Publication Year	2017
Acceptance in OA@INAF	2020-08-31T11:43:10Z
Title	The Frontier Fields lens modelling comparison project
Authors	MENEGHETTI, MASSIMO; Natarajan, P.; Coe, D.; Contini, E.; DE LUCIA, GABRIELLA; et al.
DOI	10.1093/mnras/stx2064
Handle	http://hdl.handle.net/20.500.12386/27000
Journal	MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY
Number	472

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **472,** 3177–3216 (2017)
Advance Access publication 2017 August 17

doi:10.1093/mnras/stx2064

The Frontier Fields lens modelling comparison project

M. Meneghetti,[1,2]★ P. Natarajan,[3] D. Coe,[4] E. Contini,[5] G. De Lucia,[6] C. Giocoli,[7] A. Acebron,[7] S. Borgani,[6,8,9] M. Bradac,[10] J. M. Diego,[11] A. Hoag,[10] M. Ishigaki,[12,13] T. L. Johnson,[14] E. Jullo,[7] R. Kawamata,[15] D. Lam,[16] M. Limousin,[7] J. Liesenborgs,[17] M. Oguri,[12,18,19] K. Sebesta,[20] K. Sharon,[14] L. L. R. Williams[20] and A. Zitrin[21]†

Affiliations are listed at the end of the paper

Accepted 2017 August 9. Received 2017 August 8; in original form 2016 June 9

ABSTRACT

Gravitational lensing by clusters of galaxies offers a powerful probe of their structure and mass distribution. Several research groups have developed techniques independently to achieve this goal. While these methods have all provided remarkably high-precision mass maps, particularly with exquisite imaging data from the *Hubble Space Telescope* (*HST*), the reconstructions themselves have never been directly compared. In this paper, we present for the first time a detailed comparison of methodologies for fidelity, accuracy and precision. For this collaborative exercise, the lens modelling community was provided simulated cluster images that mimic the depth and resolution of the ongoing *HST* Frontier Fields. The results of the submitted reconstructions with the un-blinded true mass profile of these two clusters are presented here. Parametric, free-form and hybrid techniques have been deployed by the participating groups and we detail the strengths and trade-offs in accuracy and systematics that arise for each methodology. We note in conclusion that several properties of the lensing clusters are recovered equally well by most of the lensing techniques compared in this study. For example, the reconstruction of azimuthally averaged density and mass profiles by both parametric and free-form methods matches the input models at the level of ∼10 per cent. Parametric techniques are generally better at recovering the 2D maps of the convergence and of the magnification. For the best-performing algorithms, the accuracy in the magnification estimate is ∼10 per cent at $\mu_{\rm true} = 3$ and it degrades to ∼30 per cent at $\mu_{\rm true} \sim 10$.

Key words: gravitational lensing: strong – gravitational lensing: weak – galaxies: clusters: general – galaxies: high-redshift.

1 INTRODUCTION

Gravitational lensing has become an increasingly popular method to constrain the matter distribution in clusters (see e.g. Soucail et al. 1987; Fort et al. 1988; Lynds & Petrosian 1989; Kneib et al. 1996; Broadhurst et al. 2005; Smith et al. 2005; Jullo et al. 2007; Limousin et al. 2007, 2008, 2012, 2016; Jullo & Kneib 2009; Richard et al. 2010a; Postman et al. 2012). Strong lensing, as it turns out, is particularly suited to probing the dense central regions of clusters. Constraining the structure of the cluster cores and their density profiles is critical to our understanding of structure formation; probing the nature of dark matter and fully comprehending the interplay between baryons and dark matter. Lensing by massive clusters has proved to be an invaluable tool to study their properties, in particular the detailed dark matter distribution within the cluster, as well as the faint, distant background population of galaxies that they bring into view (see Kneib & Natarajan 2011; Bartelmann et al. 2013; Meneghetti et al. 2013; Natarajan et al. 2017, and references therein for recent reviews). The magnification provided by lensing therefore affords the determination of the luminosity function of these high-redshift sources down to faint luminosities, thus helping inventory and identify galaxies that might have re-ionized the Universe (Seitz et al. 1998; Stark et al. 2007; Richard et al. 2008, 2011; Vanzella et al. 2012, 2014, 2015, 2016; Bouwens et al. 2014, 2015; Robertson et al. 2015; Huang et al. 2016).

Over the past two decades the *Hubble Space Telescope* (*HST*) has revolutionized the study of cluster lenses; and, with the deployment

★ E-mail: massimo.meneghetti@oabo.inaf.it
† Hubble Fellow.

© 2017 The Authors
Published by Oxford University Press on behalf of the Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

of ever more sensitive cameras from the Wide Field Planetary Camera2 to the Advanced Camera for Surveys (ACS), the data have become exquisite in terms of resolution. By 2005, mass distributions derived from lensing data were available for about 30 clusters. More recently, galaxy clusters were the primary targets of two multicycle treasury programs of the *HST* aiming at finding signatures of strong gravitational lensing in their cores. These are the 'Cluster Lensing And Supernova survey with Hubble' [CLASH; PI M. Postman (GO 12065); see Postman et al. 2012] and the ongoing Frontier Fields Initiative (FFI; PI: Lotz; see Lotz et al. 2017).

As part of the Frontier Fields programme, *HST* is currently collecting data of unprecedented depth on fields that harbour six massive clusters that act as powerful gravitational lenses. This programme utilizes orbits under the Director's Discretionary observing time. The FFI is a revolutionary deep field observing program aimed at peering deeper into the universe than ever before to not only help understand better these dramatic lenses and their properties, but also simultaneously bring into view faint, distant background galaxies that would otherwise remain unseen without the magnification provided by the foreground lens. These high-redshift sources that can be accessed due to gravitational lensing provide a first glimpse likely of the earliest galaxies to have formed in the universe, and offer a preview of coming attractions that await unveiling by the upcoming *James Webb Space Telescope*. These Frontier Fields uniquely combine the power of *HST* with that of nature's gravitational telescopes – the high magnifications produced by these massive clusters of galaxies.

Utilizing both the Wide Field Camera 3 (WFC3) and ACS in parallel in this current program, *HST* has been producing the deepest observations of clusters and the background galaxies that they lens; as well as observations of flanking blank fields that are located near these selected clusters. These images have revealed the presence of distant galaxy populations that are ∼10–100 times fainter than any previously observed (Atek et al. 2015a,b; Livermore, Finkelstein & Lotz 2017). The magnifying power of these clusters is proving to be invaluable in helping improve our statistical understanding of early galaxies that are likely responsible for the re-ionization of the universe, and are providing unprecedented measurements of the spatial distribution of dark matter within massive clusters. These six clusters span the redshift range $z = 0.3$–0.55. The program devotes 140 orbits to each cluster/blank field pair, achieving a limiting AB magnitude of $M_{\rm AB} \approx 28.7$–29 mag in the optical (ACS) and near-infrared (WFC3) bands.

The fundamental ingredient for exploiting the science outlined above is the construction of robust and reliable lens models. The ongoing FFI is an unprecedented test-bed for lens modelling techniques. Given the depth of these *HST* observations, hundreds of multiple images, covering a broad redshift range, have been newly unveiled behind each of the observed clusters (Jauzac et al. 2014; Diego et al. 2015b; Grillo et al. 2015; Jauzac et al. 2015; Wang et al. 2015; Hoag et al. 2016; Kawamata et al. 2016). In a rather unique case, even time delay measurements from a serendipitously multiply imaged supernova 'Refsdal' observed by the Grism Lens-Amplified Survey from Space (GLASS) team (Treu et al. 2015) in the FFI cluster MACSJ1149.5+2223 became available for testing and refining the lens models (Kelly et al. 2015; Jauzac et al. 2016; Rodney et al. 2016; Treu et al. 2016). Most importantly, FFI data were made publicly available immediately. Five teams were contracted by STScI to produce gravitational lensing models for all six Frontier Fields clusters to be made available to the astronomical community at large to enable wide use of this incredible data set. All teams share the latest observational constraints, including positions and redshifts of multiple images[1] before working independently to produce lensing models which are also made publicly available.[2] Several additional groups have also been working on the data and producing mass models. In short, the whole community of strong lensing modellers has been actively collaborating to maximally exploit the FFI data.

The process of converting the observed strong lensing constraints into matter distributions is called *lens inversion*. Several groups have developed algorithms that perform the lens inversion employing different methodologies and using various combinations of input constraints. These include other tracers of the cluster gravitational potential, such as weak lensing, galaxy kinematics and the X-ray emission from the intracluster medium (see e.g. Bradač et al. 2005; Donnarumma et al. 2011; Medezinski et al. 2013; Newman et al. 2013; Umetsu 2013; Umetsu et al. 2014; Merten et al. 2015). Over the years, it has become clear that while all methods are equally well motivated, they do not always converge to consistent reconstructions, even when applied to the same lens system (e.g. Smith et al. 2009; Zitrin & Broadhurst 2009). In several cases strong-lensing masses for the same cluster lens were found to be in tension (by a factor 2–3) with other independent measurements, based e.g. on the modelling of the X-ray emission by the intracluster gas (Ebeling et al. 2009; Richard et al. 2010b; Donahue et al. 2014). The constraints from strong lensing need to be combined and fit simultaneously with stellar kinematic data and with weak lensing measurements (Newman et al. 2011) to improve accuracy. Using constraints on the mass profile arising from probes other than lensing also helps break the mass-sheet degeneracy (see e.g. Treu & Koopmans 2002; Nordin et al. 2014; Zitrin, Redlich & Broadhurst 2014; Rodney et al. 2015). Finally, in several clusters, lensing data alone seems unable to discriminate between various density profiles (Shu et al. 2008). Therefore, in some clusters the data favours steep inner density profile slopes, while, in others it favours extremely shallow density profiles. This is in contrast with the predictions from the cold dark matter paradigm (Sand, Treu & Ellis 2002; Sand et al. 2005; Newman et al. 2013; Newman, Ellis & Treu 2015, but see also Bartelmann & Meneghetti 2004; Meneghetti et al. 2007) where a universal density profile is expected with minor modification due to the aggregation of baryons in the inner regions.

In this paper, we challenge these lens inversion methods to reconstruct synthetic lenses with known input mass distributions. The goals of this exercise are twofold. First, we aim to provide concrete feedback to the lens modellers on how they may improve the performance of their codes. And secondly, we aim to provide potential users of the FFI models and the astronomical community at large a sharper, more quantitative view of how robustly specific properties of lenses are recovered and the sources of error that plague each method. Such a comparison with numerical simulations and contrasting of lens mapping methodologies has not been undertaken before.

The outline of the paper is as follows. In Section 2, we outline the lens modelling challenge. In Section 3, we briefly introduce the various lens modelling techniques that were employed by participants in this study. In Section 4, we discuss the results of the reconstructions. Section 5 is dedicated to the detailed comparison

[1] The redshifts are mainly obtained in the framework of the GLASS and CLASH-VLT programs (Grillo et al. 2015; Treu et al. 2015) and with the integral field spectrograph Multi Unit Spectroscopic Explorer (MUSE) on the VLT (see e.g. Karman et al. 2015).

[2] https://archive.stsci.edu/prepds/frontier/lensmodels/

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

of the independent modelling techniques through suitably defined metrics. Finally, in Section 7, we summarize the main results of this study and present our conclusions.

2 THE CHALLENGE

The challenge that we presented to various groups of lens modellers comprised of analysing simulated observations of two mock galaxy clusters and producing magnification and mass maps for them. In generating these simulated (mock) clusters, we attempted to reproduce the depth, colour and spatial resolution of *HST* observations of the FFI cluster images including the gravitational lensing effects. While the comparison of lensing reconstructions of real clusters using the same input observational constraints strongly indicates that currently developed lens inversion techniques are robust (see e.g. Jullo et al. 2007; Grillo et al. 2015), the analysis of simulated data involving a large degree of realism where the true underlying mass distribution is known can help the lens reconstruction community to greatly improve their understanding of the modelling systematics. This view of using mocks to calibrate methodologies is widely supported by a number of extensive investigations carried out in the last few years.

There are multiple advantages to such calibration exercises. First of all, we are able to produce reasonably realistic cluster mass distributions in simulations (although up to some limit) that can be used as lensing clusters. Building on an extensive analysis of *N*-body/hydrodynamical simulations to improve the knowledge of strong lensing clusters, we have identified the important properties of the lenses which need to be taken into account during the construction of a lens model: cluster galaxies (Meneghetti et al. 2000; Meneghetti, Bartelmann & Moscardini 2003a), ellipticity and asymmetries (Meneghetti, Bartelmann & Moscardini 2003b), substructures (Meneghetti et al. 2007), baryonic physics (Puchwein et al. 2005; Killedar et al. 2012) and the dynamical state (Torri et al. 2004). In fact, we can simulate the lensing effects of galaxy clusters accounting for all these important properties, using both state-of-the-art hydrodynamical simulations and semi-analytic models. Secondly, we have developed tools to produce mock observations of these simulated lenses. Our image simulator SKYLENS (Meneghetti et al. 2008, 2010a) can mimic observations taken virtually with any telescope, but here we have used it primarily to produce simulations of *HST* images taken with both the ACS and the WFC3. In a small-scale realization of the experiment that we present here, we applied the lens inversion techniques to a limited number of simulated observations of our mock lenses. By doing so, we highlighted some key limits of the strong lensing methods. For example, we note that strong lensing alone is powerful at constraining the cluster mass within the Einstein radius ($\sim$100 kpc for a massive cluster) but the addition of further constraints at larger radii are required in order to appropriately measure the shape of the density profiles out to the cluster outskirts (Meneghetti et al. 2010a; Rasia et al. 2012). In what follows, we describe in detail how we generate the mock data set for the challenge, and what kind of high-level products were distributed to the participants.

2.1 Generation of mock cluster lenses

For the exercise reported here, we generated mass distributions for two massive cluster lenses. These two lenses are generated following substantially different approaches, as outlined below. In order to easily distinguish them, we assigned them names – *Ares* and *Hera*.

2.1.1 Ares

The mass distribution of the first simulated galaxy cluster, *Ares*, is generated using the semi-analytic code MOKA[3] (Giocoli et al. 2012a). This software package builds up mock galaxy clusters by treating them as being comprised of three components: (i) the main dark matter halo – assumed to be smooth, triaxial and well fit with an NFW profile, (ii) cluster members – subhaloes, distributed to follow the main halo and to have a truncated singular isothermal sphere profile (Hernquist 1990), and (iii) the brightest cluster galaxy (BCG) modelled with a separate Hernquist (1990) profile. The axial ratios, a/b and a/c, of the main halo ellipsoid are randomly drawn from the Jing & Suto (2002) distributions requiring $abc = 1$. We note that the observed FFI clusters typically consist of merging subclusters that cause them to be particularly efficient and spectacular lenses. As shown by Torri et al. (2004), the strong lensing cross-section is boosted significantly during the merging processes.

In the attempt to generate a mass distribution that adequately replicates the complexity of the Frontier Fields clusters, *Ares* was produced by combining two large-scale mass distributions at $z = 0.5$. The two clumps have virial masses $M_1 = 1.32 \times 10^{15}\, h^{-1}\, \mathrm{M}_\odot$ and $M_2 = 8.8 \times 10^{14}\, h^{-1}\, \mathrm{M}_\odot$ and their centres are separated by $\sim$400 h^{-1} kpc. In each of the two cases, we start by assigning the same projected ellipticity to the smooth component, to the stellar density and to the subhalo spatial distribution. This is motivated by the hierarchical clustering scenario wherein the BCG and the substructures are related to the cluster as a whole and retain memory of the directions of the accretion of repeated merging events (Kazantzidis et al. 2004, 2008, 2009; Fasano et al. 2010). In order to introduce some level of asymmetry, we then added in a small twist to the surface density contours. The degree of twisting adopted reproduces variations of the orientation of iso-surface density contours measured in numerically simulated galaxy clusters (see e.g. Meneghetti et al. 2007). The two large-scale haloes combined to create *Ares* are nearly aligned. The difference between the position angles of the two clumps is $\sim$21 deg. The central region of *Ares* contains large baryonic concentrations to mimic the presence of BCGs. We account for the possible adiabatic contraction of the dark matter caused by the presence of BCGs for *Ares* (although several empirical studies find no evidence of adiabatic contraction on these scales, see e.g. Newman et al. 2013; Dutton & Treu 2014). The adiabatic contraction as described by Keeton & Madau (2001) for Hernquist (1990) was implemented. For further details of the MOKA code, we refer to Giocoli et al. (2012a,b). MOKA also takes into account the correlation between assembly history and various halo properties that are expected in CDM: (i) less massive haloes typically tend to be more concentrated than the more massive ones, and (ii) at fixed mass, earlier forming haloes are more concentrated and contain fewer substructures. These recipes have been implemented in consonance with recent results from numerical simulations. In particular, we assume the Zhao et al. (2009) relation to link the concentration to mass and the Giocoli et al. (2010) relation for the subhalo abundance. When substructures are included we define the smooth mass as $M_{\rm smooth} = M_{\rm vir} - \sum_i m_{{\rm sub},\,i}$ and its concentration $c_{\rm s}$ are defined such that the total (smooth+clumps) mass density profile has a concentration $c_{\rm vir}$, equal to that of the total virial mass of the halo.

[3] http://cgiocoli.wordpress.com/research-interests/moka

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

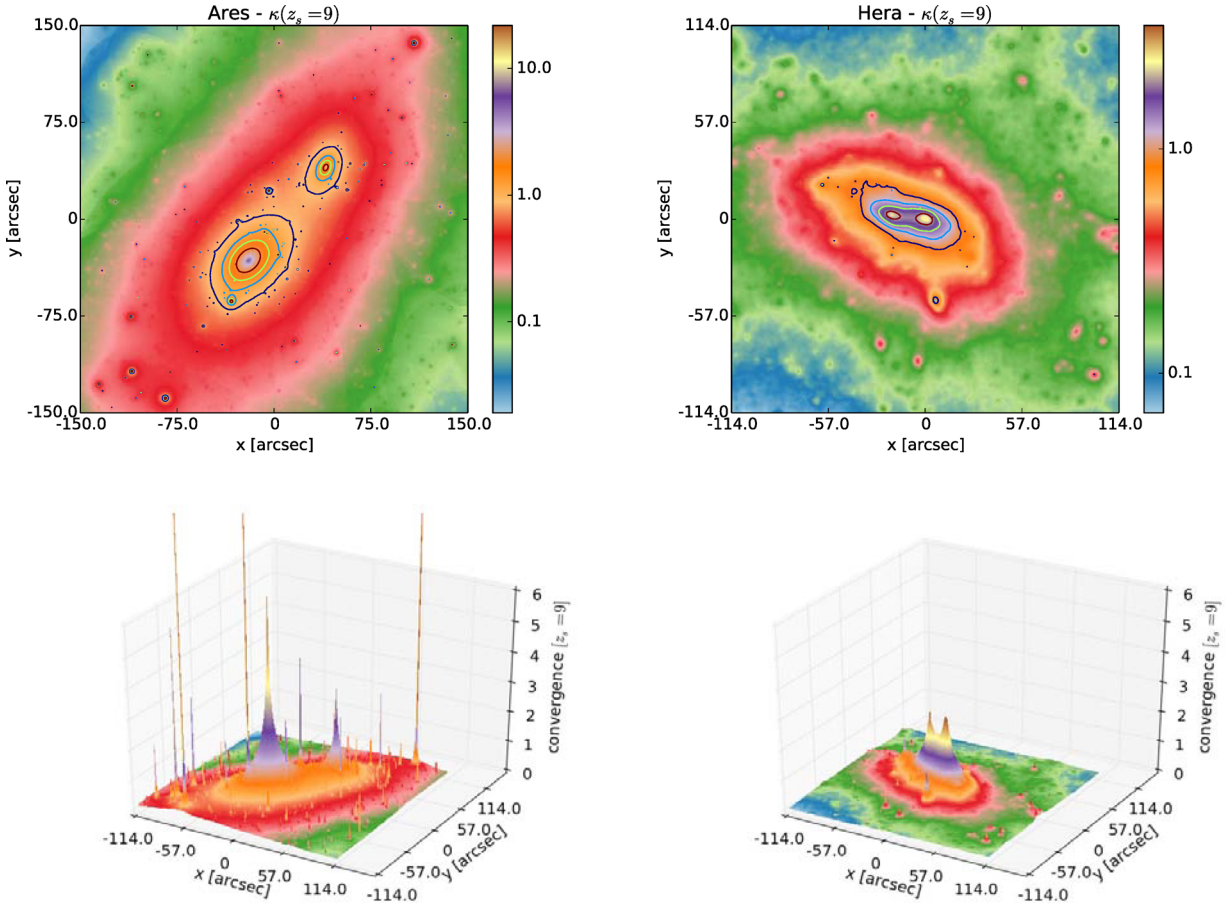

Figure 1. 2D and 3D convergence maps of *Ares* (left-hand panels) and *Hera* (right-hand panels). The maps are normalized to the source redshift $z_s = 9$.

Throughout the paper, the quoted masses and concentrations are evaluated at the virial radius, $M_{\rm vir}$ and $c_{\rm vir}$. For these definitions, we adopt derivations from the spherical collapse model:

$$M_{\rm vir} = \frac{4\pi}{3} R_{\rm vir}^3 \frac{\Delta_{\rm vir}(z)}{\Omega_{\rm m}(z)} \Omega_0 \rho_c \ , \tag{1}$$

where $\rho_c = 2.77 \times 10^{11}\, h^2\, {\rm M}_\odot\, {\rm Mpc}^{-1}$ represents the critical density of the Universe, $\Omega_0 = \Omega_{\rm m}(0)$ is the matter density parameter at the present time, $\Delta_{\rm vir}$ is the virial overdensity (Eke, Cole & Frenk 1996; Bryan & Norman 1998) and $R_{\rm vir}$ symbolizes the virial radius of the halo, i.e. the distance from the halo centre that encloses the desired density contrast; and

$$c_{\rm vir} \equiv \frac{R_{\rm vir}}{r_s} \ , \tag{2}$$

with r_s the radius at which the NFW profile approaches a logarithmic slope of -2. The concentrations assigned to the two main mass components of *Ares* are $c_1 = 5.39$ and $c_2 = 5.46$, respectively.

Ares is generated in a flat ΛCDM cosmological model with matter density parameter $\Omega_{\rm m,\,0} = 0.272$. The Hubble parameter at the present epoch is $H_0 = 70.4\ {\rm km\ s^{-1}\ Mpc^{-1}}$.

In the left-hand panels of Fig. 1, we show the convergence maps of *Ares*, calculated for a source redshift $z_{\rm s} = 9$. The cluster is elongated in the SE–NW direction and contains several massive substructures. Since *Ares* was generated using semi-analytical methods (SAMs), the small-scale substructures of its mass distribution are very well resolved, as shown in the bottom-left panel. The substructure mass function is shown in the right-hand panel of Fig. 2. As expected, this scales as $N(M) \propto M^{-0.8}$, consistent with results of numerical simulations (Giocoli et al. 2010) of the CDM model. The convergence profile, measured from the centre of the most massive clump, is shown in the left-hand panel of Fig. 2.

In the image simulations described later, we also include the light emission from cluster members. MOKA populates the dark matter subhaloes with galaxies using the Halo Occupation Distribution (HOD) technique. Stellar masses and *B*-band luminosities are subsequently assigned to each galaxy accordingly to the mass of the dark matter (sub)halo within which it formed, following Wang et al. (2006). The morphological type and the Spectral Energy Distribution (SED) of each galaxy is then defined on the basis of the stellar mass so as to reproduce the observed morphology–density and morphology–radius relations in galaxy clusters (e.g. van der Wel 2008; Ma et al. 2010).

2.1.2 Hera

The mass distribution of the second galaxy cluster, *Hera*, is instead directly derived from a high-resolution *N*-body simulation of a cluster-sized dark matter halo. More precisely, *Hera* is part of the set of simulated clusters presented in Planelles et al. (2014). The cluster halo was first identified in a low-resolution simulation box with a periodic comoving size of 1 h^{-1} Gpc for a flat ΛCDM model with present matter density parameter $\Omega_{\rm m,\,0} = 0.24$ and baryon density parameter $\Omega_{\rm b,\,0} = 0.04$. The Hubble constant adopted was

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

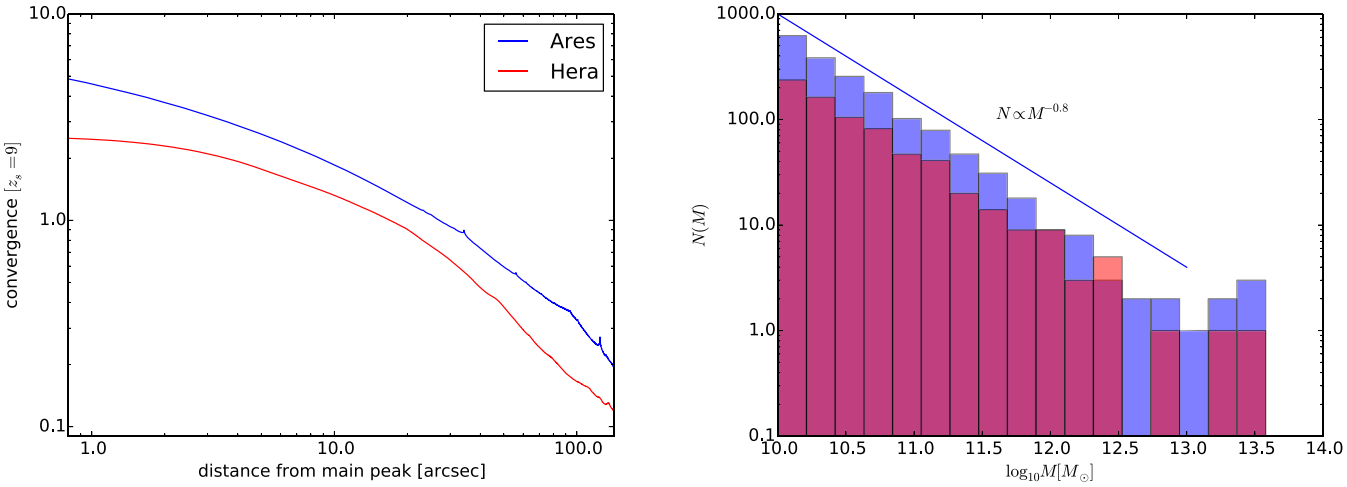

Figure 2. Key properties of *Ares* and *Hera* (blue and red colours, respectively). Left-hand panel: convergence profiles (for source redshift $z_s = 9$). In both cases, the centre has been chosen to coincide with the most massive dark matter clump in the simulation; Right-hand panel: subhalo mass function (built considering all subhaloes within $1\,h^{-1}$Mpc from the centre of the most massive clump.

$H_0 = 72$ km s^{-1} Mpc^{-1} and the normalization of the matter power spectrum $\sigma_8 = 0.8$. The primordial power spectrum of the density fluctuations is $P(k) \propto k^n$ with $n = 0.96$. The parent simulation followed 1024^3 collisionless particles in the box. *Hera* was identified at $z = 0$ using a standard *Friends-of-Friends* (FoF) algorithm, and its Lagrangian region was re-simulated at higher resolution employing the *Zoomed Initial Conditions* code (ZIC; Tormen, Bouchet & White 1997). The resolution is progressively degraded outside this region to save computational time while still providing a correct description of the large-scale tidal field. The Lagrangian region was taken to be large enough to ensure that only high-resolution particles are present within five virial radii of the cluster.

The re-simulation was then carried out using the TreePM-SPH GADGET–3 code, a newer version of the original GADGET–2 code by Springel (2005) that adopted a more efficient domain decomposition to improve the work load balance. Although, the parent *Hera* halo exists in several flavours in various simulation runs (several assumptions for the nature of dark matter particles), including several baryonic processes, the simulation used in this paper uses only the version that utilized collisionless dark matter particles. This has allowed us to increase the mass resolution by about an order of magnitude compared to the hydrodynamical versions of the simulation. The particle mass is $m_{\rm DM} = 10^8\,h^{-1}\,{\rm M}_\odot$. Therefore, the virial region of *Hera* is resolved with ∼10 million particles, with a total cluster mass of $M = 9.4 \times 10^{14}\,h^{-1}\,{\rm M}_\odot$, comparable to that inferred for observed cluster lenses. The redshift of this halo is $z_l = 0.507$. During the re-simulation, the Plummer-equivalent comoving softening length for gravitational force in the high-resolution region is fixed to $\epsilon_{\rm Pl} = 2.3\,h^{-1}$ kpc physical at $z < 2$ while being fixed to $\epsilon_{\rm Pl} = 6.9\,h^{-1}$ kpc comoving at higher redshift.

The properties of cluster galaxies used for creating the simulated observations are derived from SAMs of galaxy formation (De Lucia & Blaizot 2007). The process starts by using the algorithm SUBFIND (Springel et al. 2001) to decompose each FOF group previously found in the simulation into a set of disjoint substructures. These are identified as locally overdense regions in the density field of the background halo. Only substructures that retain at least 20 bound particles after a gravitational unbinding procedure are considered to be genuine substructures. Merging histories are constructed for all self-bound structures, using the same post-processing algorithm that has been employed for the Millennium Simulation (Springel, Frenk & White 2006). The merger-tree is then used to construct a mock catalogue of galaxies. The evolution of the galaxy population is described by a modified version of the semi-analytic model presented in De Lucia & Blaizot (2007), that included the implementation of the generation of intracluster light described in Contini et al. (2014), given by the combination of *Model Tidal Radius* and *Merger* channels presented in that paper.

Note that, even in the case of *Hera*, the galaxy positions trace reasonably well the mass. Several reconstruction methods assume that light traces the mass, a reasonable assumption that is thus satisfied both in *Ares* and in *Hera*. To increase the level of uncertainty, the galaxy shapes and orientations are chosen to be uncorrelated with the underlying mass distribution.

The convergence map of *Hera* with its complex morphology and the abundance of substructures is shown in the right-hand panels of Fig. 1. The convergence profile and the substructure mass function are displayed in Fig. 2. Compared to *Ares*, the small-scale structures of *Hera* are smoother as they are less well resolved. Nevertheless, the substructure mass function scales very similarly with halo mass. As in the case of *Ares*, *Hera* has a bimodal mass distribution. A massive substructure ($M \sim 5 \times 10^{13}\,h^{-1}\,{\rm M}_\odot$) is located ∼30 arcsec (∼$130\,h^{-1}$kpc) from the cluster centre, producing a secondary peak in the convergence map and elongating the iso-density contours in the south-west–north-easterly direction.

2.2 Ray tracing

In order to generate lensing effects in the simulated images, it is necessary to compute the deflections produced by the cluster. This allows us to then use ray-tracing methods to map the surface-brightness distribution of the sources on the camera of our virtual telescope, which is *HST* in this case. In practice, we shoot a bundle of light rays through a dense grid covering the field of view (FOV), starting from the position of the observer. Then, we use the computed deflection angles to trace the path of the light back to the sources. When simulating *HST* observations, we compute the deflection angles on a regular grid of 2048 × 2048 points, covering a FOV of 250 arcsec × 250 arcsec centred on the cluster. The pixel scale of this grid corresponds to 0.12 arcsec or ∼$0.5\,h^{-1}$ kpc.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

In the case of *Ares*, MOKA produces a map of the convergence, $\kappa(\boldsymbol{\theta})$. This can be converted into a map of the deflection angles, $\boldsymbol{\alpha}(\boldsymbol{\theta})$, by solving the equation

$$\alpha(\theta) = \frac{1}{\pi}\int \mathrm{d}^2\theta' \frac{\theta-\theta'}{|\theta-\theta'|}\kappa(\theta') \; . \tag{3}$$

Since this is a convolution of the convergence, $\kappa(\boldsymbol{\theta})$, with the kernel function

$$K(\theta) = \frac{1}{\pi}\frac{\theta}{|\theta|} \; , \tag{4}$$

this task can be achieved numerically by means of fast-Fourier-transform (FFT) methods. To do so, we use the FFT routines implemented in the *gsl* library.

In the case of *Hera*, the mass distribution of the cluster is described by a collection of dark matter particles. Instead of mapping them on a grid to construct the convergence map, we use our consolidated lensing simulation pipeline (see e.g. Meneghetti et al. 2010b, and references therein) to compute the deflections. To briefly summarize, the procedure involves the following steps:

(i) We project the particles belonging to the halo along the desired line of sight on the *lens plane*. To select particles, we define a slice of the simulated volume around the cluster, corresponding to a depth of $10\,h^{-1}$Mpc.

(ii) Starting from the position of the virtual observer, we trace a bundle of light rays through a regular grid of 2048×2048 covering a region around the halo centre on the lens plane. In the case of strong lensing simulations (e.g. for *HST* observations), we restrict our analysis to a region of $1 \times 1\,h^{-2}$Mpc^2. The pixel scale of this grid is 0.11 arcsec, corresponding to $\sim 0.47\,h^{-1}$ kpc. In the case of simulations extending into the weak-lensing regime (e.g. for Subaru-like observations), the grid of light rays covers a much wider area ($\sim 8 \times 8\,h^{-2}$ Mpc^2). In this case, we use a grid of 4096×4096 rays, whose resolution is 0.44 arcsec pixel^{-1}, corresponding to $\sim 1.87\,h^{-1}$ kpc pixel^{-1}.

(iii) Using our code GLFAST (Meneghetti et al. 2010a), we compute the total deflection $\alpha(\boldsymbol{x})$ at each light-ray position $\boldsymbol{x}$, accounting for the contributions from all particles on the lens plane. Even in the case of strong-lensing simulations, where light rays are shot through a narrower region of the lens plane, the deflections account for all particles projected out to $\sim 4\,h^{-1}$Mpc from the cluster centre. The code is based on a Tree algorithm, where the contributions to the deflection angle of a light ray by the nearby particles are summed directly, while those from distant particles are calculated using higher order Taylor expansions of the deflection potential around the light-ray positions.

(iv) The resulting deflection field is used to derive several relevant lensing quantities. In particular, we use the spatial derivatives of $\boldsymbol{\alpha}(\boldsymbol{\theta})$ to construct the shear maps, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)$, defined as

$$\gamma_1(\theta) = \frac{1}{2}\left(\frac{\partial\alpha_1}{\partial\theta_1} - \frac{\partial\alpha_2}{\partial\theta_2}\right) \; , \tag{5}$$

$$\gamma_2(\theta) = \frac{\partial\alpha_1}{\partial\theta_2} = \frac{\partial\alpha_2}{\partial\theta_1} \; . \tag{6}$$

The convergence, $\kappa(\boldsymbol{\theta})$, may also be reconstructed as

$$\kappa(\theta) = \frac{1}{2}\left(\frac{\partial\alpha_1}{\partial\theta_1} + \frac{\partial\alpha_2}{\partial\theta_2}\right) \; . \tag{7}$$

The lensing critical lines yield formally infinite magnification for a given source redshift. They are defined as the curves along which the determinant of the lensing Jacobian is zero (e.g. Schneider, Ehlers & Falco 1992):

$$\det A = (1-\kappa-|\gamma|)(1-\kappa+|\gamma|) = 0 \; . \tag{8}$$

In particular, the *tangential* critical line is defined by the condition $(1-\kappa-|\gamma|)=0$, whereas the *radial* critical line corresponds to the line along which $(1-\kappa+|\gamma|)=0$. In the following sections, we will often use the term *Einstein radius* to refer to the size of the tangential critical line. As discussed in Meneghetti et al. (2013), there are several possible definitions for the Einstein radius. Here, we adopt the *effective* Einstein radius definition (see also Redlich et al. 2012) given by

$$\theta_{\mathrm{E}} \equiv \frac{1}{d_{\mathrm{L}}}\sqrt{\frac{S}{\pi}} \; , \tag{9}$$

where S is the area enclosed by the tangential critical line and d_{L} is the angular diameter distance to the lens plane.

2.3 SKYLENS

We simulate observations of galaxy cluster fields using the code SKYLENS, which is described in detail in Meneghetti et al. (2008) and in Meneghetti et al. (2010a). The creation of the simulated images involves the following steps:

(i) we generate a past light-cone populated with source galaxies resembling the luminosity and the redshift distribution of the galaxies in the Hubble Ultra-Deep-Field (HUDF; Coe et al. 2006);

(ii) we model the morphologies of the sources using shapelet decompositions of the galaxies in the HUDF (Melchior, Meneghetti & Bartelmann 2007). Their spectral energy distributions were obtained as part of the photometric redshift measurements of these galaxies described in Coe et al. (2006);

(iii) the deflection fields of the lensing clusters are used to trace a bundle of rays from a virtual CCD, resembling the properties of the ACS or of the WFC3, back to the sources;

(iv) by associating each pixel of the virtual CCD to the emitting elements of the sources, we reconstruct their lensed surface brightness distributions on the CCD;

(v) we model the shape and the surface brightness distribution of the cluster galaxies using elliptical single or double Sérsic models (Sérsic 1963). These are obtained by fitting real cluster galaxies in a set of low to intermediate redshift clusters (Gonzalez, Zabludoff & Zaritsky 2005). The match between the observed and the simulated galaxies is done via the galaxy luminosity. The BCGs all include a large-scale component used to model the intracluster light produced by the BCG stellar haloes and by free-floating stars;

(vi) the SEDs of the cluster galaxies are modelled according to prescriptions from semi-analytic models or from the HOD technique, as explained earlier;

(vii) we convert the surface brightness distributions into counts per pixel assuming a telescope throughput curve, which accounts for the optics, the camera and the filter used in carrying out the simulated observations. In each band, we simulate the exposure times (in units of *HST* orbits[4]) used to carry out the mock Frontier Fields observations;

(viii) the images are then convolved with a point spread function (PSF) model, obtained using the TINYTIM *HST* PSF modelling software (Krist, Hook & Stoehr 2011). Finally, realistic noise is added

[4] We assume an orbital visibility period of 2500 s.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

Table 1. Details of the simulated observations. First column: filters employed; second column: exposure time in terms of *HST* orbits; third column: level of the background (counts s^{-1} in circles with radius 0.2 arcsec)

Filter	No. of orbits	bkg level (10^6 counts s^{-1})
F435W	18	0.615
F606W	10	1.821
F814W	42	2.227
F105W	24	1.136
F125W	12	1.122
F140W	10	1.371
F160W	24	0.959

mimicking the appropriate sky surface brightness in the simulated bands. The noise is assumed to have a Poisson distribution, and it is calculated according to the Meneghetti et al. (2008) equation 31, assuming a stack of multiple exposures, with the number varying from band to band.[5]

2.4 Images and catalogues

This is the first phase of a comparison project, and in the next phase we intend to include additional simulations with an even greater level of realism. For this first exercise, we proceed as follows.

(i) For both *Ares* and *Hera*, we generate simulated *HST* observations in all bands that are deployed for the FFI, mimicking the same exposure times (or number of orbits) as the real observations. The level of the background is set to values provided by the ACS and WFC3 exposure time calculators in each band. The details of these simulations are provided in Table 1. Each image covers a FOV of 204×204 arcsec^2. All images are co-aligned and corotated. Effects like gaps between chips, pixel defects, charge transfer inefficiency, cosmic rays, etc. are not included. The resolution of the ACS and WFC3 simulations are 0.05 and 0.13 arcsec pixel^{-1}, respectively. These images were made available to the modellers.

(ii) In addition to the images, we provided the list of all multiple images obtained from the ray-tracing procedure (see Figs 3 and 4, upper panels). Each multiple-image system is characterized by the redshift of its source, which is also provided to the modellers. Thus, in this exercise we assume that all images can be identified without errors and that all their redshifts can be measured 'spectroscopically'. This is certainly a very optimistic assumption that will never be satisfied in the real world. In the next round of this comparison project, the assumption will be relaxed, but for the moment we decided to release this information because our objective is to determine possible systematics of the various reconstruction algorithms. Other issues related to the approaches used to search for multiple images or the impact of redshift uncertainties on the results will be studied in a future work. Some of these systematics have already been investigated for some lens modelling methods, i.e. Johnson & Sharon (2016).

(iii) We also released a catalogue of cluster members (circled in the bottom-right panels of Figs 3 and 4), containing positions and photometry in all bands. Several reconstruction methods (in particular those employing the parametric approach) build the lens model by combining smooth dark matter haloes with substructures associated with the cluster members akin to our construction of *Ares*. In this simplified test, modellers are provided with the list of all cluster members with $m_{AB,\,F814W} < 24$. Again, this is an oversimplification that will be removed in the next round of simulations, and which implicitly favours those methods that use this information. In reality, such methods have to deal with the risks of misidentification of cluster members.

(iv) For those groups that use weak lensing measurements to complement the strong lensing constraints, we produced a single Subaru-like *R*-band image of both *Ares* and *Hera* covering a much larger FOV of 30×30 arcmin^2. The provided image contained only background galaxies (i.e. lensed by the clusters) and stars, so that shape measurements could be made using any weak-lensing pipeline without worrying about the separation of background sources from the cluster members or contamination by foreground galaxies. We also use the publicly available pipeline KSBf90[6] (Heymans et al. 2006) based on the Kaiser, Squires and Broadhurst method (Kaiser, Squires & Broadhurst 1995) to derive a catalogue containing galaxy positions and ellipticities. The resulting number density of galaxies useful for the weak lensing analysis is $\sim$14 galaxies arcmin^{-2}. This is significantly smaller than the number density achievable with *HST*.

All these data for the mock cluster lenses were shared with lens modellers participating in the project via a dedicated website.[7] We emphasize that the input mass distributions of the lenses and the techniques used to generate them were initially kept blinded to all groups. The strong lensing constraints amounted to 242 multiple images produced by 85 sources in the case of *Ares* and 65 images of 19 sources in the case of *Hera*.

3 LENS MODELLING TECHNIQUES

3.1 Submission of the models

A large fraction of lens modellers currently working actively on the analysis of the FFI data accepted the challenge and participated in this project. The two cluster simulations were not released simultaneously. We initially released only the data for *Ares*, and we received reconstructed models for this cluster from seven groups. These groups performed a fully blind analysis of the data set. Two additional models were submitted by A. Zitrin after the input mass distributions were already revealed, under the assurance that the reconstruction was actually performed blindly.

In a second stage of the comparison exercise, we released the simulation of *Hera*, and received eight models from six participating groups. Also for this cluster, we received additional reconstructions after we revealed the input mass distribution of the lens. These models were submitted by A. Zitrin and by D. Lam.

There are two general classes of inversion algorithms. They comprise *parametric* models wherein the mass distribution is reconstructed by combining clumps of matter, often positioned where the BCGs are located, each of which is characterized by an ensemble of parameters including the density profile and shape. The parameter spaces of these models are explored in an effort to best reproduce the observed positions, shapes and magnitudes of the multiple images and arcs. The second approach is called *free form* (a.k.a. non-parametric): wherein now the cluster is subdivided into a mesh on to which the lensing observables are mapped, and which

[5] Two exposures per orbit per filter.

[6] http://www.roe.ac.uk/~heymans/KSBf90/Home.html
[7] http://pico.bo.astro.it/~massimo/Public/FF

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

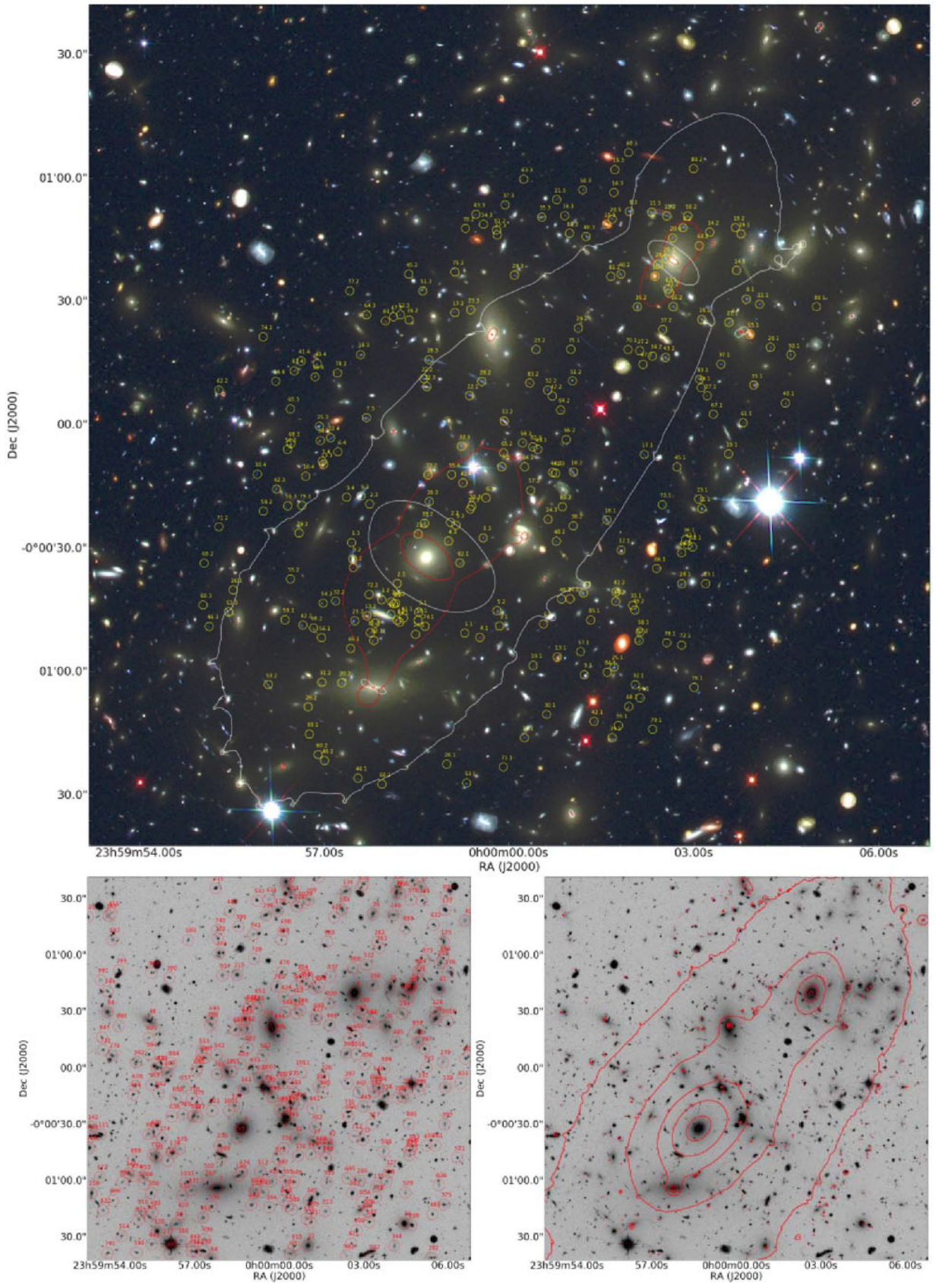

Figure 3. Colour composite image of *Ares* (upper panel). We overlay to the optical image the critical lines for $z_s = 1$ (red) and $z_s = 9$ (white). In addition, we display the location of the multiple image systems (numbered yellow circles). In the bottom panels, we show the image of the cluster in the *F*814*W* band. On the left, the galaxies identified as cluster members are indicated by red circles. On the right, we display the surface density iso-contours.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

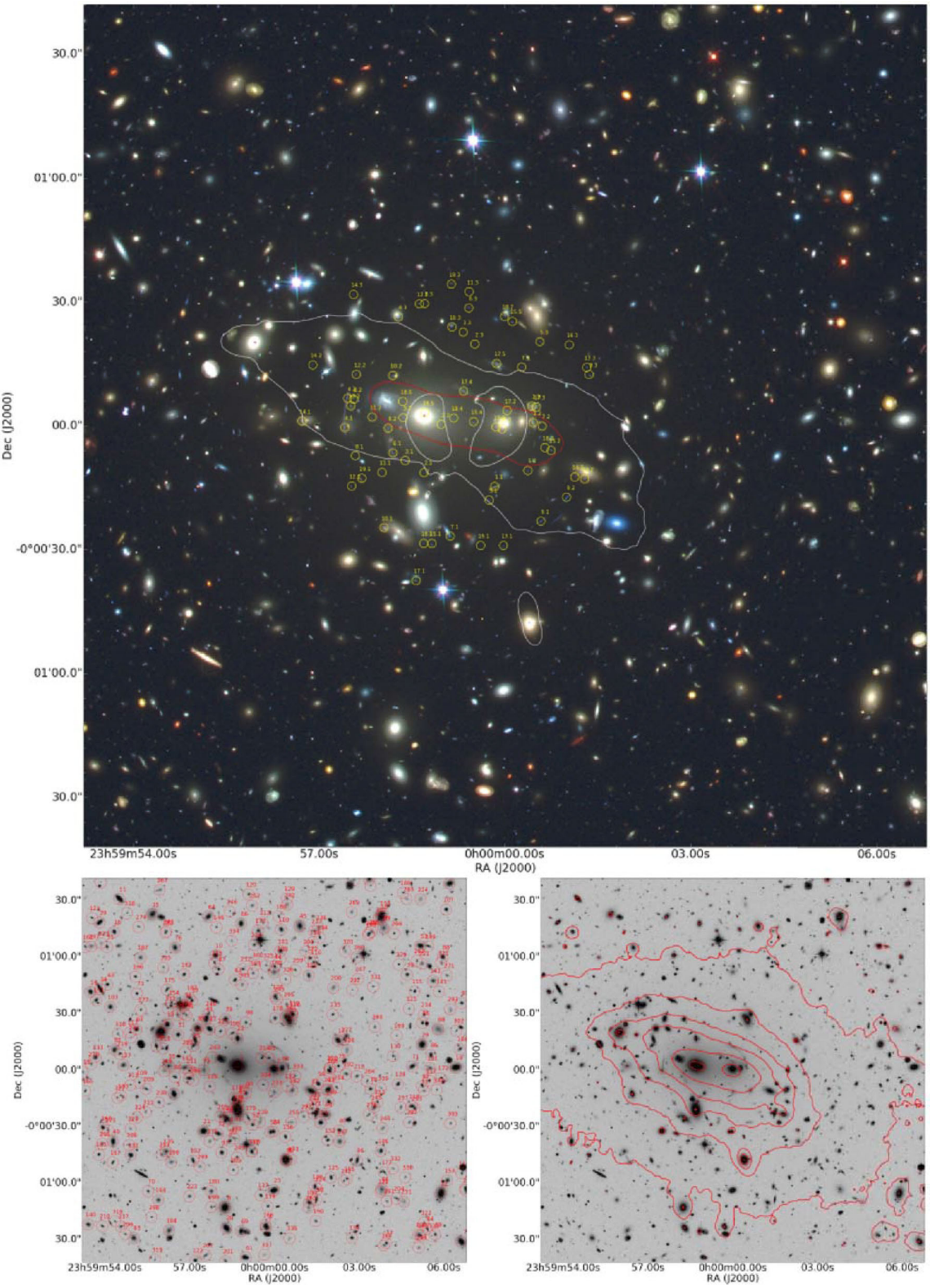

Figure 4. As in Fig. 3, but for the cluster *Hera*.

is then transformed into a pixelized mass distribution following several methods to link the observable to the underlying lens potential or deflection field.

Both these approaches were amply represented in the challenge. A summary of all submitted models, with the indication of whether they are parametric or free form and built before or after the input mass distribution of the lenses was revealed, is given in Table 2. Each model is given a reference name used throughout the paper. Each modelling technique is briefly described below. In Table 3, we summarize the details of each reconstruction of *Ares* and *Hera*.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

Table 2. Models submitted by the groups participating in the project. The table lists the name of the submitting group/author of the reconstruction, the reference name of the model, the type of algorithm (free-form, parametric or hybrid) and whether the model was submitted blind, i.e. before the input mass distribution of the lens was revealed.

Group/Author	Method	Model	Cluster	Approach	Blind
M. Bradac and A. Hoag	SWUnited	Bradac-Hoag	*Ares+Hera*	Free-form	Yes
J. Diego	WSLAP+	Diego-multires	*Hera*	Hybrid	Yes
J. Diego	WSLAP+	Diego-overfit	*Hera*	Hybrid	Yes
J. Diego	WSLAP+	Diego-reggrid	*Ares+Hera*	Hybrid	Yes
D. Lam	WSLAP+	Lam	*Hera*	Hybrid	No
J. Liesenborgs, K. Sebesta & L. Williams	GRALE	GRALE	*Ares+Hera*	Free-form	Yes
D. Coe	LensPerfect	Coe	*Ares*	Free-form	Yes
CATS	LENSTOOL	CATS	*Ares+Hera*	Parametric	Yes
T. Johnson and K. Sharon	LENSTOOL	Johnson-Sharon	*Ares+Hera*	Parametric	Yes
T. Ishigaki, R. Kawamata and M. Oguri	GLAFIC	GLAFIC	*Ares+Hera*	Parametric	Yes
A. Zitrin	LTM	Zitrin-LTM-gauss	*Ares+Hera*	Parametric	No
A. Zitrin	PIEMDeNFW	Zitrin-NFW	*Ares+Hera*	Parametric	No

3.2 SWUnited: The Bradac-Hoag model

The Bradac-Hoag model employs the method named SWUnited: Strong and weak lensing mass reconstruction on a non-uniform adapted grid. This combined strong and weak lensing analysis method reconstructs the gravitational potential $\psi_k = \psi(\boldsymbol{\theta}^k)$ on a set of points $\boldsymbol{\theta}^k$, which can be randomly distributed over the entire FOV. From the potential, any desired gravitational lensing quantity (e.g. surface mass density, deflection angle, magnification, flexion, etc.) can be readily calculated. Such an approach therefore does not require an assumption of e.g. a particular model of the potential/mass distribution. The potential is reconstructed by maximizing a log-likelihood $\log P$ (see equation 2 of Bradač et al. 2009), which uses image positions of multiply imaged sources and source plane minimization (corrected by magnification); weak lensing ellipticities, and regularization as constraints. Current implementation also includes flexion measurements, however, the data was not used in this paper.

3.2.1 Description of the method

The implementation of the method follows the algorithm first proposed by Bartelmann et al. (1996) and is described in detail in Bradač et al. (2005) and Bradač et al. (2009). It starts with a uniform grid of points θ_k at coarse resolution (25–30 pixels on a side) and refine it by two to four times in progressively smaller regions enclosing the BCGs and by eight times around multiple images, making sure there are no sharp transitions in the density of grid points anywhere. This generally results in a grid that is coarse in the outskirts of the cluster and fine in the central regions where the multiple images exist. These are the points where the potential is evaluated. From the set of potential values all observables are determined using derivatives. For example, the convergence κ is related to ψ via the Poisson equation, $2\kappa = \nabla^2\psi$ (where the physical surface mass density is $\Sigma = \kappa\ \Sigma_{\rm crit}$ and $\Sigma_{\rm crit}$ depends upon the angular diameter distances between the observer, the lens and the source). The details on how the derivatives on a non-uniform grid are evaluated can be found in Bradač et al. (2009). By using a reconstruction grid whose pixel scale varies across the field, the method is able to achieve increased resolution in the cluster centre (close to where we see strongly lensed images), and hence the magnification map in the regions of high magnification is more detailed. The posterior peak values of the potential ψ_k are found by solving the non-linear equation $\partial \log P/\partial\psi_k = 0$. This set of equations is linearized and a solution is reached in an iterative fashion (keeping the non-linear terms fixed at each iteration step). This requires an initial guess for the gravitational potential; the systematic effects arising from various choices of this initial model were discussed in Bradač et al. (2006). The choice of particular grid geometry, the regularization parameter and the hyper-parameters that set the relative weighting between the contributions to $\log P$ all become critical when weak lensing data on large scales ($\gtrsim$ 1Mpc) are included, and a full-field mass reconstruction is needed. This is not the case in this work, as we are only interested in the magnification of the inner region.

The reconstruction is performed in a two-level iteration process, outlined in Fig. 5. The inner-level iteration process described above for solving the non-linear system of equations $\partial \log P/\partial\psi_k = 0$ is solved in iterative fashion and repeated until convergence of κ. The outer-level iteration is performed for the purpose of regularization. In order to penalize small-scale fluctuations in the surface mass density, the reconstruction is started with a coarse grid (large cell size). Then, for each n_2 step the number of grid points is increased in the field and the new reconstructed $\kappa^{(n_2)}$ is compared with the one from the previous iteration $\kappa^{(n_2-1)}$ (or with the initial input value $\kappa^{(0)}$ for $n_2 = 0$), penalizing any large deviations. The second-level iterations are performed until the final grid size is reached and convergence is achieved.

Regularization is used. It is adjusted so that the model converges and the final χ^2 for weak lensing is $\sim 2N_{\rm weak}$ (number of weak lensing galaxies). If the regularization term is too small, the method fits to the noise and does not converge to a single solution.

3.2.2 Strengths and weaknesses of the method

The main strength of the method as discussed above is that instead of fitting a specific set of family of models to the data, the method is free of such an assumption. Furthermore, the positions of the points where the potential is reconstructed ($\boldsymbol{\theta}^k$) can be chosen arbitrarily, which allows us to increase the density of points in the regions where signal to noise is the highest (i.e. where multiple images are present). A coarser sampling can be employed in the areas where this is not the case (e.g. at large radii from the centre). The algorithm reconstructs the potential (rather than traditionally used surface mass density), since it locally determines both the lensing distortion (for weak lensing and flexion) as well as the deflection (for strong lensing) and there is no need to assume the surface mass density values beyond the observed field.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

Table 3. Table summarizing the main properties of the reconstructions of *Ares* and *Hera*. The columns correspond to (1) name of the model; (2) name of the cluster; (3) plane where the optimization is performed (source or image plane); (4) profile used to model the dark matter halo components; (5) number of dark matter haloes assumed; (6) type of constraints used (SL, WL); (7) number of multiple images used as constraints; (8) number of additional constraints (e.g. image ellipticity, flux, etc.); (9) number of free parameters. In the case of free-form methods, the number of grid points of the reconstructed map is given; (10) density profile used to model the cluster galaxies; (11) shape of the cluster galaxies (circular or elliptical); (12) scaling relations used to model the ensemble of cluster galaxies; (13) number of free parameters used to model the cluster galaxies; (14) catalogue of cluster galaxies used; (15) number of cluster galaxies optimized individually (i.e. independently on the scaling relations); (16) external shear (yes or no); (17) high-order perturbations (yes or no) and (18) regularization term (yes or no).

Model	Cluster	Opt. plane	Halo prof.	No. of haloes	Type of constr.	Mult. images	Other constr.	No. of free par.	Gal. profiles	Gal. shape	Scaling rel.	Free par.	Gal. catalogue	Opt. galaxies	Ext. shear	High-order pert.	Reg.
Bradac-Hoag	*Ares*	Image	–	–	SL+WL	242	574	1959*(a)*	–	–	–	–	–	–	No	No	Yes
Bradac-Hoag	*Hera*	Image	–	–	SL+WL	65	2102	5497	–	–	–	–	–	–	No	No	Yes
Diego-multires	*Hera*	Source	Gaussians*(b)*	–	SL	65	161*(c)*	304	LTM	ell.	$M/L = \text{const}$	3	Official	2	No	No	No
Diego-overfit	*Hera*	Source	Gaussians	–	SL	65	161	1027	LTM	ell.	$M/L = \text{const}$	3	Official	2	No	No	No
Diego-reggrid	*Ares*	Source	Gaussians	–	SL	242	643	1027	LTM	ell.	$M/L = \text{const}$	3	Official	2	No	No	No
Diego-reggrid	*Hera*	Source	Gaussians	–	SL	65	161	1027	LTM	ell.	$M/L = \text{const}$	3	Official	2	No	No	No
Lam	*Hera*	Source	Gaussians	–	SL	65	–	1027	NFW	circ.	$M \propto M_\star$	20	Official	10	No	No	No
GRALE	*Ares*	Source	Plummer*(d)*	–	SL	183*(e)*	–	>600*(f)*	–	–	–	–	–	–	No	No	No
GRALE	*Hera*	Source	Plummer	–	SL	65	–	>600	–	–	–	–	–	–	No	No	No
Coe	*Ares*	Image	RBFs*(g)*	–	SL	242	–	170	–	–	–	–	–	–	No	No	Yes
CATS	*Ares*	Source	PIEMD	2	SL	242	–	35	PIEMD	circ.	$\sigma \propto L^{1/4}\, r_t \propto L^{1/2}$	2	CATS Ares*(h)*	5	No	No	No
CATS	*Hera*	Source	PIEMD	2	SL	65	–	34	PIEMD	circ.	$\sigma \propto L^{1/4}\, r_t \propto L^{1/2}$	2	Official	6	No	No	No
Johnson-Sharon	*Ares*	Source	PIEMD	2	SL	242	–	18	PIEMD	ell.	$\sigma \propto L^{1/4}\, r_t \propto L^{1/2}$	1	Official	2	No	No	No
Johnson-Sharon	*Hera*	Source	PIEMD	2	SL	65	–	13	PIEMD	ell.	$\sigma \propto L^{1/4}\, r_t \propto L^{1/2}$	1	Official	2	No	No	No
GLAFIC	*Ares*	Source	eNFW	5	SL	242	–	32	PIEMD	ell.	$\sigma \propto L^{1/4}\, r_t \propto L^{1/2}$	2	GLAFIC*(i)*	–	No	No	No
GLAFIC	*Hera*	Source	eNFW	2	SL	65	–	23	PIEMD+Hernquist*(j)*	ell.	$\sigma \propto L^{1/4}\, r_t \propto L^{\eta}$	3	GLAFIC	2	Yes	Yes	No
Zitrin-LTM-gauss	*Ares*	Image	LTM	–	SL	242	–	19	LTM	circ.	Same LTM scaling	12	Official	4	Yes	No	No
Zitrin-LTM-gauss	*Hera*	Image	LTM	–	SL	65	–	17	LTM	circ.	Same LTM scaling	10	Official	4	Yes	No	No
Zitrin-NFW	*Ares*	Image	eNFW	2	SL	242	–	25	PIEMD	circ.	$\sigma \propto L^{1/4}\, r_t \propto L^{1/2}$	13	Official	4	No	No	No
Zitrin-NFW	*Hera*	Image	eNFW	3	SL	65	–	24	PIEMD	circ.	$\sigma \propto L^{1/4}\, r_t \propto L^{1/2}$	11	Official	4	No	No	No

Notes. (a) This number corresponds to the total number of grid points at the end of the refinements in the map of the lensing potential. Note that these points are highly correlated, as the lensing constraints measure higher order derivatives of the potential; *(b)* The diffuse, halo component of the cluster mass distribution in the Diego and Lam models is fitted with a combination of Gaussians on a regular or adaptive grid; *(c)* In the Diego and Lam models, constraints other than image positions are derived from the sizes and shapes of the spatially resolved images; *(d)* The cluster mass distribution in the GRALE models is given by a combination of Plummer spheres; *(e)* Systems composed by two images only were not used in the reconstruction; *(f)* The number of free parameters depends on the number of Plummer spheres effectively used. This is not known a priori as it is determined by the genetic algorithm during each run. The starting number of Plummer spheres is 600–1000 depending on the run; *(g)* The cluster mass distribution is described by a combination of RBFs whose form is given in equations 17, 20 and 21 of Coe et al. (2008); *(h)* The CATS *Ares* catalogue of cluster members is obtained by selecting the galaxies with $m_{F160W} < 22$; *(i)* Positions and luminosities were taken from the official catalogue of cluster members distributed to all teams. Ellipticity and position angle of each galaxy were measured using SEXTRACTOR; *(j)* The two brightest galaxies are modelled with Hernquist profiles and optimized individually. The other galaxies are modelled with PIEMDs and their masses and sizes scale with luminosity.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

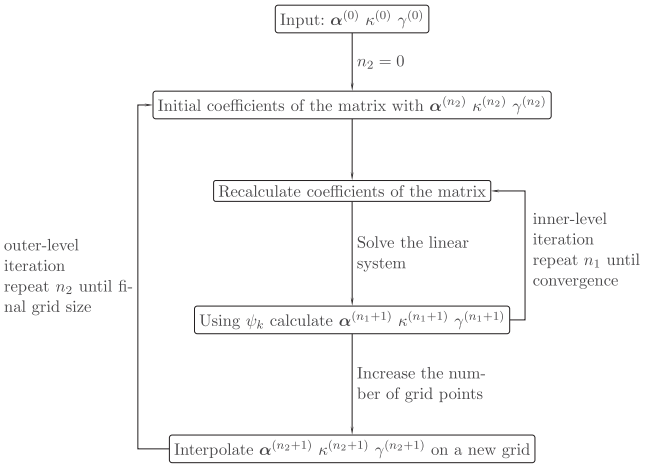

Figure 5. The outline of the method used to build the Bradac-Hoag model and the two-level iteration process.

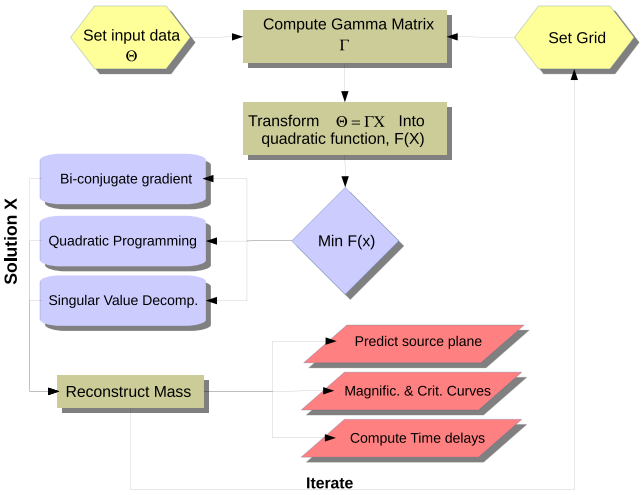

Figure 6. Diagram showing the work flow of WSLAP+.

The main weakness of the method on the other hand is the fact that it maximizes a function with a large number of parameters. Therefore, the method is inherently unstable. The inversion of the matrix satisfying the equation $\partial \log P / \partial \psi_k = 0$ is also very noisy. The method is therefore very likely to diverge or land in a secondary minimum. Regularization needs to be employed, which adds additional parameters (relative weighting of regularization term) to the rest of $\log P$ and a choice of regularization method itself. The optimal choices need to be determined using simulation data.

3.2.3 Improvements in progress

A recent improvement to the method is the addition of the measurement of flexion to the input constraints. The code has been adapted (see also Cain, Bradac & Levinson 2016) and tested on simulated data. This improvement is currently being tested using *HST* data. In the future, the code will be ported in PYTHON, to make the interface user friendly, and released to the community.

3.3 WSLAP+: the Diego and the Lam models

All Diego models (Diego-multires, Diego-overfit and Diego-reggrid models) and the Lam model are built using WSLAP+, a free-form or non-parametric method that includes also a compact mass component associated with the cluster members (thus, classified as hybrid in this paper). The main part of the code is written in FORTRAN and compiles with standard compilers (like gfortran) included in the most common linux distributions. Plotting routines written in IDL are available to display the intermediate results as the code runs. A script interface allows the user to define the input and output files, select the parts of the code to be run and control the plotting routines. A detailed description of the code and of its features can be found in the papers by Diego et al. (2005), Diego et al. (2007), Sendra et al. (2014) and Diego et al. (2016). The code is not publicly available yet but a companion code LENSEXPLORER is available. LENSEXPLORER allows the user to easily explore the lens models derived for the Frontier Fields clusters, search for new counter images, compute magnifications or predict the location and shape of multiple images. The WSLAP+ code is still under development. It already includes certain features that were not used in the analysis presented in this paper. They will be included in the future 'unblinded' version of this work. Among these features, the code incorporates spatial information about knots in resolved systems greatly improving the accuracy and robustness of the results (see Diego et al. 2016, for practical demonstration). In this work, only long elongated arcs were considered as resolved systems. Multiple knots identified in these arcs are assumed to originate from a very compact region in the source plane. For the shorter arcs, only the central knot was used.

3.3.1 Description of the method

A diagram describing the steps involved in the analysis performed by WSLAP+ is shown in Fig. 6. The algorithm divides the mass distribution in the lens plane into two components. The first is a compact one and is associated with the member galaxies. The member galaxies are selected from the red sequence. The second component is diffuse and is distributed as a superposition of Gaussians on a regular (or adaptive) grid. For the compact component, the mass associated with the galaxies is assumed to be proportional to their luminosity. If all the galaxies are assumed to have the same mass-to-light (M/L) ratio, the compact component (galaxies) contributes with just one ($N_g = 1$) extra free parameter that corresponds to the correction that needs to be applied to the fiducial M/L ratio. In some particular cases, some galaxies (like the BCG or massive galaxies very close to an arclet) are allowed to have their own M/L ratio adding additional free parameters to the lens model but typically no more than a few ($N_g \sim \mathrm{O}(1)$). For this component associated with the galaxies, the total mass is assumed to follow either an NFW profile (with fixed concentration and scale radius scaling with the fiducial halo mass) or be proportional to the observed surface brightness. The diffuse or 'soft' component is described by as many free parameters as grid (or cell) points. This number (N_c) varies but is typically between a few hundred to one thousand ($N_c \sim$ O(100)–O(1000)) depending on the resolution and/or use of the adaptive grid. The size of the adaptive grid, when used, is based on a solution obtained with the regular grid. Areas with more mass result in smaller grid cells. In addition to the free parameters describing the lens model, the method includes as unknowns the original positions of the lensed galaxies in the source plane. For the clusters included in the FFI program the number of background sources, N_s, is typically a few tens ($N_s \sim \mathrm{O}(10)$), each contributing with two unknowns (β_x and β_y). All the unknowns are then combined into a single array X with N_x elements ($N_x \sim \mathrm{O}(1000)$).

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

The observables are both strong lensing and weak lensing (shear) measurements. For strong lensing data, the inputs are the pixel positions of the strongly lensed galaxies (not just the centroids). In the case of long elongated arcs near the critical curves with no features, the entire arc is mapped and included as a constraint. If the arclets have individual features, these can be incorporated as semi-independent constraints but with the added condition that they need to form the same source in the source plane. Incorporating this information acts as an *anchor* constraining the range of possible solutions and reducing the risk of a bias due to the minimization being carried in the source plane. For the weak lensing, we use shear measurements (γ_1 and γ_2). The weak lensing constraints normally complement the lack of strong lensing constraints beyond the central region allowing for a mass reconstruction on a wider scale. When weak lensing information is used, the code typically uses an adaptive grid to extend the range up to the larger distances covered by the weak lensing data (Diego et al. 2015a)

The solution, X, which is an array containing the unknowns in the problem (the mass in each grid cell, the re-scaling factor of the assumed light-to-mass ratio for the member galaxies and the positions of the lensed galaxies in the source plane) is obtained after solving the system of linear equations

$$\Theta = \Gamma X, \tag{10}$$

where the N_o observations (strong lensing, weak lensing, time delays) are included in the array Θ and the matrix Γ is known and has dimension $N_o x(N_c + N_g + 2N_s)$. This matrix contains the deflection field and shear components (if weak lensing data is available) produced by each grid cell with a fiducial mass and the fiducial distribution of member galaxies at each observed location of an arc. In practice, X, is obtained by solving the set of linear equations described in equation (10) via a fast bi-conjugate algorithm, or inverted with a singular value decomposition (after setting a threshold for the eigenvalues) or solved with a more robust quadratic algorithm (slower). The quadratic algorithm is the preferred method as it imposes the physical constrain that the solution X must be positive. This eliminates un-physical solutions with negative masses and reduces the space of possible solutions. Errors in the solution are derived by minimizing the quadratic function multiple times, after varying the initial conditions of the minimization process, and/or modifying the grid, and/or changing the fiducial deflection field associated with the member galaxies.

3.3.2 *Strengths and weaknesses of the method*

The code implements a free-form modelling component. This implies that no strong assumptions are necessary about the distribution of dark matter. This is particularly useful if DM is not linked to the galaxies or if the baryons are also dissociated from the galaxies. The latter seems to be the case in the FFI clusters which are in a merging phase. Evidence that the solution obtained by the algorithm may be sensitive to the mass of the X-ray emitting plasma was presented in Lam et al. (2014) and Diego et al. (2015b,c, 2016).

The algorithm is very fast. Several methods are implemented to search for a solution. Using the bi-conjugate gradient algorithm a solution can be obtained in seconds. Using the slower, but more reliable, quadratic optimization approaches a robust solution that can be obtained in minutes. Other fast approaches have been implemented as well like singular value decomposition.

An adaptive grid can be used that transforms the method into a multiresolution code. Different adaptive grids can be implemented that introduce a small degree of freedom but also allows us to explore other possible solutions and hence constrain better the variability of the solution.

The code is prepared to combine weak and strong lensing. The relative weight of the two data sets is given by the intrinsic errors in the data sets (typically small in the strong lensing regime and larger in the weak lensing regime). Correlations between the lensing data can be incorporated through a covariance matrix that naturally weights the different data sets.

The minimization is made in the source plane which may result in biases towards larger magnifications (e.g. Kochanek 1991). To avoid this, the minimization algorithm needs to be stopped after a given number of iterations. Even better, including information about the size and shape of the sources in the source plane seems to solve this problem and the solution remains stable and unbiased even after a very large number of iterations. These *prior* information on the size and shape of the source galaxies is only possible when well-resolved lensed images are available and at least one of the multiple images is not highly magnified.

The compact component is *pixelized* usually into a 512×512 image that covers the FOV. For the small member galaxies, this pixelization results in a loss of resolution that have a small impact on lensed images that happen to be located near this small member galaxies. A possible solution to alleviate this problem is to pre-compute the deflection field of these galaxies prior to the minimization at higher resolution and later interpolate at the position of the observed lensed galaxies. This approach has not been implemented yet but it is expected to eliminate this problem.

In the *Lam* model of *Hera*, the cluster members are modelled using analytical mass profiles (NFW). In this case, there is no need to produce a pixelated map of the deflection angles. Instead, these are computed analytically.

A weakness of this approach is that the code can also predict more multiple images than observed. This is not being factored in at the moment but will be the subject of the *null space* implementation described in Section 3.3.3. One systematic bias is known to affect the results at large distances from the centre. The reconstructed solution systematically underpredicts the mass (and magnification) in the regions where there is no lensing constraints. These regions are normally located beyond the corresponding Einstein radius for a high-redshift background source. Addition of weak lensing to then constraints can reduce or eliminate this problem.

3.3.3 *Improvements in progress*

The addition of time delays is being implemented to the reconstruction of the solution. Time delays will be included in a similar footing as the other observables (weak and strong lensing observables) with a weight that is proportional to their associated observational error.

The addition of the *null space* was proven to be a useful and powerful way of improving the robustness of the derived solution (Diego et al. 2005). This direction has not been explored fully and we plan to incorporate the *null space* as an additional constraint. This will eliminate additional counter images that are predicted by the model but not observed in the data.

3.3.4 *Modelling of* Ares *and* Hera

The Diego models use both a regular grid with $32 \times 32 = 1024$ grid points (Diego-reggrid model) and a multiresolution grid with approximately half the number of grid points (Diego-multires model).

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

A comparison between these grids is shown in Diego et al. (2016). The compact component of the defection field is constructed based on the brightest elliptical galaxies in the cluster. We include 50 such bright ellipticals for each cluster. The mass profile for each galaxy is taken either as an NFW with scale radius (and total mass) scaling with the galaxy luminosity or directly as the observed surface brightness. This choice plays a small role in the final solution.

Depending on the number of iterations, different solutions can be obtained. Earlier work based on simple simulations (Sendra et al. 2014) showed how in a typical situation (similar to the one in *Ares* and *Hera*), after 10 000 iterations of the code, the solution converges to a stable point. The code can be left iterating longer reaching a point that we refer as 'overfit' where the observed constraints are reproduced with great accuracy but sometimes at the expense of a model with fake structures. In the case of Hera, we computed the solution also in the overfit regime (90 000 iterations) for comparison purposes (Diego-overfit model). As explained in Sendra et al. (2014), the compact component helps also in regularizing the solution since it acts as an 'anchor' for the entire solution.

The Lam model differs from the Diego models as follows. A regular grid of Gaussian functions is used instead of a multiresolution grid (as in Diego-multires). The code runs a minimization algorithm. In each step, a new solution is obtained and updated in an iterative way. The algorithm starts with an initial guess (a random realization of the iterative solution) and the minimization converges to a different solution each time. By varying the initial condition 100 times, 100 different lens models were obtained that are equally consistent with the data. The submitted model is an average of these 100 individual models. With the exception of the 10 BCGs, the relative masses of all cluster galaxies are fixed, and are derived using a stellar mass–dark matter mass relation found in the EAGLE cosmological hydrodynamical simulation (Schaller et al. 2015). The stellar masses of cluster galaxies are derived from fitting synthesized spectra to the measured photometry using FAST (Kriek et al. 2009). The contribution from cluster galaxies are parametrized by NFW haloes with scale radii derived from the dark matter mass using again a relation found in the same simulation.

3.4 GRALE: the GRALE model

The GRALE models are based on the reconstruction code GRALE.[8] GRALE is a flexible, free-form method, based on a genetic algorithm, which uses an adaptive grid to iteratively refine the mass model. As input it uses only the information about the lensed images, and nothing about cluster's visible mass (Liesenborgs, De Rijcke & Dejonghe 2006). This last feature sets GRALE apart from many other lens mass reconstruction techniques, and gives it the ability to test how well mass follows light on small and large scales within clusters. GRALE has been used to reconstruct mass distributions in a number of clusters (Liesenborgs et al. 2008, 2009; Mohammed et al. 2014; Mohammed, Saha & Liesenborgs 2015), quantify mass/light offsets in Abell 3827 (Mohammed et al. 2014; Massey et al. 2015), derive projected mass power spectra and compare to those of simulated clusters (Mohammed et al. 2016) and to study the relation between mass and light in MACS0416 (Sebesta et al. 2016). These papers used strong lensing constraints only, and so the analysis was confined to the central regions of galaxy clusters.

[8] GRALE's description, software and installation instructions are available at http://research.edm.uhasselt.be/~jori/grale.

3.4.1 *Description of the method*

GRALE starts out with an initial coarse uniform grid in the lens plane which is populated with a basis set, such as projected Plummer density spheres. A uniform mass sheet covering the whole modelling region is also added to supplement the basis set. As the code runs, the denser regions are resolved with a finer grid, with each cell given a Plummer with a proportionate width. The initial trial solution, as well as all later evolved solutions are evaluated for genetic fitness, and the fit ones are cloned, combined and mutated. The final map consists of a superposition of a mass sheet and many Plummers, typically several hundred to a couple of thousand, each with its own size and weight, determined by the genetic algorithm. Critical curves, caustics and magnifications for any given source redshift are automatically available.

Multiple fitness measures are used in GRALE. These are as follows: (a) Image positions – a successful mass map would lens image-plane images of the same source back to the same source location and shape. A mass map has a better fitness measure if the images have a greater fractional degree of overlap. Using fractional overlap of extended images ensures against overfocusing or overmagnifying images. (b) Null space – regions of image plane that definitely do not contain any lensed features belong to the null space. (c) Critical lines – in some cases, it is known on astrophysical grounds that a critical line cannot go through certain image regions, but must pass between them. GRALE can incorporate this type of constraint, but we have not used this fitness measure in the Frontier Fields work so far. (d) Time delay measurements – though not used in this work, time delay measurements can also be incorporated into the fitness (Liesenborgs et al. 2009; Mohammed et al. 2015).

Each GRALE run with the same set of images, but a different random seed, will produce a somewhat different final map. The dispersion between these quantifies mass uncertainties that are due to mass degeneracies present when all image information is held fixed. The best known among these, the mass sheet degeneracy, is broken in most clusters because of the multiple redshifts of background sources. The other, more numerous and less known degeneracies – documented (Saha 2000; Liesenborgs & De Rijcke 2012) and not documented – are the ones that contribute to the uncertainties.

The clusters *Ares* and *Hera* were modelled with multiple images as inputs, and using two fitness measures: (a) image positions and (b) null space for each source (image set) separately. For image sets where it was not entirely clear if or where the counter images might be present, the nulls were allowed to have large holes corresponding to the regions of possible additional images. GRALE can operate in two modes: with lensed images represented by points, or by extended images. The present reconstruction were done using the extended image mode.

3.4.2 *Strengths and weaknesses*

The main advantage of GRALE is its flexibility, and hence ability to explore a wide range of lensing mass degeneracies. Another important feature, which can be viewed as strength, is that GRALE does not use cluster galaxies, or any information about the distribution of luminous matter to do the mass reconstruction. This is useful if one wants to test how well mass follows light (Mohammed et al. 2014; Sebesta et al. 2016).

GRALE's main weakness is that it is not an optimal tool for identifying lensed images. This is a direct consequence, or, one may say, the flip side of GRALE's flexibility. A technical feature of GRALE worth

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

mentioning is that it requires significant computational resources: GRALE runs on a supercomputer.

3.4.3 Improvements in progress

The GRALE team has carried out numerous tests of the code, to optimize the set of genetic algorithm and other code parameters. In the near future, GRALE will be extended to include fitness measure constraints from weak shear and flexion.

3.5 LensPerfect: the Coe model

The Coe model for *Ares* uses LensPerfect[9] (Coe et al. 2008, 2010). LensPerfect makes no assumptions about light tracing mass. The lens models perfectly reproduce the input observed positions of all strongly lensed multiple images. Redshifts may be either fixed to input spectroscopic redshifts or included in the model optimization based on input photometric redshifts and uncertainties.

The image positions, redshifts and estimated source positions define the lensing deflection field sparsely at the multiple image positions. LensPerfect interpolates this vector field, obtaining a smooth model that exactly reproduces the image deflections at the input image positions. Based on this 2D deflection map, the mass distribution, magnification and all other model quantities may be derived.

3.5.1 Description of the method

The curl-free vector interpolation scheme (Fuselier 2006; Fuselier 2008) uses direct matrix inversion to obtain a model composed of radial basis functions (RBFs) at the positions of the input vectors (our multiple image locations). Each 2D RBF has two free parameters – amplitude and rotation angle – which are determined uniquely by the matrix inversion (Coe et al. 2008; see their fig. 3 and equations 17, 20 and 21).

After setting the width of the RBF, the free parameters are the source positions and any uncertain redshifts. LensPerfect performs an optimization routine searching for those parameters which yield the most 'physical' mass model according to a set of criteria described in Coe et al. (2008, 2010). Briefly, these require that the mass should be positive everywhere and, less restrictively, prefer relatively smooth variations decreasing outward from the centre on average with rough azimuthal symmetry. Penalties for deviating from these ideals are calculated based on the mass map evaluated on an 81×81 pixel grid as detailed in Coe et al. (2008; their section 2.4 and appendix).

3.5.2 Strengths and weaknesses

In high-resolution *HST* ACS images, strongly lensed multiple image locations are observed and measured with accuracies of $\sim$1 pixel, or $\sim$0.05 arcsec. By fully utilizing this information, LensPerfect is able to obtain relatively high-resolution maps of galaxy cluster substructure without relying on any assumptions about light tracing mass. Large numbers of multiple images may be input, and the number of free parameters is always roughly equal to the number of constraints. The mass model spatial resolution increases with the density of multiple images on the sky.

[9] http://www.stsci.edu/~dcoe/LensPerfect/

Given current numbers of multiple images (up to $\sim$100 or so) for a single cluster (e.g. Coe et al. 2010), LensPerfect can accurately recover cluster mass profiles along with some larger subhaloes. Magnifications, however, are influenced by local mass density gradients, which are not accurately reproduced by LensPerfect given current constraints. Furthermore, LensPerfect mass models are only well constrained within the area enclosed by the multiple images and should generally be disregarded outside this region.

3.5.3 Future improvements

LensPerfect is well suited to future data sets such as *JWST* imaging revealing still greater numbers of multiple images. Initial tests with hundreds to a thousand multiple images show great potential for resolving many individual cluster galaxy haloes without assuming light traces mass (LTM). The biggest hurdle (seen in tests with up to 10 000 multiple images) may be accounting for multiple lens plane deflections due to mass along the line of sight.

One potential improvement would be to develop a hybrid method combining LTM assumptions with LensPerfect adding deviations to the mass distribution.

3.6 LENSTOOL : the CATS and Johnson-Sharon models

LENSTOOL as an inversion algorithm deploys both strong and weak lensing data as input constraints. Below, we first briefly outline the available capabilities of the LENSTOOL software package and then describe the specific versions and assumptions that were used to reconstruct *Ares* and *Hera* by two groups: CATS (Cluster As Telescopes) and Johnson-Sharon. The CATS collaboration developed the LENSTOOL algorithm collectively over two decade. The code utilizes the positions, magnitudes, shapes, multiplicity and spectroscopic redshifts for the multiply imaged background galaxies to derive the detailed mass distribution of the cluster. The overall mass distribution in cluster lenses is modelled in LENSTOOL as a superposition of smoother large-scale potentials and small-scale substructure that is associated with the locations of bright, cluster member galaxies. Individual cluster galaxies are always described by parametric mass models, whereas the smoother, large-scale mass distribution can be flexibly modelled non-parametrically or with specific profiles. This available multiscale approach is optimal, inasmuch as the input constraints required for this inversion exercise are derived from a range of scales. Further details of the methodology are outlined in Jullo & Kneib (2009). In its current implementation in LENSTOOL, the optimization of the combined parametric and non-parametric model is computationally time intensive. And some degeneracies persist, despite the large number of stringent input constraints from the positions, shapes, brightnesses and measured spectroscopic redshifts of several families of multiple images. However, these degeneracies are well understood, in particular for specific parameters of models used to characterize the mass distribution. In order to tackle this challenge, an iterative strategy has been developed wherein initial models are derived with the best-fitting values solely from the parametric model, which are then optimized using the underlying multiscale grid. Both the multiscale and the parametric models are adjusted in a Bayesian way, i.e. their posterior probability density is probed with a Markov Chain Monte Carlo (MCMC) sampler. This process allows an easy and reliable estimate of the errors on derived quantities such as the amplification maps and the mass maps.

The CATS and the Johnson-Sharon models are built using the LENSTOOL public modelling software (see e.g. Jullo et al. 2007). The

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

public version of LENSTOOL deployed by Johnson-Sharon adopts the original modelling approach developed by Natarajan & Kneib (1997) wherein a small-scale dark matter clump is associated with each bright cluster galaxy and a large-scale dark matter clump with prominent concentrations of cluster galaxies. This technique of associating mass and light has proven to be very reliable and results in mass distributions that are in very good agreement with theoretical predictions from high-resolution cosmological *N*-body simulations. The Johnson-Sharon models follow the methods described in Sharon et al. (2012) and Johnson et al. (2014).

3.6.1 *Description of the method*

Typically, cluster lenses are represented by a few cluster-scale or group-scale haloes (representing the smooth component, with σ in the range of hundreds to $\sim$1500 km s^{-1}), with contribution from galaxy-scale haloes (see below). Large-scale dark matter haloes are parametrized as pseudo-isothermal elliptical mass distribution (PIEMD),

$$\rho(r) = \frac{\rho_0}{(1 + r^2/r_{\rm core}^2)(1 + r^2/r_{\rm cut}^2)}, \tag{11}$$

where ρ_0 is a normalization, and $r_{\rm core}$ and $r_{\rm cut}$ define a region $r_{\rm core} \lesssim r \lesssim r_{\rm cut}$ in which the mass distribution is isothermal, i.e. $\rho \propto r^{-2}$. In LENSTOOL, PIEMD has seven free parameters: x, y are the coordinates on which the halo is centred, e and θ are the ellipticity and the position angle, respectively, $r_{\rm core}$, $r_{\rm cut}$ and effective velocity dispersion σ_0 that determines the normalization (we note that the σ_0 is not exactly the observed velocity dispersion, see Elíasdóttir et al. 2007). The parameters of cluster-scale haloes are kept free, with the exception of $r_{\rm cut}$ which is usually unconstrained by the strong lensing data, and is thus fixed at an arbitrary value (typically 1500 kpc).

CATS also model galaxies as PIEMD, whereas Johnson-Sharon model galaxies as circular isothermal distributions (see Section 3.6.5). To keep the number of free parameters reasonably small, the parameters of galaxy-scale haloes are determined from their photometric properties through scaling relations assuming a constant mass-to-light ratio for all galaxies,

$$\sigma_0 = \sigma_0^* \left(\frac{L}{L^*}\right)^{1/4} \text{and} \; r_{\rm cut} = r_{\rm cut}^* \left(\frac{L}{L^*}\right)^{1/2}. \tag{12}$$

The positional parameters, x, y, e and θ, are fixed to their observed values as measured from the light distribution in the imaging data.

CATS used the simulated strong lensing catalogues and the LENSTOOL software to perform a mass reconstruction of both simulated clusters, assuming a parametric model for the distribution of dark matter. The model is optimized with the Bayesian Markov chain Monte Carlo sampler, described in detail in Jullo et al. (2007). The mass distribution is optimized in the image plane by minimizing the distance between the observed and predicted multiple image positions. Weak lensing information is not taken into account. The image-plane root-mean-squared (rmsi) distance of the images predicted by the model were used to compare with the observed positions as an accuracy estimator of the model (Limousin et al. 2007).

The CATS collaboration has modelled both clusters, *Ares* and *Hera* have been modelled as bimodal clusters with two smooth dark matter clumps and two BCGs lying in the centre of those main clumps. Each smooth component is modelled using a PIEMD profile. Cluster member galaxies are taken from the given simulated catalogues up to a magnitude of $m_{\rm F160W} < 22.0$ for *Ares* and $m_{\rm F814W} < 24.0$ for *Hera*. These are modelled with PIEMD profiles under the assumption that (i) their positions correspond to the observed ones, (ii) they have circular shape and (iii) their mass is proportional to the galaxy magnitudes in the *F*160*W* band. In the provided models, it is assumed that they all have the same *M*/*L* ratio, following the scaling relations in equation (12). Two free parameters are used to describe the ensemble of cluster galaxies, namely σ_0^* and $r_{\rm cut}^*$. At the end of the optimization, the best-fitting values for these parameters are $\sigma_0^* = 98 \pm 1.2$ km s^{-1} and $r_{\rm cut}^* = 262 \pm 0.66$ kpc for *Ares* and $\sigma_0^* = 96.7 \pm 3.3$ km s^{-1} and $r_{\rm cut}^* = 33 \pm 0.07$ kpc for *Hera*. The reference luminosities are set to $m^* = 18.5$ and $m^* = 19.82$ for *Ares* and *Hera*, respectively. All multiple images provided were used in this model. In addition, a few (massive) cluster galaxies in both clusters were more carefully modelled in order to improve the rmsi of nearby multiple images. The RMSi is defined as

$$rms = \sqrt{\frac{1}{N}\sum_{i=1}^{n} |\theta_i^{\rm obs} - \theta_i^{\rm pred}|^2}, \tag{13}$$

where $\theta_i^{\rm obs}$ and $\theta_i^{\rm pred}$ are the positions of the multiple images observed and predicted by the model; N is the total number of images.

In addition to the two BCGs, four other central cluster galaxies were modelled in this way in *Hera* (of which one is considered to be a foreground) and three, also central, galaxies in *Ares*.

These reconstructions have a resulting rms in the image plane of 0.87 arcsec for *Ares* and 0.95 arcsec for *Hera*.

The Johnson-Sharon models for *Ares* and *Hera* were constructed using techniques similar to those in Johnson et al. (2014), using the catalogues of multiple images that were provided to the lens modellers as positional constraints. The redshifts of the background sources were assumed to be known spectroscopically with no uncertainty or outliers. Both clusters were modelled with two PIEMD haloes, to represent the smooth dominant dark matter components, each centred close to the two peaks in the light distribution in the mock *HST* images with their exact positions set by the MCMC minimization process.

Individual PIEMD haloes were assigned to each galaxy in the provided catalogue, with positional parameters, x, y, e and θ, fixed to their observed values as measured from the light distribution in the mock imaging data. The parameters that describe the slope of the projected mass density were scaled with the light in the *F*125*W* band assuming a constant *M*/*L* ratio for all the galaxies, following the scaling relations in equation (12). As both clusters are at $z = 0.5$, the same scaling relations were used for the cluster member galaxies: $\sigma_0^\star = 120$ km s^{-1}, $r_{\rm core} = 0.15$ kpc and $r_{\rm cut} = 30$ kpc, and $m_\star = 20.00$, 19.87 for *Ares* and Hera, respectively.

A few galaxies located near constraints were modelled independent of the scaling relations and their core radius and velocity dispersion were left as free parameters in the lens models. This includes the two bright cluster galaxies lying at the centres of the gravitational well of both clumps in the dark matter distribution in both clusters.

We note that the PIEMD functional form of the cluster galaxies used in LENSTOOL differs from the function that was used in the simulation of *Ares* in the treatment of the truncation radius. While the PIEMD profile transitions smoothly from isothermal and asymptotes to zero at large radii, the *simulated* mass distribution truncates the mass function sharply to zero at $r = r_{\rm cut}$. This discrepancy is what causes the sharp circular residuals seen in Fig. 9. We thus do not expect the model to accurately reconstruct the mass distribution at radii larger than the truncation radius.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

In addition, the Johnson-Sharon model assumes that the ellipticity and position angle of the light of each mock galaxy follows the underlying mass distribution. In practice, all the galaxies in the underlying simulated mass distribution had circular geometry (i.e. no ellipticity) and the galaxies were painted on with arbitrary ellipticities and position angles. This feature of the blinded analysis contributes to residuals on small scales in the mass reconstruction. Finally, the Johnson-Sharon model does not use weak lensing information and does not include cluster-scale haloes outside of the main FOV if such haloes are not required by the strong lensing constraints alone.

3.6.2 *Strength and weaknesses*

LENSTOOL strengths and weaknesses are typical of parametric models. This approach is useful in the sense that it directly compares physically motivated models to data, propagating errors in a fully consistent and Bayesian manner. It allows direct comparison with simulation outputs and the assessment of possible discrepancies. On the other hand, parametric models can significantly differ from reality and their lack of freedom introduces biases in the estimated masses, matter densities or errors. Regarding practical aspects, errors estimation implies running MCMC sampling, which can only be performed on supercomputer. LENSTOOL calculations can last for a couple of weeks on shared memory machines depending on the model complexity and the amount of multiple images. In the case of *Hera* and *Ares*, optimization lasts for about 10 h.

3.6.3 *Improvements in progress*

CATS is currently working actively on two improvements that should significantly improve the accuracy of their mass reconstructions. First, LENSTOOL in its current revision does not permit radial variation of the ellipticity for the mass distribution, and this restricts the flexibility of models that can be generated. Current code development aims to include this additional degree of freedom in the modelling. Secondly, in order to maximally extract information from the exquisite image resolution afforded by the *HST* FFI, flexion measurements will be included as input constraints in the modelling. Finally, a new MCMC engine with MPI support and a GPU-based LENSTOOL are under development to decrease the computing time.

3.7 GLAFIC : the GLAFIC models

The publicly available GLAFIC code (Oguri 2010)[10] is used for mass modelling in the GLAFIC models.

3.7.1 *Description of the method*

GLAFIC adopts the so-called parametric lens modelling in which the lens mass distribution is assumed to consist of multiple components, each of which is characterized by a small number of parameters such as the centroid position, mass, ellipticity and position angle. Mass distributions of cluster member galaxies are modelled with PIEMD models. In order to reduce the number of parameters, the velocity dispersion σ and truncation radius $r_{\rm cut}$ of each member galaxy are assumed to scale with the galaxy luminosity L as $\sigma \propto L^{1/4}$ and $r_{\rm cut} \propto L^{\eta}$, and the normalizations of the scaling relations are treated as free parameters (see e.g. Oguri 2010). Ellipticities and position angles of individual member galaxies are fixed to values measured in the image. These parameters are optimized to reproduce positions of observed multiple images, either using the downhill simplex method or Markov Chain Monte Carlo. Examples of detailed cluster mass modelling with GLAFIC are found in Oguri (2010), Oguri et al. (2012), Oguri et al. (2013), Ishigaki et al. (2015) and Kawamata et al. (2016). GLAFIC can also simulate and fit lensed extended sources. This functionality has been used to, e.g. fit a lensed quasar host galaxy (Oguri et al. 2013), estimate a selection function of lensed high-redshift galaxies (Ishigaki et al. 2015) and derive sizes of lensed high-redshift galaxies (Kawamata et al. 2015).

[10] http://www.slac.stanford.edu/~oguri/GLAFIC/

3.7.2 *Strengths and weakness*

An advantage of GLAFIC is a wide range of lens potential implemented in the code, which enables flexible modelling of cluster mass distributions. For example, in addition to the standard external shear, one can add higher order perturbations with arbitrary multipole orders (see Oguri 2010). When necessary, in addition to observed multiple image positions, GLAFIC can also include flexible observational constraints such as time delays and flux ratios between multiple images, and (reduced) shear and magnification values at several sky positions measured by weak lensing and Type Ia supernovae, respectively.

The source plane χ^2 minimization is often adopted for efficient model optimizations. In doing so, GLAFIC converts the distance between observed and model positions in the source plane to the corresponding distance in the image plane using the full magnification tensor. In appendix 2 of Oguri (2010), it has been shown that this source-plane χ^2 is accurate in the sense that it is very close to the image-plane χ^2 and therefore is sufficient for reliable mass modelling.

Of course GLAFIC supports the image plane χ^2 minimization as well. Adaptive-meshing with increased resolution near critical curves is used for efficient computations of multiple images for a given source position. Multiple images and critical curves are computed for the best-fitting model from the source plane χ^2 minimization to check the robustness of the result.

A known limitation of GLAFIC is that it can only handle single lens planes. Lens systems for which multiple deflections at different redshifts play a crucial role are difficult to be modelled by GLAFIC.

3.7.3 *Modelling* Ares *and* Hera

Each halo component is modelled by the elliptical NFW profile. For *Ares*, five halo components are included, in addition to the member galaxies modelled by PIEMDs (see above). In addition, the η parameter (see above) is fixed to 0.5. For *Hera*, two NFW halo components are placed around two brightest galaxies. These two brightest galaxies are modelled by the Hernquist profile, separately from the other member galaxies. The formula for this profile is given by

$$\rho(r) = \frac{M}{2\pi(r/r_b)(1+r/r_b)^3}\ , \tag{14}$$

where the scale radius r_b is related to the effective radius R_e by $r_b = 0.551R_e$. We introduce the ellipticity in the projected mass density. Ellipticities and position angles of the brightest galaxies are fixed to observed values. To achieve better fit, external shear and

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

third-order multipole perturbation are added for *Hera*. In modelling member galaxies, η in the scaling relation of truncation radii is fixed to 0.5 for *Ares* and is treated as a free parameter for *Hera*. Simulated *F*814*W* band images are used to measure luminosities, ellipticities and position angles of member galaxies with SEXTRACTOR for both *Ares* and *Hera*. Overall, a more elaborated lens model is adopted for *Hera* compared to *Ares*, because in the initial exploration period of mass modelling it was found that the lens potential of *Hera* appears to be much more complex. In particular, both external shear and high-order perturbations have been included in the model. The latter are simulated using a potential of the form

$$\phi = -\frac{\epsilon_{\rm pert}}{m} r^n \cos m(\theta - \theta_{\rm pert} - \pi/2) \,, \tag{15}$$

where $m = 3$, $n = 2$ and $\epsilon_{\rm pert}$ and $\theta_{\rm pert}$ are left as free parameters in the fit. The best-fitting values for the external shear amplitude and orientation are $\gamma_{\rm ext} = 0.072$ and $\theta_{\rm ext} = 32.4$ deg, respectively. For the third-order multipole perturbation, we found $\epsilon_{\rm pert} = 0.018$ and $\theta_{\rm pert} = 25.9$ deg.

The resulting best-fitting models reproduce image positions very well, with rms of ∼0.27 arcsec for *Ares* and ∼0.43 arcsec for *Hera*.

3.8 LTM: the Zitrin-LTM models

The Zitrin LTM method (Broadhurst et al. 2005; Zitrin et al. 2009) was designed primarily to be a very simple, straightforward modelling method with a minimal number of free parameters, relying only on the observable light distribution of cluster members (namely their positions and relative fluxes) to supply a well-guessed and highly predictive solution to the mass distribution of the lens and the location of multiple image systems (e.g. Zitrin et al. 2012, 2013). Previous to the design of this method, it had been shown that (a) cluster galaxies must be included, typically with a mass in proportion to their luminosity, in order the solution to have predictive power to find multiple images and that (b) a dark matter component should be added (see Kneib et al. 2004; Broadhurst et al. 2005). This simple parametrization, as we detailed further below, has allowed us to identify systems of multiple images in an unprecedented number of clusters, where the images are physically matched also by the initially guessed model (which is then refined), and are not only matched by eye based on their colour information as is often accustomed.

3.8.1 *Description of the method*

As mentioned above, this method was designed to include both a galaxy component and a dark matter component, yet to successfully do so with a minimal number of free parameters. To form the galaxy component, cluster galaxies (found following the red sequence in a colour–magnitude diagram) are assigned each with a power-law mass density distribution, where the normalization of each galaxy's weight is proportional to its (relative) flux, and the exponent is the same for all galaxies and is the first free parameter of this method. The superposition of all galaxy power-law mass distribution then constitutes the lumpy, galaxy component of the model. To describe the dark matter distribution, the galaxy component is smoothed with either a Spline interpolation or usually a Gaussian kernel whose degree or width is the second free parameter of this method. The smoothing yields a diffuse, smooth dark matter component that depends on the initial light distribution; therefore, the method is dubbed LTM as both the galaxy *and* dark matter components roughly follow simply the light distribution. Next, the two components are added with a relative weight (typically around few to a couple dozen percents for the galaxies), which is another free parameter in the modelling. The fourth parameter is an overall normalization of the lens model to a certain redshift or multiple-image system. In addition to the four parameters, we often introduce several other parameters that add some flexibility and help in refining the final solution given the set of input multiple images. These include a core and two-parameter ellipticity for the BCG(s), two parameter external shear (which mimics ellipticity for the critical curves) and chosen galaxies whose weights (or fluxes) are left free to be optimized in the modelling, meaning that they are allowed to deviate from the adopted mass-to-light relation. The minimization for the best-fitting solution and related errors, given a set of multiple images (often found with the aid of the initially guessed map from this method), is performed with a χ^2 criterion comparing the positions of multiple images with the predicted ones, in the image plane, via a few-dozen thousand MCMC steps with Metropolis-Hastings algorithm. The conversion between mass distribution and deflection angles is done in Fourier space, using FFT techniques.

3.8.2 *Strengths and weaknesses*

The resulting lens model from this procedure, as its name suggests, is strongly coupled to the input light distribution of the lens (cluster members positions and luminosities). This entails various strengths and weaknesses. The fact that the solution is coupled to the light distribution is what grants this method with the unprecedented prediction power to delineate the critical curves and locate multiple images in advance, even if no multiple-image system is used as constraint (see Zitrin et al. 2012). In fact, most of the free parameters in the initial solution are relatively well known, so that as a first step (i.e. to find multiple images) we can reduced these to one free parameter – namely the normalization of the lens, and obtain a well-guessed solution, that we have shown is not much different than the resulting solution for the same clusters when using many multiple images as constraints (Zitrin et al. 2012). This means that the method is capable to supply a well-guessed solution also in cases where *HST* high-quality data is lacking.

The simplistic nature of this method also means that the solution is often faster to converge and compute than other grid-based methods or parametrizations, allowing the analysis of many dozens of cluster lenses in a relatively short time.

Another advantage that this method encompasses is that the same very simple procedure applies to all clusters – from relaxed, small clusters and groups (such as the relatively smaller cluster lenses A383, MS2137 or A611, see Zitrin et al. 2015, for recent modelling), to the most complex merging clusters such as M0416, M1149 or M0717 (Zitrin et al. 2015), that often require multihalo fits in other parametric methods.

But the coupling to the light distribution also means that the spatial flexibility of the model is small. While our parametrization does allow for a flexible mass profile in the sense that it is not limited to a certain analytic form, the solution is limited spatially by the light distribution. This means that the multiple-image reproduction accuracy is often smaller than in other more flexible parametric methods that model the dark matter independently of the light (such as other well-known methods listed in this work including our own second method listed below). This is manifested usually in clusters that have a large number of multiple images spread across the field; for these, the LTM method often reaches a finite *rms* value of ∼1–2 arcsec on which it cannot improve further.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

A second disadvantage stemming from our parametrization – since we do not model the dark matter independently of the light and since the critical curve's ellipticity in our modelling is for the most part generated by the external shear – is that there is no ellipticity assigned directly to the mass distribution. This creates some discrepancy between the *lens* and *mass* models: the mass distribution can be often significantly rounder than implied by the critical curves, whose ellipticity comes from the external shear that does not contribute ellipticity to the convergence map. In simple words, this reveals a degeneracy regarding the true ellipticity of the mass distribution – as the ellipticity of the lens can be attributed to intrinsic ellipticity or to external shear.

To summarize, we thus consider this method very *reliable and robust*, supplying especially well-guessed initial maps for any given cluster regardless of its complexity, and with unprecedented prediction power to find multiple images, but it can also be in some cases *less accurate and spatially flexible*. Also, given this is a light-tracing method, we do not expect this method to describe well numerical simulations whose mass-to-light relations are not completely representative.

3.8.3 *Improvements in progress*

The code is currently being optimized in order to speed up the minimization procedure. We are also testing whether replacing the galaxy component with the more well-behaved PIEMD (see below), despite having a fixed isothermal slope, would be sufficient for our purposes. We have also implemented an option of smoothing with an elliptical Gaussian which then introduces ellipticity into the matter distribution itself.

Note that our calculations are performed on an input grid matching an actual image of the field, with its native pixel scale. To speed the minimization procedure, we often reduce the resolution (especially in the case of *HST* that has high spatial resolution) by factors 4 to 10 on each axis. This contributes to the finite, non-negligible rms obtained often in this method (e.g. due to pixel coordinates round-ups etc.). We intend to investigate this further and try to improve the resolution in the crucial places, such as near the critical curves and when delensing to the source plane, where this lower resolution might prevent a further improved solution.

3.8.4 *Modelling of* Ares *and* Hera

To model *Ares* and *Hera*, we use the following setting in our LTM pipeline. We create a grid of 4080×4080 pixels covering the FOV, with an angular resolution of 0.5 arcsec pixel^{-1}. The calculation in practice is performed in two stages – first, we run many individual random MC chains with a grid resolution lower by factor 10 on each axis. From this we find the global minimum area and extract the covariance matrix. A proper, long MCMC is then run with a grid of four times lower resolution than the original input image. The final solution is then interpolated to match the original pixel-scale map. Errors were derived using 50 random models form the MC chain, with a positional uncertainty of 1.4 arcsec for the χ^2 term. We use the input list of galaxies supplied by the simulators scaled by their light. In *Ares*, we allow five galaxies to deviate from the nominal mass-to-light ratio and be freely weighted by the MC chain, and for two of them – especially important where radial images are seen in the data – we allow for a free core radius as well. In the case of *Hera*, only three galaxies were modelled in this way. The ellipticity (and direction) of these bright galaxies are also left as free parameters. As constraints, we use the full list of multiple images. No weak lensing constraints were used. The final rms of the model is 1.8 and 1.2 arcsec for *Ares* and *Hera*, respectively, which, as we mention above, is in part limited by the finite lower resolution of the grid we work on. The best-fitting external shear amplitude and orientation are $\gamma_{\rm ext} = 0.007$, $\theta_{\rm ext} = 6.5$ deg for *Ares* and $\gamma_{\rm ext} = 0.114$, $\theta_{\rm ext} = 72.1$ deg for *Hera*.

3.9 PIEMDeNFW: the Zitrin-NFW models

Zitrin et al. (2013) expanded their pipeline to also allow for a fully parametric solution. This method in essence is similar to the other parametric techniques mentioned here such as LENSTOOL and GLAFIC. The main motivation for adding this parametric pipeline was to (a) allow for further flexibility and improved fits by having a semi-independent solution in which the dark matter is modelled independently of the light, and (b) test for the magnitude of systematic differences between these methods (Zitrin et al. 2015).

3.9.1 *Description of the method*

As is usually accustomed in parametric modelling, in order to describe well the multiple-image positions with enough prediction power, this method also relies on a combination of galaxy and dark matter. The red sequence cluster galaxies are modelled each as PIEMDs based on the prescription and scaling relations used in LENSTOOL, and typically with a fixed mass-to-light ratio. Usually two or three parameters are left free to describe the galaxy component: the velocity dispersion, core radius and truncation radius, of an M^* galaxy, which is used as reference for the scaling relations. The dark matter component is modelled also with an analytic, fully parametric recipe. We can choose either an elliptical Navarro et al. (1996) profile (eNFW), or, also PIEMD for the cluster's dark matter halo. Therefore, in this method, the dark matter is modelled with a symmetric analytic form, independent from the light distribution. Similar to our LTM method, the same minimization engine is used here: a long MCMC with a χ^2 image-plane criteria. Also here we can add other parameters to be optimized in the minimization, such as the ellipticities of the BCGs, their mass can be allowed to deviate from the adopted scaling relation, and so forth.

3.9.2 *Strengths and weaknesses*

Compared to our LTM technique, for example, we have found that the fully parametric technique is more spatially flexible and can thus often supply a more accurate solution with a (somewhat) smaller image-plane rms. On the other hand the method is less efficient at finding new sets of multiple images (especially before the model is initially constrained).

In a similar sense, another main disadvantage of such parametrizations is the need to add dark matter haloes to model subhaloes for complex structures (such as merging clusters), without knowing if these are fundamental parameters, e.g. accounting truly for additional dark matter haloes, or just nuisance parameters that help add flexibility and refine the fit. Additionally, each such added halo adds several (usually four to six) free parameters to the minimization procedure rendering it significantly more cumbersome.

Note that since we developed this method with the same infrastructure used for our LTM method, and in part, for comparison with it, the solutions given by the PIEMDeNFW method, despite being

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

analytic in nature, are also calculated on a grid, the size of the input *.fits image, similar to our LTM procedure. This results in a somewhat slower procedure compared, for example, to our LTM technique, and also here to achieve higher converging speed we lower the grid resolution by factors of a few on each axis. Again this leads to a finite rms due to e.g. numerical round-ups in high-magnification regions.

3.9.3 *Improvements in progress*

The main improvement we wish to implement is to speed up the procedure. This for start can be achieved if part of the calculation is done completely analytically/numerically (say, only around the positions of multiple images) rather than on a full-frame grid. We intend to explore such possibilities. Also, recently we added the possibility for an external shear to allow for further flexibility.

3.9.4 *Modelling of* Ares *and* Hera

To model *Ares* and *Hera*, we use the following setting in our PIEMDeNFW pipeline. As done with the LTM-gauss method, we create a grid of 4080×4080 pixels covering the cluster. We start by running many individual random MC chains with a grid resolution lower by factor 20 on each axis. From this we find the global minimum area and extract the covariance matrix. A proper, long MCMC is then run with a grid of four times lower resolution than the original input image. The final solution is then interpolated to match the original pixel scale map. Errors were derived using 50 random models to form the MC chain, with a positional uncertainty of 1.4 arcsec for the χ^2 term. We use the input list of galaxies, scaled by their light. The brightest galaxies are optimized individually as done with the LTM-gauss pipeline. In this case, however, the ellipticity (and direction) of the four brightest galaxies in both clusters are also left as free parameters. In *Ares*, two cluster-scale DM haloes in the form of elliptical NFW mass densities are introduced, with fixed centring on the respective BCGs. In *Hera*, we used three such large haloes. As constraints we use the full list of multiple images. No weak lensing constraints were used. The final rms of the model is 1.8 arcsec, which as we mentioned above is in part limited by the finite lower resolution of the grid we work on.

4 RESULTS

In this section, we describe how the different methods perform at recovering several properties of the lenses.

4.1 Convergence maps

The reconstructed convergence maps of *Ares* and *Hera* are shown for all models in Figs 7 and 8, respectively. The maps are all normalized to $z_S = 9$. In both figures, the maps derived from the free-form algorithms are shown first (beginning from the upper-left panel). The last panel in each figure shows the true convergence map, for easy comparison. All maps cover the same FOV. This does not correspond to the size of the simulated images that were made available to the modellers. Indeed, for several technical reasons inherent to each methodology employed, the submissions by the different groups were different in size. To carry out a proper comparison between the models, we restrict our analysis to the area around each of the two lenses, which is covered by all the reconstructions. More precisely, we used as footprints for identifying the area of analysis the submissions by the GLAFIC and by the GRALE teams for *Ares* and *Hera*, respectively. In the first case, the FOV is $\sim$180 arcsec $\times$ 180 arcsec. In the second, the reconstructed area is $\sim$110 arcsec $\times$ 110 arcsec wide.

Since *Ares* was constructed parametrically with light-tracing mass, it is particularly well suited for reconstruction by parametric techniques. The parametric CATS, GLAFIC, Johnson-Sharon, and Zitrin models and the hybrid Diego model all include mass substructure at the observed positions of cluster galaxies, recovering the *Ares* mass distribution with high fidelity. The free-form GRALE, Bradac-Hoag and Coe models do not assume LTM, reconstructing the mass distribution solely based on the observed lensing. They recover the main mass peaks, but smaller substructures are not constrained by the lensing data. The GRALE model accurately reproduces the cluster bimodality. The Bradac-Hoag and Coe models are less smooth, including noisy smaller substructure, especially outside the region constrained by strongly lensed multiple images.

The *Hera* cluster, obtained from an *N*-body simulation, is less ideal for being reconstructed using parametric methods. Indeed, the performance of the parametric algorithms appears more consistent with that of the free-form ones. *Hera* is constructed assuming light traces its massive substructure, as assumed by the parametric and hybrid methods. The Bradac-Hoag and GRALE models do not make that assumption and thus recover fewer small subhaloes.

The major differences between the models and the true convergence maps are found near substructures, but also the shape of the mass distributions, especially at large distances from the centre, show inconsistencies. We will discuss them in more detail in the next sections.

To better highlight the differences between the maps, we show the relative differences $(\kappa - \kappa_{\rm true})/\kappa_{\rm true}$ between the reconstructed and the true convergence maps for *Ares* and *Hera* in Figs 9 and 10, respectively.

4.2 One-dimensional mass and convergence profiles

We begin discussing the results on the mass and convergence (or surface density) profiles. Meneghetti et al. (2010a) already showed using only one of the methods employed in this paper (LENSTOOL, employed by both the CATS and the Johnson-Sharon teams) that strong lensing can potentially measure the mass inside the Einstein radius with an accuracy of the order of a few per cent. In the cases of *Ares* and *Hera*, the sizes of the Einstein radii are significantly different. In Fig. 11, we show how $\theta_{\rm E}$ grows as a function of the source redshift $z_{\rm s}$. The Einstein radius of *Ares* is $\sim$20 arcsec at $z_{\rm s} = 1$. Its size at $z_{\rm s} \sim 2$ is more than double and it grows asymptotically to $\sim$55 arcsec at higher redshift. The reason of the steep rise between $z_{\rm s} = 1$ and $z_{\rm s} = 2$ is that *Ares* has a bimodal mass distribution. For sources at low redshift ($z_{\rm s} \sim 1$), each of the two mass clumps has its own critical lines. These are shown by the red curves in the upper panel of Fig. 3. To draw the plot in Fig. 11, we use the centre of the most massive mass clump as reference, and only the critical line enclosing this point is used to measure $\theta_{\rm E}$. By increasing the source redshift, the critical lines around the two mass clumps merge into a single, very extended critical line (see the white lines in the upper panel of Fig. 3, which shows the critical lines for sources at $z_{\rm s} = 9$).

In the case of *Hera*, the Einstein radius grows from $\sim$12 arcsec at $z_{\rm s} = 1$ to $\sim$30 arcsec at $z_{\rm s} = 9$. The critical lines for these two source redshifts are shown in the upper panel of Fig. 4.

In Fig. 11, we also show the redshift distributions of the multiple images identified in the background of the two clusters (red and blue histograms). These multiple images are marked with numbered

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

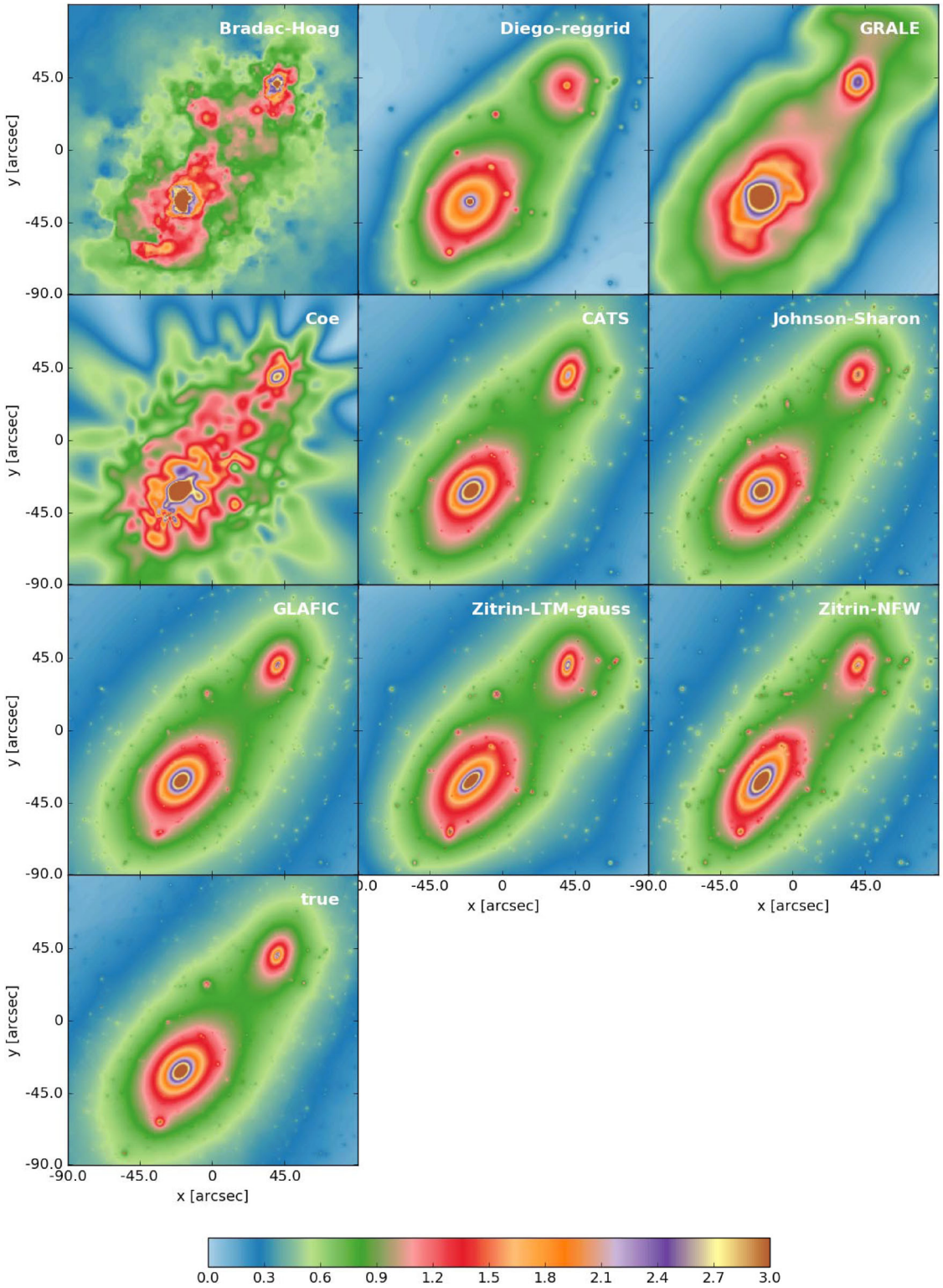

Figure 7. Convergence maps ($z_s = 9$) of *Ares*. The first nine panels show the results of the reconstructions, beginning with the free-form methods (panels 1–4) and concluding with the parametric models (panels 5–9). The lower left panel shows the true convergence map, for comparison.

circles in the upper panels of Figs 3 and 4. The labels of each image are constructed as X.Y, where X is the ID of the source and Y is the ID of the multiple images belonging to the same system. Being such a powerful lens, *Ares* produces many more multiple images than *Hera*, some of which originate from galaxies at redshift $z_s \sim 6$. The most distant multiple image system in the field of *Hera* is only at $z_s \sim 3.5$. In both cases, however, the redshift distribution of the multiple images overlaps with the redshift range where the size of the Einstein radii has the strongest growth. Indeed, the relative variation of θ_E between $z_s = 3$ and $z_s = 9$ is only $\lesssim 10$ per cent. Thus,

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

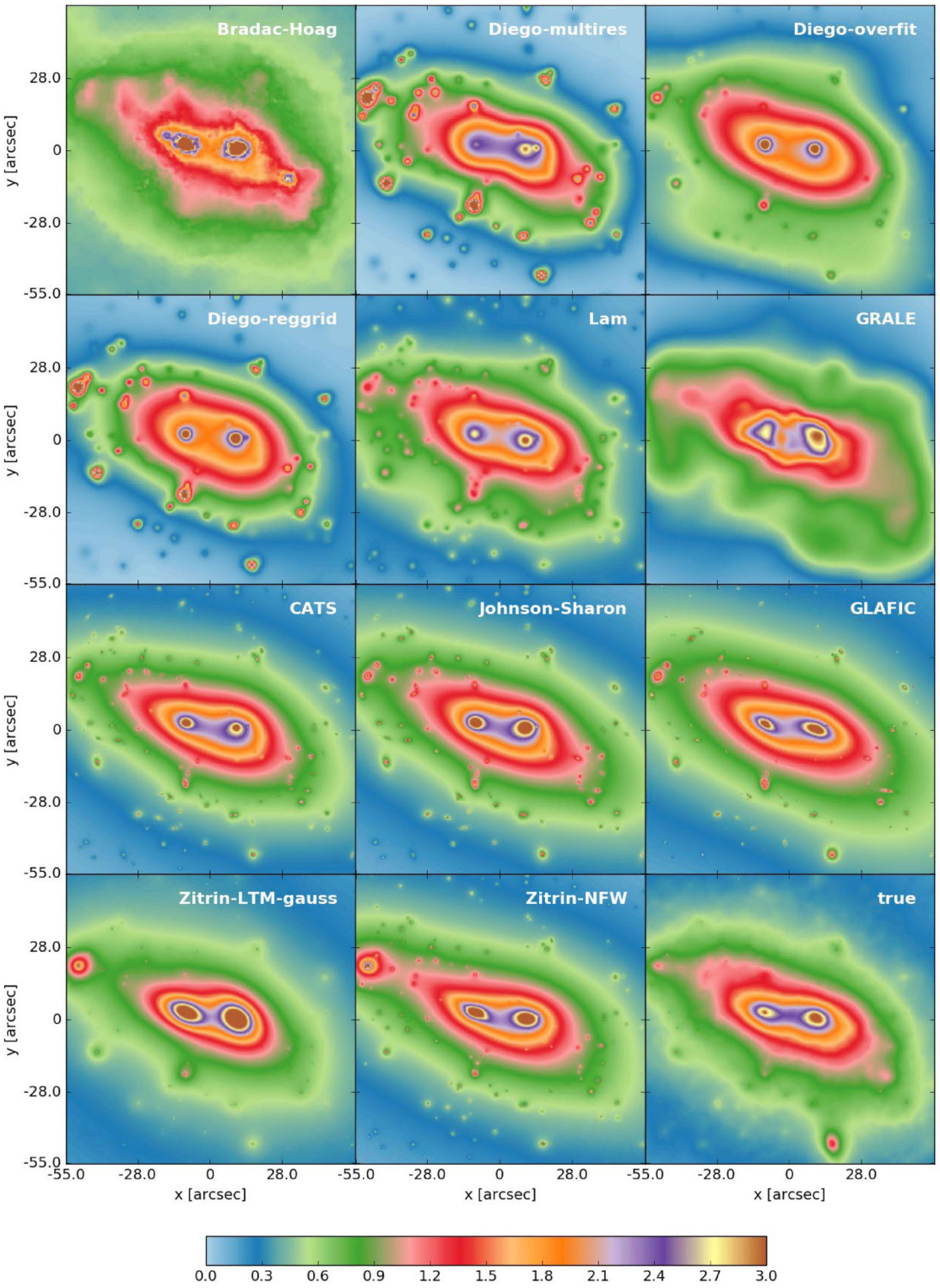

Figure 8. Convergence maps ($z_s = 9$) of *Hera*. The first 11 panels show the results of the reconstructions, beginning with the free-form methods (panels 1–6) and concluding with the parametric models (panels 7–11). The lower right panel shows the true convergence map, for comparison.

we expect that the models constructed using these constraints can be safely used to trace the growth of the cluster strong lensing region up to very high redshifts. Analogously, we expect that the mass profiles are recovered with higher precision in the radial ranges $20 \lesssim \theta \lesssim 60$ arcsec and $10 \lesssim \theta \lesssim 30$ arcsec for *Ares* and *Hera*, respectively.

This is consistent with our findings. The upper panels of Figs 12 and 13 show the projected enclosed mass profiles of *Ares* and *Hera*, respectively. The bottom panels show the projected mass density profiles in units of convergence κ for $z_S = 9$. The profiles are computed with respect to the centre of the most massive subclump in each cluster field. To facilitate the comparison between parametric

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

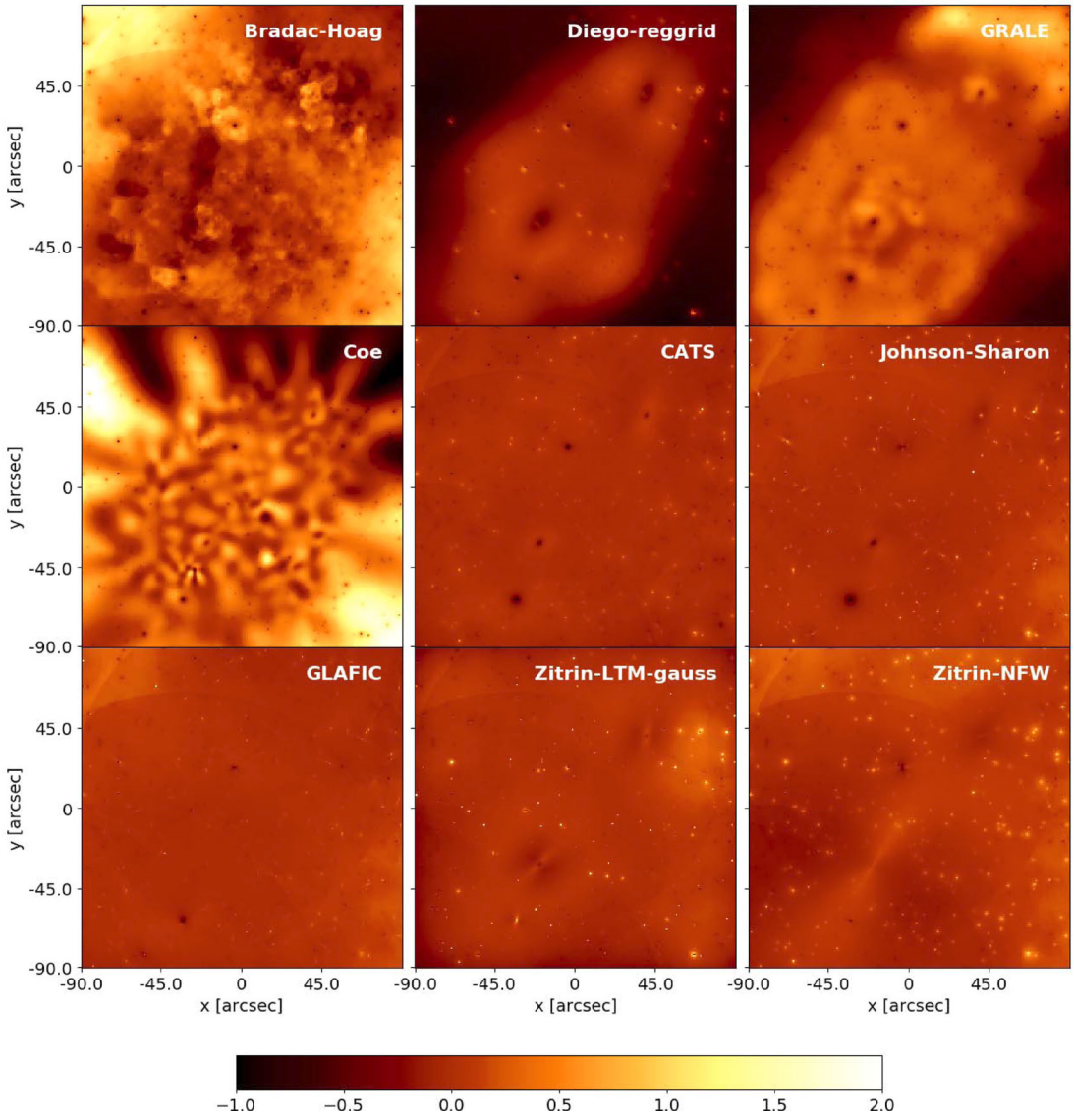

Figure 9. Relative differences $(\kappa - \kappa_{\rm true})/\kappa_{\rm true}$ between the mass reconstructions and true *Ares* mass distribution.

and free-form methods, we show the results for these two classes of models separately (left-hand and right-hand panels).

The mass distribution of *Ares* is generated in a very similar manner as used by the parametric techniques (except Zitrin-LTM) to model the lenses – as a combination of parametrized mass components, including subhaloes at the positions of cluster galaxies. Therefore, it is not surprising that these methods recover the true mass profile of *Ares* with very good accuracy. For example, the CATS, Johnson-Sharon and GLAFIC profiles differ from the true mass profile by $\lesssim \pm$ 2 per cent. Larger differences are found for the Zitrin-LTM-gauss and the Zitrin-NFW approaches (perhaps because these are calculated on a lower resolution grid, see discussion in Sections 3.8 and 3.9), but even for these models, in the region probed by strong lensing, the deviations from the true mass profiles are within $\sim\pm$10 per cent.

It is noteworthy that neither the CATS nor the Johnson-Sharon teams used the NFW density profile to model the smooth DM haloes of the two main mass components of Ares. On the contrary, they used cored isothermal profiles, which can of course be tweaked to match the lensing properties of NFW haloes. This is consistent with the findings of Shu et al. (2008), who showed that, in several cases, strong lensing clusters are equally well modelled with cored-isothermal and NFW density profiles. The additional constraints provided by complementary analysis, such as stellar-kinematics in the BCG, could help break this degeneracy (Newman et al. 2013). Moreover, the adoption of an isothermal profile with core instead of the NFW profile does not prevent several models from recovering the correct slope of the surface mass density (i.e. convergence) profile over a relatively broad range of distances from the cluster centre. The constraints available to carry out the reconstructions include both radial and tangential features, with the former particularly sensitive to the slope of the projected density profile.

Among the free-form methods, the reconstructed profiles generally deviate by $\lesssim$ 5–15 per cent from the true mass and convergence profiles. Some models (e.g. GRALE) have a very similar performance to parametric methods. The best agreement between the true profile and the models is found between 20 and 60 arcsec from the lens centre, which nicely corresponds to the size of the Einstein radius, as shown in Fig. 11.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

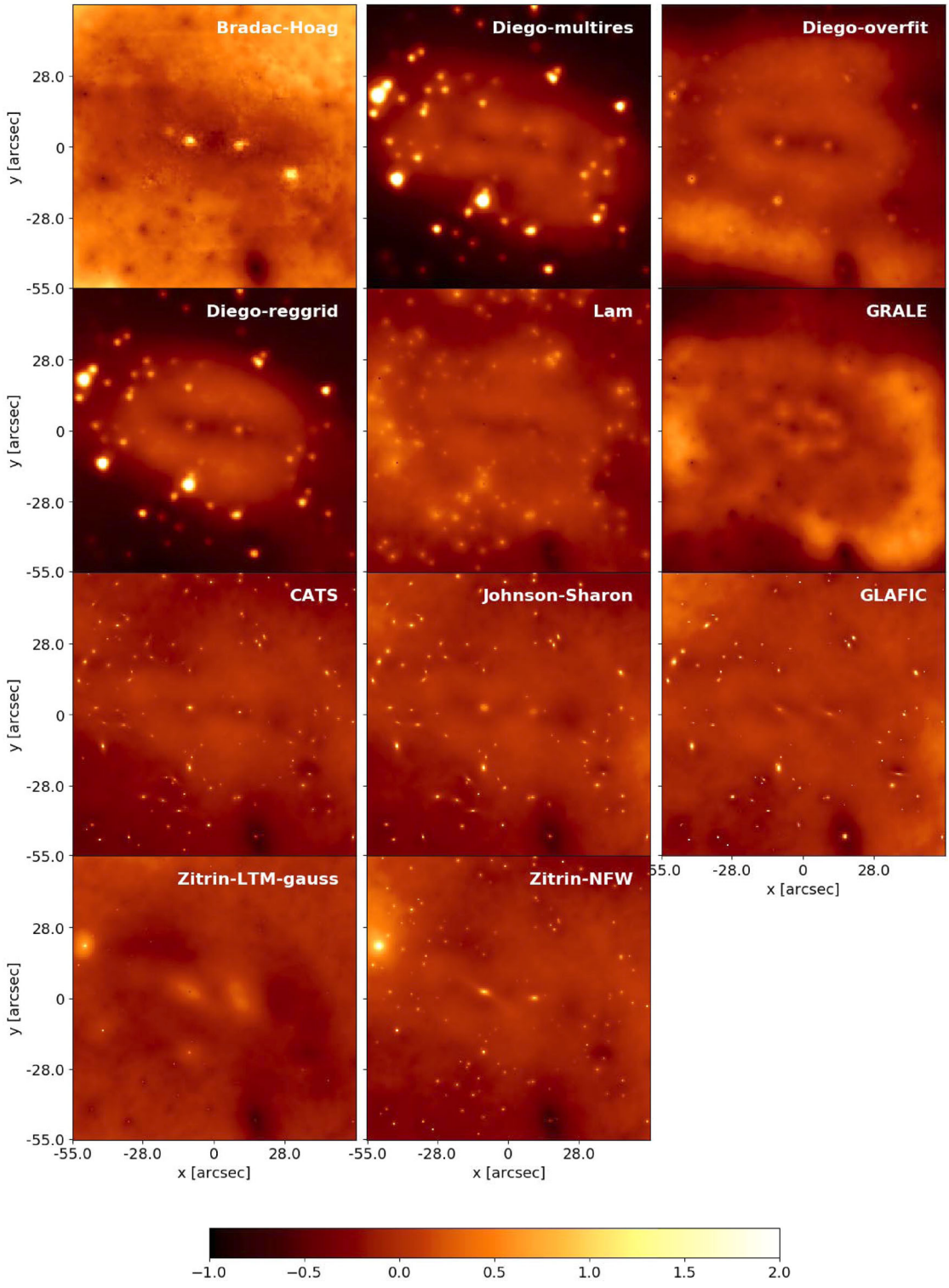

Figure 10. Relative differences $(\kappa - \kappa_{\rm true})/\kappa_{\rm true}$ between the mass reconstructions and true *Hera* mass distribution.

Hera is a less idealized test case for most of the parametric methods, but it still assumes that light traces the mass substructure. So also for this lens, the parametric models reproduce the input mass profiles more closely than the free-form methods, though the differences between the two approaches are now reduced. We find that the mass profiles obtained with the parametric methods differ from the input mass profile by less than 10 per cent within $\sim$80 arcsec from the assumed centre. The same level of accuracy is reached by the free-form methods within $10 \lesssim r \lesssim 30$ arcsec. This radial range corresponds to the size of the region probed by strong lensing. Both parametric and free-form methods clearly converge to the true mass profiles within this range of distances, where the relative differences are of the order of few per cent.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

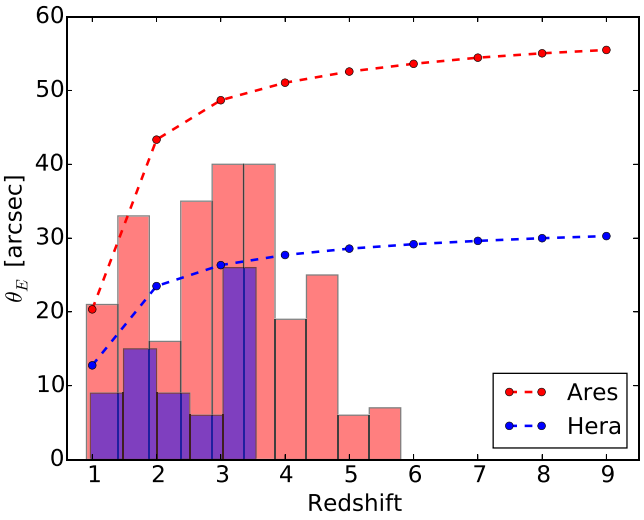

Figure 11. Size of the Einstein radii of *Ares* and *Hera* as a function of the source redshift (red and blue dashed lines, respectively). The histograms show the redshift distribution of the multiple images for the two clusters.

4.3 Shape and orientation

Having quantified the performance of the methods to reconstruct 1D mass profiles, we discuss now their ability to recover the 2D mass distributions of the lenses.

To be more quantitative about how well the methods employed recover the true shape and orientation of the two clusters, we consider the projected mass distributions of the lenses in terms of their iso-surface-density (or convergence, κ) contours. We use the following procedure:

(i) From the convergence maps, we extract the contours corresponding to κ-levels in the range 0.5–3.0. Since both *Ares* and *Hera* have bimodal mass distributions, we use the centre of the largest mass clump as the reference centre for this analysis and we consider only the contours enclosing it.

(ii) We fit an ellipse to each contour and measure its ellipticity and position angle. We also measure the size of each contour by means of an *equivalent radius* r_κ, defined as

$$r_\kappa = \sqrt{ab}\,, \tag{16}$$

where a and b are the semi-axes of the best-fitting ellipse.

(iii) Finally, we draw the radial profiles of both the ellipticity and the position angle. The radius used to produce the profiles is the equivalent radius of the iso-density contours.

The procedure outlined above is shown in Fig. 14 for the cluster *Ares*.

The radial profiles of the ellipticity and of the position angle for the two clusters are shown in Figs 15 and 16. As done in Figs 12 and 13, the results for parametric and free-form methods are displayed separately (left-hand and right-hand panels, respectively).

In each panel, the true profile is given by the black dashed line. The two clusters investigated in this work exhibit quite different ellipticity profiles. Indeed, due to the larger spatial separation between the two mass clumps, *Ares* has a less elongated inner core ($e = 1 - b/a \sim 0.3$) compared to *Hera* ($e \sim 0.7$). *Ares*'s ellipticity increases with radius, while *Hera* shows the opposite trend.

Despite the fact that we have introduced some modest radial variation of the ellipticity of two main mass clumps in *Ares*, the largest jumps in the ellipticity profile of this cluster are produced by massive substructures. These variations of ellipticity are generally well reproduced in the parametric reconstructions, and, to some extent, also in the free-form model of GRALE. Clearly, the parametric techniques produce better measurements of the core

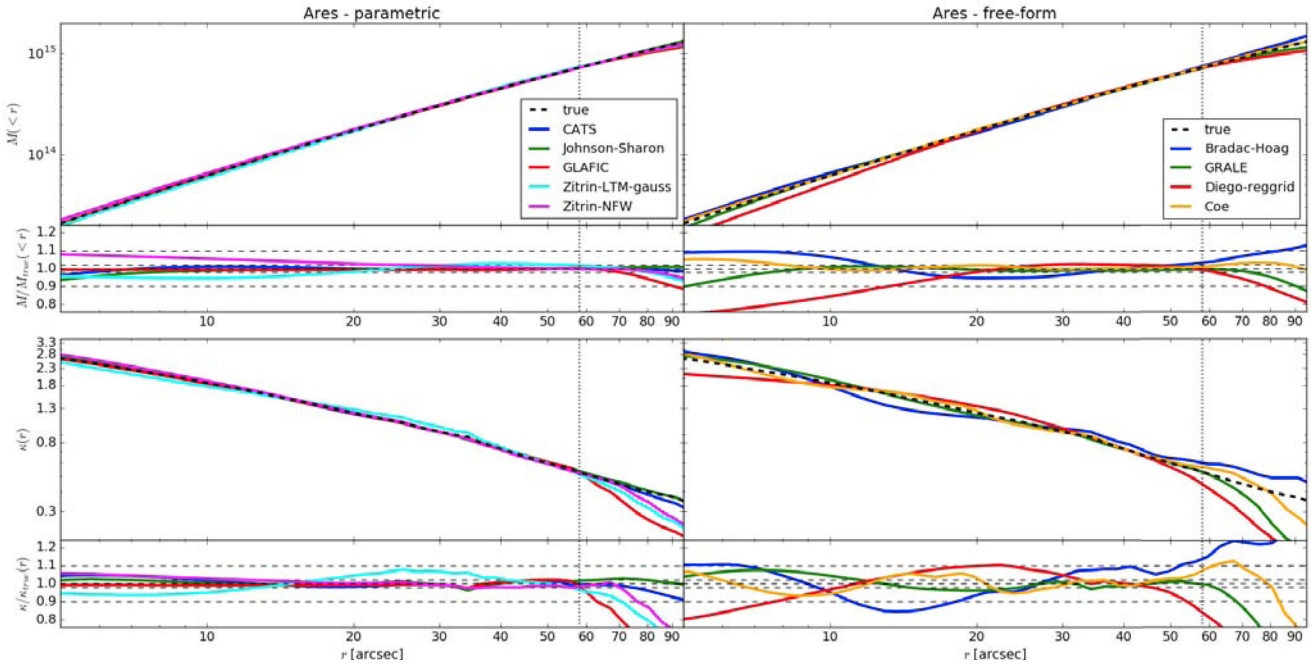

Figure 12. Mass profiles in the inner 100 arcsec of Ares: enclosed mass (upper panels) and mass surface density (lower panels). Results for parametric and free-form methods are shown in the left-hand and in the right-hand panels, respectively. The insets on the bottom of each panel show the ratio between the reconstructed and the true mass profiles. The horizontal dashed lines correspond to ±2 per cent and ±10 per cent differences between lens models and input mass distribution.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

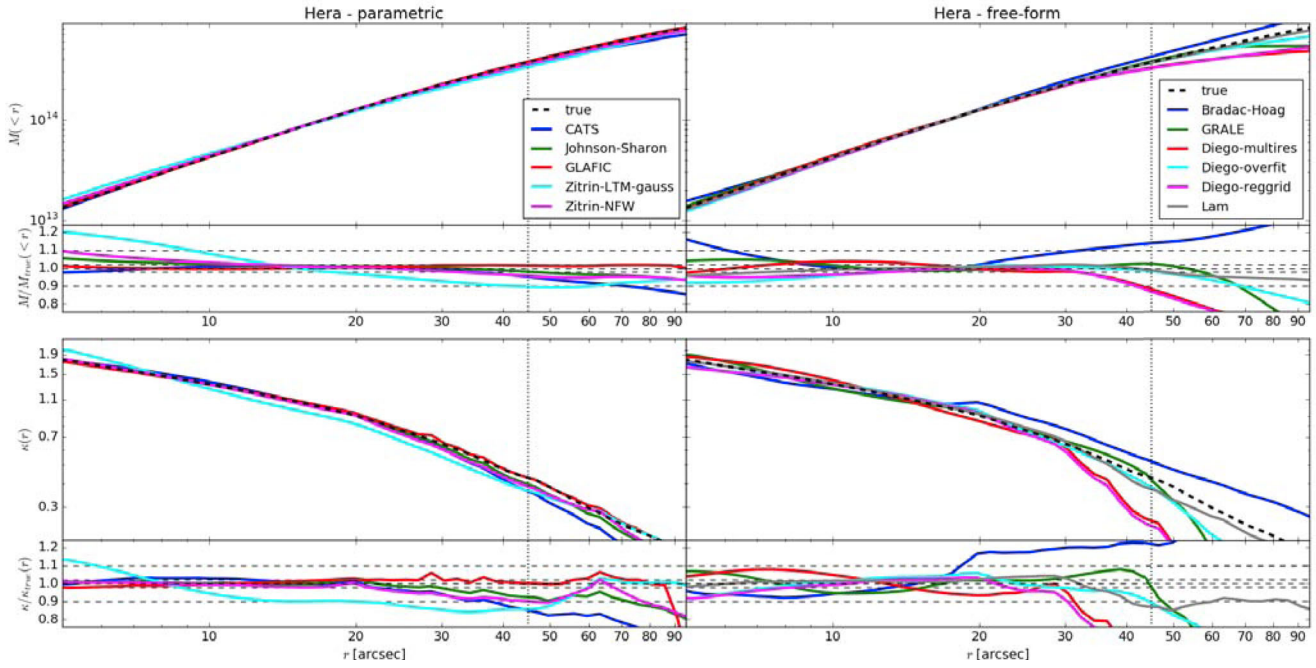

Figure 13. Mass profiles as in Fig. 12 but for *Hera*.

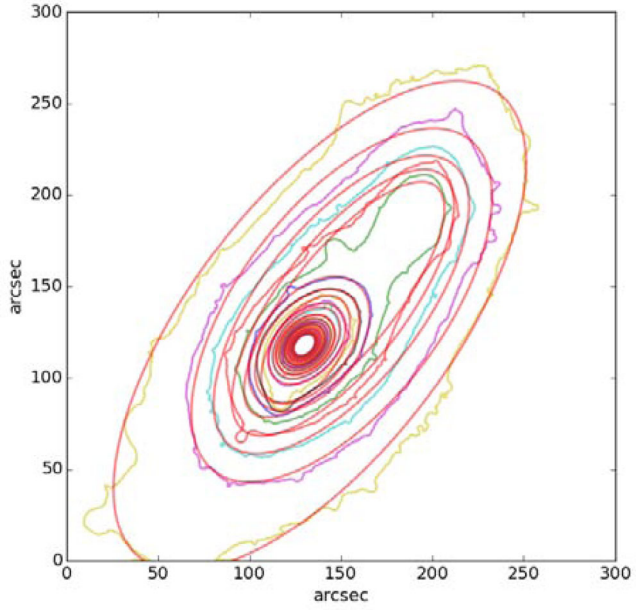

Figure 14. *Ares* mass iso-surface-density contours $\kappa = 0.5$, 1.0, 1.5, 2.0, 2.5, 3.0 for $z_s = 9$ (jagged lines) and elliptical fits in red.

shapes, both in the cases of *Ares* and *Hera*. Indeed, due to resolution limits, the convergence maps produced by the free-form methods are noisier, resulting in more irregular iso-density contours. Under these circumstances, the ellipticity measurements are more uncertain.

Among the parametric reconstructions of *Ares*, the largest deviations from the true ellipticity profile are found for the Zitrin-NFW and for the Zitrin-LTM-gauss models within ∼40 and ∼20 arcsec, respectively. Interestingly, these same algorithms provide some of the most accurate measurements of the core shape in the case of *Hera*. These algorithms generally find higher halo ellipticities compared to the other parametric methods. Such behaviour is consistent with the results of Zitrin et al. (2015), where the Zitrin-NFW and Zitrin-LTM-gauss methods are both employed in the reconstruction of the galaxy clusters in the CLASH sample. As shown in their fig. 3, the first of these two methods leads to more elliptical mass distributions. The most likely interpretation of this behaviour is that external shear compensates the smaller ellipticity of the LTM models.

All parametric methods except the Zitrin-LTM-gauss tend to overestimate the ellipticity of the mass distribution at large radii in the case of *Hera*. We shall recall that all these algorithms fit the data by combining multiple mass components, each of which has a fixed ellipticity. The results show that, within the region probed by strong lensing ($\lesssim$ 40 arcsec for *Hera*), the combination of multiple mass clumps is effective in reproducing the overall ellipticity of the cluster. At larger radii, though, the models are unconstrained and the ellipticity is extrapolated from the inner region. Free-form methods do not show the same trend; their ellipticity profiles are more noisy.

Also the orientation angles of the iso-density contours in the parametric reconstructions deviate from *Hera*'s true orientations at large radii. Being a numerically simulated cluster, *Hera* is characterized by asymmetries and twists of the iso-density contours that result to be much stronger than in *Ares*. For example, the position angle of the iso-density contours changes by ∼20 deg between the very inner region of the cluster and a distance of ∼50 arcsec.

As a result of the not perfectly reproduced shape and orientation of the cluster at large radii, the CATS, Johnson-Sharon, GLAFIC and Zitrin-NFW models have an excess of mass along the major axis of the cluster with respect to the true mass distribution of *Hera* (and consequently they lack mass in the perpendicular direction). Such peculiarities can be seen in Fig. 10, where the ratios between reconstructed and true convergence maps of *Hera* are shown.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

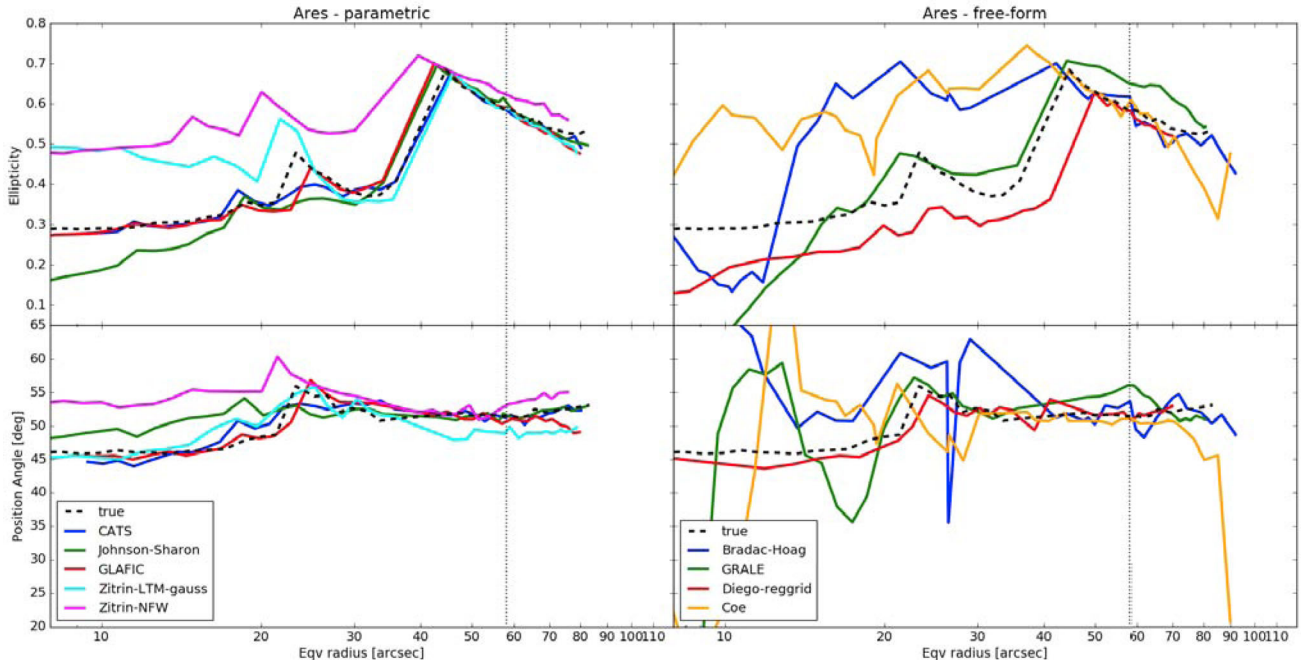

Figure 15. Ellipticity (upper panels) and position angle (lower panels) as a function of the equivalent radius of the convergence contours of *Ares*. In each plot, we show the true profile as a dashed black line. The profiles obtained from the reconstructions are given by the solid coloured lines. The left-hand and the right-hand panels refer to the parametric and to the free-form methods, respectively. The vertical dotted lines indicate the maximum radius covered by all reconstructions.

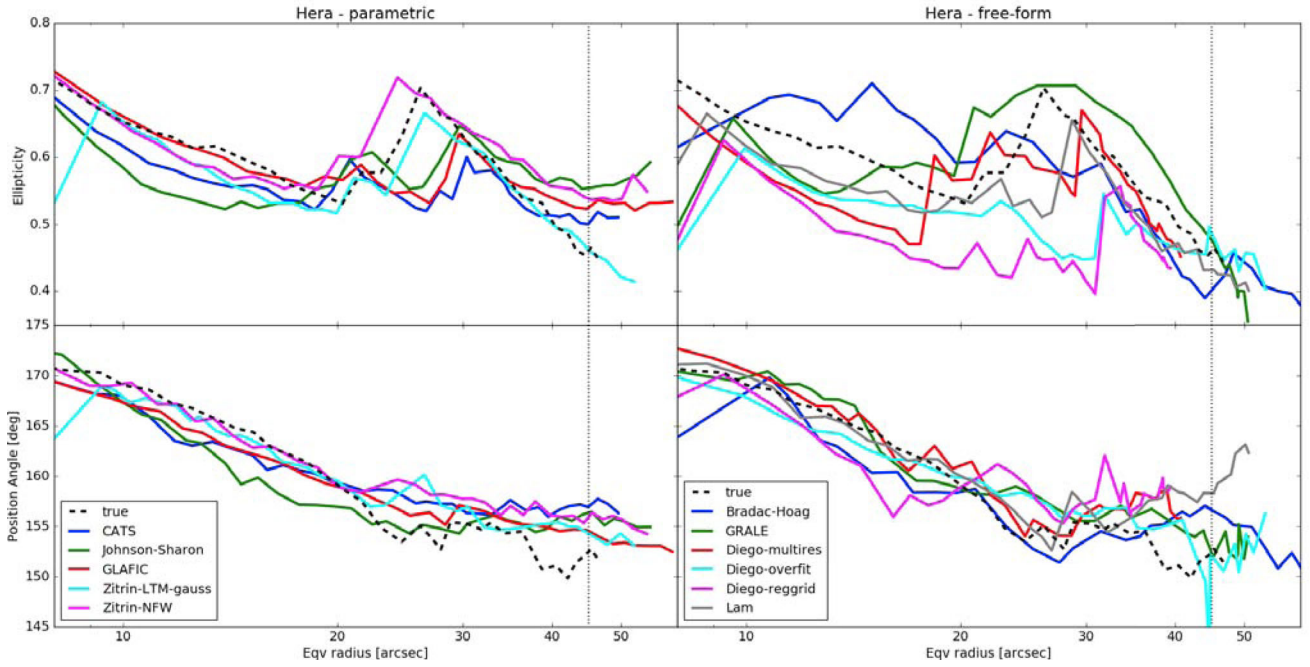

Figure 16. Ellipticity (upper panels) and position angle (lower panels) profiles as in Fig. 15 but for *Hera*.

4.4 Substructure

Figs 9 and 10 show that significant differences exist between the models near substructures. Measuring the mass of substructures is an important task that several authors have performed via strong lensing (see e.g. Natarajan, De Lucia & Springel 2007; Natarajan et al. 2009; Grillo et al. 2015, and references therein). Therefore, it is interesting to quantify the lens model precision near these secondary mass clumps.

From the perspective of strong lensing, substructures are often identified as massive haloes around cluster galaxies. This is particularly true for parametric methods: they use the luminous galaxies as tracers of the underlying mass distribution. Instead, free-form

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

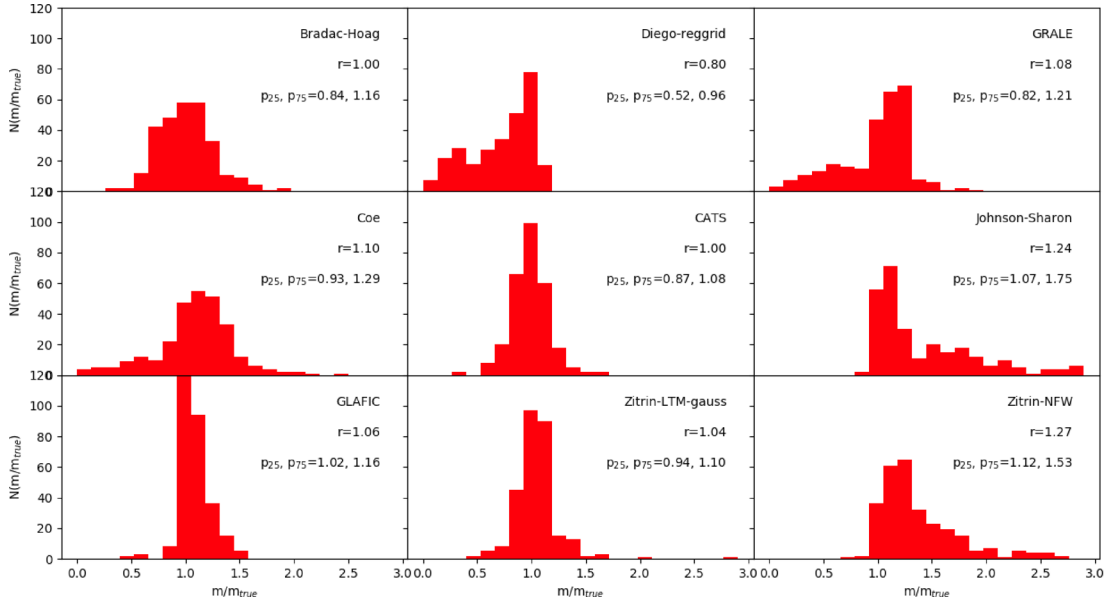

Figure 17. Distributions of the ratios between measured and true substructure masses for *Ares*. The total number of substructures is 282. Each panel shows the results for a mass model. In each panel, we indicate the median r and the 25th and 75th percentiles of the distribution.

methods can in principle detect any kind of mass substructure, even if not traced by light. However, they cannot distinguish between the projected mass belonging to the cluster halo and bound to the substructures.

Indeed, as part of their submissions, the groups did not provide estimates of the masses in substructures, nor substructure catalogues. Here, we perform the following analysis:

(i) We start from the assumption that galaxies trace the substructures. This is not a strong assumption given the method employed to generate the galaxy populations of *Ares* and *Hera*. In both cases, galaxies tend indeed to coincide with dark matter substructures. In the case of *Ares*, there is a one-to-one correspondence between luminous galaxies and dark matter subhaloes. In the case of *Hera*, we have excluded from the image simulations those galaxies that had their dark matter haloes stripped off in the course of the cluster evolution.

(ii) We create apertures centred on the cluster galaxies with $m_{AB,\,F814W} < 24$, with radii equal to twice the effective radius of the galaxies, and we measure the projected mass within each aperture from both the reconstructed and the true convergence maps.

(iii) In the following, we will refer to these masses as substructure masses, keeping in mind that these are, however, the sum of the substructure mass and of the projected mass of the underlying cluster dark matter halo.

In Figs 17 and 18, we show the distributions of the ratios between measured and true substructure masses. The two figures refer to *Ares* and *Hera*, respectively, and show the results for all the models. We characterize the distributions of the ratios r by means of their median r and of their 25th and 75th percentiles, p_{25} and p_{75}. The analysis is carried out on the same areas covered by the maps in Figs 7 and 8. Therefore, the same number of substructures have been used to build the histograms (282 and 278 for *Ares* and *Hera*, respectively).

The results found for *Ares* show that several methods recover nearly unbiased substructure masses with good accuracy. For example, the interpercentile range found for the CATS model is only 0.21 and the median is $r = 1$. Similar results are found for the Zitrin-LTM-gauss model, although with a median slightly larger than unity. Some parametric models, such as those of Johnson-Sharon and Zitrin-NFW and marginally GLAFIC, have skewed distributions with tails extending towards ratios larger than unity. Interestingly, Johnson-Sharon's model is based on the same modelling software employed by the CATS group.

Among the free-form models, the distributions are generally broader than for the parametric methods. The distribution for the Bradac-Hoag model has median $r = 1$ and interpercentile range 0.32. Similar or slightly larger scatter is found for the GRALE and Coe models. The ratio distribution obtained for the Diego-reggrid model has a tail extending towards small values and its median is $r = 0.8$.

The results found for *Hera* are quite in agreement with those found for *Ares*. Parametric methods perform very similarly among each other, providing mass measurements accurate at the level of few per cent. The Zitrin-LTM-gauss model has a median $r = 0.89$. The dispersions of the ratio distributions, as quantified by the interpercentile ranges, are ∼0.2–0.25. This is quite remarkable given the very different methods used to populate *Ares* and *Hera* with substructures and the significant differences between the density profiles of the substructures themselves in the two simulations, as shown in Fig. 1. This seems to indicate that the methods are flexible enough to account for even large variations in the substructure properties, provided they are traced by light. It is less surprising that the flexible free-form methods also behave so similarly in *Hera* and *Ares*.

4.5 Magnification maps

As one of the major goals of the Hubble Frontier Fields is to use the lensing power of galaxy clusters to detect and characterize very high redshift galaxies, we focus now on the magnification. Of course, the results shown in this section are not independent of those discussed

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

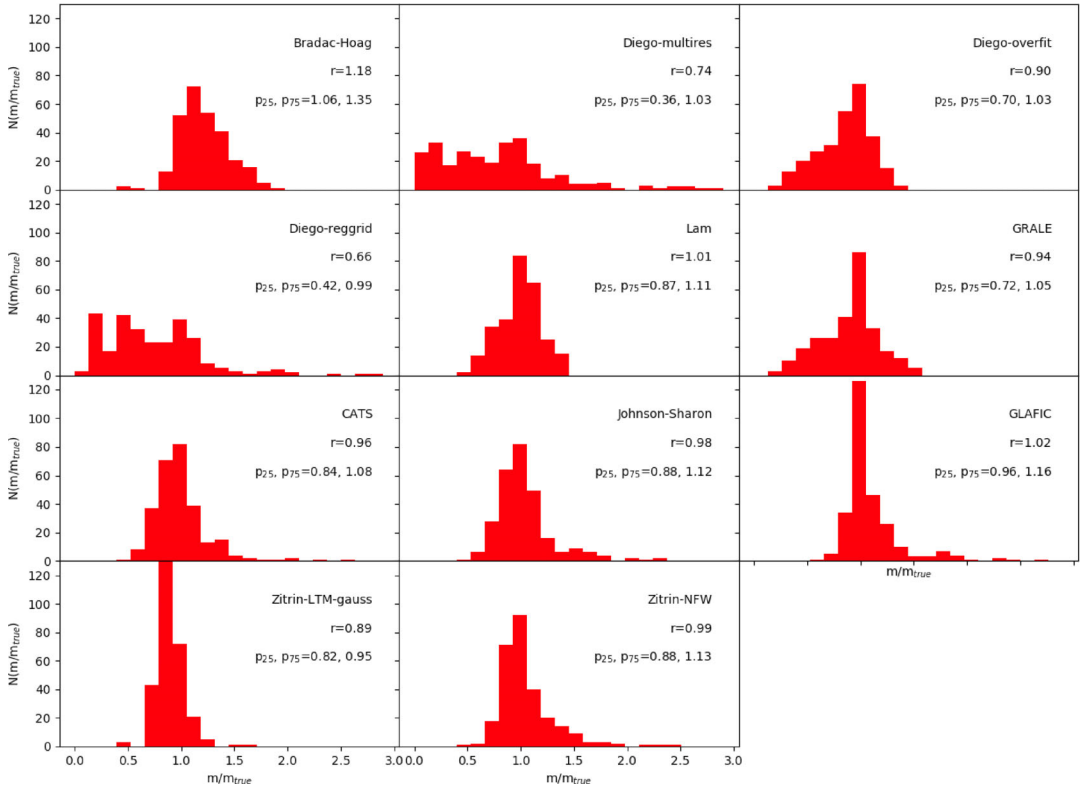

Figure 18. Distributions of the ratios between measured and true substructure masses for *Hera*. The total number of substructures is 282. Each panel shows the results for a mass model. In each panel, we indicate the median r and the 25th and 75th percentiles of the distribution.

earlier, since the convergence is one of the two quantities entering the definition of magnification. The other quantity is the shear, which was not discussed so far.

In Figs 19 and 20, we show the magnification maps for $z_S = 9$ obtained for *Ares* and *Hera*. As done previously, the results for each model are displayed in different panels. The last panel on the bottom shows the true magnification. The relative differences between each reconstructed magnification map and the true magnification maps are shown in Figs 21 and 22.

The largest discrepancies between reconstructed and true magnifications appear around the lens critical lines. These are the loci where the magnification formally diverges. Therefore, even a small misalignment of the true and reconstructed critical lines will result in potentially large magnification differences. Most of the models recover the shape and the size of the critical lines well. Others, as the Bradac-Hoag, the Coe and the Diego-multires models are characterized by critical lines with very irregular shapes.

In Figs 23 and 24, the measured magnifications are plotted as a function of the true magnifications. As anticipated, the scatter around the median increases as a function of the true magnification for all models. The scatters for parametrically reconstructed models of *Hera* are factors of 2–3 larger than for the corresponding models of *Ares*, the mock cluster that was generated parametrically. Besides, we note that *Hera* was inherently less well constrained as the cluster had fewer multiple images than *Ares*. So a slightly lower fidelity in the reconstruction was anticipated and found as expected.

In the best scenario obtained for *Hera* (i.e. the GLAFIC model, see also Fig. 25), we find very high accuracy (a few per cent bias at most) and precision: ∼10 per cent uncertainty for $\mu = 3$, growing to ∼30 per cent at $\mu = 10$, and increasing further at higher magnifications.

In other cases, median magnifications are biased low or high by as much as ∼40–50 per cent. Some of these biases are due to the models' inability to reproduce the correct magnification patterns interior to the tangential critical lines. In other cases, the gradient of the magnification around the critical lines is significantly different from that in the true magnification maps, reflecting the incorrect shape and orientation of the projected mass distribution or the incorrect slope of the convergence profile.

Regions around substructures sometimes are characterized by large uncertainties on magnification estimates. For example, the large substructure located south of the cluster *Hera* is not well constrained by any of the models, which all systematically underestimate the magnification around it. As shown in the upper panel of Fig. 4, there are no multiple images located near this substructure, which may explain why no model is able to constrain its mass properly.

In Fig. 25, we plot the precision and the accuracy of the magnification measurements attained by each model as a function of the model magnification. The results are shown for *Hera* only, and they do not account for the location of the lensing constraints. Fig. 26 shows that the precision achieved by the model at the location of the constraints is higher than in other regions with similar magnifications. The distributions of the relative differences between model and true magnifications at the location of the 65 multiple images in the cluster *Hera* are shown using green boxplots for all the lens models. The median true magnification at the location of the lensing constraints is $\hat{\mu}(z_s = 9) \sim 6$. We sample each model magnification maps at other 65 positions, making sure that the selected points have magnification distributions identical to those of the multiple images. The distributions of the relative differences between model and true magnifications at these locations are shown by the blue boxplots.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

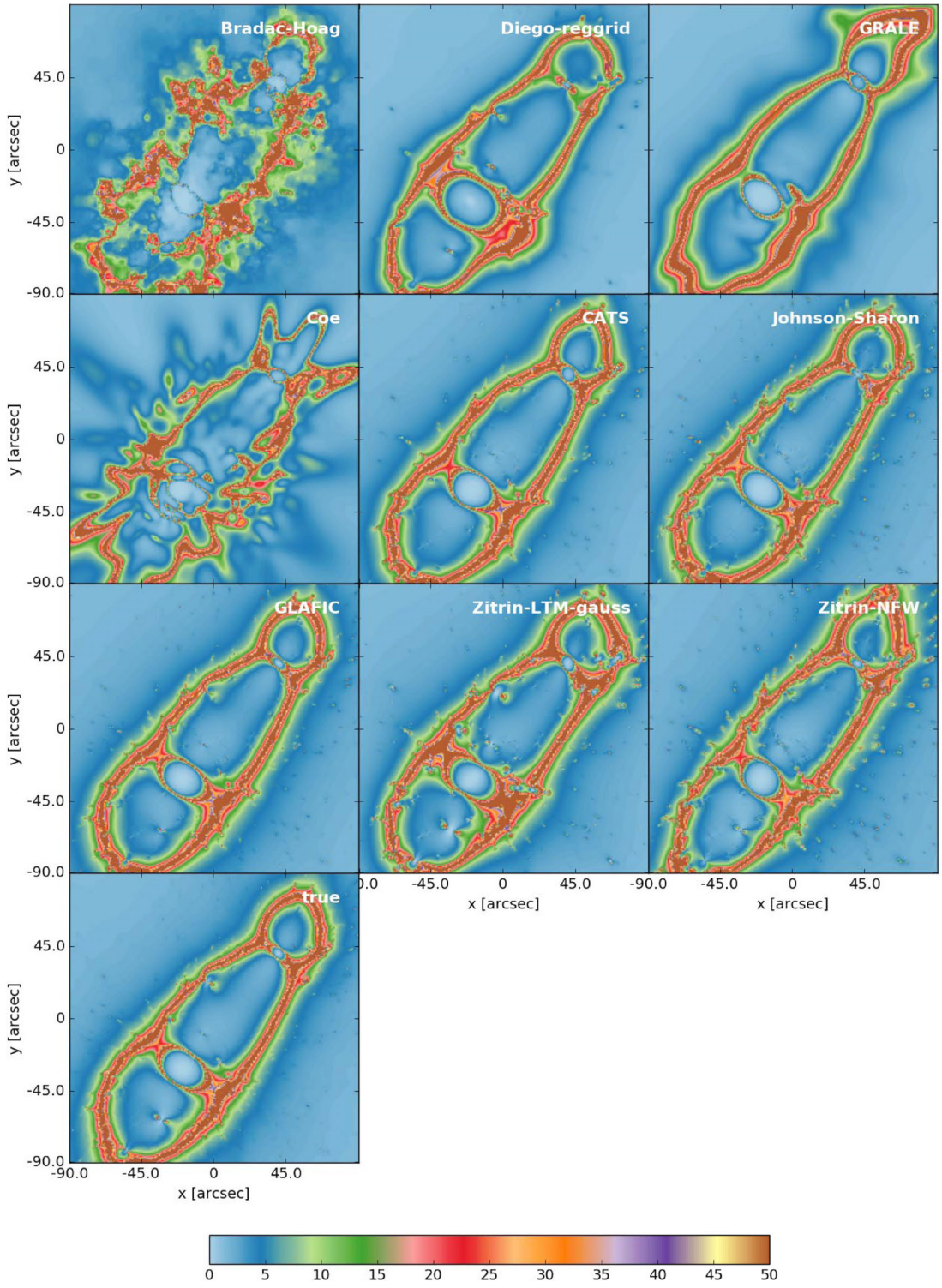

Figure 19. Magnification maps for sources at $z_s = 9$ lensed by each *Ares* reconstruction and, at bottom, the true simulated cluster.

As said above the precision of the models is usually better near the lensing constraints, although the accuracy does not generally improve.

5 RECONSTRUCTION METRICS

In order to be more quantitative in estimating the ability of the different methods employed in this work to measure several relevant properties of *Ares* and *Hera*, we have defined metrics for the lens properties discussed above. More precisely, we introduce metrics for the 1D radial profiles of the following:

(i) the 2D projected mass enclosed within radius R,
(ii) the surface mass density, or convergence $\kappa(R)$,
(iii) the ellipticity, as fit to iso-density contours,

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

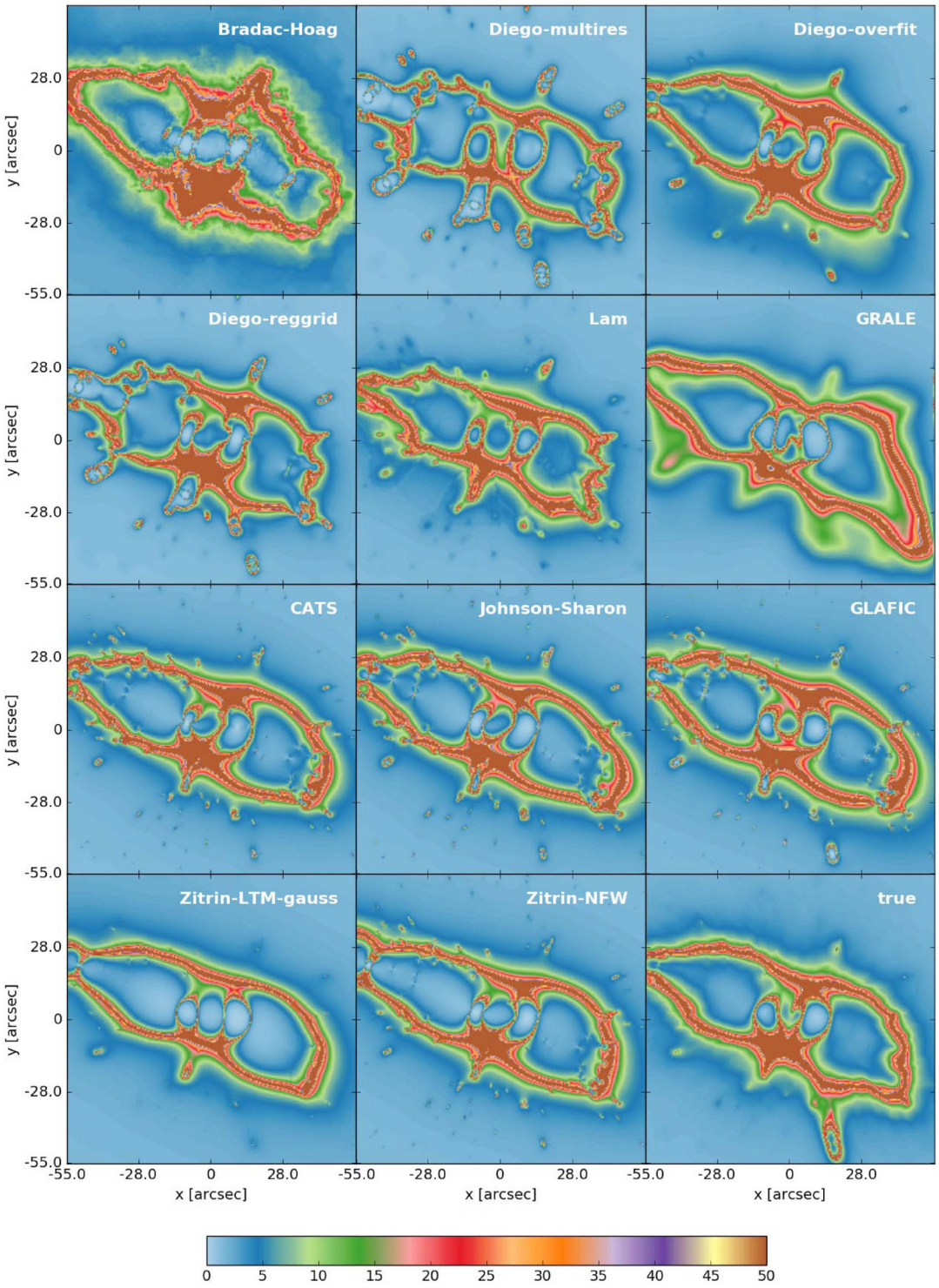

Figure 20. Magnification maps for sources at $z_s = 9$ lensed by each *Hera* reconstruction and the true simulated cluster (bottom right).

(iv) and the orientation, as given by the position angle of the convergence contours.

We also define metrics to quantify the goodness of the reconstruction of the 2D convergence and magnification maps. Finally, we define a metric for the projected subhalo masses in apertures centred on the cluster galaxies.

Thus, we have seven metrics that can be used for a more quantitative comparison between the lens models of both clusters. We can also evaluate how the performance of each algorithm changes when switching from a simulation based on a lens obtained from SAMs (*Ares*) and one obtained from a fully numerical simulation (*Hera*).

The metrics are defined as follows. Given a set of measured values v and a set of true values $v_{\rm true}$, we derive the distribution of $v/v_{\rm true}$. Then, we compute the median, ζ and the 25th and 75th percentiles of the distribution, p_{25} and p_{75}. The metric is finally defined as

$$Q_v = \log_{10}\{[(p_{75} - p_{25})|\zeta - 1|]^{-1}\} \, . \quad (17)$$

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

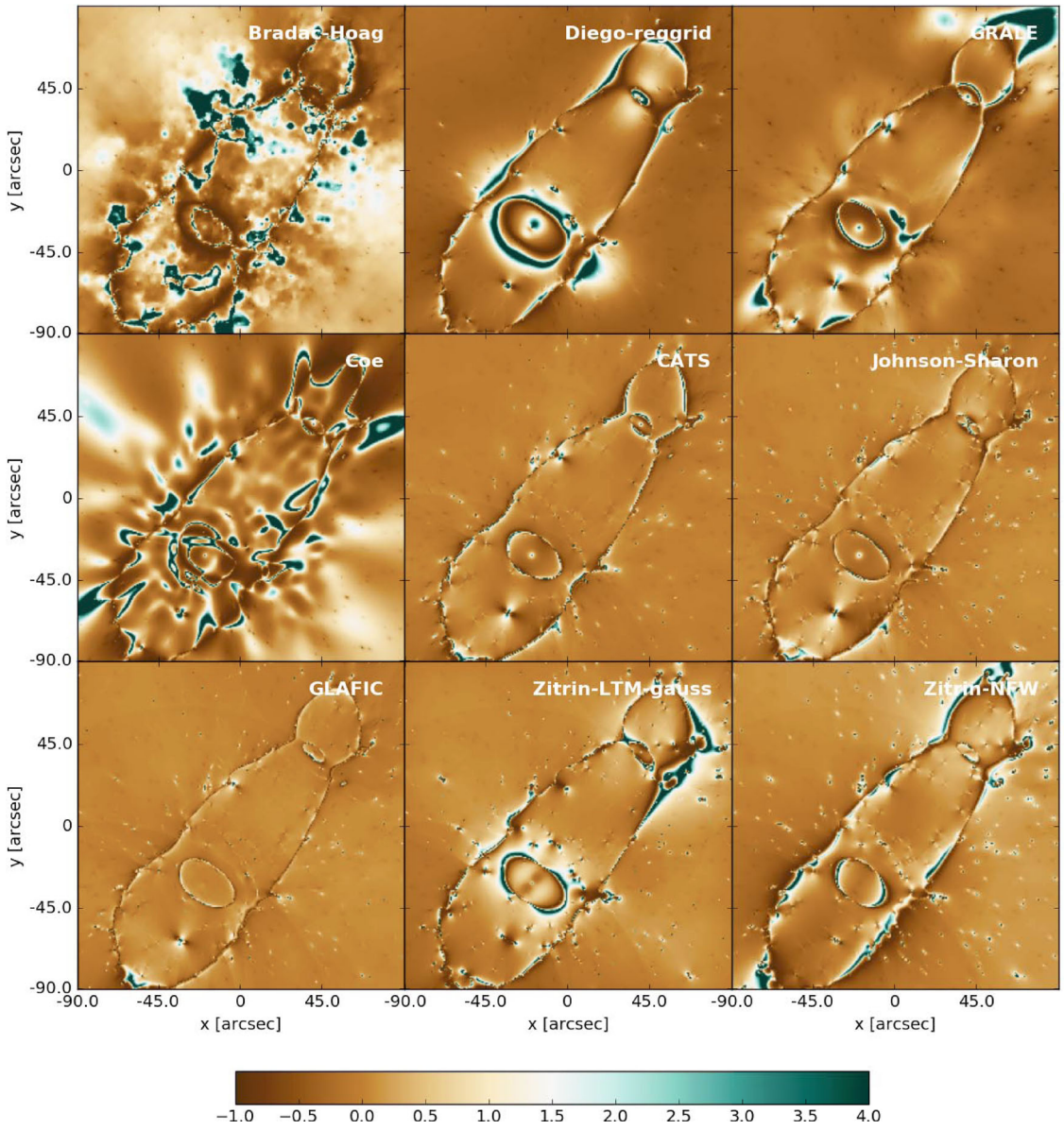

Figure 21. Relative difference between model and true magnifications for *Ares*.

By adopting this definition, the metric penalizes those reconstructions that are biased and/or affected by a large scatter.

Of course, the metrics are not fully independent. For example, a model that is able to reproduce the convergence profile of the lens with a good accuracy will also provide a robust measurement of the mass profile. Similarly, models whose reconstructed convergence maps show little deviation from the true convergence maps will also provide a good match with the simulation in terms of converge profile or shape (ellipticity and position angle). Nevertheless, the ranking among the models with respect to correlated lens properties is not always the same. For example, the Johnson-Sharon reconstruction of *Hera* ranks second in terms of convergence profiles and fourth in terms of mass profiles. In addition, the different lens properties that are discussed here are often used individually, and it may be interesting for the reader to establish which modelling technique is better suited to their scientific purposes.

In Fig. 27, we show radar plots that summarize the metric values recorded by each reconstruction. The overall performance of each model corresponds to the area of each polygon. When one model is good at measuring some of the lens properties, but less effective at capturing others, the polygon appears elongated towards one or more of the chart vertexes.

The first eight charts correspond to free-form or hybrid methods. The remaining five charts refer to parametric techniques. As we have pointed out several times earlier, there is larger discrepancy between the performances of parametric and non-parametric methods in the case of *Ares* than in the case of *Hera*. This leads us to the conclusion that, despite our attempts to make the *Ares* mass distribution less ideal for parametric methods (e.g. by simulating adiabatic contraction or by introducing some twist of the iso-density contours, including some radial dependence), the simple fact that this cluster is assembled by combining mass components traced by the cluster galaxies, consistently with the basic assumptions of most parametric techniques, gives a huge advantage to these methods. The good news, in this case, is the following. First, these algorithms work as they are supposed to. Secondly, they provide very accurate

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

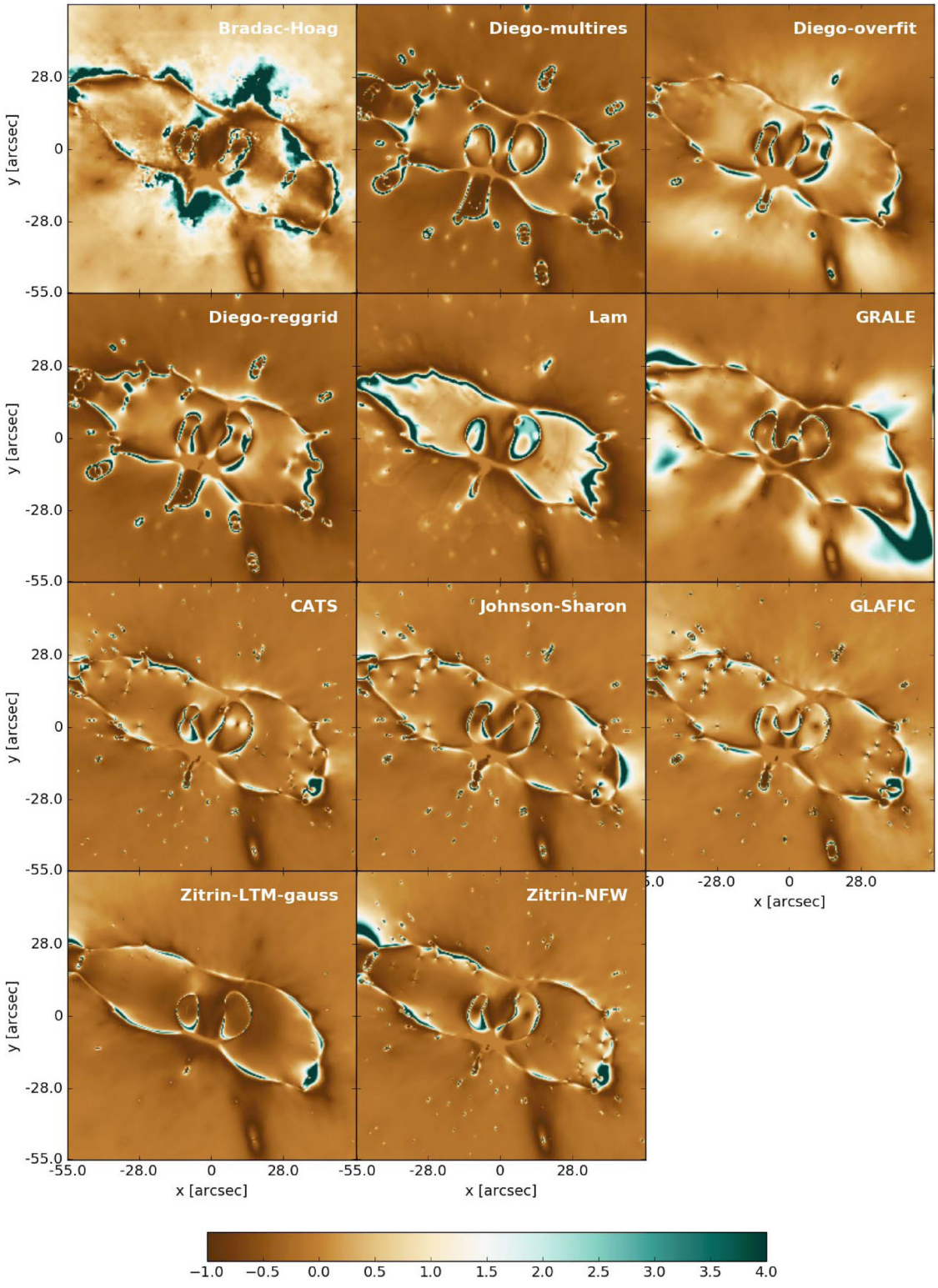

Figure 22. Relative difference between model and true magnifications for *Hera*.

reconstructions even if the parametrization chosen for the lens halo density profile is not fully consistent with the true profile of the lens. For example, none of the parametric techniques, except the Zitrin-NFW method, used the NFW profile for fitting the smooth dark matter halo components of *Ares*. Even so, models such as those submitted by the CATS, Johnson-Sharon and GLAFIC teams produce an overall better fit to the input mass distribution compared to the Zitrin-NFW reconstruction. This suggests that pseudo-elliptical, cored halo models provide the right flexibility to account for most of the effects we have introduced in the simulation, such as the adiabatic contraction, which steepens the density profile in the central region of the cluster. Alternatively, these results may be interpreted

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

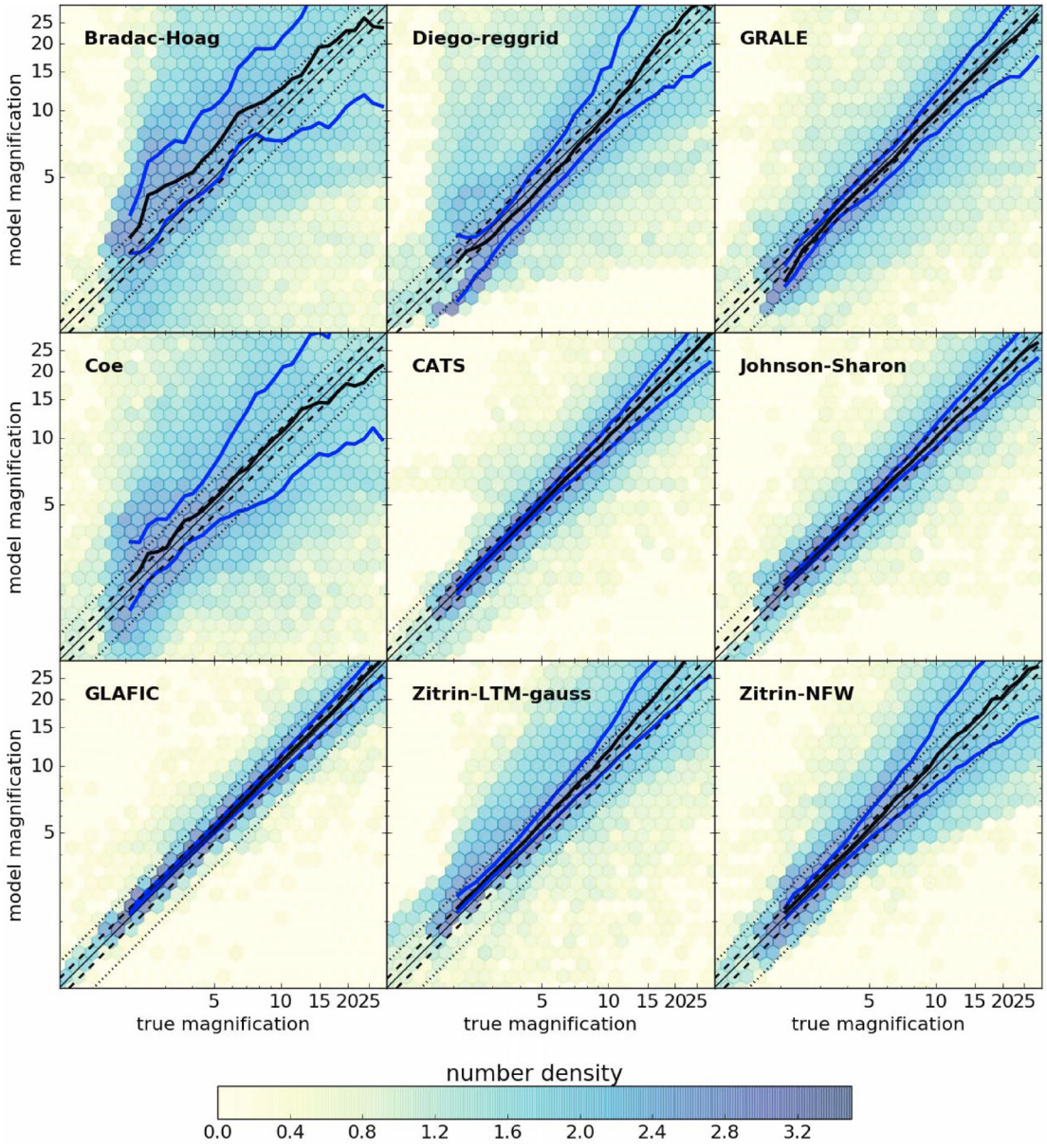

Figure 23. Model versus true magnifications ($z_s = 9$) for *Ares*. The underlying 2D histograms show the distributions of the pixel values on the μ–$\mu_{\rm true}$ plane after sampling the magnification maps on a grid of 256 × 256 pixels. The black and the blue solid lines show the median and the 25th and 75th percentiles of the measured magnifications in bins of $\mu_{\rm true}$. The dashed and the dotted lines parallel to the diagonal in each panel denote the limits of ±10 per cent and ±30 per cent deviations from the relation $\mu = \mu_{\rm true}$.

as evidence for a lack of sensitivity of lensing alone to the precise share of the halo density profiles, being mostly sensitive to the mass enclosed within the Einstein radius rather than the slope of the density profile. Another possible cause may be that the Zitrin's models are calculated on a low-resolution grid and perhaps their accuracy is limited by this resolution compared to higher resolution or completely analytic parametrizations.

When switching to a fully numerical simulation, the differences between parametric and free-form methods become weaker. At least for some of the metrics, some free-form/hybrid reconstructions of *Hera* (see e.g. the GRALE or Lam models) appear to be as good as the best parametric reconstructions of this cluster. This indicates that several parametric methods still cannot fully account for deviations of the mass distributions from a symmetric shape, which are, instead, more naturally captured by free-form methods. Asymmetries could be mimicked by suitable combinations of substructures in parametric models. Indeed, a degeneracy exists between these two properties of the mass distribution. However, the number of constraints in these simulations is high enough that this degeneracy is partially broken, as shown by how well the mass is constrained around the cluster galaxies in at least some of the parametric reconstructions.

The model provided by the CATS team for *Hera* has significantly smaller values of all metrics (except for the cluster orientation), compared to the model submitted by the same team for *Ares*. The metrics agree with those of other parametric reconstructions of the same cluster (e.g. Johnson-Sharon). On the contrary, the reconstructions provided by the GLAFIC team for the two clusters

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

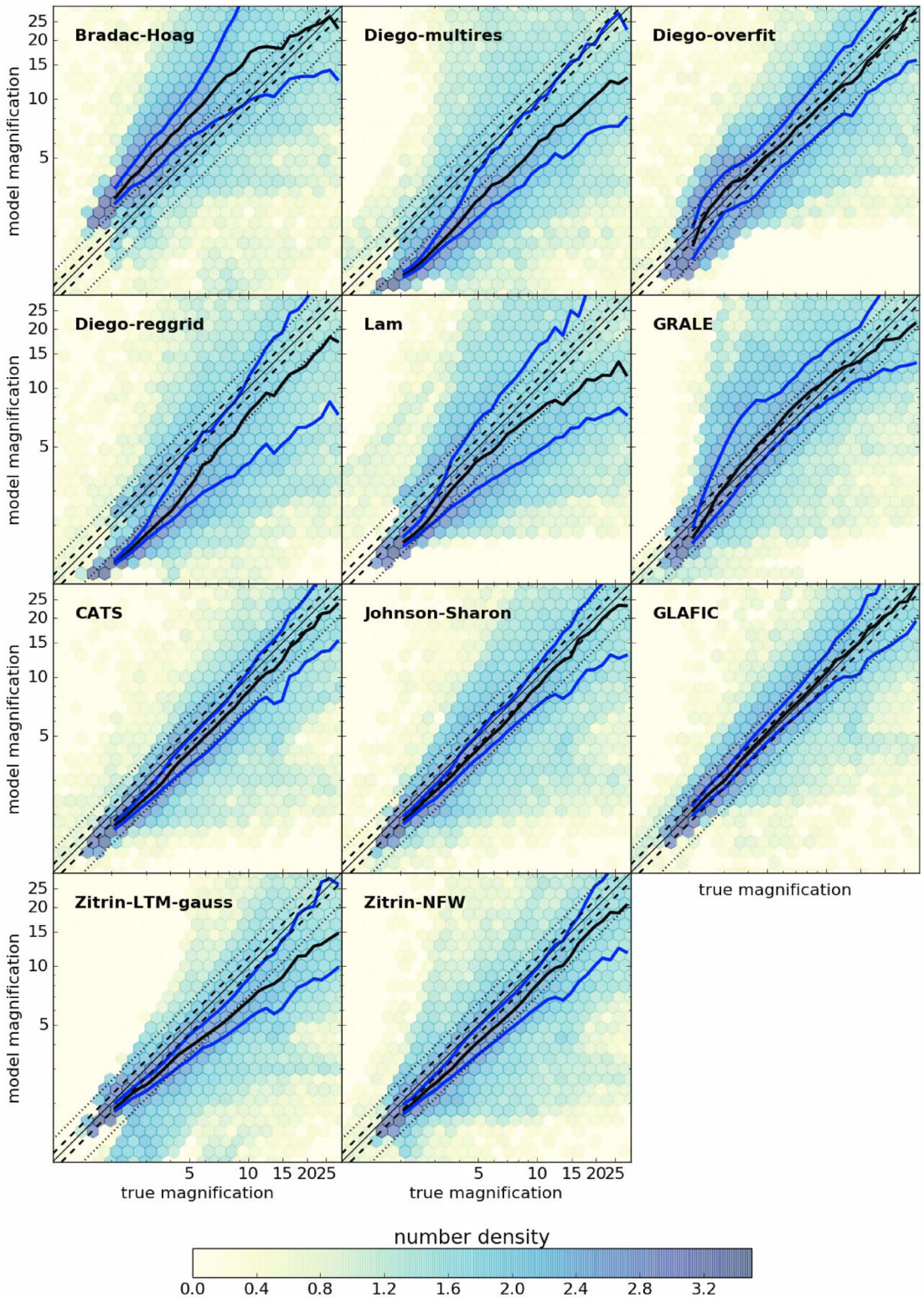

Figure 24. Model versus true magnifications ($z_s = 9$) for *Hera*, as in Fig. 23.

have quite consistently high metric values. One feature of GLAFIC, which was enabled in the reconstruction of *Hera*, is the inclusion of external shear and third-order multipoles of the mass distribution. Apparently, these additional ingredients have provided the GLAFIC model extra degrees of freedom to properly account for the asymmetric mass distribution of *Hera*.

The comparison between the metrics of parametric and free-form methods also shows that the latter techniques are generally less accurate in reconstructing the 2D maps of convergence and magnification and in measuring the mass around substructures. In fact, the spatial resolution that can be achieved with these methods is generally lower. On the contrary, radial profiles of the convergence

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

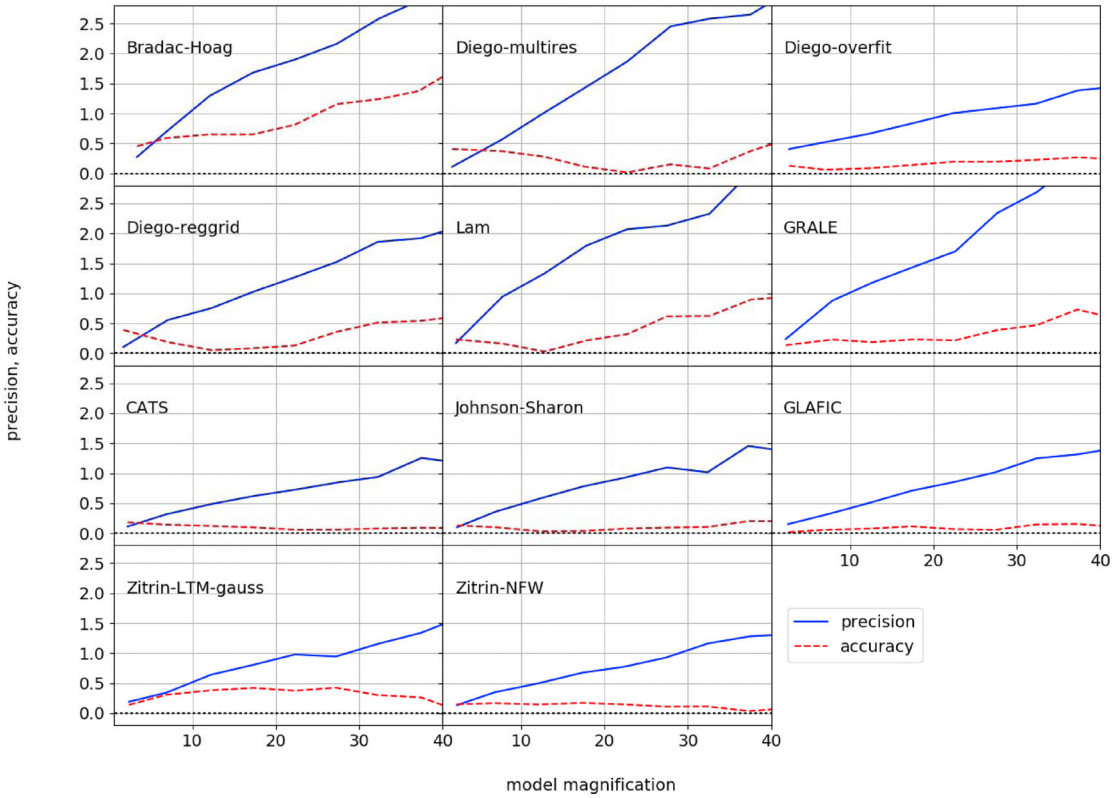

Figure 25. Magnification accuracy (dashed line) and precision (solid line) as a function of the magnification from the strong lensing constraints for the GLAFIC reconstruction of *Hera*. The precision is quantified by the difference between the 75th and 25th percentiles of the distribution of $\mu - \mu_{\rm true}$, sampled on a 256 × 256 pixel grid. The accuracy is given by the median of $\mu - \mu_{\rm true}$.

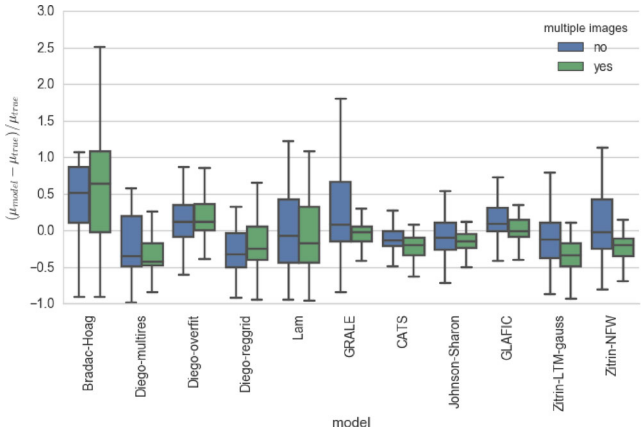

Figure 26. Comparison between accuracy and precision of the magnification measurements at the location of the multiple images used to build the lens models (green) and at other locations with similar model magnifications (blue).

and of the enclosed mass are measured by several of the free-form methods employed in this experiment with accuracy comparable to parametric techniques.

6 LIMITATIONS OF THIS TEST

We would like to remark that the tests outlined in this paper suffer of some limitations. First of all, we make the assumption that the simulations reproduce the properties of real clusters. While some methods (e.g. the free-form ones) do not care about the correlation between dark matter and baryons, other methods strongly rely on the assumption that LTM. Both *Ares* and *Hera* implement this property, which, at least in some cases, has been questioned by observations (Hoag et al. 2015; Wang et al. 2015). In particular, the results we report on substructures are sensitive to this assumption. In a recent paper, Harvey, Kneib & Jauzac (2016) have explored how assuming that LTM in strong gravitational models can lead to systematic errors in the predicted positions of multiple images. They find that images can be shifted by up to ∼1 arcsec, assuming physically motivated offsets between dark matter and stars. They quote a ∼0.5 arcsec rms error in the position of the multiple images due to breaking the assumption that mass traces light. Note, however, that, to some extent, we introduced some misalignment between matter and light in both *Ares* and *Hera*, by assigning to the observed galaxies a shape and an orientation that are not correlated with the underlying dark matter distribution.

Other limitations regard some observable properties of the galaxies in the simulated observations (e.g. luminosities and sizes) and their correlation with their halo masses. It is known that the SAMs are not fully consistent with observations in this respect (see e.g. González et al. 2009; Ascaso, Mei & Benítez 2015; Xie et al. 2015; Hirschmann, De Lucia & Fontanot 2016), and thus the standard scaling relations adopted by some parametric techniques to translate the light into the mass or the size of the host halo might not equally applicable to observations and simulations.

7 SUMMARY AND CONCLUSIONS

In this paper, we used simulated observations of two synthetic galaxy clusters to evaluate the performance of several algorithms

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

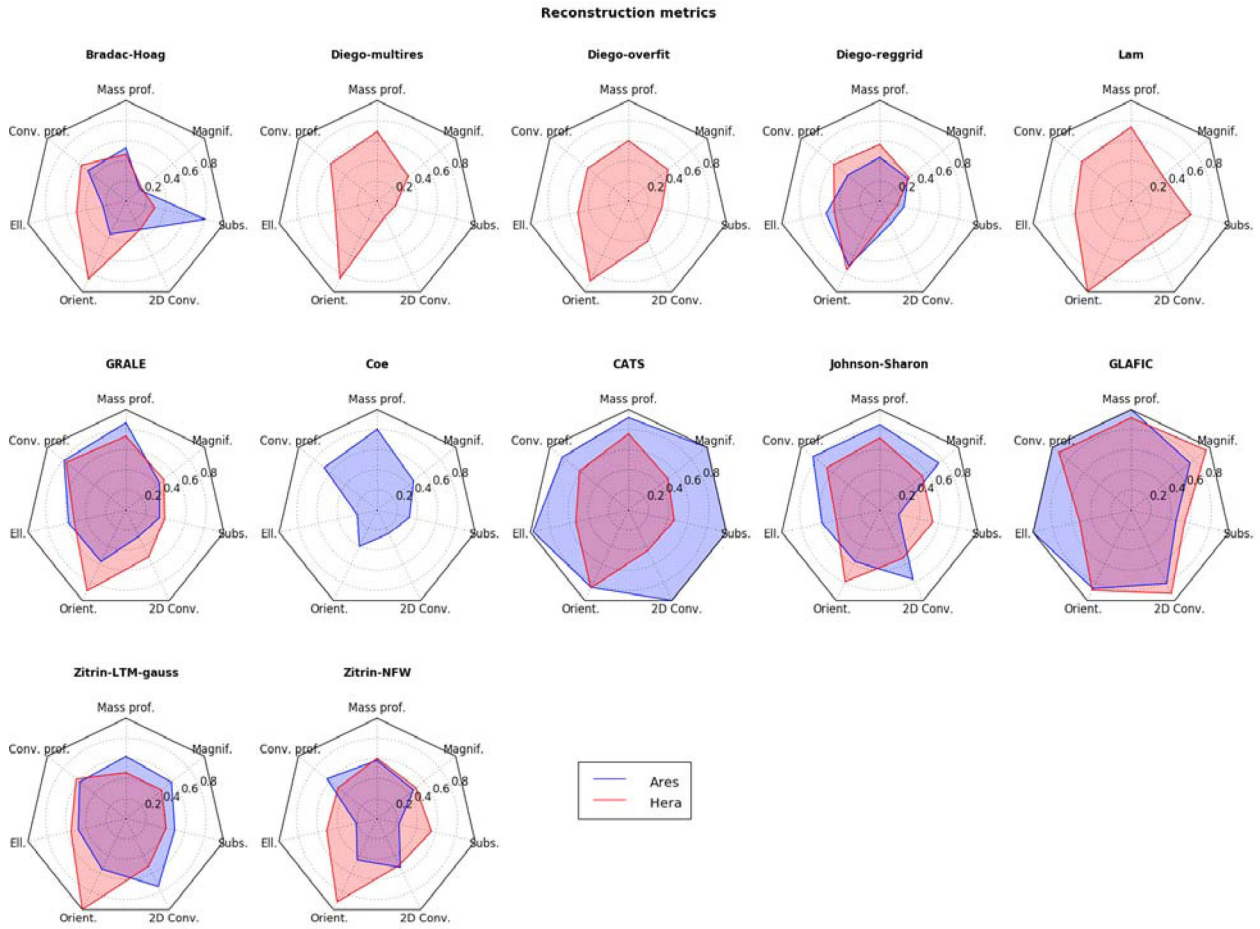

Figure 27. Radar plot showing the scores of each model for all metrics discussed in the paper. Larger polygons correspond to better overall performance. Each chart corresponds to a different lens model (see labels on the top) and shows results for both *Ares* (blue) and *Hera* (red), or whichever is available. The seven metrics are shown on the vertices of each chart. For each metric, the scores range from 0 (worst; plotted at the centre of the chart) to 1 (best; plotted at the vertex), normalized to the maximum value recorded by all models. A filled polygon is obtained by connecting the plotted scores of all metrics for each reconstruction.

for mass reconstruction with strong lensing. Such algorithms are currently being used to deliver to the community the lens models for the six galaxy clusters being observed in the Frontier Fields programme of the *HST*.

The two clusters used in this study were obtained using very different techniques. *Ares* was generated using the semi-analytical code MOKA. *Hera* is instead the output of a cosmological *N*-body simulation at high resolution. The observable properties of the cluster galaxies are modelled using HOD and SAM techniques in *Ares* and *Hera*, respectively. In both cases, the clusters have complex mass distributions, characterized by disturbed and bimodal morphology, similar to those of the FFI clusters.

We used the code SKYLENS to simulate *HST* observations of the two mock clusters with both the ACS and the WFC3-IR camera. We produced images in all photometric bands used in the FFI, calibrating the exposure times such to reach the depth of the FFI observations. These *HST* simulated data were distributed to several groups of lens modellers for a blind analysis, i.e. without unveiling the true mass distribution of the lenses, neither the method used to simulate them.

The simulated observations include lensing effects on a realistic distribution of background galaxies. We identified many strongly lensed galaxies and built a catalogue of multiple image systems, which was delivered together with the simulated observations. The catalogues also include the redshift of all the sources.

We complemented the *HST* simulations with a simulated observation in the R_c band with the Subaru telescope. The main purpose of this additional simulation was to allow the inclusion of weak-lensing constraints at larger distances from the cluster centre than those probed by *HST*. Together with the image, we also distributed a shear catalogue obtained by processing the Subaru simulation through a public KSB pipeline.

We received 9 reconstructions of *Ares* and 11 reconstructions of *Hera*, submitted by 10 different groups. Seven groups employed their techniques to reconstruct both clusters. The remainder groups reconstructed just one of the two clusters or submitted reconstructions based on different set-ups of their methods. This is the first time that such a large number of algorithms have been tested against known mass distributions. Similar to the spirit of our experiment, in the recent collaborative effort presented in Treu et al. (2016), several of the methods used to reconstruct the galaxy cluster MACSJ1149.5+2223 and to estimate the time delays between the multiple images of the SN 'Refsdal were compared. The recent reappearance of the SN, reported by Kelly et al. (2016), enabled the blind test of various model predictions, which were found to be in very accurate for several reconstructions. In addition, Rodney

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

et al. (2015) compared the magnification predictions from 17 mass models of Abell 2744 using a lensed Type Ia supernova.

The methods compared here include both parametric and free-form algorithms. We have investigated how they perform at recovering several properties of the lenses, namely: the radial profiles of the convergence and of the enclosed mass, the mass in substructures, the maps of the convergence and of the magnification. For each of these properties, we defined a metric aimed at quantifying the performance of the method.

The key results of this phase of the comparison exercise of lens mapping methodologies can be summarized as follows.

(i) Parametric methods are generally better at capturing 2D properties of the lens cores (shape, local values of the convergence and of the magnification). The free-form methods are as competitive as the parametric methods to measure convergence and mass profiles. It is worth mentioning, however, that, in both *Ares* and the *Hera*, the cluster galaxies were good tracers of the cluster mass distributions.

(ii) The accuracy and precision of strong lensing methods to measure the mass within the Einstein radius (or more generally within the region probed by the strong lensing constraints) is very high. The measured profiles deviate from the true profiles by only a few per cent at these scales. Of course, larger deviations are found at radii larger and smaller than the Einstein radius. The determination of the mass enclosed within the Einstein radius was extremely robust for all methods.

(iii) The largest uncertainties in the lens models are found near substructures and around the cluster critical lines. For some of the parametric models, the total mass around substructures (identified by cluster galaxies) is constrained with an accuracy of $\sim$10 per cent. However, other methods have much larger scatter. Uncertainties on the magnification grow as a function the magnification itself and are therefore more pronounced near the cluster critical lines. For the best-performing methods, the accuracy in the magnification estimate is $\sim$10 per cent at $\mu_{\rm true} = 3$ and degrades to $\sim$30 per cent at $\mu_{\rm true} = 10$.

(iv) Switching from *Ares* to *Hera*, i.e. from a purely parametric to a more realistic lens mass distribution, the gap between parametric and free-form methods becomes smaller. Algorithms such as that used by the GLAFIC team, which include third-order multipoles in the lens mass distribution, have extra degrees of freedom which allow them to better reproduce asymmetries. These asymmetries, and possible variations of the halo ellipticity as a function of radius, seem to be the strongest limitations of parametric methods. The adoption of an hybrid approach, where parametric and free-form methods are combined also to describe the large-scale component of the clusters, could lead to a significant improvement of the mass reconstructions.

(v) Some of the participating groups used the same code but adopted different set-ups to run them. For example, two groups (CATS and Johnson-Sharon) use the public code LENSTOOL with slight modifications. Similarly, Diego submitted several models of *Hera* using WSLAP+, which is the same code used by Lam et al. (2014). Despite using the same algorithms and using the same inputs (i.e. families of multiple images and redshifts), the reconstructions obtained by these groups are different, indicating that some choices made by the modellers when ingesting the data and hence set up priors influence the results. For example, the differences between the reconstructions performed with LENSTOOL by the CATS and Johnson-Sharon teams may be in part due to the priors on the normalization of the scaling relations used to model the cluster galaxies.

This is the first of a series of papers in which we address the issue of the accuracy of lens modelling. In a second paper, currently in preparation, we will discuss the results of the unblinded modelling of *Ares* and *Hera*. The feedback from the unblinding was used by modellers to not only tweak their best fits to reach the best possible match to the input mass distributions of the lenses but to also incorporate and instigate improvements in their modelling procedure. This will provide information on the accuracy limits achievable by each method and will also give further hints on the steps that need to be taken to optimize reconstructions.

Despite their complexity and the inclusion of several observational effects, the simulations used in this paper are still idealized in many respects. For example, the lenses are isolated and no additional lensing by matter along the line of sight is included. In addition, we alleviated the work of the lens modellers by identifying the strongly lensed sources and even providing redshifts for all of them. In the case of *Ares*, the number of available multiple images with known redshifts exceeds by a factor of $\sim$3–4 what is available in any of the Frontier Fields (e.g. MACSJ0416). We will include the uncertainties due to possible misidentification of multiple images and photometric redshifts as well as the noise added in by the intervening matter distribution along the line of sight in the next phase of this project in future work.

ACKNOWLEDGEMENTS

We thank T. Treu for the helpful discussion. MM acknowledges support from the Italian Ministry of Foreign Affairs and International Cooperation, Directorate General for Country Promotion, from INAF via PRIN-INAF 2014 C.R.A. 1.05.01.94.02, and from ASI via contract ASI/INAF/I/023/12/0. This work was supported in part by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan and JSPS KAKENHI Grant Number 26800093 and 15H05892. AZ is supported by NASA through Hubble Fellowship grant #HST-HF2-51334.001-A awarded by STScI, which is operated by the Association of Universities for Research in Astronomy, Inc. under NASA contract NAS 5-26555. JMD acknowledges support of the consolider project CSD2010-00064 and AYA2012-39475-C02-01 funded by the Ministerio de Economia y Competitividad, Spain. We acknowledge the lens modelling community for enthusiastically participating in this collaborative project to compare and contrast mass models. Finally, we want to thank the anonymous referees for their comments and suggestions that helped to improve the quality of the manuscript significantly.

REFERENCES

Ascaso B., Mei S., Benítez N., 2015, MNRAS, 453, 2515
Atek H. et al., 2015a, ApJ, 800, 18
Atek H. et al., 2015b, ApJ, 814, 69
Bartelmann M., Meneghetti M., 2004, A&A, 418, 413
Bartelmann M., Narayan R., Seitz S., Schneider P., 1996, ApJ, 464, L115
Bartelmann M., Limousin M., Meneghetti M., Schmidt R., 2013, Space Sci. Rev., 177, 3
Bouwens R. J. et al., 2014, ApJ, 793, 115
Bouwens R. J., Illingworth G. D., Oesch P. A., Caruana J., Holwerda B., Smit R., Wilkins S., 2015, ApJ, 811, 140
Bradač M., Schneider P., Lombardi M., Erben T., 2005, A&A, 437, 39
Bradač M. et al., 2006, ApJ, 652, 937
Bradač M. et al., 2009, ApJ, 706, 1201
Broadhurst T. et al., 2005, ApJ, 621, 53
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Cain B., Bradac M., Levinson R., 2016, MNRAS, 463, 4287

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, AJ, 132, 926
Coe D., Fuselier E., Benítez N., Broadhurst T., Frye B., Ford H., 2008, ApJ, 681, 814
Coe D., Benítez N., Broadhurst T., Moustakas L. A., 2010, ApJ, 723, 1678
Contini E., De Lucia G., Villalobos Á., Borgani S., 2014, MNRAS, 437, 3787
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
Diego J. M., Protopapas P., Sandvik H. B., Tegmark M., 2005, MNRAS, 360, 477
Diego J. M., Tegmark M., Protopapas P., Sandvik H. B., 2007, MNRAS, 375, 958
Diego J. M. et al., 2015a, MNRAS, 446, 683
Diego J. M., Broadhurst T., Molnar S. M., Lam D., Lim J., 2015b, MNRAS, 447, 3130
Diego J. M., Broadhurst T., Zitrin A., Lam D., Lim J., Ford H. C., Zheng W., 2015c, MNRAS, 451, 3920
Diego J. M. et al., 2016, MNRAS, 456, 356
Donahue M. et al., 2014, ApJ, 794, 136
Donnarumma A. et al., 2011, A&A, 528, A73
Dutton A. A., Treu T., 2014, MNRAS, 438, 3594
Ebeling H., Ma C. J., Kneib J., Jullo E., Courtney N. J. D., Barrett E., Edge A. C., Le Borgne J., 2009, MNRAS, 395, 1213
Eke V. R., Cole S., Frenk C. S., 1996, MNRAS, 282, 263
Elíasdóttir Á. et al., 2007, preprint (arXiv:0710.5636)
Fasano G. et al., 2010, MNRAS, 404, 1490
Fort B., Prieur J. L., Mathez G., Mellier Y., Soucail G., 1988, A&A, 200, L17
Fuselier E. J., 2006, PhD thesis, Texas A&M
Fuselier E., 2008, Adv. Comput. Math., 29, 263
Giocoli C., Tormen G., Sheth R. K., van den Bosch F. C., 2010, MNRAS, 404, 502
Giocoli C., Meneghetti M., Bartelmann M., Moscardini L., Boldrin M., 2012a, MNRAS, 421, 3343
Giocoli C., Meneghetti M., Ettori S., Moscardini L., 2012b, MNRAS, 426, 1558
Gonzalez A. H., Zabludoff A. I., Zaritsky D., 2005, ApJ, 618, 195
González J. E., Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., 2009, MNRAS, 397, 1254
Grillo C. et al., 2015, ApJ, 800, 38
Harvey D., Kneib J. P., Jauzac M., 2016, MNRAS, 458, 660
Hernquist L., 1990, ApJ, 356, 359
Heymans C. et al., 2006, MNRAS, 368, 1323
Hirschmann M., De Lucia G., Fontanot F., 2016, MNRAS, 461, 1760
Hoag A. et al., 2015, ApJ, 813, 37
Hoag A. et al., 2016, ApJ, 831, 182
Huang K.-H. et al., 2016, ApJ, 823, L14
Ishigaki M., Kawamata R., Ouchi M., Oguri M., Shimasaku K., Ono Y., 2015, ApJ, 799, 12
Jauzac M. et al., 2014, MNRAS, 443, 1549
Jauzac M. et al., 2015, MNRAS, 452, 1437
Jauzac M. et al., 2016, MNRAS, 457, 2029
Jing Y., Suto Y., 2002, ApJ, 574, 538
Johnson T. L., Sharon K., 2016, ApJ, 832, 82
Johnson T. L., Sharon K., Bayliss M. B., Gladders M. D., Coe D., Ebeling H., 2014, ApJ, 797, 48
Jullo E., Kneib J.-P., 2009, MNRAS, 395, 1319
Jullo E., Kneib J.-P., Limousin M., Elíasdóttir Á., Marshall P. J., Verdugo T., 2007, New J. Phys., 9, 447
Kaiser N., Squires G., Broadhurst T., 1995, ApJ, 449, 460
Karman W. et al., 2015, A&A, 574, A11
Kawamata R., Ishigaki M., Shimasaku K., Oguri M., Ouchi M., 2015, ApJ, 804, 103
Kawamata R., Oguri M., Ishigaki M., Shimasaku K., Ouchi M., 2016, ApJ, 819, 114
Kazantzidis S., Kravtsov A. V., Zentner A. R., Allgood B., Nagai D., Moore B., 2004, ApJ, 611, L73
Kazantzidis S., Bullock J. S., Zentner A. R., Kravtsov A. V., Moustakas L. A., 2008, ApJ, 688, 254
Kazantzidis S., Zentner A. R., Kravtsov A. V., Bullock J. S., Debattista V. P., 2009, ApJ, 700, 1896
Keeton C. R., Madau P., 2001, ApJ, 549, L25
Kelly P. L. et al., 2015, Science, 347, 1123
Kelly P. L. et al., 2016, ApJ, 819, L8
Killedar M., Borgani S., Meneghetti M., Dolag K., Fabjan D., Tornatore L., 2012, MNRAS, 427, 533
Kneib J.-P., Natarajan P., 2011, A&AR, 19, 47
Kneib J.-P., Ellis R. S., Smail I., Couch W. J., Sharples R. M., 1996, ApJ, 471, 643
Kneib J.-P., Ellis R. S., Santos M. R., Richard J., 2004, ApJ, 607, 697
Kochanek C. S., 1991, ApJ, 382, 58
Kriek M., van Dokkum P. G., Franx M., Illingworth G. D., Magee D. K., 2009, ApJ, 705, L71
Krist J. E., Hook R. N., Stoehr F., 2011, in Kahan M. A., ed., Proc. SPIE Conf. Ser. Vol. 8127, Optical Modeling and Performance Predictions V. SPIE, Bellingham, p. 81270J
Lam D., Broadhurst T., Diego J. M., Lim J., Coe D., Ford H. C., Zheng W., 2014, ApJ, 797, 98
Liesenborgs J., De Rijcke S., 2012, MNRAS, 425, 1772
Liesenborgs J., De Rijcke S., Dejonghe H., 2006, MNRAS, 367, 1209
Liesenborgs J., de Rijcke S., Dejonghe H., Bekaert P., 2008, MNRAS, 389, 415
Liesenborgs J., de Rijcke S., Dejonghe H., Bekaert P., 2009, MNRAS, 397, 341
Limousin M. et al., 2007, ApJ, 668, 643
Limousin M. et al., 2008, A&A, 489, 23
Limousin M. et al., 2012, A&A, 544, A71
Limousin M. et al., 2016, A&A, 588, A99
Livermore R. C., Finkelstein S. L., Lotz J. M., 2017, ApJ, 835, 113
Lotz J. M. et al., 2017, ApJ, 837, 97
Lynds R., Petrosian V., 1989, ApJ, 336, 1
Ma C.-J., Ebeling H., Marshall P., Schrabback T., 2010, MNRAS, 406, 121
Massey R. et al., 2015, MNRAS, 449, 3393
Medezinski E. et al., 2013, ApJ, 777, 43
Melchior P., Meneghetti M., Bartelmann M., 2007, A&A, 463, 1215
Meneghetti M., Bolzonella M., Bartelmann M., Moscardini L., Tormen G., 2000, MNRAS, 314, 338
Meneghetti M., Bartelmann M., Moscardini L., 2003a, MNRAS, 340, 105
Meneghetti M., Bartelmann M., Moscardini L., 2003b, MNRAS, 346, 67
Meneghetti M., Bartelmann M., Jenkins A., Frenk C., 2007, MNRAS, 381, 171
Meneghetti M., Argazzi R., Pace F., Moscardini L., Dolag K., Bartelmann M., Li G., Oguri M., 2007, A&A, 461, 25
Meneghetti M. et al., 2008, A&A, 482, 403
Meneghetti M., Rasia E., Merten J., Bellagamba F., Ettori S., Mazzotta P., Dolag K., Marri S., 2010a, A&A, 514, A93
Meneghetti M., Fedeli C., Pace F., Gottlöber S., Yepes G., 2010b, A&A, 519, A90
Meneghetti M., Bartelmann M., Dahle H., Limousin M., 2013, Space Sci. Rev., 177, 31
Merten J. et al., 2015, ApJ, 806, 4
Mohammed I., Liesenborgs J., Saha P., Williams L. L. R., 2014, MNRAS, 439, 2651
Mohammed I., Saha P., Liesenborgs J., 2015, PASJ, 67, 21
Mohammed I., Saha P., Williams L. L. R., Liesenborgs J., Sebesta K., 2016, MNRAS, 459, 1698
Natarajan P., Kneib J.-P., 1997, MNRAS, 287, 833
Natarajan P., De Lucia G., Springel V., 2007, MNRAS, 376, 180
Natarajan P., Kneib J.-P., Smail I., Treu T., Ellis R., Moran S., Limousin M., Czoske O., 2009, ApJ, 693, 970
Natarajan P. et al., 2017, MNRAS, 468, 1962
Navarro J., Frenk C., White S., 1996, ApJ, 462, 563
Newman A. B., Treu T., Ellis R. S., Sand D. J., 2011, ApJ, 728, L39
Newman A. B., Treu T., Ellis R. S., Sand D. J., Nipoti C., Richard J., Jullo E., 2013, ApJ, 765, 24
Newman A. B., Ellis R. S., Treu T., 2015, ApJ, 814, 26
Nordin J. et al., 2014, MNRAS, 440, 2742
Oguri M., 2010, PASJ, 62, 1017

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020

Oguri M., Bayliss M. B., Dahle H., Sharon K., Gladders M. D., Natarajan P., Hennawi J. F., Koester B. P., 2012, MNRAS, 420, 3213
Oguri M. et al., 2013, MNRAS, 429, 482
Planelles S., Borgani S., Fabjan D., Killedar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, MNRAS, 438, 195
Postman M. et al., 2012, ApJS, 199, 25
Puchwein E., Bartelmann M., Dolag K., Meneghetti M., 2005, A&A, 442, 405
Rasia E. et al., 2012, New J. Phys., 14, 055018
Redlich M., Bartelmann M., Waizmann J.-C., Fedeli C., 2012, A&A, 547, A66
Richard J., Stark D. P., Ellis R. S., George M. R., Egami E., Kneib J.-P., Smith G. P., 2008, ApJ, 685, 705
Richard J. et al., 2010a, MNRAS, 402, 44
Richard J. et al., 2010b, MNRAS, 404, 325
Richard J., Jones T., Ellis R., Stark D. P., Livermore R., Swinbank M., 2011, MNRAS, 413, 643
Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, ApJ, 802, L19
Rodney S. A. et al., 2015, ApJ, 811, 70
Rodney S. A. et al., 2016, ApJ, 820, 50
Saha P., 2000, AJ, 120, 1654
Sand D., Treu T., Ellis R., 2002, ApJ, 574, L129
Sand D. J., Treu T., Ellis R. S., Smith G. P., 2005, ApJ, 627, 32
Schaller M. et al., 2015, MNRAS, 451, 1247
Schneider P., Ehlers J., Falco E. E., 1992, Gravitational Lenses. Springer-Verlag, Heidelberg
Sebesta K., Williams L. L. R., Mohammed I., Saha P., Liesenborgs J., 2016, MNRAS, 461, 2126
Seitz S., Saglia R. P., Bender R., Hopp U., Belloni P., Ziegler B., 1998, MNRAS, 298, 945
Sendra I., Diego J. M., Broadhurst T., Lazkoz R., 2014, MNRAS, 437, 2642
Sérsic J. L., 1963, Bol. Asociacion Argentina Astron. La Plata Argentina, 6, 41
Sharon K., Gladders M. D., Rigby J. R., Wuyts E., Koester B. P., Bayliss M. B., Barrientos L. F., 2012, ApJ, 746, 161
Shu C., Zhou B., Bartelmann M., Comerford J. M., Huang J., Mellier Y., 2008, ApJ, 685, 70
Smith G. P., Kneib J.-P., Smail I., Mazzotta P., Ebeling H., Czoske O., 2005, MNRAS, 359, 417
Smith G. P. et al., 2009, ApJ, 707, L163
Soucail G., Fort B., Mellier Y., Picat J. P., 1987, A&A, 172, L14
Springel V., 2005, MNRAS, 364, 1105
Springel V., White S., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137
Stark D. P., Ellis R. S., Richard J., Kneib J.-P., Smith G. P., Santos M. R., 2007, ApJ, 663, 10
Tormen G., Bouchet F. R., White S. D. M., 1997, MNRAS, 286, 865
Torri E., Meneghetti M., Bartelmann M., Moscardini L., Rasia E., Tormen G., 2004, MNRAS, 349, 476
Treu T., Koopmans L. V. E., 2002, MNRAS, 337, L6
Treu T. et al., 2015, ApJ, 812, 114
Treu T. et al., 2016, ApJ, 817, 60
Umetsu K., 2013, ApJ, 769, 13
Umetsu K. et al., 2014, ApJ, 795, 163
van der Wel A., 2008, ApJ, 675, L13
Vanzella E. et al., 2012, MNRAS, p. L474
Vanzella E. et al., 2014, ApJ, 783, L12
Vanzella E. et al., 2015, A&A, 576, A116
Vanzella E. et al., 2016, ApJ, 825, 41
Wang L., Li C., Kauffmann G., De Lucia G., 2006, MNRAS, 371, 537
Wang X. et al., 2015, ApJ, 811, 29
Xie L., Guo Q., Cooper A. P., Frenk C. S., Li R., Gao L., 2015, MNRAS, 447, 636
Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2009, ApJ, 707, 354
Zitrin A., Broadhurst T., 2009, ApJ, 703, L132
Zitrin A. et al., 2009, MNRAS, 396, 1985
Zitrin A., Broadhurst T., Bartelmann M., Rephaeli Y., Oguri M., Benítez N., Hao J., Umetsu K., 2012, MNRAS, 423, 2308
Zitrin A. et al., 2013, ApJ, 762, L30
Zitrin A., Redlich M., Broadhurst T., 2014, ApJ, 789, 51
Zitrin A. et al., 2015, ApJ, 801, 44

[1]*Osservatorio Astronomico di Bologna, Via Gobetti 93/3, 40129, Bologna, Italy*
[2]*INFN, Sezione di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italy*
[3]*Department of Astronomy, Yale University, New Haven, CT 06511, USA*
[4]*Space Telescope Science Institute, Baltimore, MD 21218, USA*
[5]*Purple Mountain Observatory, the Partner Group of MPI für Astronomie, 2 West Beijing Road, Nanjing 210008, China*
[6]*INAF – Astronomical Observatory of Trieste, via G.B. Tiepolo 11, I-34143 Trieste, Italy*
[7]*Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, F-13388 Marseille, France*
[8]*Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, I-34131 Trieste, Italy*
[9]*INFN–National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy*
[10]*Department of Physics, University of California, One Shields Ave, Davis, CA 95616, USA*
[11]*IFCA, Instituto de Fisica de Cantabria (UC-CSIC), Av. de Los Castros s/n, E-39005 Santander, Spain*
[12]*Department of Physics, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[13]*Institute for Cosmic Ray Research, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8582, Japan*
[14]*Department of Astronomy, University of Michigan, 1085 South University Avenue, Ann Arbor, MI 48109, USA*
[15]*Department of Astronomy, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[16]*Department of Physics, The University of Hong Kong, 0000-0002-6536-5575, Pokfulam Road, Hong Kong*
[17]*Expertisecentrum voor Digitale Media, Universiteit Hasselt, Wetenschapspark 2, B-3590 Diepenbeek, Belgium*
[18]*Research Center for the Early Universe, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[19]*Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), University of Tokyo, Chiba 277-8582, Japan*
[20]*School of Physics and Astronomy, University of Minnesota, 116 Church Street SE, Minneapolis, MN 55455, USA*
[21]*Cahill Center for Astronomy and Astrophysics, California Institute of Technology, MC 249-17, Pasadena, CA 91125, USA*

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article-abstract/472/3/3177/4083623 by INAF Trieste (Osservatorio Astronomico di Trieste) user on 16 July 2020