Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **508,** 5602–5611 (2021) https://doi.org/10.1093/mnras/stab2858
Advance Access publication 2021 October 06

A new framework for understanding systematic errors in cluster lens modelling – II. Constraint selection

Dhruv T. Zimmerman, Charles R. Keeton ★ and Catie A. Raney

Department of Physics and Astronomy, Rutgers, The State University of New Jersey, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA

Accepted 2021 September 24. in original form 2021 August 31

ABSTRACT

Cluster lens models are affected by a variety of choices in the lens modelling process. We have begun a programme to develop a systematic error budget for cluster lens modelling. Here, we examine the selection of image constraints as a potential systematic effect. For constraining the mass model, we find that it is more important to have images be spatially distributed around the cluster than to have them distributed in redshift. We also find that some image sets appear to be more important than others in terms of how well they constrain the models; the 'important' image sets typically include an image that lies close to a lensing critical curve as well as an image that is relatively isolated from other images (providing constraints in a region that would otherwise lack lensing information). These conclusions can help guide observing programmes that seek follow-up data for candidate lensed images.

Key words: gravitational lensing: strong – galaxies: clusters: individual: Abell 2744 – galaxies: clusters: individual: MACS J0416.1-2403.

1 INTRODUCTION

Gravitational lensing is a general relativistic effect that results in the bending of the path of light by mass. As the most massive bound structures in the universe at $\sim 10^{14}$–10^{15} $M_{\odot}$, galaxy clusters function as very strong lenses, especially near their dense centres. Since lensing can magnify the background sources, cluster lenses can be used as cosmic telescopes to examine high-redshift galaxies (e.g. Bouwens et al. 2014; Salmon et al. 2018, 2020).

However, before we can fully study the imaged galaxies, we need to know the details of exactly how they are lensed. This means we need to construct models of the distribution of mass within the lensing galaxy clusters (see the review by Kneib & Natarajan 2011). The Hubble Frontier Fields (HFFs) project (Lotz et al. 2017) provided extensive observational data for six lensing clusters and invited multiple teams to produce lens models for the cluster mass distributions and resulting magnification maps.

The lens modelling process is complicated and requires a variety of choices and assumptions. For example, teams must identify galaxies in the cluster and decide how to treat them in models. They must also decide how to handle the dark matter that constitutes the majority of mass in the cluster. They also have to determine whether and how to account for effects from mass in the foreground or background of the main cluster. Each choice has a different systematic effect on the modelling (e.g. D'Aloisio, Natarajan & Shapiro 2014; Harvey, Kneib & Jauzac 2016; Acebron et al. 2017; Chirivì et al. 2018; Raney, Keeton & Brennan 2020a).

Another set of choices relates to the constraints on lens models. Since clusters are large on the sky, there may be a significant amount of lensing information available, but determining robust constraints can be tricky. The HFF project benefitted from deep observations with the *Hubble Space Telescope* and supporting observations from many other facilities, which allowed many potential sets of lensed images to be identified. Given these image candidates, the lens modelling teams came together and assigned ratings: images assigned a 'gold' rating had spectroscopic redshifts and good confidence that they were a result of lensing; images with a 'silver' rating had confidence that they were the results of lensing but lacked spectroscopic redshifts; and images with a 'bronze' rating lacked a clear consensus among the teams as to whether they were the results of lensing. Some images were identified later and thus lacked a rating. All of the HFF lens modelling teams had access to this information and were able to choose which image sets to use in their analysis.

As the quantity and quality of data have improved, statistical uncertainties in cluster lens models have decreased and systematic effects have become a more dominant source of error. The HFF lens modelling programme[1] was explicitly structured to explore the systematic effects by having different teams make their own choices, assumptions, and approximations. Raney et al. (2020b) scrutinized results from the most recent round of HFF lens modelling (version 4) and found that the differences between teams' magnifications exceeded their stated statistical errors.

Therefore, it is clear that systematic errors in lens modelling are now very relevant. However, it is not clear which of the many choices involved in lens modelling contribute most to the systematic effects. Our team has undertaken to examine different choices and quantify their relative importance. Raney et al. (2020a) previously examined systematic effects related to the treatment of galaxies along the line

★ E-mail: keeton@physics.rutgers.edu

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

© 2021 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/508/4/5602/6382152 by Rutgers University Libraries user on 29 August 2022

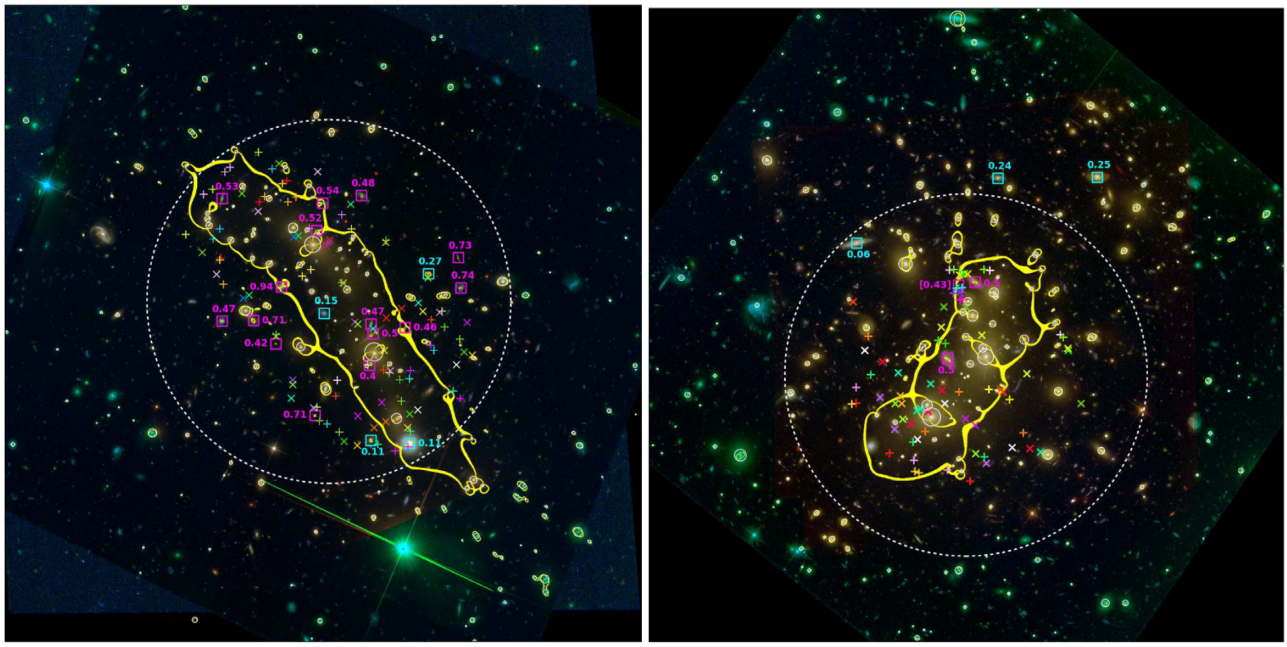

Figure 1. Sky plots of MACS 0416 (left-hand panel) and Abell 2744 (right-hand panel) along with accompanying information for lens modelling, from Raney et al. (2020a). Each plot is 3.5 arcmin on a side. The crosses and plus signs represent lensed images, with images from the same source identified by the same shape and colour. Circled objects represent cluster member galaxies. Cyan and magenta boxed objects represent selected LOS galaxies incorporated in the models; cyan galaxies are in the foreground of the cluster, and magenta galaxies are in the background. The yellow lines are areas of high magnification for the fiducial 2D lens model at source redshift $z = 9$.

of sight (LOS). In a paper accompanying this one, Raney, Keeton & Zimmerman (2021) discuss systematic effects related to the selection and treatment of cluster member galaxies.

In this paper, we study how the selection of image constraints affects lens models. Johnson & Sharon (2016) performed a similar analysis with simulated clusters created by Meneghetti et al. (2017). They randomly sampled the image constraints (with and without spectroscopic redshift information) and found that the systematic error was small with $\gtrsim$10 image sets and $\gtrsim$5 spectroscopically confirmed redshifts. Ghosh, Williams & Liesenborgs (2020) explored how large increases in the number of image constraints can affect models in two simulated clusters and found increased accuracy in the mass distributions. Since Johnson & Sharon used *random* subsamples of images and Ghosh et al. used *randomly generated* images, their analyses did not necessarily reflect *systematic* decisions that teams might have made with the available data. In this study, we examine real rather than simulated systems, and instead of using random selections, we make specific choices about which image sets to include or exclude.

We begin with an overview of our methods in Section 2, detail and analyse results in Section 3, and offer conclusions in Section 4.

2 METHODOLOGY

2.1 Clusters and data

The analysis described in this paper uses two of the HFF clusters. MACS J0416.1−2403 is drawn from the Massive Cluster Survey (MACS) catalogue and is at a redshift of $z = 0.396$ (Ebeling, Edge & Henry 2001). Our fiducial modelling uses 95 images of 35 sources in the gold sample; we omit two image sets classified as gold: One has an image that is blended with a galaxy, and the other is a triplet around a galaxy, so it mainly constrains that galaxy. Abell 2744 is drawn from the Abell catalogue and is at a redshift of $z = 0.308$ (Abell, Corwin & Olowin 1989). Our fiducial modelling uses 71 images of 24 sources in the gold sample; we omit three image sets classified as gold because it is not clear how to match counter-images. Full details about the image constraints and other model inputs can be found in Raney et al. (2020a).

2.2 Lens modelling framework

Although gravitational lensing is naturally a three-dimensional (3D) effect, for modelling it is usually projected on to a 2D plane in the 'thin lens' approximation. This approximation is valid because the scale of the cluster is much smaller than the cosmological distances over which the light travels. Additionally, this provides some degree of computational simplicity and efficiency.

Raney et al. (2020a) present the original version of our lens modelling framework, and Raney et al. (2021) describe updates designed to help us examine systematic effects. Briefly, we use parametric models for the dark matter distribution, MACS 0416 has three dark matter haloes to accommodate complexities within the main mass distribution, and Abell 2744 has two dark matter haloes in the main mass distribution and a third halo representing a structure to the north-west of the main field. We include cluster member galaxies and assign mass using mass–luminosity scaling relations normalized by the brightest cluster galaxy. We also include galaxies along the LOS, but make an approximation in which they are projected on to the main lens plane (i.e. we use the '2-D' models discussed by Raney et al. 2020a); the LOS galaxies are calibrated using redshift-dependent scaling relations. The galaxy masses include uncertainty from scatter in the scaling relations. The sky plots in Fig. 1 indicate the information that goes into our models.

The mass model is used to compute the lens potential, ϕ, which then enters the lens equation relating the position $\boldsymbol{x}$ of an image in

Downloaded from https://academic.oup.com/mnras/article/508/4/5602/6382152 by Rutgers University Libraries user on 29 August 2022

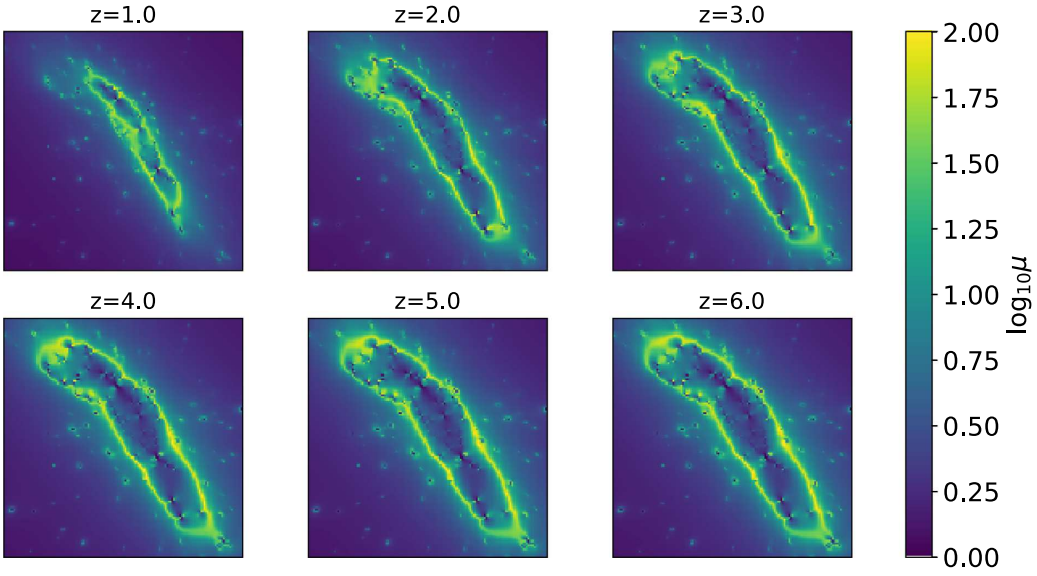

Figure 2. An example of magnification maps for MACS 0416 computed at different source redshifts. The bright yellow lines represent the critical curves, which are the areas of high magnification. This illustrates the dependence of the critical curves on the source redshift.

the lens with the position $\boldsymbol{y}$ of the source in its source plane:

$$\boldsymbol{y} = \boldsymbol{x} - \frac{D_{\rm ls}}{D_{\rm s}} \nabla\phi(\boldsymbol{x}), \tag{1}$$

Here, $D_{\rm ls}$ and $D_{\rm s}$ represent angular diameter distances between the lens and source and the observer and source, respectively, which are computed from the redshifts of the lens and source. We use a cosmology with $\Omega_{\rm M} = 0.3$ and $\Omega_\Lambda = 0.7$. The lensing magnification is then given by

$$\mu = \frac{1}{\det\left[\mathbf{I} - \mathbf{J}(\phi_{\rm eff})\right]}, \qquad \text{where} \qquad \phi_{\rm eff}(\boldsymbol{x}) = \frac{D_{\rm ls}}{D_{\rm s}}\phi(\boldsymbol{x}), \tag{2}$$

where $\mathbf{I}$ represents the 2×2 identity matrix and $\mathbf{J}(\phi_{\rm eff})$ denotes the Jacobian matrix of the effective lens potential $\phi_{\rm eff}$. The magnification depends on the redshift of the source through the $D_{\rm ls}$ and $D_{\rm s}$ factors, as depicted in Fig. 2.

We use Markov chain Monte Carlo (MCMC) techniques to explore the range of acceptable fits and characterize the statistical uncertainties in the models. We can then change our choice of inputs, run a new set of models, and compare the results to quantify systematic effects. Full details of our methodology are given by Raney et al. (2021).

2.3 Quantifying model differences

Given two sets of MCMC results, we quantify the difference between the models using a generalized two-sample test. Suppose we have samples $\boldsymbol{x}_i$ for $i = 1, ..., n_X$ drawn from distribution X and samples $\boldsymbol{y}_i$ for $i = 1, ..., n_Y$ drawn from distribution Y. For this paper, X and Y represent Bayesian posterior distributions in our model parameter space. The root-mean-square (rms) distance between points in X is

$$D_X^2 = \frac{1}{n_X(n_X - 1)} \sum_{i=1}^{n_X} \sum_{j=1}^{n_X} ||\boldsymbol{x}_i - \boldsymbol{x}_j||_{\mathbf{C}}^2, \tag{3}$$

where $||\boldsymbol{x}_i - \boldsymbol{x}_j||_{\mathbf{C}} = \left[(\boldsymbol{x}_i - \boldsymbol{x}_j)^T \mathbf{C}^{-1} (\boldsymbol{x}_i - \boldsymbol{x}_j)\right]^{1/2}$ is the distance between points computed with metric $\mathbf{C}$. The rms distance between points in Y can be written similarly. The distance between points in different samples can be written as

$$D_{XY} = \frac{1}{n_X n_Y} \sum_{i=1}^{n_X} \sum_{j=1}^{n_Y} ||\boldsymbol{x}_i - \boldsymbol{y}_j||_{\mathbf{C}}^2. \tag{4}$$

We then define the difference between the distributions X and Y as

$$\Delta^2 = D_{XY}^2 - D_X D_Y. \tag{5}$$

Raney et al. (2021) show that this distance metric can be simplified in terms of the means $\boldsymbol{m}_X, \boldsymbol{m}_Y$ and covariance matrices $\mathbf{C}_X, \mathbf{C}_Y$ for the two samples as

$$\Delta^2 = (\boldsymbol{m}_X - \boldsymbol{m}_Y)^T \mathbf{C}^{-1} (\boldsymbol{m}_X - \boldsymbol{m}_Y) + \left[\sqrt{\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_X)} - \sqrt{\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_Y)}\right]^2, \tag{6}$$

where $\mathbf{C}$ denotes the pooled covariance $\mathbf{C} = (\mathbf{C}_X + \mathbf{C}_Y)/2$. The first term measures a shift in the means, and is identical to Hotelling's t^2 statistic. The second term measures a shift in the shape of the distributions.

3 RESULTS

Our overall goal is to understand how constraint selection affects lens modelling. We do not attempt to replicate all of the specific choices made by various modelling teams; rather, we seek to understand in a general way how decisions about image constraints could systematically affect the resulting models. We group images into sets based on their source, choose subsamples for rerunning the models, and compare the results with our fiducial models using the full sample of images. We consider random subsamples (Section 3.1) before examining deliberate cuts based on position (Section 3.2) and redshift (Section 3.3), and we then find it instructive to omit image sets one by one (Section 3.4) as a way to understand which images have the most significant effect on lens models.

3.1 Random subsamples

We first explore how the models change when we use random subsamples of the constraints. This is similar to the methodology

Downloaded from https://academic.oup.com/mnras/article/508/4/5602/6382152 by Rutgers University Libraries user on 29 August 2022

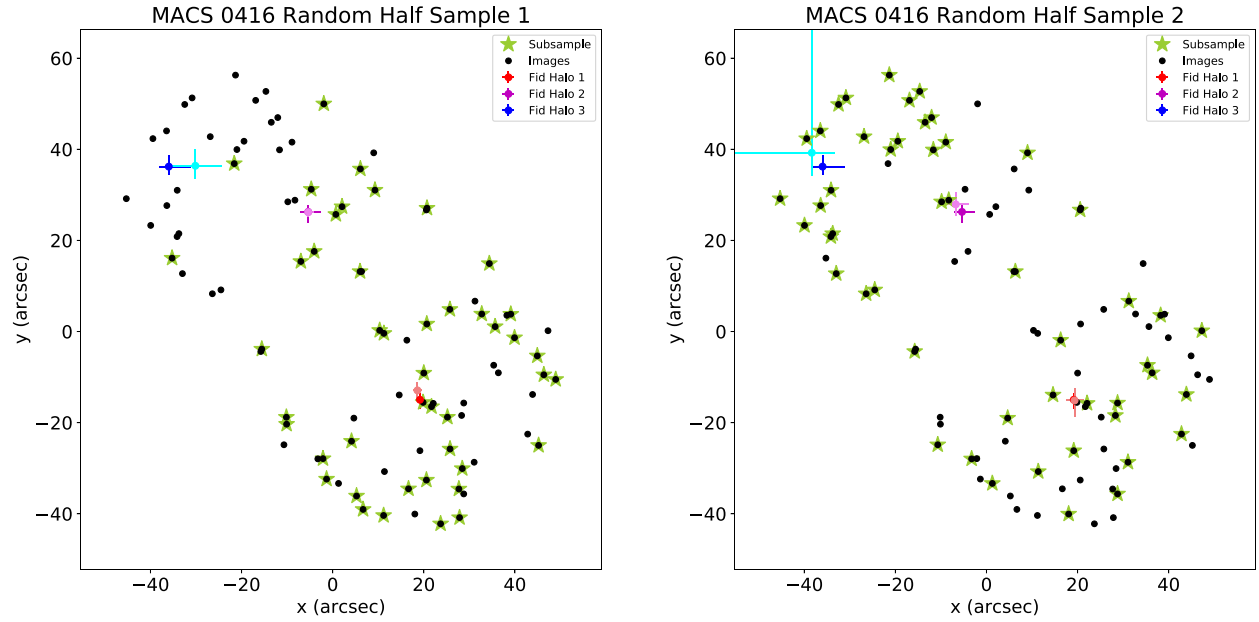

Figure 3. Results for one set of complementary random half samples for MACS 0416. In each sample, roughly half the image sets are included to constrain the model. In this plot, and similar plots to follow, the black dots represent the full sample of images used by our team, while the green stars represent the subsample of images constraints used for the particular analysis. The red, blue, and magenta points represent the locations of the dark matter haloes in the fiducial model, with error bars indicating the 68 per cent confidence region. The corresponding lighter points to each halo (e.g. violet for magenta) represent the locations of the model haloes within the subsample run. The shifts relative to the fiducial model are $\Delta^2 = 21.1$ for the first half and $\Delta^2 = 12.3$ for the second half. (Note that Δ^2 is computed using the full parameter space, not just the halo positions shown here.)

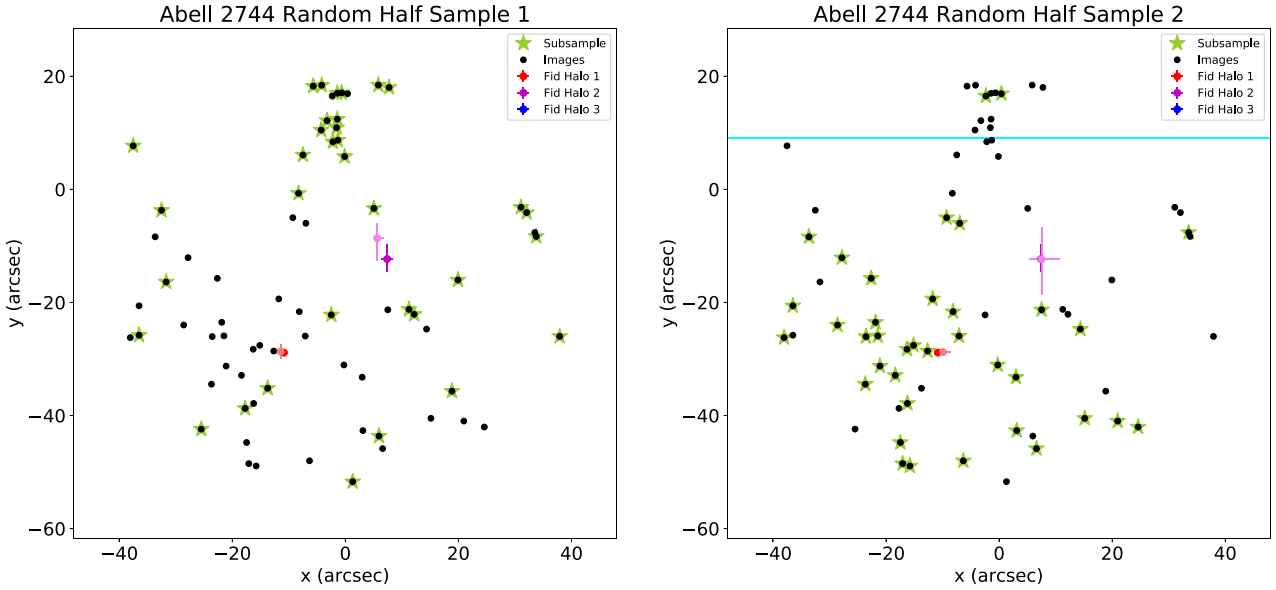

Figure 4. Similar to Fig. 3, but for a complementary pair of random half image samples for Abell 2744. The shifts relative to the fiducial model are $\Delta^2 = 15.0$ for the first half and $\Delta^2 = 14.9$ for the second half. (In the right-hand panel, the cyan line represents a large error bar for the third halo.)

used by Johnson & Sharon (2016). We randomly divide the image sets into two halves, construct lens models using each half, and compute Δ^2 to quantify the shifts relative to the fiducial model.

Fig. 3 shows results for the pair of random subsamples for MACS 0416. We find shifts of $\Delta^2 = 21.1$ for the first half and $\Delta^2 = 12.3$ for the second half. In the first half subsample, the positions and uncertainties for the cluster haloes do not change very much, but the cluster member galaxies tend to be less massive than in the fiducial model. In the second half subsample, several parameters become less well constrained; in particular, the third halo becomes much more uncertain, which is curious because this subsample retains many of the images in the vicinity of the third halo. We will see more about the behaviour of the third halo in later subsamples below.

Fig. 4 shows the corresponding results for the pair of random subsamples for Abell 2744. Here, the shifts are nearly identical at $\Delta^2 = 15.0$ and $\Delta^2 = 14.9$. In the first half subsample, the first halo tends to be less massive than in the fiducial model; that change may arise because many of the images in the vicinity of the first halo have

Downloaded from https://academic.oup.com/mnras/article/508/4/5602/6382152 by Rutgers University Libraries user on 29 August 2022

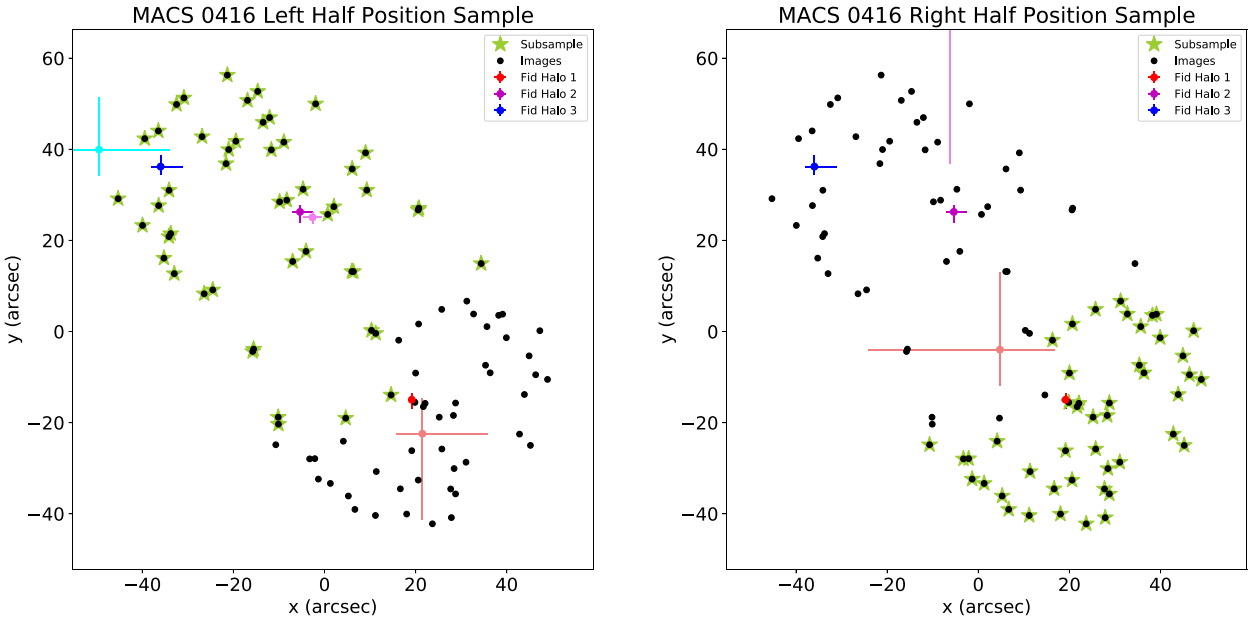

Figure 5. Similar to Fig. 3, but now we divide the images for MACS 0416 into 'left' and 'right' halves based on the typical position among images in a set. The shift relative to the fiducial model is $\Delta^2 = 49.6$ for the left half subsample and $\Delta^2 = 265.1$ for the right half subsample.

been removed. Also, the third halo tends to be more massive in this case. In the second subsample, many parameters have modest shifts, but the most dramatic change is that the third halo becomes much less constrained in both position and mass.

From this analysis, we infer that Δ^2 can reach the level of $\sim$10–20 just from having fewer images providing constraints.

3.2 Positional subsamples

The HFF lens systems have complicated spatial configurations with images spanning more than an arcminute and multiple mass components used to model the cluster dark matter distribution. Therefore, it is interesting to consider how the global arrangement of images affects the modelling. Here, we divide the image sets into 'left' and 'right' halves based on the average position of images in a set.

Fig. 5 shows the results for MACS 0416. The left half position sample has a shift of $\Delta^2 = 49.6$ relative to the fiducial model. The uncertainties for halo #1 increase, which is perhaps not surprising because most of the images around it have been removed. Halo #3 also shifts and becomes less constrained, even though the images around it are still part of the subsample; this halo appears to be compensating for the changes in halo #1. Indeed, there is a general decrease in the mass of the first halo and a corresponding increase in the mass parameters for the second and third haloes. The right half position subsample has a significantly larger shift of $\Delta^2 = 265.1$. Visually, it is clear that this image subsample has a limited spatial extent. With the loss of spatial information, haloes #2 and #3 move away from the fiducial model and become largely unconstrained in position. Interestingly, the mass normalization parameter for halo #2 is consistent with the fiducial model, but for halo #3 it appears unconstrained. Even though the subsample retains images in the vicinity of halo #1, this component also becomes much less constrained.

Fig. 6 shows the corresponding results for Abell 2744. For the left half position subsample, the shift relative to the fiducial model is $\Delta^2 = 18.5$. The second halo becomes less well constrained because it lies among the images that have been removed. The third halo, which is off the plot to the upper right-hand side, becomes almost completely unconstrained. For the right half subsample, the shift is $\Delta^2 = 65.6$. In this case, the first halo loses most of the information around it to constrain it and shifts noticeably. The second halo becomes less constrained, although its changes are fairly modest. Interestingly, the third halo's properties do not appear to exhibit much change with the right half position subsample, perhaps because this subsample contains the images that are closer to this mass component.

The deliberate spatial cuts clearly have much more significant effects on the models than the random cuts.

3.3 Redshift subsamples

Since lensing depends on redshift (cf. Fig. 2), it is interesting to consider redshift-selected subsamples. Fig. 7 shows results for the lower and upper half redshift subsamples in MACS 0416. The Δ^2 values are smaller than we saw for the positional cuts, and more similar to the random cuts: $\Delta^2 = 9.4$ and 24.4 for the lower and upper half redshift subsamples, respectively. A key difference is that the redshift cuts retain a reasonable spread of images throughout the cluster, so the shifts in halo positions are relatively modest. The shifts are somewhat more apparent in the upper half redshift subsample, and in examining other details of this model, we see other changes involving the second halo. We will revisit these results in Section 3.4 as the subsamples considered there offer helpful insights.

Fig. 8 shows the corresponding results for Abell 2744. Here, the shifts are $\Delta^2 = 23.6$ and 17.0 for the lower and upper half redshift samples, respectively. In the lower half redshift sample, the second halo moves as several of the nearby images constraining it are removed from the sample. In this system, many of the images in the left half position sample are also in the lower half redshift sample, which may explain why we see similar higher Δ^2 values for those redshift samples in this cluster.

Downloaded from https://academic.oup.com/mnras/article/508/4/5602/6382152 by Rutgers University Libraries user on 29 August 2022

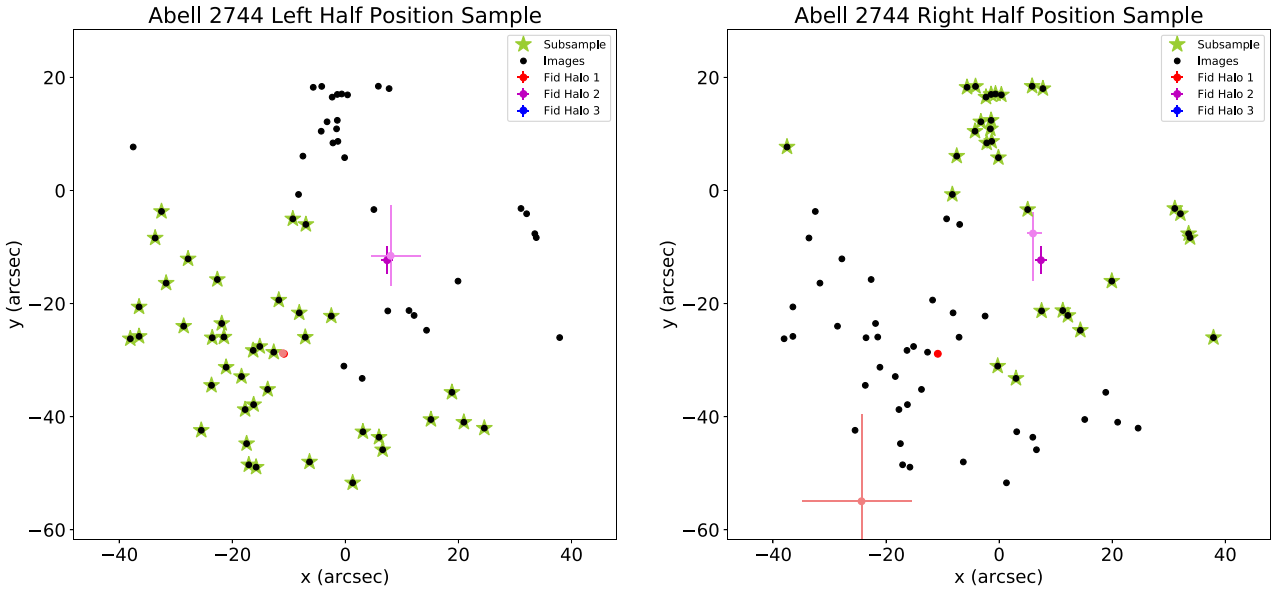

Figure 6. Similar to Fig. 5, but showing the two position subsamples for Abell 2744. Here, the shifts relative to the fiducial model are $\Delta^2 = 18.5$ and 65.6 for the left and right half position subsamples, respectively.

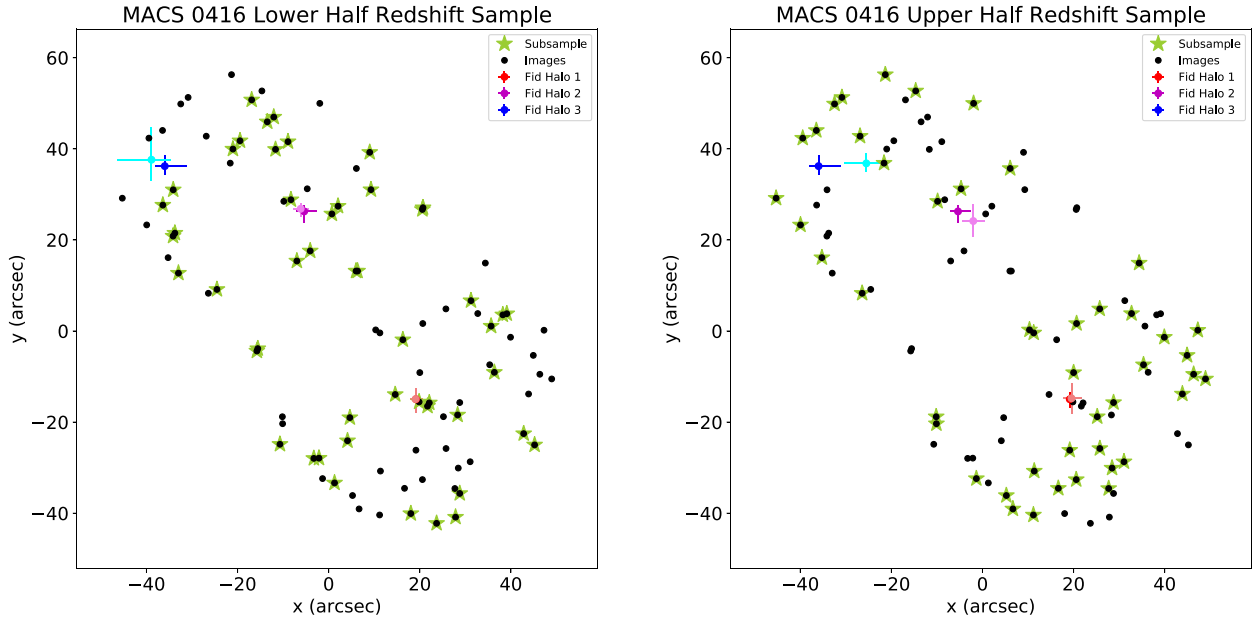

Figure 7. Similar to Fig. 5, but now comparing the lower and upper half redshift samples for MACS 0416. The shifts are $\Delta^2 = 9.4$ and 24.4 for the lower and upper half redshift subsamples, respectively.

Comparing results from the positional and redshift subsamples, we conclude that it is more important to have images be spatially distributed around the cluster than to have them distributed in redshift, at least for constraining the mass model.

3.4 Leave-one-out (LOO) subsamples

We now dig in and try to understand in detail why removing some image sets has more impact than others. We adopt a LOO approach in which we remove sets of images one by one (with replacement) and perform modelling with each resulting subsample. (This can also be viewed as jackknife resampling.) In this analysis, Δ^2 quantifies the shift that arises from not having a given image set in the sample, so it can be considered to indicate the importance of that image set for the overall model. Across both clusters, the majority of individual Δ^2 values are $\lesssim$1.5, most are below 3, and all are below 10. Fig. 9 shows the four image sets with the highest Δ^2 values for MACS 0416, with the lensing critical curves (computed at the appropriate source redshift) shown for comparison. Qualitatively, these image sets all appear to have a close pair of images near the critical curve.

To quantify the results, we compare Δ^2 to various features of the image set and look for interesting relations. Fig. 10 contains a summary of these comparisons. We first consider the redshift of the source (panel a). There is no clear relation, with low and high Δ^2 values distributed across the redshift range. One interesting feature of the sample for MACS 0416 is that many of the most important

Downloaded from https://academic.oup.com/mnras/article/508/4/5602/6382152 by Rutgers University Libraries user on 29 August 2022

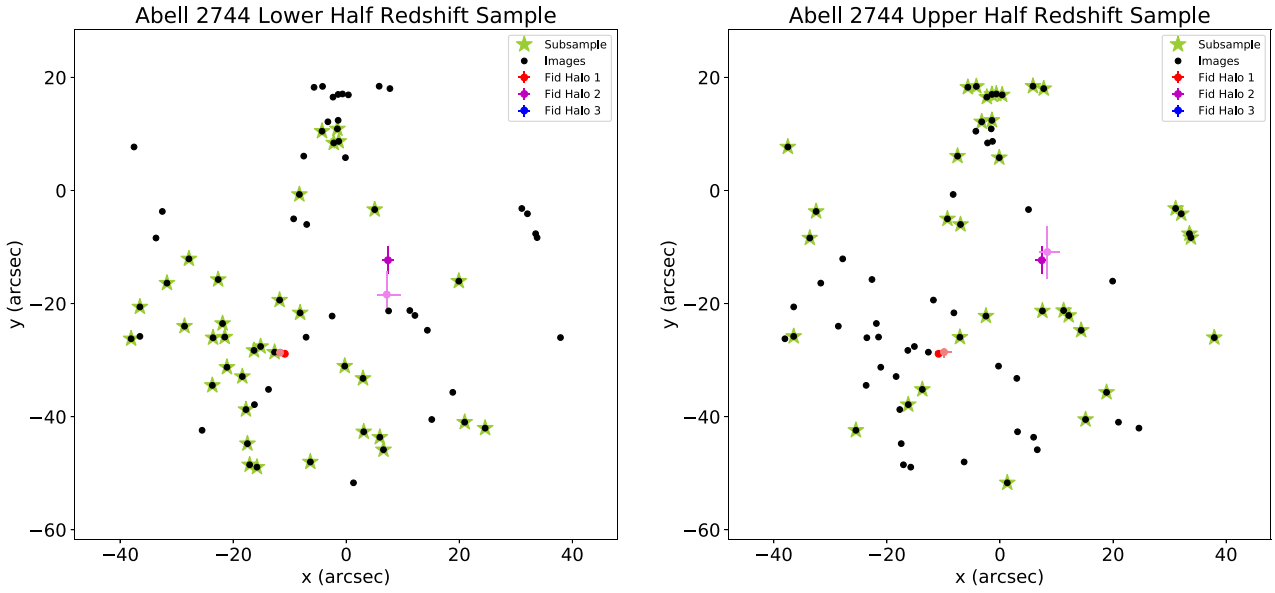

Figure 8. Similar to Fig. 7, but now showing the redshift subsamples for Abell 2744. The shifts are $\Delta^2 = 23.6$ and 17.0 for the lower and upper half redshift subsamples, respectively.

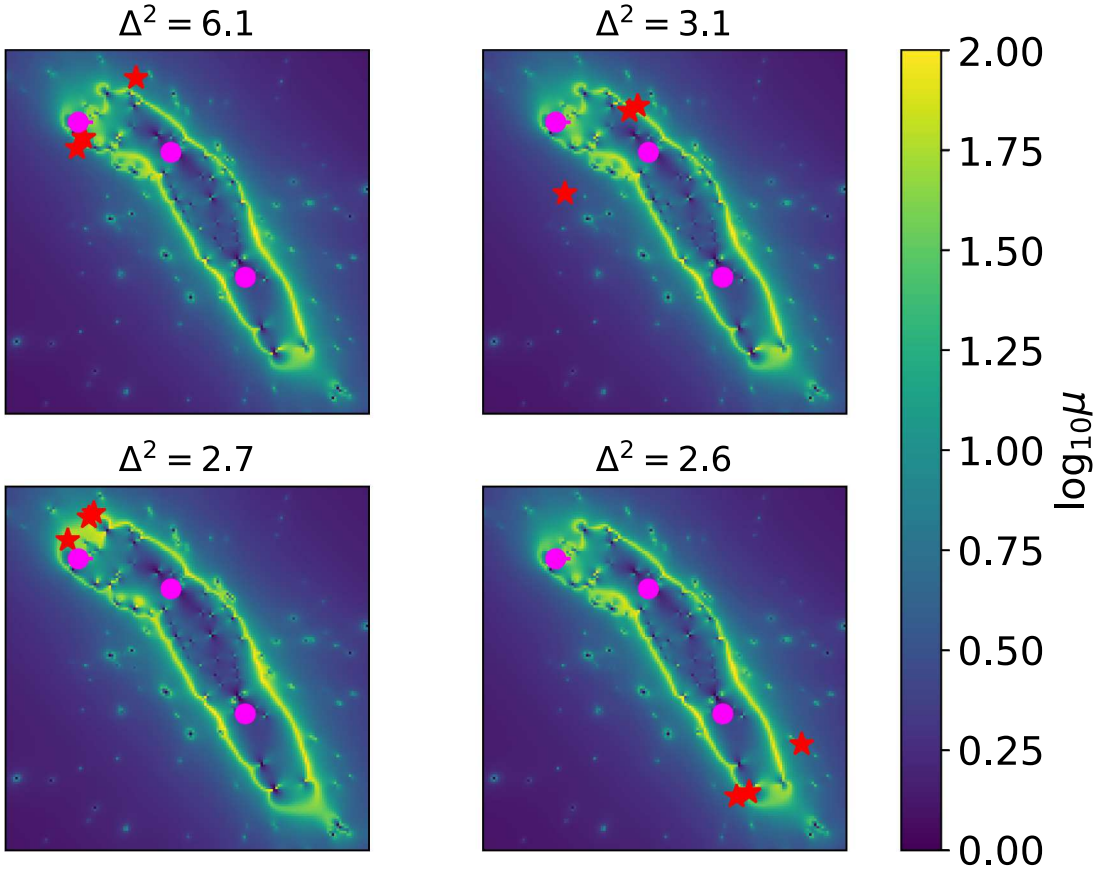

Figure 9. Illustration of the four image sets in MACS 0416 that yield the largest LOO Δ^2 values. The magnification maps are computed at the redshift of the source. Red stars mark the locations of the images in each set, while purple dots represent the placement of the different dark matter haloes. We note several interesting similarities in these image sets: At least two of the images are close together, and all the sets have at least one image that appears to be very close to a critical curve.

Downloaded from https://academic.oup.com/mnras/article/508/4/5602/6382152 by Rutgers University Libraries user on 29 August 2022

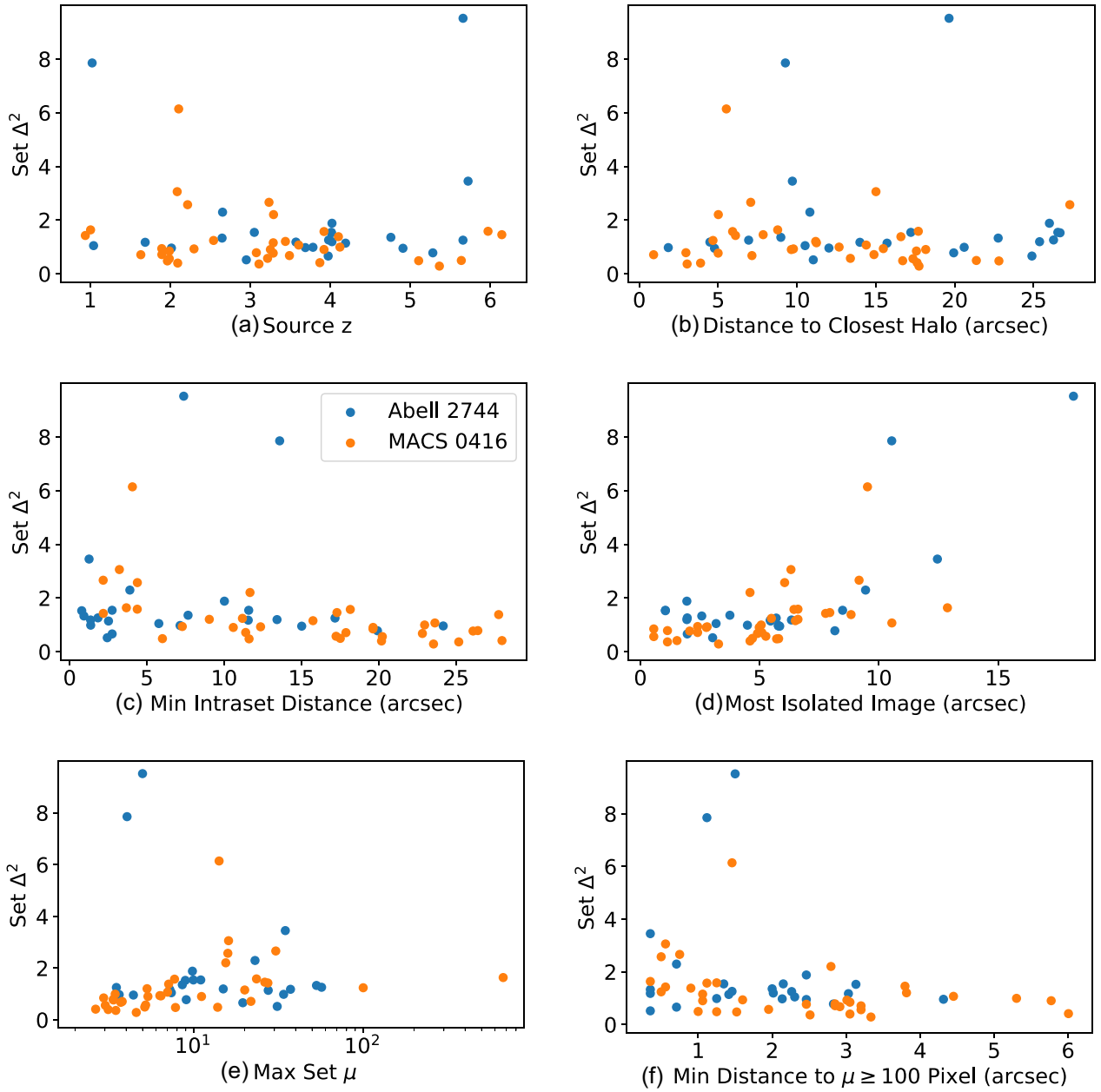

Figure 10. Illustration of how Δ^2 for each image set depends on properties of the set: (a) redshift of the source; (b) minimum distance between the image set and one of the model halo positions; (c) minimum distance between images within a set; (d) isolation criterion, defined to be the distance of the most isolated image in the set to the nearest image in any other set; (e) maximum magnification of the images in the set; and (f) minimum distance to high magnification pixels ($\mu \geq 100$), as a proxy for the distance to the critical curve. In each panel, results are shown for both Abell 2744 (blue) and MACS 0416 (orange). See text for a discussion of the results.

images seem to be in the lower half of the redshift sample. This may explain why our redshift subsample for MACS 0416 yielded a noticeably large Δ^2 value when we removed all of the lower redshift sources (q.v. Fig. 7).

Next, we consider the smallest distance of an image set to one of the haloes in the fiducial model. Fig. 10(b) shows that there is no apparent relation between this distance and Δ^2.

Based on what we saw in Fig. 9, we examine the distance between the closest pair of images within the set (Fig. 10(c)). Although we do not necessarily see a correlation between these two variables, we do note that for MACS 0416 most of the high-Δ^2 image sets appear to have a pair of images that are relatively close together. For Abell 2744, it generally seems that some of the mid-range Δ^2 values are associated with close image pairs, but we note two sets of images with high Δ^2 values that do not follow this trend. It is certainly fair to say that widely separated image sets have low values of Δ^2.

We also consider the degree to which images are isolated from other images. Specifically, we compute the distance for each image in a set to the nearest image outside the set; we take the maximum of these distances as a measure of the degree of isolation for the *most isolated image* in the set. Fig. 10(d) shows that Δ^2 tends to increase as this isolation criterion increases, suggesting that having an isolated image contributes to making an image set important. This makes sense because an isolated image provides information about

Downloaded from https://academic.oup.com/mnras/article/508/4/5602/6382152 by Rutgers University Libraries user on 29 August 2022

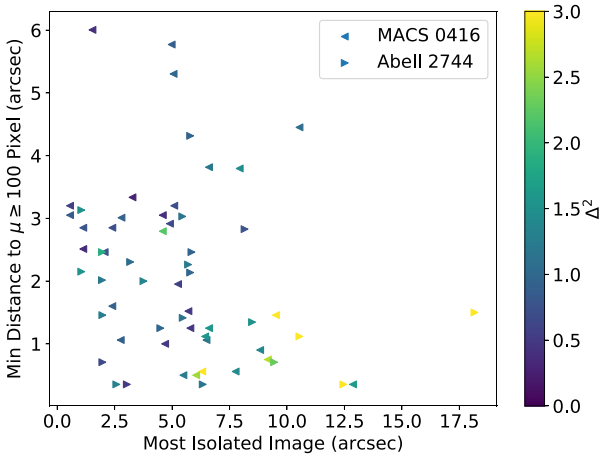

Figure 11. This figure shows how the Δ^2 value for each image set (indicated by colour) depends jointly on the set's isolation criterion and distance to the critical curve. The two clusters are distinguished by different symbols. In general, the highest Δ^2 values are located in the lower right-hand region of the plot, corresponding to image sets that are isolated and near a critical curve. See text for more details.

a part of the sky in which we otherwise would not have information to constrain the behaviour of the lens.

Last but not least, we consider proximity to the lensing critical curve. We compute the magnifications of the images (adjusted for the appropriate redshift) using our fiducial model, and plot Δ^2 versus the maximum magnification in the image set in Fig. 10(e). In general, image sets with $\Delta^2 \gtrsim 2$ have reasonably high magnifications, $\mu \gtrsim 10$. Interestingly, however, the most important images for both clusters do not appear to have especially high magnifications. Upon visual inspection, we find that the two image sets in Abell 2744 with $\Delta^2 > 7$ appear to have an image near the critical curve but its magnification is suppressed by the presence of a galaxy nearby.[2] Therefore, as a complementary measure, we compute the geometric distance to the critical curve; for computational simplicity, we actually compute the median distance from an image to the nearest pixel with $\mu > 100$ in our magnification maps (again adjusted for redshift). Fig. 10(f) shows that the most important image sets have an image within about 2 arcsec of a critical curve, and even apart from the three extreme values of Δ^2, there is a general trend that the more important image sets have smaller distances from a critical curve.

One might wonder whether image sets that contain more images tend to be more important. We find that most of the important image sets have three images, yet there are plenty of image sets containing three images that are not important. In Abell 2744, there are also four sets with more than three images (two sets have four and the others have five images), but none appears to be as important to the overall model as some of the other sets.

While Fig. 10 has revealed some characteristics that help indicate whether image sets are important, there does not seem to be a *single* characteristic that dominates. It is therefore worth considering whether *combinations* of these characteristics offer deeper understanding. Based on the discussion above, and on examining various scatter plots, we find that the most instructive analysis compares Δ^2 to an image set's isolation criterion and distance to the critical curve, as shown in Fig. 11. Overall, the higher Δ^2 values appear to congregate in the lower right region of this plot. While there is not a perfect gradient, it seems as though most of the important image sets in both clusters are in that area, while virtually none of the darker, less important points lies in the same area. We infer that the most important image sets for lens modelling are those that contain both isolated images and images near critical curves.

[2]Inside the critical curve, a mass concentration can create a region of reduced magnification (see Schechter & Wambsganss 2002).

4 CONCLUSIONS

The first major conclusion we draw is that reducing the number of image sets leaves models less well constrained. While that is not at all surprising, it is instructive to see the details. We have found, for example, that it is more important to have images well distributed on the sky than to have images drawn from a wide range in redshift. It is worth noting that we have retained the full complexity of our mass models while reducing the number of constraints. A complementary approach, which is more typical of early stages of lens modelling when few constraints are available, would be to use less complicated lens models. Simple models must be used with care because they miss much of the deeper complexity of real clusters. (Even our 'complex' models are no doubt simpler than reality.)

We have found that certain image sets are more important than others for constraining lens models. While the details differ between clusters, in general, the most important image sets contain both an image that is fairly isolated, which provides constraints in a region that would otherwise lack lensing information, and an image that is close to the critical curve. This conclusion suggests that observing programmes seeking follow-up data for lensed images should prioritize candidates in regions of the cluster that are not already mapped, as well as candidates suspected of being near a critical curve; ideally, colour and/or redshift information could be used to identify image systems that include *both* features. We recognize that the criterion of being near a critical curve is not directly observable. However, having a close pair of image candidates can be used as a proxy. Also, if lens modelling and follow-up observations are done iteratively (as in the HFF programme), early models can provide estimates of the critical curves (although care must be taken to avoid circular reasoning).

The Δ^2 metric allows us to make a direct, quantitative comparison of the constraint selection studied here and the selection/treatment of member galaxies studied by Raney et al. (2021). Broadly speaking, constraint selection effects are stronger in MACS 0416 than in Abell 2744, while the opposite is true of galaxy selection effects. The difference presumably arises because MACS 0416 has the larger number of image constraints, so the models are well constrained when all of the images are included and can degrade significantly when images are excluded. Omitting scatter in the galaxy luminosity–mass scaling relations is on the same scale as cutting the image sample in half randomly or by redshift. (As we saw in Section 3.2, cutting the image sample in half by position can have fairly dramatic effects.)

There are some natural extensions of this work. It would of course be interesting to examine the other four clusters in the HFF programme. Additionally, while this analysis primarily used 'gold' images, some of the modelling teams drew more heavily from the 'silver' and 'bronze' image candidates that lack spectroscopic redshifts and/or are less secure. It would be interesting to explore how using these image candidates affects the models. While it seems obvious that having more image constraints is better, our work suggests ways to prioritize image candidates that could be most helpful in improving constraints on cluster lens models.

Downloaded from https://academic.oup.com/mnras/article/508/4/5602/6382152 by Rutgers University Libraries user on 29 August 2022

ACKNOWLEDGEMENTS

We thank everyone associated with the HFF programme for making it such a valuable resource for research on cluster lensing. We acknowledge support from the US National Science Foundation through grant number AST-1909217.

DATA AVAILABILITY

The software and data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Abell G. O., Corwin Harold G. J., Olowin R. P., 1989, ApJS, 70, 1
Acebron A., Jullo E., Limousin M., Tilquin A., Giocoli C., Jauzac M., Mahler G., Richard J., 2017, MNRAS, 470, 1809
Bouwens R. J. et al., 2014, ApJ, 795, 126
Chirivì G., Suyu S. H., Grillo C., Halkola A., Balestra I., Caminha G. B., Mercurio A., Rosati P., 2018, A&A, 614, A8
D'Aloisio A., Natarajan P., Shapiro P. R., 2014, MNRAS, 445, 3581
Ebeling H., Edge A. C., Henry J. P., 2001, ApJ, 553, 668
Ghosh A., Williams L. L. R., Liesenborgs J., 2020, MNRAS, 494, 3998
Harvey D., Kneib J. P., Jauzac M., 2016, MNRAS, 458, 660
Johnson T. L., Sharon K., 2016, ApJ, 832, 82
Kneib J.-P., Natarajan P., 2011, A&AR, 19, 47
Lotz J. M. et al., 2017, ApJ, 837, 97
Meneghetti M. et al., 2017, MNRAS, 472, 3177
Raney C. A., Keeton C. R., Brennan S., 2020a, MNRAS, 492, 503
Raney C. A., Keeton C. R., Brennan S., Fan H., 2020b, MNRAS, 494, 4771
Raney C. A., Keeton C. R., Zimmerman D. T., 2021, MNRAS, in press, https://doi.org/10.1093/mnras/stab2857
Salmon B. et al., 2018, ApJ, 864, L22
Salmon B. et al., 2020, ApJ, 889, 189
Schechter P. L., Wambsganss J., 2002, ApJ, 580, 685

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

Downloaded from https://academic.oup.com/mnras/article/508/4/5602/6382152 by Rutgers University Libraries user on 29 August 2022