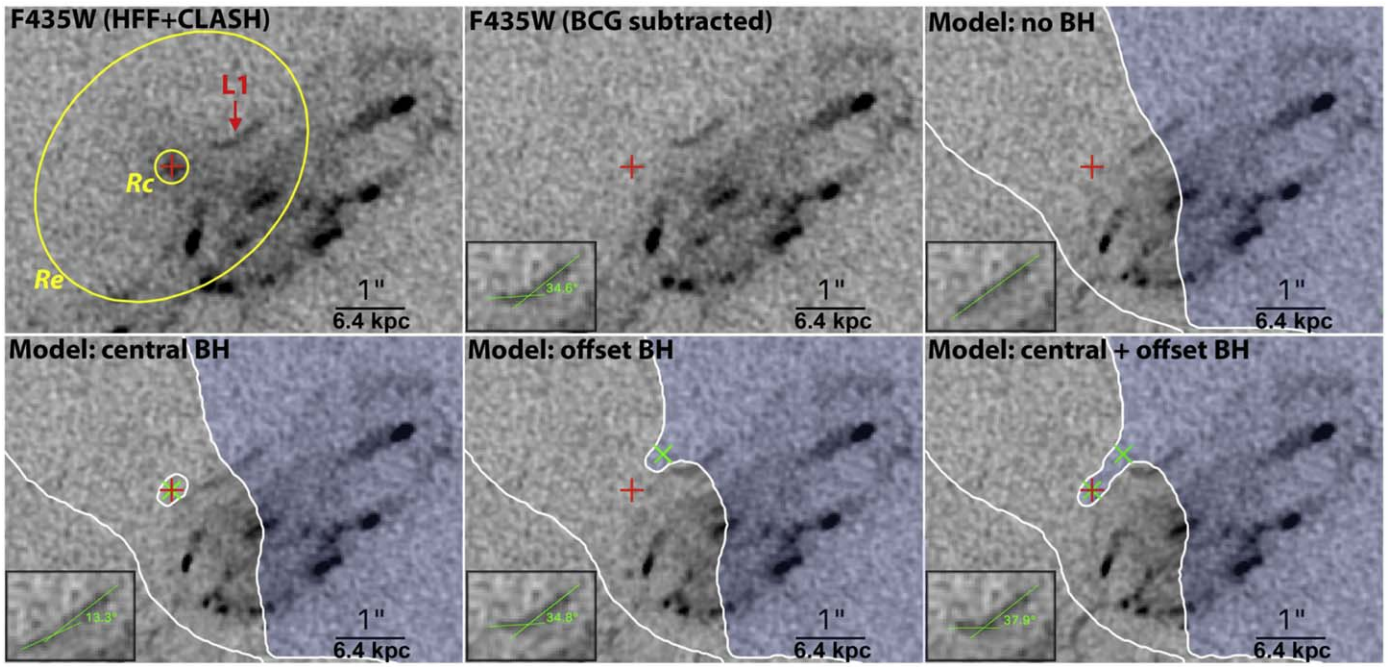
The simplest explanation for a dark point mass necessary to explain L1 is an SMBH hosted by the BCG. The central light profile of the BCG galaxy has a pronounced flattening within a few kpc of the light center, similar to other well known BCG galaxies where a stellar “core” is claimed (Postman et al. 2012). A convincing explanation for such cores involves prolonged binary SMBH merging so that stars are scattered away, particularly stars on radial orbits. This idea is supported by stellar dynamical data for the cluster NGC 1399 (Gebhardt et al. 2007). The BCG stellar core radius that we derive of 1.5 kpc (see Figure 15) is in line with other similarly luminous elliptical galaxies.

The location of the compact mass beyond the stellar core may imply ejection of the SMBH by the “rocket effect” (Bekenstein 1973; Begelman et al. 1980), where a preceding binary phase that may be responsible for the flattened core has resulted in the coalescence and ejection of the resulting merged black hole. In the calculations to follow, we estimate the initial kick velocity after the coalescence of the black hole binary and the time elapsed since the kick.

We first assume that the mass associated with the BCG is spherically symmetrical. We take the two-dimensional BCG mass distribution and de-project it using the inverse Abel transform to its three-dimensional mass distribution:

$$\rho(r) = \frac{1}{\pi} \int_{R_{\max}}^r \frac{1}{\sqrt{R^2 - r^2}} \frac{dI(R)}{dR} dR, \quad (11)$$

where $\rho(r)$ is the de-projected three-dimensional mass density and $I(R)$ is the two-dimensional surface mass density obtained from our lens model. We assume that the offset SMBH is at rest at the best-fit position of our model (~ 4.4 kpc from the centroid



Co-added model predictions using all four sets of lensed images:

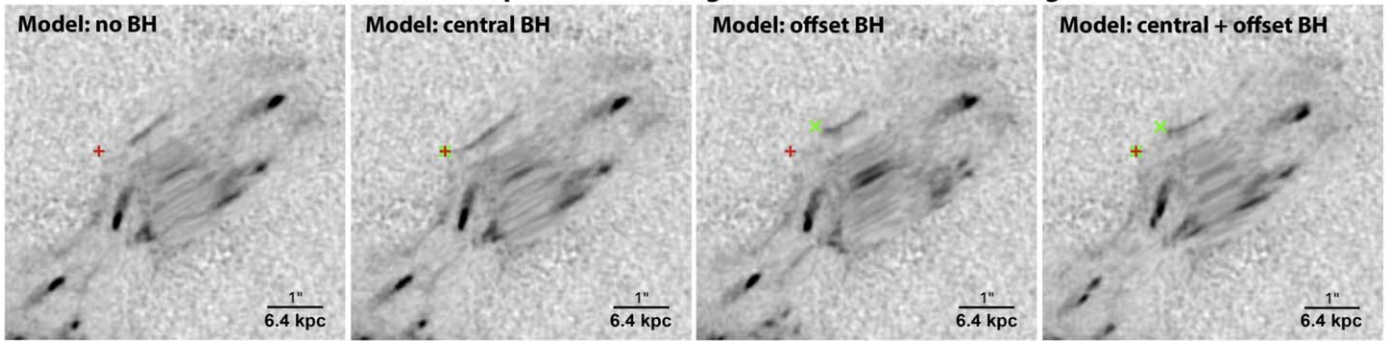


Figure 12. Comparison between data and lens model predictions. Upper and middle rows: image in the F435W band that provides the best contrast between the BCG (at its faintest in this band) and the multiply lensed H II regions of the spiral galaxy (at their brightest in this band). The projected light center of the BCG (determined in the infrared) is marked with a red cross in all the panels. The BCG’s stellar light core, R_c , and effective radius, R_e , are also plotted for reference. Subtraction of the BCG light in the B-band (derived by scaling from its infrared light) is shown in the upper middle panel, enhancing the appearance of nearby lensed features, including the “banana” shaped feature L1 in the spiral galaxy. We apply the free-form lens model to relens all of the data pixels highlighted by the blue area lying to the right of the critical lensing curve (white line), so that everything to the left of this line is the model prediction. The predicted data in “Model: no BH” is in very good agreement with the observed data except for the curved feature L1, for which we predict a linear feature symmetrically folded over a critical curve. Adding a central black hole does not produce the observed curvature but instead lengthens the predicted image of L1. A good match is obtained by placing the black hole closer to L1, at the position indicated by the green cross, with an Einstein radius of $0''.15$. Adding a second black hole at the center does not improve the fit, but is not excluded by the data. The inserts in the model-related panels indicate the deviation of the predicted image L1 from a straight line, as indicated by the two green lines. Lower row: the combined predicted images, summing over all four delensed and relensed images Sp1-4, for each of the above models.

of the BCG), and that the distance between the SMBH and the centroid of the BCG is the projected distance observed (i.e., the minimal true distance). We then calculate the minimum initial kick velocity by considering that the SMBH is only slowed down by gravitational force due to the gravitational potential of the BCG. Specifically, we obtain the gravitational deacceleration that the SMBH experienced at a radial position r by

$$\frac{1}{2}v(r)^2 = 4\pi G \left(\frac{1}{D} \int_0^D \rho(r') r'^2 dr' - \frac{1}{r} \int_0^r \rho(r') r'^2 dr' \right), \quad (12)$$

where $D = 4.4$ kpc. The right-hand side of Equation (12) is the difference between the gravitational potential energy at radius r and radius $D = 4.4$ kpc. When r approaches 0, we obtain the

initial kick velocity by

$$\frac{1}{2}v_{\text{kick}}^2 = \frac{4\pi G}{D} \int_0^D \rho(r') r'^2 dr'. \quad (13)$$

We calculate the time elapsed since the kick using

$$t = - \int_0^D \frac{1}{a(r)} \frac{dv(r)}{dr} dr, \quad (14)$$

where $v(r)$ is obtained from Equation (12) and the gravitational deacceleration $a(r)$ is obtained by

$$a(r) = \frac{GM(r)}{r^2} = \frac{4\pi G}{r^2} \int_0^r \rho(r') r'^2 dr'. \quad (15)$$

We find $v_{\text{kick}} = 314 \text{ km s}^{-1}$ and $t = 2.0 \times 10^7$ years. Theoretical studies suggest a kick velocity of a few hundred km s^{-1}

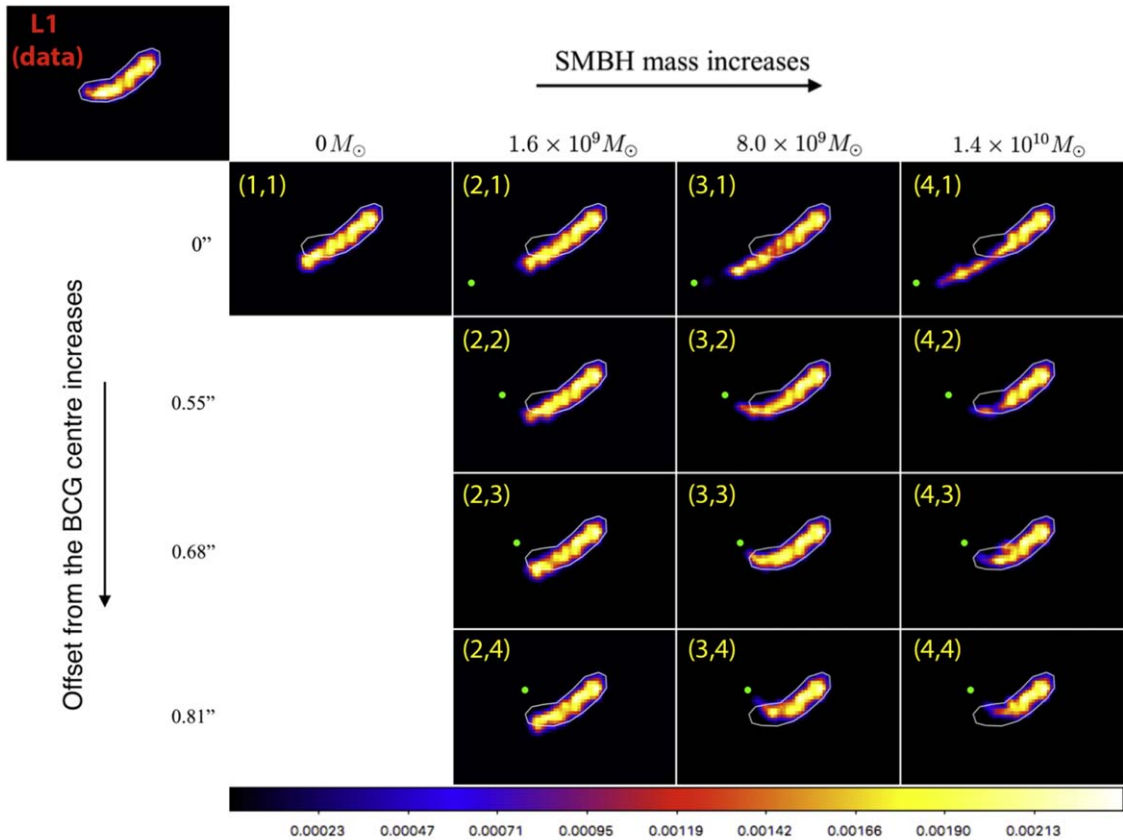


Figure 13. Observed vs. predicted images of L1 as a function of position and mass of a supermassive black hole. The uppermost panel shows L1, enclosed within an isophotal contour that is repeated in the remaining model panels. We explore a grid of positions and masses for a point mass that when added to the WSLAP+ lens model can produce the observed curvature of L1, where the offset of the point mass from the stellar light centroid of the BCG is shown on the vertical scale and its mass shown on the horizontal scale. The position of the point mass is indicated by the green dot in each model panel. In these panels, the best match is provided by a mass of $\sim 8 \times 10^9 M_\odot$ with an offset of $\sim 0''.7$ from the projected stellar light center. Color bar has units of e^-/s in the F435W band.

(Favata et al. 2004), demonstrating that the “rocket effect” is a viable explanation for the offset of the SMBH from the centre of its host galaxy. Empirically, several cases of spatially offset SMBH are now known in the SDSS survey (Lena et al. 2014; Kim et al. 2016; Chiaberge et al. 2017), linked to host galaxy merging, with a median offset of 4:6 kpc (Barrows et al. 2016) that is similar to our measured offset.

7.2. Implication for SMBH-galaxy Co-evolution

The BCG is a very luminous galaxy with a V-band magnitude of $M_V = -24.1$ and an integrated luminosity of $3.9 \times 10^{11} L_{V\odot}$ within an aperture of radius 30 kpc. Assuming a standard initial mass function viewed at a look back time of 5 Gyr (Chabrier 2003), we obtain a stellar mass of $6.3 \times 10^{11} M_\odot$. By comparison, in the WSLAP+ lens model, the total projected mass within a cylinder of $r < 30$ kpc centered on the BCG is $7.1 \times 10^{12} M_\odot$, contributed by both the BCG and other cluster matter along the line of sight. The SMBH mass of $M_{\text{BH}} = 8.4_{-1.8}^{+4.3} \times 10^9 M_\odot$ is in agreement with the local $M_{\text{BH}}-L_V$ relation between the SMBH mass and host galaxy V-band luminosity (Lauer et al. 2007).

This SMBH mass measurement seems to show that the $M_{\text{BH}}-L_V$ relation extends to an epoch when the Universe was half of its present age. Therefore, at face value, our work provides support for the co-evolution between SMBHs and their host galaxies. The BCG studied in this paper is a giant elliptical

galaxy residing at the center of a galaxy cluster that may not grow much more in mass over the next 5 Gyr. Therefore, that this object follows the $M_{\text{BH}}-L_V$ relation may not provide a strong argument that SMBHs necessarily grow in step with their host galaxies over cosmic times; i.e., it may already closely resemble BCGs in the local universe. To shed light on the co-evolution of SMBHs and their host galaxies, what is needed are measurements of SMBH masses for distant galaxies that will still grow substantially in mass to the current epoch.

8. Conclusion

Direct and reliable determinations of the masses of SMBHs are restricted to the low-redshift Universe ($z < 0.3$), where stellar or gas kinematics in the close vicinity of the SMBH can be spatially resolved (Ferrarese & Ford 1999; Kormendy 2004). At larger distances, we have to resort to reverberation mapping. The veracity of the masses thus inferred for SMBHs remains poorly understood. For galaxies where both of these techniques are not possible, we have to resort to scaling relations—the veracity of which is even more questionable (Graham 2016). Nevertheless, based on these measurements, the masses of local SMBHs are found to correlate with the masses of the bulge components of their host galaxies, indicating a co-evolution history of SMBHs and their host galaxies (Kormendy & Ho 2013). However, bright QSOs at high redshifts are inferred to harbor extremely massive SMBHs, suggesting a rapid early growth of SMBH. This is in contradiction with the idea of a co-evolution with their host galaxies (Trakhtenbrot et al. 2015;

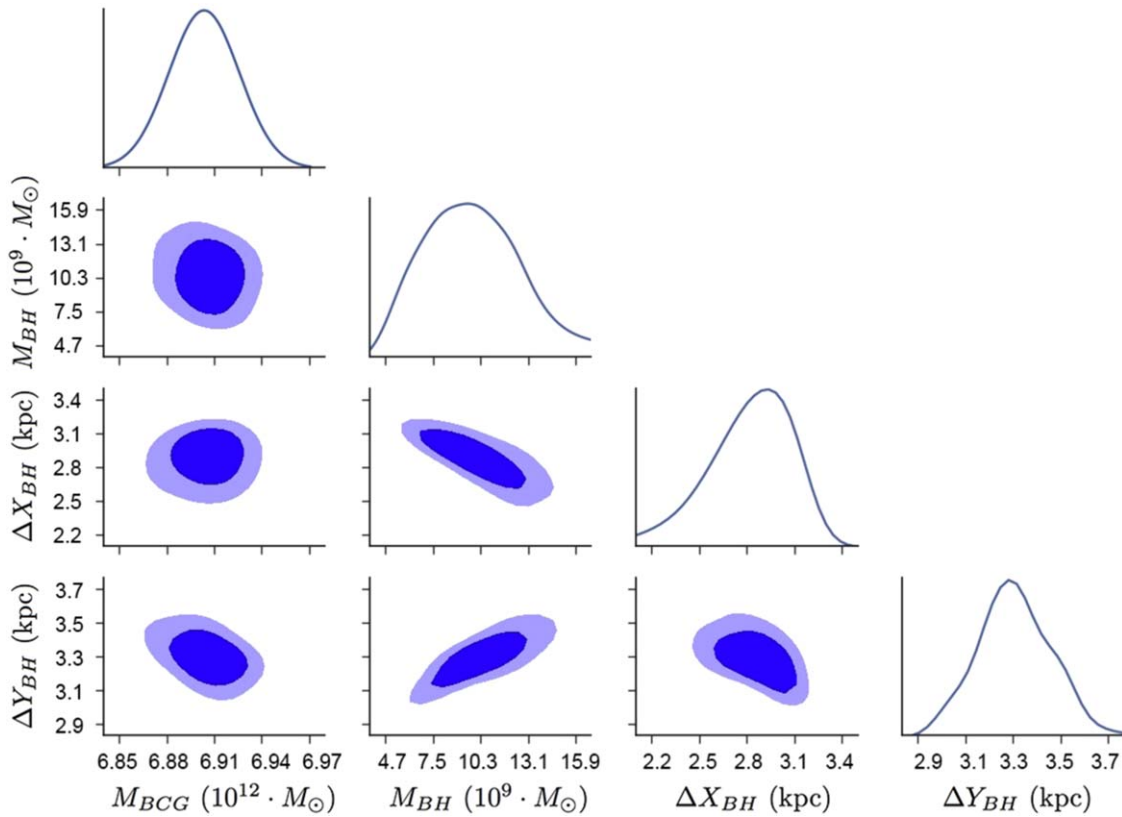


Figure 14. Posterior probability distributions of BCG mass and the three parameters describing the black hole, its position x and y , and its mass M_{BH} . Contours represent 68% and 95% confidence levels. The BH parameters are constrained to be (within in 68% CI) $M_{\text{BH}} = 8.4_{-1.8}^{+4.3} \times 10^9 M_{\odot}$ and $\Delta_{\text{BH}} = 4.4 \pm 0.3$ kpc.

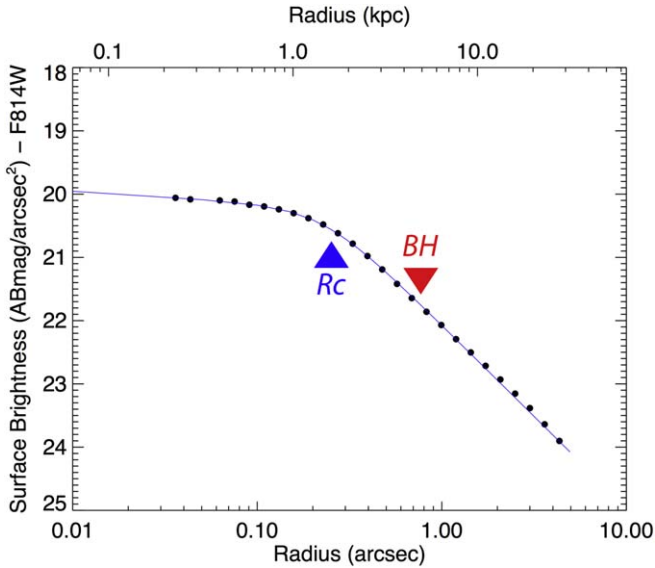


Figure 15. Standard “Nuker” profile fitted to the BCG light distribution. The profile shows a distinct break at radius of 1.5 kpc. The radial location that we derive for the black hole is marked (at ~ 5 kpc), lying beyond the break.

Wu et al. 2015). To understand the growth of SMBH masses over cosmic time, what is clearly needed is the ability to accurately measure SMBH masses over cosmic history.

Gravitational lensing is a promising method to directly measure SMBH masses beyond the local Universe. In the case of lensing by a single foreground galaxy with near-perfect alignment between the source and the lensing galaxy, a central de-magnified image is generic to lensing, generated within the Einstein radius of

the lensing galaxy; such that the larger the SMBH mass, the more this central image is attracted toward the SMBH and is de-magnified (Mao et al. 2001; Hezaveh et al. 2015; Rusin et al. 2005). However, in these cases there is an inherent degeneracy between the mass of an SMBH and the central mass profile of its host galaxy. In situations where a background galaxy is lensed by a foreground galaxy cluster so that an individual lensed feature appears close to the center of a cluster member, as in the case presented in this paper, an unambiguous case can be made for an SMBH and its parameters can be directly determined.

Specifically, based on a “banana-shaped” lensed feature L1 that is located close to the center of the BCG in MACS 1149, we constrain the position and mass of an offset SMBH hosted by this BCG. L1 is the closest image to the center of the BCG and it exhibits a radius of curvature of only $0''.6$. L1 is consistently predicted to have a straight appearance owing to the large shear at its position. To bend L1 on the small scale as observed in the data, a local deflector is required in addition to the combined cluster and member galaxy lens model to alter the shape of the critical curve cutting through L1. By adding a point mass of $M_{\text{BH}} = 8.4_{-1.8}^{+4.3} \times 10^9 M_{\odot}$ offset by ~ 4.4 kpc from the center of the BCG, we successfully reproduce the curvature of L1. A highly truncated extended halo (i.e., a galaxy) with a velocity dispersion $> 100 \text{ km s}^{-1}$ can also bend L1 to the degree required. However, such a galaxy deflects neighboring lensed images from the same background galaxy in an undesirable way. Furthermore, the halo is required to have a large mass-to-light ratio ($> 50(M/L)_{\odot}$) to be invisible in the deep HFF data; however, no galaxy with such a large mass and a high mass-to-light ratio is known. The point-mass solution can be interpreted as an offset SMBH that has been

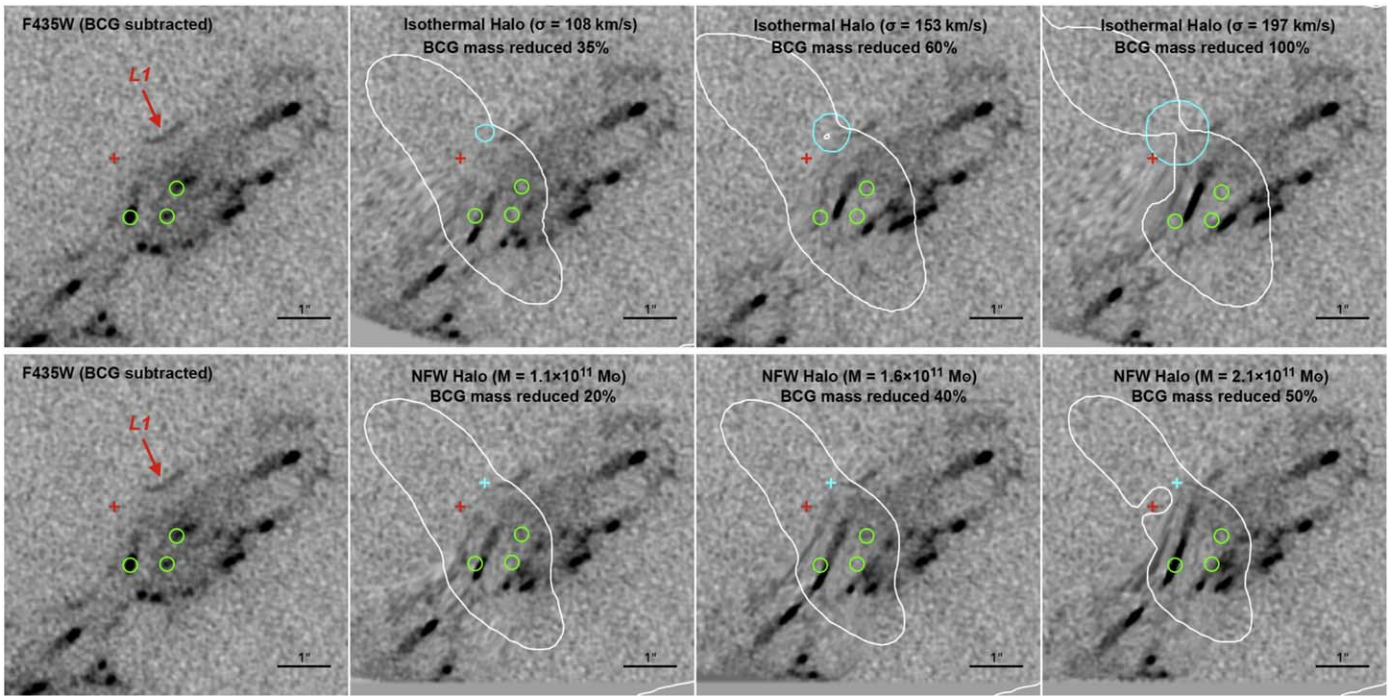


Figure 16. Lensing effect of an extended small substructure. Red crosses indicate the light center of the BCG. Green circles indicate the three major lensed features in the released region, same as in Figure 2. Upper row: replacing the offset point mass with a halo comprising a singular isothermal sphere with a velocity dispersion, and hence mass, as indicated in each panel. The cyan circles indicate the corresponding isothermal sphere’s Einstein radius for a source at $z = 1.4888$. The BCG mass is scaled down as the halo mass is raised, which is indicated in each panel. Although such a lens can reproduce the observed curvature of L1, its relatively extended deflection field has the effect of significantly deflecting other lensed features so that we no longer achieve good agreement between these features and the data. Lower row: as in the upper row but now but with an NFW profile representing the extended halo. Cyan crosses indicate the centers of NFW halos. Once again, although we can reproduce the observed curvature of L1, the other lensed features are perturbed so that they no longer agree with the data.

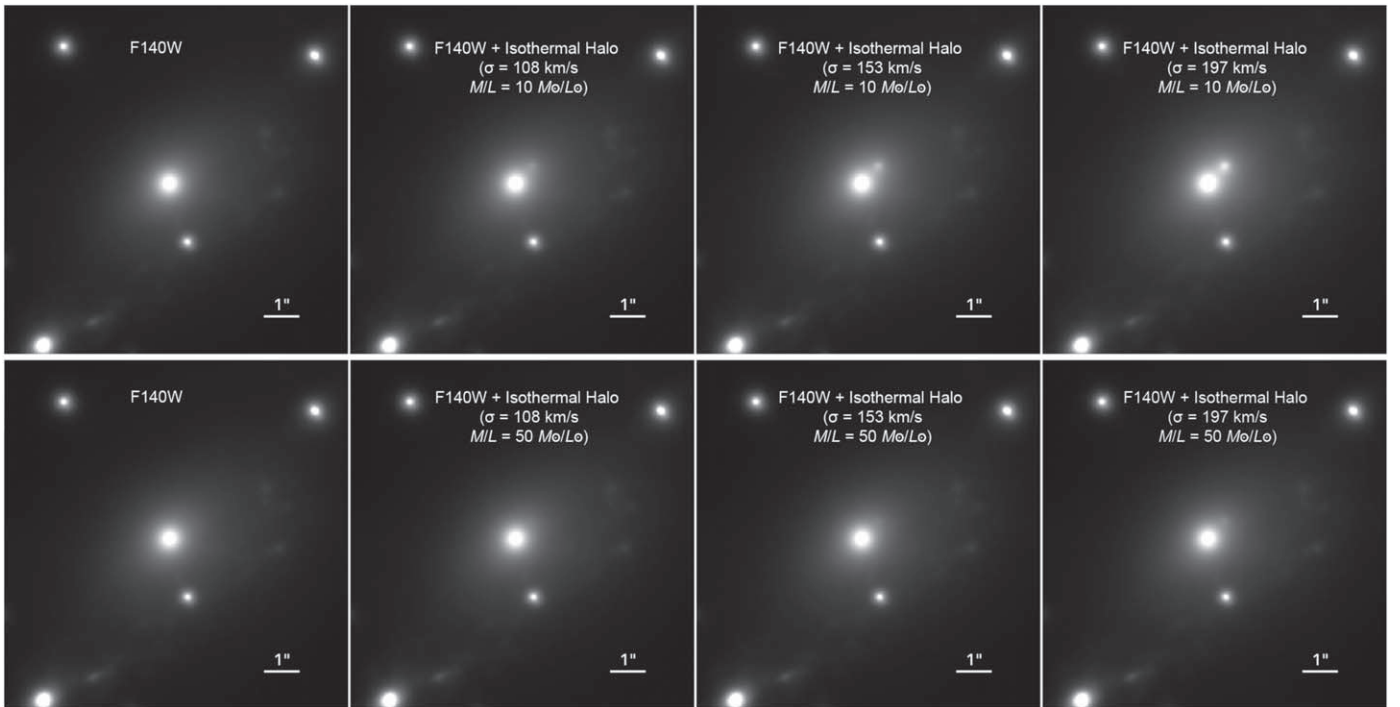


Figure 17. Visibility of a halo comprising a singular isothermal sphere as described in Figure 16 for different mass-to-light (M/L) ratios. Upper row: $M/L = 10(M/L)_{\odot}$, as is the case for a typical elliptical galaxy. Three different velocity dispersions and hence masses for the halo are shown in their corresponding luminosity, with the latter computed from the bolometric correction in Buzzoni (2005) and convolved with a PSF adopted from an isolated star in F140W band data. Lower row: same as the upper row but with a high mass-to-light ratio of $50(M/L)_{\odot}$. Despite the glare from the BCG, one should be able to see the starlight from the halo, unless it has a high M/L ratio above 50.

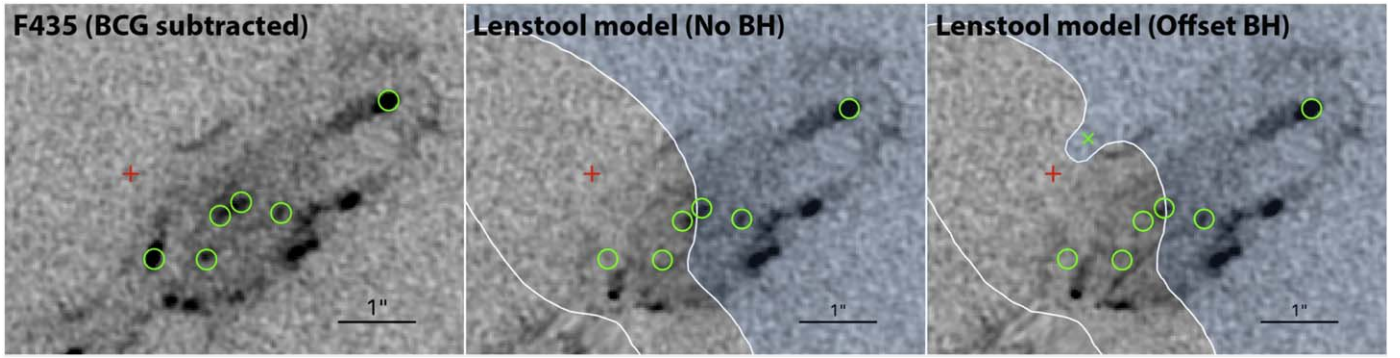


Figure 18. Released predictions from *Lenstool* model. Green circles point out three major features close to the center of the BCG, and they are plotted in both model panels to indicate the offset of the model predicted positions. Without adding an offset dark point mass, the parametric package *Lenstool* produces a consistent model with WLSAP+. A critical curve passes through L1, leaving the prediction from the right half of L1 very straight (middle panel). We later add a point mass in the lens modeling with its position and mass being a free parameter, and we then re-run the model construction while leaving all of the parameters of the BCG free and holding the other parameters fixed. *Lenstool* produces a similar result as WLSAP+ with an extra point mass, with offset from the BCG of $0''.66 \pm 0.11$ and a higher best-fit mass of $1.2^{+0.19}_{-0.22} \times 10^{10} M_{\odot}$ (the released image is shown in the right-hand panel).

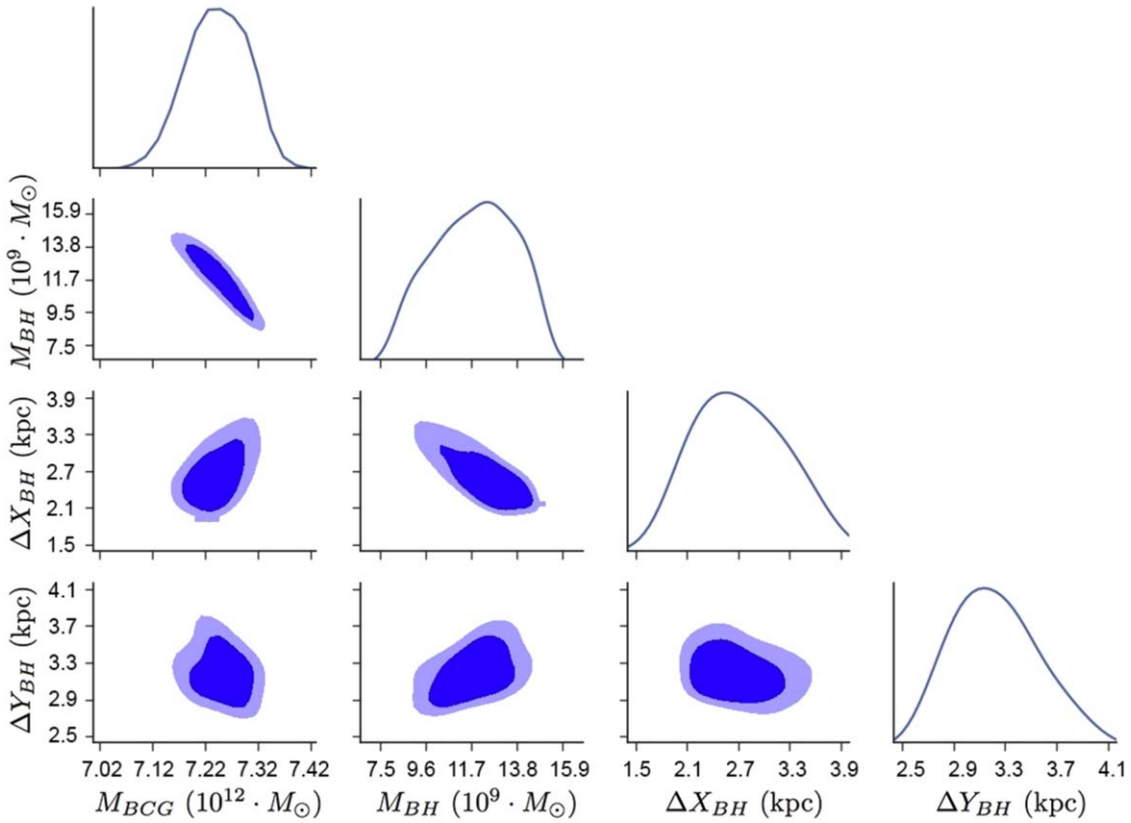


Figure 19. Posterior probability distributions of BCG mass and the three parameters describing the black hole, its position x and y , and its mass M_{BH} from *Lenstool* model. Similar to Figure 14, contours represent 68% and 95% confidence levels. In this *Lenstool* solution, the BH parameters are constrained to be (within in 68% CI) $M_{\text{BH}} = 1.25^{+0.14}_{-0.27} \times 10^{10} M_{\odot}$ and $\Delta_{\text{BH}} = 4.2^{+0.7}_{-0.5}$ kpc.

ejected from the BCG center due to the asymmetric gravitational wave radiation in a merger event (the “rocket effect”). To bring the SMBH to its present position, the kick should have occurred at least 2.0×10^7 years ago and it should have had an initial velocity of $>314 \text{ km s}^{-1}$. These estimations are in agreement with theoretical predictions for typical “rocket effect” events.

The authors would like to thank the anonymous referee for providing detailed and constructive suggestions to improve this paper. J.L. acknowledges support from the Research Grants

Council of Hong Kong through grant 17319316. J.L. also acknowledges a Seed Fund for Basic Research from the University of Hong Kong. T.B. was supported by a Visiting Research Professor Scheme from the University of Hong Kong, during which major parts of this work were conducted. J.M.D. acknowledges the support of projects AYA2015-64508-P (MINECO/FEDER, UE), AYA2012-39475-C02-01, and the consolder project CSD2010-00064 funded by the Ministerio de Economia y Competitividad. Y.O. is supported by the Ministry of Science and Technology (*MOST*) of Taiwan, *MOST* 106-2112-M-001-008.

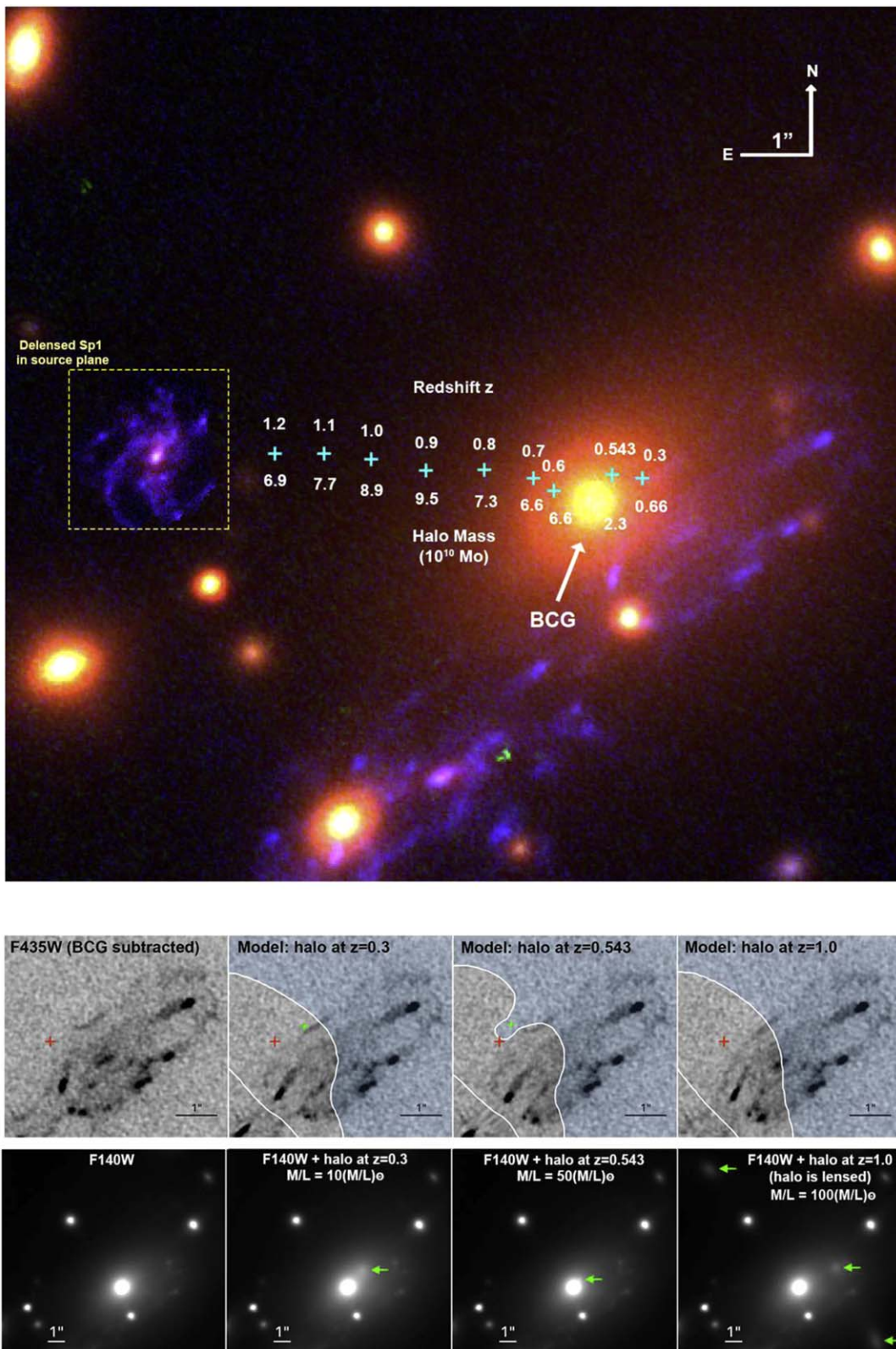


Figure 20. Adding an extended PIEMD halo at different redshifts. Upper row: best-fit position and mass of an extended halo at different redshifts. The delensed image of Sp1 in the source plane is overlotted on this color data image, and the cyan crosses indicate the best-fit position of the extended halos with their redshifts marked on the top and best-fit mass marked on the bottom. Those extended halos align in the region between the projected source position and the position of L1, with generally bigger masses toward higher redshifts while moving closer to the source position. Middle row: relensed predictions with an extended halo at redshift 0.3, 0.543 and 1.0. The red cross is the position of the BCG’s light center and the green cross is the position of the extended halo. Note that in the case of $z = 1.0$, the extended halo is located outside the field of view showing in this model prediction region. Lower row: halo starlight added to F140W data. The extended halo with mass shown in panel a is converted to starlight (with $M/L = 10(M/L)_{\odot}$, $M/L = 50(M/L)_{\odot}$, and $M/L = 100(M/L)_{\odot}$ respectively) by the method described in Chapter 6, similar to Figure 17. Note that when the halo is at $z = 1.0$, it is strongly lensed by the cluster and forms three multiple images.

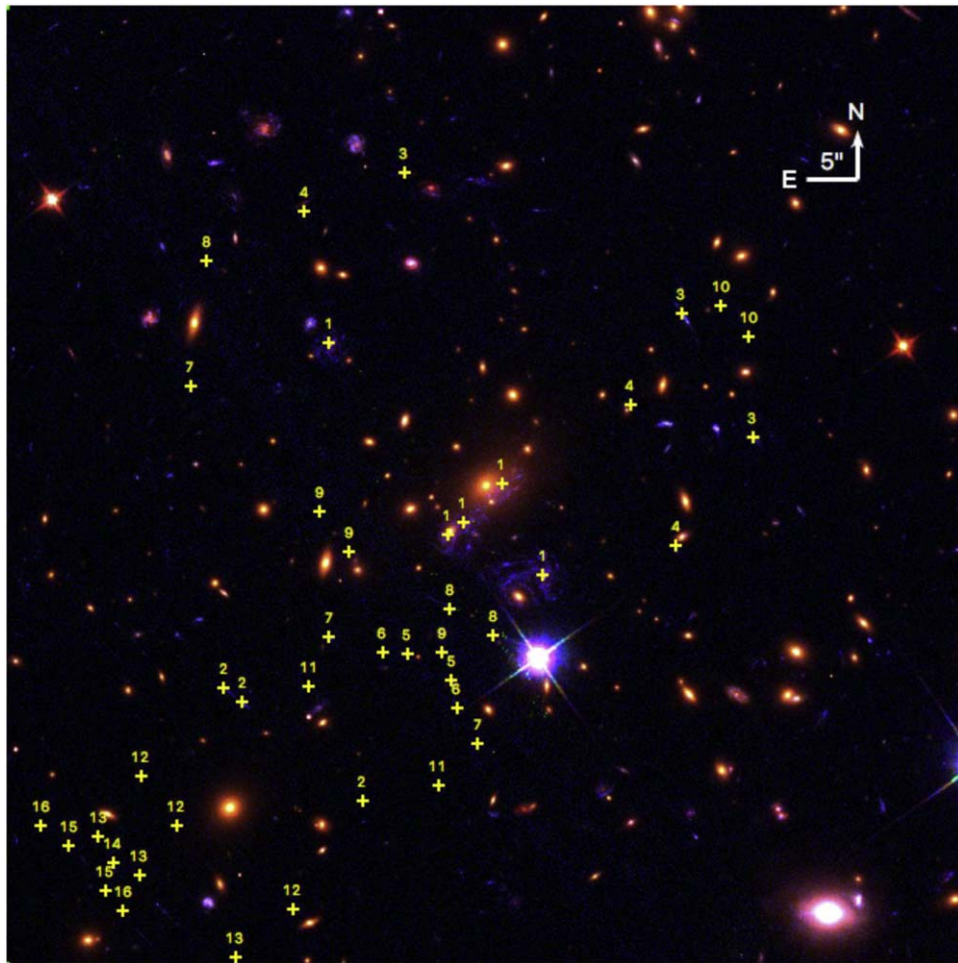






Figure 21. Locations of the 16 multiply lensed image systems used as constraints for constructing both WSLAP+ and *Lenstool* models. The image shown here is with the WSLAP+ modeling FOV of $1'.6 \times 1'.6$. System 1 corresponds to the background spiral galaxy at $z = 1.4888$. We used more than 20 individual features from system 1 as constraints, as listed in Table 1.

Appendix Information of Multiply Lensed Images

In Figure 21, we plot the locations of the multiply lensed images that are used as constraints for lens modeling and we tabulate their coordinates.

ORCID iDs

Mandy C. Chen  <https://orcid.org/0000-0002-8739-3163>
 Jose M. Diego  <https://orcid.org/0000-0001-9065-3926>
 Youichi Ohya  <https://orcid.org/0000-0001-9490-3582>
 Holland Ford  <https://orcid.org/0000-0001-8505-9535>

References

- Barrows, R. S., Comerford, J. M., Greene, J. E., & Pooley, D. 2016, *ApJ*, 829, 2016
- Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, *Natur*, 287, 307
- Bekenstein, J. D. 1973, *ApJ*, 183, 657
- Bentz, M. C., Peterson, B. M., Netzer, H., Pogge, R. W., & Vestergaard, M. 2009, *ApJ*, 697, 160
- Broadhurst, T., Benítez, N., Coe, D., et al. 2005, *ApJ*, 621, 53
- Buzzoni, A. 2005, *MNRAS*, 361, 725
- Chabrier, G. 2003, *PASP*, 115, 763
- Chiaberge, M., Ely, J. C., Meyer, E. T., et al. 2017, *A&A*, 600, A57
- Diego, J. M., Broadhurst, T., Chen, C., et al. 2016, *MNRAS*, 456, 356
- Diego, J. M., Kaiser, N., Broadhurst, T., et al. 2018, *ApJ*, 857, 25
- Diego, J. M., Protopapas, P., Sandvik, H. B., & Tegmark, M. 2005, *MNRAS*, 360, 477
- Diego, J. M., Tegmark, M., Protopapas, P., & Sandvik, H. B. 2007, *MNRAS*, 375, 958
- Ebeling, H., Barrett, E., Donovan, D., et al. 2007, *ApJL*, 661, L33
- Elíasdóttir, Á, Limousin, M., Richard, J., et al. 2007, arXiv:0710.5636
- Emsellem, E. 2013, *MNRAS*, 433, 1862
- Favata, M., Hughes, S. A., & Holz, D. E. 2004, *ApJL*, 607, L5
- Ferrarese, L., & Ford, H. C. 1999, *ApJ*, 515, 583
- Ferrarese, L., & Merritt, D. 2000, *ApJL*, 539, L9
- Gebhardt, K., Lauer, T. R., Pinkney, J., et al. 2007, *ApJ*, 671, 1321
- Graham, A. W. 2016, *ASSL*, 418, 263
- Grillo, C., Karman, W., Suyu, S. H., et al. 2016, *ApJ*, 822, 78
- Halkola, A., Hildebrandt, H., Schrabback, T., et al. 2008, *A&A*, 481, 65
- Hezaveh, Y. D., Marshall, P. J., & Blandford, R. D. 2015, *ApJL*, 799, L22
- Jullo, E., Kneib, J.-P., Limousin, M., et al. 2007, *NJPh*, 9, 447
- Kaspi, S., Brandt, W. N., Maoz, D., et al. 2007, *ApJ*, 659, 997
- Kaspi, S., Maoz, D., Netzer, H., et al. 2005, *ApJ*, 629, 61
- Kassiola, A., & Kovner, I. 1993, *ApJ*, 417, 450
- Kassiola, A., Kovner, I., & Fort, B. 1992, *ApJ*, 400, 41
- Kelly, P. L., Diego, J. M., Rodney, S., et al. 2018, *NatAs*, 2, 334
- Kelly, P. L., Rodney, S. A., Treu, T., et al. 2015, *Sci*, 347, 1123
- Kelly, P. L., Rodney, S. A., Treu, T., et al. 2016, *ApJL*, 819, L8
- Kim, D.-C., Evans, A. S., Stierwalt, S., & Privon, G. C. 2016, *ApJ*, 824, 122
- Kneib, J.-P., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M. 1996, *ApJ*, 471, 643

- Kormendy, J. 2004, in Carnegie Carnegie Observatories Astrophysics Ser., *Coevolution of Black Holes and Galaxies*, ed. L. C. Ho (Cambridge: Cambridge Univ. Press), 1
- Kormendy, J., & Ho, L. C. 2013, *ARA&A*, 51, 511
- Lauer, T. R., Faber, S. M., Richstone, D., et al. 2007, *ApJ*, 662, 808
- Lena, D., Robinson, A., Marconi, A., et al. 2014, *ApJ*, 795, 146
- Mao, S., Witt, H. J., & Koopmans, L. V. E. 2001, *MNRAS*, 323, 301
- Markwardt, C. B. 2009, in ASP Conf. Ser. 411, *Astronomical Data Analysis Software and Systems XVIII*, ed. D. A. Bohlender, D. Durand, & P. Dowler (San Francisco, CA: ASP), 251
- Meyer, L., Ghez, A. M., Schödel, R., et al. 2012, *Sci*, 338, 84
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, *ApJ*, 462, 563
- Oguri, M. 2015, *MNRAS*, 449, L86
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, *AJ*, 139, 2097
- Ponente, P. P., & Diego, J. M. 2011, *A&A*, 535, A119
- Postman, M., Lauer, T. R., Donahue, M., et al. 2012, *ApJ*, 756, 159
- Quinn, J., Jackson, N., Tagore, A., et al. 2016, *MNRAS*, 459, 2394
- Rau, S., Vegetti, S., & White, S. D. M. 2014, *MNRAS*, 443, 957
- Rusin, D., Keeton, C. R., & Winn, J. N. 2005, *ApJL*, 627, L93
- Schwarz, G. 1978, *AnSta*, 6, 461
- Sendra, I., Diego, J. M., Broadhurst, T., & Lazkoz, R. 2014, *MNRAS*, 437, 2642
- Sharon, K., & Johnson, T. L. 2015, *ApJL*, 800, L26
- Simon, J. D., & Geha, M. 2007, *ApJ*, 670, 313
- Smith, G. P., Ebeling, H., Limousin, M., et al. 2009, *ApJL*, 707, L163
- Smith, R. J., Lucey, J. R., & Edge, A. C. 2017, *MNRAS*, 467, 836
- Suyu, S. H., Hensel, S. W., McKean, J. P., et al. 2012, *ApJ*, 750, 10
- Tamura, Y., Oguri, M., Iono, D., et al. 2015, *PASJ*, 67, 72
- Trakhtenbrot, B., Urry, C. M., Civano, F., et al. 2015, *Sci*, 349, 168
- Treu, T., Brammer, G., Diego, J. M., et al. 2016, *ApJ*, 817, 60
- van den Bosch, R. C. E., Gebhardt, K., Gültekin, K., et al. 2012, *Natur*, 491, 729
- Winn, J. N., Rusin, D., & Kochanek, C. S. 2003, *ApJ*, 587, 80
- Winn, J. N., Rusin, D., & Kochanek, C. S. 2004, *Natur*, 427, 613
- Wong, K. C., Suyu, S. H., & Matsushita, S. 2015, *ApJ*, 811, 115
- Wu, X.-B., Wang, F., Fan, X., et al. 2015, *Natur*, 518, 512
- Zheng, W., Postman, M., Zitrin, A., et al. 2012, *Natur*, 489, 406
- Zitrin, A., & Broadhurst, T. 2009, *ApJL*, 703, L132