A Highly Magnified Gravitationally Lensed Red QSO at $z = 2.5$ with a Significant Flux Ratio Anomaly

Eilat Glikman[1] , Cristian E. Rusu[2,3] , Geoff C.-F. Chen[4] , James Hung-Hsu Chan[5] , Cristiana Spingola[6,7] ,
Hannah Stacey[8] , John McKean[9] , Ciprian T. Berghea[10] , S. G. Djorgovski[11] , Matthew J. Graham[11] , Daniel Stern[12] ,
Tanya Urrutia[13] , Mark Lacy[14] , Nathan J. Secrest[10] , and John M. O'Meara[15,16]

[1] Department of Physics, Middlebury College, Middlebury, VT 05753, USA
[2] Subaru Fellow, Subaru Telescope, National Astronomical Observatory of Japan, 650 N. Aohoku Place, Hilo, HI 96720, USA
[3] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-0015, Japan
[4] Department of Physics, University of California, Davis, CA 95616, USA
[5] Institute of Physics, Laboratory of Astrophysique, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290, Versoix, Switzerland
[6] Dipartimento di Fisica e Astronomia, Università degli Studi di Bologna, Via Gobetti 93/2, I-40129 Bologna, Italy
[7] INAF—Istituto di Radioastronomia, Via Gobetti 101, I-40129, Bologna, Italy
[8] Max Planck Institute for Astrophysics, Karl-Schwarzschild Str. 1, D-85748 Garching bei München, Germany
[9] Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, The Netherlands
[10] U.S. Naval Observatory, 3450 Massachusetts Avenue NW, Washington, DC 20392, USA
[11] California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA
[12] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
[13] Leibniz Institut für Astrophysik, An der Sternwarte 16, D-14482 Potsdam, Germany
[14] National Radio Astronomy Observatory, Charlottesville, VA 22903, USA
[15] Department of Physics, Saint Michael's College, One Winooski Park, Colchester, VT 05439, USA
[16] W. M. Keck Observatory, 65-1120 Mamalahoa Highway, Kamuela, HI 96743, USA

Abstract

We present the discovery of a gravitationally lensed dust-reddened QSO at $z = 2.517$, identified in a survey for QSOs by infrared selection. Hubble Space Telescope imaging reveals a quadruply lensed system in a cusp configuration, with a maximum image separation of $\sim1\rlap{.}''8$. We find that, compared to the central image of the cusp, the neighboring brightest image is anomalous by a factor of ~7–10, which is the largest flux anomaly measured to date in a lensed QSO. Incorporating high-resolution Very Large Array radio imaging and submillimeter imaging with the Atacama Large Millimeter/submillimeter Array, we conclude that a low-mass perturber is the most likely explanation for the anomaly. The optical through near-infrared spectrum reveals that the QSO is moderately reddened with $E(B - V) \simeq 0.7$–0.9. We see an upturn in the ultraviolet spectrum due to $\sim1\%$ of the intrinsic emission being leaked back into the line of sight, which suggests that the reddening is intrinsic and not due to the lens. The QSO may have an Eddington ratio as high as $L/L_{\rm Edd} \approx 0.2$. Consistent with previous red QSO samples, this source exhibits outflows in its spectrum, as well as morphological properties suggestive of it being in a merger-driven transitional phase. We find a host galaxy stellar mass of $\log M_\star/M_\odot = 11.4$, which is higher than the local $M_{\rm BH}$ versus $M_\star$ relation but consistent with other high-redshift QSOs. When demagnified, this QSO is at the knee of the luminosity function, allowing for the detailed study of a more typical moderate-luminosity infrared-selected QSO at high redshift.

1. Introduction

Models of galaxy evolution that invoke major mergers (e.g., Sanders et al. 1988; Di Matteo et al. 2005; Hopkins et al. 2005) have been highly successful at incorporating the growth of supermassive black holes (SMBHs) in galactic nuclei and explaining various scaling relations between the two, such as the M–σ relation (Ferrarese & Merritt 2000; Gebhardt et al. 2000). These models predict a phase during the merger process in which the growing SMBH is enshrouded by dust. In addition, while at its peak luminosity, this active galactic nucleus (AGN; or the more luminous QSO[17]) is heavily obscured and thus elusive to most

AGN and QSO surveying techniques, especially at visible wavelengths. Recent work in the near-infrared has revealed a population of QSOs with moderate amounts of dust extinction that appear to be transitioning from a heavily dust-enshrouded phase to a typical, unobscured QSO (e.g., Banerji et al. 2012; Glikman et al. 2012; Brusa et al. 2015).

Combining near-infrared and radio data has proven to be a very effective method for finding quasars in this transitional state (Glikman et al. 2004, 2007; Urrutia et al. 2009; Glikman et al. 2012, 2013). These efforts have resulted in a sample of $\gtrsim120$ dust-reddened quasars from the combined Faint Images of the Radio Sky at Twenty cm (FIRST)+Two Micron All Sky Survey (2MASS) surveys (F2M) with reddenings in the range $0.1 < E_{B-V} < 1.5$. F2M red quasars are found to be predominantly driven by major mergers (Urrutia et al. 2008; Glikman et al. 2015), are accreting at very high rates ($L/L_{\rm Edd} \simeq 0.69$; Kim et al. 2015), and exhibit broad absorption lines associated with outflows and feedback (Urrutia et al. 2009). These properties are consistent with buried quasars expelling their

---

[17] In this work, we adopt the canonical nomenclature that distinguishes quasars, radio-detected luminous AGNs whose radio emission is essential to their selection, from QSOs, the overall class of luminous AGNs.

Table 1
Integrated Photometry of W2M J1042+1641 from Available Surveys

Band	AB Mag	Band	AB Mag
u^a	20.93 ± 0.10	H^b	16.87 ± 0.13
g^a	20.40 ± 0.03	$K_s{}^b$	15.87 ± 0.06
r^a	20.26 ± 0.03	W1^c	15.52 ± 0.02
i^a	20.04 ± 0.04	W2^c	14.97 ± 0.02
z^a	18.99 ± 0.05	W3^c	13.00 ± 0.02
J^b	17.84 ± 0.17	W4^c	11.84 ± 0.04

Notes.
[a] SDSS model magnitudes from de Vaucouleurs profile fitting.
[b] 2MASS magnitudes.
[c] AllWISE magnitudes.

dusty shrouds in an evolutionary phase predicted by merger-driven coevolution models.

Among the sources in the F2M sample, two gravitationally lensed systems were found. F2M J0134−0931 is a radio-loud red quasar at $z = 2.216$ that is lensed into at least five images, possibly by two galaxies at $z = 0.7645$ (Gregg et al. 2002; Hall et al. 2002; Winn et al. 2002). This scenario (Keeton et al. 2003) proposes that the lenses are both spiral galaxies, which may then also be responsible for the reddening. F2M J1004 +1229, at $z = 2.65$, is a rare low-ionization broad absorption line quasar (LoBAL) that includes strong absorption from metastable Fe II (FeLoBAL; Becker et al. 1997). The location of the reddening in this system is unclear (Lacy et al. 2002).

Based on distinct color differences between optically selected lensed QSOs and those selected in the radio or infrared, Malhotra et al. (1997) suggest that reddening by dust in the lensing galaxy is biasing surveys for lensed QSOs, underestimating their numbers. Alternatively, if the reddening of the lensed QSOs is intrinsic, then their larger presence suggests that the population of red quasars may be significantly underestimated.

Both of the lensed F2M quasars were selected requiring a radio detection, a wavelength that is largely insensitive to dust reddening, which may have made them easier to find. However, since radio-loud and radio-intermediate quasars make up only $\sim$10% of the overall quasar population (Ivezić et al 2002), the radio restriction also limited the sample to a rarer class of quasars. In this paper, we report the discovery of a quadruply lensed radio-quiet red QSO discovered in a search for red QSOs using Wide-field Infrared Survey Explorer (WISE) color selection and no radio criterion.

Throughout this work we quote magnitudes on the AB system, unless explicitly stated otherwise. When computing luminosities and any other cosmology-dependent quantities, we use the ΛCDM concordance cosmology: $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.30$, and $\Omega_\Lambda = 0.70$.

2. Discovery and Observations

2.1. Selection

We recently constructed a sample of radio-quiet dust-reddened QSOs selected by their infrared colors in WISE and 2MASS (W2M), applying well-established color cuts in WISE color space (Lacy et al. 2004; Stern et al. 2005; Donley et al. 2012; Stern et al. 2012; Assef et al. 2013, 2018; Glikman et al. 2018) and the infrared-to-optical KX color space (Warren et al. 2000). Our survey covers $\sim$2000 deg^2 with a relatively shallow near-infrared flux limit ($K < 16.7$) and has resulted in 37 newly

identified red QSOs (Glikman et al. 2022). Among the sources was W2M J104222.11+164115.3,[18] whose infrared luminosity, based on WISE photometry, $L_{IR} \simeq 10^{14} L_\odot$, was more luminous than any other known radio-quiet QSO and implied extreme properties suggestive of gravitational lensing.

The source is undetected in FIRST, implying that its 20 cm flux density is below 1 mJy. There are also multiepoch Swift/ X-ray Telescope (XRT) observations of this object revealing moderately high absorption ($N_H \sim 10^{23}$ cm^{-2}) and exhibiting variability (Matsuoka et al. 2018). Table 1 lists the broadband magnitudes of this source from the surveys used in its discovery.

2.2. Initial Spectroscopy

W2M J1042+1641 was observed with the MODS1B Spectrograph on the Large Binocular Telescope (LBT) observatory for 1200 s with the red and blue arms simultaneously, with a 0″.6-wide slit on UT 2013 March 14, covering the wavelength range 3300–10100 Å. After removing the CCD signatures (modsCCDred), spectral extraction, wavelength and flux calibration, and telluric correction were done with the IRAF apall task.

On UT 2013 March 19, we observed the source with the SpeX spectrograph (Rayner et al. 2003) on the NASA Infrared Telescope Facility (IRTF) for 32 minutes using an 0″.8-wide slit covering a wavelength range of 0.808–2.415 μm. The seeing was 1″, and sky conditions were clear. An A0 V star was observed within an air mass difference of 0.1 immediately after the object spectrum was obtained to correct for telluric absorption. The data were reduced using the Spextool software (Cushing et al. 2004), and the telluric correction was conducted following Vacca et al. (2003).

Figure 1 shows the combined optical through infrared spectrum of W2M J1042+1641. The near-infrared spectrum shows strong broad Hα and Hβ plus the narrow [O III] $\lambda\lambda$4959, 5007 doublet, while the optical spectrum shows narrow emission lines in permitted as well as forbidden species, securing a QSO redshift identification of $z = 2.517$. The blue curve represents an unreddened QSO spectrum, made out of the UV composite QSO template of Telfer et al. (2002) combined with the optical-to-near-infrared composite spectrum from Glikman et al. (2006), illustrating the large amount of UV light lost. We fit this curve to the spectrum following the technique outlined in Glikman et al. (2007) and find that a suitable fit can only be achieved if the rest-frame UV emission below 2275 Å ($\lambda_{obs} < 8000$ Å) is ignored. This best fit is achieved with a QSO template reddened by $E(B - V) = 0.68$ (corresponding to $A_V = 5.4$ mag in the QSO rest frame; red line).

The excess UV flux, blueward of $\sim$8000 Å, could be explained if the dust were placed close to the AGN, between the broad- and narrow-line-emitting regions. This interpretation is also consistent with a model for the UV spectrum of Mrk 231 (a nearby, dusty, luminous QSO in a merger) suggested by Veilleux et al. (2013) in which the broad-line region is reddened by a dusty and patchy outflowing gas. This model predicts a small "leakage" fraction of a few percent, which is consistent with a similar degree of leakage seen in the X-ray spectra of other red quasars (Glikman et al. 2017). A similar conclusion was reached by Assef et al. (2016) for a hot

---

[18] The source name is shortened to W2M J1042+1641 hereafter.
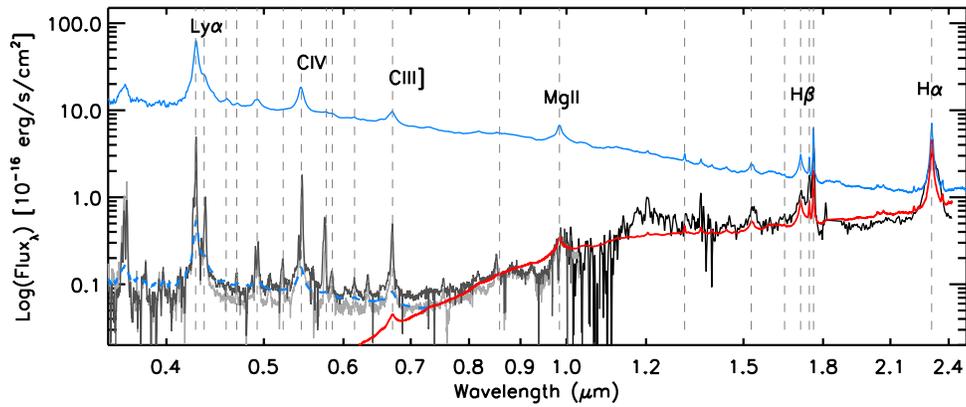
Figure 1. Optical through near-infrared spectrum of W2M J1042+1641, plotted on logarithmic wavelength and flux axes. The black line at $\lambda > 1\ \mu m$ is the near-infrared spectrum, showing broad Balmer line emission shifted to $z = 2.517$. The dark-gray line is the LBT optical spectrum, and the light-gray line is the LRIS spectrum (Section 2.4), showing the consistency between the two spectra taken 5 yr apart (observed frame). Vertical dashed lines mark the locations of strong emission lines seen in the spectrum. The red line is the best-fit QSO template (shown in blue) reddened by $E(B-V) = 0.68$ to the LBT spectrum combined with the near-infrared spectrum. In both cases, the rest-frame UV part of the spectrum deviates from the reddened template, implying that the obscuring dust is shielding most of the central region but allows ~0.8% of the intrinsic emission to enter our line of sight (blue dashed line).

dust-obscured galaxy (DOG) that displayed blue excess in its spectral energy distribution (SED). The authors arrive at leaked intrinsic QSO light through a patchy obscuring medium, or by reflection, as the best explanation.

We plot in Figure 1 the QSO template scaled to 0.8% of the intrinsic spectrum (with a dashed blue line) and find that it fits well the spectral shape. We note that the UV emission lines have a higher equivalent width but are narrower than the template, similar to "extremely red" QSOs in the Sloan Digital Sky Survey (SDSS) studied by Hamann et al. (2017). These arguments lead us to conclude that the dust is local to the lensed QSO and that the QSO is not reddened by the lens.

2.3. Hubble Imaging

We obtained Hubble Space Telescope (HST) imaging of W2M J1042+1641 with the WFC3/IR camera in Cycle 24 as part of a program to study the host galaxies of W2M red QSOs. We used the F160W and F125W filters, which were chosen to straddle the 4000 Å break. We observed the source over two visits, UT 2017 February 26 and UT 2017 May 7, covering both filters in a single orbit observation per visit. We observed our sources in MULTIACCUM mode using the STEP100 sampling, which is designed to provide a broad dynamic range while avoiding saturation. We performed a four-point box dither pattern with 400 (224) s at each position for the F160W (F125W) filter. We reduced the images using the `Drizzle-Pac` software package to a final pixel scale of $0\rlap{.}''06\ \mathrm{pixel}^{-1}$.

The top row of Figure 2 shows the reduced, color-combined images for the two HST visits, with the first visit shown on the left and the second visit shown on the right. The image reveals four point sources surrounding an extended-appearing source at the center, in a geometry suggestive of quadruply lensed system with a cusp configuration. The bottom row of Figure 2 shows the results of our profile fits. The top row labels the four lensed components A, B, C, D, in decreasing order of flux, as well as nearby galaxy G1, which we also modeled (see Section 3).

Figure 3 shows the image of the Einstein ring after the QSO components (A, B, C, D), modeled lens galaxy, and G1 have been subtracted. The F125W image is shown in the top row, and the second row shows the residuals once the $\sim 0\rlap{.}''9$-radius Einstein ring is also subtracted. The third and fourth rows show
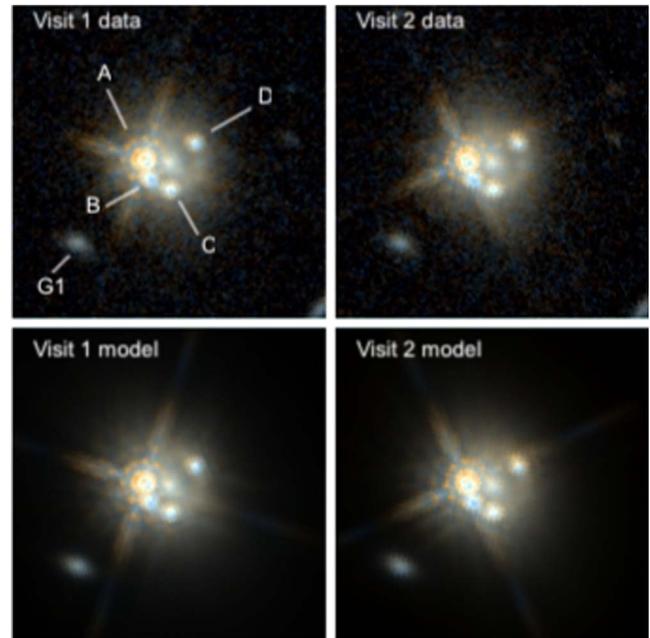

Figure 2. HST WFC3/IR F125W and F160W color-combined images of W2M J1042+1641 over two visits, along with output from a morphological analysis with `hostlens`. The top row shows the observed, drizzled image. The second row shows the best-fit model consisting of four PSFs, Sérsic profiles for the lensing galaxy (located in between the four PSFs) and G1, and the Einstein ring image of the source host. The modeled components are labeled and referenced in the text. The images are oriented with north pointing up and east to the left, although the two visits were each taken rotated by $47\rlap{.}°2$ with respect to each other. The scale is $10'' \times 10''$. Left: visit 1, UT 2017 February 26. Right: visit 2, UT 2017 May 7.

the same but for F160W. In Tables 2, 3, and 4 we list the resulting photometry, relative astrometry of each component, and morphological parameters of the lensing galaxy and its companion, respectively. All the HST data used in this paper can be found in MAST: doi:10.17909/rw0m-7191.

2.4. Follow-up Keck Spectroscopy along Multiple Position Angles

On UT 2018 March 19, we obtained a follow-up spectrum with the LRIS spectrograph on the Keck I telescope, orienting
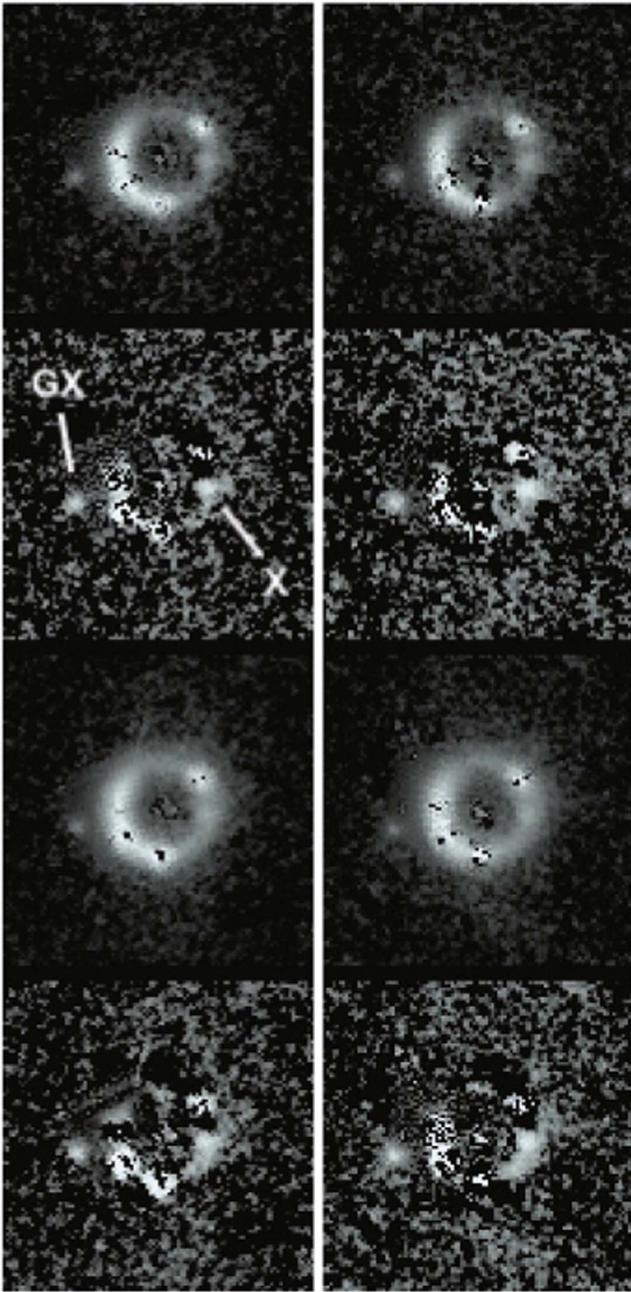
Figure 3. HST residual images of W2M J1042+1641 over two visits, after subtracting the best-fit morphological models determined by `hostlens` shown in the bottom panel of Figure 2. The top row shows the F125W image with the PSFs and Sérsic profiles subtracted, revealing a clear Einstein ring made up of the lensed QSO's host galaxy light. The second row shows the full residuals of the same fitted model for the F125W image. The third and fourth rows show the same but for F160W. The labeled components are the nearby object (GX), which we consider as a galaxy perturber associated with the lensing galaxy in our modeling, and a source that is likely associated with the QSO host galaxy (X). The images are scaled to $6'' \times 6''$ and oriented to match those in Figure 2. Left: visit 1, UT 2017 February 26. Right: visit 2, UT 2017 May 7.

the slit along different position angles (PAs), aiming to disentangle the emission from the different components. We placed a slit along the parallactic angle ($79°$) centered on the brightest component for two 600 s exposures. Another two 600 s exposures were taken with the slit placed along the A, B, and C components, at a PA of $41°.9$. Finally, a fifth 600 s

exposure was performed with a PA of $128°.2$ along components B and D including the lensing galaxy with the intention of identifying the redshift of the lens. Although the seeing was $\sim 1''$, precluding our ability to cleanly separate the different components along the position axis of the slit, the 2D spectrum is clearly extended beyond the width of a point-spread function (PSF). Specifically, the data taken at PA $= 128°.2$ show two clear lensed AGN components separated by $\sim 1''.6$; however, we detect no obvious signal from the lens itself.

The combined LRIS spectrum is shown in light gray in Figure 1. The best-fit reddened QSO template to the combined LRIS plus near-infrared spectrum, considering only $\lambda > 8000$ Å, finds $E(B - V) = 0.73$ (corresponding to $A_V = 5.8$ mag in the QSO rest frame). The LRIS and LBT spectra are remarkably similar, suggesting that not much has changed in this source between the two spectroscopic epochs, 5 yr apart in the observed frame, or 1.4 yr in the rest frame.

2.4.1. Lens Redshift

We identify absorption consistent with Mg II $\lambda\lambda 2796, 2803$ at $z = 0.5985$ (see Figure 4), which is in excellent agreement with the photometric redshift expected from the color of the lensing galaxy, and we thus adopt this as a tentative redshift for the lensing galaxy.[19]

2.5. VLA Radio Follow-up

We obtained radio data of W2M J1042+1641 with the Very Large Array (VLA; ID: 19A–430; PI: N. Secrest) on 2019 August 26 (epoch 1) and 2019 September 28 (epoch 2) at C band (6 cm) in A configuration, as part of a program to follow up quadruply lensed quasars with sensitive VLA observations. The two observations use the new 3-bit sampler, which gives a 4 GHz bandwidth, divided into 32 spectral windows, with 128 MHz bandwidth and 32 channels each using dual polarization. At the beginning of both epochs we observed J1331+3030 for the amplitude and bandpass calibration, while J1051+2119 was the phase-reference calibrator. The scans on the target were ~ 10 minutes each, which were interleaved by ~ 2-minute scans on the phase-reference calibrator. The total exposure time for each epoch was 90 minutes. The data were reduced with the Common Astronomy Software Application package (CASA; McMullin et al. 2007) following the standard calibration procedures (e.g., Spingola et al. 2020a, 2020b). We detected and CLEANed the target using natural weights. The signal-to-noise ratio was too low to perform self-calibration.

Only emission corresponding to images A and B is detected and resolved. The contour maps of two epochs are shown in Figure 5. The beam size in the first epoch is $0''.300 \times 0''.223$ at a PA of $52°.581$ (east of north), and the total integrated flux densities of images A and B were 51 ± 8 μJy and 13 ± 7 μJy, respectively; the off-source rms noise is 6.7 μJy beam^{-1}. In the second epoch, the beam size is $0''.293 \times 0''.204$ at a PA of $-49°.727$ (east of north). Here the flux densities increased by 50%, being 82 ± 9 μJy and 19 ± 5 μJy (images A and B, respectively). The off-source rms noise level was 4.6 μJy beam^{-1}. We consider the uncertainty due to calibration

---

[19] We used the `mag2mag` routine from Auger et al. (2009), available at https://github.com/tcollett/LensPop/tree/master/stellarpop/, to check that for a Coleman et al. (1980) E/S0 galaxy template, redshifted to $z = 0.599$, 19.19 mag in F160W corresponds to 19.58 mag in F125W, in excellent agreement with the lensing galaxy photometry we measured in Table 2.

Table 2
Photometry of W2M J1042+1641

Filter	A (mag)	B (mag)	C (mag)	D (mag)	G (mag)	G1 (mag)	GX (mag)	S (mag)
F125W (1; w/ GX)	18.26 ± 0.001	20.48 ± 0.03	21.13 ± 0.02	21.84 ± 0.02	19.57 ± 0.01	23.29 ± 0.03	25.43 ± 0.05	23.30 ± 0.02
F125W (1; w/o GX)	18.24 ± 0.0005	20.47 ± 0.03	21.09 ± 0.03	21.78 ± 0.02	19.53 ± 0.01	23.29 ± 0.02	25.45 ± 0.05	23.47 ± 0.02
F125W (2; w/ GX)	18.26 ± 0.002	20.46 ± 0.004	20.92 ± 0.01	21.95 ± 0.03	19.60 ± 0.01	23.23 ± 0.03	25.43 ± 0.05	23.53 ± 0.02
F125W (2; w/o GX)	18.25 ± 0.001	20.46 ± 0.01	20.90 ± 0.01	21.91 ± 0.03	19.56 ± 0.01	23.23 ± 0.02	25.30 ± 0.05	23.66 ± 0.02
F160W (1; w/ GX)	17.62 ± 0.0003	20.22 ± 0.04	20.66 ± 0.05	21.49 ± 0.02	19.19 ± 0.01	23.03 ± 0.02	25.22 ± 0.04	22.34 ± 0.02
F160W (1; w/o GX)	17.60 ± 0.0004	20.22 ± 0.03	20.60 ± 0.06	21.41 ± 0.02	19.18 ± 0.004	23.03 ± 0.01	25.44 ± 0.05	22.20 ± 0.01
F160W (2; w/ GX)	17.72 ± 0.001	20.24 ± 0.02	20.43 ± 0.01	21.37 ± 0.004	19.16 ± 0.005	22.97 ± 0.02	25.10 ± 0.04	22.38 ± 0.02
F160W (2; w/o GX)	17.70 ± 0.001	20.22 ± 0.01	20.39 ± 0.005	21.31 ± 0.01	19.19 ± 0.004	22.98 ± 0.01	25.42 ± 0.05	22.23 ± 0.01

Note. Photometry has been measured with `hostlens`. "S" stands for the best-fit magnitude of the delensed QSO host galaxy, for which a de Vaucouleurs profile is used. Visits: (1) UT 2017 February 26; (2): UT 2017 May 7. Here "w/ GX" stands for the lens mass model that accounts for GX as a perturber, whereas "w/o GX" stands for the mass model without a perturber. See Section 3.2 for details.

Table 3
Relative Astrometry of W2M J1042+1641

Component	Model with GX E →W (arcsec)	Model with GX S →N (arcsec)	Model without GX E →W (arcsec)	Model without GX S →N (arcsec)
A	0.000 ± 0.0004	0.000 ± 0.0004	0.000 ± 0.0004	0.000 ± 0.0004
B	0.151 ± 0.005	-0.561 ± 0.008	0.147 ± 0.006	-0.566 ± 0.006
C	0.812 ± 0.005	-0.911 ± 0.006	0.812 ± 0.004	-0.913 ± 0.006
D	1.590 ± 0.005	0.539 ± 0.009	1.593 ± 0.005	0.536 ± 0.009
G	0.777 ± 0.004	-0.079 ± 0.003	0.775 ± 0.002	-0.077 ± 0.003
G1	-2.140 ± 0.009	-2.642 ± 0.004	-2.140 ± 0.010	-2.643 ± 0.005
GX	-0.841 ± 0.029	-0.403 ± 0.031	-0.827 ± 0.024	-0.399 ± 0.021

Note. Similar to Table 2, "w/ GX" stands for the lens mass model that accounts for GX as a perturber, whereas "w/o GX" stands for the mass model without a perturber. We report the medians and the standard deviations of the values measured in the two filters, in both visits, relative to image A. For image A itself, we report representative MCMC uncertainties.

Table 4
Morphology of the Lens and Nearby Galaxy

| Object and Filter	n	R_e (arcsec)	b/a	PA (deg)
G F125W (w/ GX)	4.38 ± 0.02	1.08 ± 0.01	0.66 ± 0.00	27.40 ± 0.12
G F125W (w/o GX)	4.70 ± 0.03	1.17 ± 0.01	0.67 ± 0.00	26.24 ± 0.16
G F160W (w/ GX)	4.42 ± 0.02	1.00 ± 0.01	0.66 ± 0.00	$[27.40 \pm 0.12]$
G F160W (w/o GX)	4.74 ± 0.02	1.03 ± 0.01	0.66 ± 0.00	$[26.24 \pm 0.16]$
G1 F125W (w/ GX)	1.58 ± 0.09	0.22 ± 0.00	0.51 ± 0.02	67.56 ± 0.60
G1 F125W (w/o GX)	1.56 ± 0.10	0.23 ± 0.00	0.50 ± 0.01	68.05 ± 0.65
G1 F160W (w/ GX)	1.57 ± 0.08	0.24 ± 0.01	0.43 ± 0.01	$[67.56 \pm 0.60]$
G1 F160W (w/o GX)	1.53 ± 0.07	0.25 ± 0.01	0.41 ± 0.01	$[68.05 \pm 0.65]$

Note. Morphology has been measured with `hostlens`. Similar to Tables 2 and 3 "w/ GX" stands for the lens mass model that accounts for GX as a perturber, whereas "w/o GX" stands for the mass model without a perturber. Angles are positive E of N. The effective radius is measured along the semimajor axis. For each filter, the modeling was done by enforcing the match of each morphological parameter between the two observing epochs. The PA was also enforced to match between the two filters (represented here by the use of square brackets for F160W).

(estimated using the scatter on the amplitude gains) to be on the order of 10%. Finally, a 2D Gaussian fit using the task IMFIT to the second-epoch CLEANed image (because of the more robust detection) found that both lensed images are consistent with a point source. We discuss the impact of these data on our analysis in Section 4.1.
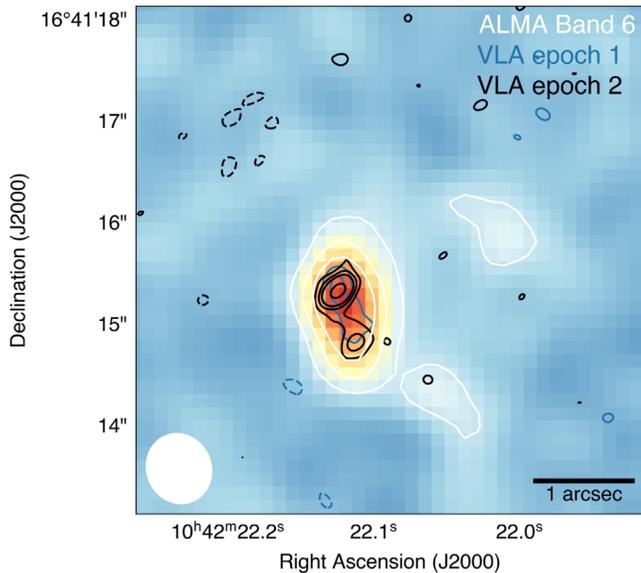
2.6. ALMA Millimeter Imaging

W2M J1042+1641 was observed with the Atacama Large Millimeter/submillimeter Array (ALMA) on 2019 October 30 under project code 2019.1.00964.S (PI: Stacey). The target data were correlated in four spectral windows centered on 247, 249, 262, and 264 GHz, each with 2 GHz bandwidth and 128 channels. One of the spectral windows covers the redshifted rest frequency of a CO line, not reported here. J1058+0133 was observed as a flux and spectral bandpass calibrator. J1045+1735 was used to correct time-dependent phase variations. The total integration time on target was 5 minutes. The data were calibrated using the ALMA pipeline within CASA, and the data were inspected to confirm the quality of the calibration. The continuum-only spectral channels were imaged and deconvolved using a Briggs
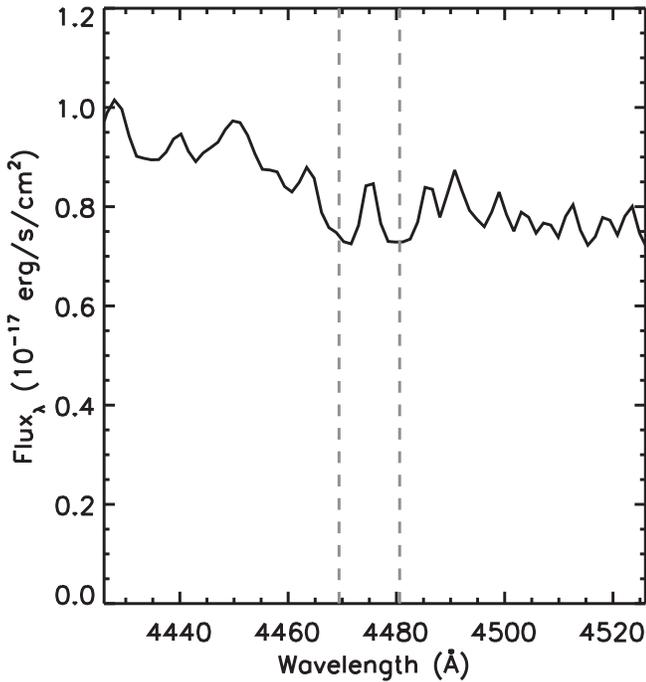
Figure 4. We identify a Mg II $\lambda\lambda2796$, 2803 absorption feature, marked by dashed vertical gray lines, at a redshift of $z = 0.5985$.

Figure 5. Radio (VLA) and submillimeter (ALMA) follow-up imaging. The blue image and black contours indicate the VLA emission at the first and second epochs, respectively. The white contours indicate the ALMA Band 6 emission. The contours are drawn at $(-3, 3, 6, 9, 18, 36, 72,$ and $144)$ times the off-source noise of each map, which is 6.7, 4.6, and 62 μJy beam^{-1} for the VLA epoch 1, VLA epoch 2, and ALMA observations, respectively. The restoring Gaussian beam of the ALMA observations is shown in white in the lower left corner and is $0\rlap{.}''64 \times 0\rlap{.}''57$ with a PA of 23°.

weighting of the visibility data, resulting in a synthesized beam of $0\rlap{.}''64 \times 0\rlap{.}''57$ and an rms noise of 66 μJy. The task IMFIT within CASA was used to fit a PSF to each lensed image. No significant residuals remain after fitting, suggesting that the lensed images are not resolved.

We overplot in Figure 5 flux density measurements of the rest-frame 330 μm continuum, obtained with ALMA. All four

QSO images are detected, with flux densities A = 835 ± 66 μJy, B = 519 ± 66 μJy, C = 296 ± 66 μJy, and D = 281 ± 66 μJy, consistent with thermal dust emission (e.g., Stacey et al. 2018). We find no significant evidence of misalignment between the radio (VLA) and submillimeter (ALMA) emissions, indicating a cospatial origin. Table 5 lists the positions and flux densities of the radio and millimeter sources.

3. Modeling of the System

The Einstein ring shown in Figure 3 is relatively bright, and therefore any morphological fitting of the system that does not account for it may bias the quantities of interest: the relative astrometry and photometry of each light source, as well as the morphology of the lensing galaxy. We therefore chose to fit the system using hostlens (Rusu et al. 2016), which incorporates, along with the other morphological components, a model of the Einstein ring as an analytical Sérsic profile concentric with the QSO light, lensed through a lensing mass model and convolved with the PSF. We focused on a $16'' \times 19''$ cutout around the system starting with a newly constructed PSF (following the method described in Glikman et al. 2015, which involved combining a few dozen bright stars in each HST filter) for each filter. The images from the two visits were taken at different angles, making their combined-image PSF difficult to model; we therefore model these independently, although some of the parameters are treated as coupled, as we describe in the next section.

3.1. Coupled Light and Lens Modeling

There are three reasons why a naive, direct modeling of this system with hostlens would be suboptimal: (1) Employing hostlens with a single smooth lens mass model (and without multiple mass substructures whose parameters are highly degenerate) cannot account for the flux ratio anomalies found in this system (see Section 3.2); this is because, for a given mass model, hostlens does not allow one to arbitrarily change the flux ratios of the images, which are determined by the relative position of the source with respect to the lens. (2) The PSFs, although carefully constructed in Section 3, were found to produce significant residuals when used to fit the QSO images, particularly the bright image A. (3) hostlens can only model a single input image cutout at a time. But given that we have images from two HST visits in two filters, it is desirable to model the system using the joint information from the cutouts of all available data, in order to better constrain some of the morphological parameters we derive.

To tackle the issues above, we wrote a custom wrapper code around hostlens, which uses an iterative approach:

1. We first run hostlens without a lensing model, fitting the four QSO images as point sources characterized by their positions and fluxes; the light of the lensing galaxy G and that of the nearby galaxy G1 are fitted with one Sérsic profile each, convolved with our original PSFs, in each of the two visits and two filters (four cutouts). During the fitting, our wrapper performs the parameter optimization by minimizing the sum of the quality-of-fit χ^2 reported by hostlens for the four cutouts, using the Nelder−Mead algorithm (Gao & Han 2012). The sky pedestals, relative astrometry, and fluxes of each

Table 5
Radio and Millimeter Source Positions and Flux Densities

Source	R.A. (J2000)	Decl. (J2000)	Flux Density (μJy)
	JVLA Epoch 1		
A	$10:42:22.1245 \pm 0.0013$	$+16:41:15.3127 \pm 0.0254$	51 ± 5
B	$10:42:22.111 \pm 0.022$	$+16:41:14.942 \pm 0.169$	13 ± 1
	JVLA Epoch 2		
A	$10:42:22.1252 \pm 0.0005$	$+16:41:15.3371 \pm 0.0073$	82 ± 8
B	$10:42:22.1113 \pm 0.0021$	$+16:41:14.830 \pm 0.0294$	19 ± 2
	ALMA		
A	$10:42:22.1208 \pm 0.0013$	$+16:41:15.3465 \pm 0.0225$	835 ± 66
B	$10:42:22.1183 \pm 0.0021$	$+16:41:14.8356 \pm 0.0362$	519 ± 66
C	$10:42:22.0441 \pm 0.0037$	$+16:41:14.3003 \pm 0.0634$	296 ± 66
D	$10:42:22.0204 \pm 0.0039$	$+16:41:16.1130 \pm 0.0669$	281 ± 66

light component are optimized independently for each cutout.[20] For each cutout, we constrain the orientations of the two Sérsic profiles, accounting for the different rotation angles of the two visits. Their ellipticity, effective radius, and Sérsic index are constrained to matching values in the two visits, but not in the two filters.

2. We fit a shared lens mass model (details are provided in Section 3.2) with `glafic` (Oguri 2010) using the relative astrometry we derived in the previous step. `glafic` solves the lens equation and computes the lensed point-source images using an adaptive grid algorithm; it compares these positions to the observed ones via χ^2 minimization, in order to optimize the mass model. We then repeat the optimization from the previous step, but also fitting for the extended QSO host galaxy, responsible for the Einstein ring, with a circular Sérsic profile (we found that allowing for ellipticity did not improve the fit), lensed through the lens mass model. We fixed the Sérsic index to the fiducial value for an early-type galaxy, $n = 4$ (de Vaucouleurs 1948), as we found that this parameter would otherwise diverge to large values ($n > 10$) without producing visually improved residuals or modifying significantly the photometry of the other components. The fluxes of the host are fitted independently to the four cutouts,[21] while the effective radius can vary between the filters but not between visits.

3. At this point, we found that if we optimize for all light components at the same time, the parameters of the Sérsic profiles of the lens and QSO host galaxies are affected by the significant residuals at the location of the QSO image. We therefore hold fixed the best-fit models of the QSO

images and mask them using circular masks of $\sim 0{.}''3$ in radius.[22] Next, we optimize the parameters of the Sérsic profiles and also the astrometry and photometry of object GX (see Figure 3), which we model as a point source.[23] We then hold fixed the parameters of the profiles we just fitted at this step, remove the masks, and fit again for the QSO images, to allow them to adjust in response to the profiles mentioned above.

4. We follow the approach in Chen et al. (2016), developed for the analysis of gravitational lenses observed with adaptive optics, where the PSFs are a priori unknown and must be derived from the data. This approach improves the PSF of the four cutouts under the assumption that the PSF should not vary among the QSO images.

5. We now proceed in an iterative fashion, where we first refit the parameters of the shared lens mass model with `glafic` using the improved relative astrometry[24] from step 3 and repeating steps 3 and 4. We do this in 30 steps, where the PSF correction box size is increased from 7 to 35 pixels on a side, by 2 pixels every second step. The gradual increase is adopted in order to improve the convergence of the PSF correction, by preventing it from being dominated by noise outside the core of the PSF. Figure 3 shows the residuals after the final iteration.[25] Following the iterations, the χ^2 is much improved, whether we measure it by first masking the pixels corresponding to the QSO images cores or not.

Once we have inferred the best-fit profile parameters as described above, we determine the corresponding uncertainties by combining five independent MCMC chains of 15,000–30,000 steps. We ensure their convergence by monitoring the change in the parameter values over time and removing the "burn-in"

---

[20] The reason we do not enforce that the relative positions of each component match between cutouts is that the uncertainties on these positions (especially that of the position of the lensing galaxy) have a dominant effect on the best-fit lens mass models we derive in the next steps, and we found that Markov Chain Monte Carlo (MCMC) approaches to determine this uncertainty can significantly underestimate it. We therefore prefer to use the scatter in relative astrometry between the four cutouts as a measure of uncertainty.

[21] While we do not expect the flux of the host to vary between the two visits, we nevertheless obtained a better fit (fewer residuals) by allowing the host flux to be a free parameter between visits. The final difference is up to ~ 0.2 mag (see Table 2).

[22] We also mask the luminous blob on top of the Einstein ring, which we describe in Appendix C.

[23] We attempted to fit GX with a Sérsic profile, but we measured a vanishingly small effective radius $\lesssim 0{.}''01$. We therefore consider this component to be unresolved.

[24] We checked that the difference between the positions predicted by the lens model and the ones actually measured stays within a fraction of a pixel size for the four cutouts.

[25] The shape of the PSF correction box may be seen around image A. As this image is much brighter than the other ones, it has more weight in the improved PSF.

steps. Due to our modeling approach, we need to run MCMC separately, first for the QSO images and then for the other profiles.[26] While our reconstructed PSF is superior to the original one, it is not perfect, as shot noise is present in the core of the bright point sources. If we integrate the residual flux (positive or negative) in the pixels corresponding to the core of each of the point sources, it is not exactly zero for a given point source. Therefore, for the photometry of the QSO images and the lensing galaxy reported in Table 2, we add to the uncertainties the contribution of this residual flux.

3.2. Lensing Analysis

The relative astrometries of the four QSO images and of the lensing galaxy, from the HST data, provide the most robust constraints to determining a gravitational lens model for W2M J1042+1641. The procedure outlined in Section 3.1 is applied to two different lensing models, which we will explain later in this section. However, we will first analyze the "definitive" lens models constructed using the relative astrometry obtained by the iterative modeling described in the previous section and listed in Table 3. This will serve to motivate the choice of the two lens models mentioned above.

A commonly used model to fit gravitationally lensed QSOs, when the main constraints are astrometric, is the singular isothermal ellipsoid with external shear (SIE+γ). In this model the "strength" of the lens is characterized by a velocity dispersion, expected to be close to the central velocity dispersion of the stars in the lensing galaxy (e.g., Kochanek 1994).[27] This is one of two types of lens models we used in Section 3.1 to fit the imaging data, and we report its best-fit parameters in Table 6. However, this model results in a statistically very poor fit with $\chi^2 \sim 62.5$ for a single degree of freedom. The reason for the poor fit is that the model is unable to reproduce the observed locations of the QSO images. If we remove the constraint on the lens location, we obtain a perfect fit, although the model becomes underconstrained. We refer to this model as "free SIE+γ" in Table 6.

Inspired by the work on lens galaxy environments by Sluse et al. (2012), we next looked at the nearby environment of the system for clues that might explain the poor fit of our SIE+γ model. Figures 2 and 3 reveal two structures near the lensing galaxy G: galaxy G1, located $3\farcs90$ from G, and GX, a structure much fainter but closer to the system ($1\farcs67$ from G and $0\farcs94$ from A). Including G1 in the fit as a second singular isothermal sphere (SIS) does not result in a significant improvement. In fact, its impact on the model based on its luminosity compared to that of the elliptical lensing galaxy and scaled by the Faber–Jackson law (Faber & Jackson 1976) is negligible, and is expected to be even smaller in reality, since it has a Sérsic index of $n \simeq 1.5$ suggestive of spiral morphology (see Table 4). On the other hand, GX is a compact object whose morphology we are unable to resolve, but whose existence as a real object as opposed to a PSF artifact is validated by its presence at the same location in both filters and both visits. If we include it in the fit as an SIS at the observed location and at the redshift of

---

Table 6
Lensing Mass Models

Parameter	Free SIE+γ	SIE+γ	SIE+γ+GX
σ_G	$222.0^{+0.9}_{-0.6}$	229.0 ± 0.8	$223.2^{+1.3}_{-1.1}$
$e = 1 - b/a$	$0.17^{+0.14}_{-0.08}$	0.55 ± 0.02	$0.43^{+0.04}_{-0.03}$
θ_e	$49.2^{+8.7}_{-3.8}$	28.0 ± 1.0	$24.9^{+1.8}_{-2.2}$
γ	0.03 ± 0.01	0.15 ± 0.01	$0.08^{+0.02}_{-0.01}$
θ_γ	$-5.7^{+25.0}_{-23.7}$	-66.7 ± 1.5	$-75.5^{+5.3}_{-4.3}$
σ_{GX}	$-$	$-$	$46.5^{+4.2}_{-3.9}$
Δt (days)	$12.9^{+7.4}_{-3.9}$	26.1 ± 1.1	22.7 ± 1.5
χ^2/ν	$0/-1$	$62.5/1$	$0/0$

Note. Based on fitting with glafic. Uncertainties are determined from five MCMC runs with 10^5–10^6 steps each, for each mass model. We assume astrometric errors of $0\farcs001$ for image A, and for the other images we use the errors reported in Table 3. (For the SIE+γ model we use uncertainties 10 times smaller than measured. Otherwise, we find that in order to match the observed position of the lensing galaxy, the predicted QSO images deviate too much from the observed positions. As a consequence, although the reported χ^2 for this model was computed using the observed error bars, the error bars on the parameters may be artificially small.) Both G and GX are assumed to be at $z_l = 0.599$ (see Section 2.4.1), and the source is fixed at $z_s = 2.517$. Angles are positive E of N. ν is the number of degrees of freedom. Image D leads, and all other images have similar time delays with respect to it, with differences of $\lesssim 1$ day.
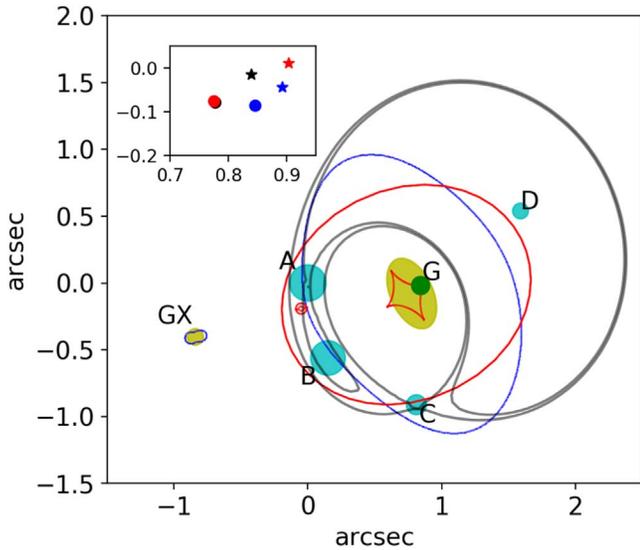
---

the lens, with a velocity dispersion free to vary, we obtain a perfect fit for zero degrees of freedom (see Table 6). This is the second type of model we used for the iterative fitting in Section 3.1.

We note that, in a program to study a sample of 30 lensed QSOs with HST (Cycle 26, Program ID 15652; PI: Treu), W2M J1042+1641 was imaged in the F475X and F814W filters with WFC3/UVIS. Schmidt et al. (2023) modeled this system as part of the sample using an automated pipeline using the LENSTRONOMY software package (Birrer & Amara 2018) with a power-law elliptical mass distribution plus external shear. The fitting was done simultaneously for the two UVIS bands plus the F160W data from this work, and the F125W data were not used. The astrometric positions derived for the QSO components differ from what we find here by up to $\sim$40 mas, which is still at the subpixel level but results in a different lens model. The automated nature of this analysis does not include the finer structures such as sources GX and X; in addition, the reported astrometry is fine-tuned to the specific choice of lensing model (private communication), making a direct comparison to our results infeasible.

3.3. The Fiducial Model

In addition to the quality of the fit, we list the following arguments as to why the SIE+γ+GX model is more realistic:

1. While the velocity dispersion of GX was a free parameter during the fit, its best-fit value of $\sim$46.5 km s^{-1} is in good agreement with the predicted value of $\sim$50 km s^{-1} based on its relative luminosity compared to G and the Faber–Jackson law. Though we can only estimate a rough photometric redshift for GX based on a single color, it is in good agreement with our best estimate for G (Section 2.4.1), but only if we assume a late-type galaxy spectral template (GX appears bluer than G in Figure 3).

---

[26] An alternative modeling approach that can model all profiles at the same time is presented, e.g., in Wong et al. (2017). It works by rescaling the weights of the pixels corresponding to large residuals.

[27] The other parameters of the model are (1–2) the coordinates of the lensing galaxy; (3–4) the coordinates of the source; (5–6) the orientation of the major axis of the ellipsoid (projected on the plane of the sky), as well as its axis ratio; and (7–8) the orientation and strength of the external shear.

Figure 6. Lensing configuration for the SIE+γ+GX model. The two lenses (G and GX) are marked with yellow ellipses whose sizes scale in proportion to the modeled velocity dispersions. The ellipsis for G shows the orientation and axis ratio predicted by the mass model. Critical lines are drawn in blue and caustics in red. The green circle marks the location of the source QSO in the source plane, and the cyan circles mark the locations of the observed images. The size of the circles is in proportion to the predicted flux ratios. Time delay surface isochrones at the location of the saddle point images A and C are drawn in gray. In the inset, the source QSO position (stars) and lens position (circles) are marked for the SIE+γ+GX model (black), the SIE+γ model (red), and the free SIE+γ model (blue). The black circle coincides with the light profile centroid of the lensing galaxy.

2. Previous lensing studies find that there is a good alignment between the axes of the light and mass distributions in lensing galaxies, within ∼10° (e.g., Shajib et al. 2019; Keeton et al. 1998; Sluse et al. 2012). We find that when GX is added to the model, the mass and light profiles of the main lens G do indeed show excellent alignment (the light PA for G in Table 4 is ∼27° and the mass profile $\theta_e = 24°.9$ in Table 6), while the models without GX show misalignment by ∼20° ($\theta_e = 45°–49°$ in Table 6). This supports the SIE+γ and the SIE+γ+GX models over the free SIE+γ model.

3. The best-fit location of the lensing galaxy in the free SIE+γ model is 0″072 away from the center of the observed light profile, toward the west (see inset in Figure 6). However, previous work by Yoo et al. (2006) found that typical offsets between the mass and light centroids are on the order of a few mas. This result was further confirmed by Sluse et al. (2012), who find similar values for 11/14 quads, once they include nearby luminous perturbers in their models.[28] As with the previous point, this supports the SIE+γ and the SIE+γ+GX models over the free SIE+γ model. It should be noted, however, that not only is the SIE+γ a poor fit, but its parameters are a more extreme version of those of the SIE+γ+GX model. In particular, from Table 6, both the ellipticity of the lens (e) and the value of the shear (γ) have increased, and their directions are within 5° of being perpendicular, a potential sign of degeneracies in the model.

4. Like any cusp configuration in a smooth lensing mass model, the central image (in this case B) is expected to have a flux equal to the sum of the two surrounding images (A and C; e.g., Keeton et al. 2003). Depending on which filter/visit is used to measure the observed flux, this makes the observed flux of image A, in particular, anomalous by more than an order of magnitude compared to the SIE+γ model (see Section 4.1). Such an anomalous flux ratio exceeds the largest previously recorded anomaly in the optical, for SDSS J0924+0219 (Inada et al. 2003).[29] The addition of GX boosts the predicted flux of image A in relation to B by a factor of ∼2 compared to the free SIE+γ model and by a smaller amount compared to the SIE+γ model, thus mitigating some of the discrepancy (see Section 4.1 for a comprehensive analysis of the flux ratios).

We therefore adopt the SIE+γ+GX model as our fiducial model, which we used to fit the morphology in Figures 2 and 3, and for which we plot the image configuration, critical lines, caustics, and time delay surface in Figure 6. We show the equivalent of Figures 2 and 3, but employing the free SIE+γ model, in Appendix A. We note that we do not find the difference in the residuals after fitting these two models with hostlens to be large enough to rule out one of the models, so our choice of the fiducial model is based entirely on the four arguments given above.

4. Discussion

With our fiducial (SIE+γ+GX) model in hand, we explore in this section the unique properties of this system. We investigate possible explanations for the flux anomaly seen among the QSO components. We also revisit the nature of this object as a red QSO and study its BH and host galaxy properties in the context of their coevolution.

4.1. Flux Ratio Analysis and Total Magnification

A crucial quantity needed for the purpose of characterizing the physical properties of the source QSO is its total magnification. However, due to the large flux anomalies present in this system, this quantity is not trivial to determine. We show these anomalies in Figure 7, by comparing the measured flux ratios with the histograms of the ratios of image magnifications, corresponding to the lens models from Table 6.[30] The fiducial lens model, SIE+γ+GX, is shown in the left column of Figure 7, the middle column shows the SIE+γ model, and the right column shows the free SIE+γ model. As we noted in Section 3.2, all flux ratios related to image A are highly anomalous. To compute a robust magnification, we require at

---

[28] We note that Shajib et al. (2019) find offsets larger by about a magnitude in their modeling of 13 quads with HST imaging; however, their exploratory models do not account for potential nearby perturbers.

[29] In the case of SDSS J0924+0219, the anomalous flux is suppressed rather than enhanced, in contrast to the present case.

[30] The magnification histograms for each image, which go into the flux ratios in this figure, as well as into the subsequent discussion on the total magnification, were computed using the MCMC chains used in Table 6, which assumed a point source. However, we can check that they hold for a realistic size of the accretion disk as follows. The standard deviation of the source position with respect to the lens position, obtained from MCMC, is ∼12 mas (for the SIE+γ+GX mas model). At the source redshift, this corresponds to ∼100 pc. However, the accretion disk size we estimate in Section 4.1.2 is ∼100 lt-day, in agreement with estimates of the size of the broad-line region in the literature (e.g., Gravity Collaboration et al. 2018). Since this is on the order of 1000 times smaller than the measured standard deviation of the source position, we conclude that the physical size of the source has no effect on the magnifications and flux ratios.
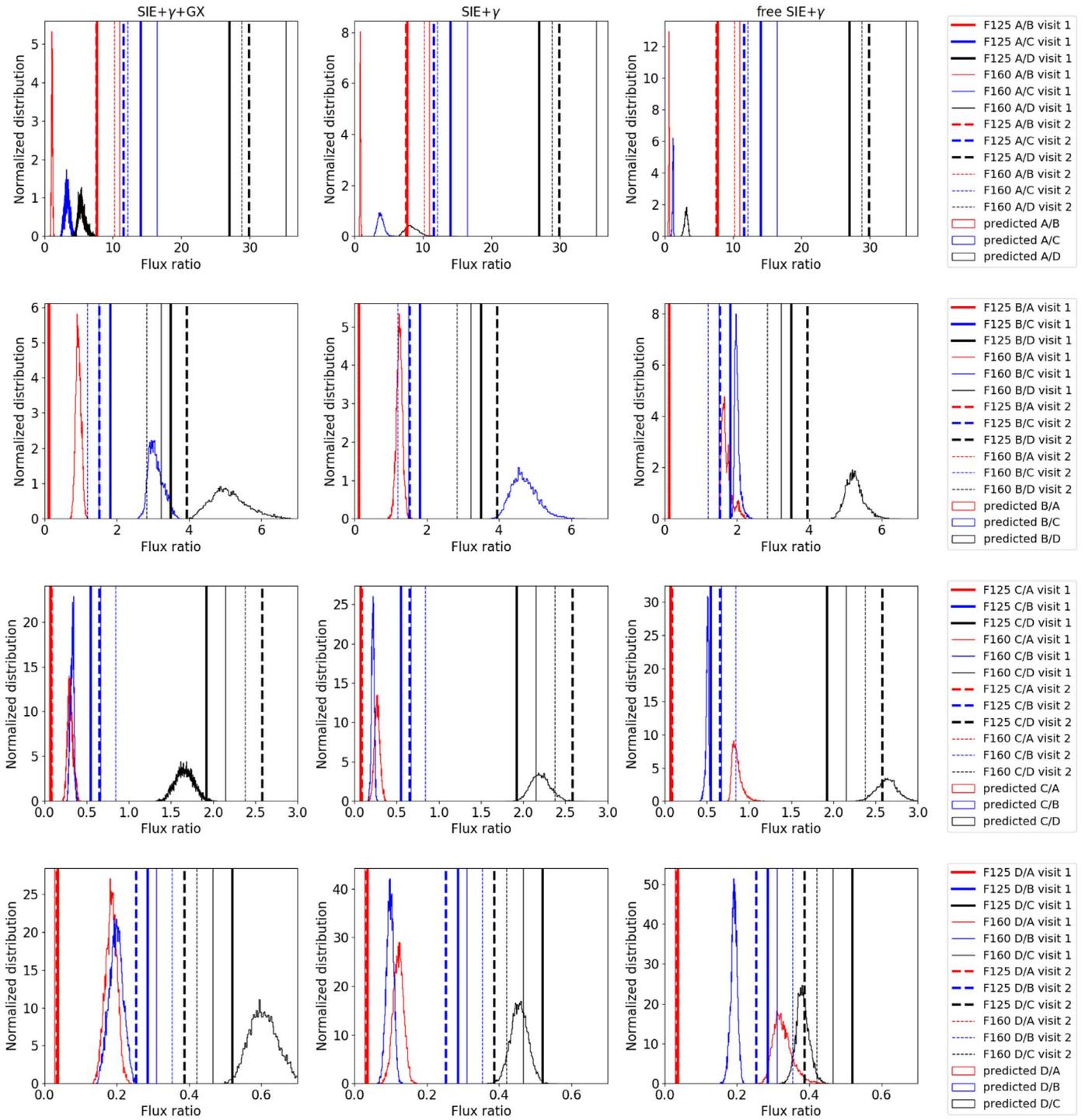
Figure 7. Comparison of observed and predicted flux ratios. The lens mass models, from left to right, are SIE+γ+GX, SIE+γ, and free SIE+γ. The histograms consist of 10,000 points from the MCMC chains computed with `glafic`, and the vertical lines correspond to the photometry measured by `hostlens`. Visits: (1) UT 2017 February 26; (2): UT 2017 May 7. The fluxes of image D in visit 2 may be unreliable owing to a diffraction spike from image A falling on top of it.

least one match between an observed and a model-predicted flux ratio.

The least anomalous flux ratios are those involving images C and D. The observation–prediction overlap in Figure 7 is small for the fiducial model but larger for the free SIE+γ model and even larger for the SIE+γ model. We note, however, the following caveats of the observed flux ratios: First, Figure 2 shows that in both filters of visit 2 image D is covered by a diffraction spike of image A. It is difficult to assess what level

of systematics this may introduce for the photometry of image D, reported in Table 2. A clue that the photometry of D might be problematic is that Figure 8 shows different directions of variation for the magnitude of image D in the two filters, as opposed to image C, which varies in a single direction, and also against expectations if the change was caused by microlensing or intrinsic variability. By comparing with the direction of variation of image C, the flux of D in F125W visit 2 is more likely to be affected by systematics. Second, both microlensing
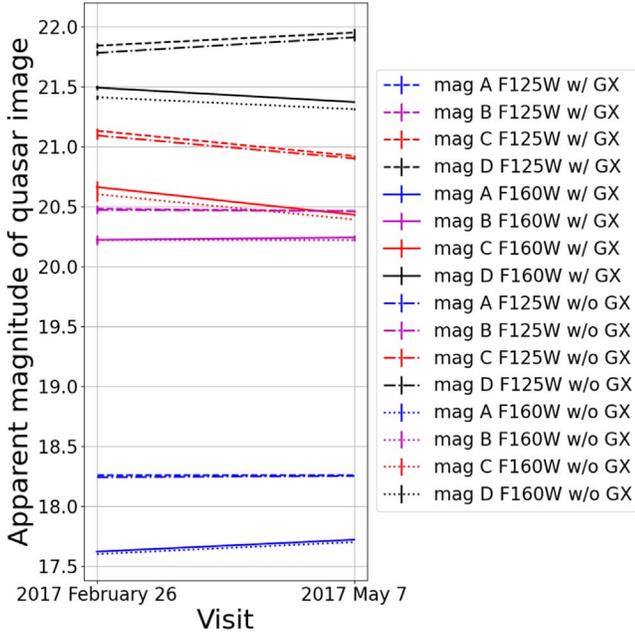
Figure 8. Variation of the QSO image magnitudes with observing epoch.

and intrinsic variability would cause variations of smaller amplitude at longer wavelengths, so the flux ratios in F160W should give a better estimate of the intrinsic flux ratio.

Based on the arguments we presented in Section 3, we construct an "anchor" by using the fiducial lens model, in spite of the fact that the SIE+γ model shows a better match, and intersecting the C/D model prediction distribution with a Gaussian fit to the observed fluxes in filter F160W only.[31] We checked that, although this intersection samples from the tail of the C/D model prediction distribution, the resulting samples correspond overall to the distribution of χ^2 values of the entire MCMC chain we used in Table 6, and therefore not to particularly poor fits. Finally, the total magnification we compute, i.e., integrated over the flux of the four images, accounting for flux ratio anomalies, is[32]

$$\mu_{\text{total}}^{\text{anomalous A,B}} = \text{obs}_{A/C} \times |\mu_{C,C/D\cap}| + \text{obs}_{B/C} \times |\mu_{C,C/D\cap}|$$
$$+ |\mu_{C,C/D\cap}| + \mu_{D,C/D\cap} = 117^{+22}_{-14}. \quad (1)$$

Here obs refers to the observed flux ratios and μ is the model-predicted magnification for each individual image. C/D∩ denotes that the magnifications are subject to the intersection condition introduced above.

The total magnification computed above is what we would use to study the physical properties of this system based on its observed, unresolved, longer-wavelength data, if we expect

---

[31] We have a direct measurement of C/D = 1.05 ± 0.34 from ALMA imaging of host galaxy thermal dust emission in Section 2.6, which is lower by a factor of ~2 than the HST-derived value, as well as the lens model predictions. However, this submillimeter emission is extended (see Figure 5) and expected to be primarily associated with dust in the QSO host galaxy (see Section 4.1.4; Stacey et al. 2018, 2021). These characteristics make the ALMA-based measurement unsuitable to use as an anchor.

[32] If we include the observed flux ratios in F125W when computing the intersection, we obtain 107^{+15}_{-13}. If we compute $\text{obs}_{A/D} \times |\mu_{D,C/D\cap}| + \text{obs}_{B/D} \times |\mu_{D,C/D\cap}| + |\mu_{C,C/D\cap}| + \mu_{D,C/D\cap}$ instead, we obtain 145^{+16}_{-20}. Alternatively, we get $obs_{A/B} \times \mu_{B,B/C\cap} + \mu_{B,B/C\cap} + \mu_{C,B/C\cap} + obs_{D/B} \times \mu_{B,B/C\cap} = 221^{+23}_{-21}$. Finally, alternatively, if we use the lens model free SIE+γ, we obtain 196^{+90}_{-57}.

---

that the flux anomalies we see in the HST images still hold at those wavelengths and the moment in time when those data were collected. On the other hand, if images A and B were not anomalous, the fiducial model predicts a total magnification of[33,34]

$$\mu_{\text{total}}^{\text{predicted A,B}} = (|\mu_A| + \mu_B + |\mu_C| + \mu_D)_{C/D\cap} = 52^{+5}_{-4}. \quad (2)$$

This is what we must use if we expect the flux anomalies to disappear at longer wavelengths. In the following sections, we explore one by one the three physical phenomena known to be responsible for flux anomalies in general.

4.1.1. Extinction

In addition to the intrinsic reddening that the QSO experiences from dust in its own host galaxy, discussed in Section 2.2, we consider the possibility that the different lensed lines of sight may be reddened as well. To date, the largest sample of lensed QSOs (23 systems) used to study the extinction properties in these systems remains the one of Falco et al. (1999). The median differential extinction was found to be $E(B - V) = 0.04$, the median total extinction $E(B - V) = 0.08$, and the median R_V in particular, for the early-type sample, was consistent with the Galactic value of 3.1. The consistency with the Galactic extinction parameter was later confirmed by Elíasdóttir et al. (2006) and Østman et al. (2008). Using the extinction curve from Cardelli et al. (1989),[35] for the assumed lens redshift in Section 2.4.1, we find a median differential extinction of 0.08 mag in F125W and 0.05 mag in F160W and a median total extinction of 0.15 mag in F125W and 0.10 mag in F160W. The differential extinction is too small to significantly impact our measured flux ratios, and accounting for total extinction would only correct the inferred total magnification upward by 10%–15%. Because these median values are small and their parent distributions are broad, we do not implement extinction corrections.

We note that the lens galaxy in W2M J1042+1641 is an early-type galaxy and thus is expected to be dust- and gas-poor, hence producing smaller extinction. While Falco et al. (1999) do not find a correlation with the impact parameter, we report that the effective radius is ~1″, and the QSO images are located at an impact parameter between 0.″78 and 1.″02, with images C and D located farthest away. Finally, we note that Peng et al. (2006) also found that extinction is small enough to ignore in their study of the BH–QSO host coevolution from a sample of lensed QSOs.

4.1.2. Microlensing Analysis, Intrinsic Variability, and Time Delays

We conducted microlensing simulations in order to investigate the plausibility of microlensing as the reason for the large flux anomaly of image A. We report the relevant values of convergence and shear at the location of each image in Table 7, for all three best-fit lens mass models. We show the details of the computation, for which we employ the BH properties derived in Section 4.3, in Appendix B. We show our results for the SIE+γ +GX and SIE+γ mass models in Figure 9. In the left panel, we show for each of the four images the microlensing magnification

---

[33] Removing the intersection constraint, the total magnification is $49.0^{+3.7}_{-3.1}$.

[34] Following the argument in Appendix C, the uncertainty on the radial slope of the host also introduces an additional uncertainty of $\sigma_\mu/\mu \sim 0.24$ on the total magnification, here as well as in the paragraph above.

[35] We perform the calculation using `extinction` (https://extinction.readthedocs.io).
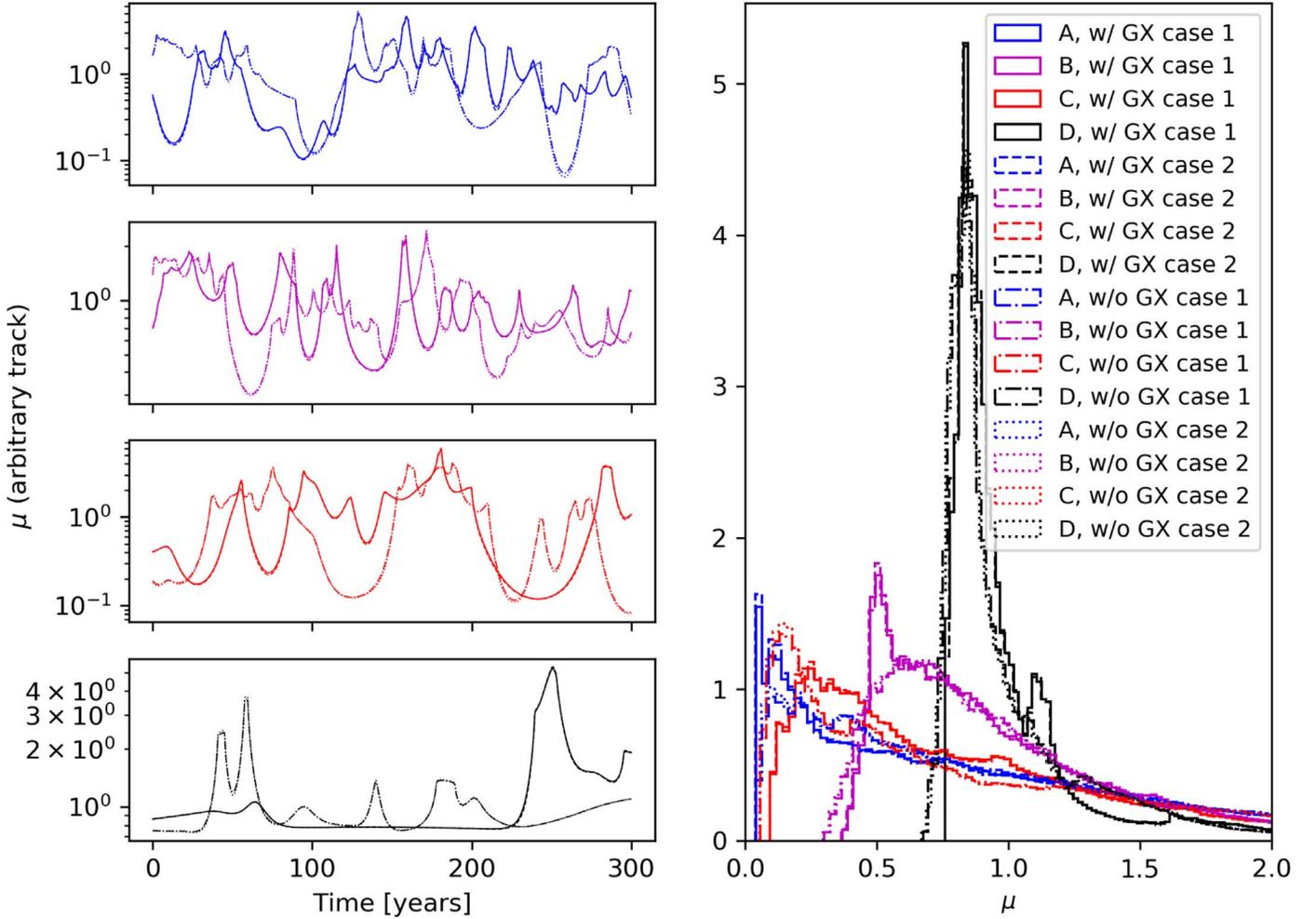
Figure 9. Simulated microlensing light curves in F160W. Left: magnification curves along an arbitrary track of the magnification maps of the four QSO images. The magnification of each image in the absence of microlensing has been normalized to unity by dividing over the magnifications predicted by two best-fit lensing models, SIE+γ+GX and SIE+γ with free lens position. The timescale is chosen to show the typical frequency of microlensing events for each image. Right: histograms of the microlensing magnification over the entire microlensing map. The histograms continue to decrease smoothly beyond the limit displayed. Case 1: $\log M_{\rm BH}[M_\odot] = 9.06$, $L/L_{\rm Edd} = 0.21$, $\eta = 0.43$, $R_0 = 2.76e+15$ m; Case 2: $\log M_{\rm BH}[M_\odot] = 9.56$, $L/L_{\rm Edd} = 0.046$, $\eta = 1.23$, $R_0 = 2.52e+15$ m. Both cases use $1.0 R_0$, $i = 60°$, and PA $= 90°$. See Section B for further details.

Table 7
Convergence and Shear at the Location of Each QSO Image

Image	Free SIE+γ		SIE+γ		SIE+γ+GX		
	κ	γ	κ	γ	κ	γ	$\kappa_\star$
A	0.536	0.514	0.465	0.563	0.496	0.551	0.053
B	0.490	0.480	0.400	0.543	0.454	0.493	0.045
C	0.517	0.540	0.641	0.534	0.639	0.545	0.078
D	0.425	0.418	0.310	0.455	0.343	0.402	0.028

Note. To compute $\kappa_\star$ we first modeled the system with a de Vaucouleurs (de Vaucouleurs 1948) + γ + GX mass profile using as priors the morphological parameters from Table 4. We then reduced the mass in the best-fit de Vaucouleurs profile to a stellar fraction of 0.2 (Dai et al. 2010) and computed the resulting convergence at the location of the QSO images.

for a random microlens track relative to the source image, where the base value of zero corresponds to the magnification value of the fiducial macrolens mass model. We plot two of our simulated cases, corresponding to two different sets of physical properties ($M_{\rm BH}$, $L/L_{\rm Edd}$) that we derive in Section 4.3. In the right panel we show the histograms of the magnifications over all simulated tracks. From these histograms we find that images A and C have

the largest probability of being affected by microlensing, as they correspond to the broadest histograms, though most of the time they would be demagnified, being saddle points of the time delay surface, in agreement with, e.g., Schechter & Wambsganss (2002). On the other hand, the flux of image D is the least impacted by microlensing. However, the probability that image A is magnified by a factor as large as required to explain the observed anomaly is only $0.22^{+0.55}_{-0.18}\%$ (for the SIE+γ+GX model, in case 2 from the plot; $0.17^{+0.48}_{-0.14}\%$ for case 1. For the SIE+γ model the probability is even smaller). We therefore conclude that while microlensing can explain part of the observed anomaly, it is highly unlikely to explain all of it. Flux changes this large would take on the order of a decade to complete. Finally, we note that while the integrated spectrum does show an enhancement in the flux in the blue, we have argued in Section 2.2 that this can be explained by factors intrinsic to the QSO and does not require chromatic microlensing.

Since the QSO images correspond to the same physical source, if we ignore microlensing, any variation due to intrinsic variability has to be reflected in all images, shifted in time by the time delays. The order of arrival of the images is D, B, A, and C. The time delay between images A, B, and C is $\lesssim 1$ day, while between D

and the other images it is ∼23 days. As the time delay between images A, B, and C is so short, we do not expect to see variations due to intrinsic variability between these images, whose amplitude are expected to be governed by the structure function (e.g., Vanden Berk et al. 2004). Indeed, the flux ratio A/B is preserved at two epochs, as we showed in Section 2.5. Nevertheless, Figure 8 shows variations in images A (blue lines; ∼0.1 mag) and C (red lines; ∼0.2 mag) on the timescale of 70 days (20 days in the QSO rest frame). Because the flux of image B (magenta lines) is constant during this timescale, we interpret this to imply the lack of QSO intrinsic variability (or at least to imply variations resulting in the same start and end points), and also as a check of our absolute photometric calibration (which we have also checked independently, finding consistent photometry for the bulk of the sufficiently bright objects in the HST field of view). More puzzlingly, image A varies only in the longer-wavelength filter, F160W, and image D (black lines) varies in opposite directions between the two filters, by ∼0.1 mag, with the variation in F160W tracking that of image C in both filters. We find these variations difficult to reconcile with microlensing, which affects shorter wavelengths more prominently, although microlensing can be responsible for variations of ≲0.15 mag on a similar timescale (see Figure 13 in Eigenbrod et al. 2008). We note also that these variations are far outside the error bars we measure in Table 2 with our careful light profile fitting technique.

In parallel to the microlensing simulations where we explored flux changes, we also looked into the effect that microlensing can have on QSO time delays, assuming the lamp-post model (e.g., Tie & Kochanek 2018). We find that the rms impact on the time delays for images A, B, and C is as large as ∼18, ∼7, and ∼13 days, respectively (for the largest accretion disk we considered, of size $2R_0$). These values are significantly larger than the ≲1-day intrinsic time delay between these images, and when plugged into the structure function (after conversion to rest frame), they are large enough to factor in a hypothetical explanation of the 0.1 mag stochastic variations mentioned above. It should be noted that while a larger disk size has an increased impact on the time delays, it also smooths out the microlensing signal, decreasing the magnitude of flux ratio anomalies. We refrain from speculating on the exact mechanisms of these flux variations, as we do not have enough data to constrain them.

Finally, we have also looked for direct evidence of variability in the longer, 5 yr baseline spanned by Pan-STARRS1 (Chambers et al. 2016), which we show for completeness in Figure 10. However, the uncertainties of the automatic single-component fit to the photometry provided by the Pan-STARRS pipeline are likely to be underestimated, since the system consists of four QSO images and a bright lensing galaxy. We therefore can only conclude that there is no evidence for a monotonous variation recorded by Pan-STARRS1, which would correspond to long-term variability dominated by image A.

4.1.3. Substructure in the Lensing Galaxy

ΛCDM substructure, either dark or luminous, has long been invoked to explain flux ratio anomalies in lensed quads (e.g., Dalal & Kochanek 2002). If the flux anomalies we measure in the HST data are due to substructure, we expect them to persist at long wavelengths as well, as long as the emitting region is not very extended (i.e., AGN radio core emitting region, typically with a parsec-scale length, is small enough to be affected by substructure, but kiloparsec-scale stellar emission is not). In Figure 11 we demonstrate with a toy model that a relatively low
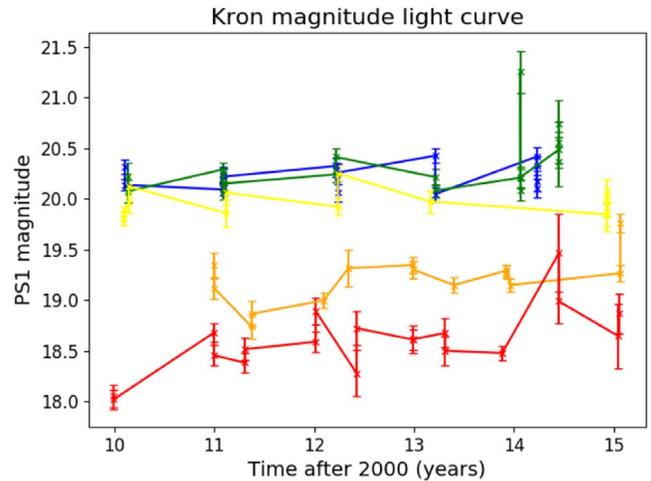
Figure 10. Variability of W2M J1042+1641 in individual Pan-STARRS1 exposures. Blue, green, yellow, orange, and red colors represent, in order, *grizy*. Mismatches in photometry between points very close in time suggest that the photometric errors are underestimated, weakening evidence for variability. Note that the SDSS magnitudes from Table 1, taken earlier in time, are in relatively good agreement.

mass perturber, placed on the order of milliarcseconds away from image A, can produce enough magnification to cause the highly anomalous flux observed in this image, while leaving all other observables unchanged.[36]

4.1.4. Clues from the Radio Data

Radio and far-IR photometry of lensed QSOs is considered to be unaffected by microlensing (e.g., Koopmans et al. 2003), as well as dust extinction. This is because the emission region in the source plane, on scales of milliarcseconds, corresponds to radii much larger than the projected Einstein radius of individual stars in the lensing galaxy (scales of micro- or nanoarcseconds). If the emission in radio and far-IR is due to the AGN, we would expect to measure a similarly anomalous flux ratio to what we measure with HST if it is the case that the anomaly is due to substructure (we use here the term to refer to a low-mass perturber), but we should measure something comparatively closer to the nonano-malous prediction of the best-fit lens model, if it is due to microlensing. We would not expect perfect matches to either of these values because (1) as mentioned above, the size of the emission region in the radio (as well as in the submillimeter) is not the same as that of the optical accretion disk, for example, the isophotes in Figure 5 show extended structure; and (2) the image separations measured from HST (Table 3) and from ALMA (Table 5) are inconsistent, for example, the separation between images A and B is 0″58 and 0″51, respectively, and the separation between the other images is even more discrepant. We note these caveats upfront and point out that the data are not constraining enough to estimate to what extent they can affect the analysis we present below.

The VLA images reveal a flux ratio of A/B ∼ 4.0, compared to the lens model prediction A/B ∼ 1 and the HST measure-ment of A/B ∼ 6.6–10.2 (depending on the measurement filter and visit). The integrated emission that we measure from the two epochs of VLA data is ∼60–100 μJy, 6 − 10 times larger

---

[36] We used five MCMC chains of 10,000 points each. The fluxes of the images predicted by the SIE+γ+GX model were used as constraints (the flux of image A being boosted), with 5% uncertainty. All other lensing parameters except those characterizing the substructure were held fixed.
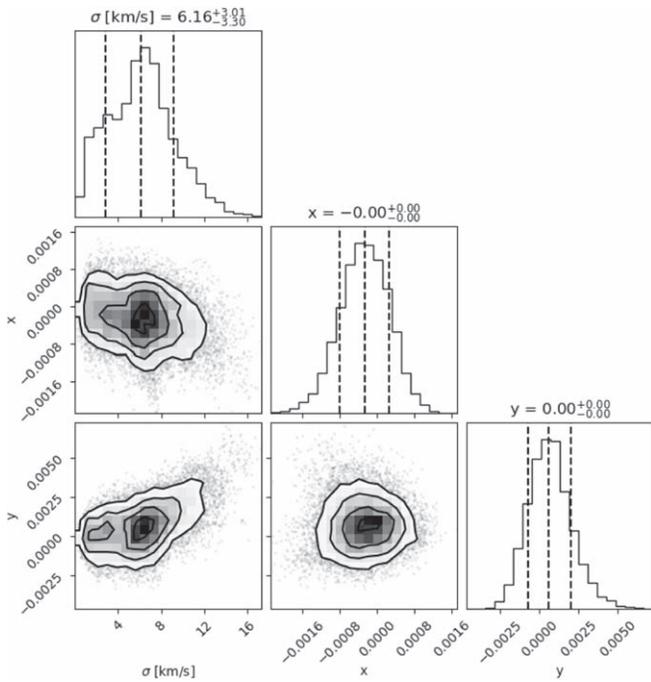
Figure 11. Velocity dispersion and position (in arcsec) of an SIS substructure added to the SIE+γ+GX model around image A, in order to boost its flux by a factor of $\sim$5. The 16th and 84th velocity dispersion percentiles correspond to a mass (inside the substructure's Einstein radius) of $4.1 \times 10^3 M_\odot$ and $4.3 \times 10^5 M_\odot$, respectively.

compared to the prediction of $\sim$10 μJy based on the ALMA data, assuming a typical effective dust temperature of 40 K (see Stacey et al. 2018, 2019) and using the known far-IR$-$radio correlation of star-forming galaxies (e.g., van der Kruit 1971; Yun et al. 2001; Ivison et al. 2010a, 2010b). The excess of measured flux at radio wavelengths suggests that we are seeing AGN emission in addition to synchrotron emission from star formation.[37] This is in line with current models of QSO evolution, which suggest that coexistence of dust-obscured star formation and AGN activity is typical of most QSOs (e.g., Condon et al. 2013; Stacey et al. 2018). We can explain the mismatch in the HST-derived and radio/submillimeter-derived A/B flux ratios if the radio emission consists of a combination of point-like AGN emission and star-formation-related emission. Furthermore, if the anomalous HST flux ratio is due to substructure, extended star formation emission is expected to be comparatively much less affected by a low-mass perturber, which preferentially magnifies the comparatively much more compact AGN emission.

By accounting for the amount of AGN emission relative to the emission from star formation, we can reproduce the intermediate flux ratio that we measure in the VLA data. Let $A_{\rm SF}$ and $B_{\rm SF}$ be the flux densities measured in the radio from star formation in images A and B, respectively, and let $A_{\rm AGN}$ and $B_{\rm AGN}$ be the flux densities measured from compact AGN emission. Using the constraints $A_{\rm SF}/B_{\rm SF} \sim 1$, $A_{\rm AGN}/B_{\rm AGN} \sim 6.6$–10.2, and $(A_{\rm AGN} + A_{\rm SF})/(B_{\rm AGN} + B_{\rm SF}) \sim 4.0$, we can solve for $A_{\rm AGN}/A_{\rm SF} \sim 4.7$–20.4 and $B_{\rm AGN}/B_{\rm SF} \sim 1.4$–0.5. Assuming that the difference between the ALMA-predicted flux densities for the VLA emission and the detected values is also due to further emission from

---

[37] Another source for enhanced radio emission in red QSOs is shocks from winds, as proposed in Glikman et al. (2022).

AGNs, we measure $(A_{\rm AGN} + A_{\rm SF} + B_{\rm AGN} + B_{\rm SF})/(A_{\rm SF} + B_{\rm SF}) \sim$ 6–10, and we calculate a very consistent range of $\sim$4–11.5.

In light of the results above, we conclude that the radio data are more consistent with the HST-measured flux ratio anomaly of image A persisting at longer wavelengths, and therefore being mostly due to substructure rather than microlensing. Therefore, the total magnification to use for determining the source properties is the large value we computed in Section 4.1 ($\mu \sim 117$) from the observed flux ratios in the HST data, and we adopt this value when demagnifying fluxes to consider the QSO's intrinsic properties, in the sections that follow.

Finally, we conclude this section with two notes. First, in the submillimeter emission measured by ALMA, A/B = 1.61 ± 0.24, which is somewhat larger than the model prediction of $\sim$1. One possible explanation would be that of thermal dust emission from a sufficiently compact region located close enough in projection to the substructure to be sufficiently magnified. Indeed, Stacey et al. (2021) showed that such emission can be as compact as a few hundred parsecs. Second, it is worth mentioning that similar analyses combining near-IR, submillimeter, and radio flux ratio measurements were recently performed by Badole et al. (2020) for the highly optically anomalous SDSS J0924+0219, concluding that the long-persisting anomaly is most likely caused by microlensing, and by Stacey et al. (2018) for MG J0414+0534, suggesting that the anomaly is caused by substructure.

4.2. Spectral Characteristics

The LRIS spectrum of W2M J1042+1641, taken across multiple PAs, for a total of 50 minutes (Section 2.4) is our best data set for investigating the QSOs' emission-line properties. Figure 12 shows the combined LRIS spectrum of W2M J1042 +1641 (black line) plotted on a logarithmic y-axis in order to enhance features across its dynamic range.

To identify and study line features, we require a better determination of the QSO continuum. Following our interpretation that W2M J1042+1641 is a reddened QSO with some leakage of the intrinsic spectrum, we model the QSO spectrum as a power law, with spectral index $\alpha_\nu = -0.5$ ($F_{\lambda 0} \propto \lambda^{-(\alpha_\nu + 2)}$). We then fit the line-free portions of the spectrum with a two-component power law, one reddened and one pure, both with the same power-law index:

$$F_\lambda = AF_{\lambda 0} + BF_{\lambda 0}e^{-\tau_\lambda}. \qquad (3)$$

The best fit is shown with a purple dashed–dotted line, along with the unreddened component, plotted with a blue dashed–dotted line, and the reddened component, plotted with a red dashed–dotted line. For comparison, we also show the best-fit reddened QSO template (red solid line) that was determined from the near-infrared spectrum combined with the LRIS spectrum, using only wavelengths longer than 8000 Å (Section 2.4). While the template-based fit results in $E(B - V) = 0.73$ mag, the two-component power-law fit yields $E(B - V) = 0.89$ mag.

We see absorption in the Mg II line, plotted in velocity space in the top right panel of Figure 12. The purple line is the continuum from our fit. We see two distinct absorption systems that are well fit by a double Gaussian, with FWHM speeds of 4348 and 1838 km s^{-1} and systemic outflow speeds of -1407 and -350 km s^{-1}. A feature peaking at $+3000$ km s^{-1} from the Mg II position is unidentified and may be part of the Mg II line itself.
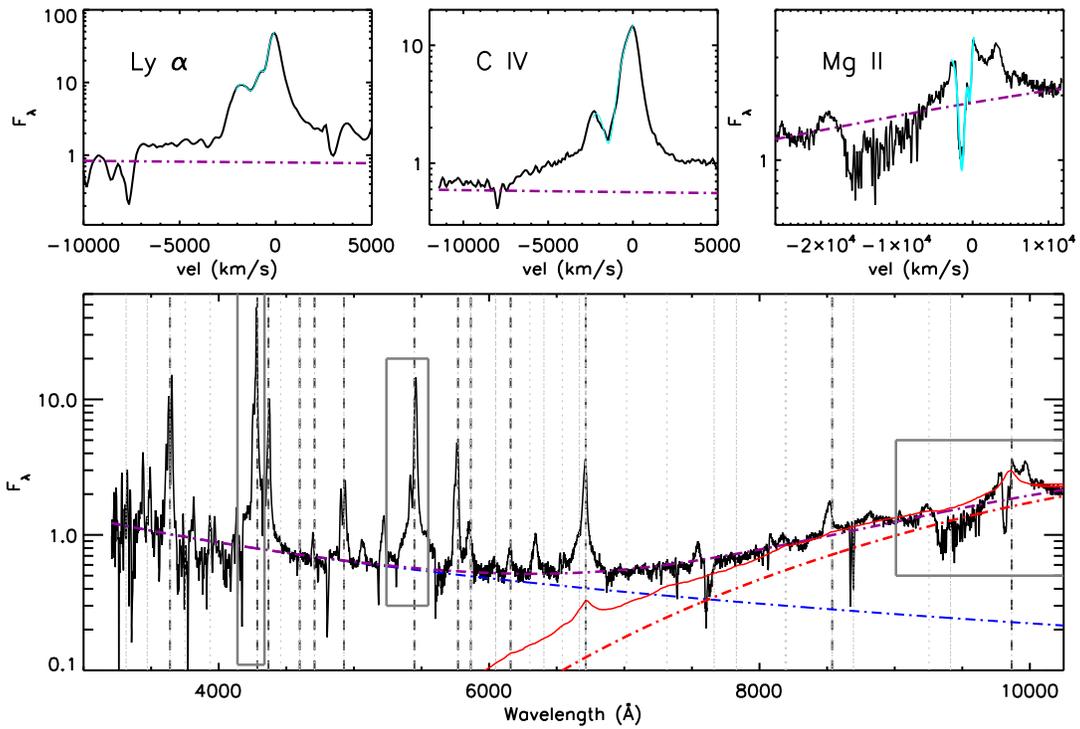
Figure 12. Deep LRIS spectrum of W2M J1042+1641 (bottom panel; black line), with typical (light gray) and prominent (dark gray) QSO emission lines marked with vertical dotted and dashed lines, respectively. A best-fit reddened power-law continuum plus an unreddened leakage component is shown with a purple dashed–dotted line, with the reddened ($E(B − V) = 0.89$) and leaked components plotted in red and blue, respectively. The solid red line is the best-fit reddened template with $E(B − V) = 0.73$. *Top row:* profiles of prominent emission lines Mg II, C IV, and Lyα (left to right) plotted in velocity space. All three lines show a distinct absorption feature that is well fit by a single or double Gaussian profile, with an outflow speed of ∼1000–1500 km s^{-1}. We note that the absorption feature blueward of Mg II is due to telluric absorption. The line regions plotted in the squares at the top are also marked with gray boxes in the bottom panel.

The other two panels along the top of Figure 12 display the same for C IV (middle) and Lyα (left), which both show a blueshifted absorption feature, also well fit by a Gaussian. The feature at C IV is best fit by a single Gaussian with an FWHM velocity width of 3060 km s^{-1} and an outflow velocity of -1576 km s^{-1}. The absorption at Lyα is best fit by two Gaussians with FWHM velocity widths of 4184 and 1306 km s^{-1}, outflowing at -1013 and -515 km s^{-1}. These features are indicative of outflowing gas.

4.3. Black Hole Mass

The left panel of Figure 13 shows that the Hα line in the near-infrared spectrum is well fit by a single Gaussian, with an FWHM in velocity space of $v_{FWHM} = 8479$ km s^{-1}, which we use to estimate the BH mass of W2M J1042+1641. We combine the line width with an estimate of the QSO's intrinsic luminosity at 5100 Å and apply those values to the single-epoch virial BH mass estimator ($M_{BH,vir}$) following the formalism of Shen & Liu (2012),

$$\log\left(\frac{M_{BH,vir}}{M_\odot}\right) = a + b\log\left(\frac{L_{5100}}{10^{44}\text{erg s}^{-1}}\right) + c\log\left(\frac{v_{FWHM}}{\text{km s}^{-1}}\right),$$
(4)

adopting the values $a = 0.774$, $b = 0.520$, $c = 2.06$ for single-epoch measurements of FWHM$_{Hα}$ and L_{5100}, based on the calibration of Assef et al. (2011). We choose the Hα line because it is in a region of minimal reddening in our spectrum and the more commonly used Hβ is strongly blended with [O III]. The values derived from our procedure are listed in Table 8.

To determine L_{5100}, we must account for reddening and magnification, as well as an absolute spectrophotometric calibration of our near-infrared spectrum. Our procedure is shown in the right panel of Figure 13. Because the magnification of the QSO is derived from the F160W image, demagnifying the observed luminosity at this wavelength ($\lambda_{F160W,rest} = 4370$ Å) will give the most internally consistent results. We first estimate the continuum flux, represented by Equation (3), through the F160W band and scale it to the summed flux of the four QSO components listed in Table 2 (orange circle). We then find the intrinsic continuum at this wavelength by dereddening the second term in Equation (3) and adding to it the first term (cyan dotted–dashed line). We scale the total observed flux from the four QSO components to match the dereddened continuum and then shift their flux value at $\lambda_{F160W,rest} = 4370$ Å (red triangle) to the dereddened flux at $\lambda_{rest} = 5100$ Å (green square). Finally, we divide the flux by the magnification factor of $\mu = 117$ (Section 4.1) and compute a luminosity of $\log(L_{5100}$ [erg s^{-1}]) = 45.35$. These values yield $\log(M_{BH,vir}$ [$M_\odot$]) = 9.56$.

To estimate L/L_{Edd}, we use a bolometric correction (BC) at 5100 Å of 9.5 from Richards et al. (2006). This gives $\log(L_{bol}$ [erg s^{-1}]) = 46.33$ and a corresponding Eddington ratio $\log(L/L_{Edd}) = -1.34$, which is a relatively low accretion rate compared with previously studied red quasars, most of which have $\log(L/L_{Edd}) \gtrsim -0.3$ (Kim et al. 2015).

For comparison, recompute these physical parameters, using the observed WISE W4 (22 μm) instead of the F160W flux, as the rest wavelength of W4 corresponds to 6.28 μm, which suffers minimal extinction compared with the F160W band and is dominated by AGN emission (Stern et al. 2014). For
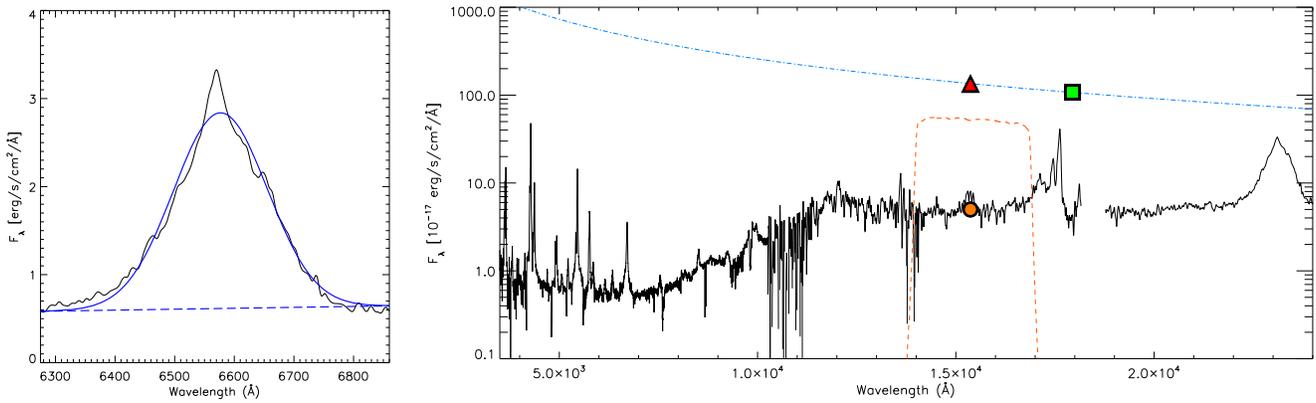
Figure 13. Left: we show the Hα line, seen in the near-infrared spectrum (black line), fit with a single Gaussian plus a linear continuum fit (solid blue line) to determine the FWHM needed to compute the BH mass in Equation (4). The linear continuum is shown with a dashed blue line. Right: we scale our spectrum to the total F160W flux from the QSO (orange circle). We plot the dereddened continuum from Equation (3) (cyan dotted–dashed line) and shift from the observed flux through the F160W band to match the dereddened continuum (red triangle) to the flux at 5100 Å flux (green square). We use this value to find L_{5100} and compute the BH mass in Equation (4).

Table 8
Black Hole Mass and Accretion Rate Estimates

Parameter	Value	Unit
	QSO Properties	
Redshift, z	2.517	…
Magnification, μ	117	…
	Hα Line Fit	
FWHM$_{H\alpha}$	8340 ± 330	km s^{-1}
$F_{H\alpha}$	$4.4 \pm 0.2 \times 10^{-14}$	erg s^{-1} cm^{-2}
$\log(L_{H\alpha})$	44.09	erg s^{-1}
$\log(M_{BH})$	9.51	$M_{\odot}$
	Continuum Fit	
F_{5100}	1.07×10^{-15}	erg s^{-1} cm^{-2} Å^{-1}
$\log(L_{5100})$	45.35	erg s^{-1}
$\log(M_{BH})$	9.56	$M_{\odot}$
$\log(L_{bol})$	46.33	erg s^{-1}
$\log(L/L_{Edd})$	-1.34	…
	WISE Based	
$\log(L_{W4})$	45.60	erg s^{-1}
$\log(L_{5100})$	44.37	erg s^{-1}
$\log(M_{BH})$	9.06	$M_{\odot}$
$\log(L_{bol})$	46.47	erg s^{-1}
$\log(L/L_{Edd})$	-0.68	…

Note. All luminosities presented are demagnified.

example, an extinction of $E(B - V) = 0.7$ results in $A_{W4} = 0.06$ mag (or 6% of the flux) at 6.28 μm and thus avoids the need for a reddening correction. In Section 4.1.3, we concluded that the flux anomaly seen in this object is due at least in part to substructure, and therefore the magnification factor relevant at this wavelength would be the same one observed in the HST/WFC3 F160W images ($\mu \sim 117$).[38] We scale the observed WISE W4 luminosity to the luminosity at 5100 Å using the SED from Richards et al. (2006) for optically red QSOs. We

demagnify this flux by a factor of 117 and find $\log(L_{5100} [\text{erg s}^{-1}]) = 44.37$, resulting in $\log(M_{BH,vir} [M_{\odot}]) = 9.06$.

Using a BC of 7.5 from Richards et al. (2006) based on the optically red QSO SED[39] for a frequency corresponding to W4 (22 μm) in the rest frame, $\log(\nu [\text{Hz}]) = 13.68$, we compute $\log(L_{bol} [\text{erg s}^{-1}]) = 46.47$—consistent with the estimate based on the dereddened and F160W flux. We find an Eddington ratio of $\log(L/L_{Edd}) = -0.68$, which is more consistent with the values found for other red quasars.

4.4. The Relation between the Black Hole Mass and the Host Galaxy Luminosity and Stellar Mass

Gravitational lensing provides a unique opportunity to study the coevolution of SMBH host galaxies at high redshifts where, in the absence of lensing, the AGN outshines the host, making it nearly impossible to cleanly separate the two components. Lensed QSOs offer the advantage of magnifying and stretching out the host galaxy while preserving surface brightness. At the same time, the AGN light, being a point source, appears in multiple distinct spots and is easier to separate and subtract than in the absence of lensing. Early work, using 31 lensed QSOs at high redshifts ($1 < z < 4.5$) to study their host galaxies and investigate coevolution, was reported in Peng et al. (2006). More recently, Ding et al. (2017, 2021) analyzed an additional eight lensed systems at $0.65 < z < 2.32$, including a direct measure of their stellar masses, $M_{\star}$, using SED modeling. Both studies observe an increase in the $M_{BH}/M_{\star}$ ratio toward earlier times, though the number of sources is statistically small above $z \sim 2.5$ and Peng et al. (2006) do not directly measure $M_{\star}$, but rather infer the ratio via the evolution of host luminosities, L_R, toward high redshift.

We have determined the demagnified magnitude of the host galaxy of W2M J1042+1641 in both HST bands (Section 3.1), reported in the last column of Table 2. When combined with the BH mass that we estimated in Section 4.3, we can include W2M J1042+1641 in these high-redshift investigations of galaxy and SMBH coevolution.

Following Peng et al. (2006), we use the demagnified source magnitude in F160W to convert to absolute magnitude R band, which is commonly used in the literature, as the K-correction is

---

[38] If the anomaly was instead due entirely to microlensing, the magnification would be much closer to the value of $\mu \sim 53$, as microlensing is expected to become insignificant at long wavelengths.

[39] We note that using the template for all SDSS QSOs (BC $\simeq 7.9$) does not significantly affect the results.
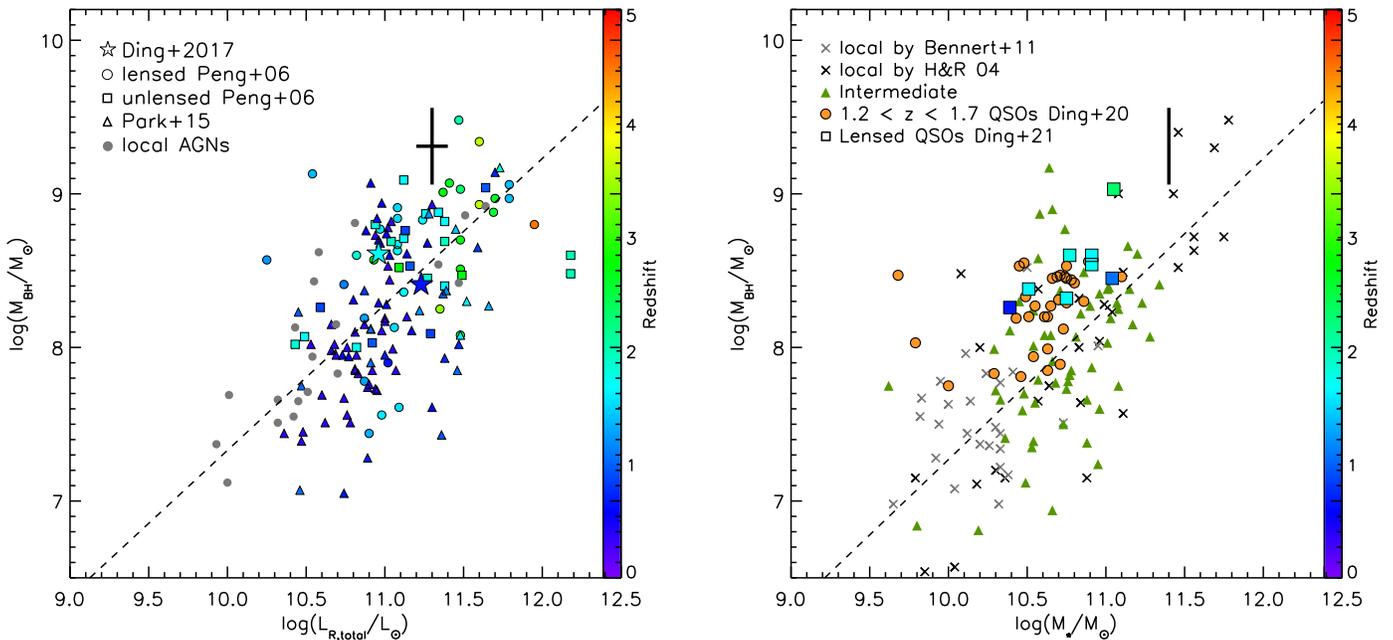
Figure 14. Left: 0bserved BH mass vs. total R-band host galaxy luminosity using data from lensed and unlensed AGNs across a broad redshift range (reproducing Figure 3 of Ding et al. 2017, which has not been corrected for passive evolution). Right: BH masses vs. stellar mass, as shown in Figure 3 of Ding et al. (2021). Sources are color-coded by redshift, and W2M J1042+1641 is represented by a thick black line spanning the possible values according to our calculations. The dashed black line is a fit to the local AGNs. In both figures, W2M J1042+1641 is above the local relation, and although it is the highest-redshift object with a stellar mass estimate, its offset from the local relation is consistent with the next-highest redshift source, HE 1104−1805 at $z = 2.32$ (green circle; Ding et al. 2021).

less dependent on the galaxy SED template.[40] We note that while the de Vaucouleurs profile, used to fit the radial light profile of early-type galaxies, provides a better fit to the lensed host galaxy light, the Sbc template actually provides a better fit to the color (the galaxy is bluer in F125W than predicted from the E template). We find $\log(L_R/L_\odot) \sim 11.4$ for the early-type template and $\log(L_R/L_\odot) \sim 11.2$ for the Sbc template.[41]

The left panel of Figure 14 plots $M_{\rm BH}$ versus L_R for samples of galaxies with these available measurements, following the structure of Figure 3 of Ding et al. (2017) for ease of comparison. Lensed and unlensed sources from Peng et al. (2006) and Park et al. (2015) are plotted with circles and squares, respectively, and are color-coded by redshift as indicated by the color bar. Two additional lensed QSOs at $z = 0.654$ and $z = 1.693$ from Ding et al. (2017) are plotted with stars, local AGNs from Bennert et al. (2010) are shown with gray circles, and the dashed black line represents the best fit to the local relation. W2M J1042+1641 is plotted with a black cross representing the range of BH masses as estimated in Section 4.3 and host luminosities from the K-correction described above. W2M J1042+1641 lies above the relation, in the region of the diagram consistent with other galaxies in the $z = 2$–3 range (green symbols).

To investigate the growth of the host galaxy's stellar mass, given only two photometric measurements of the host galaxy flux, we cannot use SED modeling as was done in Ding et al. (2021). However, at the redshift of W2M J1042+1641, the F125W and F160W effective wavelengths shift nearly perfectly to the SDSS u and g passbands, respectively. Therefore, we can use the F125W − F160W color as rest-frame $u - g$ and estimate the galaxy's stellar mass using mass-to-light ratios

from a single color as determined by Bell et al. (2003) (Table 4). Using the rest-frame g-band luminosity, we estimate $\log(M_\star/M_\odot) = 11.4$ with the model that includes GX as a perturber.[42]

We caution that this stellar mass is based on only two photometric measurements and is further subject to the uncertainty propagated from the magnification factor, but it is the best that can be estimated with the current data. An estimate of the stellar mass in W2M J1042+1641 was also conducted by Matsuoka et al. (2018), which found $\log(M_\star/M_\odot) = 13.55$ using an SED fit (with the Code Investigating GALaxy Emission, CIGALE; Noll et al. 2009) to the observed photometry that includes both the AGN and host galaxy light. They then demagnify the SED by $\mu = 122$[43] to arrive at $\log(M_\star/M_\odot) = 11.46$. Despite the coincidental agreement between our values, we caution that the SED fitting method in Matsuoka et al. (2018) contains several inaccuracies in this approach. First, the CIGALE SED fitting code does not include a component for a reddened AGN and only allows for either an unreddened AGN (Type 1) or a completely obscured AGN (Type 2), neither of which fits the shape of the moderately reddened yet AGN-dominated continuum of this red QSO. This leads CIGALE to attribute the rest-frame UV–optical light to a heavily reddened host galaxy, which will naturally over-estimate $M_\star$. Furthermore, the CIGALE-reported $M_\star$ is based on an SED whose galaxy component was demagnified by the same factor as the AGN. However, our analysis shows that the (extended) host galaxy is only magnified by a factor of 15 (computed with `hostlens` for the SIE+γ+GX model), which is off by a factor of ∼8 from the AGN's magnification (from

---

[40] We employ the Coleman et al. (1980) spectral templates.
[41] To implement the K-correction, we use the `mag2mag` routine (see Section 2.4.1).

[42] $\log(M_\star/M_\odot) = 11.5$ is estimated for the model without a perturber.
[43] This was an initial estimate for the magnification, reported in Glikman et al. (2018), but which is superseded by the more thorough analysis presented in this work.

Equation (1)). It appears that the overestimate of the stellar mass from the SED fitting and the overcorrection of the host galaxy's demagnification conspire to yield the same value that we derive here. Depending on the value for M_{BH} used, from Table 8, we find $\log M_{BH}/M_\star = -2.3$ to -1.8.
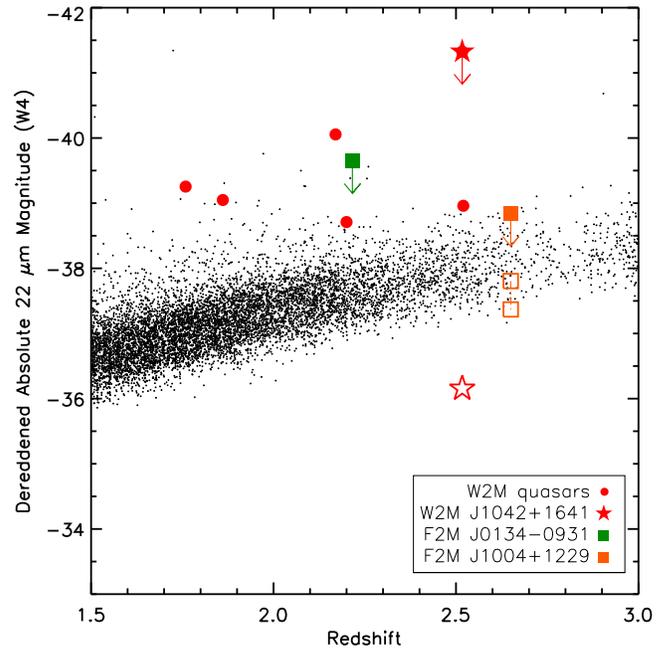
The right panel of Figure 14 plots M_{BH} versus $M_\star$, following the structure of Figure 3 of Ding et al. (2021) for ease of comparison. This plot has fewer high-redshift sources owing to the necessity of multiple bands for estimating $M_\star$. Here the local AGNs are represented by gray and black crosses for the samples of Bennert et al. (2011a) and Häring & Rix (2004), respectively. Intermediate-redshift sources spanning $0.34 < z < 1.90$ are plotted with green triangles (Bennert et al. 2011b; Cisternas et al. 2011; Schramm & Silverman 2013). Intermediate-redshift, unlensed QSOs from Ding et al. (2020) are shown with orange circles, and the sample of lensed QSOs analyzed in Ding et al. (2017, 2021) are plotted with squares and colored according to their redshift coded by the color bar. W2M J1042+1641 appears as a vertical black line, spanning the range of possible BH masses at the stellar mass of $\log(M_\star/M_\odot) = 11.4$ that we derived. W2M J1042+1641 is the highest-redshift source among the objects in this figure and is shifted by a similar amount from the local relation (black dashed line) to the next-highest redshift source, HE 1104 −1805 at $z = 2.32$.

4.5. Magnification and Population Analysis

Figure 15 shows W2M J1042+1641 (red filled star) on a WISE W4 luminosity versus redshift diagram. The filled red circles are the other high-redshift W2M QSOs found in our study, which is spectroscopically complete (Glikman et al. 2022). We compare them with ∼20,000 QSOs from SDSS with matches in the UKIDSS and WISE catalogs with spectroscopic redshifts from Peth et al. (2011; black circles). The demagnified position of W2M J1042+1641 is shown as an open star, below the lower envelope of SDSS QSOs detected, which represents the WISE W4 detection limit. We see that this red QSO exists among a population whose luminosity is too low to be discovered in the wide-field surveys that yield the vast majority of QSOs in the literature.

We also plot in Figure 15 the positions of the two other lensed F2M red quasars: F2M J0134−0931 (green square; Gregg et al. 2002) and F2M J1004+1229 (orange square; Lacy et al. 2002). We modeled F2M J1004+1229 by fitting SIE+γ models to the relative astrometry reported in CASTLES,[44] in order to determine its magnification factor. We do not compute a magnification factor for F2M J0134−0931, as this system has a complex lensing configuration (Keeton et al. 2003), whose modeling is beyond the scope of this paper. Their demagnified positions are shown with corresponding open symbols, and they, too, lie at the edge of or below the faintest luminosities accessible to SDSS, UKIDSS, and WISE.

The W2M survey finds seven high-redshift ($z > 1.7$) QSOs, including the lensed quasar F2M J1004+1229, which we recover from our previous FIRST-selected sample. That means that the lensing fraction of this complete, flux-limited sample is $2/7 = 28\%$—three orders of magnitude higher than the lensing fraction for luminous unobscured QSOs in typical surveys (Oguri & Marshall 2010). The lensing fraction of the F2M survey (Section 1) is also large, at ∼2%, but smaller than that

---

[44] CfA-Arizona Space Telescope LEns Survey, C. S. Kochanek, E. E. Falco, C. Impey, J. Lehar, B. McLeod, H.-W. Rix, https://www.cfa.harvard.edu/castles/.

Figure 15. Absolute magnitude at 22 μm in the observed frame, from WISE, for the sample of red QSOs in the W2M survey. The observed luminosity of W2M J1042+1641 is indicated by a star, with a downward-pointing arrow pointing toward its demagnified intrinsic luminosity (using $\mu = 117$). This figure indicates that W2M J1042+1641 is a reddened QSO drawn from a population of sources that is not accessible to the wide-area QSO samples currently being studied. We show, for comparison, the observed and magnification-corrected luminosities of the other high-redshift lensed F2M red quasars discussed in the text, using only an upper limit for F2M J0134−0931 (see text) and the luminosity range for F2M J1004+1229 corresponding to magnification $3.15^{+0.71}_{-0.55}$.

of the W2M survey. This is likely due to the F2M survey being about a magnitude deeper than W2M.

We note that the previously discovered lensed red quasars are all radio sources. Some of them are intrinsically reddened, while others are likely reddened by the lensing galaxy itself. Malhotra et al. (1997) showed that lensed quasars found in radio-selected samples have redder colors, implying that radio selection finds dusty systems that may be lost in optical QSO samples that often impose a blue color cut. Our surveying of red QSOs in shallow, wide-field surveys (i.e., W2M) is beginning to remedy this incompleteness by recovering radio-quiet reddened lensed QSOs. In addition, shallow flux-limited surveys benefit from increased magnification bias, making these lenses easier to find.

The demagnified luminosity density computed from the W4 flux density of W2M J1042+1641 ($\lambda_{rest} = 6.2$ μm) is $\log(L) = 32.46$ in units of erg s^{-1} Hz^{-1} for magnification = 117. We interpolate between the WISE bands and find the demagnified $\log(L_{5\mu m}) = 32.29$. The 5 μm luminosity function at $z = 2.5$ for red QSOs in deep Spitzer fields is $\log(L_0^*) = 31.92$ (Table 8 and Figure 19 of Glikman et al. 2018). Thus, when demagnified, W2M J1042+1641 is near the knee but on the bright-end side of the red QSO luminosity function (QLF) derived in Glikman et al. (2018). The sources that enabled a determination of this QLF were derived from deep Spitzer fields that covered relatively small areas. Although we cannot use this QSO to investigate the faint end of the red QLF, finding such a highly magnified red QSO offers a unique opportunity to study a population of QSOs whose luminosity is intrinsically lower than the QSOs found in wide-field surveys such as SDSS, which make up the vast majority of known QSOs.

5. Conclusions

We have discovered a quadruply lensed radio-quiet QSO at $z = 2.517$ identified through HST imaging of dust-reddened QSOs selected by their WISE colors. Using optical and near-infrared spectroscopy, we determine that the QSO is reddened by $E(B - V) \simeq 0.7$–0.9 from dust intrinsic to the QSO's environment, as opposed to dust in the lensing galaxy. Our lensing analysis finds a magnification factor of $\sim$52 for the best-fit model, but boosted to $\sim$117 owing to strong flux anomaly. Using photometric data from near-IR to radio, we conclude that substructure is the most likely cause for the anomaly.

We estimate the QSO's BH mass to be in the range $\log(M_{\rm BH}\,[M_\odot]) = 9.06$–$9.56$, depending on how the unreddened continuum luminosity is computed. The QSO's rest-frame infrared luminosity is $\log(L_{5\mu{\rm m}}\,[{\rm erg\,s^{-1}\,Hz^{-1}}]) = 32.29$, which is near the knee of the QLF, representing more typical quasar luminosities that are difficult to access at high redshifts. The QSO's Eddington ratio could be as high as $L/L_{\rm Edd} = 0.21$, if the bolometric luminosity is estimated from the 22 μm flux, but could be as low as $L/L_{\rm Edd} = 0.05$ if the HST F160W band is used. The former would be consistent with accretion rates seen for more luminous red quasar samples. These characteristics, in addition to evidence for outflowing gas seen in absorption in the QSO's spectrum, point to a system in a transitional phase following a major merger, as is seen in the hosts of more luminous red quasars.

In the future, on a long timescale, monitoring observations would be required to separate microlensing and intrinsic flux variations. Deeper radio observations to robustly measure the flux of all four QSO images, or observations probing narrow-line emission from the source, too spatially extended to be affected by microlensing (Nierenberg et al. 2014, 2017), would be required to further separate the effects of microlensing and substructure and provide a more robust total magnification factor. In the shorter term, a dedicated effort to completing the pixelated modeling, described in Appendix C, with both filters and both visits will improve our understanding of the QSO host galaxy and its apparent companion (source X). In addition, a more thorough analysis, adding the two WFC3/UVIS bands to the WFC3/IR bands, will constrain the colors and SED of the host galaxy to better probe coevolution of QSOs and their hosts at the height of cosmic noon.

Appendix A
SIE+γ Model

We show in Figures 16 and 17 the results of our fitting using the SIE+γ model discussed in Section 3.2 and Table 6. This
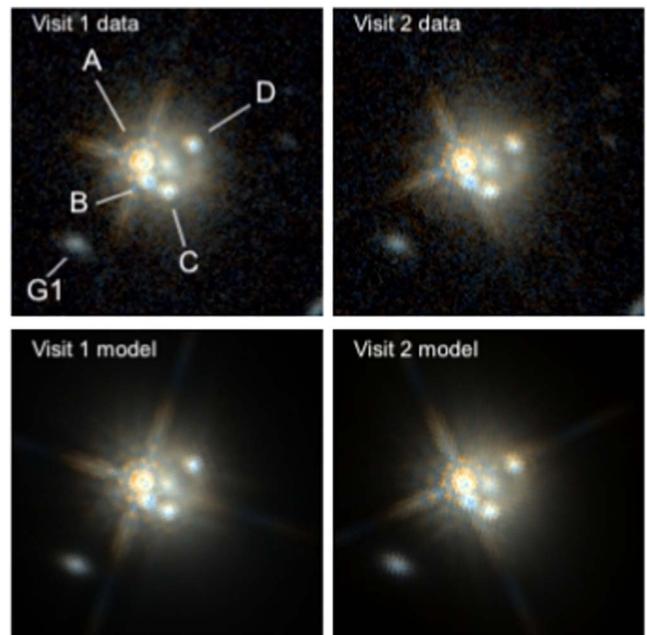
Figure 16. HST WFC3/IR F125W and F160W color-combined images of W2M J1042+1641 over two visits, along with output from a morphological analysis with `hostlens`, as in Figure 2, but with a "free SIE+γ" model.
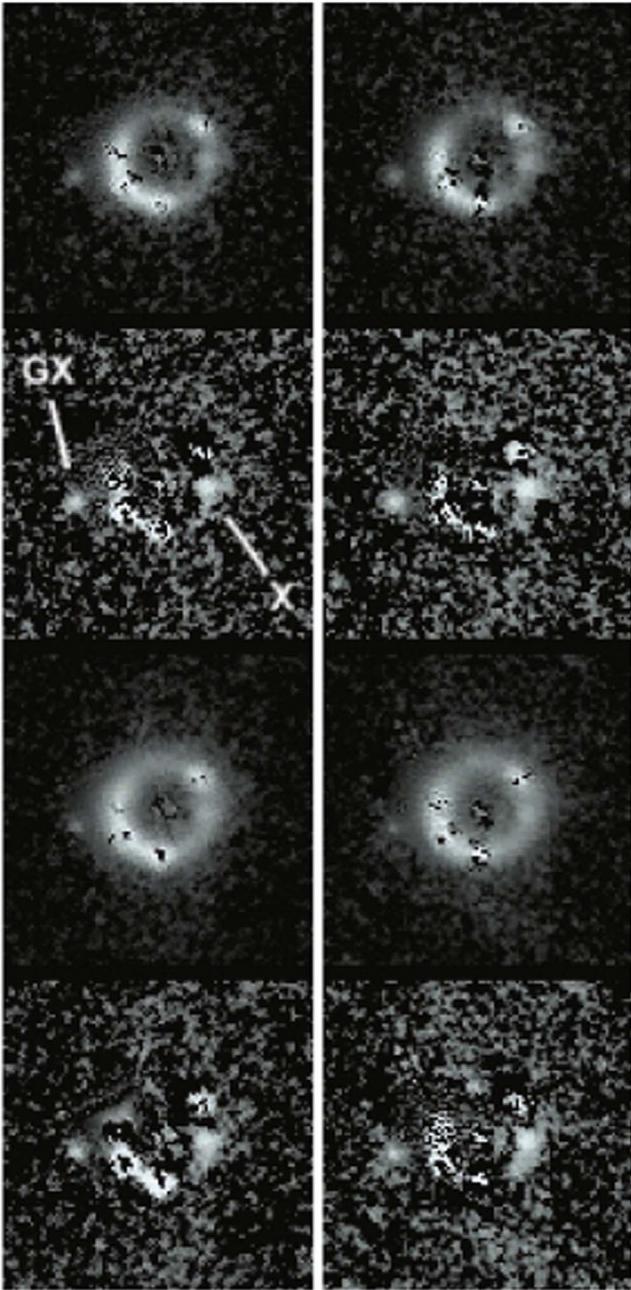
Figure 17. HST Residual images of W2M J1042+1641 over two visits, after subtracting the best-fit morphological models determined by `hostlens` shown in the bottom panel of Figure 16, as in Figure 3, but with a "free SIE+γ" model.

model does not include object GX as a perturber. Despite this, the difference between images is barely distinguishable by eye, and our reasoning for adopting the model with GX as a perturber (SIE+γ+GX) is laid out in Sections 3.2 and 4.

Appendix B
Details of the Microlensing Calculation

We follow the simulation technique in Chen et al. (2018), which itself is built on the approach in Kochanek (2004) to compute microlens tracks. The technique is implemented in the code `GPU-D`[45] (Vernardos & Fluke 2014), further modified by

---

[45] https://gerlumph.swin.edu.au/software/#gpud

Chan et al. (2021) to account for the initial mass function (IMF). For the population of lenses we assume a Salpeter (Salpeter 1955) IMF and a mean solar mass of 0.3 $M_\odot$. To characterize the QSO accretion disk, we consider the two values of the BH mass, corresponding bolometric luminosities, and Eddington ratios computed in Section 4.3 (cases 1 and 2 in Figure 9), and we compute the radiation efficiency η by combining Equations (11) and (13) in Davis & Laor (2011). For the size of the accretion disk, R_0, in the observed F160W filter we use Equation (2) from Morgan et al. (2010) and assume inclination $i = 60°$ and PA = 90°. To compute the microlensing timescale represented in the horizontal axis of the left panel of Figure 9, we follow the methodology in Kochanek (2004) and compute a typical microlens relative speed of ~ 340 km s^{-1}. To compute the microlensing probability, we used the observed magnification of image A and divided it by its predicted SIE+γ+GX model magnification, anchored by C/D in Section 4.1. The magnification excess is 4.9 ± 1.0, and we measure the fraction of pixels from the microlensing maps that satisfy these constraints.

Appendix C
The QSO Host Galaxy, Structure along the Einstein Ring, and Pixelated Modeling

As mentioned in Section 3, we fitted the source QSO host galaxy with `hostlens`, using a circular de Vaucouleurs profile. The delensed magnitude of the source is given in Table 2. As noted in Marshall et al. (2007), a dominant source of uncertainty on the magnitude of the source is due to the distribution of power-law slope values in the radial profile of lenses (we have assumed isothermal profiles in Section 3.2). Following the Appendix of Marshall et al. (2007), this uncertainty amounts to $\sigma_{\mathrm{mag}} \sim 0.26$ mag; the same uncertainty also affects the inferred effective radius r_e as $\sigma_{r_e}/r_e \sim 0.12$. The effective radii inferred by `hostlens` are $0\rlap{.}''15$ or 1.2 kpc at the QSO redshift (F125W), $0\rlap{.}''34$ or 2.7 kpc (F160W) for the SIE+γ+GX model, and $0\rlap{.}''16$ or 1.3 kpc (F125W) and $0\rlap{.}''50$ or 4.0 kpc (F160W) for the SIE+γ model. Note that we find r_e to be significantly larger in F160W.

The QSO host galaxy, lensed into an Einstein ring, does not appear to be smooth, but instead shows at least one bluish peak present in both visits and both filters, south of image D (denoted by "X" in Figure 3). This object shows what appears to be a tail following the curvature of the Einstein ring, which makes it unlikely to be a projected structure or one that is associated with the lensing galaxy. In fact, its location on top of the Einstein ring suggests that X is most likely to be a source located at the same redshift with the QSO host. Thus, X must be an extended structure, part of the QSO host or of its close environment. The position of X on top of the Einstein ring implies that it must produce counterimage(s) (mirror-like image (s) located on the other side of the Einstein ring, with respect to the lens). Our approach in Section 3.1 is not designed to account for these, and they would be treated as noise around the bright QSO images during the PSF reconstruction. To identify such counterimages, we conduct an exploratory analysis where we model the first visit in the F125W filter by reconstructing the source on a grid of pixels, with the code `glee` (Suyu & Halkola 2010; Suyu et al. 2012). The advantage of using a pixelated grid is that it does not need to assume a particular morphology of the source object.
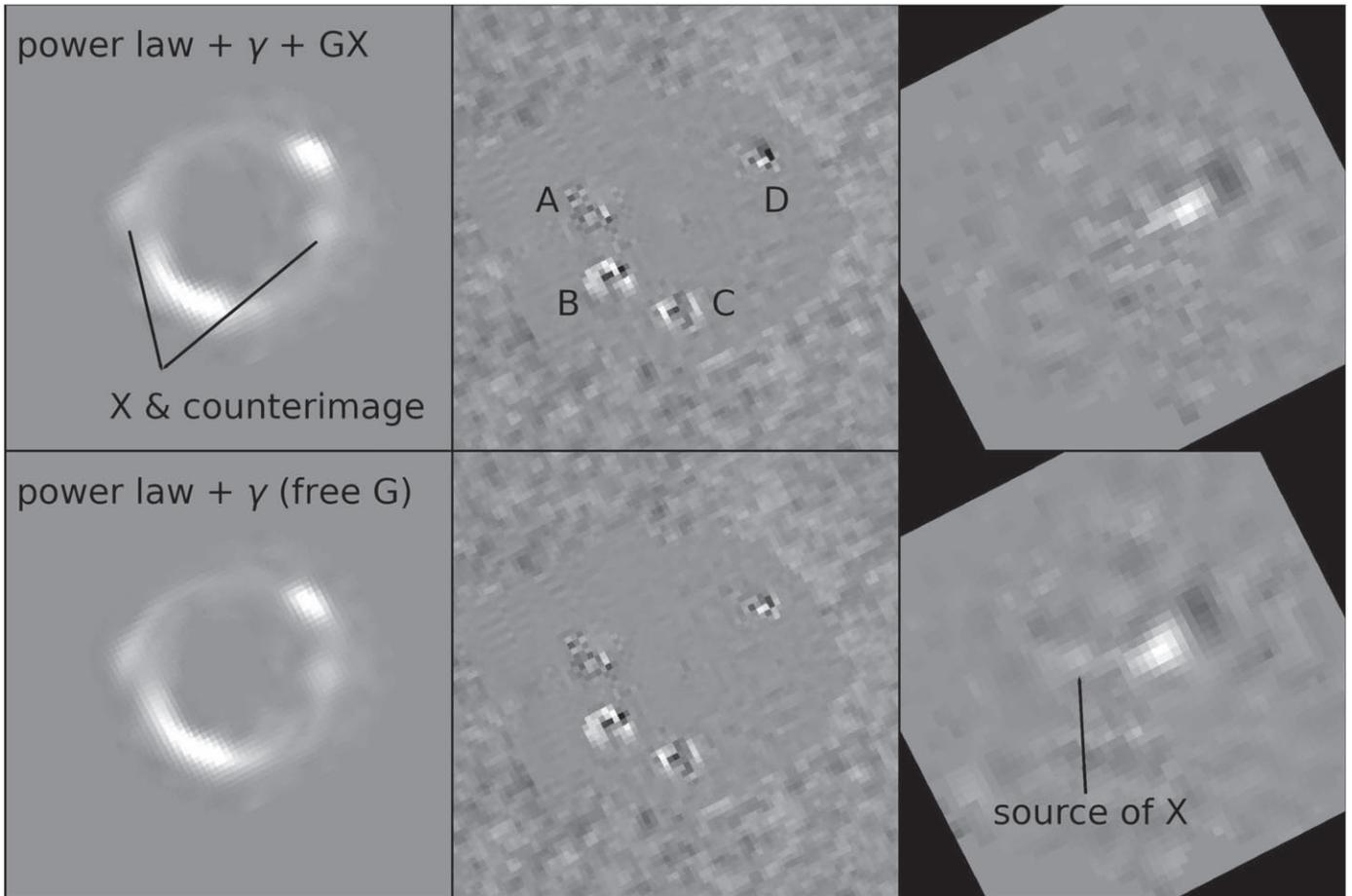
Figure 18. Pixel-based modeling of F125W visit 1 with `glee` (Suyu & Halkola 2010; Suyu et al. 2012). A power-law $+ \gamma$+GX mass model was used for the top panels, and a power-law $+ \gamma$ model with free lens position was used for the bottom panels. From left to right: the best-fit model of the extended arcs, the residuals after subtracting the best-fit model (the gray scale is linear, from -10σ to 10σ), and the source reconstruction on a 50×50 pixel grid. The pixel scales in the source plane averaged between the two axes are $0{.}''030$ (top) and $0{.}''024$ (bottom). North is up, and east is to the left. Note in the arc the luminous blob, X, south of image D, and its fainter counterimage situated on the diametrically opposed part of the arc. The QSO host appears elongated with a possible tidal tail extending toward the less luminous substructure on the left, responsible for the luminous blob seen in the arc.

Since `glee` uses a power law rather than SIE for the mass model, we perform the modeling with a power-law $+ \gamma$+GX mass model with a slope of 2.22 and a power-law $+ \gamma$ model with free lens position with a slope of 1.90. We show the structure of the pixel-based Einstein ring model, the residuals of the model, and the reconstructed source structure in the source plane in Figure 18, for lensing models w/ and w/o GX. The Einstein ring model shows a clear extended counterimage of X to the east, and the reconstructed source appears elongated (Figure 18, top row). Since we were able to fit the Einstein ring with a circular analytical source profile in Section 3.1, we interpret the source elongation to be due to the blending (given the reconstructed source resolution) between the QSO host galaxy and the source responsible for X and its counterimage. In the bottom row of Figure 18, rightmost panel, the source of X itself may even be resolved and possibly connected to the host galaxy by a tidal tail. The existence of this extended source is in agreement with the observation that most red quasars are hosted by major mergers (Urrutia et al. 2008; Glikman et al. 2015). However, in this case, the reconstructed source resolution does not allow us to unambiguously reject the possibility that X could be a minor merger or even a starburst or star-forming region in the host, caused by a merger.

ORCID iDs

Eilat Glikman ⓘ https://orcid.org/0000-0003-0489-3750
Cristian E. Rusu ⓘ https://orcid.org/0000-0003-4561-4017
Geoff C.-F. Chen ⓘ https://orcid.org/0000-0002-2580-6321
James Hung-Hsu Chan ⓘ https://orcid.org/0000-0001-8797-725X
Cristiana Spingola ⓘ https://orcid.org/0000-0002-2231-6861
Hannah Stacey ⓘ https://orcid.org/0000-0002-8999-9636
Ciprian T. Berghea ⓘ https://orcid.org/0000-0001-5538-5903
S. G. Djorgovski ⓘ https://orcid.org/0000-0002-0603-3087
Matthew J. Graham ⓘ https://orcid.org/0000-0002-3168-0139
Daniel Stern ⓘ https://orcid.org/0000-0003-2686-9241
Tanya Urrutia ⓘ https://orcid.org/0000-0001-6746-9936
Mark Lacy ⓘ https://orcid.org/0000-0002-3032-1783
Nathan J. Secrest ⓘ https://orcid.org/0000-0002-4902-8077
John M. O'Meara ⓘ https://orcid.org/0000-0002-7893-1054

References

Assef, R. J., Denney, K. D., Kochanek, C. S., et al. 2011, ApJ, 742, 93
Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, ApJ, 772, 26
Assef, R. J., Stern, D., Noirot, G., et al. 2018, ApJS, 234, 23
Assef, R. J., Walton, D. J., Brightman, M., et al. 2016, ApJ, 819, 111

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Auger, M. W., Treu, T., Bolton, A. S., et al. 2009, ApJ, 705, 1099
Badole, S., Jackson, N., Hartley, P., et al. 2020, MNRAS, 496, 138
Banerji, M., McMahon, R. G., Hewett, P. C., et al. 2012, MNRAS, 427, 2275
Becker, R. H., Gregg, M. D., Hook, I. M., et al. 1997, ApJL, 479, L93
Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289
Bennert, V. N., Auger, M. W., Treu, T., Woo, J.-H., & Malkan, M. A. 2011a, ApJ, 726, 59
Bennert, V. N., Auger, M. W., Treu, T., Woo, J.-H., & Malkan, M. A. 2011b, ApJ, 742, 107
Bennert, V. N., Treu, T., Woo, J.-H., et al. 2010, ApJ, 708, 1507
Birrer, S., & Amara, A. 2018, PDU, 22, 189
Brusa, M., Bongiorno, A., Cresci, G., et al. 2015, MNRAS, 446, 2394
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
Chan, J. H. H., Rojas, K., Millon, M., et al. 2021, A&A, 647, A115
Chen, G. C.-F., Chan, J. H. H., Bonvin, V., et al. 2018, MNRAS, 481, 1115
Chen, G. C.-F., Suyu, S. H., Wong, K. C., et al. 2016, MNRAS, 462, 3457
Cisternas, M., Jahnke, K., Bongiorno, A., et al. 2011, ApJL, 741, L11
Coleman, G. D., Wu, C. C., & Weedman, D. W. 1980, ApJS, 43, 393
Condon, J. J., Kellermann, K. I., Kimball, A. E., Ivezić, Ž., & Perley, R. A. 2013, ApJ, 768, 37
Cushing, M. C., Vacca, W. D., & Rayner, J. T. 2004, PASP, 116, 362
Dai, X., Kochanek, C. S., Chartas, G., et al. 2010, ApJ, 709, 278
Dalal, N., & Kochanek, C. S. 2002, ApJ, 572, 25
Davis, S. W., & Laor, A. 2011, ApJ, 728, 98
de Vaucouleurs, G. 1948, AnAp, 11, 247
Di Matteo, T., Springel, V., & Hernquist, L. 2005, Natur, 433, 604
Ding, X., Silverman, J., Treu, T., et al. 2020, ApJ, 888, 37
Ding, X., Treu, T., Birrer, S., et al. 2021, MNRAS, 501, 269
Ding, X., Treu, T., Suyu, S. H., et al. 2017, MNRAS, 472, 90
Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, ApJ, 748, 142
Eigenbrod, A., Courbin, F., Sluse, D., Meylan, G., & Agol, E. 2008, A&A, 480, 647
Elíasdóttir, Á., Hjorth, J., Toft, S., Burud, I., & Paraficz, D. 2006, ApJS, 166, 443
Faber, S. M., & Jackson, R. E. 1976, ApJ, 204, 668
Falco, E. E., Impey, C. D., Kochanek, C. S., et al. 1999, ApJ, 523, 617
Ferrarese, L., & Merritt, D. 2000, ApJL, 539, L9
Gao, F., & Han, L. 2012, Comp. Opt. Appl., 51, 259
Gebhardt, K., Bender, R., Bower, G., et al. 2000, ApJL, 539, L13
Glikman, E., Gregg, M. D., Lacy, M., et al. 2004, ApJ, 607, 60
Glikman, E., Helfand, D. J., & White, R. L. 2006, ApJ, 640, 579
Glikman, E., Helfand, D. J., White, R. L., et al. 2007, ApJ, 667, 673
Glikman, E., Lacy, M., LaMassa, S., et al. 2018, ApJ, 861, 37
Glikman, E., Lacy, M., LaMassa, S., et al. 2022, ApJ, 934, 119
Glikman, E., LaMassa, S., Piconcelli, E., Urry, M., & Lacy, M. 2017, ApJ, 847, 116
Glikman, E., Simmons, B., Mailly, M., et al. 2015, ApJ, 806, 218
Glikman, E., Urrutia, T., Lacy, M., et al. 2012, ApJ, 757, 51
Glikman, E., Urrutia, T., Lacy, M., et al. 2013, ApJ, 778, 127
Gravity Collaboration, Sturm, E., Dexter, J., et al. 2018, Natur, 563, 657
Gregg, M. D., Lacy, M., White, R. L., et al. 2002, ApJ, 564, 133
Hall, P. B., Richards, G. T., York, D. G., et al. 2002, ApJL, 575, L51
Hamann, F., Zakamska, N. L., Ross, N., et al. 2017, MNRAS, 464, 3431
Häring, N., & Rix, H.-W. 2004, ApJL, 604, L89
Hopkins, P. F., Hernquist, L., Martini, P., et al. 2005, ApJL, 625, L71
Hunter, J. D. 2007, CSE, 9, 90
Inada, N., Becker, R. H., Burles, S., et al. 2003, AJ, 126, 666
Ivezić, Ž., Menou, K., Knapp, G. R., et al. 2002, AJ, 124, 2364
Ivison, R. J., Alexander, D. M., Biggs, A. D., et al. 2010a, MNRAS, 402, 245
Ivison, R. J., Magnelli, B., Ibar, E., et al. 2010b, A&A, 518, L31
Keeton, C. R., Gaudi, B. S., & Petters, A. O. 2003, ApJ, 598, 138

Keeton, C. R., Kochanek, C. S., & Falco, E. E. 1998, ApJ, 509, 561
Kim, D., Im, M., Glikman, E., Woo, J.-H., & Urrutia, T. 2015, ApJ, 812, 66
Kochanek, C. S. 1994, ApJ, 436, 56
Kochanek, C. S. 2004, ApJ, 605, 58
Koopmans, L. V. E., Biggs, A., Blandford, R. D., et al. 2003, ApJ, 595, 712
Lacy, M., Gregg, M., Becker, R. H., et al. 2002, AJ, 123, 2925
Lacy, M., Storrie-Lombardi, L. J., Sajina, A., et al. 2004, ApJS, 154, 166
Malhotra, S., Rhoads, J. E., & Turner, E. L. 1997, MNRAS, 288, 138
Marshall, P. J., Treu, T., Melbourne, J., et al. 2007, ApJ, 671, 1196
Matsuoka, K., Toba, Y., Shidatu, M., et al. 2018, A&A, 620, L3
McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in ASP Conf. Ser. 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell (San Francisco, CA: ASP), 127
Morgan, C. W., Kochanek, C. S., Morgan, N. D., & Falco, E. E. 2010, ApJ, 712, 1129
Nierenberg, A. M., Treu, T., Brammer, G., et al. 2017, MNRAS, 471, 2224
Nierenberg, A. M., Treu, T., Wright, S. A., Fassnacht, C. D., & Auger, M. W. 2014, MNRAS, 442, 2434
Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793
Oguri, M. 2010, PASJ, 62, 1017
Oguri, M., & Marshall, P. J. 2010, MNRAS, 405, 2579
Østman, L., Goobar, A., & Mörtsell, E. 2008, A&A, 485, 403
Park, D., Woo, J.-H., Bennert, V. N., et al. 2015, ApJ, 799, 164
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Peng, C. Y., Impey, C. D., Rix, H.-W., et al. 2006, ApJ, 649, 616
Peth, M. A., Ross, N. P., & Schneider, D. P. 2011, AJ, 141, 105
Rayner, J. T., Toomey, D. W., Onaka, P. M., et al. 2003, PASP, 115, 362
Richards, G. T., Strauss, M. A., Fan, X., et al. 2006, AJ, 131, 2766
Rusu, C. E., Masamune, O., Minowa, Y., et al. 2016, MNRAS, 458, 2
Salpeter, E. E. 1955, ApJ, 121, 161
Sanders, D. B., Soifer, B. T., Elias, J. H., et al. 1988, ApJ, 325, 74
Schechter, P. L., & Wambsganss, J. 2002, ApJ, 580, 685
Schmidt, T., Treu, T., Birrer, S., et al. 2023, MNRAS, 518, 1260
Schramm, M., & Silverman, J. D. 2013, ApJ, 767, 13
Shajib, A. J., Birrer, S., Treu, T., et al. 2019, MNRAS, 483, 5649
Shen, Y., & Liu, X. 2012, ApJ, 753, 125
Sluse, D., Chantry, V., Magain, P., Courbin, F., & Meylan, G. 2012, A&A, 538, A99
Spingola, C., Dallacasa, D., Belladitta, S., et al. 2020a, A&A, 643, L12
Spingola, C., McKean, J. P., Vegetti, S., et al. 2020b, MNRAS, 495, 2387
Stacey, H. R., McKean, J. P., Jackson, N. J., et al. 2019, A&A, 622, A18
Stacey, H. R., McKean, J. P., Powell, D. M., et al. 2021, MNRAS, 500, 3667
Stacey, H. R., McKean, J. P., Robertson, N. C., et al. 2018, MNRAS, 476, 5075
Stern, D., Assef, R. J., Benford, D. J., et al. 2012, ApJ, 753, 30
Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, ApJ, 631, 163
Stern, D., Lansbury, G. B., Assef, R. J., et al. 2014, ApJ, 794, 102
Suyu, S. H., & Halkola, A. 2010, A&A, 524, A94
Suyu, S. H., Hensel, S. W., McKean, J. P., et al. 2012, ApJ, 750, 10
Telfer, R. C., Zheng, W., Kriss, G. A., & Davidsen, A. F. 2002, ApJ, 565, 773
Tie, S., & Kochanek, C. S. 2018, MNRAS, 473, 80
Urrutia, T., Becker, R. H., White, R. L., et al. 2009, ApJ, 698, 1095
Urrutia, T., Lacy, M., & Becker, R. H. 2008, ApJ, 674, 80
Vacca, W. D., Cushing, M. C., & Rayner, J. T. 2003, PASP, 115, 389
Vanden Berk, D. E., Wilhite, B. C., Kron, R. G., et al. 2004, ApJ, 601, 692
van der Kruit, P. C. 1971, A&A, 15, 110
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Veilleux, S., Trippe, M., Hamann, F., et al. 2013, ApJ, 764, 15
Vernardos, G., & Fluke, C. J. 2014, A&C, 6, 1
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Warren, S. J., Hewett, P. C., & Foltz, C. B. 2000, MNRAS, 312, 827
Winn, J. N., Lovell, J. E. J., Chen, H.-W., et al. 2002, ApJ, 564, 143
Wong, K. C., Suyu, S. H., Auger, M. W., et al. 2017, MNRAS, 465, 4895
Yoo, J., Kochanek, C. S., Falco, E. E., & McLeod, B. A. 2006, ApJ, 642, 22
Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, ApJ, 554, 803