



On the Accuracy of Time-delay Cosmography in the Frontier Fields Cluster MACS J1149.5+2223 with Supernova Refsdal

C. Grillo^{1,2,3} , P. Rosati^{4,5}, S. H. Suyu^{6,7,8}, G. B. Caminha⁹, A. Mercurio¹⁰, and A. Halkola

¹ Dipartimento di Fisica, Università degli Studi di Milano, via Celoria 16, I-20133 Milano, Italy; claudio.grillo@unimi.it

² Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, DK-2100 Copenhagen, Denmark

³ INAF—IASF Milano, via A. Corti 12, I-20133 Milano, Italy

⁴ Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Saragat 1, I-44122 Ferrara, Italy

⁵ INAF—Osservatorio Astronomico di Bologna, via Gobetti 93/3, I-40129 Bologna, Italy

⁶ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85748 Garching, Germany

⁷ Physik-Department, Technische Universität München, James-Frank-Straße 1, D-85748 Garching, Germany

⁸ Academia Sinica Institute of Astronomy and Astrophysics (ASIAA), 11F of ASMA, No.1, Section 4, Roosevelt Road, Taipei 10617, Taiwan

⁹ Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, The Netherlands

¹⁰ INAF—Osservatorio Astronomico di Capodimonte, Via Moiariello 16, I-80131 Napoli, Italy

Received 2019 December 18; revised 2020 June 4; accepted 2020 June 6; published 2020 July 27

Abstract

We study possible systematic effects on the values of the cosmological parameters measured through strong lensing analyses of the Hubble Frontier Field galaxy cluster MACS J1149.5+2223. We use the observed positions of a large set of spectroscopically selected multiple images, including those of supernova “Refsdal” with their published time delays. Starting from our reference model in a flat Λ CDM cosmology, published in Grillo et al. (2018), we confirm the relevance of the longest measurable time delay, between SX and S1, and an approximately linear relation between its value and that of H_0 . We perform true blind tests by considering a range of time delays around its original estimate of 345 ± 10 days, as an accurate measurement of this time delay is still not known at the time of analysis and writing. We investigate separately the impact of a constant sheet of mass at the cluster redshift, of a power-law profile for the mass density of the cluster main halo and of some scatter in the cluster member scaling relations. Remarkably, we find that these systematic effects do not introduce a significant bias on the inferred values of H_0 and Ω_m and that the statistical uncertainties dominate the total error budget: a 3% uncertainty on the time delay of image SX translates into approximately 6% and 40% (including both statistical and systematic 1σ) uncertainties for H_0 and Ω_m , respectively. Furthermore, our model accurately reproduces the extended surface brightness distribution of the supernova host. We also present the interesting possibility of measuring the value of the equation-of-state parameter w of the dark energy density, currently with a 30% uncertainty.

Unified Astronomy Thesaurus concepts: Strong gravitational lensing (1643); Cosmological parameters (339); Dark matter (353); Dark energy (351); Galaxy clusters (584); Hubble constant (758)

1. Introduction

After nearly a century from its first estimates (Lemaître 1927; Hubble 1929), the exact value of the cosmic expansion rate, the Hubble constant H_0 , is still hotly debated. The most recent measurements of H_0 , based on the distance ladder from the SH0ES program (Riess et al. 2019) and from the Planck satellite (Planck Collaboration VI 2018), cannot be reconciled. Accurate and precise estimates from independent methods and calibrations (e.g., Reid et al. 2013; Abbott et al. 2017; Freedman et al. 2019; Wong et al. 2020) are thus crucial to clarify whether an extension of the standard cosmological model is needed. Time delays between the multiple images of supernovae (SNe) and quasars (QSOs) strongly lensed by galaxy clusters can place interesting new constraints on the values of H_0 and of the parameters defining the global geometry of the universe (e.g., Refsdal 1964; Chen et al. 2019; Shajib et al. 2020; Wong et al. 2020).

In this paper, we use a full strong lensing model of the Hubble Frontier Fields (HFFs) galaxy cluster MACS J1149.5+2223 (hereafter MACS 1149; $z = 0.542$), where the first multiply imaged and spatially resolved SN “Refsdal” ($z = 1.489$) was discovered (Kelly et al. 2015, 2016a). We extend our previous analysis (Grillo et al. 2018, hereafter G18) and show that a large set of spectroscopically confirmed multiple images from

different sources (Grillo et al. 2016, hereafter G16, Treu et al. 2016) and the measured time delays between the multiple images of a variable source (Kelly et al. 2016b; Rodney et al. 2016) in a lens galaxy cluster can provide accurate measurements of the values of the Hubble constant and of cosmological parameters, which are competitive with those from other cosmological probes. Alternative methods exploiting SN Refsdal for cosmography were presented in other studies (Vega-Ferrero et al. 2018; Williams & Liesenborgs 2019).

2. Methods

Our reference strong lensing model of MACS 1149 (labeled as r/REF in the following tables and figures) is the one detailed in G16 and G18. We refer the reader to those papers for a comprehensive description. Here, we summarize briefly its main characteristics. To reconstruct the cluster total mass distribution and infer the values of the considered cosmological parameters, we always use the observed positions of 89 multiple images from 28 different point-like sources, with redshifts between 1.240 and 3.703, and the values of the time delays of the images S2, S3, and S4, relative to S1, of SN Refsdal, as measured by using polynomial fitting by Rodney et al. (2016). The value of the SX–S1 time delay ($\Delta t_{\text{SX:S1}}$) is based on the first results by Kelly et al. (2016b) and is varied

between 315 and 375 days, with a constant uncertainty of 10 days. The cluster total mass distribution is modeled with a combination of 3 cored elliptical pseudo-isothermal and 300 dual pseudo-isothermal mass density profiles for, respectively, the extended dark-matter halos and the galaxy members of the cluster. Flat ($\Omega_m + \Omega_\Lambda = 1$) Λ CDM models with uniform priors on the values of H_0 , between 20 and 120 $\text{km s}^{-1} \text{Mpc}^{-1}$, and Ω_m , between 0 and 1, are considered. The software GLEE (Suyu & Halkola 2010; Suyu et al. 2012) is used to conduct the entire strong-lensing study.

First, to analyze the dependence of the inferred value of H_0 on the observed value of the time delay between SX and S1, we perform full statistical analyses with the reference model and nine different values for this time delay (i.e., 315, 322, 330, 337, 345, 353, 360, 368, or 375 days), keeping its uncertainty fixed to 10 days. We also check the importance of the SX–S1 time delay in estimating the values of H_0 and Ω_m by comparing the results of the reference model with and without this observable quantity.

Then, to test the effect of the most important sources of systematic uncertainties on the measurement of the values of the cosmological parameters, we consider additional models, where we (1) add to the reference model (with 315, 345, or 375 ± 10 days for the SX–S1 time delay) a constant sheet of mass at the cluster redshift with a uniform prior between -0.2 and 0.2 on the value of the convergence k_0 (models labeled as κ); (2) substitute (with 315, 345, or 375 ± 10 days for the SX–S1 time delay) the pseudo-isothermal profile (i.e., $\rho(r) \sim r^{-2}$, in the outer regions $r \gg r_c$) of the cluster main halo with a more general power-law profile (i.e., $\rho(r) \sim r^{-\gamma}$, in the outer regions $r \gg r_c$) with a uniform prior between 1.4 and 2.6 on the value of γ (models labeled as γ); (3) use five different realizations (all with 345 ± 10 days for the SX–S1 time delay) of the values of the Einstein angles (ϑ_E) and truncation radii (r_t) of the cluster members (models labeled as s1–s5), randomly scattered with normal distributions centered on the values obtained with total mass-to-light ratios increasing according to the galaxy near-IR luminosities and with standard deviations equal to 10% of their values (see Bergamini et al. 2019 for recent results on cluster member modeling with stellar kinematics information).

Next, to quantify the goodness of our best-fitting reference model (with a time delay between SX and S1 of 345 ± 10 days), we compare the observed and model-predicted surface brightness distribution of the multiple images of the SN Refsdal host. We note that in this model all the sources are approximated as point-like objects and the extended surface brightness information has not been exploited yet. We then use the best-fitting reference model, based on the point images, to reconstruct the SN host galaxy on a grid of pixels on the source plane from the lensed and distorted multiple images of the SN host galaxy. To suppress the contaminating contribution of the cluster members' light to the lensed SN host galaxy light, we use a linear combination of the Hubble Space Telescope (HST) F606W and F435W bands, with a pixel size of $0''.06$, from the HFF project. In each pixel, the observational uncertainty is estimated by propagating those measured in the two HST bands through the ASTRODRIZZLE software (Koekemoer et al. 2002). Due to the very large number of pixels over which the multiple images of the SN Refsdal host galaxy extend, our computing memory sets a maximum size of 75×75 pixels for the reconstruction grid of the source surface brightness. From the

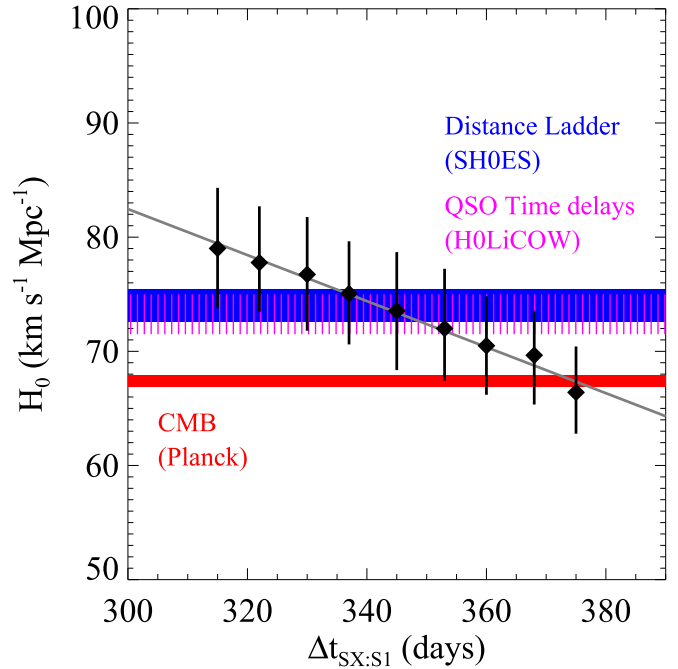


Figure 1. Inferred median values (diamonds) and 68% CL intervals (bars) of H_0 for reference (r) models with different adopted values for the time delay between SX and S1 (315, 322, 330, 337, 345, 353, 360, 368, or 375 days) and a constant uncertainty of 10 days (the time delays of the images S2, S3, and S4, relative to S1, measured by Rodney et al. 2016, are also included). The gray line shows the linear fit of the inferred values. Flat Λ CDM models ($\Omega_m + \Omega_\Lambda = 1$) with uniform priors on the values of the cosmological parameters ($H_0 \in [20, 120] \text{ km s}^{-1} \text{Mpc}^{-1}$ and $\Omega_m \in [0, 1]$) are considered. Final MCMC chains have 2×10^6 samples for each model. The blue, purple, and red bands indicate, respectively, the credible intervals, at 1σ CL, from SH0ES (Riess et al. 2019), H0LiCOW (Wong et al. 2020), and Planck (Planck Collaboration VI 2018).

reconstructed SN host galaxy surface brightness on the source plane, we then lens it with the mass model to the image plane to compare with the observed lensed images.

Finally, we consider models (with 315, 345, or 375 ± 10 days for the SX–S1 time delay) in a flat ($\Omega_m + \Omega_\Lambda = 1$) w CDM cosmology, in which the dark energy density is time dependent, with an equation-of-state parameter w . Here, the value of w is also free to vary between -2 and 0 , with a uniform prior.

3. Results

We remark that all the results presented here come from a blind analysis, because the exact value of the time delay $\Delta t_{\text{SX:S1}}$ was not known at the time of analysis and writing.

We show in Figure 1 the median values of H_0 and the 68% confidence level (CL) intervals obtained by varying the value of the time delay between SX and S1 from 315 to 375 days, with linear steps of 7–8 days, as described in Section 2. The inferred value of H_0 decreases from 79.0 to 66.4 $\text{km s}^{-1} \text{Mpc}^{-1}$ as the time-delay value increases. The assumed constant uncertainty of 10 days on the SX–S1 time delay translates into (1σ) statistical errors for H_0 between 5.8 and 7.0%, with a median value of 6.1%, consistent with the results presented in G18. A simple linear regression analysis of the H_0 values with errors provides the following scaling result: $[73.4 - 0.202 \text{ day}^{-1} \times (\Delta t_{\text{SX:S1}} - 345 \text{ days})] \text{ km s}^{-1} \text{Mpc}^{-1}$. This is

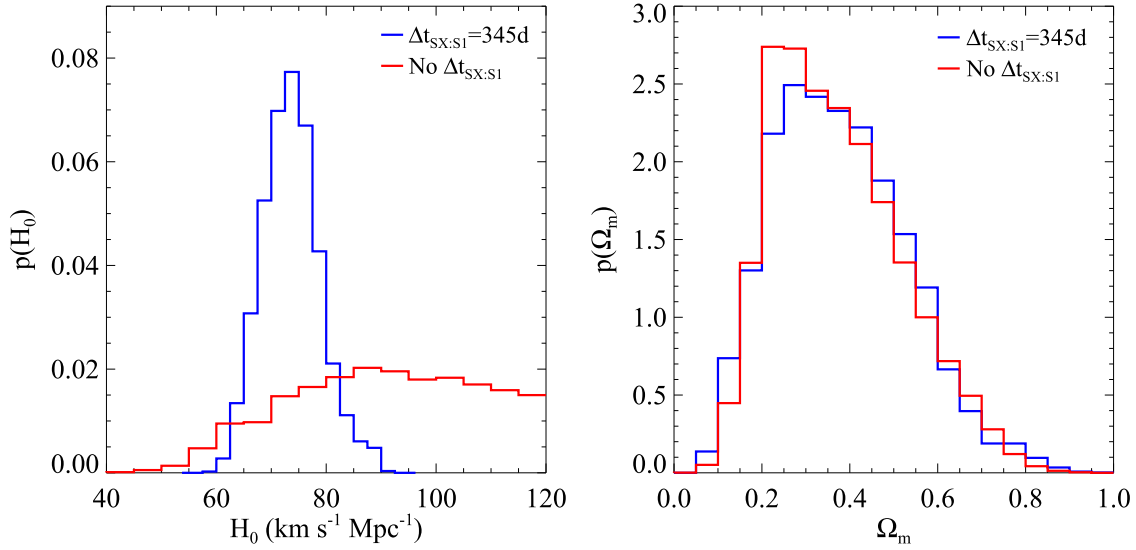


Figure 2. Marginalized probability distribution functions of H_0 (on the left) and Ω_m (on the right) for reference models with a time delay between SX and S1 of 345 ± 10 days (blue) and no time delay (red). The observed time delays (and their statistical errors) of the images S2, S3, and S4, relative to S1, are included in both cases. Flat Λ CDM models ($\Omega_m + \Omega_\Lambda = 1$) with uniform priors on the values of the cosmological parameters ($H_0 \in [20, 120]$ $\text{km s}^{-1} \text{Mpc}^{-1}$ and $\Omega_m \in [0, 1]$) are considered. Final MCMC chains have 2×10^6 samples for each model.

slightly less steep, but consistent, given the errors, with that published in G18.¹¹

We have also checked that an error of 2% on the value of $\Delta t_{\text{SX:S1}} = 345$ days (i.e., 7 days) results into an approximately 1% smaller error on the value of H_0 (similarly to G18) and that a smaller uniform prior between 0.05 and 0.5 for the value of Ω_m (as adopted, for instance, in Wong et al. 2020) reduces by more than 10% the error on its value and by less than 1% the error on the value of H_0 . This confirms that the present uncertainty on the reconstructed cluster total mass distribution (i.e., on the Fermat potential) represents the largest contribution ($\sim 5\%$) to the total error budget on the measurement of H_0 .

In Figure 2, we illustrate explicitly the relevance of the longest measurable time-delay value between the multiple images of SN Refsdal, i.e., the SX–S1 time delay, to the cosmological inference within our reference model. We contrast the inferred results for H_0 and Ω_m with a 345 ± 10 day time delay and without this time delay. We obtain that the probability distribution function of H_0 changes dramatically, from unimodal with a well-defined peak and small dispersion to almost flat over the largest values considered within the uniform prior. Remarkably, the probability distribution function of Ω_m is instead essentially unaltered, with very similar (within 3%) median values and (1σ) statistical errors.

¹¹ We have checked that the median (maximum) of the absolute values of the differences between the nine measured values of H_0 and those obtained from the linear fit is 0.3 (0.9) $\text{km s}^{-1} \text{Mpc}^{-1}$, corresponding to 0.3% (1.4%) of the estimated H_0 values. We have also checked that the median (maximum) value of the absolute values of the differences between the measured values and those obtained from the linear fit of H_0 divided by the errors obtained from our analysis (i.e., the error bars shown in Figure 1) is less than 0.1 (~ 0.2). We can conclude that the linear relation between the values of H_0 and $\Delta t_{\text{SX:S1}}$ is supported by the data and that the very small differences between the measured and fitted values of H_0 are on average less than $1/300$ of the measured values of H_0 and approximately $1/20$ of the measurement errors. We remark though that the most accurate and precise measurements of the values of the cosmological parameters and their errors from MACS 1149 and SN Refsdal will require an additional run of the strong lensing models, which include the final measurement of the value and uncertainty of the SX–S1 time delay.

Table 1
Median Values and 1σ CL Uncertainties of the Hubble Constant H_0 (in $\text{km s}^{-1} \text{Mpc}^{-1}$) and Ω_m for the Models Shown in Figure 3

$\Delta t_{\text{SX:S1}}$	Model	H_0	Ω_m
315 days	r	79.0 ± 5.3	0.32 ± 0.14
	κ	79.7 ± 5.2	0.39 ± 0.16
	γ	78.7 ± 4.8	0.33 ± 0.12
345 days	r	73.5 ± 5.2	0.37 ± 0.16
	κ	73.1 ± 4.3	0.38 ± 0.15
	γ	73.9 ± 3.9	0.39 ± 0.13
375 days	r	66.4 ± 3.8	0.48 ± 0.18
	κ	66.4 ± 4.0	0.49 ± 0.21
	γ	65.5 ± 3.7	0.47 ± 0.10
345 days	s1	73.5 ± 4.3	0.37 ± 0.15
	s2	73.5 ± 4.7	0.40 ± 0.15
	s3	72.3 ± 4.3	0.34 ± 0.13
	s4	73.5 ± 4.2	0.36 ± 0.15
	s5	72.9 ± 4.3	0.38 ± 0.17

We summarize in Table 1 and in Figures 3 and 4 the outcomes of our tests to estimate the systematic uncertainties on the values of H_0 and Ω_m , as detailed in Section 2. The most important result is that, for a fixed value of the SX–S1 time delay, the inferred values of the cosmological parameters are not significantly affected by the modeling details (considering their errors). For instance, when we use a time-delay value of 345 ± 10 days for $\Delta t_{\text{SX:S1}}$, the median values of H_0 and Ω_m vary, respectively, between 72.9 and 73.9 $\text{km s}^{-1} \text{Mpc}^{-1}$ and 0.34 and 0.40 and the results of our reference model are not biased low or high. The combination of the eight final MCMC chains of the models r, κ , γ , and s1–5 with $\Delta t_{\text{SX:S1}} = 345 \pm 10$ days provides a global MCMC chain with 1.6×10^7 samples with median values and 1σ uncertainties of $73.3 \text{ km s}^{-1} \text{Mpc}^{-1}$ and 0.37 and 6% and 40% for H_0 and Ω_m , respectively. By comparing these results with those published in G18, we can conclude that the error budget on the estimate of H_0 is strikingly dominated by the

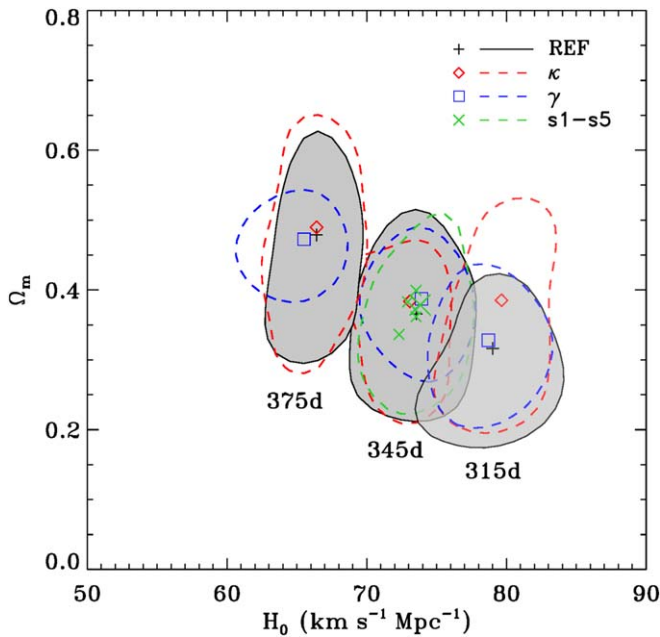


Figure 3. Median values and contour levels, at 1σ CL, of H_0 and Ω_m for different lensing models. Time delays between SX and S1 of 315 ± 10 , 345 ± 10 , and 375 ± 10 days are adopted (the time delays of the images S2, S3, and S4, relative to S1, measured by Rodney et al. 2016, are also included). Reference (REF/r) models are shown with black pluses and solid contours filled in gray. The addition of a constant sheet of mass (κ) at the cluster redshift (diamonds and dashed red contours), the use of a variable central density slope (γ) for the cluster main halo (squares and dashed blue contours), and 10% scatter (s) on the values of ϑ_E and r_t of the cluster members (crosses and dashed green averaged contours) are tested. Flat Λ CDM models ($\Omega_m + \Omega_\Lambda = 1$) with uniform priors on the values of the cosmological parameters ($H_0 \in [20, 120]$ km s $^{-1}$ Mpc $^{-1}$ and $\Omega_m \in [0, 1]$) are considered. Final MCMC chains have 2×10^6 samples for each model.

statistical uncertainties and the known systematics play only a secondary role. We notice that the latter are slightly more relevant for the measurement of Ω_m .

We remark that very similar general conclusions hold when the other two values, i.e., 315 and 375 ± 10 days, of the SX–S1 time delay are used. As discussed above, the inferred median values of H_0 decrease on average with increasing values of $\Delta t_{\text{SX:S1}}$. The results of these tests might also suggest that higher values of Ω_m are preferred by higher values of $\Delta t_{\text{SX:S1}}$. Moreover, the 1σ errors of H_0 and Ω_m decrease mildly with increasing values of the SX–S1 time delay. This reflects the decreasing relative errors of the assumed constant 10 day uncertainty on longer time delays.

The juxtaposition of the observed and model-predicted surface brightness distributions of the SN Refsdal host shown in Figure 5, similarly to the comparison shown in Figure 7 by G16, strongly supports the validity of our results. Over the selected regions, composed of more than 3×10^4 HST pixels, the mean and standard deviation values of the normalized residuals (i.e., the difference between the observed and model-predicted surface brightness values divided by the observational uncertainties) are, respectively, equal to -0.002 and 1.08 . This means that the multiple images of the SN Refsdal host, with their many star-forming regions or knots, are very well reconstructed, both in their positions and fluxes (we note that the 63 point-like images from the 18 knots in the

SN Refsdal host used in the models are reconstructed with a remarkably good accuracy, i.e., with an rms difference between the observed and model-predicted positions of only $\sim 0''.1$). In Figure 6, we show the probability distribution of our normalized residuals and that of a normal distribution with the same standard deviation. We notice that the former has an almost perfectly symmetric shape and we measure that 78.5% (96.5%) of the total number of pixels have normalized residuals smaller than 1 (2), to be compared to the 68.3% (95.5%) probability of a standard normal distribution. The fraction of pixels with normalized residuals larger than 5 is 0.5% ($6 \times 10^{-5}\%$ for a standard normal distribution). We have checked that the vast majority of these pixels are located around the SN host galaxy’s peaks of emission, where the finite resolution of the grid on the source plane, remaining small offsets between the observed and model-predicted image positions, and the so-far neglected contribution of the point-spread function make accurate image reconstructions more difficult to achieve. Moreover, we have measured that in random empty regions of the HST original data, far from the SN Refsdal extended images (not shown in Figure 5), the mean and standard deviation values of the flux are 0.000 and 0.002, respectively. The fractions of pixels in the HST image, shown in the left panel of Figure 5, with flux values larger than 0.002, 0.004, and 0.006 are, respectively, 68.3%, 43.9%, and 25.7%. Therefore, a significant fraction of the selected HST image, which was purposely chosen to include the SN Refsdal host images and avoid contamination from other sources, contains relevant information (i.e., signal). We have repeated the statistical analysis discussed above considering only the pixels with flux values larger than 0.006 and have obtained that 61.3% (90.6%) of the total number of pixels have normalized residuals smaller than 1 (2) and that the fraction of pixels with normalized residuals larger than 5 is 1.7%. These results are not significantly different from the previous ones, when all the pixels were considered, and confirm the goodness of the model on the high signal-to-noise pixels. Considering the relatively simple subtraction of the cluster member flux (that can still be improved), the residuals at the positions of the S1–S4 and SX multiple images of SN Refsdal are remarkably small.

Finally, in Figure 7 we show the inference on the values of the cosmological parameters H_0 , Ω_m , and w . If we consider our reference model (with $\Delta t_{\text{SX:S1}} = 345 \pm 10$ days) and vary at the same time all the strong lensing model parameters, we obtain the following noteworthy 1σ CL constraints: $73.6^{+4.6}_{-4.1}$ km s $^{-1}$ Mpc $^{-1}$ for H_0 , $0.41^{+0.15}_{-0.15}$ for Ω_m , and $-1.07^{+0.22}_{-0.27}$ for w . The achieved precision on the values of Ω_m and w is linked to the combined information coming from the time delays and the positions of the multiple images of sources at different redshifts.

4. Discussion

From the results presented in the previous section, it is clear that the most accurate and precise measurements of the values of the cosmological parameters from MACS 1149 and SN Refsdal will require an additional run of the strong lensing models that include the final measurement of the value and uncertainty of the SX–S1 time delay. Although some scaling relations between the value of $\Delta t_{\text{SX:S1}}$ and those of the cosmological parameters exist (as highlighted in G18 and

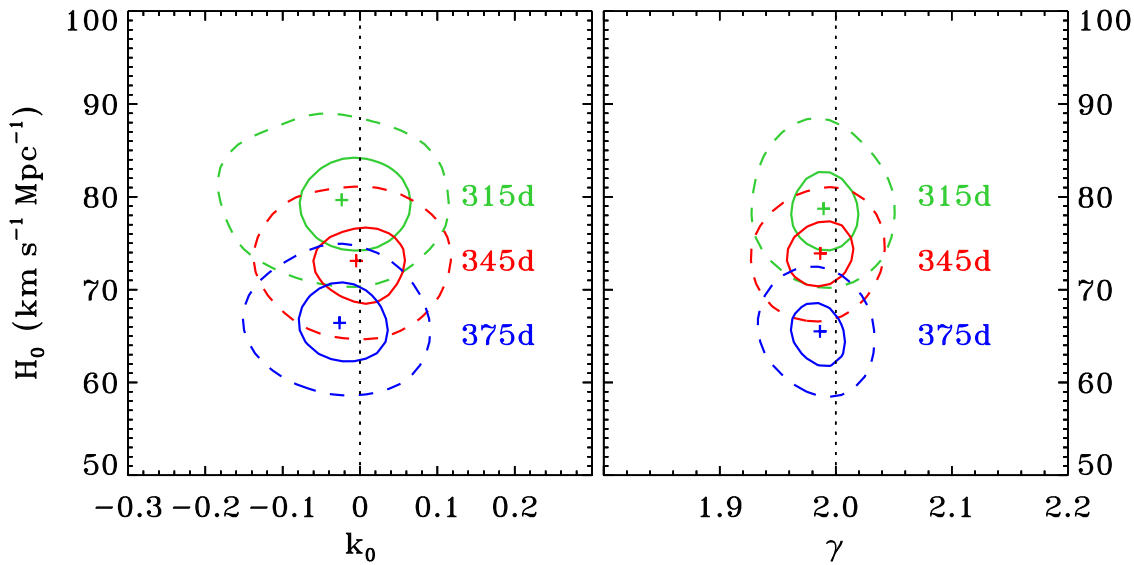


Figure 4. Median values (pluses) and contour levels, at 1 (solid) and 2 (dashed) σ CL, of H_0 , k_0 (value of the convergence of a constant sheet of mass at the cluster redshift) and γ (value of the slope of a power-law profile for the mass density of the cluster main halo) for different (κ and γ) lensing models. Time delays between SX and S1 of 315 ± 10 (green), 345 ± 10 (red), and 375 ± 10 (blue) days are adopted (the time delays of the images S2, S3, and S4, relative to S1, measured by Rodney et al. 2016, are also included). The fixed values of $k_0 = 0$ and $\gamma = 2$ used in the reference (r) models are shown with vertical dotted lines. Flat Λ CDM models ($\Omega_m + \Omega_\Lambda = 1$) with uniform priors on the values of the cosmological parameters ($H_0 \in [20, 120]$ $\text{km s}^{-1} \text{Mpc}^{-1}$ and $\Omega_m \in [0, 1]$) and on the values of $k_0 \in [-0.2, 0.2]$ and $\gamma \in [1.4, 2.6]$ are considered. Final MCMC chains have 2×10^6 samples for each model.

above), the best option to infer the most rigorous values of H_0 and Ω_m and their errors, which aim to be competitive with the most recent ones from other cosmological probes, is definitely not through a simple interpolation of the results already obtained with the current plausible but approximate time delay. We have also demonstrated that a precise measurement of the value of H_0 is not achievable without some information on $\Delta t_{\text{SX:S1}}$ and that the value of Ω_m is mainly constrained by the so-called family ratios of spectroscopically confirmed multiple images of sources located at different redshifts (see the right panel of Figure 2; for more details, see Caminha et al. 2016 and G18). The nonexact linear relation between the values of $\Delta t_{\text{SX:S1}}$ and H_0 can be associated with the fact that (1) in our full strong models the time-delay distance (of the SN Refsdal multiple images) and the family ratios (of the 89 multiple images) are all used at the same time to infer the values of H_0 and Ω_m and of the total mass distribution of the cluster (see G18 for further discussion), and (2) the time delays between the images S1–S4 measured by Rodney et al. (2016) are used in addition to the SX delay relative to S1 as constraints for our model parameters. Furthermore, we have considered an additional set of models with (1) only the time delay between SX and S1 (i.e., excluding those from the other SN Refsdal multiple images), (2) fixed values of all the cosmological parameters, except for that of H_0 , which is free to vary, (3) only the parameter values of the main extended dark-matter halo of the cluster free to vary (to speed up convergence) and have found a linear relation between the values of $\Delta t_{\text{SX:S1}}$ and H_0 .¹²

By comparing different results of the strong lens time delay method, with SN Refsdal in MACS 1149 (as shown here) and with lensed quasars in the galaxy-scale systems of the

H0LiCOW program (Suyu et al. 2017), we can conclude that (1) the relative error on the inferred value of H_0 from a single (galaxy or cluster) strong lensing system is similar (mean value of 6.4% in Figure 2 of Wong et al. 2020), (2) in a single lens cluster, we have the additional interesting possibility of estimating the value of Ω_m (and w), thanks to the observations of several multiple-image families with spectroscopically confirmed redshifts and to the measurement of the time-delay values between the multiple images of time-varying sources, and (3) the observed positions of many spectroscopic multiple images (some of which are crucial to constraining the lens tangential and radial critical curves at different redshifts) provide precise calibrations of all the different mass components (i.e., extended dark-matter halos, cluster members, and possibly hot gas) included in the modeling of a galaxy cluster and, thus, also a good approximation of the effect of the “environment” in the regions adjacent to where the time delays are measured. This is supported by the almost negligible impact of the inclusion in the (κ) models of a constant sheet of mass at the lens redshift (Figure 3). In particular, as shown in Figure 4, the inferred value of k_0 is approximately 0 and not correlated with that of H_0 , whereas this is not the case when only multiple images at a single redshift (i.e., that of SN Refsdal) are used (see the Appendix).

Finally, we notice that relatively large residuals have already been reported in close proximity to the most luminous pixels of multiply imaged QSOs with measured time delays in the image reconstruction of the lensed quasar and host galaxy surface brightness distributions (e.g., Wong et al. 2017). As in the case of SN Refsdal host, this can be ascribed mainly to the finite grid resolution on the source plane. We have checked that by lowering the resolution of the pixel grid on the source plane the distribution of the normalized residuals worsens. We suggest that a comparison between the observed and model-predicted surface brightness distributions of SN Refsdal host, as that shown in Figures 5 and 6, should be used to test the goodness

¹² The median of the absolute values of the differences between the measured values of H_0 and those obtained from the linear fit is less than $0.2 \text{ km s}^{-1} \text{Mpc}^{-1}$. We remark that numerical effects in the final MCMC chains can explain variations on the value of H_0 up to $0.1 \text{ km s}^{-1} \text{Mpc}^{-1}$.

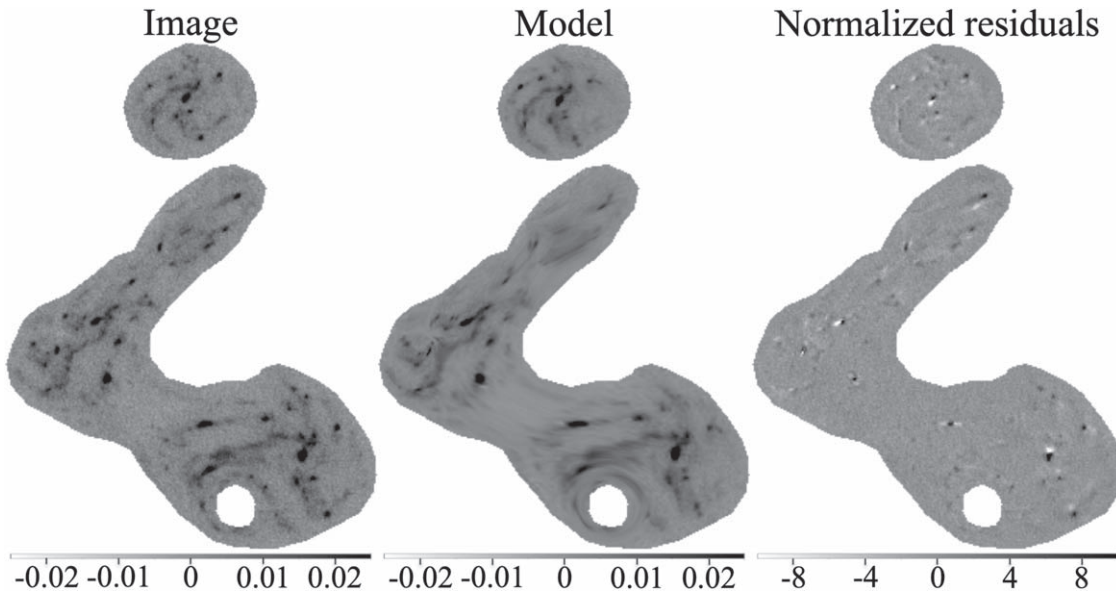


Figure 5. Observed (on the left), model-predicted (in the middle), and normalized residual (on the right) images of the surface brightness distributions of the SN Refsdal host for the reference model with a time delay between SX and S1 of 345 ± 10 days (the time delays of the images S2, S3, and S4, relative to S1, measured by Rodney et al. 2016, are also included). The original data (in units of counts per second) is the linear combination of the HST F606W and F435W final images from the HFF project, optimized to suppress the flux contamination by the cluster member galaxies. The images have a pixel size of $0''.06$. The source surface brightness is reconstructed on a 75×75 pixel grid.

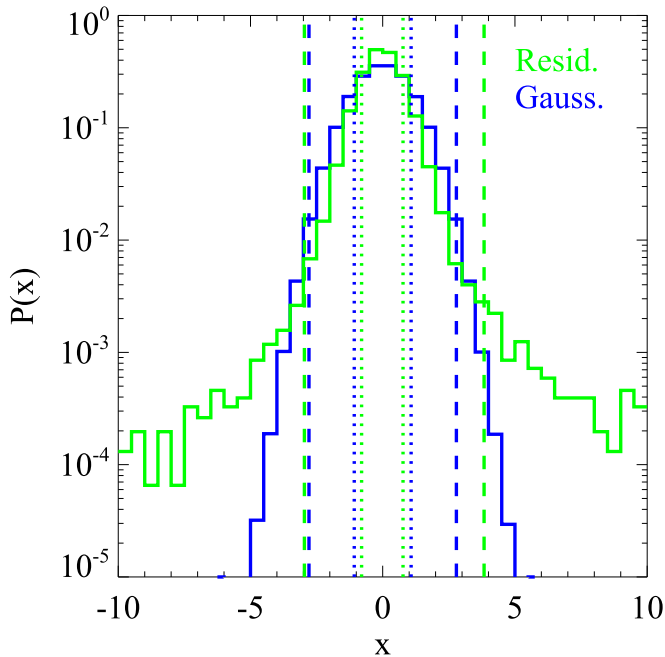


Figure 6. Probability distribution functions of the normalized residuals (green) between the observed and model-predicted surface brightness distributions of the SN Refsdal host for the reference model with a time delay between SX and S1 of 345 ± 10 days (see the right panel in Figure 5) and of a normal distribution (blue) with standard deviation equal to that of the normalized residuals. Vertical dotted and dashed lines represent, respectively, the 68% and 99% probability intervals.

of any strong lensing model which provides estimates of the values of cosmological parameters from the time delays of the SN Refsdal multiple images. More in general, we believe that the extended surface brightness modeling of multiply imaged QSO and SN hosts, although computationally expensive, will help to make progress in time-delay cosmography with lens galaxy clusters, as already demonstrated with lens galaxies.

5. Conclusions

We have shown that accurate and precise estimates of the value of the Hubble constant can be obtained in lens galaxy clusters, when a large set of spectroscopically confirmed multiple images of sources at different redshifts and the time delays of the multiple images of a time-varying source are measured. In flat Λ CDM models, we have demonstrated that a full and blind strong lensing analysis of the HFF galaxy cluster MACS J1149.5+2223, with an error of 10 days (i.e., $\sim 3\%$) on the longest measurable time delay between the multiple images of SN Refsdal provides approximately 6% and 40% total uncertainties for H_0 and Ω_m , respectively. A range of values (315–375 days) for this time delay has been explored, since its accurate measurement has not been published to date and it is not known at the time of writing. We have tested several possible sources of systematic uncertainties and confirmed that their contribution to the estimated values of the cosmological parameters is significantly less important than that of the statistical uncertainties. Specifically, we have found that the impact of the systematic effects on the measurement of the value of the Hubble constant, from the lens models explored here, does not exceed $1 \text{ km s}^{-1} \text{ Mpc}^{-1}$. We have verified that the effect of the “mass-sheet degeneracy” is considerably reduced, thanks to the presence of tens of multiply imaged sources at different distances behind the lens. To illustrate the goodness of our models, we have shown the excellent reconstruction of the extended surface brightness distribution of the SN host, over more than 3×10^4 HST pixels. When the values of H_0 and Ω_m are measured, we remark the importance of considering the full covariance between all the parameters defining the cosmological model and the total mass distribution of the cluster lens. We conclude that time-delay cluster lensing will turn into a new valuable cosmological tool, once more high-quality data of cluster strong lenses become available from the next deep and wide surveys.

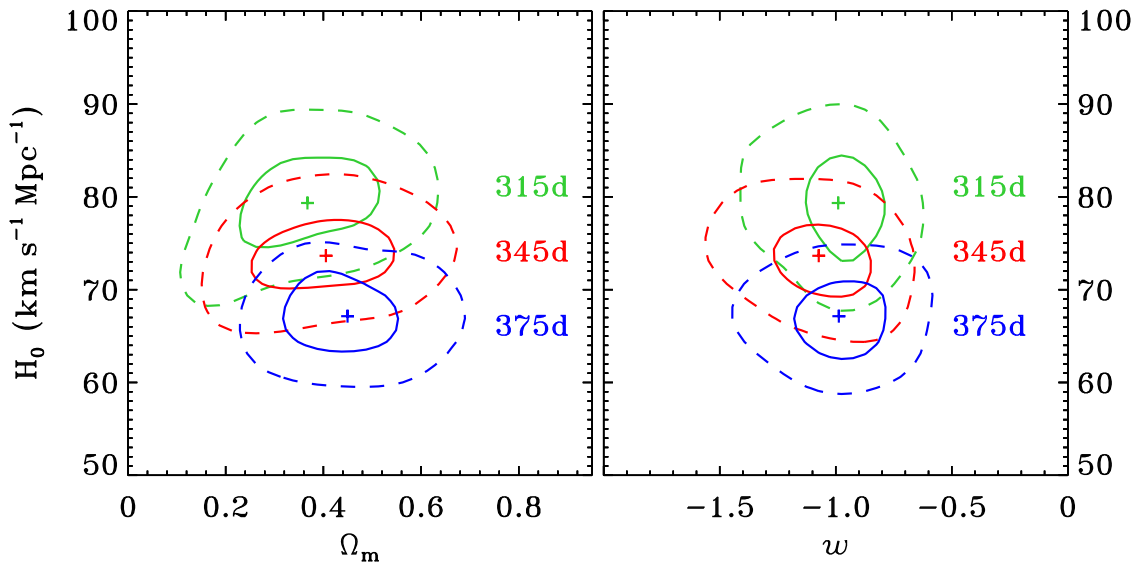


Figure 7. Median values (pluses) and contour levels, at 1 (solid) and 2 (dashed) σ CL, of H_0 , Ω_m , and w for the reference (r) models. Time delays between SX and S1 of 315 ± 10 (green), 345 ± 10 (red), and 375 ± 10 (blue) days are adopted (the time delays of the images S2, S3, and S4, relative to S1, measured by Rodney et al. 2016, are also included). Flat w CDM models ($\Omega_m + \Omega_\Lambda = 1$) with uniform priors on the values of the cosmological parameters ($H_0 \in [20, 120]$ $\text{km s}^{-1} \text{Mpc}^{-1}$, $\Omega_m \in [0, 1]$ and $w \in [-2, 0]$) are considered. Final MCMC chains have more than 1.2×10^6 samples for each model.

C.G. and P.R. acknowledge support through grant MIUR2017 WSCC32 and C.G. also through grant No. 10123 of the VILLUM FONDEN Young Investigator Programme. S.H.S. thanks the Max Planck Society for support through the Max Planck Research Group. A.M. acknowledges funding from the INAF PRIN-SKA 2017 program 1.05.01.88.04. This work is based in large part on data collected in service mode at ESO VLT, through the Director’s Discretionary Time Programme 294.A-5032. The CLASH Multi-Cycle Treasury Program is based on observations made with the NASA/ESA Hubble Space Telescope. The Space Telescope Science Institute is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. ACS was developed under NASA Contract NAS 5-32864.

Appendix The Mass-sheet Degeneracy

We compare here the results of two (κ) models with a constant sheet of mass at the cluster redshift and with $\Delta t_{\text{SX:S1}} = 345 \pm 10$ days, where we consider either the entire sample of 89 multiple images from 28 sources at different redshifts (as presented and discussed above) or only the 63 multiple images from SN Refsdal and its host, all at $z = 1.489$. In Figure A1, we show the inferred values of H_0 and k_0 . We find that the use of many multiple images belonging to sources located at different distances reduces significantly the effect of the so-called “mass-sheet degeneracy” (Falco et al. 1985; Schneider 2019). At the 1σ CL, the credible interval of k_0 varies from $[-0.30, 0.24]$ (extending over a large fraction of the adopted uniform prior, $[-0.5, 0.5]$, chosen here to illustrate better the parameter degeneracy), when sources at a single redshift are included in the model, to $[-0.08, 0.06]$, when all sources at different redshifts are considered. This translates into an approximately 9% difference in the median value of H_0 (and of Ω_m) and in a remarkable reduction by a factor of more than 3, from $\sim 21\%$ to $\sim 6\%$, for its uncertainty (from $\sim 63\%$ to $\sim 40\%$ for the uncertainty on Ω_m).

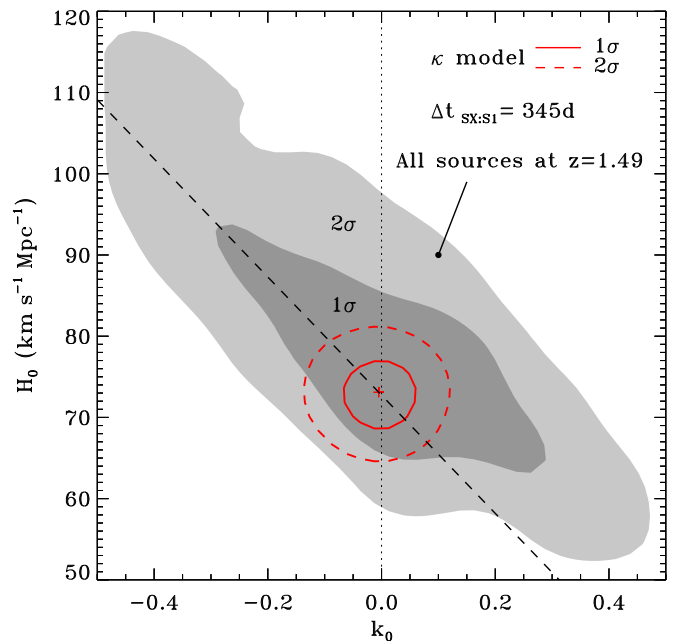


Figure A1. Median values (plus) and contour levels, at 1 and 2σ CL, of H_0 and k_0 for the κ lensing models with all (89) multiple images (red) and with only those (63) belonging to SN Refsdal and its host (gray). A time delay between SX and S1 of 345 ± 10 days is adopted (the time delays of the images S2, S3, and S4, relative to S1, measured by Rodney et al. 2016, are also included). The fixed value of $k_0 = 0$ used in the reference (r) models is shown with the vertical dotted line. The dashed line illustrates the theoretical effect of the so-called “mass-sheet degeneracy” (Schneider & Sluse 2013). Flat Λ CDM models ($\Omega_m + \Omega_\Lambda = 1$) with uniform priors on the values of the cosmological parameters ($H_0 \in [20, 120]$ $\text{km s}^{-1} \text{Mpc}^{-1}$ and $\Omega_m \in [0, 1]$) and on the value of $k_0 \in [-0.2, 0.2]$ or $[-0.5, 0.5]$ are considered. Final MCMC chains have 2×10^6 samples for each model.

ORCID iDs

C. Grillo <https://orcid.org/0000-0002-5926-7143>

A. Halkola <https://orcid.org/0000-0001-5001-8020>

References

- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, *Natur*, **551**, 85
- Bergamini, P., Rosati, P., Mercurio, A., et al. 2019, *A&A*, **631**, 130
- Caminha, G. B., Grillo, C., Rosati, P., et al. 2016, *A&A*, **587**, 80
- Chen, G. C.-F., Fassnacht, C. D., Suyu, S. H., et al. 2019, *MNRAS*, **490**, 1743
- Falco, E. E., Gorenstein, M. V., & Shapiro, I. I. 1985, *ApJ*, **289**, 1
- Freedman, W. L., Madore, B. F., Hatt, D., et al. 2019, *ApJ*, **882**, 34
- Grillo, C., Karman, W., Suyu, S. H., et al. 2016, *ApJ*, **822**, 78, (G16)
- Grillo, C., Rosati, P., Suyu, S. H., et al. 2018, *ApJ*, **860**, 94, (G18)
- Hubble, E. 1929, *PNAS*, **15**, 168
- Kelly, P. L., Brammer, G., Selsing, J., et al. 2016a, *ApJ*, **831**, 205
- Kelly, P. L., Rodney, S. A., Treu, T., et al. 2015, *Sci*, **347**, 1123
- Kelly, P. L., Rodney, S. A., Treu, T., et al. 2016b, *ApJ*, **819**, 8
- Koekemoer, A. M., Fruchter, A. S., Hook, R., & Hack, W. 2002, in The 2002 HST Calibration Workshop: Hubble after the Installation of the ACS and the NICMOS Cooling System, ed. S. Arribas, A. Koekemoer, & B. Whitmore (Baltimore, MD: STScI), 337
- Lemaître, G. 1927, *ASSB*, **47**, 49
- Planck Collaboration VI 2018, arXiv:1807.06209
- Refsdal, S. 1964, *MNRAS*, **128**, 307
- Reid, M. J., Braatz, J. A., Condon, J. J., et al. 2013, *ApJ*, **767**, 154
- Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019, *ApJ*, **876**, 85
- Rodney, S. A., Strolger, L.-G., Kelly, P. L., et al. 2016, *ApJ*, **820**, 50
- Schneider, P. 2019, *A&A*, **624**, 54
- Schneider, P., & Sluse, D. 2013, *A&A*, **559**, 37
- Shajib, A. J., Birrer, S., Treu, T., et al. 2020, *MNRAS*, **494**, 6072
- Suyu, S. H., Bonvin, V., Courbin, F., et al. 2017, *MNRAS*, **468**, 2590
- Suyu, S. H., & Halkola, A. 2010, *A&A*, **524**, A94
- Suyu, S. H., Hensel, S. W., McKean, J. P., et al. 2012, *ApJ*, **750**, 10
- Treu, T., Brammer, G. B., Diego, J. M., et al. 2016, *ApJ*, **817**, 60
- Vega-Ferrero, J., Diego, J. M., Miranda, V., & Bernstein, G. M. 2018, *ApJ*, **853**, 31
- Williams, L. L. R., & Liesenborgs, J. 2019, *MNRAS*, **482**, 5666
- Wong, K. C., Suyu, S. H., Auger, M. W., et al. 2017, *MNRAS*, **465**, 4895
- Wong, K. C., Suyu, S. H., Chen, G. C.-F., et al. 2020, *MNRAS*, in press