The Magnificent Five Images of Supernova Refsdal: Time Delay and Magnification Measurements

Patrick L. Kelly[1], Steven Rodney[2], Tommaso Treu[3], Simon Birrer[4], Vivien Bonvin[5], Luc Dessart[6], Ryan J. Foley[7], Alexei V. Filippenko[8,9], Daniel Gilman[3,10], Saurabh Jha[11], Jens Hjorth[12], Kaisey Mandel[13,14], Martin Millon[5], Justin Pierel[2], Stephen Thorp[13], Adi Zitrin[15], Tom Broadhurst[16], Wenlei Chen[1], Jose M. Diego[17], Alan Dressler[18], Or Graur[19,20], Mathilde Jauzac[21,22,23], Matthew A. Malkan[3], Curtis McCully[24,25], Masamune Oguri[26,27,28], Marc Postman[29], Kasper Borello Schmidt[30], Keren Sharon[31], Brad E. Tucker[32,33,34], Anja von der Linden[35], and Joachim Wambsganss[36,37]

[1] Minnesota Institute for Astrophysics, University of Minnesota, 116 Church Street SE, Minneapolis, MN 55455, USA
[2] Department of Physics and Astronomy, University of South Carolina, 712 Main St., Columbia, SC 29208, USA
[3] Department of Physics and Astronomy, University of California, Los Angeles, CA 90095, USA
[4] Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA
[5] Laboratoire d'Astrophysique, Ecole Polytechnique Federale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland
[6] Institut d'Astrophysique de Paris, CNRS-Sorbonne Université, 98 bis boulevard Arago, F-75014 Paris, France
[7] Department of Astronomy and Astrophysics, UCO/Lick Observatory, University of California, 1156 High Street, Santa Cruz, CA 95064, USA
[8] Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA
[9] Miller Institute for Basic Research in Science, University of California, Berkeley, CA 94720, USA
[10] Department of Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada
[11] Department of Physics and Astronomy, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA
[12] Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark
[13] Institute of Astronomy and Kavli Institute for Cosmology, Madingley Road, Cambridge, CB3 0HA, UK
[14] Statistical Laboratory, DPMMS, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WB, UK
[15] Physics Department, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 8410501, Israel
[16] Ikerbasque Foundation, University of the Basque Country, DIPC Donostia, Spain
[17] IFCA, Instituto de Física de Cantabria (UC-CSIC), Av. de Los Castros s/n, E-39005 Santander, Spain
[18] The Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA
[19] Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK
[20] Department of Astrophysics, American Museum of Natural History, Central Park West and 79th Street, New York, NY 10024, USA
[21] Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham DH1 3LE, UK
[22] Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK
[23] Astrophysics and Cosmology Research Unit, School of Mathematical Sciences, University of KwaZulu-Natal, Durban 4041, South Africa
[24] Las Cumbres Observatory, 6740 Cortona Dr., Suite 102, Goleta, CA 93117, USA
[25] Department of Physics, University of California, Santa Barbara, CA 93106-9530, USA
[26] Research Center for the Early Universe, University of Tokyo, Tokyo 113-0033, Japan
[27] Center for Frontier Science, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan
[28] Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8583, Japan
[29] Space Telescope Science Institute, 3700 San Martin Dr., Baltimore, MD 21218, USA
[30] Leibniz-Institut fur Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany
[31] University of Michigan, Department of Astronomy, 1085 South University Avenue, Ann Arbor, MI 48109-1107, USA
[32] The Research School of Astronomy and Astrophysics, Mount Stromlo Observatory, Australian National University, Canberra, ACT 2611, Australia
[33] National Centre for the Public Awareness of Science, Australian National University, Canberra, ACT 2611, Australia
[34] The ARC Centre of Excellence for All-Sky Astrophysics in 3 Dimension (ASTRO 3D), Australia
[35] Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA
[36] Astronomisches Rechen-Institut (ARI), Zentrum für Astronomie der Universität Heidelberg (ZAH), Mönchhofstr. 12-14, D-69120 Heidelberg, Germany
[37] International Space Science Institute (ISSI), Hallerstrasse 6, 3012 Bern, Switzerland

Abstract

In late 2014, four images of supernova (SN) "Refsdal," the first known example of a strongly lensed SN with multiple resolved images, were detected in the MACS J1149 galaxy-cluster field. Following the images' discovery, the SN was predicted to reappear within hundreds of days at a new position $\sim 8''$ away in the field. The observed reappearance in late 2015 makes it possible to carry out Refsdal's original proposal to use a multiply imaged SN to measure the Hubble constant H_0, since the time delay between appearances should vary inversely with H_0. Moreover, the position, brightness, and timing of the reappearance enable a novel test of the blind predictions of galaxy-cluster models, which are typically constrained only by the positions of multiply imaged galaxies. We have developed a new photometry pipeline that uses DOLPHOT to measure the fluxes of the five images of SN Refsdal from difference images. We apply four separate techniques to perform a blind measurement of the relative time delays and magnification ratios between the last image SX and the earlier images S1–S4. We measure the relative time delay of SX–S1 to be $376.0^{+5.6}_{-5.5}$ days and the relative magnification to be $0.30^{+0.5}_{-0.3}$. This corresponds to a 1.5% precision on the time delay and 17% precision for the magnification ratios and includes uncertainties due to

millilensing and microlensing. In an accompanying paper, we place initial and blind constraints on the value of the Hubble constant.

Unified Astronomy Thesaurus concepts: Hubble constant (758); Gravitational lensing (670); Type II supernovae (1731); Cosmology (343)

Supporting material: machine-readable table

1. Introduction

By deflecting the trajectories of photons and increasing their travel time, a sufficiently massive object can produce multiple, magnified images of a well-aligned background object. Einstein (1936) first described the possibility of a strong gravitational lens system, envisioning a hypothetical foreground star that creates multiple images of a second, more distant star. Zwicky (1937a, 1937b) subsequently conjectured that external galaxies ("extragalactic nebulae") and galaxy clusters should magnify and produce multiple images of background galaxies far beyond the Milky Way.

Even before the first multiply imaged source had been discovered, Refsdal (1964) explored the idea that a supernova (SN) could explode in a strongly lensed galaxy. The light curves of the SN observed in each of its multiple images could be used to measure the relative time delays among the images. Sjur Refsdal's analysis demonstrated that these relative time delays could be used to infer the Hubble constant (H_0) given an accurate model for the mass distribution of the lens.

Despite the scientific potential of such a strongly lensed SN with multiple resolved images, none were identified in the 50 yr after Refsdal's paper was published. In Hubble Space Telescope (HST) images taken on 2014 November 10 (UT dates are used throughout this paper), Kelly et al. (2015) identified the first example of such a strongly lensed SN with multiple resolved images. This event, dubbed SN Refsdal, appeared behind the MACS J1149.5+2223 cluster (Ebeling et al. 2001, 2007) in imaging collected for the Grism Lens-Amplified Survey from Space (GLASS; GO-13459; PI Treu; Schmidt et al. 2014; Treu et al. 2015). When SN Refsdal was detected in 2014, it appeared as four images (named S1–S4) in an Einstein-cross configuration around an early-type galaxy in the foreground, as shown in Figure 1. Lens models predicted that it should subsequently reappear on a timescale of hundreds of days in a trailing image (named SX) of its host galaxy that forms close to the cluster's center (Kelly et al. 2015; Oguri 2015; Sharon & Johnson 2015; Diego et al. 2016; Grillo et al. 2016; Jauzac et al. 2016; Kawamata et al. 2016; Treu et al. 2016). As demonstrated by Rodney et al. (2016) in their analysis of the light curves of the Einstein-cross images, SN Refsdal is the first example of a strongly lensed SN where relative time delays can be measured. When the SN reappeared in HST imaging taken on 2015 December 11 (Kelly et al. 2016c; GO-14199; PI Kelly), SN Refsdal became the first (and only, to date) SN whose appearance on the sky was predicted in advance. The MACS J1149 galaxy cluster produces a third, complete image of SN Refsdal's host galaxy, and lens models indicate that an image of the SN appeared within that galaxy image in the late 1990s.

The first multiply imaged SN turned out to be even more scientifically useful than originally envisioned by Refsdal (see Oguri 2019 for a review of applications of strongly lensed transients). From the very first observations of the SN's reappearance, Kelly et al. (2016b) found approximate and preliminary constraints on image SX's relative time delay and magnification. Using a Bayesian framework and available model predictions, Vega-Ferrero et al. (2018) obtained a ~30% constraint on the value of H_0 from the multiply imaged SN across a 68% confidence interval. In addition to its potential to measure the Hubble constant, the SN can be used to test the accuracy of lens models on the scale of a galaxy cluster (Treu et al. 2016). Furthermore, the monitoring of SN Refsdal has led to the characterization of a rare core-collapse explosion of a blue supergiant star at redshift $z = 1.49$ (Kelly et al. 2016a) and to the serendipitous discovery of an individual star in SN Refsdal's host galaxy extremely magnified by microlensing (Kelly et al. 2018).

In this paper, we present a blinded analysis for measurement of the time delay and relative magnifications between SX and S1–S4. In an accompanying paper (Kelly et al. 2023), we then use these high-precision measurements to obtain an estimate of H_0 from SN Refsdal using pre-reappearance lens models updated using only the reappearance's position. In the accompanying paper, we also use the measurements presented here to evaluate MACS J1149 cluster lens models.

This paper is organized as follows. Section 2 summarizes the observations and previous analyses of SN Refsdal. The methods used in data processing and photometry are presented in Section 3. In Section 4, we include a description of our simulations of microlensing and millilensing that affect images S1–SX. Section 5 describes the process of fitting the SN Refsdal light curves to determine time delays and relative magnifications. Our measurements of the time delays and relative magnifications are presented in Section 6. We summarize the results and discuss future work in Section 7. Throughout this paper, we use AB magnitudes (Oke & Gunn 1983).

2. Follow-up Observations and Early Analysis

Following the discovery of the four images S1–S4 of SN Refsdal in GLASS imaging (Kelly et al. 2015), the Hubble Frontier Fields (HFF) program (DD-13504; PI Lotz; Lotz et al. 2017) fortuitously began its already-scheduled imaging campaign of the MACS J1149 galaxy-cluster field. To complement the early HFF imaging, we acquired additional imaging as part of the FrontierSN program (GO-13386; PI Rodney) to characterize better the spectral energy distribution (SED) of the SN from photometry of images S1–S4. These early measurements of the SN color and light curve left open the possibility that the SN could be a Type Ia SN whose luminosity could be calibrated within ~10%, and we obtained 30 orbits of WFC3 IR G141 grism spectroscopy to classify spectroscopically the SN before it reached maximum brightness as part of a Director's Discretionary Time program (DD-14041; PI Kelly; Kelly et al. 2016a). Independent analysis of the SN's explosion site using MUSE spectroscopy showed that it erupted in a highly ionized region with low metallicity (Yuan et al. 2015; Karman et al. 2016).
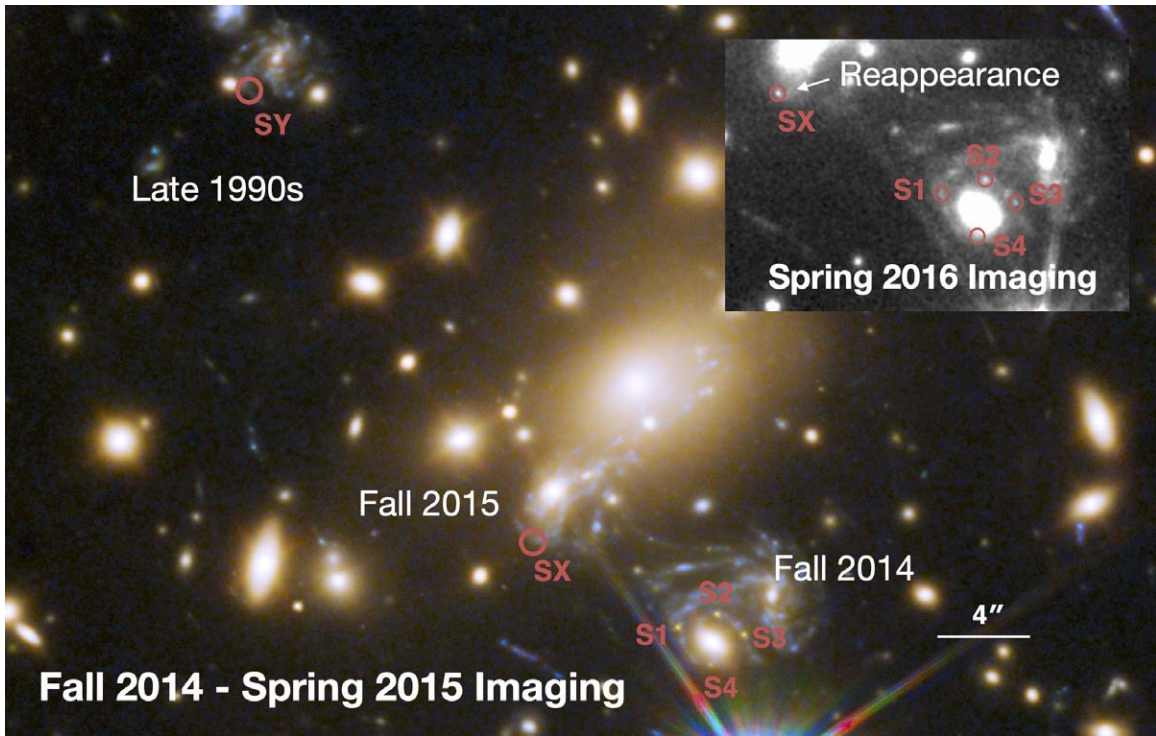
Figure 1. Image of the MACS J1149 galaxy-cluster field and the locations and timing of the appearances of SN Refsdal. A first appearance occurred according to model predictions in the host galaxy image at the upper left in the late 1990s but was missed. We detected images S1–S4 in an Einstein-cross configuration in 2014 November (Kelly et al. 2015). Following this appearance, lens modelers predicted the SN's reappearance, which was detected in late 2015 (Kelly et al. 2016c). Inset at upper right shows a co-added WFC3 IR F125W image of the field following the reappearance of the SN in image SX.

Table 1
HST WFC3 IR Imaging Data

Dates	Short Program Name	HST Program ID+PI	Reference
2010-12-05..2011-03-10	CLASH	GO-12068; M. Postman	Postman et al. (2012)
2013-11-02..2015-01-06	HFF	DD/GO-13504; J. Lotz	Lotz et al. (2017)
2014-11-03..2014-11-11	GLASS	GO-13459; T. Treu	Treu et al. (2015)
2014-11-30..2015-07-21	FrontierSN	GO-13790; S. Rodney	Rodney et al. (2015)
2014-12-23..2015-01-05	Refsdal-DD	DD-14041; P. Kelly	Kelly et al. (2015)
2015-10-30..2016-10-30	Refsdal Redux	GO-14199; P. Kelly	Kelly et al. (2016c)
2016-07-11..2016-07-23	Icarus-DD1	DD-14528; P. Kelly	Kelly et al. (2018)
2016-11-28..2017-01-04	Icarus-DD2	DD-14872; P. Kelly	Kelly et al. (2018)

Three of the images of SN Refsdal that form the Einstein cross (S1–S3) are magnified by factors of up to $\sim$15–22, while the fourth (S4) was predicted to have a magnification of $\sim$1–4 (Kelly et al. 2015; Oguri 2015; Sharon & Johnson 2015; Diego et al. 2016; Grillo et al. 2016; Jauzac et al. 2016; Kawamata et al. 2016; Treu et al. 2016). By fitting photometry of imaging acquired from 2014 November through 2015 December, Rodney et al. (2016) measured the time delays and magnification ratios of the four images S1–S4 that form the Einstein cross. Those measurements used both polynomials and templates of SN 1987A–like SNe and compared the resulting measurements to lens model predictions.

To detect the reappearance of SN Refsdal in image SX, as shown in Figure 1, and measure its time delay and magnification relative to S1–S4, we executed a multicycle "Refsdal Redux" monitoring campaign (GO-14199; PI Kelly). This program continued to monitor the MACS J1149

field at a cadence of approximately every 2 weeks whenever the field was visible to HST (from approximately late October through July). In 2016 May these observations captured a microlensing peak of a highly magnified blue supergiant star (nicknamed Icarus) in SN Refsdal's host galaxy (Kelly et al. 2018). This prompted additional HST follow-up observations of the MACS J1149 field (DD-14528; DD-14872; PI Kelly).

Table 1 provides a summary of all these observation programs. As many of these programs were designed to help measure the relative time delays and magnifications of the images of SN Refsdal, a large fraction of the observing time used a consistent instrumental setup. The HST Wide Field Camera 3 (WFC3) infrared (IR) detector was employed, using filters F125W (J) and F160W (H), which correspond to approximately the rest-frame V and R bands. This filter choice facilitates comparison with observations of nearby SNe.

3. Image Processing and Photometry

In this section we describe the processing of the WFC3 IR imaging, an approach to subtract stellar diffraction spikes, and photometry of the five SN Refsdal images.

3.1. Astrometry and Alignment

We processed all HST images using the `DrizzlePac` software suite.[38] All images were first aligned to a common world coordinate system (WCS) using `TweakReg`. We adopted the WCS used for the CLASH, GLASS, and HFF images and catalogs,[39] which is also the WCS generally used by published lens models for this cluster. We then co-added and resampled each image set to a scale of $0\overset{''}{.}03\ \mathrm{pixel}^{-1}$ using `AstroDrizzle` (Fruchter et al. 2010).

3.2. Subtraction of Stellar Diffraction Spikes

In the WFC3 IR imaging of the MACS J1149 field, a bright foreground star ($J = 14.0$ mag Vega) offset by $\sim 6''$ from images S1–S4 produces four bright diffraction spikes, two of which can be seen near the bottom of Figure 1. We requested telescope roll angles for which the stellar diffraction spikes fell away from the SN images. However, for a significant fraction of observations no such angle was available.

The diffraction spikes require additional effort to obtain a measurement of the SN flux when they overlap the SN image, but we have found that it is possible to remove the diffraction spikes almost entirely. We performed all photometry on "difference images" that were constructed by subtracting a pre-SN "template" image from the co-added images of each visit containing the SNe. To create a template image without diffraction spikes, we constructed masks that enclose pixels contaminated by the diffraction spikes. We next constructed template images by combining all WFC3 IR images acquired before 2014 October 30 (preceding the first detection of the SN) and applying the pixel masks. The images that we use to construct the template include data acquired at multiple telescope roll angles, so the co-added template images have no regions that are completely masked.

As shown by Rodney et al. (2016), the similarity among the star's four diffraction spikes and rotation of the diffraction spikes make it possible to remove effectively if not entirely diffracted light at SN positions. We first co-add the single exposures taken during each visit and then subtract the template image from the resulting co-added image. As described above, the template images we created lack diffraction spikes. By subtracting the template from each single-epoch image, we removed the light from all galaxies and the sky background. Consequently, the difference images only show light collected during the follow-up epoch, and only from two sources: the multiply imaged SN and the diffraction spikes.

We next rotate the difference images by 90°, after setting all regions outside of those contaminated by the diffraction spikes to equal zero. Given the similarity among the four diffraction spikes, each rotated spike provides an excellent model for the adjacent spikes. After subtracting this masked and rotated image of the spikes, we are left with a difference image that has

Table 2
J2000 Positions of the Five Detected Images of SN Refsdal

Image	R.A.	Decl.
S1	$11^{\mathrm{h}}49^{\mathrm{m}}35\overset{\mathrm{s}}{.}575$	$22°23'44\overset{''}{.}25$
S2	$11^{\mathrm{h}}49^{\mathrm{m}}35\overset{\mathrm{s}}{.}454$	$22°23'44\overset{''}{.}82$
S3	$11^{\mathrm{h}}49^{\mathrm{m}}35\overset{\mathrm{s}}{.}370$	$22°23'43\overset{''}{.}92$
S4	$11^{\mathrm{h}}49^{\mathrm{m}}35\overset{\mathrm{s}}{.}474$	$22°23'42\overset{''}{.}63$
SX	$11^{\mathrm{h}}49^{\mathrm{m}}36\overset{\mathrm{s}}{.}022$	$22°23'48\overset{''}{.}10$

an increase in noise in a narrow region along the spike, but without substantial contamination.

3.3. Positions of the SN Images

In Table 2, we list the coordinates of images S1–SX measured in the CLASH astrometric system from co-additions of WFC3 IR F125W imaging acquired in 2015 and 2016.

3.4. Photometry with `DOLPHOT`

Fitting a model of the point-spread function (PSF) to an unresolved point source such as S1–SX should, in principle, yield a measurement of the source's flux with an optimal signal-to-noise ratio (S/N). For a point source, pixels *closest* to the source's center have the greatest S/N. When fitting a PSF model to the data, one can assign greater weights to these central pixels. By contrast, when measuring instead an aperture magnitude, the flux within a fixed radius is summed and an aperture correction is applied, and *all* pixels in the aperture are assigned effectively the same weight.

We have adapted the `DOLPHOT` (Dolphin 2000) PSF-fitting package to be able to measure photometry from WFC3 IR difference imaging. When applying `DOLPHOT`, we use the default sky-fitting setting (FitSky=1) and model images S1–SX as point sources (Force1=1). When fitting images S1–SX, we instruct `DOLPHOT` to assign additional weight to central pixels (PSFPhot=2). Finally, we enforce the additional constraint that the flux of S1–SX acquired does not change between exposures taken within each imaging epoch, except for variation due to noise (ForceSameMag=0).

We estimate the uncertainty of the measured fluxes by using `DOLPHOT` to inject fake point sources into individual exposures. We next measure the fluxes of the fake sources and estimate their uncertainty by computing the standard deviation of measured fluxes.

For the purpose of comparison, we use `PythonPhot`[40] (Jones et al. 2015) to measure the light curves of images S1–SX of SN Refsdal from co-added imaging. The `PythonPhot` package includes an implementation of PSF-fitting photometry measurement based on the `DAOPHOT` package (Stetson 1987). We have modified the code to inject fake stars at locations at which the local background is similar to the locations of images S1–SX.

The measured photometry for all five images is shown in Figure 2, along with a piecewise polynomial model fit (which will be described in Section 3.6). A full table of the measured photometry is available in the online version of the journal. Table 3 provides an excerpt.
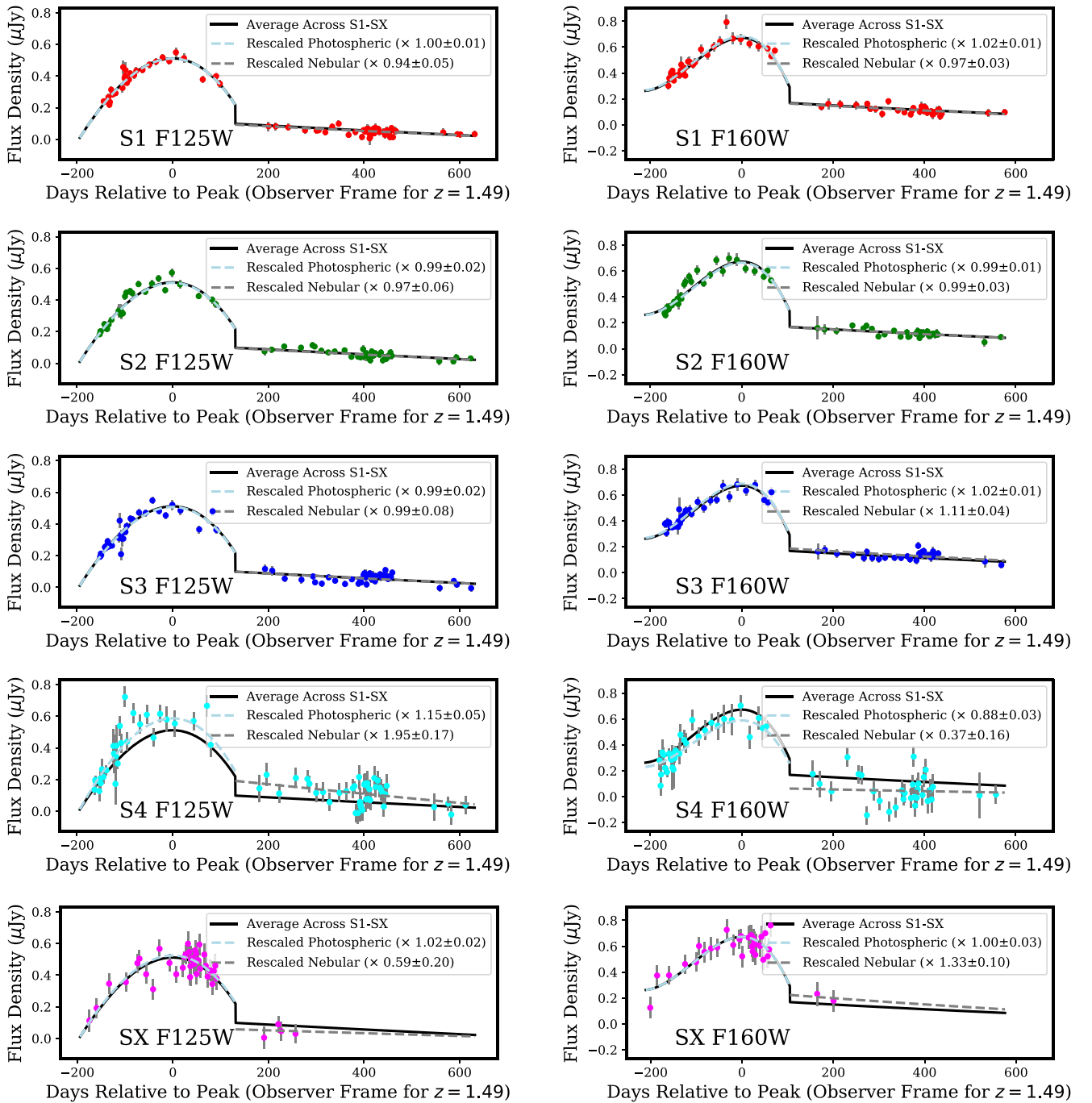
Figure 2. Photometry of the five images of SN Refsdal and light-curve fits using the fiducial light-curve model. The fiducial piecewise polynomial model is shown fit simultaneously to the F125W and F160W (rest-frame approximately *V* and *R*) light curves of images S1–SX (solid). The dashed line shows the same model fit after rescaling to match the photospheric and nebular light curves of each image separately. Each legend lists the respective best-fitting factor used to rescale each light-curve component, and we list the ratios between the photospheric- and nebular-phase factors in Table 4. Microlensing that is not achromatic could cause the apparent color of each image to become comparatively bluer or redder. Although not expected given that the galaxy-scale lens is an early-type galaxy, extinction by dust would cause the color to become redder. The color of image S4 is much *bluer* than the average color of Refsdal images.

3.5. Light Curves of Stellar Sources in Cluster Field

To assess the performance of the photometry pipelines, we selected a set of stars from a co-added image of the field acquired with the ACS WFC F606W broadband filter and required that the stars' F125W and F160W fluxes were similar to those of S1–SX over the evolution of their observed light curves. We next ran DOLPHOT in an image section around the location of each star and determined the star's position in each image. Using these positions, we then fit for the position, proper motion, and annual parallax. Finally, we ran DOLPHOT and PythonPhot on each of the comparison stars and calculated the standard deviation and reduced χ-squared statistics of the flux measurements.

We estimate the photometric uncertainties by injecting 100 fake point sources using DOLPHOT and PythonPhot. In the case of DOLPHOT, we select areas near each star without significant contamination. In Figures A1–A5, we plot the photometry of the stars, showing a comparison between the results of the photometry codes. From the DOLPHOT tests we measure a reduced χ^2 of 0.59 −1.82 with a median of 1.06, which affirms the accuracy of our

Table 3
Photometry of SN Refsdal

Obj.	Filt.	MJD	Flux	Flux Err.
S1	F125W	57,004.89	0.240	0.012
S1	F125W	57,006.34	0.237	0.009
S1	F125W	57,007.1	0.239	0.011
S2	F125W	57,013.746	0.200	0.011
S2	F125W	57,015.152	0.221	0.009
S2	F125W	57,015.96	0.218	0.010
S3	F125W	56,957.17	0.187	0.011
S3	F125W	56,958.68	0.206	0.010
S3	F125W	56,959.38	0.197	0.010
S4	F125W	56,994.383	0.042	0.010
S4	F125W	56,995.83	0.061	0.009
S4	F125W	56,996.59	0.059	0.011
SX	F125W	57,004.527	0.009	0.010
SX	F125W	57,006.027	−0.004	0.008
SX	F125W	57,006.74	−0.002	0.009

Note. A portion of the table is shown here. The full table is available online.

(This table is available in its entirety in machine-readable form.)

Table 4
Anomalous Color of Image S4 of SN Refsdal

Object	Filter	Ratio	Uncertainty
S1	F125W	0.968	0.031
S2	F125W	0.982	0.035
S3	F125W	0.991	0.046
S4	F125W	1.561	0.110
SX	F125W	0.979	0.035
S1	F160W	0.996	0.017
S2	F160W	0.989	0.016
S3	F160W	1.063	0.023
S4	F160W	0.709	0.066
SX	F160W	1.016	0.028

Note. Ratios between measured fluxes and model light curve that best fits the light curves of all images. The best-fitting model accounts for the fluxes of all images only by changing the relative magnifications of the image.

uncertainty estimates for our DOLPHOT photometry of the SN images. These test cases show that the DOLPHOT photometry provides more reliable photometry and uncertainty estimates. We therefore adopt the DOLPHOT photometry for all SN images in the following analysis.

3.6. The Fiducial Light-curve Model

To estimate the uncertainties and to evaluate and combine estimates of the time delay fitting methods that we will apply to the SN Refsdal data (Section 5), we create simulated photometric data that approximate the real light-curve data. To achieve this, we adopt a piecewise polynomial model as our "fiducial" light-curve model.

Spectroscopic follow-up observations and the peculiar light curve of SN Refsdal have shown that it was an SN 1987A–like SN with a peak B-band luminosity inferred to be a factor of 5–10 brighter than that of SN 1987A, given magnification favored by current cluster models (Kelly et al. 2016a). As shown in Figure 3, a piecewise polynomial function with separate, low-order polynomial functions for the photospheric and nebular phases can describe the light curves of both SN
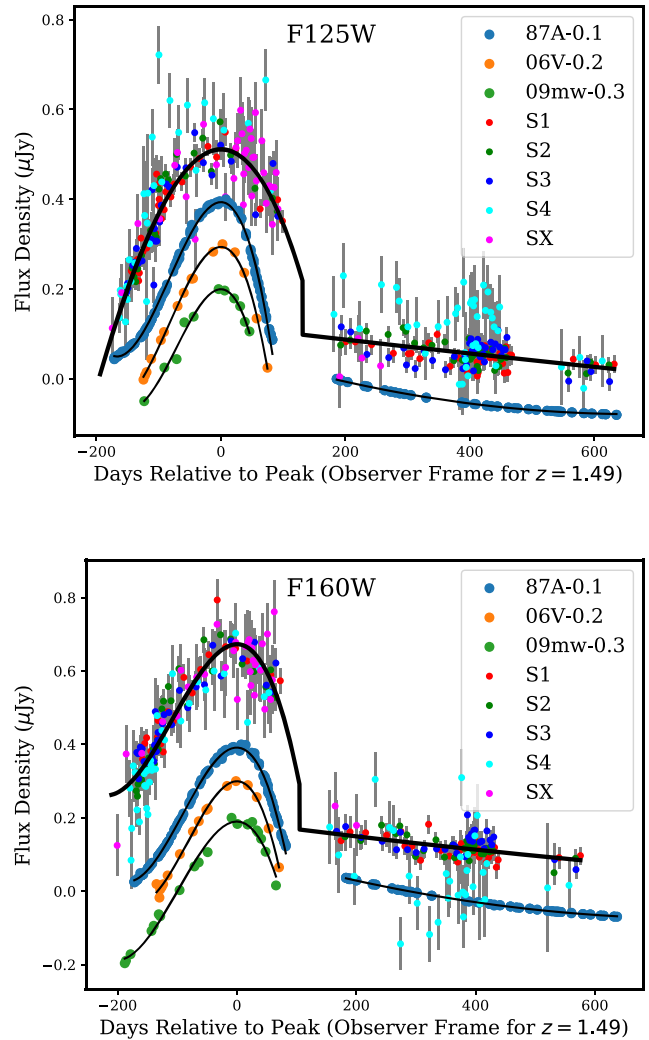

Figure 3. The observer-frame $(z = 1.49)$ broadband light curves at 5000–7000 Å of SN Refsdal and those of low-redshift SN 1987A–like SNe with a similar shape stretched in time by a factor of 2.49. All light curves are well fit by a piecewise polynomial function. In the fits shown above, the photospheric phase light curve is modeled using a third-order polynomial and the nebular-phase light curve with a second-order polynomial. The uppermost points and best-fitting model are of SN Refsdal's F125W and F160W combined light curves constructed by shifting and rescaling the measured photometry of its five appearances according to the best-fitting time delays and magnifications relative to that of S1. The F125W and F160W bandpasses correspond approximately to rest-frame V and R bandpasses at $z = 1.49$, and we plot in respective panels above the measured V and R for the low-redshift examples of SN 1987A–like SNe. For visual clarity we rescale and then subtract constant flux values from the light curves of the nearby SNe.

Refsdal and nearby SN 1987A–like SNe whose light curves have broadly similar shapes. Indeed, Rodney et al. (2016) found that a low-order polynomial could describe the SN Refsdal data taken through the photospheric phase.

Table 5 describes fit statistics, including Akaike information criterion (AIC; Akaike 1974) values, for possible low-order models for each phase. While a fourth-order or fifth-order model for the photospheric phase and a first-order model for the nebular phase are preferred according to the AIC, we select a third-order model for the photospheric phase and a second-order model for the nebular phase. We choose the lower-order model for the photospheric phase because time-variable

Table 5
Piecewise Model Comparison

Phot.	Neb.	χ^2	n_{par}	AIC	χ^2_ν	Δt_{X1}
2	1	993.1	18	1029.1	1.66	372.254953
3	1	984.4	20	1024.4	1.65	374.454781
4	1	958.0	22	1002.0	1.61	379.730981
5	1	952.9	24	1000.9	1.61	375.789274
6	1	1037.9	26	1089.9	1.76	365.728932
2	2	993.1	20	1033.1	1.66	372.251261
3	2	1009.7	22	1053.7	1.70	374.956267
4	2	980.9	24	1028.9	1.65	381.557306
5	2	1036.8	26	1088.8	1.75	367.115460
6	2	993.1	28	1049.1	1.69	380.966961

Note. Results from fitting piecewise models with photospheric and nebular components, each of varying order.

Table 6
Lensing Parameters

Name	κ	γ	μ	Parity	μ_t	μ_r
S1	0.722	0.093	14.6	$+$	5.4	2.7
S2	0.733	0.361	16.8	$-$	-10.6	1.5
S3	0.675	0.230	19.0	$+$	10.5	1.8
S4	0.727	0.466	7.0	$-$	-5.2	1.4
SX	0.966	0.488	4.2	$-$	-2.2	1.9

Note. We measure these parameters from the v3 GLAFIC (Oguri 2010; Kawamata et al. 2016) maps posted online at https://archive.stsci.edu/pub/hlsp/frontier/macs1149/models/glafic/v3/.

microlensing magnification may introduce structure into the observed light curves that we do not want to include as part of the fiducial model. Moreover, as Figure 3 demonstrates, nearby SN 1987A–like SN light curves (which lack microlensing) in similar rest-frame bands are well described by polynomials with this order. We select the higher, second-order model for the nebular phase, given that its evolution with time is expected to follow the nonlinear, exponential rate of energy deposition from the decay of radioactive ^{56}Co.

4. Microlensing and Millilensing Simulations

With a model for the SN light curve, we next simulate the effects of microlensing and millilensing on the light curves of SN Refsdal. In order to include these effects, we also need an estimate of the convergence (κ) and shear (γ) at the positions of the SN images. Furthermore, an estimate for the stellar mass density and the substructure fraction along the line of sight to each of those images are needed. Here we explain how we infer these and use them to simulate the effects of microlensing and millilensing.

4.1. Convergence and Shear

We adopt the GLAFIC (Oguri 2010) estimates for the lensing parameter from a model that was developed for the MACS J1149 field and updated to predict SN Refsdal's reappearance (Kawamata et al. 2016). This model was among the most accurate in matching the observed time delays and magnifications for the images S1–S4 (Rodney et al. 2016). After unblinding, we found that it best matches the full set of observations of SN

Refsdal in our accompanying paper (Kelly et al. 2023). Many models have been developed of the cluster lens and could also have been used for this purpose. Table 6 reports the relevant lens parameters from this model, used in the subsequent principal set of simulations. To investigate the importance of these parameters, we repeat our entire analysis using alternative lens parameters listed in Table 7.

4.2. Density of Stellar Mass along Sight Lines

The stars along the line of sight to each image of SN Refsdal comprise multiple components. These can include (1) the stellar populations in the early-type galaxy lens responsible for the Einstein-cross images S1–S4, (2) the MACS J1149 brightest cluster galaxy (BCG), and (3) the intracluster light (ICL). We model the BCG and ICL together as a single underlying stellar population with the same properties (age, initial mass function (IMF), etc.). Here we describe our models for these stellar populations, and we then describe the ray-tracing calculations that we use to compute simulated magnification maps in the source plane at $z = 1.49$. Finally, we briefly describe our simulations of millilensing by expected dark matter subhalos.

4.2.1. Models of the Early-type Lens, BCG, and ICL

To model the light distribution of the early-type lens and the BCG+ICL, we first extracted a $33'' \times 39''$ (east–west × north–south, respectively) subsection of the HFF co-added images. Using SExtractor, we next identified a set of pixels associated with each detected object (a "segmentation" image). We assembled a merged mask of all sources except for the BCG and used GALFIT (Peng et al. 2002) to fit a Sérsic model centered on the BCG to the WFC3 F125W image subsection using the mask. During the fit, the sky is fixed to be zero, since the HFF imaging background is already removed. Morishita et al. (2017) found that the background estimate used to construct the HFF co-added images was accurate for the MACS J1149 field.

The best-fitting Sérsic model for the single profile fit to the light distribution of the BCG, which transitions to that of the ICL, in the F125W bandpass has a Sérsic index of 4.86, an effective radius of $19''.6$, and an F125W magnitude of 16.18 AB. Holding all parameters except for the total flux fixed, we next perform fits to subsections of the ACS WFC F435W, F606W, and F814W and WFC3 IR F105W, F125W, F140W, and F160W HFF co-additions. We subtract the best-fitting model for the BCG+ICL in each bandpass from the $33'' \times 39''$ image subsection. The top panel of Figure 4 shows the region around the SN images S1–S4 after subtracting this BCG+ICL model.

The next step is to fit the early-type galaxy lens in the BCG+ICL-subtracted F125W image. We create a mask that excludes all pixels except those inside of an elliptical aperture with the same position angle, center, and ellipticity as the early-type galaxy. The aperture's semimajor axis is $1.8''$. Circular masks with radii of $0''.15$ are also placed over the locations of images of S1–S4 of SN Refsdal. Our GALFIT model consists of a Sérsic model for the early-type galaxy lens and a background level that accounts for flux from the superposed lensed host galaxy.

From the F125W image we find a Sérsic index of 2.22, an effective radius of $0''.26$, and a total magnitude of

Table 7
Lens Model Parameters Used for Alternative Microlensing Simulations

Model	Reference	S1		S2		S3		S4		SX	
		κ	γ	κ	γ	κ	γ	κ	γ	κ	γ
Oguri-a	HFF Lens Models (v3)	0.722	0.093	0.733	0.361	0.675	0.230	0.727	0.466	0.966	0.488
Grillo-g	CLASH-VLT Clusters	0.726	0.078	0.771	0.330	0.678	0.154	0.796	0.431	0.963	0.458
Sharon-g	HFF Lens Models (v4corr)	0.775	0.087	0.780	0.292	0.725	0.177	0.837	0.399	0.997	0.490

Note. Values of κ and γ at the positions of the five observed images of SN Refsdal. The Sharon values correspond to those from the v4corr model. For the Grillo-g and Sharon-g models, we linearly interpolate the κ and γ maps.
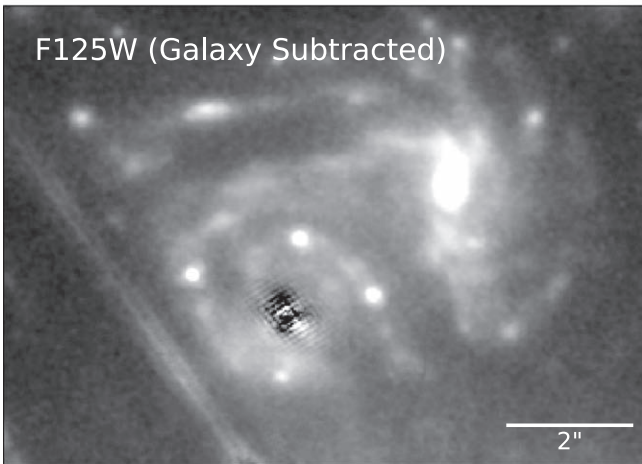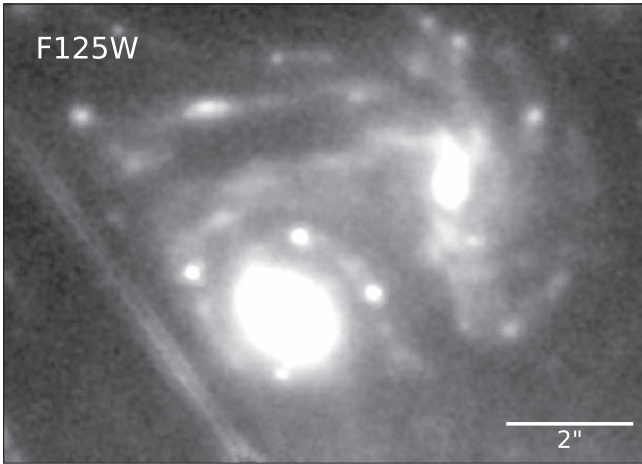
Figure 4. HST WFC3 F125W co-added image before and after the best-fitting Sérsic model of the early-type galaxy lens ($z = 0.54$) is subtracted. The top panel shows the image before subtraction of the early-type galaxy, and the bottom panel shows the result after subtraction. In a previous and separate step, the cluster's BCG has already been modeled and subtracted to create the "before" image. The four point sources in the Einstein-cross configuration are images S1–S4 of SN Refsdal, and the underlying distorted spiral galaxy is the host galaxy of the SN ($z = 1.49$).

F125W = 20.16 AB for the early-type lens. The bottom panel of Figure 4 shows the S1–S4 region after this early-type galaxy has been subtracted.

We next repeat the fit for the BCG+ICL-subtracted co-added image sections in all other available HST bandpasses from 4000 Å to 1.6 μm, holding all parameters fixed except for the galaxy's total flux and the local surface brightness. Figure 5 shows the resulting measured SEDs for the BCG+ICL, the early-type galaxy lens, and the host galaxy of SN Refsdal. The inferred SED of the BCG+ICL matches that of the early-type galaxy.

4.2.2. SED Fitting to Estimate the Stellar Mass Density at the SN Image Positions

We next want to model the measured SEDs to derive estimates for the physical properties of the BCG+ICL and the elliptical galaxy lens. To that end, we correct the fluxes for Galactic extinction using Schlafly & Finkbeiner (2011) and then use the Fitting and Assessment of Synthetic Templates (FAST) software tool (Kriek et al. 2009) and the Bruzual & Charlot (2003, hereafter BC03) stellar population model to model the SEDs shown in Figure 5. Our fitting uses the BC03 isochrones, a Chabrier (2003) IMF, and an exponentially declining star formation history; moreover, it assumes no extinction due to dust associated with the BCG, ICL, or early-type lens. The stars are set to have initial masses that are between 0.1 and 100 $M_\odot$, and we include models with metallicities Z of 0.004, 0.008, 0.02, and 0.05. We measure $\log(M_\odot/f_{F125W})$, where f_{F125W} has units of μJy, to be 9.01 and 9.15 dex for the combined BCG and ICL model fit and the early-type galaxy lens, respectively. Table 8 reports the inferred stellar mass and total mass (i.e., including dark matter) along the line of sight to each of the five SN Refsdal images.

4.3. Simulated Microlensing Magnification Maps

Using the estimated shear and convergence along the lines of sight to each image, we generated caustic magnification patterns in the source plane with an adapted version of the MULE inverse ray-tracing code (Dobler et al. 2015). However, we obtain similar results instead using the microlens code (Wambsganss 1990, 1999). We next use a spherically symmetric model for SN 1987A (Dessart & Hillier 2019) at phases of 10.9, 21.3, 31.1, 41.5, 50.1, 66.8, 80.8, 97.8, 117.0, 177.0, 220.0, and 300.0 days to generate the microlensing simulations. We rescale the radius of the photosphere models linearly with time and interpolate between the rescaled model to approximate the effect of the expansion of the SN during its photospheric phase.

In our simulation, each pixel has a length of 0.002 Einstein radii for a solar-mass lens. For the photospheric expansion velocity of -8356 ± 1105 km s^{-1} at a phase of -47 ± 8 days relative to maximum brightness measured by Kelly et al. (2016a), the SN photosphere would expand in radius by a single pixel each 1.25 days in the rest frame at $z = 1.49$. Each ray-tracing simulation creates a map of the magnification across a grid of 1024×1024 pixels, and we place the simulated SN
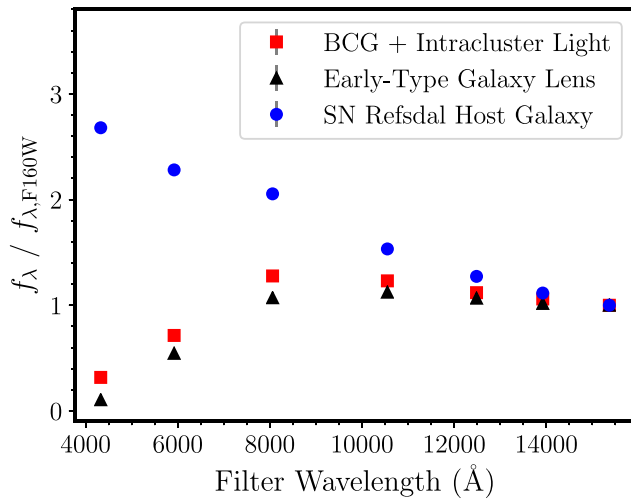
Figure 5. Inferred SEDs of three separate underlying components at the positions of the images of SN Refsdal. Measured flux densities are shown relative to the F160W band. The BCG+ICL are shown as red squares, the early-type galaxy that lenses the SN images S1–S4 are black triangles, and the background flux density that accounts for the SN Refsdal host galaxy are blue circles. All flux densities were measured using GALFIT (Peng et al. 2002) modeling, treating the BCG+ICL as a single Sérsic profile and the elliptical galaxy lens as a separate Sérsic profile. The light from the SN Refsdal host galaxy (the background spiral) is accounted for as a background flux level above the uniform sky brightness. Photometry is corrected for foreground Milky Way extinction using Schlafly & Finkbeiner (2011). Measurement uncertainties are shown but are smaller than the size of the points. See the main text for further details of the fitting procedure.

photosphere at the center of the pixel grid. Figure 6 plots example source-plane magnification maps at the positions of S1–SX. The fact that we use a spherically symmetric explosion model may present a limitation of the simulation. A more sophisticated treatment would use a two-dimensional or even three-dimensional explosion model (e.g., Goldstein et al. 2018; Bonvin et al. 2019). To approximate an asymmetric explosion, we perturb the circular isophotes of the symmetric model to become elliptical with a ratio of the major axes of $b/a = 0.8$. The isophotal brightness of each elliptical isophote matches that of the circular isophote whose diameter is equal to the length of the ellipse's major axis. As shown in Table 9, the elliptical photosphere yields almost identical constraints on the SX–S1 time delay.

In Figure 7, we show the simulated impact of microlensing on the F125W–F160W color. This indicates that stellar microlensing of the photosphere of an SN 1987A–like SN affects its apparent F125W and F160W magnitudes equally within ≲0.05 mag. Furthermore, Figure 8 shows that microlensing is not expected to strongly affect the shape of the light curve of SN 1987A–like SNe during the photospheric phase. Microlensing can also impact the overall magnification of each of the images, causing either increases or decreases to the apparent total magnification of individual SN images. Figure 9 shows the distributions of total magnifications inferred from our simulations, including effects from both "macro" lensing and microlensing.

4.4. Millilensing Simulations

Using the ray-tracing technique presented by Gilman et al. (2019), we compute the expected millilensing due to dark matter subhalos for a projected mass density in substructure bracketing the range of expected values based on theoretical

and observational arguments (Gilman et al. 2020). In practice, we explore $10^7 \, M_\odot \, \mathrm{kpc}^{-2}$ in the main lens plane, corresponding to 0.5% substructure mass fraction, as well as $4 \times 10^7 \, M_\odot \, \mathrm{kpc}^{-2}$, or a 2% mass fraction in substructure. We model the line-of-sight structure using a Sheth–Tormen halo mass function (Sheth et al. 2001). In these simulations, halos and subhalos are rendered in a range of 10^6–$10^{10} \, M_\odot$. Masses outside of this range are either massive enough to contain a visible counterpart or too small to matter statistically (Gilman et al. 2020). Table 10 lists the 16% and 84% quantiles of the fractional change in magnification due to 0.5% and 2% substructure fractions, and Figure 10 plots the distributions we measure for the case where the substructure fraction is 0.5%. These show that magnification due to millilensing ranges from a ~4% to ~18% effect among the five SN images.

4.5. Modeling Correlation Structure in Light Curves

The light curve of any SN will not, of course, be described exactly as a piecewise polynomial. Baklanov et al. (2021) modeled the light curves of SN Refsdal measured through 2015 presented by Rodney et al. (2016) and Kelly et al. (2016c) using a radiation hydrodynamics code. We first fit the Baklanov et al. (2021) F125W and F160W model using the same piecewise polynomial model that we apply to the data. We next model these residuals as a realization of a Gaussian process (GP; see Rasmussen & Williams 2006) with a squared exponential covariance function and estimate correlated deviations from the piecewise polynomial with a timescale of 10 days. We then model residuals of the piecewise model from the actual data as a combination of a correlated component common to all images and an uncorrelated ("white-noise") component whose variance differs between images. The correlated component is modeled as a realization of a GP with a squared exponential kernel and fixed timescale of 10 days. For each filter, we estimate the standard deviations of the correlated and uncorrelated perturbations, which are listed in Table 11. We use our GP model to generate realizations of perturbations to the best-fitting piecewise polynomial model that are informed by both the Baklanov et al. (2021) model and the residual variation seen in the real data.

4.6. Simulated SN Refsdal Light Curves

We now have all the components we believe are needed to assemble simulated light curves that mimic the SN Refsdal data. These simulations will then be used to evaluate our light-curve fitting algorithms (Section 5) and define weights when combining the inferences from all models together (Section 5.4).

Figure 11 shows the steps to produce a single simulated data set, which was chosen at random from the 1000 primary simulated light curves. We use the fiducial light-curve model (a piecewise polynomial, described in Section 3.6) to generate 1000 simulated light curves with the same cadence and photometric uncertainties as for SN Refsdal, as well as similar apparent magnitudes to those of SN Refsdal. We apply additional perturbations from our GP model of the expected variation around the fiducial model. We next apply our simulations of the microlensing magnifications derived from the inferred stellar mass densities, shear, and convergence along the lines of sight to each image (as described in Section 4.3). Since the uncertainties in our measurements are background dominated, we do not change photometric
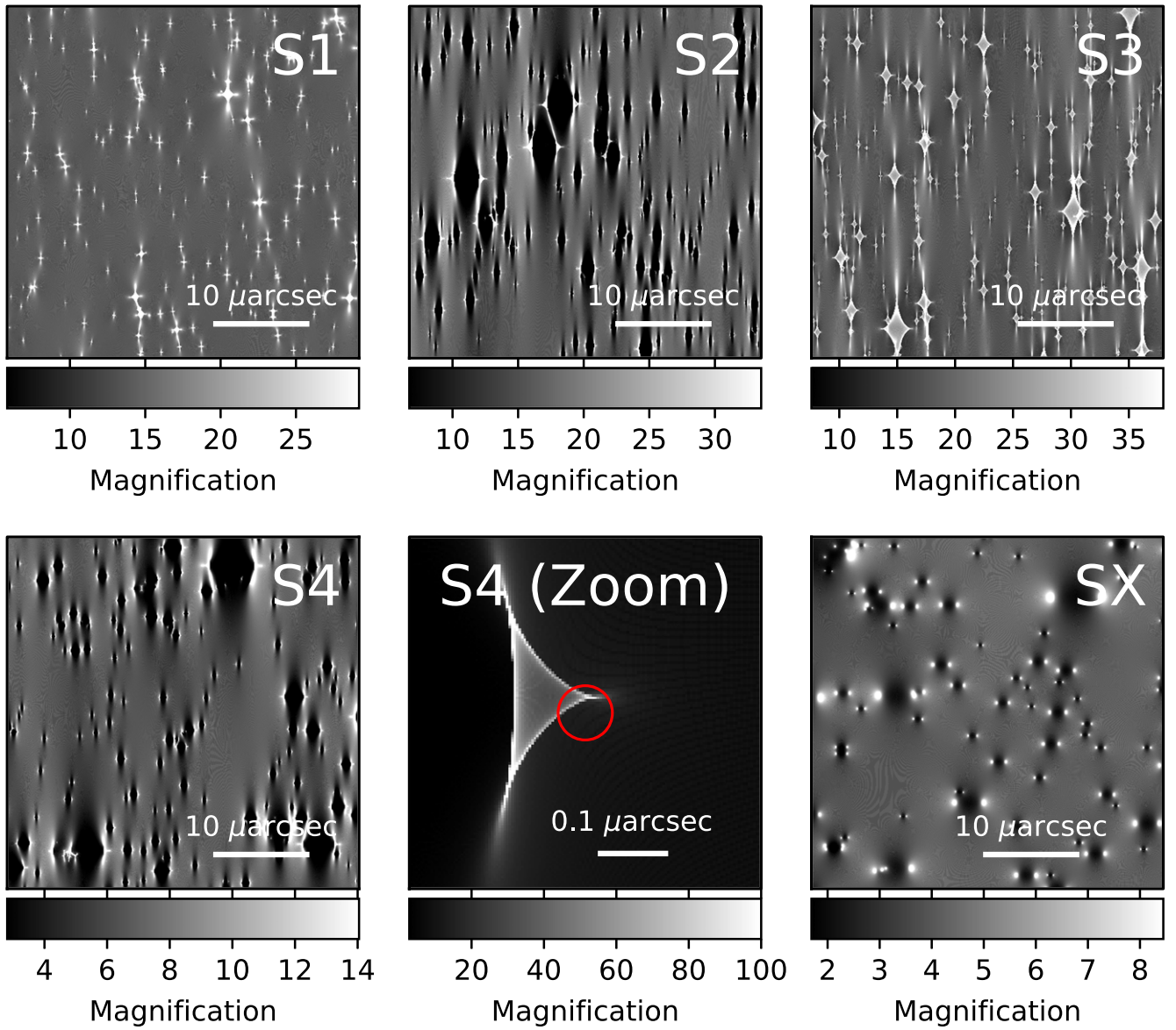
Figure 6. Simulated source-plane magnification maps for random realizations of stellar microlenses for SN Refsdal. Here we show simulations we have generated using the κ and γ values listed in Table 6 from the GLAFIC v3 model (Oguri 2010; Kawamata et al. 2016). The red circle shows the size of a photosphere after 10 days of expansion with a velocity of 6000 km s^{-1}. All panels (except for the zoomed-in panel) are 36 mas (0.31 pc) on a side. The Einstein radius of a solar-mass star at the lens's redshift is 0.015 pc.

Table 8
Stellar and Total Mass Density along Multiple Lines of Sight to SN Refsdal

SN Image	$\Sigma_*/(10^7\,M_\odot\,\mathrm{kpc}^{-2})$			$\Sigma_{\mathrm{tot}}/(10^7\,M_\odot\,\mathrm{kpc}^{-2})$	$\kappa_\star/\kappa_{\mathrm{tot}}$
	BCG+ICL	Early-type Gal.	Total		
S1	0.434 ± 0.005	0.724 ± 0.008	1.158 ± 0.009	185.7 ± 9.9	0.0062 ± 0.0003
S2	0.403 ± 0.004	0.484 ± 0.005	0.887 ± 0.007	188.3 ± 11.4	0.0047 ± 0.0003
S3	0.297 ± 0.003	0.419 ± 0.005	0.716 ± 0.006	176.7 ± 12.5	0.0041 ± 0.0003
S4	0.273 ± 0.003	1.520 ± 0.017	1.793 ± 0.017	190.4 ± 11.4	0.0094 ± 0.0006
SX	1.392 ± 0.022	0.000 ± 0.000	1.392 ± 0.022	231.5 ± 13.4	0.0060 ± 0.0004

uncertainties when rescaling by microlensing fluctuations. We include the effects of millilensing (Section 4.4) until after fitting each microlensed light curve, since its effect is only a constant change in magnification.

In Figure 12, for the F125W and F160W light curves of SX, we compare the piecewise polynomial model that best fits the simulated light curves in Figure 11, against the same model shifted to reproduce the actual relative time delay. In the case of this simulated light curve, the best-fitting time delay is within 3 days of the true delay. Indeed, the estimated time delay can only be substantially affected by microlensing if the apparent *shape* of the light curve is altered, which requires a significant change in

Table 9
SX–S1 Time Delays for Alternative Assumptions

Method	Time Delay (Δt; days)
κ and γ from Grillo-g (see Table 7)	376.62 (365.18, 370.72, 376.14, 381.67, 389.91)
κ and γ from Sharon-g (see Table 7)	376.42 (364.71, 370.97, 376.55, 383.12, 393.22)
Elliptical Photosphere ($b/a = 0.8$) and κ and γ from Oguri-a	375.82 (363.57, 370.39, 376.19, 382.29, 392.17)
Double $\kappa_\star$ and κ and γ from Oguri-a	376.22 (363.17, 370.17, 376.22, 382.03, 391.19)

Note. Constraints on the SX–S1 time delay using microlensing simulations carried out using alternative assumptions. These include convergence and shear values for the Grillo and Sharon models, the Oguri shear and convergence with an elliptical photosphere having a major axis ratio of $b/a = 0.8$, and an assumption of twice the inferred mass density in the intracluster medium. Comparing the inferences to each other and with Table 15 shows that the SX–S1 time delay constraints are not sensitive to these alternate assumptions. Each measurement is reported as the measured (maximum likelihood) value, followed in parentheses by the 2.27th, 16th, 50th, 84th, and 97.73th percentiles of the probability distribution.
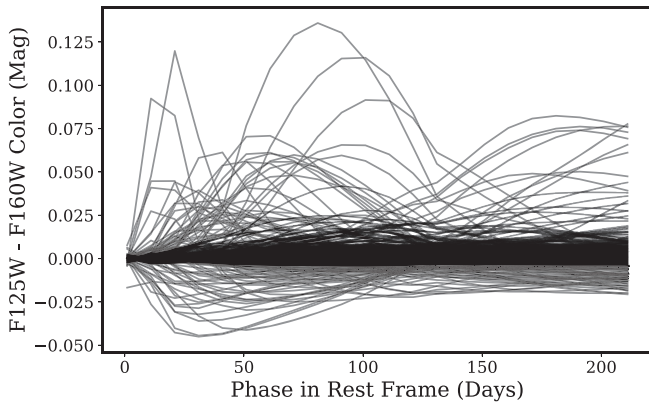
Figure 7. Effect of microlensing on the apparent F125W–F160W color ($-2.5 \times \log_{10}(\text{F125W}/\text{F160W})$) of SN Refsdal from simulations, as a function of rest-frame time since explosion. Shown are 1000 random realizations of an SN 1987A–like photosphere expanding linearly in a stellar microlensing field. These curves demonstrate that microlensing magnification will affect the total F125W and F160W flux from the photosphere almost equally, resulting in color changes of <0.05 mag.
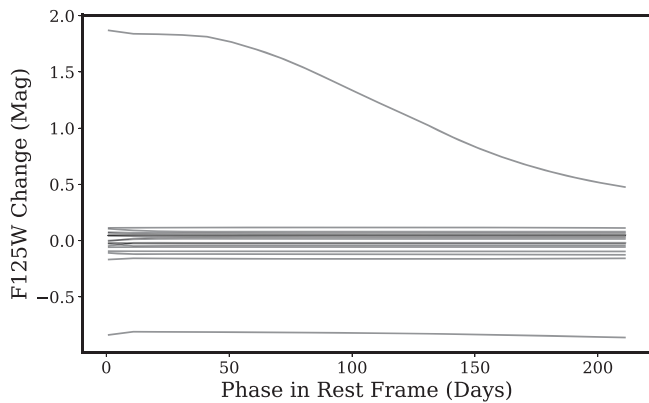
Figure 8. Simulated magnification of SN photosphere in the F125W bandpass due to microlensing as a function of SN phase. Here we show light curves corresponding to different random configurations of masses and positions of foreground microlenses. These simulations demonstrate that, in almost all cases, expected microlensing will not appreciably affect the shape of the light curve.

magnification with time. Given the low values of $\kappa_\star$ for the images of SN Refsdal, the optical depth for microlensing is small, so these events should be comparatively rare.

Four separate teams attempted to recover the relative time delays and magnification ratios from the 1000 sets of simulated light curves, using the methods described in Section 5. The fitting was fully blind, meaning that the true values of the simulated time delays and magnification ratios were hidden from the light-curve fitting teams. Section 5.4 presents the analysis of these simulation tests and their application when we combine their results.

5. Time Delay and Magnification Measurement Methods

In this section, we describe the four algorithms that we have applied to recover the relative time delays and magnifications of images S1–SX. We then explain how we used the simulated light curves to derive probability distributions for each of the observable lensing properties of SN Refsdal, and then we use those probability distributions to combine the estimates from all models.

5.1. Light-curve Fitting Algorithms

5.1.1. Piecewise Photospheric and Nebular Polynomial Model

Our first light-curve fitting algorithm uses the same piecewise polynomial apparatus that we used for generating the simulated light curves, as shown in Figure 3 and described in Section 4.6. For the reasons described in Section 3.6, we adopt a third-order polynomial for the photospheric phase and a second-order polynomial for the nebular phase. We include separate models for the F125W and F160W light curves (i.e., assuming no fixed rule for color variation) but simultaneously fit the light curves for all images S1–SX. Table 5 presents the best-fitting parameters of this model for various polynomial orders, as well as the AIC (Akaike 1974) values and the inferred time delays from image S1 to SX.

5.1.2. Bayesian Gaussian Process Model

We construct a Bayesian GP model for the light curves, motivated by the approaches of Tak et al. (2017) and Hojjati et al. (2013) for modeling lensed quasar time delays. Our framework describes the intrinsic SN light-curve shape, perturbations due to microlensing, and relative time delays and magnification due to strong lensing, within a single probabilistic model. In this model, the time series data of each image are treated as noisy measurements of the same underlying light curve, shifted in time and magnitude and modulated by a microlensing curve. The latent light curve, assumed to be the same for all images, is modeled as a realization of a GP for smooth functions (see Rasmussen & Williams 2006, for a background on GPs). The microlensing perturbations are also
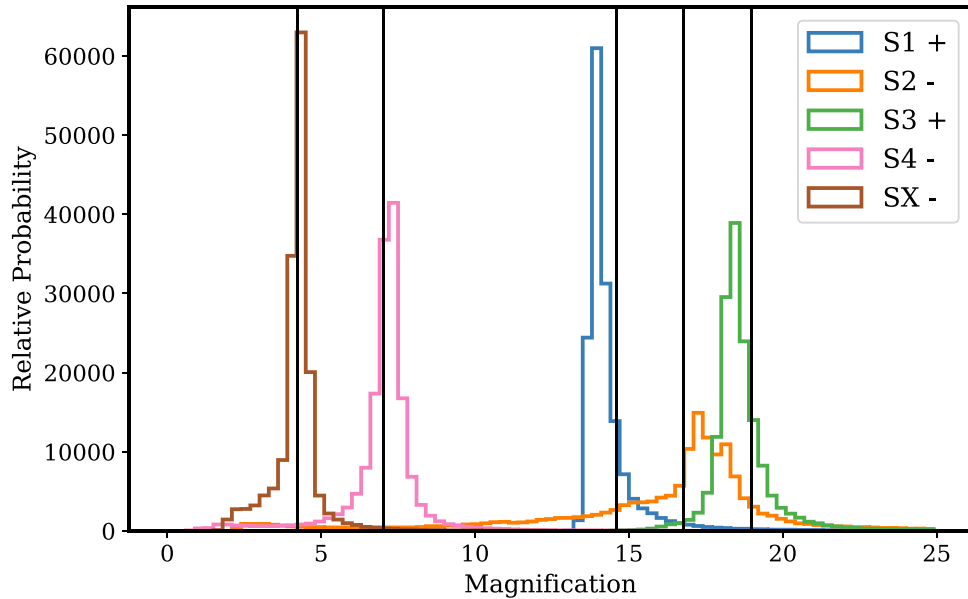
Figure 9. Distributions of total magnification of images S1–S4 and SX of SN Refsdal from microlensing simulations. The shape of the distributions depends on the parity of the SN image, the shear γ and convergence κ of the local lensing potential, and the mass density of the stellar microlenses listed in Tables 8 and 6. The respective vertical lines show the magnification for the parameters in Table 6 that we use to carry out our simulation. The negative-parity images (S2, S4, and SX) have significant tails extending to low magnification. We have convolved the simulated magnification maps with a circular kernel with a radius of 0.0017 pc equal to that expected for a photosphere with a velocity of 6000 km s^{-1} that has expanded for 100 days in its rest frame.

Table 10
Millilensing Simulations

cImage	0.05% of Dark Matter		2% of Dark Matter	
	16%	84%	16%	84%
S1	−0.12	0.09	−0.14	0.09
S2	−0.17	0.12	−0.17	0.14
S3	−0.18	-0.14	−0.08	0.08
S4	−0.07	0.07	−0.08	0.08
SX	−0.04	0.03	−0.04	0.04

Note. Effect of millilensing by populations of dark matter subhalos on magnification of images S1–SX. Here we list the 16% and 84% quantiles of $(\mu - \mu_{macro})/\mu$, where μ is the magnification measured from each simulation and μ_{macro} is the magnification that the SN images would have owing to the galaxy cluster and cluster members in the absence of millilensing or microlensing. The columns at left show the results of simulations where 0.5% of the dark matter along the line of sight consists of substructure, and the columns at right show results where substructure accounts for 2% of the dark matter along the line of sight.

Table 11
Parameters of Residual Model

Filter	Correlated	Uncorrelated				
	K	S1	S2	S3	S4	SX
F125W	0.00670	0.00576	0.01463	0.02097	0.03788	0
F160W	0.01066	0.00998	0.00416	0	0.02492	0

Note. Standard deviations of correlated and uncorrelated components of the model used to generate residuals from the underlying piecewise polynomial.

modeled as GP realizations, but with no correlation between the multiple images. Our overall model has 12 parameters—two latent light-curve hyperparameters, two microlensing hyperparameters, four time delays, and four magnitude shifts. We sample the joint posterior distribution of these parameters, conditioning simultaneously on the light-curve observations of all the images, using the dynesty nested sampling package (Speagle 2020) implementation of nested sampling with ellipsoidal bounds (Skilling 2004, 2006; Feroz et al. 2009). We obtain posterior distributions of the time delays by marginalizing over all other parameters.

5.1.3. *Supernova Time Delays Parametric Model*

We fit the light-curve data using version 2 of the Supernova Time Delays (SNTD) software package (Pierel &

Rodney 2019). Here we model SN Refsdal's flux as a function of time using

$$F(t) = A \frac{e^{-(t-t_0)/\tau_{\rm fall}}}{1 + e^{-(t-t_0)/\tau_{\rm rise}}} + B, \tag{1}$$

presented by Bazin et al. (2009) as a generic model for SN light curves, where $\tau_{\rm rise}$ and $\tau_{\rm fall}$ characterize the relative rise and decline times of the light curve, respectively, and A and B set its normalization. While this model is flexible, its simple shape minimizes sensitivity to microlensing features in the light curves of S1–SX.

In order to measure the time delays of the simulated light curves described in Section 4.6, we employed SNTD's parallel method according to which the light curves of S1–SX are first fit separately before deriving joint likelihood estimates for the light-curve model parameters (see Pierel & Rodney 2019). We note that SNTD was designed to use spectrophotometric models of SNe, to leverage, when available, prior knowledge of the intrinsic color and shape of well-studied SN species including Type Ia SNe or common types of core-collapse SNe. SN Refsdal's light curve, however, was peculiar and similar to that of SN 1987A (Kelly et al. 2016c), the explosion of a blue
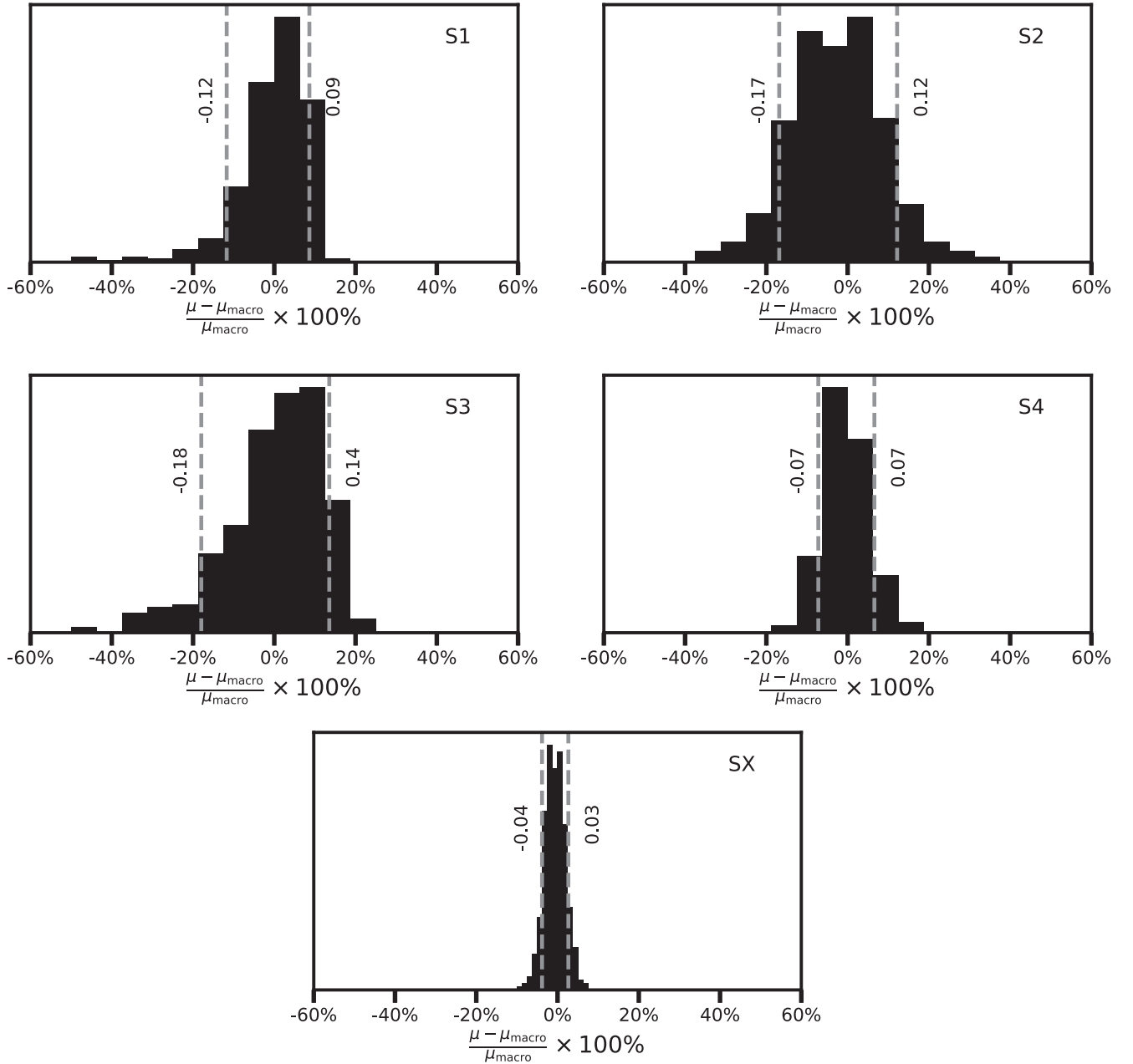
Figure 10. Simulation of effect of millilensing dark matter subhalos on the magnification of images S1–SX. Here we show results where 0.5% of the cluster dark matter is in the form of substructure, but Table 10 shows results where 2% of the cluster dark matter consists of subhalos.

supergiant, and a useful spectrophotometric model was unfortunately not available.

We applied the parallel algorithm to the F125W and F160W light curves separately, since Equation (1) only describes the flux in a single band. For each filter, we did two iterations. We first identified the time τ_{pk} as the epoch with the maximum observed flux and carried out a "coarse" fit subject to the following bounds on the fitting parameters:

$$\tau_{pk} - 150 < t_0 < \tau_{pk} + 150$$
$$20 < t_{rise} < 80$$
$$60 < t_{fall} < 300$$
$$0.001 < A < 4$$
$$0 < B < 2.$$

From this coarse run we identified the set of parameters that returned the best χ^2 as $(t_{0,c}, t_{rise,c}, t_{fall,c}, A_c, B_c)$. We adopted the first three of those parameter values as estimates for the time of peak, rise time, and fall time, respectively.

Our next step was to carry out a "fine" fit using the same bounds for A and B as in the coarse run while applying the following constraints:

$$t_{0,c} - 50 < t_0 < t_{0,c} + 50$$
$$t_{rise,c} - 3 < t_{rise} < t_{rise,c} + 3$$
$$t_{fall,c} - 10 < t_{fall} < t_{fall,c} + 10.$$

After completing this procedure for the F125W and F160W bands separately, we defined the final time delay and
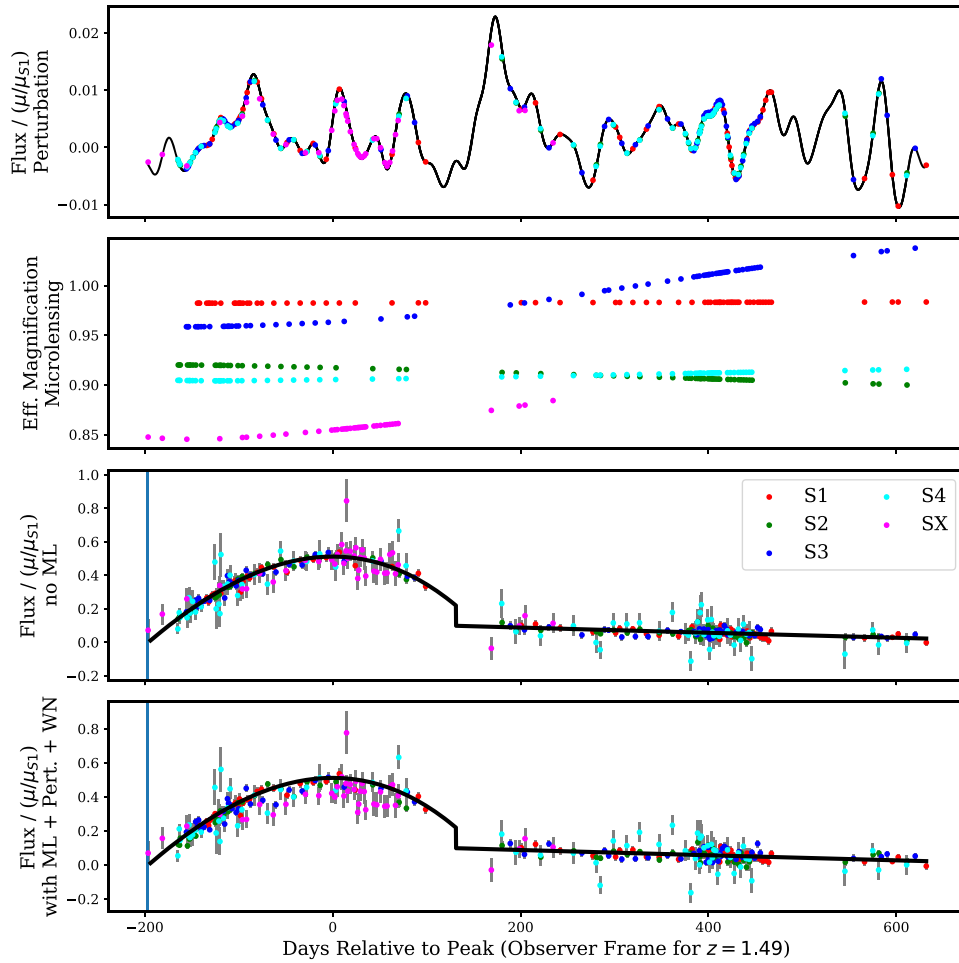
Figure 11. Components of a single set of simulated light curves of the five appearances of SN Refsdal. The top panel shows the perturbation applied to the fiducial piecewise polynomial model. The perturbation is a realization of a GP model of the residuals of the Baklanov et al. (2021) light curves from the best-fitting piecewise polynomial model. The next panel plots the change in apparent magnification in the F125W band due to microlensing by a randomly placed set of stars and stellar remnants. The middle panel plots the generative light-curve model without microlensing (ML) included. Motivated by fits to the light curves of low-redshift SN 1987A–like SNe with light curves similar to that of SN Refsdal (see Figure 3), the generative model consists of separate polynomials for the photospheric and nebular SN phases. In the bottom panel, we show the simulated light curves after adding the microlensing plotted in the top panel. In both the middle and bottom panels, we divide the observed flux by μ/μ_{S1} to remove the relative magnification ratios among the light curves due to smoothly distributed matter (i.e., in the absence of microlensing), and we remove their relative offsets.
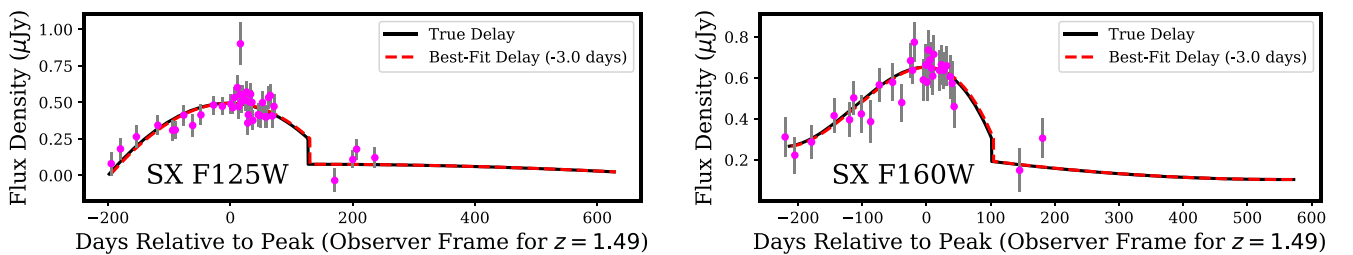
Figure 12. Comparison between recovered and actual delay of SX relative to S1 for the example simulation shown in Figure 11. We show the best-fitting piecewise polynomial model (red dashed) and the same fit after shifting the model according to the true delay used to create the simulated light curve, leaving its normalization unchanged. For this realization, the time delay is recovered within 3 days. Estimated time delays are only affected by time-varying changes in total magnification arising from microlensing that modify the shape of the light curve (e.g., Dobler & Keeton 2006). As a consequence, the inferred time delay requires strong microlensing events, which should not occur frequently, given the relatively small value of κ_s.

magnification estimates from the band preferred by the Bayesian evidence for the fine-run fitting results.

During the testing phase with simulated light curves (i.e., before applying any fitting method to the actual Refsdal data), we experimented with more aggressive constraints on

the parameter values. Specifically, we experimented with tighter bounds on the model parameters when fitting to the light curves of S2, S3, and S4 and then using these parameters to model the light curves of S1 and SX. However, this procedure resulted in a ∼7-day bias in the
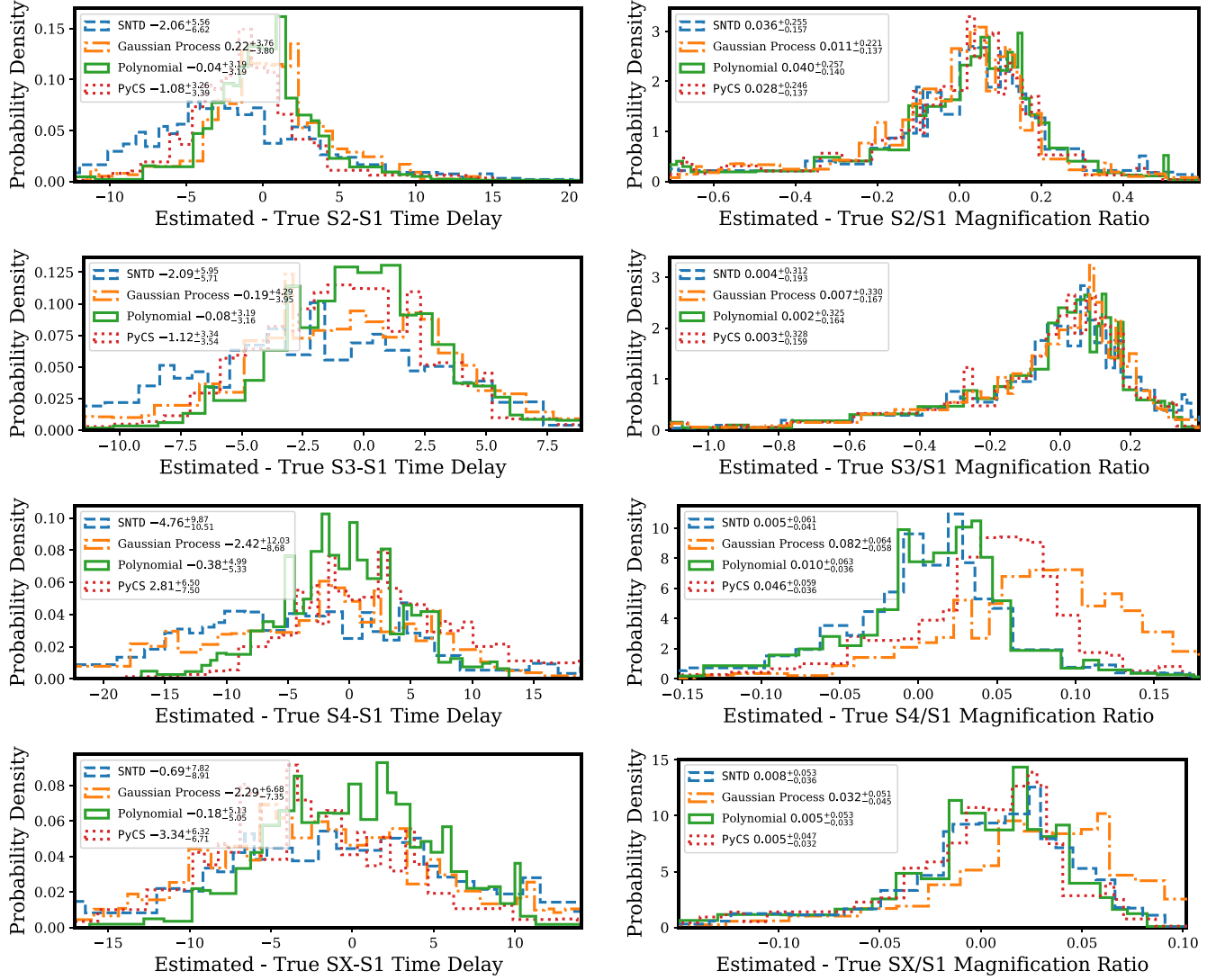
Figure 13. Distributions of residuals of time delay (left) and magnification ratio (right) estimated from simulated light curves by four separate algorithms. As indicated, each row shows results from the simulated data that mimic one pair of images, from S2:S1 in the top tow to SX:S1 in the bottom row. The 16th, 50th, and 84th percentiles are listed in the legends.

Table 12
RMS SCATTER

Method	Time Delay				Magnif. Ratio			
	S2–S1	S3–S1	S4–S1	SX–S1	S2/S1	S3/S1	S4/S1	SX/S1
Polynomial	4.6	4.3	5.8	5.5	0.29	0.33	0.08	0.08
Gaussian process	5.2	5.0	10.3	7.7	0.24	0.32	0.11	0.08
SNTD	7.1	6.6	11.1	8.6	0.30	0.35	0.08	0.08
PyCS	4.9	4.6	8.1	7.6	0.27	0.33	0.08	0.08

Note. The rms scatters in the time delay columns are given in units of days. The magnification ratio columns are unitless and report the scatter of flux magnification ratios (i.e., no conversion into magnitudes has been done). When computing the scatter, we have excluded > 30-day outliers for the time delays, and 1.5 for the magnification.

estimated time of peak for the S1 and SX light curves, and we discarded this approach.

5.1.4. Python Curve Shifting3 Free-knot Spline Model

Python Curve Shifting3 (PyCS3) is a new version of the PyCS package, which was originally designed for

application to lensed quasar light curves, and includes several independent methods for measuring time delays and estimating uncertainties (Tewes et al. 2013; Millon et al. 2020b). For application to the SN Refsdal data, we use the free-knot spline estimator, which fits third-order splines to the light-curve data. The intrinsic variations of the source, common to all lensed images, are fit with a single spline. The

Method	Time Delay				Magnif. Ratio			
	S2–S1	S3–S1	S4–S1	SX–S1	S2/S1	S3/S1	S4/S1	SX/S1
Polynomial	0.038	0.019	0.019	0.009	0.028	0.034	0.019	0.016
Gaussian process	0.031	0.015	0.026	0.006	0.019	0.035	0.005	0.014
SNTD	0.029	0.008	0.019	0.009	0.022	0.034	0.019	0.016
PyCS	0.048	0.019	0.030	0.040	0.023	0.035	0.010	0.016

Note. Outliers are defined as time delay measurements that differ from the mean by more than 30 days, or measured magnification ratios that deviate by more than 1.

variability of the spline is controlled by the initial spacing of the knots. The position of the knots, coefficients of the spline, and time and magnitude shifts of all the light curves are iteratively adjusted to minimize a global χ^2, following the bounded optimal knots algorithm (Molinari et al. 2004). The details of the fitting process are given by Tewes et al. (2013), and assessments of the robustness of the method when applied to lensed quasars are presented by Liao et al. (2015) and Bonvin et al. (2016).

We explore a set of plausible knot spacing values with the automated pipeline described by Millon et al. (2020a) and select the value providing the most precise results, according to the PyCS3 uncertainty estimation procedure. Note that we do not use this procedure for our final estimates of the uncertainties. Instead, we use the free-knot spline fitting to estimate a single value for the relative time delays, and we compute the uncertainty using the residuals of the PyCS3 estimates, compared to the simulated light curves as described in Section 5.4. Finally, we obtain our final time delay estimates by measuring the time delays and flux ratios separately in the F125W and F160W band and by taking the weighed mean, with the weight being proportional to the inverse of the estimated uncertainties in each band.

5.2. Probability Distributions for Each Observable, from Each Method

Each of the four methods described in Section 5, when given the light curves of images S1–SX, can yield a single estimate for eight lensing observables $\mathcal{O}$. These observables consist of the four time delays between images S2, S3, S4, and SX relative to S1 ($\Delta t_{2,1},...,\Delta t_{x,1}$) and the four respective magnification ratios ($\mu_2/\mu_1,...,\mu_X/\mu_1$). The choice of image S1 as the reference image is arbitrary. In order to combine the estimates from multiple methods with appropriate weighting for each method, we need a consistent way to define a probability distribution $p(\mathcal{O}|D; M_k)$ for each method M_k given the light-curve data D. We address this by performing fits to the 1000 simulated light curves, which were described in Section 4.6.

For the ith simulated light curve, we calculate the residual of each *input* value $\mathcal{O}_{\text{input};i}$ from the *estimated* value $\mathcal{O}_{\text{est};i}$,

$$\delta\mathcal{O}_{\text{resid};i} = \mathcal{O}_{\text{est};i} - \mathcal{O}_{\text{input};i}. \qquad (2)$$

We use a histogram of these residuals to construct a probability distribution P_{resid}.

We next compare the ability of the four fitting methods to recover the input relative time delays. As shown in Figures 13, the polynomial method has the least bias, and the 16th, 50th, and 84th percentiles of its residual distribution are most closely separated. The rms scatter of the polynomial method, as listed

in Table 12, is also smallest. In some cases, the performance of the GP and PyCS fitting methods approaches that of the polynomial method. The light curve of S4 has a lower S/N than those of the other images, however, and in that case, both methods are not as effective in recovering its relative time delay as the polynomial method.

Considering the codes' ability to recover magnification ratios, Figure 13 and Table 12 demonstrate that all fitting codes yield comparably accurate and precise magnification ratios S2/S1 and S3/S1. In the case of the S4/S1 ratio, the GP and PyCS methods show a greater bias than the polynomial and SNTD methods. In Table 13, we report the fraction of simulated light curves for which the difference between measurement and truth ($\delta\mathcal{O}_{\text{resid};i}$) is more than 30 days (for time delays) or >1 (for magnification ratios).

The IMF of the stellar populations in the early-type galaxy lens (e.g., Treu 2010; Conroy & van Dokkum 2012) and the intracluster medium may have a Salpeter (1955) IMF instead of the Chabrier (2003) IMF that we have assumed for our simulations. In Tables A1 and A2, we report the same set of statistics but instead using microlensing simulations computed for a Salpeter (1955) IMF. In Figure A6, we plot the residual distributions given a Salpeter (1955) IMF. The residual distributions for simulations that adopt a Salpeter IMF show a very modest increase ($\lesssim 10\%$) in the difference between the 16th and 84th percentiles in the case of the SX–S1 time delay and larger ($\sim 30\%$) increases for the S2–S1, S3–S1, and S4–S1 images.

We next want to compute the probability distribution for a property of the light curve $\mathcal{O}$, given a single-point estimate of the same observable $\mathcal{O}_{\text{est};D;k}$, derived from the actual light-curve data D using method M_k, to compute this probability distribution. To accomplish this, we use the distribution of residuals $P(-\delta\mathcal{O}_{\text{resid}})$ assembled from fits to the simulations. This probability distribution $P(\mathcal{O}|D; M_k)$ is given by

$$P(\mathcal{O}|D; M_k) = P(\mathcal{O}|\mathcal{O}_{\text{est};D;k}) = P_{\text{resid}}(\mathcal{O} - \mathcal{O}_{\text{est};D;k}). \qquad (3)$$

5.3. Results from Bootstrapping Samples

We computed the relative time delays and magnifications of the SN light curves after creating 1000 light curves after performing bootstrapping with replacement. Since the uncertainties we measured were smaller than those we computed from the simulated light curves, which included perturbations from the piecewise polynomial model, microlensing, and millilensing, we adopt the results from the simulations.
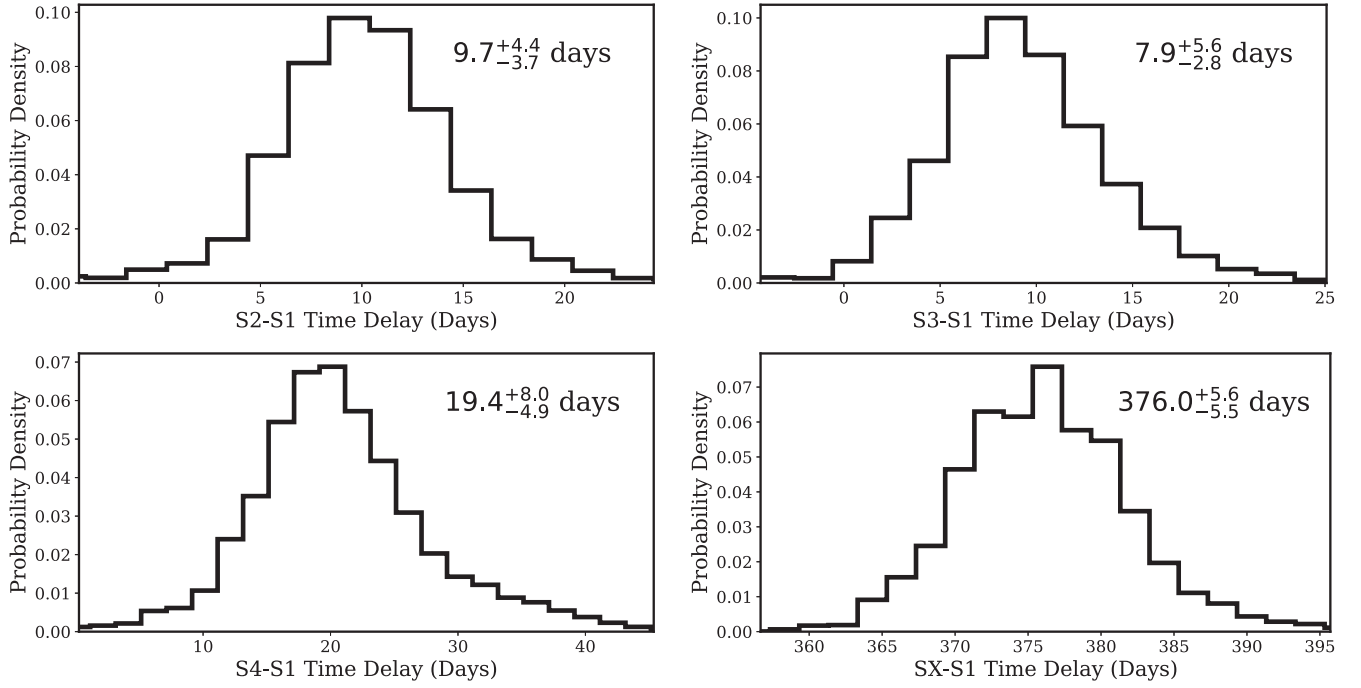
Figure 14. Relative time delays between images S2–SX and S1. Constraints shown are the weighted combination of results from four different algorithms where weights are determined by maximizing the likelihood of the actual time delays in simulated light curves. The values in the upper right corner are the 16th percentile confidence level, maximum likelihood value, and 84th percentile confidence level, as listed in Table 15.

Table 14
Weights of Methods to Maximize Likelihood of True Parameters

Method	Time Delay				Magnif. Ratio			
	S2–S1	S3–S1	S4–S1	SX–S1	S2/S1	S3/S1	S4/S1	SX/S1
Polynomial	0.540	0.598	0.840	0.885	0.128	0.373	0.583	0.252
Gaussian process	0.217	0.173	0.015	0.038	0.872	0.535	0.054	0.014
SNTD	0.005	0.027	0.023	0.053	0.000	0.092	0.308	0.000
PyCS	0.238	0.203	0.122	0.024	0.000	0.000	0.055	0.734

Note. Weight assigned to each method determined by maximizing the likelihood of each input time delay or magnification.

5.4. Combination of Estimates from Distinct Light-curve Algorithms

For each observable, we sought to identify an optimal combination of the set of four probability distributions $P(\mathcal{O}|D; M_k)$ corresponding to the respective methods. We find the nonnegative linear combination of probability distributions $P(\mathcal{O}|D; M_k)$ that *maximizes the likelihood* of the input observable values used to construct the simulations,

$$\max_{w \in \mathcal{S}_1^K} \frac{1}{n}\sum_{i=1}^{n} \log \sum_{k=1}^{K} w_k P(\mathcal{O}_{\text{input};i}|D_i; M_k), \qquad (4)$$

subject to $\mathcal{S}_1^K = \{w \in [0,1]^K : \sum_{k=1}^{K} w_k = 1\}$. When computing these weights using fits to the simulated light curves, we use a modified probability function $P_{\text{penal}}(\mathcal{O}|D; M_k)$ that has the effect of penalizing bias. The penalized probability distribution is given by

$$P_{\text{penal}}(x) = P_{\text{resid}}(x - \text{median}(\Delta\mathcal{O}_{\text{resid}})). \qquad (5)$$

For a distribution where the median residual, $\text{median}(\delta\mathcal{O}_{\text{resid}})$, is equal to zero, $P_{\text{penal}} = P_{\text{resid}}$. Given best-fitting weights w_k,

the probability distribution for the actual light curve becomes

$$P(\mathcal{O}|D) = \sum_{k=1}^{K} w_k P(\mathcal{O}|D; M_k). \qquad (6)$$

We note that maximizing the likelihood of the relative time delays and magnifications used to construct the simulated light curve corresponds to minimizing the Kullback–Leibler divergence between a single-valued distribution where the residual $\delta\mathcal{O}_{\text{resid};i} = 0$ and the stacked set of probability distributions $p(\mathcal{O}|D)$. The use of the Kullback–Leibler divergence as a means to identify an optimal combination instead of *predictive* models was developed by Yao et al. (2018).

The Kullback–Leibler divergence allows us to consider full probability distributions, as opposed to single-point estimates. For this analysis, we use the residual distribution for all simulated light curves to construct the probability for every light curve as described in Equation (2); the framework we describe could accommodate a separate probability distribution for each simulated light curve.
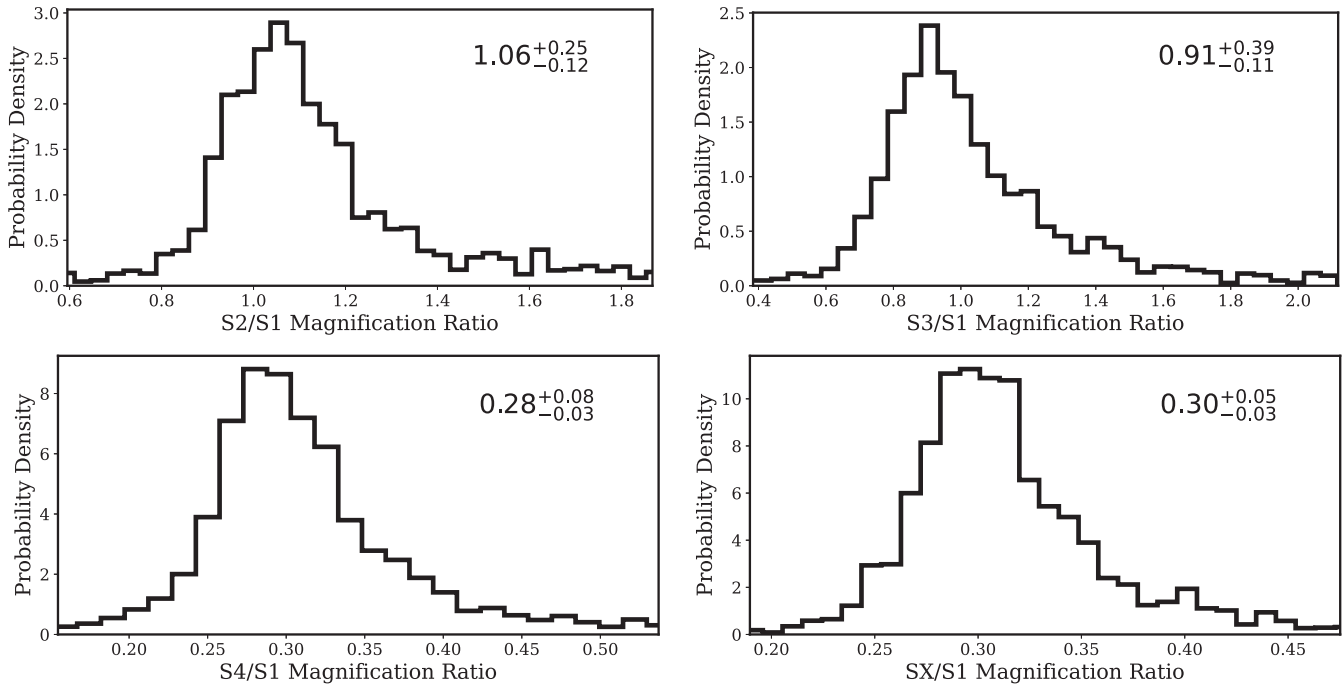
Figure 15. Magnification ratios between images S2–SX and S1. Constraints shown are a weighted combination of results from four different algorithms, where weights are determined by maximizing the likelihood of the actual magnification ratios in simulated light curves. The values in the upper right corner are the 16th percentile confidence level, maximum likelihood value, and 84th percentile confidence level, as listed in Table 16.

Table 15
Measured Values of Time Delays

Method	Time Delay (Δt; days)	
	S2–S1	S3–S1
Polyn.	8.49 (−6.67, 5.49, 8.57, 11.80, 17.75)	7.52 (−0.07, 4.12, 7.46, 10.72, 17.02)
GPR	11.09 (−0.80, 7.95, 11.61, 15.83, 23.48)	11.13 (2.24, 8.04, 11.74, 16.25, 23.06)
SNTD	4.28 (−13.97, −3.10, 3.23, 8.74, 17.61)	12.93 (−0.85, 7.23, 13.52, 19.58, 28.07)
PyCS	12.89 (−9.96, 9.60, 12.80, 15.93, 20.95)	11.73 (2.99, 8.15, 11.65, 14.77, 19.83)
Combined	9.69 (−8.97, 5.96, 9.92, 14.06, 20.56)	7.92 (0.47, 5.16, 8.95, 13.56, 20.19)
	S4–S1	SX–S1
Polyn.	19.24 (4.92, 14.19, 19.29, 24.36, 31.75)	376.02 (365.31, 370.63, 375.74, 380.99, 387.37)
GPR	55.48 (38.12, 49.51, 58.52, 70.77, 85.43)	383.83 (367.04, 375.96, 383.62, 390.23, 400.15)
SNTD	16.84 (−15.09, 5.24, 16.27, 27.06, 36.89)	380.83 (360.53, 370.68, 379.58, 386.68, 394.58)
PyCS	31.25 (0.65, 24.37, 31.22, 37.83, 44.79)	382.43 (367.43, 376.45, 382.53, 389.14, 450.11)
Combined	19.44 (5.94, 14.56, 20.28, 27.45, 43.15)	376.02 (365.12, 370.50, 376.03, 381.65, 390.23)

Note. Each measurement is reported as the measured (maximum likelihood) value, followed in parentheses by the 2.27th, 16th, 50th, 84th, and 97.73th percentiles of the probability distribution.

To calculate the likelihood of the input light-curve parameters in Equation (4), we need to choose how to construct a probability density function from the set of residuals $\delta\mathcal{O}_{\mathrm{resid};i}$. The probability density must always be nonzero, with the bin widths sufficiently large to avoid small number statistics. We create bins such that each bin includes three residuals (i.e., each contains three values) in each of the <5% and >95% quantiles, 15 residuals in the 5%–30% and 70%–95% quantiles, and 30 residuals in the 30%–70% quantiles. Moreover, we adjust these bins so that the bin width is at least 0.5 days for the time delays, or 0.01 for the magnification ratios.

For each observable, we carry out an optimization using Equation (4) to identify the weights that maximize the likelihood of the parameters used to construct the light curves.

Table 14 lists the weights that we obtain. The piecewise polynomial method receives the majority of the weight for all of the relative time delay measurements, while multiple codes receive the majority of the weight for the estimates of magnification ratios. Marginal probability distributions from the combination of four methods are shown as a "corner plot" in Figure 17.

In Figure A7, we show the separate constraints from all the methods.

6. Relative Time Delays and Magnification Ratios

With a procedure for optimally combining estimates from the light-curve fitting methods and estimating uncertainty, we turn to measurements of the relative time delays and

Table 16
Measured Values of Magnification Ratios

Method	Magnification Ratio (μ/μ_1)	
	S2–S1	S3–S1
Polyn.	1.021 (0.618, 0.915, 1.048, 1.287, 1.921)	0.907 (0.357, 0.799, 0.969, 1.250, 2.021)
GPR	1.057 (0.767, 0.944, 1.081, 1.296, 1.794)	0.919 (0.506, 0.803, 0.977, 1.285, 2.025)
SNTD	1.021 (0.451, 0.910, 1.056, 1.308, 1.949)	0.889 (0.281, 0.721, 0.933, 1.206, 2.008)
PyCS	1.051 (0.667, 0.952, 1.079, 1.315, 1.885)	0.931 (0.396, 0.827, 0.996, 1.279, 2.044)
Combined	1.06 (0.63, 0.94, 1.08, 1.30, 1.78)	0.91 (0.50, 0.80, 0.97, 1.30, 2.07)
	S4–S1	SX–S1
Polyn.	0.296 (0.199, 0.268, 0.303, 0.363, 0.529)	0.287 (0.211, 0.264, 0.297, 0.350, 0.459)
GPR	0.357 (0.239, 0.298, 0.356, 0.420, 0.525)	0.311 (0.203, 0.264, 0.311, 0.362, 0.455)
SNTD	0.278 (0.176, 0.247, 0.285, 0.345, 0.495)	0.275 (0.198, 0.253, 0.289, 0.343, 0.447)
PyCS	0.333 (0.248, 0.310, 0.344, 0.397, 0.538)	0.305 (0.194, 0.274, 0.309, 0.355, 0.437)
Combined	0.28 (0.18, 0.26, 0.30, 0.37, 0.52)	0.30 (0.22, 0.27, 0.31, 0.35, 0.45)

Note. Each measurement is reported as the measured (maximum likelihood) value, followed in parentheses by the 2.27th, 16th, 50th, 84th, and 97.73th percentiles of the probability distribution.

magnification ratios from the SN Refsdal light-curve data. We used DOLPHOT photometry, restricted to data acquired after 2013 December 24 (MJD = 56,650) to fit the light curves of S1–S4 and 2015 October 10 (MJD = 57,300) for SX. To enable the time delay fitting to be blind, we generated a random number for each light curve and added that random number to the Modified Julian Date corresponding to each photometric measurement. Each of the four fitting methods was applied to the shifted light curves and combined using the weighted sum described in Equation (6).

In Figure 14, we plot the combined constraints on the time delays relative to image S1. The combined constraints on the magnification ratios between these sets of images are shown in Figure 15. In Tables 15 and 16, we list the measured maximum likelihood values and the 2.27th, 16th, 50th, 84th, and 97.73th percentiles inferred from these probability distributions. Marginal probability distributions from the combination of four methods are shown as a "corner plot" in Figure 16.

In Figures 17 and 18, we plot the light curves of SX and those of S1, S2, and S3, the images in the Einstein cross with the greatest magnification, after removing SX–S1 relative time delays of 376 days (most probable value), 360 days (3σ offset from 376 days), and 349 days (5σ offset from 376 days). The plotted light curves of S2 and S3 are shifted by their time delays relative to S1 as listed in Table 15. In Figure 19, we show the χ^2 value of the best-fitting piecewise polynomial model against the SX–S1 relative time delay.

6.1. Comparison of Time Delays from F125W and F160W Light Curves

We next apply the piecewise polynomial modeling code to the F125W and F160W light curves separately in order to evaluate whether the inferred SX–S1 time delays are consistent. To estimate the uncertainties associated with the measurement for each filter, we use the simulated sets of light curves in the respective filters. We find that the F125W light curve yields a constraint on the SX–S1 relative time delay of $377.6^{+5.13}_{-5.04}$ days, while the F160W light curve, whose measurements have, on average, lower S/Ns, yields $370.6^{+9.78}_{-10.07}$ days. The estimates

from the F125W and F160W light curves are in 0.6σ agreement.

6.2. Alternative Assumptions

We explore the effects of alternative assumptions about the lens model, the mass density of stars making up the ICL, and the symmetry of the SN photosphere. In Table 7, we list the convergence and shear at the positions of images S1–SX according to the Grillo-g and Sharon-g models. Table 9 lists the constraints on the SX–S1 time delay when we instead adopt the Grillo-g and Sharon-g lensing parameters. We also obtain inferences on the SX–S1 time delay after constructing an elliptical photosphere with a ratio of major axes of $b/a = 0.8$ and, separately, after doubling the inferred value of $\kappa_\star$, the density of the stars that account for the intracluster medium. Table 9 shows that adopting these alternative assumptions yields almost identical constraints on the relative time delay between images SX and S1.

6.3. Covariance Matrix

We next derive a covariance matrix for Δt and μ/μ_1. The covariance matrix can be used to compute likelihoods for model predictions, although we develop a more precise approach with less sensitivity to outlying values in the accompanying paper. To compute the measurement covariance matrix, we use the 1000 simulations of the light curves of images S1–SX, which include the effects of microlensing and photometric noise.

We can compute the covariance matrix as

$$K_{X_iX_j} = \text{cov}[X_i, X_j] = E[(X_i - E[X_i])(X_j - E[X_j])]. \quad (7)$$

Here we have X and Y correspond to the distributions of residuals of the estimated value subtracted from the actual value for the simulation,

$$\delta\Delta t^{\text{est}}_{i,1} = \Delta t^{\text{est}}_{i,1} - \Delta t^{\text{true}}_{i,1}. \quad (8)$$

As described in Section 5.4, we have identified the weights $w_{j,\text{obs}}$ for each model M_j and observable such that the linear
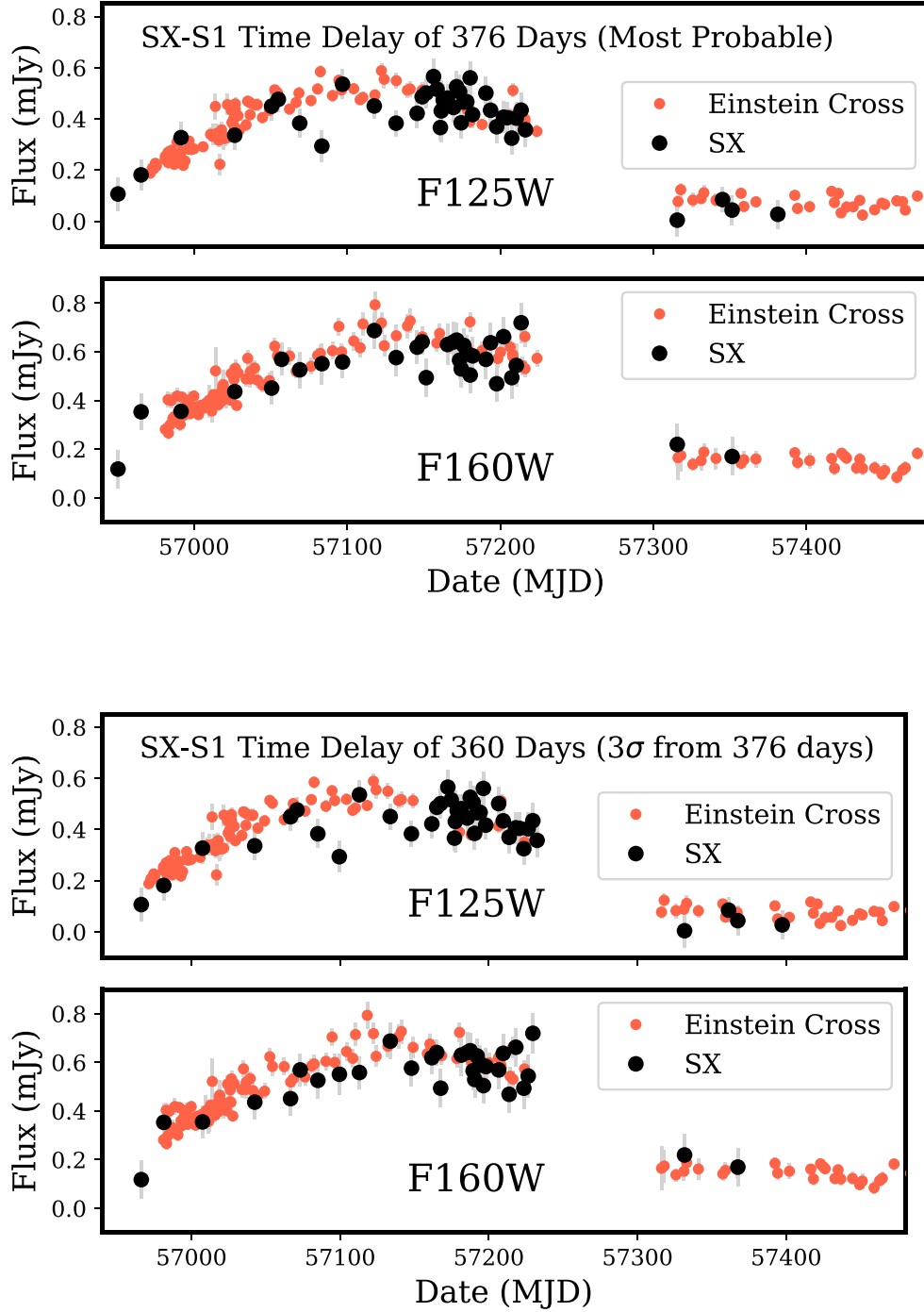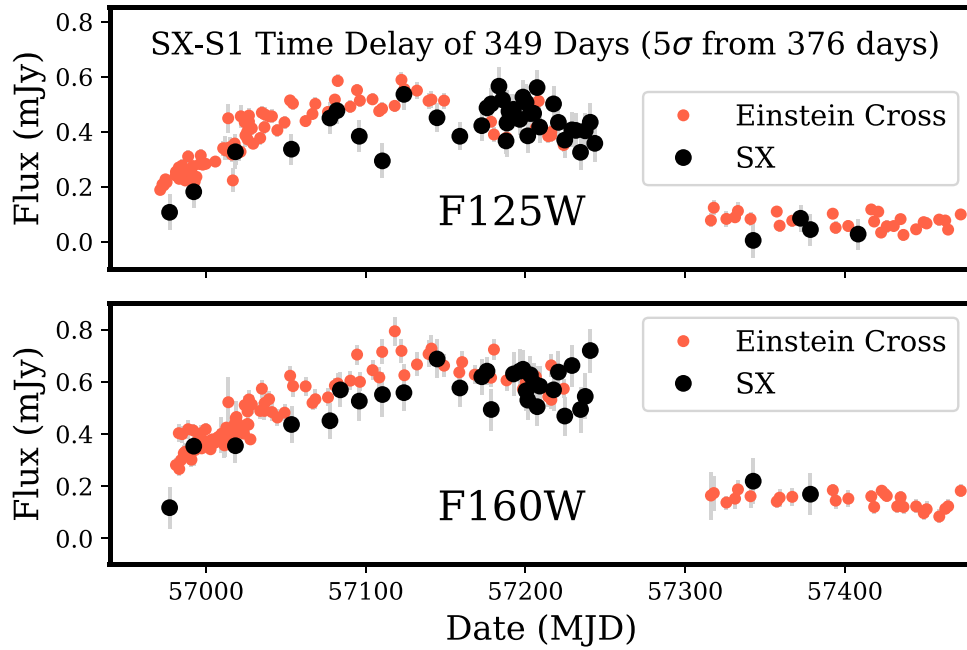
Figure 16. The light curves of the most highly magnified Einstein-cross images S1–S3 and that of image SX, after removing the best-fitting SX–S1 time delay of 376 days (top panel), and, for comparison, an SX–S1 time delay of 360 days (bottom panel). Given our constraints on the SX–S1 time delay of 376.0 + 5.6 − 5.5 days, 360 days is offset by 3σ from the best-fitting value. We have also shifted the plotted light curves of images S2 and S3 to remove their time delays relative to S1, using the values listed in Table 15. We do not include the light curve of S4 in this comparison, given its smaller S/N. Figure 19 plots the χ^2 value of the best-fitting piecewise polynomial model as a function of the SX–S1 relative time delay.

combination of the probability distributions,

$$p(\mathcal{O}|D) = \sum_{j=1}^{K} w_{j,\mathrm{obs}}\, p(\mathrm{observable}|D; M_j), \qquad (9)$$

maximizes the likelihood of the value of the observable used to create the simulated light curve. Consequently, we have a set of

weights $w_{j,\mathrm{obs}}$ where $\sum_j w_{j,\mathrm{obs}} = 1$ for each relative delay $\Delta t_{1,j}$ and magnification ratio μ_k/μ_1 between the jth image and S1.

In general, considering a discrete set of possible realizations (x_i, y_i) of (X, Y) with probability p_i, the covariance is

$$\mathrm{cov}(X, Y) = \sum_{i=1}^{n} p_i (x_i - E(X))(y_i - E(Y)). \qquad (10)$$
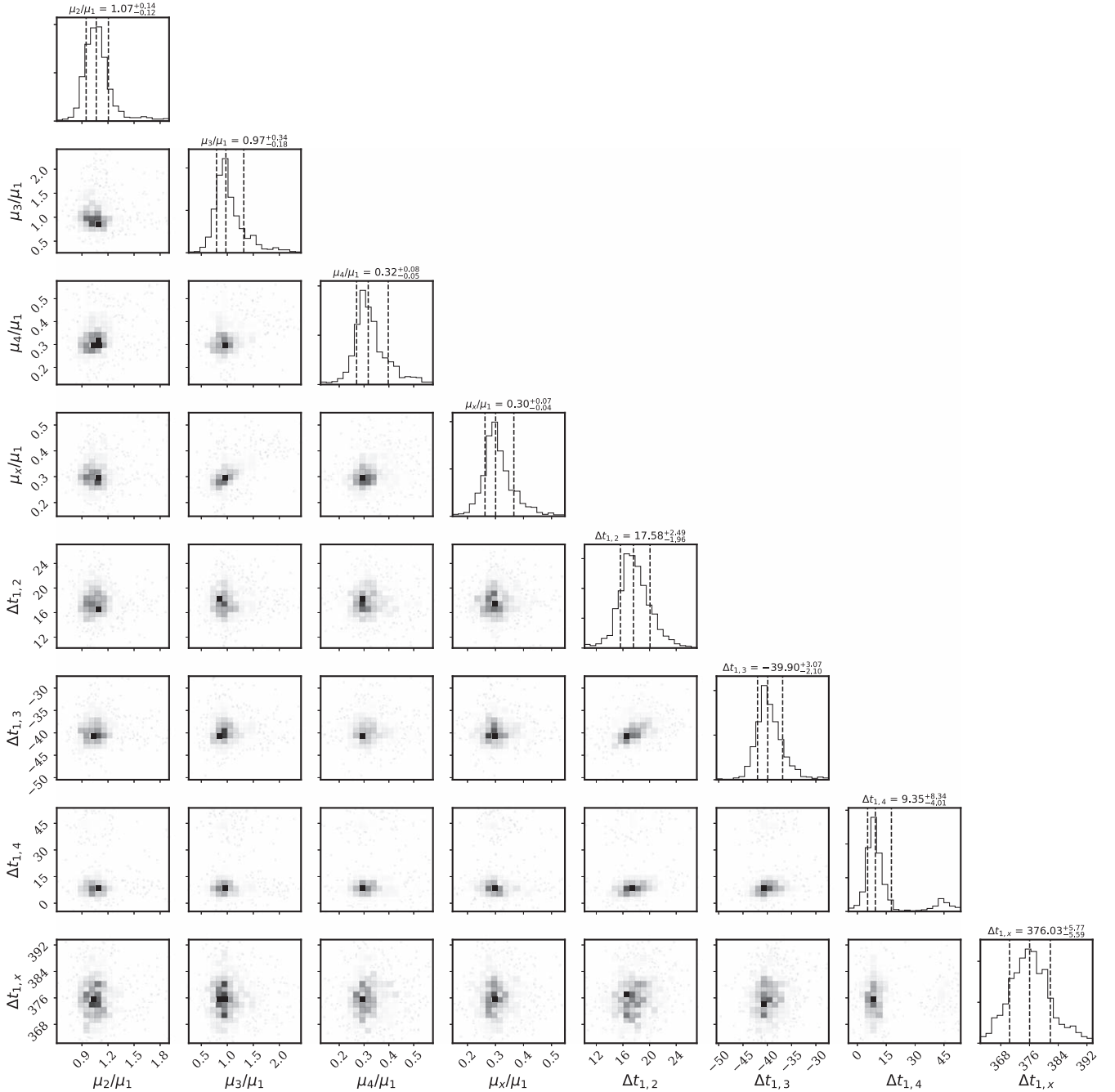
Figure 17. Relative time delays and magnification ratios among images S2–SX and S1. Constraints shown are weighted combination of results from four different algorithms, where weights are determined by maximizing the likelihood of the actual time delay used to construct simulated light curves. The values shown above each histogram correspond to the 16th, 50th, and 84th percentiles.

In our case, the probability p_i depends on the fitting method from which the x_i or y_i are drawn. The probability $p_i = (1/n_{\rm sim}) w_{i,X} \times w_{j,Y}$.

In our case, to create our realizations of the stacked probability distribution, we need to draw from the probability distributions. The probability of each pair of realizations corresponds to the product of the weights $w_{i,X} \times w_{j,Y}$ that the

respective fitting methods receive for the observable. We can write

$$\text{cov}(X, Y) = \sum_{i=1}^{n_{\rm sim}} \sum_{j=1}^{n_{\rm mod}} \sum_{k=1}^{n_{\rm mod}} \frac{1}{n_{\rm sim}} w_{j,X} \times w_{k,Y} \\ \times (x_{i,j} - E(X))(y_{i,k} - E(Y)), \quad (11)$$

Figure 18. Continuation of Figure 18. The light curves of the most highly magnified Einstein-cross images S1–S3 and that of image SX, after removing an SX–S1 time delay of 349 days (bottom panel). Given our constraints on the SX–S1 time delay of $376.0 + 5.6 − 5.5$ days, 349 days is offset by 5σ from the best-fitting value. Figure 19 plots the χ^2 value of the best-fitting piecewise polynomial model as a function of the SX–S1 relative time delay.
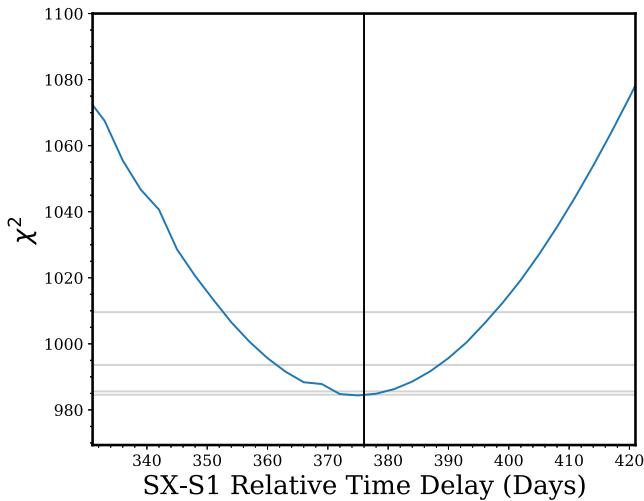
Figure 19. The χ^2 value of the best-fitting piecewise polynomial model as a function of the relative time delay between SX and S1. Here we hold the value of the SX–S1 time delay fixed while allowing all other time delay and magnification ratio parameters for all images, as well as the parameters of the light-curve model, to vary. The χ^2 values of the best-fitting model that we obtain for each value of the SX–S1 time delay are plotted above in blue. The black vertical line shows the most probable value of the SX–S1 time delay of 376 days. The gray horizontal lines show, from bottom to top, the minimum χ^2 value of 984.6 and added increments of 1, 9, and 25.

$$\text{cov}(X, Y) = \sum_{j=1}^{n_{\text{mod}}} \sum_{k=1}^{n_{\text{mod}}} w_{j,X} \times w_{k,Y}$$
$$\times \sum_{i=1}^{n_{\text{sim}}} \frac{1}{n_{\text{sim}}} (x_{i,j} - E(X))(y_{i,k} - E(Y)). \quad (12)$$

While there may be model covariance between measured or predicted image separations and the time delay and magnification, we do not include it here.

Unlike the effects of microlensing, the effects of millilensing were not included when we created simulated light curves. Since the Einstein radii of stars in the foreground cluster can be comparable in size to the expanding SN photosphere, microlensing can cause changes in the shape of the light curve, as well as the mean magnification of each image (e.g., Dobler & Keeton 2006). Given the much greater masses of subhalo deflectors expected to be responsible for millilensing, their effect should be to change the mean magnification but not affect the shape of the light curve.

We are interested in the covariance of random draws x_i and y_i from the linear combinations of distributions X and Y corresponding to two observables. The probability of drawing from the distribution X_j corresponding to the jth model is $w_{j,X}$, and the probability of drawing from the distribution Y_k corresponding to the kth model is $w_{k,Y}$. Therefore, the probability of drawing from both is $p = w_{j,X} \times w_{k,Y}$. We present this estimate for the covariance matrix in Table 17 and Figure 20. When computing the covariance matrix, we have excluded light curves for which a combined estimate from all fitting codes using the codes' respective weights represents a >30-day outlier for a time delay, or a 1.5 outlier for a magnification.

7. Summary and Discussion

We have described our blind measurements of the relative time delays and magnification ratios of the first known example of a multiply imaged SN. SN Refsdal provides a powerful and unique probe of the value of H_0 and a novel test of galaxy-cluster models. Our constraints employ all available HST WFC3 IR imaging data collected during multiple follow-up campaigns. To arrive at robust, as precise as possible, and accurate inferences about the relative time delays and magnification of the five observed images, we have used four separate light-curve fitting algorithms. Moreover, to minimize human bias, we have performed

Table 17
Covariance Matrix for Relative Time Delays and Magnification Ratio Measurements Calculated by Estimating Input Parameters from Simulated Light Curves That Include Microlensing, Millilensing, and Photometric Noise

	$\Delta t_{2,1}$	$\Delta t_{3,1}$	$\Delta t_{4,1}$	$\Delta t_{X,1}$	μ_2/μ_1	μ_3/μ_1	μ_4/μ_1	μ_X/μ_1
$\Delta t_{2,1}$	16.85806	4.60409	6.26366	4.74173	−0.07571	−0.01513	0.00951	−0.00102
$\Delta t_{3,1}$	4.60409	18.74664	4.21964	4.16457	−0.05818	−0.06709	−0.03741	−0.01790
$\Delta t_{4,1}$	6.26366	4.21964	34.00929	3.87574	0.02516	0.09217	−0.04550	0.01435
$\Delta t_{X,1}$	4.74173	4.16457	3.87574	30.64843	−0.00417	−0.04247	0.01111	0.01576
μ_2/μ_1	−0.07571	−0.05818	0.02516	−0.00417	0.05467	−0.00385	0.00659	0.00240
μ_3/μ_1	−0.01513	−0.06709	0.09217	−0.04247	−0.00385	0.10805	0.00578	0.01033
μ_4/μ_1	0.00951	−0.03741	−0.04550	0.01111	0.00659	0.00578	0.00663	0.00131
μ_X/μ_1	−0.00102	−0.01790	0.01435	0.01576	0.00240	0.01033	0.00131	0.00376

Note. Figure 20 shows a graphical representation of these values. The plotted covariance estimates do not include lens model uncertainties, which are added before computing likelihood values. When computing the covariance, we have excluded light curves for which a combined estimate from all fitting codes using the codes' respective weights represents a >30-day outlier for a time delay, or a 1.5 outlier for a magnification.
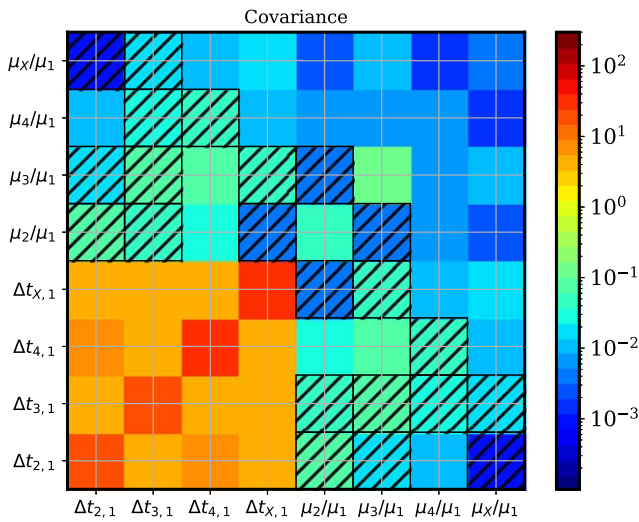
Figure 20. Expected covariance between relative time delays and magnification ratios among the images of SN Refsdal from microlensing, millilensing, photometric uncertainty, and measurement. Hashed squares denote negative values. The plotted covariance estimates do not include lens model uncertainties, which are added before computing likelihood values. When computing the covariance, we have excluded light curves for which a combined estimate from all fitting codes using the codes' respective weights represents a > 30-day outlier for a time delay, or a 1.5 outlier for a magnification.

the analysis in a blind manner by adding random offsets to the light-curve dates. As shown in Figure 2, we find evidence that image S4 has experienced chromatic microlensing, which is not expected given our simulations with a spherically symmetric model photosphere.

To determine the uncertainties of the inferred delays and magnifications and identify the optimal combination of inferences from the light-curve fitting techniques, we constructed simulated light curves with the same cadence and S/N ratios as the observed light curves that include the effects of microlensing, millilensing, and intrinsic variation in the light curves of SN 1987A–like SNe. We identified a powerful new statistical framework for optimally combining measurements from multiple separate light-curve fitting algorithms.

In the accompanying paper (Kelly et al. 2023), we use our measurements to obtain a first, blind measurement of the value of H_0 from a multiply imaged SN and compare the performance of

lens model predictions for the relative magnification and time delays of the five images of SN Refsdal.

In the upcoming decade, deep imaging surveys that use newly constructed, wide-field ground-based and space-based telescopes will discover hundreds of strongly lensed SNe that will be multiply imaged by both galaxy-scale and cluster-scale gravitational lenses. The Legacy Survey of Space and Time (LSST) with the Vera Rubin Observatory will discover hundreds of strongly lensed SNe and provide measurements of their light curves for wide-separation systems (e.g., Oguri & Marshall 2010; Goldstein & Nugent 2017; Huber et al. 2019; Wojtak et al. 2019). After launch in the mid-2020s, the Nancy Grace Roman Space Telescope will be used for a planned 2 yr SN Ia program, which may yield well-measured light curves of dozens of strongly lensed SNe (Oguri 2010; Pierel et al. 2021). While a large majority of the strongly lensed SNe will have foreground galaxy lenses, a significant fraction will be multiply imaged by foreground galaxy clusters (Petrushevska et al. 2018).

As the sample of strongly lensed SNe grows to hundreds of objects, we can anticipate at least two additional avenues of exploration. First is time delay cosmography, not only to measure H_0 but also to place constraints on additional cosmological parameters such as the dark energy equation-of-state parameter (Linder 2011; Treu & Marshall 2016; Suyu et al. 2018; Grillo et al. 2018). A second possibility is to anticipate the appearance of a trailing SN image and acquire observations very shortly after the time of explosion. This method to obtain early-time spectroscopy and photometry may provide powerful constraints on the properties of SN progenitor (e.g., Kasen 2010; Foxley-Marrable et al. 2020). Early spectroscopy of lensed SNe at $z > 1$ could enable tests for evolution of the SN progenitor populations across cosmic time.

We anticipate that the methods we have developed to extract the time delay and magnification ratios for SN Refsdal may be helpful to the analysis of the expected set of discoveries. In particular, our adaptation of DOLPHOT for difference imaging should be useful for the Roman Space Telescope. Moreover, the framework for the combination of estimates from multiple light-curve fitting algorithms, informed by their performance on simulated data, we hope will be useful.

SN Refsdal was surprising given that it was a peculiar, SN 1987A–like SN explosion of a blue supergiant star, which are not common in the nearby universe (Kelly et al. 2016a). Intriguingly, a second, highly magnified blue supergiant star was discovered in the same host galaxy (Kelly et al. 2018). The data set that we assembled for SN Refsdal required a large investment of observing time with HST. While the light-curve evolution of SN Refsdal is simple, comparably precise delays with much more modest light-curve sampling may be possible for spectroscopic types of SNe that are very well characterized, such as the multiply imaged SN iPTF16geu (Goobar et al. 2017).

Appendix

Figures A1–A5 show photometry for the five test stars in the F125W and F160W bands and measured using the two fitting codes (`DOLPHOT` and `PythonPhot`).
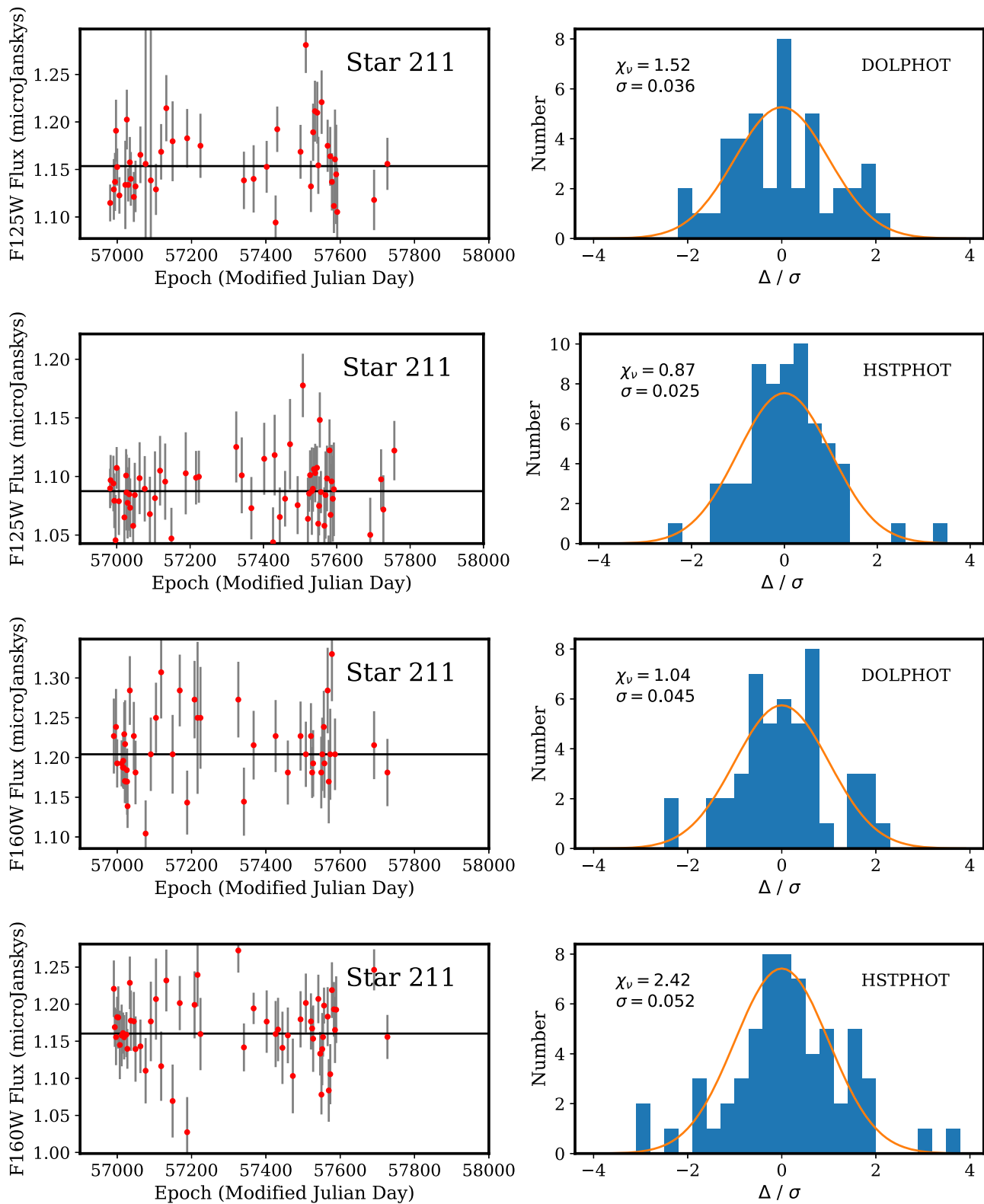
Figure A1. WFC3 IR F125W and F160W photometry of Star 211 in the MACS J1149 galaxy-cluster field with the DOLPHOT PSF-fitting code and aperture fluxes with HSTPHOT. The left panel shows flux measurements and median flux values. The right panel plots residuals from the median flux value normalized by the flux uncertainty estimate.
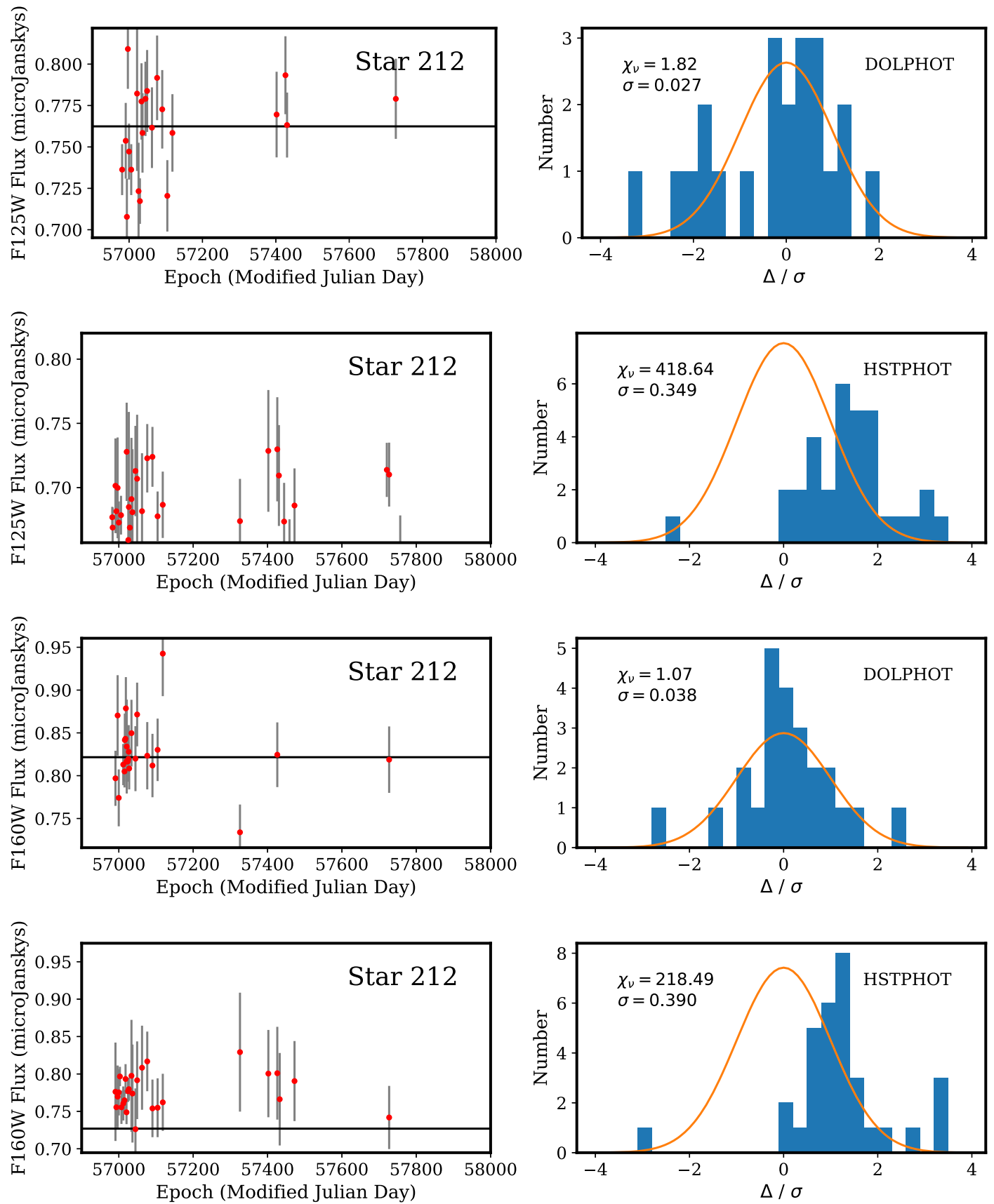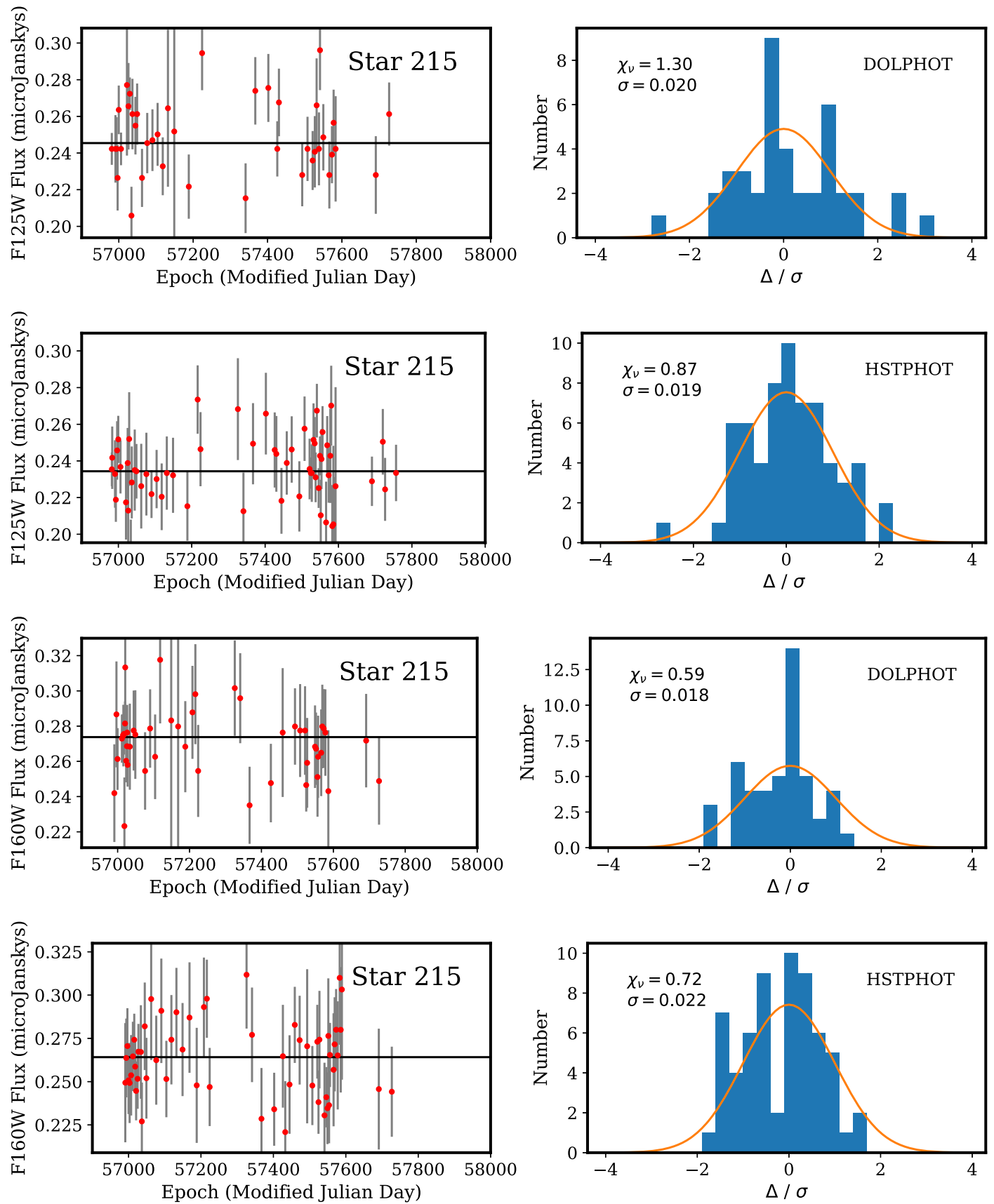
Figure A2. WFC3 IR F125W and F160W photometry of Star 212 in the MACS J1149 galaxy-cluster field with the DOLPHOT PSF-fitting code and aperture fluxes with HSTPHOT. The left panel shows flux measurements and median flux values. The right panel plots residuals from the median flux value normalized by the flux uncertainty estimate.
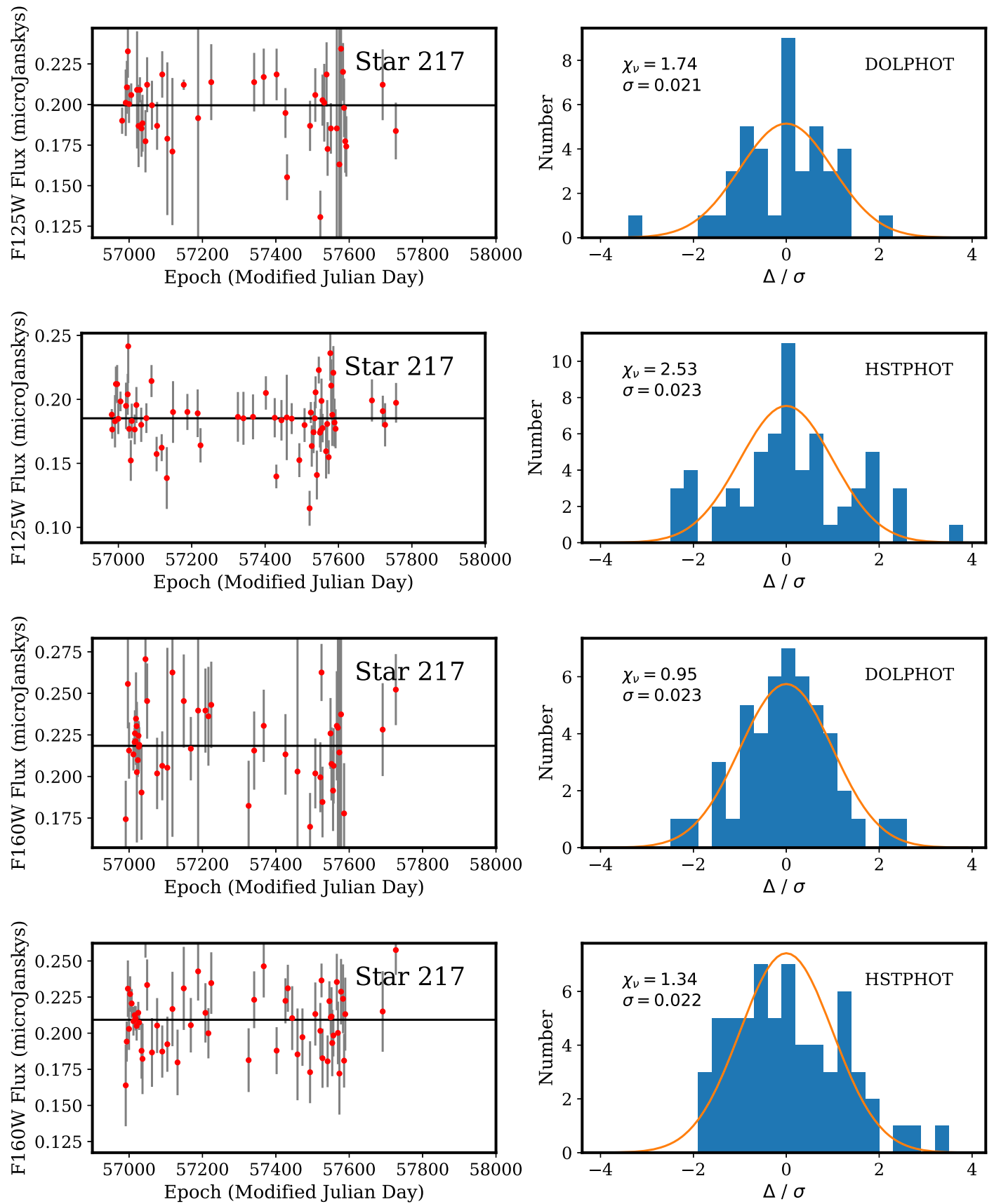
Figure A3. WFC3 IR F125W and F160W photometry of Star 215 in the MACS J1149 galaxy-cluster field with the DOLPHOT PSF-fitting code and aperture fluxes with HSTPHOT. The left panel shows flux measurements and median flux values. The right panel plots residuals from the median flux value normalized by the flux uncertainty estimate.

Figure A4. WFC3 IR F125W and F160W photometry of Star 217 in the MACS J1149 galaxy-cluster field with the DOLPHOT PSF-fitting code and aperture fluxes with HSTPHOT. The left panel shows flux measurements and median flux values. The right panel plots residuals from the median flux value normalized by the flux uncertainty estimate.
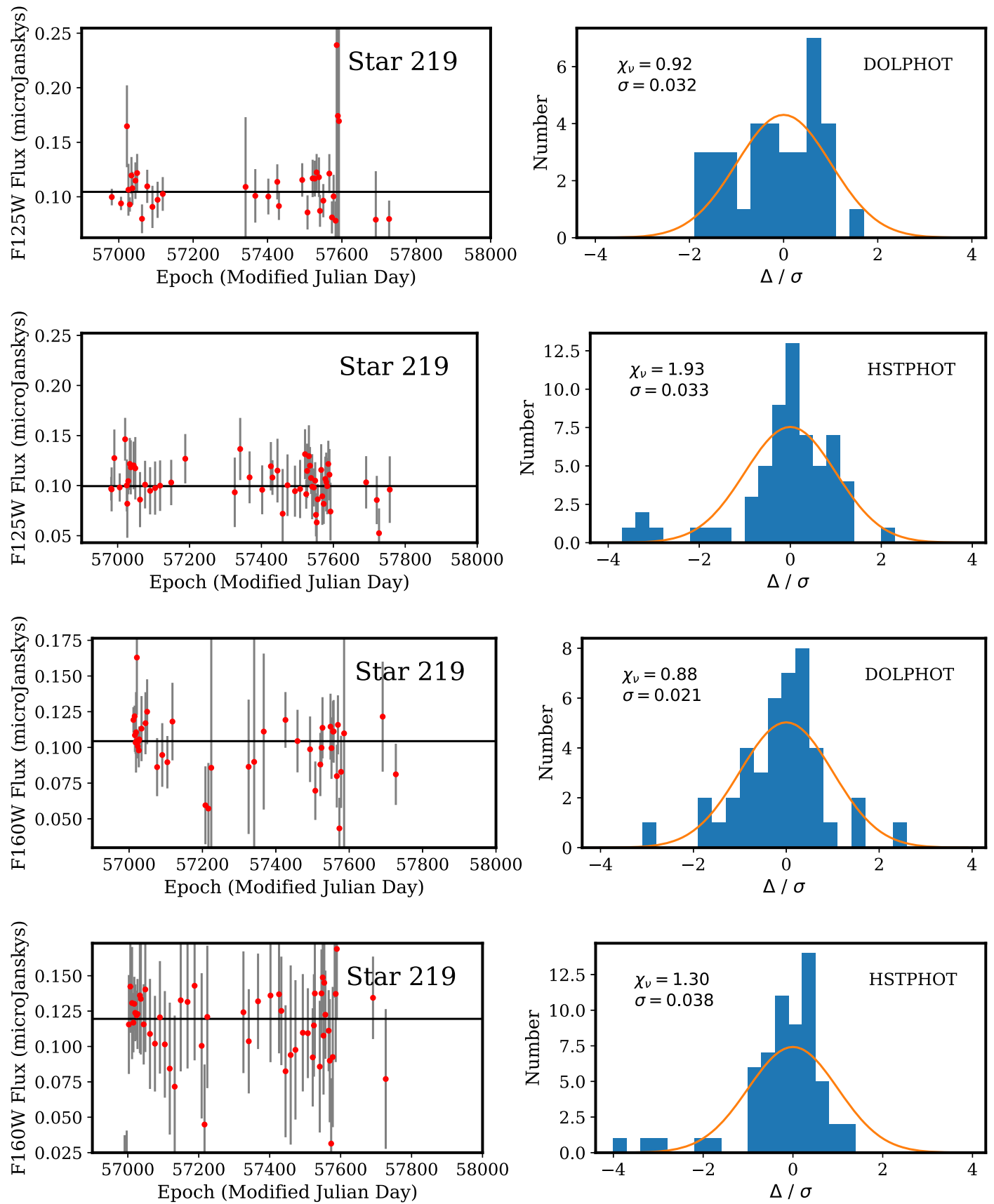
Figure A5. WFC3 IR F125W and F160W photometry of Star 219 in the MACS J1149 galaxy-cluster field with the DOLPHOT PSF-fitting code and aperture fluxes with HSTPHOT. The left panel shows flux measurements and median flux values. The right panel plots residuals from the median flux value normalized by the flux uncertainty estimate.
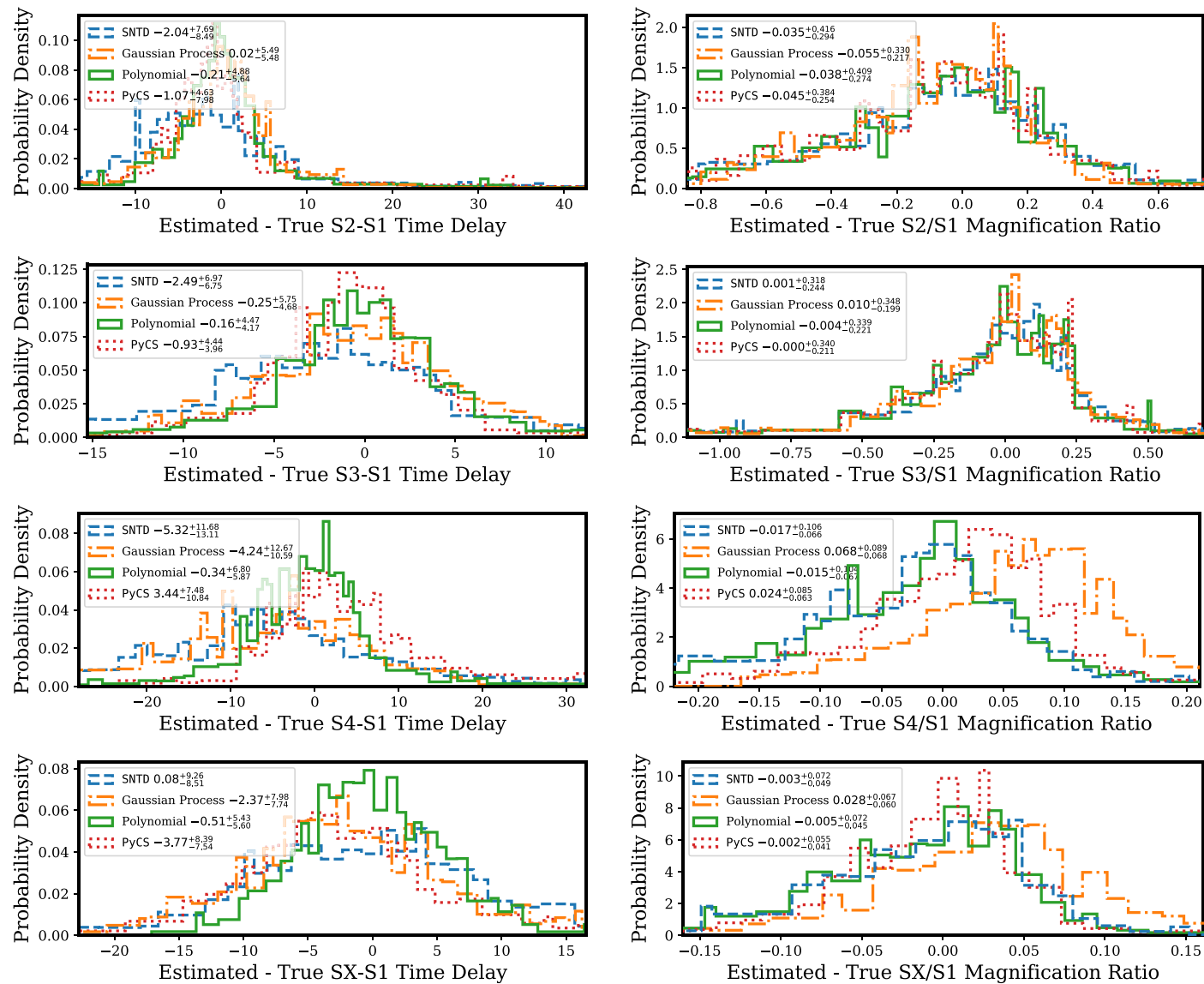
Figure A6. Distributions of residuals of time delay (left) and magnification ratio (right) estimated from simulated light curves by four separate algorithms given the assumption of a Salpeter (1955) IMF. As indicated, each row shows results from the simulated data that mimic one pair of images, from S2:S1 in the top tow to SX:S1 in the bottom row. The 16th, 50th, and 84th percentiles are listed in the legends.
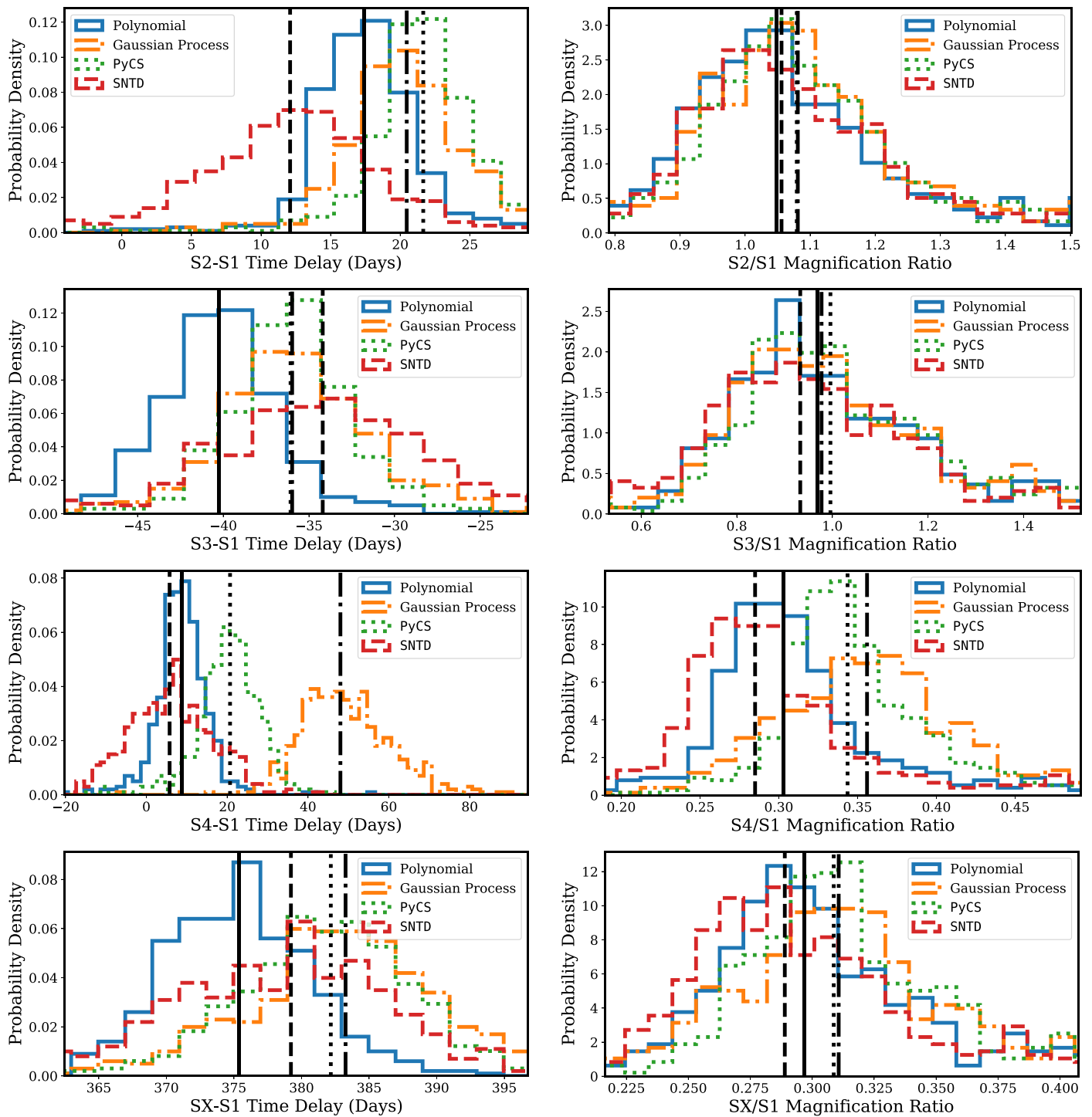
Figure A7. Constraints on the relative time delays and magnifications of the five images of SN Refsdal for the four fitting methods. The greatest tension is between the estimates of the S4–S1 time delay from the GP method and the piecewise polynomial method. The vertical black lines correspond to the 50th percentile of the respective distributions, and each vertical line has the same line style as that of the corresponding distribution.

Table A1
Salpeter rms Scatter

Method	Time Delay				Magnification Ratio			
	S2–S1	S3–S1	S4–S1	SX–S1	S2/S1	S3/S1	S4/S1	SX/S1
Polynomial	7.3	5.9	7.8	6.0	0.38	0.36	0.11	0.09
Gaussian process	6.7	6.0	11.5	8.5	0.32	0.34	0.11	0.10
SNTD	9.0	7.7	12.3	9.2	0.40	0.37	0.11	0.08
PyCS	7.6	6.1	9.6	8.7	0.35	0.36	0.09	0.09

Note. The rms scatter in the time delay columns is given in units of days. The magnification ratio columns are unitless and report the scatter of flux magnification ratios (i.e., no conversion into magnitudes has been done). When computing the scatter, we have excluded >30-day outliers for the time delays, and 1.5 for the magnification.

Table A2
Salpeter $>3\sigma$ Outlier Fraction

Method	Time Delay				Magnification Ratio			
	S2–S1	S3–S1	S4–S1	SX–S1	S2/S1	S3/S1	S4/S1	SX/S1
Polynomial	0.068	0.027	0.042	0.014	0.012	0.032	0.011	0.014
Gaussian process	0.045	0.015	0.048	0.011	0.010	0.027	0.010	0.015
SNTD	0.055	0.011	0.024	0.016	0.014	0.026	0.009	0.018
PyCS	0.087	0.030	0.060	0.078	0.012	0.031	0.011	0.016

Note. Outliers are defined as time delay measurements that differ from the mean by more than 30 days, or measured magnification ratios that deviate by more than 1.

ORCID iDs

Patrick L. Kelly ⓘ https://orcid.org/0000-0003-3142-997X
Tommaso Treu ⓘ https://orcid.org/0000-0002-8460-0390
Simon Birrer ⓘ https://orcid.org/0000-0003-3195-5507
Vivien Bonvin ⓘ https://orcid.org/0000-0003-1471-3952
Luc Dessart ⓘ https://orcid.org/0000-0003-0599-8407
Ryan J. Foley ⓘ https://orcid.org/0000-0002-2445-5275
Alexei V. Filippenko ⓘ https://orcid.org/0000-0003-3460-0103
Daniel Gilman ⓘ https://orcid.org/0000-0002-5116-7287
Saurabh Jha ⓘ https://orcid.org/0000-0001-8738-6011
Jens Hjorth ⓘ https://orcid.org/0000-0002-4571-2306
Kaisey Mandel ⓘ https://orcid.org/0000-0001-9846-4417
Martin Millon ⓘ https://orcid.org/0000-0001-7051-497X
Justin Pierel ⓘ https://orcid.org/0000-0002-2361-7201
Adi Zitrin ⓘ https://orcid.org/0000-0002-0350-4488
Wenlei Chen ⓘ https://orcid.org/0000-0003-1060-0723
Jose M. Diego ⓘ https://orcid.org/0000-0001-9065-3926
Alan Dressler ⓘ https://orcid.org/0000-0002-6317-0037
Or Graur ⓘ https://orcid.org/0000-0002-4391-6137
Mathilde Jauzac ⓘ https://orcid.org/0000-0003-1974-8732
Matthew A. Malkan ⓘ https://orcid.org/0000-0001-6919-1237
Curtis McCully ⓘ https://orcid.org/0000-0001-5807-7893
Masamune Oguri ⓘ https://orcid.org/0000-0003-3484-399X
Marc Postman ⓘ https://orcid.org/0000-0002-9365-7989
Kasper Borello Schmidt ⓘ https://orcid.org/0000-0002-3418-7251
Keren Sharon ⓘ https://orcid.org/0000-0002-7559-0864
Brad E. Tucker ⓘ https://orcid.org/0000-0002-4283-5159
Anja von der Linden ⓘ https://orcid.org/0000-0002-3881-7724
Joachim Wambsganss ⓘ https://orcid.org/0000-0002-8365-7619

References

Akaike, H. 1974, ITAC, 19, 716
Baklanov, P., Lyskova, N., Blinnikov, S., & Nomoto, K. 2021, ApJ, 907, 35
Bazin, G., Palanque-Delabrouille, N., Rich, J., et al. 2009, A&A, 499, 653
Bonvin, V., Tewes, M., Courbin, F., et al. 2016, A&A, 585, A88
Bonvin, V., Tihhonova, O., Millon, M., et al. 2019, A&A, 621, A55
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Chabrier, G. 2003, ApJL, 586, L133
Conroy, C., & van Dokkum, P. G. 2012, ApJ, 760, 71
Dessart, L., & Hillier, D. J. 2019, A&A, 622, A70
Diego, J. M., Broadhurst, T., Chen, C., et al. 2016, MNRAS, 456, 356
Dobler, G., Fassnacht, C. D., Treu, T., et al. 2015, ApJ, 799, 168
Dobler, G., & Keeton, C. R. 2006, ApJ, 653, 1391
Dolphin, A. E. 2000, PASP, 112, 1383
Ebeling, H., Barrett, E., Donovan, D., et al. 2007, ApJL, 661, L33
Ebeling, H., Edge, A. C., & Henry, J. P. 2001, ApJ, 553, 668
Einstein, A. 1936, Sci, 84, 506
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Foxley-Marrable, M., Collett, T. E., Frohmaier, C., et al. 2020, MNRAS, 495, 4622
Fruchter, A. S., Hack, W., Dencheva, N., Droettboom, M., & Greenfield, P. 2010, in 2010 Space Telescope Science Institute Calibration Workshop (Baltimore, MD: STScI), 382
Gilman, D., Birrer, S., Nierenberg, A., et al. 2020, MNRAS, 491, 6077
Gilman, D., Birrer, S., Treu, T., Nierenberg, A., & Benson, A. 2019, MNRAS, 487, 5721
Goldstein, D. A., & Nugent, P. E. 2017, ApJL, 834, L5
Goldstein, D. A., Nugent, P. E., Kasen, D. N., & Collett, T. E. 2018, ApJ, 855, 22
Goobar, A., Amanullah, R., Kulkarni, S. R., et al. 2017, Sci, 356, 291
Grillo, C., Karman, W., Suyu, S. H., et al. 2016, ApJ, 822, 78
Grillo, C., Rosati, P., Suyu, S. H., et al. 2018, ApJ, 860, 94
Hojjati, A., Kim, A. G., & Linder, E. V. 2013, PhRvD, 87, 123512
Huber, S., Suyu, S. H., Noebauer, U. M., et al. 2019, A&A, 631, A161
Jauzac, M., Richard, J., Limousin, M., et al. 2016, MNRAS, 457, 2029
Jones, D. O., Scolnic, D. M., & Rodney, S. A. 2015, PythonPhot: Simple DAOPHOT-type photometry in Python, Astrophysics Source Code Library, ascl:1501.010
Karman, W., Grillo, C., Balestra, I., et al. 2016, A&A, 585, A27
Kasen, D. 2010, ApJ, 708, 1025

Kawamata, R., Oguri, M., Ishigaki, M., Shimasaku, K., & Ouchi, M. 2016, ApJ, 819, 114
Kelly, P. L., Brammer, G., Selsing, J., et al. 2016a, ApJ, 831, 205
Kelly, P. L., Diego, J. M., Rodney, S., et al. 2018, NatAs, 2, 334
Kelly, P. L., Rodney, S., Diego, J. M., et al. 2016b, ATel, 9097, 1
Kelly, P. L., Rodney, S., Treu, T., et al. 2023, Sci, in press
Kelly, P. L., Rodney, S. A., Treu, T., et al. 2015, Sci, 347, 1123
Kelly, P. L., Rodney, S. A., Treu, T., et al. 2016c, ApJL, 819, L8
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Liao, K., Treu, T., Marshall, P., et al. 2015, ApJ, 800, 11
Linder, E. V. 2011, PhRvD, 84, 123529
Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, ApJ, 837, 97
Millon, M., Courbin, F., Bonvin, V., et al. 2020a, A&A, 640, A105
Millon, M., Courbin, F., Bonvin, V., et al. 2020b, A&A, 642, A193
Molinari, N., Durand, J.-F., & Sabatier, R. 2004, Comp. Stat. Data Anal., 45, 159
Morishita, T., Abramson, L. E., Treu, T., et al. 2017, ApJ, 846, 139
Oguri, M. 2010, PASJ, 62, 1017
Oguri, M. 2015, MNRAS, 449, L86
Oguri, M. 2019, RPPh, 82, 126901
Oguri, M., & Marshall, P. J. 2010, MNRAS, 405, 2579
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Petrushevska, T., Goobar, A., Lagattuta, D. J., et al. 2018, A&A, 614, A103
Pierel, J. D. R., & Rodney, S. 2019, ApJ, 876, 107
Pierel, J. D. R., Rodney, S., Vernardos, G., et al. 2021, ApJ, 908, 190
Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25
Rasmussen, C. E., & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning (Cambridge, MA: MIT Press)
Refsdal, S. 1964, MNRAS, 128, 307
Rodney, S. A., Patel, B., Scolnic, D., et al. 2015, ApJ, 811, 70

Rodney, S. A., Strolger, L.-G., Kelly, P. L., et al. 2016, ApJ, 820, 50
Salpeter, E. E. 1955, ApJ, 121, 161
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schmidt, K. B., Treu, T., Brammer, G. B., et al. 2014, ApJL, 782, L36
Sharon, K., & Johnson, T. L. 2015, ApJL, 800, L26
Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1
Skilling, J. 2004, in 24th Int. Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer et al. (Melville, NY: AIP), doi:10.1063/1.1835238
Skilling, J. 2006, Bayesian Anal., 1, 833
Speagle, J. S. 2020, MNRAS, 493, 3132
Stetson, P. B. 1987, PASP, 99, 191
Suyu, S. H., Chang, T.-C., Courbin, F., & Okumura, T. 2018, SSRv, 214, 91
Tak, H., Mandel, K., van Dyk, D. A., et al. 2017, AnApS, 11, 1309
Tewes, M., Courbin, F., & Meylan, G. 2013, A&A, 553, A120
Treu, T. 2010, ARA&A, 48, 87
Treu, T., Brammer, G., Diego, J. M., et al. 2016, ApJ, 817, 60
Treu, T., & Marshall, P. J. 2016, A&ARv, 24, 11
Treu, T., Schmidt, K. B., Brammer, G. B., et al. 2015, ApJ, 812, 114
Vega-Ferrero, J., Diego, J. M., Miranda, V., & Bernstein, G. M. 2018, ApJL, 853, L31
Wambsganss, J. 1990, PhD thesis, Fakultät für Physik
Wambsganss, J. 1999, JCoAM, 109, 353
Wojtak, R., Hjorth, J., & Gall, C. 2019, MNRAS, 487, 3342
Yao, Y., Vehtari, A., Simpson, D., & Gelman, A. 2018, Bayesian Anal., 13, 917
Yuan, T., Kobayashi, C., & Kewley, L. J. 2015, ApJL, 804, L14
Zwicky, F. 1937a, PhRv, 51, 290
Zwicky, F. 1937b, PhRv, 51, 679