A new high-precision strong lensing model of the galaxy cluster
MACS J0416.1−2403

Robust characterization of the cluster mass distribution from VLT/MUSE deep
observations

P. Bergamini1,⋆,⋆⋆, P. Rosati3,1, E. Vanzella1, G. B. Caminha2, C. Grillo5,6, A. Mercurio4, M. Meneghetti1,
G. Angora3,4, F. Calura1, M. Nonino7, and P. Tozzi8

1 INAF – OAS, Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Gobetti 93/3, I-40129 Bologna, Italy
2 Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, The Netherlands
3 Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, via Saragat 1, I-44122 Ferrara, Italy
4 INAF – Osservatorio Astronomico di Capodimonte, Via Moiariello 16, I-80131 Napoli, Italy
5 Dipartimento di Fisica, Università degli Studi di Milano, via Celoria 16, I-20133 Milano, Italy
6 Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200 Copenhagen, Denmark
7 INAF – Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143, Trieste, Italy
8 INAF – Osservatorio Astrofisico di Arcetri, Largo E. Fermi, I-50125, Firenze, Italy

Received February 14, 2020; accepted February 14, 2020

ABSTRACT

We present a new high-precision parametric strong lensing model of the galaxy cluster MACS J0416.1−2403, at $z = 0.396$, which takes advantage of the MUSE Deep Lensed Field (MDLF, Vanzella et al. 2020b), with 17.1h integration in the North-East region of the cluster, and Hubble Frontier Fields data. We spectroscopically identify 182 multiple images from 48 background sources at $0.9 < z < 6.2$, and 171 cluster member galaxies. Several multiple images are associated to individual clumps in multiply lensed resolved sources. By defining a new multiplet, which is sensitive to the gradients of the deflection field, we show that we can accurately reproduce the positions of these star-forming knots despite their vicinity to the model critical lines. The high signal-to-noise spectra of the MDLF enables the measurement of the internal velocity dispersion of 64 cluster galaxies, down to $m_{F606W} = 22$. This allows us to independently estimate the contribution of the subhalo mass component of the lens model from the measured Faber-Jackson scaling relation. Our best reference model, which represents a significant step forward compared to our previous analyses, is selected from a comparative study of different mass parametrizations. The r.m.s. displacement between the observed and model-predicted image positions is only $0.40''$, 33% smaller than in all previous models. The mass model appears to be particularly well constrained in the MDLF region. We characterize the robustness of the magnification map at varying distances from the model critical lines and the total projected mass profile of the cluster.

Key words. Galaxies: clusters: general – Gravitational lensing: strong – cosmology: observations – dark matter – galaxies: kinematics and dynamics

1. Introduction

In recent years, strong gravitational lensing has become one of the most effective techniques to characterize the total mass distribution in the inner regions of galaxy clusters, where multiple images of background sources are formed and provide crucial constraints on lens models. The total mass profile of galaxy clusters from kpc up to Mpc scales can be derived by combining strong lensing with other mass tracers, such as weak lensing (e.g. Umetsu et al. 2014; Hoekstra et al. 2015; Melchior et al. 2015), dynamical methods using inner stellar kinematics of the central Brightest Cluster Galaxy (BCG, Sartoris et al. 2020) and the phase space of cluster galaxies (Biviano et al. 2013; Stock et al. 2015), as well as X-ray hydro-static analysis (e.g. Ettori et al. 2013).

By comparing the observed mass density profiles with those predicted by N-body and hydrodynamical simulations one can test the ΛCDM paradigm of structure formation (e.g. Merten et al. 2015). In particular, a robust characterization of the mass distribution in the cores of galaxy clusters, separating the baryonic and the dark matter (DM) components, can reveal missing physical ingredients in cosmological simulations or possibly constrain DM physical properties (e.g. Newman et al. 2013; Grillo et al. 2015; Natarajan et al. 2017; Annunziatella et al. 2017; Bonamigo et al. 2018; Meneghetti et al. 2020).

High-precision strong lensing models can be effectively used to study the mass distribution of cluster substructures (or subhalos), particularly when internal kinematics of cluster galaxies is taken into account (see Bergamini et al. 2019, hereafter B19). A recent study by Meneghetti et al. (2020) finds an inconsistency between the number of observed Galaxy-Scale Strong Lensing Systems (GSSLS) in massive clusters with those predicted by state-of-the-art cosmological simulations. This study relies on
high-precision strong lensing models, such as the one presented here, to dissect the subhalo population from the total projected mass distribution of clusters.

In addition to total mass distributions, magnification maps obtained from strong lensing models of galaxy clusters are fundamental tools to study the astrophysical properties of lenses and highly magnified background sources. Specifically, the lensing magnification allows one to resolve details on scales of a few tens of parsecs at redshift $2–6$ (see Vanzella et al. 2017a,b; Johnson et al. 2017; Rigby et al. 2017; Cava et al. 2018; Vanzella et al. 2019, 2020b). Thus, until the advent of new observational facilities, such as the James Webb Space Telescope (JWST) and the Extremely Large Telescope (ELT), strong lensing represents the only available technique to characterize the high redshift population of faint sources, i.e. the progenitors of present day galaxies and globular clusters, which are expected to play an important role in the re-ionization process of the Universe (e.g. Boylan-Kolchin 2018; Ma et al. 2020; He et al. 2020; Vanzella et al. 2020a).

Over the last decade, we have witnessed important progress in cluster lens models thanks to several observational campaigns that have provided high-quality photometric and spectroscopic data on a sizable sample of massive galaxy clusters, which are generally powerful gravitational lenses. The first of these campaigns, the Cluster Lensing And Supernova survey with Hubble (CLASH, Postman et al. 2012) provided panchromatic data on 25 massive clusters with the WFC3 and ACS cameras on Hubble Space Telescope (HST). The Re-ionization Lensing Cluster Survey (RELICS, Coe et al. 2019) extended this HST study to 41 clusters. Significantly deeper HST observations, in seven clusters in the Re-ionization Lensing Cluster Survey (RELICS, Coe et al. 2019) extended this HST study to 41 clusters. Significantly deeper HST observations, in seven clusters in the Re-ionization Lensing Cluster Survey (RELICS, Coe et al. 2019) extended this HST study to 41 clusters. Significantly deeper HST observations, in seven clusters in the Re-ionization Lensing Cluster Survey (RELICS, Coe et al. 2019) extended this HST study to 41 clusters. Significantly deeper HST observations, in seven clusters in the Re-ionization Lensing Cluster Survey (RELICS, Coe et al. 2019) extended this HST study to 41 clusters.

In addition to the HST imaging campaigns, spectroscopic follow-up programs were carried out on a sub-sample of clusters. The CLASH-VLT Large Programme (Rosati et al. in prep.), for example, collected $\sim 30000$ redshifts in the fields of 13 CLASH clusters with the VLT/VIMOS multi-slit spectrograph. Approximately two hundred lensed background sources and multiple images were also identified. The CLASH-VLT spectroscopic campaign on MACS 0416 was presented in Balestra et al. (2016). The ability to identify large number of multiple images in the core of galaxy clusters improved dramatically with the advent of the the Multi Unit Spectroscopic Explorer (MUSE) on the VLT (Bacon et al. 2012, see e.g. Caminha et al. 2017; Lagattuta et al. 2017).

The accuracy of strong lensing modeling relies critically on the number of spectroscopically confirmed multiple images (Grillo et al. 2015; Johnson et al. 2014; Caminha et al. 2019). In fact, spectroscopic information reduces the risk of multiple image misidentifications and breaks the degeneracy between source distances and the total mass of the lens. In addition, integral field spectroscopy allows the identification of a complete and pure sample of cluster galaxies which are another crucial ingredient for lens models. As demonstrated in B19, MUSE data can also be used to measure the inner stellar kinematics of numerous cluster member galaxies, which can be incorporated into lens models, thus adding an important constraint on the mass of the subhalo population (see also Monna et al. 2017).

The dynamical analysis of MACS 0416 by Balestra et al. (2016), using ~900 spectroscopic members, combined with the Chandra X-ray and VLA radio observations, revealed a complex, mostly bi-modal mass distribution, which is likely the result of a pre-merging phase. The 30 spectroscopically confirmed multiple images identified in the CLASH-VLT campaign were used in the lens model of Grillo et al. (2015), thus improving on previous photometric studies (Zitrin et al. 2013). A free-form lens model was presented by Hoag et al. (2016), using an extended spectroscopic coverage of MACS 0416 based on the GLASS survey (Treu et al. 2015). A significant step forward was then made by Caminha et al. (2017) (hereafter C17) by exploiting the first MUSE observations of MACS 0416. This study presented a high-precision lens model based on a new large sample of 102 spectroscopic multiple images (from 37 background sources) and an extended catalog of cluster galaxies (including 144 spectroscopic members). In Bonamigo et al. (2018), we refined the C17 lens model by including the mass component associated to the hot-gas, which is traced by deep Chandra X-ray observations (Ogrean et al. 2015) and dominates the cluster baryonic content.

In this paper, we further improve the lens model of MACS 0416 by exploiting additional MUSE observations in the North-East part of the cluster, the MUSE Deep Lens Field (MDLF), which is presented in a accompanying paper by Vanzella et al. (2020b). When compared to our previous models (C17 and B19), the combination of the MDLF and the HFF data lead us to identify ~80% more multiple images, ~20% more spectroscopic galaxies and to measure the inner velocity dispersion of cluster members one magnitude fainter.\(^1\)

The paper is organized as follows. In Sec. 2, we present the MACS 0416 dataset used to develop our new lens model. Sec. 3 describes in detail the mass parametrization of a selected number of lens models. In Sec. 4, we discuss the results of the optimization of four lens models and the criteria leading to the selection of the best reference model. The latter is used to study the robustness of the magnification values of multiple images and to characterize the projected total mass profile of MACS 0416. The main conclusions of our work are drawn in Sec. 5.

Throughout this work, we adopt a flat $\Lambda$CDM cosmology with $\Omega_m=0.3$ and $H_0=70\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}$. Using this cosmology, a projected distance of $1\arcsec$ corresponds to a physical scale of 5.34 kpc at the MACS 0416 redshift of $z=0.396$. All magnitudes are given in the AB system.

# 2. Data

In this section, we briefly describe the MACS 0416 data-set presented in Vanzella et al. (2020b), specifically the new catalogs of multiple images and cluster member galaxies used in our lens models.

## 2.1. HST photometric and MUSE spectroscopic observations of MACS 0416

This work is based on the HST multi-band imaging data from the CLASH survey and the HFF program. MACS 0416 has been the target of several spectroscopic follow-up campaigns. The redshift measurements over a wide ($\sim 20\, \text{arcmin}$ across) field were

\(^1\) When this paper was close to be submitted, Richard et al. (2020) presented the results from a lens model of MACS 0416 based on the new MUSE deep observations. A comparison between the two models will be possible when their model details will be made public.
obtained as part of the CLASH-VLT program with VLT/VIMOS and presented in Balestra et al. (2016). Grillo et al. (2015) constructed a lens model using 30 spectroscopically confirmed multiple images from this initial spectroscopic dataset. The number of multiple images with a spectroscopic confirmation increased by more than a factor of three with the advent of MUSE observations, as described by C17. MUSE data-cubes have a field-of-view of 1 arcmin$^2$ that is spatially sampled with 0.2′′ × 0.2′′ pixels. The wavelength range extends from 4700 Å to 9350 Å with a dispersion of 1.25 Å/pix, and a spectral resolution of ∼ 2.6 Å fairly constant across the entire spectral range.

The study of C17 included two MUSE pointings, one to the North-East (NE) of 2h exposure (0.6′′ seeing) and one to the South-West (SW) of 11h integration (1.0′′ seeing), the latter however provided spectra with S/N below expectations (see comment in that paper). In this work, we take advantage of the MDLF dataset, which extended the integration time in most of the NE field to a total of 17.1h, with a final seeing of 0.6′′ (Vanzella et al. 2020b). In Fig. 1, we show the footprints of the three MUSE pointings overlaid onto the HST/RGB image of the cluster.

### 2.2. Catalog of multiple images

The latest public catalog of spectroscopically confirmed multiple images in MACS0416, before this work, was released by C17. It included a total of 102 images from 37 background sources. The same catalog was used in the lens models presented by Bonamigo et al. (2018) and B19. As described in Vanzella et al. (2020b), the combined analysis of the MDLF observations and the CLASH/HFF multi-band images has led to the identification of 182 secure multiple images, at 0.9 < z < 6.2, which are used in our new lens model.

We assign to each multiple image an ID containing a number and a letter so that images with the same number but different letters belong to the same family of multiple images. In MACS 0416, we find eight resolved lensed sources, each containing two or more point-like knots (or clumps) at the same redshift, which are often embedded into a diffuse light emission. Multiple images of the same clump form a family, while we call the set of images coming from the same resolved source, a ‘system’ of multiple images. Thus, the 182 multiple images are found associated to 66 independent families, drawn from 48 different background sources. Images belonging to a system are characterized by an identical integer part of their ID number, but by a different fractional part. Our analysis demonstrates that these multiply imaged clumps are very efficient in constraining the position of the critical lines of the lens models (e.g. see Sys-12 discussed below).

Two examples of systems of multiple images are given in the cut-outs of Fig. 2. Sys-12 is made by six image families, with IDs from 12.1 to 12.6, containing two multiple images (b, c) of the same clumps at z = 0.940. Instead, Sys-5 is formed by six families (from 5.1 to 5.6) of three multiple images each (a, b, c). These resolved sources are indicated with rectangles in Fig. 2, with a different color highlighting each of the eight systems. Red crosses correspond to multiple images belonging the previous catalog by C17, while green crosses are the newly discovered images. Note that two images of the same background source at z = 3.923 (identified as 22b and 22c in the C17) have been excluded from our final catalog since the elongation of their Lyα emission does not allow a precise determination of their positions. The Northern region of the cluster includes 80 out of 82 of the new multiple images, where the deepest MUSE pointing was carried out. The remaining two images, identified as 20.3a and 20.3c, are new clumps identified in the source at z = 3.222 in the SW field.

In Fig. 3 (lower panel), we show in gray the redshift distribution of the new set of 182 multiple images, ranging from z = 0.940 to z = 6.145, while the blue histogram refers to the previous sample published in C17. The 125 images in the deep NE region are drawn in red.

### 2.3. Cluster members selection and measured stellar velocity dispersions

Exploiting the high signal-to-noise of the cluster member spectra extracted from the MDLF pointing in the NE region of MACS 0416, we extend the publicly released cluster member catalog used in the lens model of C17. The latter contained 144 spectroscopically confirmed cluster members, which are defined as those galaxies in the HST/WFC3 field-of-view, brighter than $m_{F160W} = 24$, and with a velocity within ±3000 km s$^{-1}$ in cluster rest frame centered at $z = 0.396$ (this corresponds to the redshift range [0.382-0.410]). To maximize sample completeness, this sample was complemented with 49 photometric members selected from their color space distribution based on 12 CLASH photometric bands, with $m_{F160W} < 24$, as described in Grillo et al. 2015.
Fig. 2. HST RGB image (F814W, F606W, F435W) of the central region of the galaxy cluster MACS 0416 at $z = 0.396$. The colored crosses mark the position of the 182 spectroscopic multiple images used to constrain the lens models: red crosses mark the 100 images in common with the catalog by C17, while the newly identified images are in green. The two BCGs of the cluster and the foreground galaxy at $z = 0.112$ are encircled in white. Colored rectangles highlight the systems of multiple images produced by eight background sources resolved into multiple clumps (see Sec. 2.2). The dotted cyan circle marks the position of the third predicted image of Sys-12. The bottom panels show three zoom-in images of systems 12 and 5, obtained from a median stack of the F814W, F606W, and F435W HST filters. The red and blue circles correspond to the red and green crosses in the main panel.

The revised catalog of cluster members, to be used for the new lens models, includes 19 additional spectroscopic members based on MDLF observations. Four of these galaxies are fainter than $m_{F160W} = 24$, they are however included in the sample due to their secure membership. We also add to this catalog seven extra spectroscopic members based on new HFF/F160W pho-
ometry, which is needed for the lens model. Moreover, we add to the final sample a bright galaxy (\(m_{F160W}=18.11\)) at \(z = 0.4111\) whose peculiar velocity is only \(245 \text{ km s}^{-1}\) above the upper bound of the velocity range chosen for the cluster member selection. Due to its high F160W luminosity, we expect this galaxy to have a non-negligible impact on the cluster mass model. We remove a faint cluster member (ID Gal-739 in C17) since it is below the magnitude limit based on revised photometry.

Summarizing, our final cluster member catalog counts a total of 213 galaxies, 171 have a secure spectroscopic redshift, while 42 members are still photometric members. We note that a recent method based on a convolution neural network technique, which was developed to identify cluster members using multiband HST image cutouts (Angora et al. 2020), confirms 40 out of the 42 photometric members included in this work.

Following the study by B19, we measure the line-of-sight stellar velocity dispersion, \(\sigma_{ap}^{gal}\), of a large number of cluster galaxies by taking advantage of the increased mean Signal-to-Noise \((S/N)\) of galaxy spectra in the MDLF (a factor \(\sim 2\)). We use spectral extraction apertures of 0.8ag radius, which yields the best compromise between high \((S/N)\) and low contamination from nearby bright sources. Velocity dispersions are measured by cross-correlating the observed galaxy spectra with a set of stellar templates, using the 02/2018 version of the public software pPXF (Cappellari & Emsellem 2004; Cappellari 2017), as discussed in B19. pPXF fits are performed over the wavelength range 5000-7260 Å to avoid the red part of the MUSE spectra, which is contaminated by skyline residuals, and at the same time includes all relevant galaxy absorption lines. In addition to regions containing sky emission lines, masking is applied to the wavelength range 5800-6000 Å, which is highly contaminated by the sodium laser emission of the adaptive optics system used in most of the MDLF observations.

As explained in B19, to ensure robust velocity dispersion measurements we limit our sample of galaxies with measured stellar kinematics to \((S/N)> 10\) and \(\sigma_{ap}^{gal}> 60 \text{ km s}^{-1}\). We also apply small corrections to \(\sigma_{ap}^{gal}\) and its uncertainties, \(d\sigma_{ap}^{gal}\), based on spectral simulations (see appendix A in B19). The final sample of cluster members with internal kinematics includes 64 galaxies. We note that the MDLF pointing allows us to measure reliable velocity dispersions for 15 additional galaxies down to \(m_{F160W} \sim 22\), i.e. approximately a magnitude fainter than the limit adopted in B19. The sample of cluster members with measured internal kinematics is indicated in Fig. 1 (green circles). In Fig. 4, we show the measured \(\sigma_{ap}^{gal}\) values as a function of magnitude in the F160W filter.

3. Strong lensing models

To model the mass distribution of MACS 0416 we use the public software LensTool (Kneib et al. 1996; Jullo et al. 2007; Jullo & Kneib 2009) for strong lensing modeling. Following the method...
adopted in our previous studies, we decompose the cluster total mass into several mass components, based on simple parametric models. A set of free-parameters $\xi$ is determined by constraining the mass models with the positions of the 182 multiple images, thereby maximizing the following posterior probability distribution function (see Caminha et al. 2016):

$$P(\mathbf{x}^{\text{obs}}|\xi) \propto P(\xi|\mathbf{x}^{\text{obs}}) \times P(\xi).$$  \hspace{1cm} (1)$$

where $P(\xi)$ correspond to prior probability distributions for the model free-parameters, while the likelihood $P(\xi|\mathbf{x}^{\text{obs}})$ is given by:

$$L \equiv P(\xi|\mathbf{x}^{\text{obs}}) \propto \exp\left(-\frac{1}{2} \chi^2(\xi)\right).$$  \hspace{1cm} (2)$$

The lens model $\chi^2$ quantifies the displacement on the lens plane between the observed and model-predicted positions of the multiple images $(\mathbf{x}^{\text{pred}})$, given the set of model parameters $\xi$. To compute the $\chi^2$ value an isotropic uncertainty, $\Delta x_{i,j}$, on the observed positions of the images is assumed (Jullo et al. 2007):

$$\chi^2(\xi) := \sum_{i=1}^{N_{\text{con}}} \sum_{j=1}^{N_{\text{pm}}} \left(\frac{x_{i,j}^{\text{obs}} - x_{i,j}^{\text{pred}}}{\Delta x_{i,j}}\right)^2.$$  \hspace{1cm} (3)$$

The index $i$ runs over the multiple images belonging to the same $j$-th family (e.g. $1a$, $1b$, $1c$ for system 1), while the index $j$ runs over all families (in our case $N_{\text{family}} = 66$). So $N_{\text{pm}}$ is the number of observed multiple images coming from the same $j$-th family (e.g. $N_{\text{pm}} = 3$).

Since the lens model is constrained by the position of $N_{\text{family}} = \sum_{j=1}^{N_{\text{pm}}} N_{\text{pm}}^j$ observed multiple images, by defining $N_{\text{pm}}$ as the total number of model free-parameters, we can write the number of degree-of-freedom (DoF) of the lens model as:

$$\text{DoF} = 2 \times N_{\text{pm}}^\text{tot} - 2 \times N_{\text{family}} - N_{\text{param}} = N_{\text{con}} - N_{\text{param}}.$$  \hspace{1cm} (4)$$

The term $2 \times N_{\text{family}}$ stems from the fact that the unknown positions of the $N_{\text{family}}$ background sources are additional free parameters of the model. Thus, $N_{\text{con}}$ is the effective number of available constraints.

To sample the lens model posterior distribution in Eq.(2), LensTool exploits a Bayesian Markov chain Monte Carlo (MCMC) technique. After the removal of a large burn-in phase, the final MCMC chains of our lens models contain at least 10^5 steps. The uncertainties on the model free-parameters are determined by re-sampling the posterior distribution once the error on the positions of the observed multiple images ($\Delta x_{i,j}$) is re-scaled so that the reduced $\chi^2$ is close to 1 (the initial value $\Delta x_{i,j}$ is set to 0.5" (1.0" for HST (MUSE)) detected images).

To quantify the goodness of our lens models, we use three main indicators. The first, as customary, is the root-mean-square separation between the observed and model-predicted positions of multiple images, $\Delta_{\text{rms}}$ (see e.g. Caminha et al. 2016):

$$\Delta_{\text{rms}} = \sqrt{\frac{1}{N_{\text{pm}}^\text{im}} \sum_{i=1}^{N_{\text{pm}}^\text{im}} |\Delta i|^2}, \quad \text{with } \Delta i = x_{i}^{\text{obs}} - x_{i}^{\text{pred}},$$  \hspace{1cm} (5)$$

where $\Delta i$ is the displacement between the $i$-th observed and predicted image. The second and the third indicators are the Bayesian Information Criterion (BIC, Schwarz 1978), and the Akaike Information Criterion (AIC, Akaike 1974) defined as:

$$\text{BIC} = -2 \ln(L_{\text{max}}) + N_{\text{param}} \ln(N_{\text{con}}),$$

$$\text{AIC} = -2 \ln(L_{\text{max}}) + 2 N_{\text{param}}.$$  \hspace{1cm} (6)$$

$L_{\text{max}}$ is the maximum value of the likelihood in Eq. 2.

These information criteria ensure that the introduction of extra free parameters in a lens model is justified by a corresponding increase of the model likelihood, thus avoiding over-fitting. As a general rule of thumb, the best cluster lens model has the lowest $\Delta_{\text{rms}}$, BIC and AIC values.

3.1. Mass components

The overall total mass distribution of MACS 0416 (or equally its total gravitational potential, $\phi_{\text{tot}}$) is divided into the following sum of parametric mass profiles:

$$\phi_{\text{tot}} = \sum_{i=1}^{N_{\text{halo}}} \phi_{i}^{\text{halo}} + \sum_{j=1}^{N_{\text{gal}}^{\text{star}}} \phi_{j}^{\text{gal}} + \phi_{\text{foreg}} + \phi_{\text{extra}}.$$  \hspace{1cm} (7)$$

The first term takes into account the $N_{\text{halo}}$ cluster-scale smooth halos of the cluster potential ($\phi_{\text{halo}}$), while the second one is associated to the $N_{\text{gal}}$ cluster member galaxies (or subhalos), each with gravitational potentials $\phi_{j}^{\text{gal}}$. The third term, $\phi_{\text{foreg}}$, is the contribution from a prominent foreground galaxy residing in the South-West region of the cluster. The last term, $\phi_{\text{extra}}$, refers to a possible constant convergence, or shear, associated to extra mass unaccounted for in the cluster field.

In our study, we have explored a large number of lens models with different mass parametrizations and number of subcomponents. In the following two subsections, we describe the best four lens models selected on the basis of the aforementioned criteria, which are referred to as LM-4HALOS, LM-BCGs, LM-HLBBCGs, and LM-SHEAR. In Sec. 4, we explain why the LM-4HALOS model is finally adopted as our reference model.

3.1.1. Cluster-scale mass distribution

Most of the total mass of galaxy clusters is in the form of smooth halos that extend over a scale of hundreds to thousands of kpc. These cluster-scale halos are dominated by the DM component, with a non-negligible fraction of hot-gas and stars responsible for the intra-cluster light (ICL) emission.

The cluster-scale component of our LensTool models is parametrized as a sum of elliptical dual Pseudo-isothermal mass distributions (dPIE, Eliasdóttir et al. 2007; Limousin et al. 2005; Bergamin et al. 2019). The set of free-parameters includes: two parameters for their sky position, two for the ellipticity $e = a^2-b^2/a^2+b^2$ (where $a$ and $b$ are the semi-major and semi-minor axis of the dPIE profile) and the position angle ($\theta$) measured counterclockwise from the west direction, additional three parameters are the central velocity dispersion $\sigma$, the core radius $r_{\text{core}}$, and the truncation radius $r_{\text{tr}}$.

All the lens models we have developed contain four cluster-scale dPIEs that are used to parameterize the hot-gas mass content, obtained by fitting the Chandra deep X-ray surface brightness distribution (see Bonamigo et al. 2018). These are fixed profiles and therefore do not contribute to the number of
Table 1. Main parameters of the scaling relations $\sigma_{gal}^\text{ref}$- $m_{F160W}^\text{cut}$ and $\sigma_{ap}$- $m_{F160W}$.

| N($\sigma_{gal}^\text{ref}$) | $m_{F160W}^\text{ref}$ | $\sigma_{gal}^\text{ref}$ [km s$^{-1}$] | $\alpha$ | $\Delta\sigma_{ap}$ [km s$^{-1}$] | $\beta_{\text{cut}}(y = 0.2)$ |
|-----------------------------|----------------------|-----------------------------|---------|-----------------------------|-----------------|
| 64                          | 17.02                | 295.2$^{+17.3}_{-16.6}$    | 0.30$^{+0.03}_{-0.03}$ | 33.1$^{+3.4}_{-2.9}$ | 0.60$^{+0.06}_{-0.06}$ |

Each subhalo is parametrized as a circular dPIE profile, with negligible core radius, whose position is centered on the peak of the stellar light emission. As customary, to reduce the number of free-parameters, we adopt the following scaling relations for the central velocity dispersions ($\sigma_{0}^{\text{gal}}$) and truncation radii ($r_{\text{cut}}^{\text{gal}}$), as a function of galaxy luminosity:

$$
\sigma_{LT,0}^{\text{gal}} = \sigma_{LT} L_{0}^{\alpha},
$$

(8)

$$
r_{\text{cut}}^{\text{gal}} = r_{\text{cut}}^{\text{ref}} L_{0}^{\beta_{\text{cut}}},
$$

(9)

In Eq. 8, we introduce the LensTool fiducial velocity dispersion $\sigma_{LT}$ that is related to the central velocity dispersion of the dPIE by $\sigma_{0} = \sqrt{\pi} \sigma_{LT}$. The two normalizations $\sigma_{LT}^{0}$ and $r_{\text{cut}}^{\text{ref}}$ are computed at the reference luminosity $L_{0}$. For the $L_{0}$ luminosities, we use F160W Kron magnitudes of cluster galaxies, as a good proxy of their total mass (Grillo et al. 2015). The BCG-N magnitude, $m_{F160W} = 17.02$, is used for the values of $L_{0}$ in Eqs. 8 and 9. A third scaling relation with slope $\beta_{\text{core}} = 0.5$ is implemented in LensTool to scale the dPIE core radius. However, since we assume a negligible value for the reference core radius ($r_{\text{core}}^{\text{ref}} = 1'' \times 10^{-4} = 0.5$ pc), the very inner region of the galactic dPIEs is well approximated by a singular isothermal behavior.

Based on the Bayesian technique described in B19, we fit the slope $\alpha$ and the normalization $\sigma_{LT}^{\text{gal}}$ of the $\sigma_{gal}^{\text{cut}}$-$m_{F160W}$ relation, i.e. the Faber-Jackson relation in Eq. 8 (Faber & Jackson 1976), using the 64 stellar velocity dispersions $\sigma_{gal}^{\text{cut}}$ (see Sec. 2.3). The solid green line in Fig. 4 corresponds to the best-fit scaling relation, while the light-green band represents the mean scatter ($\Delta\sigma_{ap}$) of the relation, which is a free parameter in the fit (see posterior distributions in Fig. 5).

Following B19, by assuming a fixed scaling between the total mass of the cluster galaxies and their luminosity, i.e. $M_{\text{tot}}/L_{i} \propto L_{j}$, the slope of the scaling relation in Eq. 9 can be determined by:

$$
\beta_{\text{cut}} = \gamma - 2\alpha + 1,
$$

(10)

where $\gamma = 0.2$ to be consistent with the observed fundamental plane relation. In Table 1, we quote the best-fit values of $\alpha$, $\sigma_{LT,0}^{\text{gal}}$ and $\Delta\sigma_{ap}$, and the inferred value of $\beta_{\text{cut}}$.

In our lens models, we fix the two slopes $\alpha$ and $\beta_{\text{cut}}$ to the fitted (or inferred) values, while we use a Gaussian prior for the normalization $\sigma_{LT}^{\text{gal}}$, centered on the measured $\sigma_{LT}^{\text{gal}}$, with standard deviation equal to $\Delta\sigma_{ap}$. To obtain this prior on $\sigma_{LT}^{\text{gal}}$, we deproject the measured $\sigma_{gal}^{\text{cut}}$, as detailed in B19. Conversely, a large uniform prior between 1'' and 50'' is adopted for the normalization $r_{\text{cut}}^{\text{gal}}$ in Eq. 9.

Fig. 5. Marginalized posterior distributions for the fitting parameters of the $\sigma_{gal}^{\text{cut}}$-$m_{F160W}$ scaling relation (see Fig. 4). The normalization $\sigma_{gal}^{\text{cut}}$ is computed at the reference magnitude $m_{F160W}^{\text{ref}} = 17.02$ of the Northern BCG. The slope $\beta_{\text{cut}}$ of the $\sigma_{gal}^{\text{cut}}$-$m_{F160W}$ relation is inferred from Eq. 10 (see text).

3.1.2. Galaxy-scale mass distribution

Here we describe how we model the total mass (DM plus baryons) content of cluster member galaxies.
The main results from the optimizations of the four lens models are summarized in Table 2. In addition to the three indicators of the goodness of the lens models defined above (Δrms, BIC, and AIC), we introduce a fourth figure of merits, $\chi^2_{\text{lens}}$, which quantifies the agreement between the lens predicted and measured stellar velocity dispersions of member galaxies. If we call $\sigma_{\text{ap}}(m_{F160W,i})$ the aperture-averaged projected velocity dispersion of the $i$-th galaxy (with luminosity $m_{F160W,i}$), inferred from the best-fit LensTool scaling relation, $\chi^2_{\text{lens}}$ is defined as:

$$
\chi^2_{\text{lens}} = \sum_{i=1}^{N_{\text{gal}}^m} \left( \frac{\sigma_{\text{ap}}(m_{F160W,i}) - \sigma_{\text{gal}}^{\text{ref}}}{\sigma_{\text{ap},i}} \right)^2,
$$

where $N_{\text{gal}}^m$ is the number of galaxies with a measured velocity dispersion. As general rule, the lower is the $\chi^2_{\text{lens}}$ value, the better is the agreement between the LensTool best-fit $\sigma$-mag scaling relation and the measured Faber-Jackson relation.

Based on these four figure of merits, the LM-4HALOS lens model emerges as the best model, which reproduces the positions of the multiple images with the lowest $\Delta_{\text{rms}}$, and best match the internal kinematics of the cluster member galaxies (lowest $\chi^2_{\text{lens}}$). This model also possesses the lowest values of the BIC and AIC criteria.

In the upcoming sections, we will therefore characterize in detail this reference model by discussing its ability to predict robust positions and magnifications of multiple images, specifically those close to critical lines and around selected cluster galaxies.
Table 2. Description of the four selected best lens models. \( N_{par} \) and DoF are the number of model free-parameters and degrees-of-freedom. \( \Delta_{rms} \) is the root-mean-square displacement between the positions of observed and model-predicted multiple images (see Eq. 5). The BIC (Bayesian Information Criterion) and AIC (Akaike Information Criterion) values are computed using Eq. 6. The \( \chi^2_{lin} \) value, defined in Eq. 11, quantifies the agreement between the model predicted and measured cluster member velocity dispersions. In the last column, we summarize the differences of the mass parametrization in the four lens models. The reference model selected on the basis of the best figure of merits (columns 4–7) is indicated in bold.

| Model ID     | \( N_{par} \) | DoF | \( \Delta_{rms} \) ["] | BIC | AIC | \( \chi^2_{lin} \) | Description                                      |
|--------------|----------------|-----|--------------------------|-----|-----|---------------------|-------------------------------------------------|
| LM-4HALOS    | 30             | 202 | 0.40                     | 346 | 243 | 4941                | Four cluster-scale halos, 212 cluster members including BCGs |
| LM-BCGs      | 28             | 204 | 0.45                     | 408 | 311 | 6653                | Three cluster-scale halos, 210 cluster members excluding BCGs |
| LM–HLBCGs    | 28             | 204 | 0.46                     | 412 | 316 | 7217                | Three cluster-scale halos, 212 cluster members including BCGs |
| LM–SHEAR     | 27             | 205 | 0.48                     | 413 | 320 | 5042                | Three cluster-scale halos, one shear term, one convergence term, 212 cluster members including BCGs |

4.1. **LM-4HALOS lens model: predicted positions and magnifications of the multiple images**

In Fig. 7, we show the distribution of the differences between the observed and model-predicted positions of the multiple images in the \( x \) and \( y \) directions and corresponding values of \( \Delta_{rms} \). It is worth noting that the resulting \( \Delta_{rms} = 0.40 \) is 34% smaller than our previous model (B19) despite the significantly larger number of multiple images (182 vs 102). The rms value in the MDLF field (0.37") is appreciably smaller than the one in the Southern field (0.47"). We interpret this as the result of a better constrained mass model in the NE field due to the large number of multiple images (see Vanzella et al. 2020b). The mass distribution in the SE region appears more complex, as discussed by Balestra et al. (2016) who found MACS 0416 in a pre-merging phase based on a dynamical and structural analysis of the cluster. Interestingly, we find that the inclusion of the 4-th cluster-scale halo in the LM-4HALOS lens model is needed to significantly reduce \( \Delta_{rms} \) in the SE field, thus confirming the complex structure of the southern cluster field. The bright foreground galaxy (Fig. 6), which we simply model as a circular dPIE at the cluster redshift, adds further uncertainty to the projected mass distribution in this region. In Fig. 8, we show the spatial distribution of the absolute displacements, \( \Delta_i \), between the observed and model-predicted positions of multiple images. This also shows that the multiple images in the NE field are, on average, better reproduced independently from their redshifts.

The reference lens model predicts 226 multiple images associated to 66 independent sources. Note that there are 44 additional images predicted ‘a posteriori’ by the model, which are still waiting for a secure identification (see for example Sys-12 at the end of this sub-section).

To study the statistical uncertainty (\( \Delta m \)) on the absolute magnification, \( \mu \), we show in Fig. 9 the relative error on the absolute magnification in different \( \mu \) regimes. This is particularly relevant for the study of high-z strongly magnified sources (see Vanzella et al. 2020b). The magnification uncertainties are computed as follows. Firstly, we generate 500 realizations of the LM-4HALOS lens model by randomly extracting 500 parameter samples from the LensTool MCMC chains. For each realization, we compute the absolute magnifications at the predicted position of the multiple images and derive their posterior distributions. To confidently assign the correct magnification to each image, we verify that each model realization predicts the expected parity for that image. We define \( \mu \) as the median value of the magnification distributions, and \( \Delta m \) as half of the difference between the 84-th and 16-th percentiles.

One can see in Fig. 9 that the relative magnification uncertainty progressively increases at larger \( \mu \). This is due to the fact that the most magnified images are located closer to the critical
lines of the lens model where magnification gradients becomes particularly large. In these regions we also expect that systematic errors due to model parametrization become increasingly important. This analysis is deferred to a future paper.

To better understand the origin of magnification uncertainties, we investigate how $\Delta\mu/|\mu|$ varies as a function of the distance between the multiple images and their closest cluster galaxy, and the magnification itself. The result is shown in Fig. 10. On average, the images that form close to cluster members have larger relative errors. This anti-correlation is explained by the fact that galaxy masses act as small gravitational lenses embedded into the cluster potential and introduce numerous secondary critical lines (with sizes of few arcsec) into the lens model. Images closer to critical lines tend to have larger $\Delta\mu/|\mu|$ values. The few highly magnified images with $\Delta\mu/|\mu| \sim 20\%$ at distances of $6 - 8''$ from cluster galaxies are necessarily located around the critical lines of the cluster computed for their redshifts.

We now define a new metric, which is sensitive to the gradients of the deflection field. We focus on a number of specific lensed systems to test the ability of the lens model to predict the positions of multiply imaged clumps, associated to the same resolved source, which are close to the critical lines.

We take Sys-5 (see bottom-right panels of Fig. 2) as an example. This system is composed by six clumps (5.1, 5.2, 5.3, 5.4, 5.5, and 5.6) belonging to the same extended source at distance between the multiple images and their closest clumps, associated to the same resolved source, which are close to the critical lines. The few highly magnified images with $\Delta\mu/|\mu| \sim 20\%$ at distances of $6 - 8''$ from cluster galaxies are necessarily located around the critical lines of the cluster computed for their redshifts.

We now define a new metric, which is sensitive to the gradients of the deflection field. We focus on a number of specific lensed systems to test the ability of the lens model to predict the positions of multiply imaged clumps, associated to the same resolved source, which are close to the critical lines.

We take Sys-5 (see bottom-right panels of Fig. 2) as an example. This system is composed by six clumps (5.1, 5.2, 5.3, 5.4, 5.5, and 5.6) belonging to the same extended source at distance between the multiple images and their closest clumps, associated to the same resolved source, which are close to the critical lines. The few highly magnified images with $\Delta\mu/|\mu| \sim 20\%$ at distances of $6 - 8''$ from cluster galaxies are necessarily located around the critical lines of the cluster computed for their redshifts.

We now define a new metric, which is sensitive to the gradients of the deflection field. We focus on a number of specific lensed systems to test the ability of the lens model to predict the positions of multiply imaged clumps, associated to the same resolved source, which are close to the critical lines.

We take Sys-5 (see bottom-right panels of Fig. 2) as an example. This system is composed by six clumps (5.1, 5.2, 5.3, 5.4, 5.5, and 5.6) belonging to the same extended source at distance between the multiple images and their closest clumps, associated to the same resolved source, which are close to the critical lines. The few highly magnified images with $\Delta\mu/|\mu| \sim 20\%$ at distances of $6 - 8''$ from cluster galaxies are necessarily located around the critical lines of the cluster computed for their redshifts.

We now define a new metric, which is sensitive to the gradients of the deflection field. We focus on a number of specific lensed systems to test the ability of the lens model to predict the positions of multiply imaged clumps, associated to the same resolved source, which are close to the critical lines.

We take Sys-5 (see bottom-right panels of Fig. 2) as an example. This system is composed by six clumps (5.1, 5.2, 5.3, 5.4, 5.5, and 5.6) belonging to the same extended source at distance between the multiple images and their closest clumps, associated to the same resolved source, which are close to the critical lines. The few highly magnified images with $\Delta\mu/|\mu| \sim 20\%$ at distances of $6 - 8''$ from cluster galaxies are necessarily located around the critical lines of the cluster computed for their redshifts.

We now define a new metric, which is sensitive to the gradients of the deflection field. We focus on a number of specific lensed systems to test the ability of the lens model to predict the positions of multiply imaged clumps, associated to the same resolved source, which are close to the critical lines.

We take Sys-5 (see bottom-right panels of Fig. 2) as an example. This system is composed by six clumps (5.1, 5.2, 5.3, 5.4, 5.5, and 5.6) belonging to the same extended source at distance between the multiple images and their closest clumps, associated to the same resolved source, which are close to the critical lines. The few highly magnified images with $\Delta\mu/|\mu| \sim 20\%$ at distances of $6 - 8''$ from cluster galaxies are necessarily located around the critical lines of the cluster computed for their redshifts.

We now define a new metric, which is sensitive to the gradients of the deflection field. We focus on a number of specific lensed systems to test the ability of the lens model to predict the positions of multiply imaged clumps, associated to the same resolved source, which are close to the critical lines.

We take Sys-5 (see bottom-right panels of Fig. 2) as an example. This system is composed by six clumps (5.1, 5.2, 5.3, 5.4, 5.5, and 5.6) belonging to the same extended source at distance between the multiple images and their closest clumps, associated to the same resolved source, which are close to the critical lines. The few highly magnified images with $\Delta\mu/|\mu| \sim 20\%$ at distances of $6 - 8''$ from cluster galaxies are necessarily located around the critical lines of the cluster computed for their redshifts.

We now define a new metric, which is sensitive to the gradients of the deflection field. We focus on a number of specific lensed systems to test the ability of the lens model to predict the positions of multiply imaged clumps, associated to the same resolved source, which are close to the critical lines.

We take Sys-5 (see bottom-right panels of Fig. 2) as an example. This system is composed by six clumps (5.1, 5.2, 5.3, 5.4, 5.5, and 5.6) belonging to the same extended source at distance between the multiple images and their closest clumps, associated to the same resolved source, which are close to the critical lines. The few highly magnified images with $\Delta\mu/|\mu| \sim 20\%$ at distances of $6 - 8''$ from cluster galaxies are necessarily located around the critical lines of the cluster computed for their redshifts.

We now define a new metric, which is sensitive to the gradients of the deflection field. We focus on a number of specific lensed systems to test the ability of the lens model to predict the positions of multiply imaged clumps, associated to the same resolved source, which are close to the critical lines.

We take Sys-5 (see bottom-right panels of Fig. 2) as an example. This system is composed by six clumps (5.1, 5.2, 5.3, 5.4, 5.5, and 5.6) belonging to the same extended source at distance between the multiple images and their closest clumps, associated to the same resolved source, which are close to the critical lines. The few highly magnified images with $\Delta\mu/|\mu| \sim 20\%$ at distances of $6 - 8''$ from cluster galaxies are necessarily located around the critical lines of the cluster computed for their redshifts.

We now define a new metric, which is sensitive to the gradients of the deflection field. We focus on a number of specific lensed systems to test the ability of the lens model to predict the positions of multiply imaged clumps, associated to the same resolved source, which are close to the critical lines.

We take Sys-5 (see bottom-right panels of Fig. 2) as an example. This system is composed by six clumps (5.1, 5.2, 5.3, 5.4, 5.5, and 5.6) belonging to the same extended source at distance between the multiple images and their closest clumps, associated to the same resolved source, which are close to the critical lines. The few highly magnified images with $\Delta\mu/|\mu| \sim 20\%$ at distances of $6 - 8''$ from cluster galaxies are necessarily located around the critical lines of the cluster computed for their redshifts.

We now define a new metric, which is sensitive to the gradients of the deflection field. We focus on a number of specific lensed systems to test the ability of the lens model to predict the positions of multiply imaged clumps, associated to the same resolved source, which are close to the critical lines.

We take Sys-5 (see bottom-right panels of Fig. 2) as an example. This system is composed by six clumps (5.1, 5.2, 5.3, 5.4, 5.5, and 5.6) belonging to the same extended source at distance between the multiple images and their closest clumps, associated to the same resolved source, which are close to the critical lines. The few highly magnified images with $\Delta\mu/|\mu| \sim 20\%$ at distances of $6 - 8''$ from cluster galaxies are necessarily located around the critical lines of the cluster computed for their redshifts.
Magnification (\(\Delta \mu/|\mu|\)) of their distances from the closest cluster galaxy. The colors of the relative statistical uncertainty on the absolute magnification (Fig. 9). The distributions of the relative statistical error on the absolute magnification are obtained for images (as in Fig. 2). Each dot corresponds to a different pair of clumps, colored according to the parent system. We define \(d_{ij}^{\text{obs}}\) (\(d_{ij}^{\text{pre}}\)) as the measured (predicted) distance between the \(i\)-th and \(j\)-th clump belonging to the same resolved lensed image (e.g. \(i = 12.1, j = 12.5\) in the top left panel of Fig. 12). We plot the absolute difference \(|d_{ij}^{\text{obs}} - d_{ij}^{\text{pre}}|\) along the radial axis on the right side, while the angle \((\Delta \phi)\) between \(d_{ij}^{\text{obs}}\) and \(d_{ij}^{\text{pre}}\) vectors (measured counterclockwise) is plotted across the angular coordinate. On the left side of the diagram, we draw a binned distribution of all the data-points. The 90-th, 95-th, and 99-th percentiles of the distributions along the radial and angular directions are shown as red semi-circles and sectors respectively.

Using the observed and model-predicted positions of the multiple images, we determine the ‘distance vectors’ connecting each pair of lensed clumps belonging to the same extended image: e.g. \(d_{5,1a,5,2a}^{\text{obs(pre)}}, d_{5,1a,5,2a}^{\text{obs(pre)}}, d_{5,2a,5,3a}^{\text{obs(pre)}}, \) etc. (similar combinations are obtained for images \(b\) and \(c\)). The modulus of these vectors corresponds to the distance between these knots. Moreover, we measure the angular difference in the orientation of the corresponding observed and predicted ‘distance vectors’, so that \(\Delta \phi^{ij} = \arccos \left(\frac{d_{ij}^{\text{pre}} \cdot d_{ij}^{\text{obs}}}{\left|d_{ij}^{\text{obs}}\right| \left|d_{ij}^{\text{pre}}\right|}\right)\), where we use ‘\('\)' to indicate the scalar product, while \(i\) and \(j\) are different knots (e.g. \(\phi^{5,1a,5,2a}_\text{pre} = \arccos \left(d_{5,1a,5,2a}^{\text{pre}} \cdot d_{5,1a,5,2a}^{\text{obs}}\right)\)).

On the right side of Fig. 11, we plot, on the radial axis, the absolute value of the difference between the predicted and observed ‘distance vectors’ \(|d_{ij}^{\text{obs}} - d_{ij}^{\text{pre}}|\) for the whole set of multiple images belonging to eight different systems. The values of \(\Delta \phi^{ij}\) are plotted along the angular coordinate. In this representation, a perfectly reconstructed pair of images should lie on the origin of the reference frame. On the left side of the diagram, we show the angular distribution of the \(\Delta \phi^{ij}\) values. In 90% of the cases, the difference in the distances between the observed and predicted knots is lower than 0.33°. Similarly, the \(\Delta \phi^{ij}\) values are lower than 5.9° for 90% of image pairs.

This result demonstrates that our reference lens model not only accurately predicts the observed position of the multiple images (\(\Delta \text{rms} = 0.40\)°), but it is also able to accurately reproduce the extension and the orientation of the inner structure of extended multiple images.

To conclude this session characterizing the robustness of our reference model, we briefly describe the properties of two interesting systems of multiple images, namely Sys-12 and Sys-14. The extended background source associated to Sys-12, at \(z = 0.94\), is a spiral galaxy which is lensed into three images.
located 0.05″ from the critical line and have in fact the highest magnifications among the entire sample of multiple images, \( \mu_{12.4b} = 133^{+97}_{-86} \) and \( \mu_{12.4c} = 134^{+97}_{-86} \). We emphasize how our reference lens model reproduces the 12 positions of Sys-12 with very high precision (\( \Delta m = 0.05″ \)), as well as the relative distances and orientations of all knots belonging to a given lensed image (b or c), as one can appreciate from Fig. 11 (cyan dots cluster near the origin of the diagram).

The third image of Sys-12 is blindly predicted by the lens model ~ 20″ SE of image b/c (see Fig. 2), around a cluster member which contributes to create an Einstein ring configuration (see top-right and middle-right panels of Fig. 12). This third image is not included in the catalog of multiple images as the corresponding lensed knots cannot be readily identified in HST images.

The MUSE spectra of the multiple images in Sys-12 are characterized by a prominent [OII]3726.2, 3729.1 emission lines doublet (see \( \text{Vanzella et al.} \ 2020b \)). The wavelength drift of the [OII] peaks across the lensed images can be readily detected in the MUSE datacube\(^3\). This analysis yields the velocity maps shown in the lower panels of Fig. 12. By ray-tracing the six star-forming knots in image b and c on the source plane at \( z = 0.94 \), we find them very well aligned with a symmetric rotation curve stretching from ~ −100 km s\(^{-1}\) for 12.4, up to ~ 100 km s\(^{-1}\) for 12.5, across ~ 11 kpc. We defer to a future publication a detailed study of this and other galaxy-scale systems of multiple images.

Sys-14 (see Fig. 13) was studied in detail by \( \text{Vanzella et al.} \ 2017b \). In this system, two close compact sources (14.1 and 14.2) at \( z = 3.221 \) are lensed into six multiple images each (9 out of 12 multiple images are used as constraints into the lens models). Eight images, i.e. 14.1a, b, c, d, and 14.2a, b, c, d, are located around a pair of cluster galaxies. Since the mass distribution of the brighter galaxy (Gal-8971) strongly affects the geometry of the GSSLS, its mass halo is parameterized as an elliptical dPIE profile outside the subhalo scaling relations in all lens models (see Sec. 3.1 and Table 3). Note that the three images 14.2d, 14.2e, and 14.2f, are not used in the lens models, due to the lack of a secure counterpart in the HST images. Instead, we tentatively identify the 14.2e image and use it in the lens model with the same position as 14.1c and a positional error which includes both images. Compared to our previous lens model by B19, the new model predicts an additional counter-image of the 14.1 and 4.2 pair, which however still remains not securely identified (mid panel of Fig. 13). For all images associated to Sys-14, we find \( \Delta m = 0.25″ \). We refer to \( \text{Vanzella et al.} \ 2020b \) for a discussion on the magnification and physical sizes of the multiple images of Sys-14.

4.2. MACS 0416 total projected mass distribution

As described in Sec. 3.1, the mass distribution of MACS 0416 obtained with the LM-4HALOS lens model includes a cluster-scale component and a subhalo population traced by the cluster members. The former includes a fixed hot gas component and the dominant DM halo which is modeled with four elliptical dPIE profiles. The latter includes 212 circular dPIEs with vanishing core radii and a separate elliptical subhalo to model the bright member in Sys-14 (Fig. 13). In Table 3, we quote all the input parameters and the output optimized values of each parameter obtained with of our reference model. In Fig. 8, we show the isodensity contours of the total projected mass distribution overlaid

\[^3\] We used the Python library \texttt{mpdaf} for this analysis (https://mpdaf.readthedocs.io/en/latest).
Fig. 13. Inspection of system 14 at $z = 3.221$. A pair of nearly point-like clumps (14.1 and 14.2) is imaged six times by the cluster mass distribution, around two cluster galaxies (Gal-8971 and Gal-8980), with F160W magnitudes 19.6 and 21.5, respectively. Green crosses, red ellipses, and blue lines have the same meaning as Fig. 12. The three gray-scale cutouts on the left are median stack images of F814W, F606W, and F435W bands, respectively. Green vertical segments mark the positions of the 182 multiple images computed at $z = 3.221$. The lens model predicts a fifth ($e$) and a sixth ($f$) image (middle and bottom panels, see also Fig. 2).

Fig. 14. Top: Cumulative projected total mass profile of MACS 0416 as a function of the distance ($R$) from the Northern BCG. The red curve refers to the median and the 68% confidence level obtained from the reference lens model. The previous result by B19 is plotted in blue. The total mass of the subhalos, associated to cluster galaxies, is shown in green. Black vertical segments mark the positions of the 182 multiple images. Bottom: Ratio between the new mass profile and the one from B19 (blue) and fractional contribution of the subhalo component to the total cumulative mass (in green).

The optimized position of the third halo in the NE region, whose presence is needed to reduce significantly the $\Delta_{\text{rms}}$ value, is very close to a local over-density of the projected galaxy distribution (Fig. 8), in keeping with C17 and B19.

The cumulative projected total mass profile of MACS 0416 as a function of the projected distance, $R$, from the BCG-N is shown in Fig. 14 (red thin band). The small 1-$\sigma$ uncertainties are computed by randomly extracting parameters from the MCMC chains of 500 realizations of the LM-4HALOS model. The new cumulative mass profile is found in very good agreement with that obtained with our previous B19 model, which had a somewhat different parametrization and used 80 fewer multiple images (the difference is less than 3% over the radial range where the multiple images are located).

Fig. 14 also shows the cumulative projected mass profile associated to the subhalos, and their relative contribution to the total mass. At small projected distances from the BCG-N, the subhalos contribute for more than 40% to the total mass, while this contribution does not exceed 15% at $R > 100$ kpc.

As discussed in Sec. 3.1.2, the normalization of the $\sigma_{160W}$ scaling relation for the subhalos in all the lens models is obtained adopting a Gaussian prior derived from the observed $\sigma_{160W}$ value, see magenta square). As discussed in Sec. 3.1.2, the normalization of the $\sigma_{160W}$ scaling relation for the subhalos in all the lens models is obtained adopting a Gaussian prior derived from the observed $\sigma_{160W}$ value, see magenta square). As discussed in Sec. 3.1.2, the normalization of the $\sigma_{160W}$ scaling relation for the subhalos in all the lens models is obtained adopting a Gaussian prior derived from the observed $\sigma_{160W}$ value, see magenta square).
The projected mass of a dPIE within an aperture of radius $R_E$ is given by (Elíasdóttir et al. 2007 and B19):

$$M(R_E) = \frac{\pi \sigma_0^2}{G} \left( \sqrt{r_{\text{core}}^2 + R_E^2} - r_{\text{core}} - \sqrt{r_{\text{core}}^2 + r_{\text{cut}}^2} + r_{\text{cut}} \right). \quad (12)$$

For a fixed value of the core radius, a higher values of $r_{\text{cut}}$ produces the same mass $M(R_E)$ for a smaller central velocity dispersion $\sigma_0$. Our best fit model yields an $r_{\text{cut}}$ value for this specific galaxy exceeding $R_E$, with a large uncertainty.

In Fig. 15, we show the posterior distribution of the reference central velocity ($\sigma_0^{\text{ref}}$) and the reference truncation radius ($r_{\text{cut}}^{\text{ref}}$) of the cluster member scaling relations, which clearly displays the same degeneracy discussed above. The constraints on these two parameters obtained for the new reference model are compared with the results of B19, which also used a kinematic prior in the lens models. We note that the normalization of the scaling relations ($\sigma_0^{\text{ref}}$) in the new model is consistent with B19 at 1-$\sigma$ level, whereas we find a difference of $\sim 30$ kpc in the normalization of the truncation radius. It is worth noting that such a difference could not be detected without using galaxy kinematics to constrain the scaling relations in the lens model, because in this case the $\sigma_0 r_{\text{cut}}$ degeneracy is significantly larger (see B19).

Since the B19 model assumed a larger value of the reference core radius of the subhalo dPIE profiles ($r_{\text{core}}^{\text{B19}} = 0.05''$ compared to $1'' \times 10^{-4}$ in LM-4HALOS), we re-optimize the lens model changing the core radius parameter. The result, illustrated in Fig. 15 (dashed confidence contours), shows that $\sigma_0$ scales as expected with the larger core radius, based on Eq. 12, however $r_{\text{cut}}$ does not change significantly. Therefore, the larger truncation radii of the subhalos obtained with the new model are not due to inherent degeneracies of the dPIE parametrization with $r_{\text{cut}}^{\text{ref}}$.

Considering the dependence of the total mass of the subhalos on their values of velocity dispersion and cut radius (12), the cluster members in the LM-4HALOS model are characterized by larger masses (by a factor of $\sim 2$) compared to those in the model of B19. Based on the previous model, Meneghetti et al. (2020) found that cluster members in MACS 0416 have a significantly larger probability to produce GSSLS compared to galaxy-scale subhalos in ΛCDM numerical simulations. We have verified that the GSSLS probability obtained from the LM-4HALOS model is $\sim 30\%$ higher than that measured from the B19 model. Thus, our new results confirm the tension reported by Meneghetti et al. (2020) between the observations of the inner structure of cluster galaxies and the theoretical expectations in the framework of the ΛCDM model.

We note that constraints on the mass distribution of the subhalo population with lens models can be further improved by taking into account the measured velocity dispersion for each member galaxy, thus including the intrinsic scatter of the $\sigma-L$ scaling relation (Bergamini et al. 2020), which is neglected in the current lens models. We leave to a future work a further analysis of the subhalo component, which fully exploits strong lensing and galaxy kinematics constraints.

5. Conclusions

We have presented a new high-precision strong lens model for the galaxy cluster MACS J0416.1–2403 at $z = 0.396$, which takes advantage of the MUSE deep lensed field of 17.1h, carried out in the North-East region of the cluster (Vanzella et al. 2020b). By combining this deep pointing with a careful re-inspection of the HST images, we have identified 82 additional spectroscopic multiple images, compared to the previous catalog published in C17, resulting in a new sample of 182 images from 48 different background sources. In this new sample, we have added several multiply lensed clumps, belonging to resolved extended sources, which are particularly efficient in constraining the position of the critical lines in their vicinity. Moreover, we have extended the sample of cluster members to 213 galaxies, 171 of which are spectroscopically confirmed (27 more than in our previous catalog).

By exploiting new, high signal-to-noise galaxy spectra extracted from the MDF, we have measured the inner stellar kinematics of 64 cluster members (15 more than in the previous analysis by B19), down to $m_{\text{F160W}} \approx 22$. We have used the newly measured velocity dispersions to constrain the scaling relations of the subhalo population in the lens models as in B19. The cluster-scale mass distribution is modeled with a number of DM-dominated halos in addition to the hot gas component traced by the Chandra X-ray data.

Among the four lens models in our study, we have selected a reference model, which best reproduces the positions of the observed multiple images (smallest $\Delta_{\text{ani}}$) and the observed $\sigma$-mag scaling relation for cluster members (see Table 2).

We can summarize the results obtained from our reference MACS 0416 lens model as follows:

![Fig. 15. Marginalized posterior distributions of the normalizations $\sigma_0^{\text{ref}}$ and $r_{\text{cut}}^{\text{ref}}$ of the cluster member scaling relations (see Eq. 8 and 9). Normalizations are computed at the magnitude of the BCG-$N$ (mag$_{\text{F160W}} = 17.02$). Red distributions refer to the LM-4HALOS reference lens model; results from the previous model by B19 model are in blue. Colored contours encompass the 1, 2, 3 $\sigma$ confidence levels; the vertical solid and dashed lines correspond to the 50-th, 16-th and 84-th percentiles of the marginalized distributions. The 1 and 2 $\sigma$ black dashed contours refer to the LM-4HALOS model with $r_{\text{core}}^{\text{ref}} = 0.05''$, instead of $r_{\text{core}}^{\text{B19}} = 1'' \times 10^{-4}$ of the reference model. The green and magenta lines are $\sigma_0-r_{\text{cut}}$ curves with constant projected mass, within an aperture of $R = 1''$ (= 5.34 kpc at $z = 0.396$), for a circular dPIE profile. The mass values from bottom to top are quoted in the legend. Green curves refer to $r_{\text{core}} = 1'' \times 10^{-4}$, magenta curves refer $r_{\text{core}} = 0.05''$ (as in B19).]
1. Despite the large number of observed multiple images, the new reference lens model yields $\Delta m_{\text{ns}} = 0.4''$, thus reducing by 1/3 the value obtained by B19. The model is particularly accurate in the MDLF region of MACS 0416 ($\Delta m_{\text{ns}}^E = 0.37''$), where the newly identified images are located.

2. We have studied the robustness of the magnification maps derived from our reference lens model. We have computed the uncertainties on the absolute magnification ($|\mu|$) of multiple images on their predicted positions by sampling the $|\mu|$ distributions from the MCMC chains of the lens model. We have studied how the relative error on the magnification, $\Delta |\mu|/|\mu|$, varies with the magnification and with the distance of the multiple images from secondary critical lines associated to cluster galaxies. We find that the relative error remains within $\sim 10\%$ at $|\mu| \leq 10$, at distances beyond 2'' from the center of member galaxies, while it increases up to $\sim 100\%$ close to critical lines ($|\mu| \geq 20$).

3. We have investigated the ability of the new lens model to reproduce the inner fine structure of extended sources by defining a new metric, which is sensitive to the gradients of the deflection field. We find that the new model can predict the relative distances and orientations of pairs of multiply imaged clumps, inside well resolved sources in the vicinity of critical lines, with errors less than 0.33'' and 5.9'', respectively, for 90% of image pairs. This analysis lends support to the interpretation presented by Vanzella et al. (2020b) of the nature of highly magnified clumps in resolved lensed galaxies in the MLDF. They range from young massive star clusters (e.g. in Sys-14, see also Vanzella et al. 2017b), to barely resolved knots, with sizes and luminosities similar to what observed in local star clusters (e.g. Sys-12).

The new lens model of MACS 0416 presented here, together with the new catalog of cluster members and the updated catalog of multiple images presented in Vanzella et al. (2020b), will be made publicly available upon acceptance of this paper4.

Acknowledgements. This project is partially funded by PRIM-MIUR 2017WSCC32. PB acknowledges financial support from ASI through the agreement ASI-INAF n. 2018-29-HH.0. This project is partially funded by PRIM-MIUR. MM acknowledges support from the Italian Space Agency (ASI) through contract “Euclid - Phase D” and from the agreement ASI-INAF n. 2018-29-HH.0. PB acknowledges financial support from ASI though the agreement ASI-INAF n. 2018-29-HH.0.

References

Akaike, H. 1974, IEEE Transactions on Automatic Control, 19, 716
Angora, R., Rosati, P., Brescia, M., et al. 2020, arXiv e-prints, arXiv:2009.08224
Annunziatella, M., Bonamigo, M., Grillo, C., et al. 2017, ApJ, 851, 81
Bacon, R., Accardo, M., Adjali, L., et al. 2012, The Messenger, 147, 4
Balestra, L., Mercurio, A., Sartoris, B., et al. 2016, ApJS, 224, 33
Bergamini, P., Agnello, A., & Caminha, G. 2020, arXiv e-prints, arXiv:2008.11728
Bergamini, P., Rosati, P., Mercurio, A., et al. 2019, arXiv e-prints, arXiv:1905.13236
Biviano, A., Rosati, P., Balestra, I., et al. 2013, A&A, 558, A1
Bonamigo, M., Grillo, C., Ettori, S., et al. 2018, ApJ, 864, 98
Boylan-Kolchin, M. 2018, MNRAS, 479, 332
Caminha, G. B., Grillo, C., Rosati, P., et al. 2016, A&A, 587, A80
Caminha, G. B., Grillo, C., Rosati, P., et al. 2017, A&A, 600, A90
Caminha, G. B., Rosati, P., Grillo, C., et al. 2019, arXiv e-prints, arXiv:1903.05103
Cappellari, M. 2017, MNRAS, 466, 798

4 Previous lens models and catalogs have been available at www.fe.infn.it/astro/lensing