






Galaxy mass modelling from multiwavelength JWST strong lens analysis: dark matter substructure, angular mass complexity, or both?

Samuel C. Lange ¹★, Aristeidis Amvrosiadis,¹ James W. Nightingale,^{1,2,3} Qiuhan He ¹,
 Carlos S. Frenk,¹ Andrew Robertson ⁴, Shaun Cole,¹ Richard Massey ^{1,2}, Xiaoyue Cao ^{5,6},
 Ran Li^{5,6} and Kaihao Wang^{5,6}

¹*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

²*Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

³*School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK*

⁴*Observatories, Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA*

⁵*School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China*

⁶*National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Chaoyang District, Beijing 100012, China*

Accepted 2025 March 18. Received 2025 March 2; in original form 2024 October 16

ABSTRACT

We analyse two galaxy-scale strong gravitational lenses, SPT0418-47 and SPT2147-50, using JWST NIRC*am* imaging across multiple filters. To account for angular complexity in the lens mass distribution, we introduce multipole perturbations with orders $m = 1, 3, 4$. Our results show strong evidence for angular mass complexity in SPT2147, with multipole strengths of 0.3–1.7 per cent for $m = 3, 4$ and 2.4–9.5 per cent for $m = 1$, while SPT0418 shows no such preference. We also test lens models that include a dark matter substructure, finding a strong preference for a substructure in SPT2147-50 with a Bayes factor of $\Delta \log(\epsilon) \sim 60$ when multipoles are not included. Including multipoles reduces the Bayes factor to $\Delta \log(\epsilon) \sim 11$, still corresponding to a 5σ detection of a subhalo with an NFW (Navarro–Frenk–White) mass of $\log_{10}(M_{200}/M_{\odot}) = 10.87^{+0.53}_{-0.71}$. While SPT2147-50 may represent the fourth detection of a dark matter substructure in a strong lens, further analysis is needed to confirm that the signal is not due to systematics associated with the lens mass model.

Key words: gravitational lensing: strong – galaxies: structure – dark matter – cosmology: observations.

1 INTRODUCTION

The nature of the dark matter (DM) has long been one of the most fundamental questions in cosmology. Cold dark matter (CDM) is considered the most probable explanation for this ‘missing matter’ (Davis et al. 1985; see Frenk & White 2012 for a review). This proposal, together with the assumption of a non-zero Cosmological Constant to explain the accelerated expansion of the Universe, makes up the standard Λ CDM cosmology (e.g. Aghanim et al. 2020).

Structure formation in a CDM Universe has been intensely studied in cosmological simulations (e.g. Frenk et al. 1988; Springel et al. 2005). There are still alternatives to CDM that have not been excluded, such as self-interacting dark matter (SIDM; reviewed in e.g. Tulin & Yu 2018) and warm dark matter (WDM; e.g. Bode, Ostriker & Turok 2001; Lovell et al. 2014). The formation of low-mass structures provides a key test of CDM and these alternatives; specifically, the CDM paradigm predicts a large population of low-mass structures and of substructures of main haloes, that form in a ‘bottom-up’ hierarchical merging process (e.g. Frenk et al. 1985; Diemand et al. 2008; Springel et al. 2008). By contrast, in WDM the primordial spectrum of density perturbations has a cut-off on small

scales that gives rise in a cut-off in the mass function of haloes at a mass corresponding to dwarf galaxies, $\log_{10}(M/M_{\odot}) \sim 8.5$, for the most popular values of the WDM particle mass. This mass function in both CDM and WDM models has been extensively studied – see e.g. Bode et al. (2001), Gao et al. (2004), Macciò & Fontanot (2010), Lovell et al. (2014), or Zavala & Frenk (2019) for a review. There are several methods for investigating the nature of DM. These include measuring the ‘clumpiness’ of the Ly α forest (Viel et al. 2005, 2013), the abundance of dwarf galaxies, such as the satellites of the Milky Way (e.g. Enzi et al. 2021; Newton et al. 2021, 2024).

Another method is to search directly for low-mass dark haloes using gravitational imaging of strong lenses (e.g. Vegetti & Koopmans 2009; Vegetti et al. 2024). In strong lensing, light rays from a background source are deflected by a massive foreground object, forming multiple images of the source. In the regime of galaxy–galaxy strong lensing, the background source is extended and light rays from different regions of the source are traced through different regions of the lensing galaxy allowing the structure in this lensing galaxy to be studied at high resolution. If a dark substructure happens to lie along one of the lines traced by the source light rays, it will cause a perturbation of the image in an asymmetric manner. All of this can be modelled to create a picture of the lens and source galaxy that best fits the observed image (usually in a Bayesian framework).

* E-mail: samuel.c.lange@durham.ac.uk

There have been three detections of DM structures thus far using this type of gravitational imaging (Vegetti et al. 2010, 2012; Hezaveh et al. 2016), as well as a detection of an ultramassive black hole (Nightingale et al. 2023b). The inferred halo masses from the three detections depend upon the profile of the DM halo assumed in the modelling, but correspond to Navarro–Frenk–White (NFW)-halo masses (Navarro, Frenk & White 1997) in the region 10^9 – $10^{10}M_{\odot}$.¹ Non-detections of DM haloes/subhaloes (where one would be expected) using gravitational imaging have been used to constrain the subhalo mass function (SHMF) and the possible nature of sterile neutrinos as WDM (e.g. Ritondale et al. 2019). This analysis requires a ‘Sensitivity Map’ (e.g. Despali et al. 2022) that indicates where in an image any substructures would be detectable. DM substructures may also be at any redshift along the line of sight (rather than at the redshift of the main lens) and can have concentrations that vary from the standard mass–concentration relationship (e.g. Li et al. 2016; Despali et al. 2018; Amorisco et al. 2021; Minor et al. 2021a; He et al. 2022; O’Riordan et al. 2023).

Understanding the complexity in the primary lens galaxy is crucial for analysing DM substructures, since overly simplistic mass models can result in false-positive ‘detections’ of substructure. This issue arises due to a degeneracy between the mass complexity of the main deflector and the presence (or absence) of subhaloes (e.g. He et al. 2023; O’Riordan & Vegetti 2024). Multipole perturbations (Chu et al. 2013) to the standard power-law mass profile (Tessore & Metcalf 2015) help address some of this complexity. This degeneracy is also encountered when using flux ratio anomalies (brightness differences in multiple images of a point source) to infer substructure effects and constrain DM mass models (Cohen et al. 2024; Gilman et al. 2024).

The necessity for complex mass modelling has gained further support from recent studies with high-quality lensing data from instruments like ALMA (Powell et al. 2022; Stacey et al. 2024), which recommend the inclusion of third and fourth-order multipoles in lens models. Amvrosiadis et al. (2024) examined the stellar distribution in nearby massive elliptical galaxies, finding evidence for third- and fourth-order multipoles. They also emphasized that in some cases, a first-order multipole (indicating lopsidedness in mass distribution) is crucial for accurate modelling, as seen in a recent detailed strong lens analysis survey (Barone et al. 2024). Other studies have highlighted various forms of angular complexity, such as radial ellipticity variations and twists in the mass distribution (Nightingale et al. 2019; He et al. 2024), and showed that lensing mass profiles cannot always be adequately described by simple power-law density profiles (Etherington et al. 2023). Additionally, ‘External Shear’ often used in lens modelling can account for some complexities that may be absent in the central lens mass model (Keeton, Kochanek & Seljak 1997; Witt & Mao 1997; Cao et al. 2022; Etherington et al. 2024).

JWST *NIRCam* imaging provides an unprecedented multiwavelength view of strong lenses (e.g. Bergamini et al. 2023; van Dokkum et al. 2024; Mercier et al. 2024; Rigby et al. 2024). This capability is essential for testing lens models since the source morphology varies across wavelengths (e.g. Amvrosiadis et al. 2025), producing different lensed images, while the mass model and instrumental effects remain wavelength independent. In this work, we perform independent lens model fits across multiple wavelengths for two strong lenses to assess: (i) whether angular mass complexity (in the form of multipoles) is consistently required for the lensing

galaxies across wavelengths, and (ii) whether a DM substructure is consistently favoured in each waveband. This multiwavelength approach allows a detailed exploration of the degeneracy between these two model components.

The paper is structured as follows. In Section 2, we describe the JWST data and the context of the systems studied. In Section 3, we discuss our PYAUTOLENS modelling pipeline and the details of our models. We also introduce the multipole perturbations and discuss how these and real subhaloes are considered in the data. In Section 4, we show our model results for the multipole and subhalo fits, first to SPT2147, and then to SPT0418. In Section 5, we discuss the results addressing the question of whether SPT2147 actually contains multipoles and/or a substructure. We summarize our conclusions in Section 6. Throughout this work we assume the Planck 2015 cosmological parameters (Planck Collaboration XIII 2016) unless otherwise specified. Additionally, we make use of both natural logarithm (denoted simply ‘log’) and log with base 10 (denoted specifically as ‘log₁₀’).

2 THE DATA

We study two strong gravitational lenses: SPT2147-50 and SPT0418-47. We use JWST *NIRCAM* imaging from filters f277W, f356W, and f444W (2.77, 3.56, and 4.44 μm , respectively) for SPT2147 and f356W, f444W for SPT0418 – other (bluer) wavelength images exist, but lack a sufficient source signal-to-noise ratio (SNR) to produce informative models. The lenses were imaged as part of the *TEMPLATES* survey (PI: Rigby, Rigby et al. 2024). ALMA data is also available for these lenses and the source galaxy of SPT2147 has been studied in Amvrosiadis et al. (2025), with ALMA substructure analysis of both lenses in ALMA to be released in a subsequent work.

JWST-*NIRCam* imaging is shown in Fig. 1 for SPT2147 (upper) and SPT0418 (lower), where we have saturated all pixels above 3.0 MJySr^{-1} in order to see the lensed source galaxy light, as it is strongly outshone by the lensing galaxy. Fig. 2 shows SNR maps for only the source emission, based on the model image created by our best-fitting models (see Section 4). We have normalized the colours to the same values across both lenses and all filters, to visualize the large difference between the maximum pixel SNR of the two lenses, with SPT2147 ranging from $2\times$ to $5\times$ higher SNR, depending on the filter. Specific values for maximal and average (over all pixels with $\text{SNR} > 3$) SNRs are quoted in Table 1.

2.1 SPT 2147–50

This system is composed of a lensing galaxy at redshift $z = 0.845$ and the background source galaxy at $z = 3.76$ (Weiß et al. 2013; Reuter et al. 2020). In conjunction with this work, this lens was also studied using ALMA data taken in band 8 by Amvrosiadis et al. (2025), where we discover that the background source galaxy is a spiral-barred galaxy. There is also a line-of-sight object south-west of centre, on the lower arc, marked with the white box in Fig. 1. This object is only really visible in f277W, and not visible at all in redder filters. Modelling the spectral energy distribution (SED) of this object (only visible in the *NIRCam* bands) using the BAGPIPES.² SED fitting code (Carnall et al. 2018) gives a much lower redshift than our lens galaxy, in the range $z = 0.04$ to $z = 0.10$, and best-fitting mass of the order $\log_{10}(M/M_{\odot}) \sim 6$.

¹Detections from Vegetti et al. (2010, 2012) use tidally truncated pseudo-jaffe profiles, with quoted subhalo masses $\sim 3.5 \times 10^9 M_{\odot}$ and $\sim 1.9 \times 10^8 M_{\odot}$

²<https://github.com/ACCarnall/bagpipes>

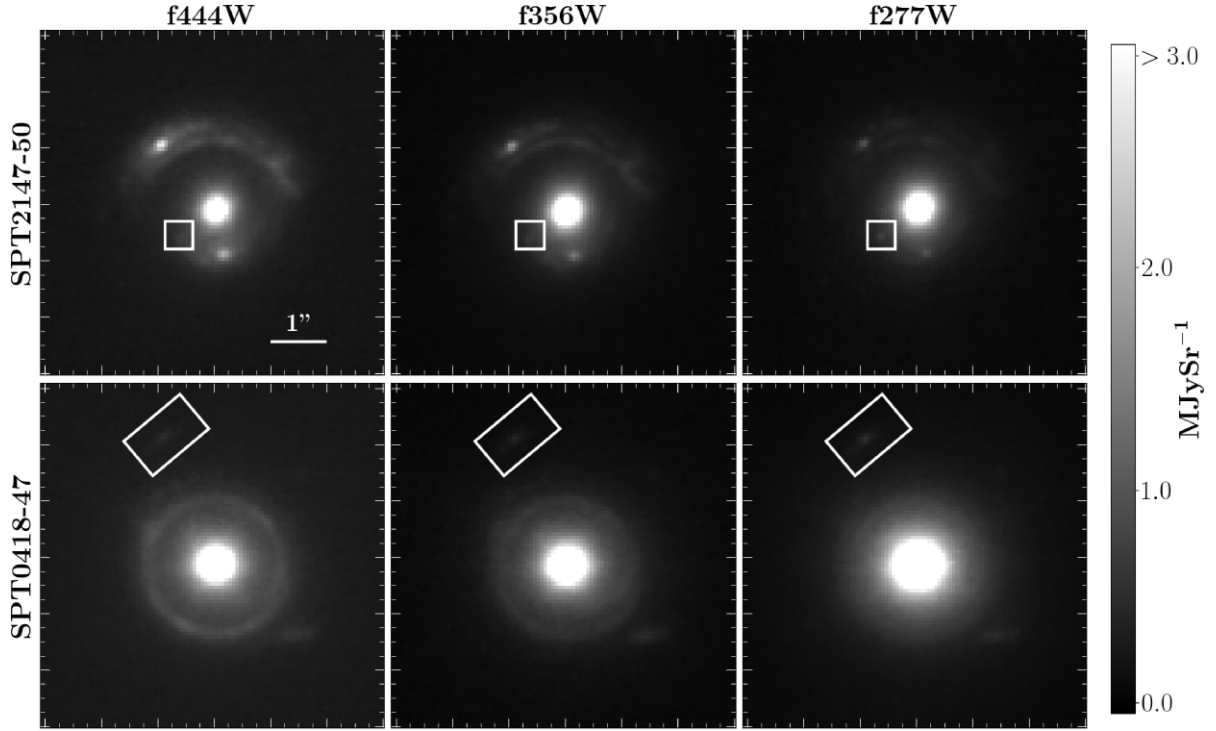


Figure 1. Data from the three reddest NIRCcam filters (f277W, f356W, f444W) for SPT2147-50 (upper panels) and SPT0418-47 (lower panels). All panels are $6'' \times 6''$ squares. The signal is saturated above 3.0 MJySr^{-1} in order to see the lensed source emission clearly. Marked with white boxes are line of sight galaxies in each lens (note the LOS object in SPT2147 is brightest in f277W).

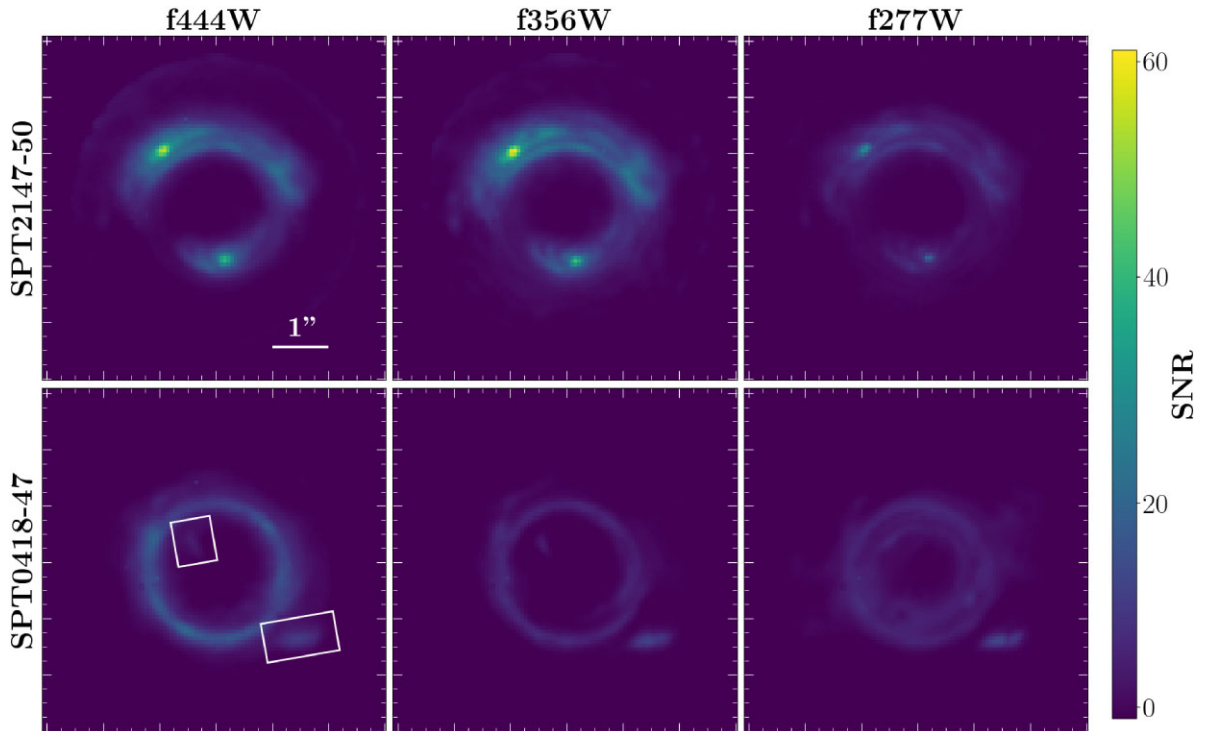


Figure 2. SNRs for SPT2147-50 (upper panels) and SPT0418-47 (lower panels), based on the model lensed source emission from our best-fitting models. All panels are $6 \text{ arcsec} \times 6 \text{ arcsec}$ squares. All panels have colour normalization to the same level ($[0, 60]$), but exact maximal and average values are shown in Table 1. The white boxes in SPT0418 (f444W) mark the image and counter-image of the source companion galaxy as this lens has a multiobject source.

Table 1. Signal information for each filter of our two strong lenses. The source parameters are based on the lens-light subtracted model images and therefore approximations based on model results. The avg. source SNR was calculated as the mean of all pixels with SNR > 3. The RMS noise was calculated over all pixels.

| | SPT0418 | | | SPT2147 | | |
|---|---------|-------|-------|---------|-------|-------|
| | f444 | f356 | f277 | f444 | f356 | f277 |
| Max. Source SNR pix^{-1} | 20.90 | 12.57 | 14.68 | 53.80 | 59.59 | 28.79 |
| Avg. Source SNR pix^{-1} | 7.67 | 5.20 | 4.65 | 9.65 | 9.61 | 5.93 |
| RMS Noise $\text{MJySr}^{-1}\text{pix}^{-1}$ | 1.606 | 1.654 | 1.455 | 1.599 | 0.930 | 1.259 |

2.2 SPT 0418–47

This system is composed of a lensing galaxy at redshift $z = 0.263$ and a background source at redshift $z = 4.22$ (Weiß et al. 2013; Reuter et al. 2020). From our modelling and the literature, the source here is composed of two merging galaxies, with Cathey et al. (2024) reporting a projected separation of ~ 4.42 kpc and a *stellar* mass ratio of 4:1. Interestingly, the source companion galaxy is not visible in *HST* or ALMA and so has only been discovered through JWST NIRCam, MIRI, and NIRSpec data (Peng et al. 2023; Cathey et al. 2024). This companion galaxy is visible in the south-east of the image, outside the Einstein ring, with its counter-image visible north-west of the centre only once the lens light has been removed. Both are marked with the white box in the f444 panel of Fig. 2. There is also a nearby or line-of-sight galaxy, marked by the white box in Fig. 1, which we mask out of our analysis as it is too far from the main system to have any appreciable gravitational effect.

3 METHODS

To perform all the modelling described in this paper we use the open source lens modelling package PYAUTOLENS³ (Nightingale & Dye 2015; Nightingale, Dye & Massey 2018; Nightingale et al. 2021b) and its associated parent packages, including PYAUTOFIT⁴ an extension which provides a statistical fitting and analysis framework to our lens modelling (Nightingale, Hayes & Griffiths 2021a) and PYAUTOGALAXY⁵ a package which includes implementations of all galaxy mass and light profiles which may be needed for modelling (Nightingale et al. 2023a).

The lens modelling procedure is similar to previous works using this software (e.g. Etherington et al. 2022) and is as follows, with parametrizations as described in Table 2, and model choices detailed in the subsections below.

- (i) **Initialize** – create a basic model for all light and lens mass using parametric profiles.
- (ii) **Source pixelization** – reconstruct the source galaxy onto a pixel-based grid in order to capture higher levels of complexity.
- (iii) **Lens refinement** – re-model the lens light based on this new pixelized source, and fit some more complex lens mass model.
- (iv) **Optional: angular mass complexity** – fit some perturbation to increase the angular mass complexity of the model (see Section 3.2).

- (v) **Subhalo search** – introduce a subhalo to the lens mass model and investigate (see Section 3.3).

For all modelling stages, we use the non-linear sampler NAUTILUS⁶ (Lange 2023), which uses Importance Nested Sampling to build a Bayesian Posterior estimate. We also remind the reader that ‘log’ refers to the natural log, and e.g. ‘log₁₀’ refers to the base-10 logarithm.

3.1 Standard mass and light profiles

All light and mass profile quantities are computed with regards to their elliptical components’

$$\begin{aligned} \epsilon_1 &= \frac{1-q}{1+q} \sin 2\phi, \\ \epsilon_2 &= \frac{1-q}{1+q} \cos 2\phi, \end{aligned} \quad (1)$$

where q is the axis ratio and ϕ the anticlockwise position angle, defined as $\phi = 0$ when aligned to the positive x -axis.

3.1.1 Lens mass: elliptical power law

We model the lens mass through the elliptical power-law (EPL) density profile (Tessore & Metcalf 2015), which is commonly used in strong lens modelling (e.g. Vegetti et al. 2012; Nightingale et al. 2024). This profile represents the total mass profile of the lens in the following form

$$\kappa(\xi) = \frac{(3 - \gamma^{\text{lens}})}{1 + q^{\text{lens}}} \left(\frac{\theta_E^{\text{lens}}}{\xi} \right)^{\gamma^{\text{lens}} - 1}, \quad (2)$$

where the superscript ‘lens’ implies the parameters relating to the primary lens mass (i.e. the ‘Macro-model’, not any additional subhalo structures). Here, q is the minor-to-major axis ratio and ξ is the elliptical coordinate defined as $\xi = \sqrt{x^2 + y^2/q^2}$. Additionally, γ^{lens} is the density slope, and θ_E^{lens} is the Einstein radius (in arcsec). $\gamma = 2.0$ gives a singular isothermal ellipsoid (SIE) which is used to initialize the lens mass model.

3.1.2 External shear

An external shear is included in all stages of modelling. Some investigations note that there can be strong degeneracy of the shear with other model parameters (e.g. power-law ellipticity, Johnson et al. 2024) and there is evidence that the shear may account for missing internal mass complexity in the lens (Etherington et al. 2024). We fit the shear as two elliptical components (γ_1, γ_2), and akin to the elliptical components of the main galaxy EPL, we relate the shear components to the magnitude, γ_{ext} and position angle ϕ_{ext} (again, with the angle defined counterclockwise from the positive x -axis):

$$\begin{aligned} \gamma^{\text{ext}} &= \sqrt{\gamma_1^{\text{ext}2} + \gamma_2^{\text{ext}2}}, \\ \tan 2\phi^{\text{ext}} &= \frac{\gamma_2^{\text{ext}}}{\gamma_1^{\text{ext}}}. \end{aligned} \quad (3)$$

³<https://github.com/Jammy2211/PyAutoLens>

⁴<https://github.com/rhayes777/PyAutoFit>

⁵<https://github.com/Jammy2211/PyAutoGalaxy>

⁶<https://github.com/johannesulf/nautilus>

Table 2. Brief overview of the model parametrizations at the various stages of our fitting pipeline.

| Stage: | Initialize | Source reconstruction | Lens light refinement | Lens main mass refinement | Multipole perturbation | Subhalo search |
|---------------------|--|---|--|----------------------------------|---|---|
| Source Light | MGE: 1x30 | Voronoi mesh with natural neighbour interpolation | FIX model: linear solve only | FIX model: linear solve only | FIX model: linear solve only | FIX model: linear solve only |
| Lens Light | MGE: 2x30 + 1x10 (+1x30 for SPT0418) | FIX model: linear solve only | NEW MGE: 2x30 + 1x10 (+1x30 for SPT0418) | FIX model: linear solve only | FIX model: linear solve only | FIX model: linear solve only |
| Lens Mass | Singular isothermal ellipsoid + external shear | SIE + shear (priors from previous) | FIX model: linear solve only | Elliptical power law + new shear | EPL + shear (priors from previous) + multipoles | EPL + shear + Multipoles (priors from previous) + spherical NFW |

3.1.3 Multi-Gaussian Expansion

To initialize the source galaxy light, and to fit the lens galaxy light, we apply a combined set of 2D elliptical Gaussian light profiles, known as a Multi-Gaussian Expansion (Cappellari 2002, hereafter MGE), used instead of the usual elliptical Sérsic profile to attain increased model flexibility and accuracy, as described in He et al. (2024). The intensity of an MGE set is given as

$$I_{\text{set}}(x, y) = \sum_i^N G_i(x, y). \quad (4)$$

with G_i the i th Gaussian profile,

$$G_i(x, y) = I_i \cdot \exp\left(-\frac{R_i^2(x, y)}{2\sigma_i^2}\right), \quad (5)$$

containing I_i as some intensity normalization factor and σ_i as the full-width at half-maximum of the Gaussian profile. $R_i(x, y)$ is then the Gaussian elliptical radius, and given as

$$R_i(x, y) = \sqrt{x'^2 + \left(\frac{y'}{q_i}\right)^2} \\ x' = \cos \phi_i \cdot (x - x_i^c) + \sin \phi_i \cdot (y - y_i^c) \\ y' = \cos \phi_i \cdot (y - y_i^c) - \sin \phi_i \cdot (x - x_i^c), \quad (6)$$

where q_i is the axis ratio, ϕ_i the position angle, and (x_i^c, y_i^c) is the centre of the Gaussian. A ‘set’ of Gaussians is used to refer to n Gaussians which share the same axis ratio and position angle, but have sigma values increasing in fixed \log_{10} intervals. We do not allow any individual Gaussian intensity to be negative, in order to avoid lens-source light degeneracy, solved using a modified⁷ fast non-negative least squares (fnnls) algorithm (Bro & De Jong 1997).

We combine 30 Gaussians for the source light initialization. The number of Gaussians used for the lens light varies between our two lenses; for SPT2147, we use $2 \times$ sets of 30 Gaussians (akin to a bulge and disc component) plus $1 \times$ set of 10 Gaussians to model central point-like emission. For SPT0418, the lens light is more blended with the lensed source emission and so we use $3 \times$ sets of 10 Gaussians (akin to a thin-disc, thick-disc and bulge, or bulge–disc–envelope system), plus $1 \times$ set of 10 Gaussians for central point-like emission. For both lenses, the centres of the sets of 30 Gaussians is fixed to be the same (i.e. only one centre is fit for all), with the point-source set free to have a different centre, and the position angle/axis ratio of Gaussians are the same throughout a set,

⁷The fnnls code we are using is modified from <https://github.com/jvendrow/fnnls>.

but different between sets. The sigma values (controlling the ‘width’ of the Gaussian) are set between a lower limit of one-fifth the pixel scale (here $0.063 \text{ arcsec}/5 = 0.0126 \text{ arcsec}$) and an upper limit of 3 arcsec for the main bulge/disc sets, and 0.126 arcsec (2 times the pixel scale) for the point-like sets.

3.1.4 Source light: Voronoi mesh reconstruction

The full formalism of the source reconstruction described and tested in detail with regards to substructure investigations on HST imaging of strong lenses in Appendix A of He et al. (2024) and in Nightingale et al. (2024) and is based on Nightingale & Dye (2015). In general, for a given set of mass model parameters, once we subtract the lens light, the remaining lensed source emission is ray-traced back to the source plane and reconstructed onto an adaptive mesh ‘pixel grid’. We use the Voronoi mesh grid, with a Natural Neighbour interpolation⁸ (Sibson 1981) and our own adaptive regularization to smooth the reconstruction based on the source luminosity at that point⁹

Equations (7) describe the numerical implementation (without intermediate steps) of the ‘goodness-of-fit’ function which optimizes the balance between the MGE lens light and pixelized source light, allowing both to be fit simultaneously:

$$G = \chi^2 + G_S + G_M, \quad (7a)$$

$$= \frac{1}{2} \left(\sum_{i=1}^{M_i} \left[\frac{\sum_{j=1}^{N_s} f_{ij} S_j + \sum_{k=1}^{N_g} I_k A_{ik} - d_i}{n_i} \right]^2 + G_L + G_M \right), \quad (7b)$$

$$= \frac{1}{2} \left\| \begin{pmatrix} Z & X \\ \sqrt{r} B_S & B_M \end{pmatrix} \cdot \begin{pmatrix} S \\ L \end{pmatrix} - \begin{pmatrix} Y \\ 0 \end{pmatrix} \right\|^2, \quad (7c)$$

Within equation (7a), G_S and G_M are regularization terms for the pixelized source and the MGE light, respectively. Equation 7b expands the chi-squared term in relation to the image–plane to source–plane mapping, where S_j is the j th source–plane pixel flux, I_k is the k th image–plane pixel flux, f_{ij} is the mapping matrix element quantifying the contribution of source pixel j to the flux of image pixel k (which includes psf and image-processing effects). In addition, the term includes A_{ik} , which is the exponential from

⁸More details about the natural neighbour interpolation technique can be found at <https://gwlucastrig.github.io/TinfourDocs/NaturalNeighborTinfourAlgorithm/index.html>.

⁹In PYAUTOLENS these schemes are the VoronoiNNBrightnessImage pixelization and the AdaptiveBrightnessSplit regularization.

the Gaussian profile of equation (5) G_k , for (x_i, y_i) , as well as the data and noise d_i and n_i at the i th pixel. All of this collapses down into equation (7c), where $Z_{ij} = f_{ij}/n_i$, $X_{ij} = A_{ij}/n_i$, r is the regularization strength, B_S and B_M are ‘square root’ matrices of the source and MGE regularizations, respectively, S and L are the source-plane and image-plane fluxes, and Y_i is the division d_i/n_i .

3.2 Angular mass complexity: multipoles

To introduce additional complexity to the ‘macro-model’, we fit an extension to the EPL profile where we include internal multipoles, building on the formalism of Chu et al. (2013). An individual multipole perturbation adds to the EPL to give a functional form of the convergence, in polar coordinates, of

$$\kappa(r, \phi) = \frac{1}{2} \left(\frac{\theta_E^{\text{lens}}}{r} \right)^{\gamma^{\text{lens}}-1} \cdot k_m \cos(m(\phi - \phi_m)), \quad (8)$$

where r is the radial position in arcseconds, ϕ is the angle of interest for the calculation (in degrees), and m is the defined multipole order. When modelling, the Einstein radius, θ_E^{lens} , slope, γ^{lens} and (y, x) centre of the multipole are linked to the values fit to the main EPL, and in addition, we fit for the multipole elliptical components (ϵ_1^{mp} , ϵ_2^{mp}) (meaning we have the six EPL parameters plus an additional two per multipole). The multipole elliptical components relate to the multipole strength, k_m and multipole position angle ϕ_m (defined counterclockwise from the positive x -axis) in a similar way to the EPL elliptical components, through

$$\phi_m = \frac{1}{m} \arctan \frac{\epsilon_2^{\text{mp}}}{\epsilon_1^{\text{mp}}}, \quad k_m = \sqrt{\epsilon_1^{\text{mp}2} + \epsilon_2^{\text{mp}2}}. \quad (9)$$

The deflection angles can also be given in polar co-ordinates (where α is the power-law slope) as

$$\alpha_r(r, \phi) = -\frac{3-\alpha}{m^2-(3-\alpha)} r_E^{(\alpha-1)} r^{(2-\alpha)} k_m \cos[m(\phi - \phi_m)], \quad \alpha_\phi(r, \phi) = \frac{m^2}{m^2-(3-\alpha)} r_E^{(\alpha-1)} r^{(2-\alpha)} k_m \sin[m(\phi - \phi_m)]. \quad (10)$$

Multipole perturbations cause the mass distribution to lose angular symmetry, but maintain rotational symmetry at different orders, represented by m (i.e. $m = 1$ has one line of symmetry, $m = 2$ as two lines of symmetry, etc.) and so in theory, any number of multipole perturbations could be added to the mass model, however, we restrict ourselves to three orders, $m = 1$, $m = 3$, and $m = 4$. For this work, we test both individual perturbations as well as combinations, in which case the perturbations sum together, giving

$$\kappa(r, \phi) = \frac{1}{2} \left(\frac{\theta_E^{\text{lens}}}{r} \right)^{\gamma^{\text{lens}}-1} \cdot \sum_m \left(k_m \cos(m(\phi - \phi_m)) \right). \quad (11)$$

Some physical motivations and examples of the effects of multipole perturbations to lensing systems are shown in Appendix A.

3.3 Subhalo modelling and Bayesian evidence

We use a Spherical Navarro–Frenk–White (NFW) profile (Navarro et al. 1997) to model DM-only subhalo additions, and only model subhaloes assuming the same redshift as the primary lensing galaxy. The density of an NFW with normalization ρ_s and scale radius r_s is

given as

$$\rho_{\text{NFW}} = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2}. \quad (12)$$

We fit only for the (x, y) centre of the subhalo and the M_{200} mass (the enclosed mass at r_{200} , which is the radius at which the average enclosed density is 200 times ρ_{crit} for the Universe) and relate M_{200} to the scale radius through the mean mass–concentration relation given by Ludlow et al. (2016).

Lens models including a dark substructure produce complex multimodal parameter spaces which are challenging for non-linear tools such as NAUTILUS to sample robustly. We therefore split the image into a 5×5 grid (each ‘pixel’ being 1 square arcsecond) and model a subhalo where the positional priors are within that grid cell only (see e.g. Nightingale et al. 2024). Following this, we can both visualize the effect of adding a subhalo at different locations, as well as use the result from the grid cell with the highest Bayesian evidence increase to initialize a new fit for a subhalo where the model has the freedom to place the subhalo anywhere, but the positional priors are Gaussians centred on the best-fitting position of the highest-evidence grid search.

3.3.1 Model discriminators: Bayesian evidence

We use the natural log Bayesian evidence $-\log(\epsilon)$ as the discriminating factor between models. A simplified way to understand the Bayesian evidence is to consider it to be the Likelihood of the model being the truth, but marginalized over the parameters that are being fit, such that say a small prior that contains many bad solutions will have a lower evidence than a larger prior that only contains good solutions. The equation

$$\epsilon = p(\mathbf{X}|\alpha) = \int p(\mathbf{X}|\theta)p(\theta|\alpha)d\theta \quad (13)$$

formalizes this description, where \mathbf{X} is some observed data, θ is a data parameter, α is a prior hyperparameter, such that the evidence is $p(\mathbf{X}|\alpha)$ (probability of data given prior parameters), which is then the integral over all data parameters of the Likelihood ($p(\mathbf{X}|\theta)$, probability of data given a parameter set) multiplied by the prior ($p(\theta|\alpha)$, probability of a certain data parameter set given the prior hyperparameter sets).

To claim a subhalo candidate against a standard EPL mass model, we would normally set an a-priori requirement of $\Delta\log(\epsilon) > 10$ for the subhalo addition, and we set the same standard for the inclusion of multipoles in a model. For reference, $\Delta\log(\epsilon) = 10$ corresponds to roughly a 5σ result,¹⁰ or a frequentist p -value of order 10^{-7} . Similarly, $\Delta\log(\epsilon) = 5$ corresponds to roughly 3σ or a p -value of ~ 0.03 . The use of $\Delta\log(\epsilon)$ can also be considered as the logarithm of the ‘Bayes Factor’ or ‘Odds’, since $\Delta\log(\epsilon) = \log(\epsilon_1) - \log(\epsilon_2) = \log(\epsilon_1/\epsilon_2)$.

Our specific implementation of evidence thresholds in this work is to hold the requirement of $\Delta\log(\epsilon) > 10$ when we add multipoles to the EPL + shear model, and if these multipole-inclusive models pass the threshold then we continue to add a subhalo to the model. If the multipole evidence is not satisfied, we would only apply the subhalo to the base model.¹¹

¹⁰A useful online source for Bayes factors in context can be found here: <https://ned.ipac.caltech.edu/level5/Sept13/Trotta/Trotta4.html>.

¹¹We add subhaloes to some models in SPT0418 despite them not passing the set evidence threshold only to attain a more thorough comparison to SPT2147.

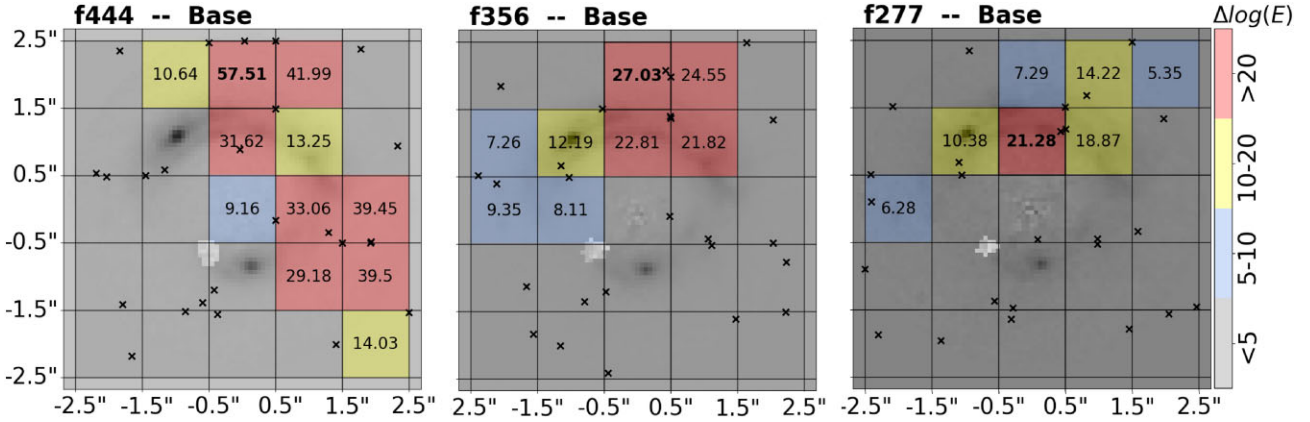


Figure 3. Results from the grid-based preliminary subhalo searches in SPT2147 for the base EPL + shear model. The values and colours correspond to the log-evidence change – $\Delta\log(\epsilon)$ – when a subhalo is modelled in a certain grid cell. Crosses mark the best-fitting position of the subhalo within each grid cell. The highest evidence grid cells are marked in bold. $\Delta\log(\epsilon) = 10$ is considered a roughly 5σ detection.

When using only single-filter imaging, if there are many different combinations of multipoles and multipoles + subhaloes that create an evidence change of $\Delta\log(\epsilon) > 10$ to the base model, it would be difficult to decide between taking just the highest evidence model, or requiring that a subhalo has a further evidence to the multipole (due to unknown degeneracies between multipoles and subhaloes). Since we have multifilter fitting available, we are able to apply the ‘highest evidence model’ approach and cross-reference by applying a $\Delta\log(\epsilon) = 5$ (3σ) confidence region below this model – comparing results which fit into the confidence region across all filters. If more than one model was within this region, we choose the highest evidence model from across all filters as the primary candidate (i.e. the highest evidence model from f444W, since this has the highest source SNR).

4 RESULTS

4.1 SPT2147 results

The SED of the line-of-sight object to the south-west of centre on the lower arc was modelled with BAGPIPES, as mentioned in Section 2.1, to obtain $0.04 \lesssim z \lesssim 0.10$, and a best-fitting mass of the order $\log_{10}(M/M_{\odot}) \sim 6$. We initially try to include this object in our lens modelling; however, the Einstein radius converges to zero in all fits, meaning the object has no appreciable lensing effect in the system. We therefore remove the object from our subsequent fits by artificially increasing the noise level over this area, so that the light at this point provides no contribution to our likelihood functions, due to the now very low SNR.

4.1.1 Base model substructure results

We show in Fig. 3 our results from the subhalo grid-search analysis (see Section 3.3), in each filter, using the EPL + shear as the base macromodel. This ‘base’ model shows multiple high-evidence ($\Delta\log(\epsilon) > 10$) grid squares repeated in similar locations in each filter. The subhalo’s location with the highest evidence is north of the extended arc, with maximal evidence increases of $\sim 57, 27, 21$ in the grid-search for f444, f356, and f277 (note that these values adjust to 60.35, 42.13, 20.54 for the final DM subhalo fit after the grid search, where the priors on its centre are not confined to a 1×1 arcsec square).

The f444 filter also gives a secondary set of high-evidence grid cells, east of the lensing galaxy, which are not seen in the bluer filters. The reason these right-hand cells do not show high evidences in the f356 and f277 filters may be related to the much lower source SNR in f277, and potentially the source light morphology in f356. We have seen from simulations that genuine subhalo signals can also produce residual signal in incorrect locations (Nightingale et al. 2024, section B2), so this could just be a residual effect of a genuine subhalo located towards the top of the image. There may also be a case where the substructure signals in f444 are not due to a genuine substructure, but rather a product of missing complexity from the EPL + shear in that area, where the model uses the inclusion of substructure to compensate for this – the combination of any slight model differences with source morphology changes between filters could then also explain why these are not seen in f356/f277.

4.1.2 Results of multipole perturbations

Fig. 4 shows the log-evidence increases – $\Delta\log(\epsilon)$ – when we include various multipole perturbations to our EPL + shear model, instead of the subhalo additions from the previous section. We see that all multipole inclusions are preferred (give a positive evidence change), and all apart from M1-only and M3-only give $\Delta\log(\epsilon) > 10$ in all three filters. From here forwards therefore, we discount the M1-only and M3-only fits unless otherwise stated, similarly discounting the f277 fits due to the low SNR (and therefore low sensitivity to model changes).

From Fig. 4, we also see that the M4 (in f356) and M3 + 4 (in f444 and f356) fits, show a lower evidence than the other multipole additions, which is a potential indication that, whilst an M1 multipole addition on its own cannot produce much difference to the base model, it may be required in order for the higher order M3 and M4 perturbations to become fully effective – a possible sign again of multipoles having the potential to compensate for missing complexity which they are not primarily intended to capture. A similar effect is shown in Amvrosiadis et al. (2024), where M1 addition changes the fit values of M3 and M4 components to galaxy light distributions.

Fig. 5 shows the recovered parameters from the high-evidence multipole models to SPT2147, as well as the base EPL + shear model. For each model, the parameters are consistent within the

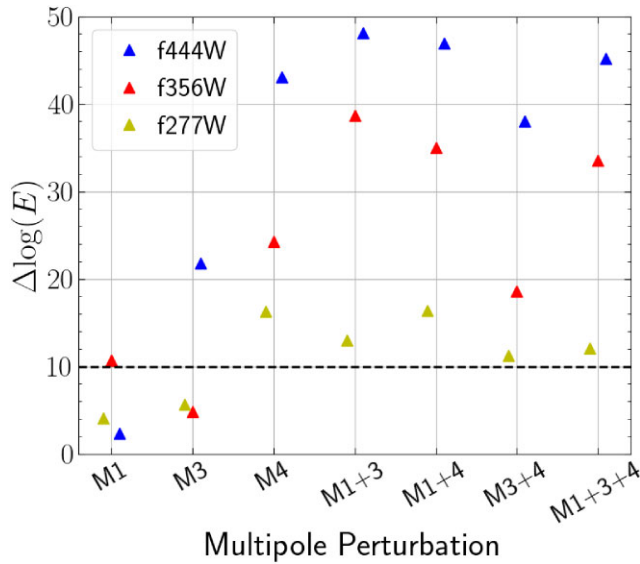


Figure 4. $\Delta\log(\varepsilon)$ (log-evidence changes) with the addition of multipole perturbations to the base model of SPT2147. The dashed line marks $\Delta\log(\varepsilon)=10$, which is our approximate 5σ significance threshold.

errors across all three filters, which shows that the fits are converging successfully since, again, the three filters offer independent source light structures and PSF instances but with the same mass distribution, and thus agreement between the three filters within a model provides an important sanity check and confidence boost in the recovered parameters. Between different models however, we see one model notably deviate from the base model (and indeed most other model) parameters – the M1 + M3 model, showing an offset centre, smaller Einstein radius, smaller power-law slope, smaller shear magnitude and higher M1 magnitude. Some of these M1 + M3 parameters are consistent within errors to at least one filter of another model but the best-fitting values are clearly offset, which is important as this model actually has the highest evidence of all the models in f444 ($\Delta\log(\varepsilon)=48.16$) and f356 ($\Delta\log(\varepsilon)=38.69$). In terms of result trends between multipoles, we do also see a relation between the M3 and M4 multipoles in this lens where the best-fitting M3 magnitude slightly decreases and its angle becomes much less constrained once we add the M4 multipole (in M3 + M4 and M1 + M3 + M4), and similarly for the M4 magnitude when the M3 is added. The M1 magnitude also shows decreases when the M4 is added compared to the M3.

4.1.3 Inclusion of both multipoles and substructure

Evidence increases from the addition of multipoles to the EPL + shear model, and the addition of substructures to this model, are very similar in magnitude, which again may be a possible indication that there is degeneracy between multipoles and substructure in terms of each compensating for the absence of the other, adding to the results of O’Riordan & Vegetti (2024). Running the subhalo searches on these models which have multipole perturbations allows us to see that both can actually co-exist. Fig. 6 shows the subhalo grid search results for the multipole models, with every multipole model shown having a grid cell with $\Delta\log(\varepsilon)>10$ in f444. No models show a cell with $\Delta\log(\varepsilon)>10$ in f277, but we do not draw strong conclusions from this filter due to the low source signal (and therefore low constraining power). The M3 + M4 model has a high-evidence

grid cell ($\Delta\log(\varepsilon)>20$) in f356, potentially accounting for the slight dip seen in the evidence for the multipole perturbation in this model. Importantly, most models in f444 and f356 (except M1 + M3-f356 and M1 + M3 + M4-f356) still favour a subhalo in the same or similar position to the subhalo suggested from the base model search, showing that even with the addition of multipole perturbations, a substructure still obtains higher Bayesian evidences for the model fit.

Fig. 7 shows the finalized multipole evidences and subhalo evidences for the base model and multipole models. We see two results here: firstly that in f444 the multipole perturbations alone cannot reach the same evidence increase that just adding a subhalo to the base model can, however in f356, the multipole perturbations can reach and surpass this level; second, we see that many models lie within a reasonable evidence range of one another, showing the difficulty we will have to choose the ‘best’ models confidently without informative physical priors and context. Following the method set out in Section 3.3.1 however, we utilize the multiwavelength fitting advantages of these data and plot the confidence region of $\Delta\log(\varepsilon)<5$ below the highest overall evidence models in each filter as the shaded regions in Fig. 7. Again, neglecting the f277 filter due to its low constraining power, we see that the M1 + M3 + subhalo and M1 + M4 + subhalo models are the only models which lie within this region in both f444 and f356. Since M1 + M3 + subhalo in f356 lies both at the edge of the region and only $\Delta\log(\varepsilon)\sim 2$ above the M1 + M3 model, we take M1 + M4 + subhalo as the preferred model. Additionally, the M4 multipole perturbation is more easily related to physical structure than the M3 multipole, providing another reason to take this as the preferred model.

Table 3 shows the evidences, positions and masses for the recovered substructure in our favoured model, and Fig. 8 shows the subhalo position and mass for all filters and models. Within individual models, all masses agree between filters, but there are some disagreements between f444 and f277 on position in the base, M1 + 3 and M3 + 4 models. Between models, on the other hand, the f444 and f356 subhalo positions all agree, with some disagreements on the subhalo mass, which is to be expected from models which have different mass distributions in the macromodel. We will discuss in Section 5.2 the potential interpretations of the model-internal disagreements.

The macromodel parameters are shown in Appendix C, Fig. C1, as well as the changes in those parameters from the no-subhalo to the with-subhalo models in Fig. C2. Generally, most changes are consistent with 0, although for best-fitting values we see a general drop in M4 strength, and that the addition of a subhalo tends to bring the power-law slope closer to Isothermal ($\gamma = 2$). The most notable changes upon addition of a subhalo are to the Einstein radius, with the base, M4 and M3 + M4 models showing a definitive drop in Einstein radius in f356 and f444. For our preferred M1 + M4 model, however, the addition of the subhalo leads to no notable change in Einstein radius – there is a potential increase in EPL slope, decrease in shear magnitude and decrease in M4 strength, but all with errors making these changes consistent with 0.

4.1.4 Model residuals

The normalized residuals (i.e. the model residuals divided by the noise, from here forward just ‘residuals’) from the fits to SPT2147 of the base, base + subhalo, base + M1 + M4, and base + M1 + M4 + subhalo models are shown in Fig. 9. There are areas of the upper arc

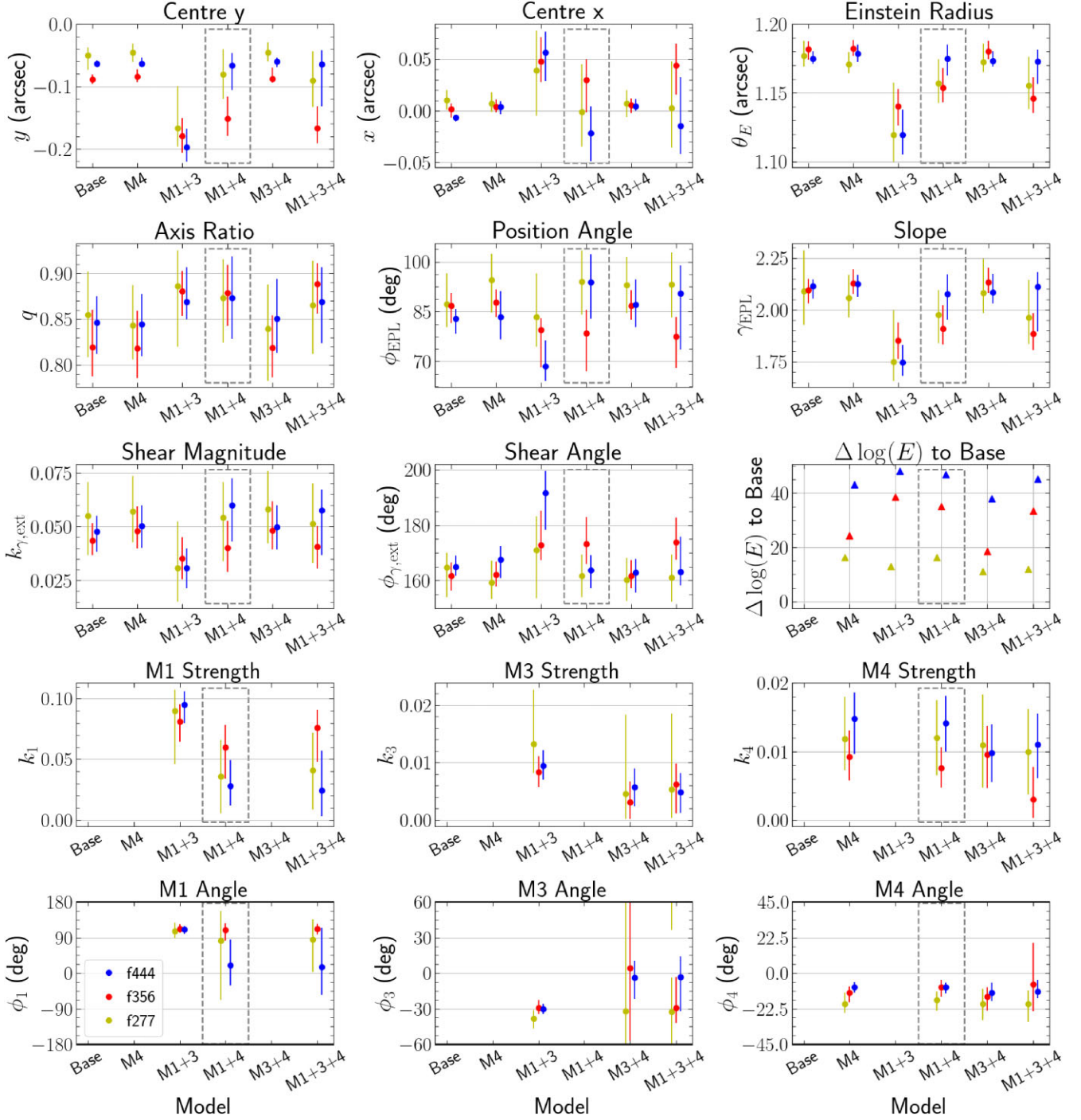


Figure 5. Reported parameters and associated 3σ errors for the highest evidence models without a subhalo in SPT2147, and the base model. Multipole angles have a symmetry through $360/n$ – i.e. a 90 deg error on the M4 multipole means it is unconstrained as this multipole has rotational symmetry through 90 deg, and so the errors are wrapped around $\pm(360/n)/2$. Evidence changes shown are in comparison to the base EPL + shear model. We highlight with the grey box our preferred final model of EPL + shear + M1 + M4.

of the lensed source (marked with back boxes) which see noticeable changes in their residuals from the base model to the model with multipoles and to the model with substructure; however, the change is much less obvious going from these models to the model with multipoles *and* substructure, apart from in the f356 filter, which is a good indicator that the increase in Bayesian evidences we see is due to improvements in the modelling of the lensed source and not the lens itself or the noise.

4.2 SPT0418 results

4.2.1 Base model substructure results

Fig.10 shows the subhalo grid-search results for the f444 and f356 base models in SPT0418. The f277 filter had too low a source signal level to produce reasonable model results, so it is not shown. There is a potential indication of substructure signal towards the lower left of

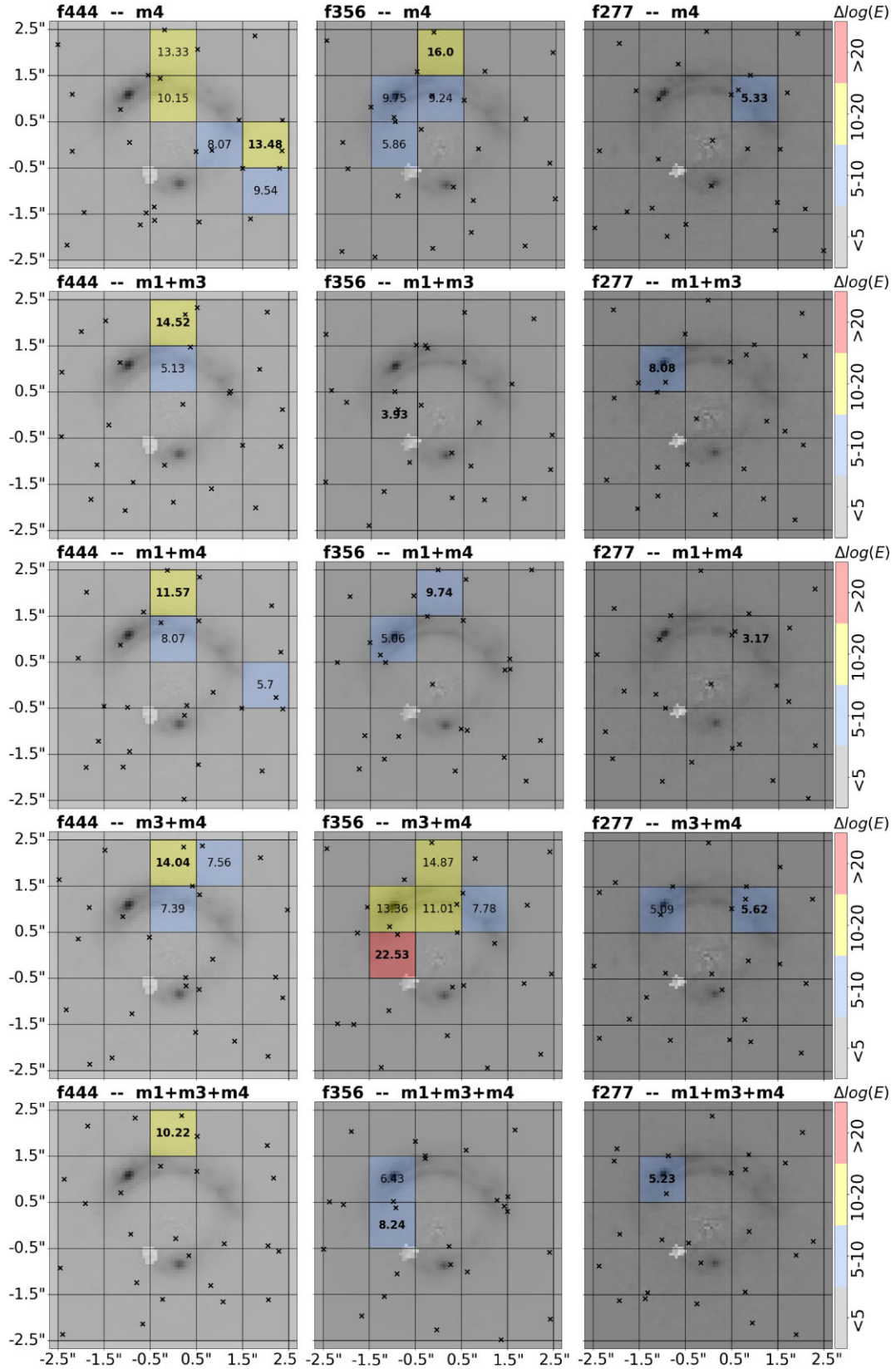


Figure 6. Results from the grid-based preliminary subhalo searches in SPT2147 for the EPL + shear + multipole models. The values and colours correspond to the log-evidence change – $\Delta \log(\epsilon)$ – when a subhalo is modelled in a certain grid cell. Crosses mark the best-fitting position of the subhalo within each grid cell. The highest evidence grid cells are marked in bold. $\Delta \log(\epsilon) = 10$ is considered a roughly 5σ detection.

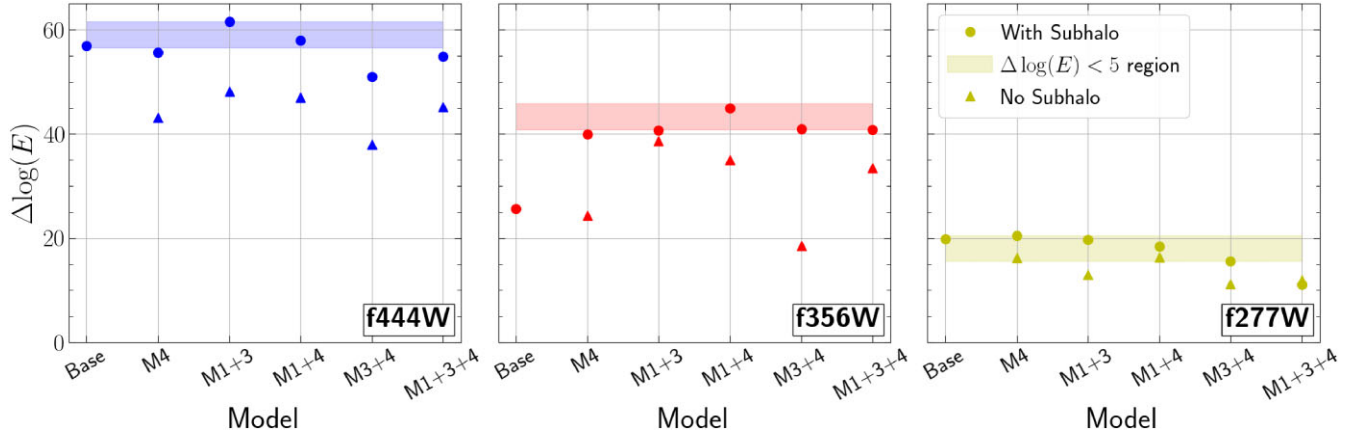


Figure 7. $\Delta\log(\varepsilon)$ (log-evidence changes) for the addition of a multipole (triangles, for reference), and then subhaloes (circles) to highest evidence models of SPT2147. All evidences are compared to the base no-subhalo model. The shaded regions correspond to the (roughly 3σ) $\Delta\log(\varepsilon) < 5$ confidence region below the highest evidence models in each filter.

Table 3. $\Delta\log(\varepsilon)$ (log-evidence changes) and subhalo parameters for our prime model of EPL + shear + M1 + M4 + subhalo. Errors quoted are 3σ . Evidences are quoted to ‘base’ – the EPL + shear model, and to ‘M1 + M4 (no-sub)’ – the EPL + shear + M1 + M4 model before adding the subhalo. All values are consistent across three filters.

| | f444 | f356 | f277 |
|--|-------------------------|-------------------------|-------------------------|
| $\Delta\log(\varepsilon)$ to base | 57.96 | 45.00 | 18.46 |
| $\Delta\log(\varepsilon)$ to M1 + 4 (no-sub) | 10.97 | 9.96 | 2.07 |
| Position y (arcsec) | $2.47^{+0.27}_{-0.33}$ | $2.59^{+0.13}_{-0.25}$ | $1.19^{+1.34}_{-1.55}$ |
| Position x (arcsec) | $-0.07^{+0.29}_{-0.19}$ | $-0.06^{+0.15}_{-0.13}$ | $0.61^{+1.43}_{-1.48}$ |
| $\log_{10}(M_{200}/M_{\odot})$ | $10.87^{+0.53}_{-0.71}$ | $10.99^{+0.33}_{-0.37}$ | $10.17^{+0.64}_{-2.12}$ |

the image, but in only one grid cell in each filter, and moving closer to the Einstein ring in the lower signal f356 filter. The grid-search evidences are $\Delta\log(\varepsilon) \sim 5$, and the finalized search evidences are $\Delta\log(\varepsilon) = 4.44, 4.31$ for f444, f356, meaning there is no appreciable evidence for substructure in this lens when using the base EPL + shear model.

4.2.2 Results of multipole perturbations

Given that there is no evidence for substructure with the base EPL + shear model, we already know that the model can fit the data reasonably well without any additional perturbations, and thus we do not expect strong evidences for multipole complexity in the mass model. Fig. 11 shows $\Delta\log(\varepsilon)$ (evidence–triangles) and $\Delta\log(L)$ (likelihood–crosses) when multipoles are added to the EPL + shear model. The small positive log-likelihood increases indicate that multipole models produce a small improvement to the overall fit. However, the decrease in evidence implies that this does not outweigh the increase in model complexity and thus the multipole models are disfavoured overall.

The f444 filter has no increase in $\Delta\log(\varepsilon)$ for any multipole addition, and in fact, as more multipoles are added, the magnitude of the evidence drop increases – as one would expect if adding further unnecessary complexity to the model. The f356 filter does show minor increases of $\Delta\log(\varepsilon) < 10$ in all additions except for M1-only and M3-only. The fit to the f356 filter however required the model to be initialized from the f444 model as the source signal was too low to provide a reasonable constraint on parameters. We therefore place

much less value on the f356 results and show them primarily for completeness, using the f444 as the primary result set. Regardless, the low evidences for multipoles in f356 coupled with the negative evidences for multipoles in f444 lead to the conclusion that there is no basis to claim multipole angular complexity in the mass of lens – an important contrast to SPT2147. We show the parameters for the fits in Appendix C, Fig. C3 and show that they are all consistent across every model and filter, with the exception of f444 shear angles being generally higher than f356. The consistency and the fact that all multipole strengths return close to zero strengthens the opinion of the base model being most appropriate as the multipole models offer little to no difference.

4.2.3 Inclusion of both multipoles and substructure

As mentioned, the f356 model fit required initializing from the f444 model results, and therefore we primarily test the f444 model with substructure additions. The f444 model, however, had all negative log-evidence changes upon multipole addition, so we therefore only test the addition of a substructure to the models with a single multipole perturbation. Fig. 12 shows the finalized log-evidence changes from the smooth EPL + shear model for the multipole additions (for reference) and the multipole + subhalo (or just subhalo for the base model) additions for the f444 base, M1, M3, M4 models and the f356 base and M1 models.

The subhalo grid-searches for the multipole models are not shown as they contain no cells with $\Delta\log(\varepsilon) > 10$, and only one cell with $\Delta\log(\varepsilon) > 5$ in f444, although interestingly this is now on the lower right of the lens rather than the lower left as seen in Fig. 10, but is likely just fitting noise or compensating for the multipole addition. Fig. 12 does show that the M1 subhalo fit in f356 has an evidence boost from the smooth M1 model of $\Delta\log(\varepsilon) = 8.25$, but considering the addition of the M1 causes a drop in evidence of -2.88 from the base EPL + shear model, we suspect the subhalo is primarily just compensating for this.

4.2.4 Model residuals

We plot the normalized residuals for SPT0418 in Fig. 13, showing the base model, base + subhalo model, and M1 + M3 + M4 model. We see that, unlike in SPT2147 (Section 4.1.4), there is no noticeable

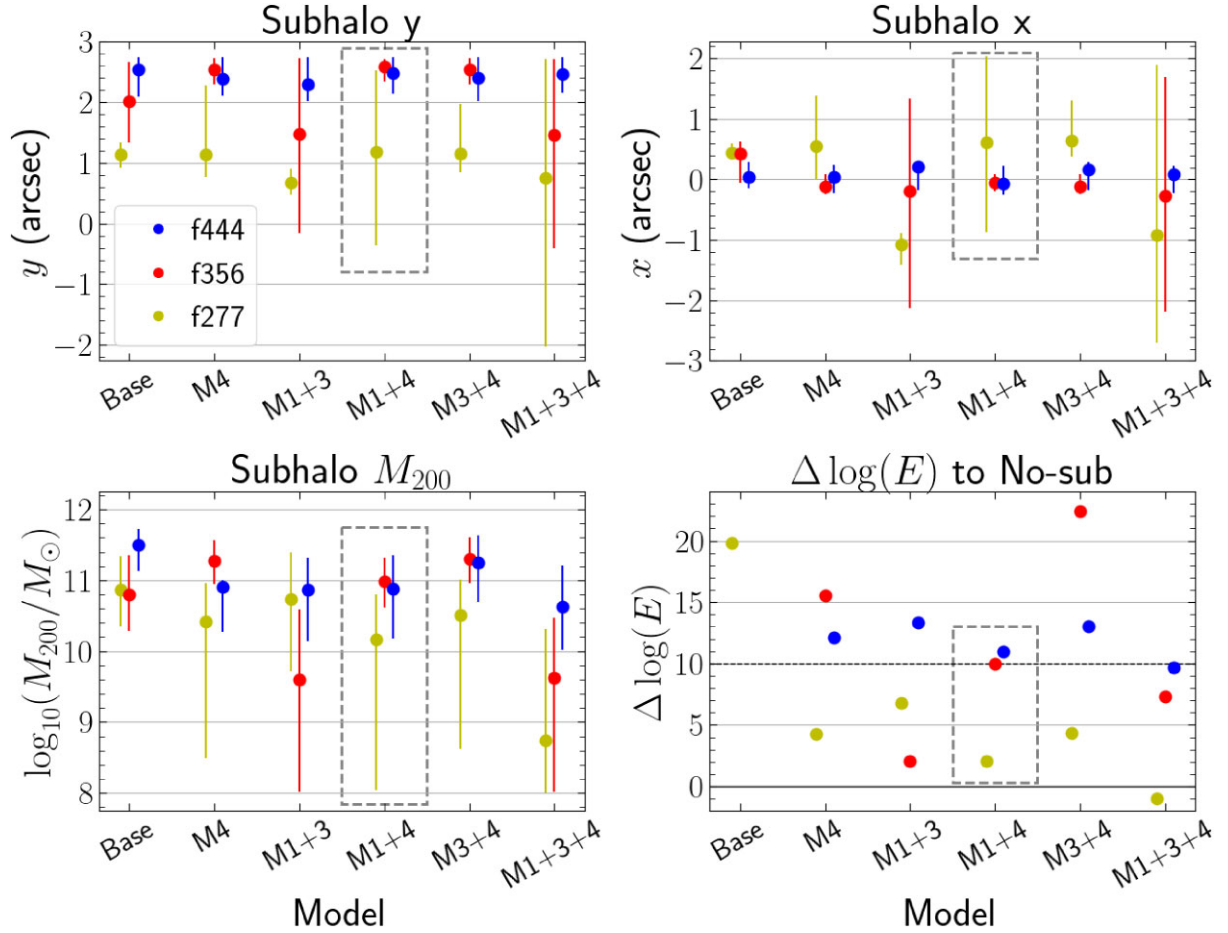


Figure 8. Reported subhalo parameters for the best-fitting substructure to the base and high-evidence multipole models. Our preferred model, EPL + shear + M1 + M4 is highlighted with the grey box. Log-evidence changes – $\Delta \log(\epsilon)$ – compared to the no-subhalo model of that type are also shown, but the base f356 and f444 evidences are above the plot limit.

change in the residuals when adding either angular complexity or substructure to the base model, which supports the narrative that this lens does not have detectable complexity.

5 DISCUSSION

5.1 Is there a dark subhalo in SPT2147?

Our analysis of SPT2147 provides strong evidence (exceeding 10σ) for both (i) models with angular complexity (multipoles) when a DM subhalo is omitted, and (ii) models including a DM subhalo when multipoles are omitted. The preferred model (see Section 3.3.1) includes both $m1$ and $m4$ multipoles and a DM subhalo, showing evidence increases of $\Delta \log E = 10.97, 9.96,$ and 4.3 for the f444W, f356W, and f277W, compared to models without a DM subhalo. This suggests that SPT2147 likely requires both angular complexity in the lens mass distribution and the presence of a DM subhalo. However, systematic uncertainties in strong lens modelling make it difficult to come to a definitive interpretation, leading us to consider three possibilities:

(i) **Lens galaxy contains only angular complexity, without a DM subhalo:** In this scenario, the DM subhalo is producing a false positive detection with an evidence increase above 50 by

compensating for missing complexity in the power-law lens mass model, an effect shown in He et al. (2023). However, the model including $m1$ and $m4$ multipoles still favours the DM subhalo with an evidence increase of 10.97, and even with all of $m1, m3,$ and $m4$, the evidence increase for substructure is 9.68, meaning that if a DM subhalo is truly not present the mass distribution must have complexity beyond what the combination of the $m1, m3,$ and $m4$ multipoles can capture.

(ii) **Lens galaxy contains a DM subhalo, without angular complexity:** Here, the multipole-only models are favoured because they are actually fitting the lensing signal of a genuine DM subhalo, a degeneracy outlined by O’Riordan & Vegetti (2024). However, some models with the combination of multipoles and substructure are consistently a higher evidence than the models with substructure only, implying that there is excess lensing potential beyond what can be captured by a single subhalo.

(iii) **Lens galaxy contains both angular complexity and a DM subhalo:** A combination of the above cases. The models including a DM subhalo but lacking multipoles see evidence increases of ~ 50 because the subhalo primarily accounts for missing complexity in the power-law model of the central mass distribution (see He et al. 2023). However, even after the multipoles are added, accounting for this missing complexity, a DM subhalo remains necessary for an improved model fit as there is a true subhalo in the system.

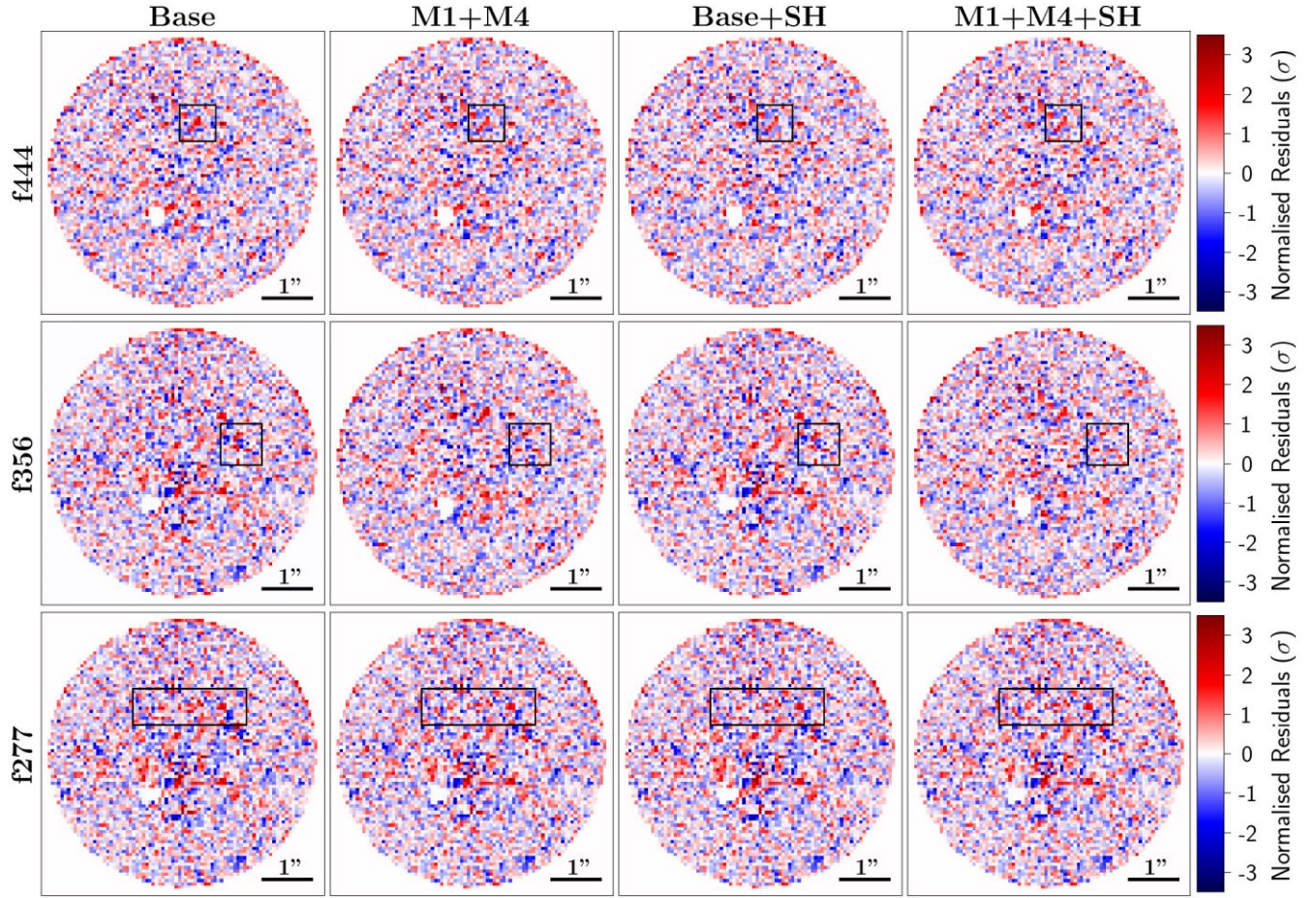


Figure 9. Normalized residuals (residuals divided by the noise in each pixel) from model fits to SPT2147 of the ‘base’ (EPL + shear) model and the model including $m = 1$ and $m = 4$ multipoles, and their corresponding fits with substructure additions (‘SH’). The black boxes highlight areas where there are notable and correlated residual changes within a filter, all corresponding to the upper arc of the lensed source.

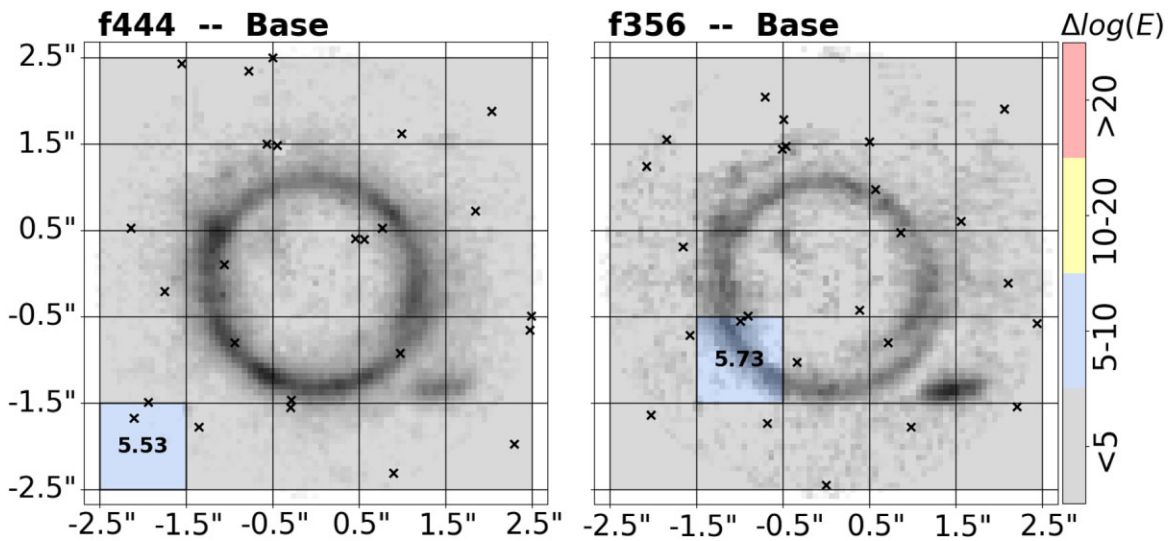


Figure 10. Results from the grid-based preliminary subhalo searches in SPT0418 for the base EPL + shear model. The values and colours correspond to the log-evidence change $-\Delta\log(\epsilon)$ when a subhalo is modelled in a certain grid cell. Crosses mark the best-fitting position of the subhalo within each grid cell. The highest evidence grid-cells are marked in bold. $\Delta\log(\epsilon) = 10$ is considered a roughly 5σ detection.

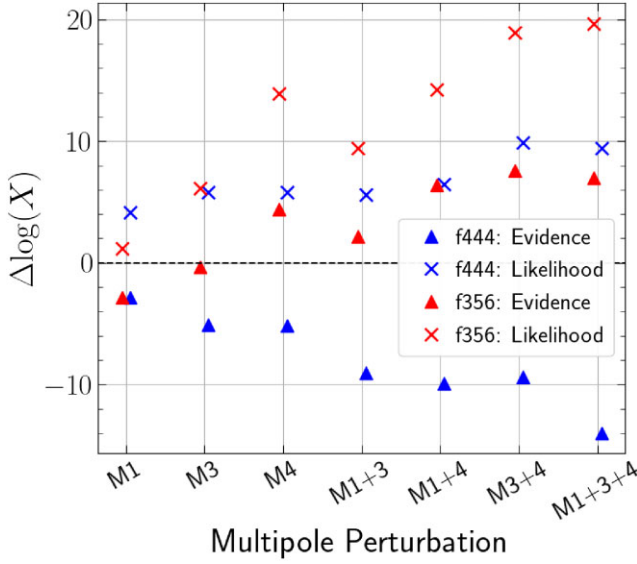


Figure 11. $\Delta\log(\epsilon)$ (log-evidence changes, triangles) and $\Delta\log(L)$ (log-likelihood changes, crosses) for the addition of multipole perturbations to the base model of SPT0418.

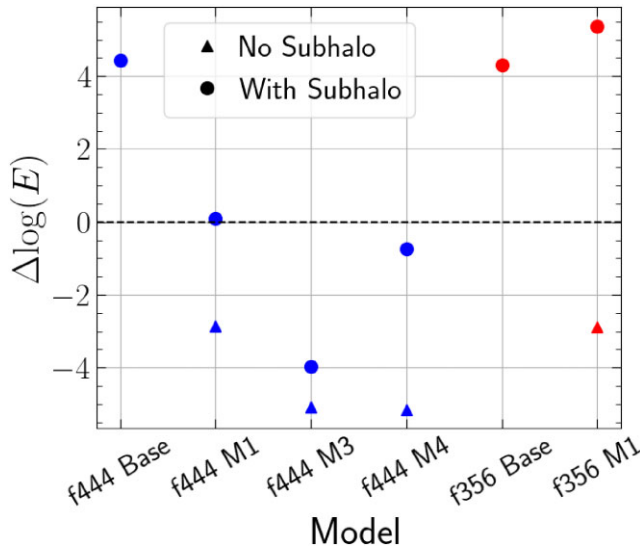


Figure 12. $\Delta\log(\epsilon)$ (log-evidence changes) for the addition of a multipole (triangles, for reference), and then subhaloes (circles) to the best models for SPT0418. All evidences are compared to the base no-subhalo smooth model. The dashed line marks $\Delta\log(\epsilon)=0$.

Option iii above, that this lens contains both substructure and multipoles, is the solution with the highest Bayesian evidence, and is a similar option to recent work by Enzi et al. (2024) on SDSSJ0946+1006, who find evidence for significant multipoles alongside the substructure in this lens, but note that the multiple source configuration of this lens is essential to reducing the degeneracy between multipoles and other parameters – a luxury that we do not have with SPT2147-50.

Given the current evidence, we consider SPT2147 as a DM subhalo candidate. Future work will aim to refine this interpretation using more flexible lens mass models, such as potential corrections (Suyu et al. 2009; Vegetti et al. 2012). SPT2147 is an ideal target to investigate this approach, given the apparent need for both angular

complexity and a DM subhalo in its modelling, along with the availability of multiwavelength data for validation. Additionally, follow-up lensing analysis with ALMA data is planned, building on Amvrosiadis et al. (2025). Whilst we are not there yet, SPT2147 suggests that separating a DM substructure from angular complexity in strongly lensed extended sources is a tractable task, ensuring that it can prove a key test of the DM particle in the near future.

5.2 Candidate position and mass

The best-fitting position for the DM subhalo in SPT2147 lies approximately 1 arcmin north of the uppermost part of the lensed arc. In the f277W filter, the subhalo’s position shifts on to the lensed arc; however, this is not in tension with the f444W and f356W results, due to greatly increased errors on the position. In the upper panels of Fig. 8 we see that all subhalo x -positions agree to 3σ except in the M3 + 4 f277W model, and subhalo y -positions agree except f277W-f444W results with the base and M1 + M3 models – the important point being that there is good agreement to 3σ in subhalo position (and masses) of our preferred model of M1 + M4. The reason for the discrepancy in precision of the f277W results is likely due to the much lower signal-to-noise ratio of the lensed source in this filter (see Table 1) and the resultant need of the model to bring the subhalo closer to the arc to achieve a noticeable perturbation with this low signal. Due to this, we assume that the true subhalo position aligns with the fits from the f444W and f356W bands, at the centre-top of the image, and primarily look at these filters (and specifically f444W) from here forward.

While the subhalo’s relatively large inferred mass helps explain how it can still cause a lensing perturbation at this distance from the arc, a DM halo of $\sim 10^{11}M_{\odot}$ would typically be expected to host a visible dwarf galaxy. However, no such galaxy is observed. In Appendix B, we estimate the upper limit for the subhalo’s brightness using the JWST Exposure Time Calculator (Pontoppidan et al. 2016), with settings that match those of the observing run (Rigby et al. 2024) for F444W. The result is an expected SNR of approximately 0.16, indicating that the object would be indistinguishable from the background. A dwarf galaxy detection is not unreasonable as an alternative to a dark subhalo; recent work from Ballard et al. (2024) showed that the substructure detection in SDSSJ0946 + 1006 is compatible with a dwarf galaxy which is too faint to be detected given the available data.

The inferred subhalo mass depends significantly on the choice of its density profile. Our analysis assumes a Navarro–Frenk–White (NFW) profile, which has a shallower central density compared to the pseudo-Jaffe profile used in other works (e.g. Vegetti et al. 2010, 2012). This assumption leads to higher subhalo mass estimates, as seen in the analysis of SDSSJ0946 + 1006 by Nightingale et al. (2024), where the inferred mass was around $\sim 10^{10.5}M_{\odot}$, compared to masses in the region of $10^{9.5}M_{\odot}$ inferred from studies using more concentrated profiles (e.g. Minor et al. 2021b). Additionally, Despali et al. (2018) demonstrated that masses derived using NFW profiles map to lower values when assuming pseudo-Jaffe profiles.

Further exploration of the DM subhalo’s density profile could suggest a mass closer to $5.0 \times 10^9 M_{\odot}$ or below if the subhalo is more concentrated. In fact, a recent study from Despali et al. (2024) found that for two detections of dark subhaloes in the literature, these subhaloes prefer a steeper density slope than the NFW profile, so it is likely that our quoted masses are an over-estimate due to their lower concentration. Additionally, if the subhalo lies along the line-of-sight rather than within the lens plane, this could affect both its detectability and the inferred mass (e.g. He et al. 2022). The

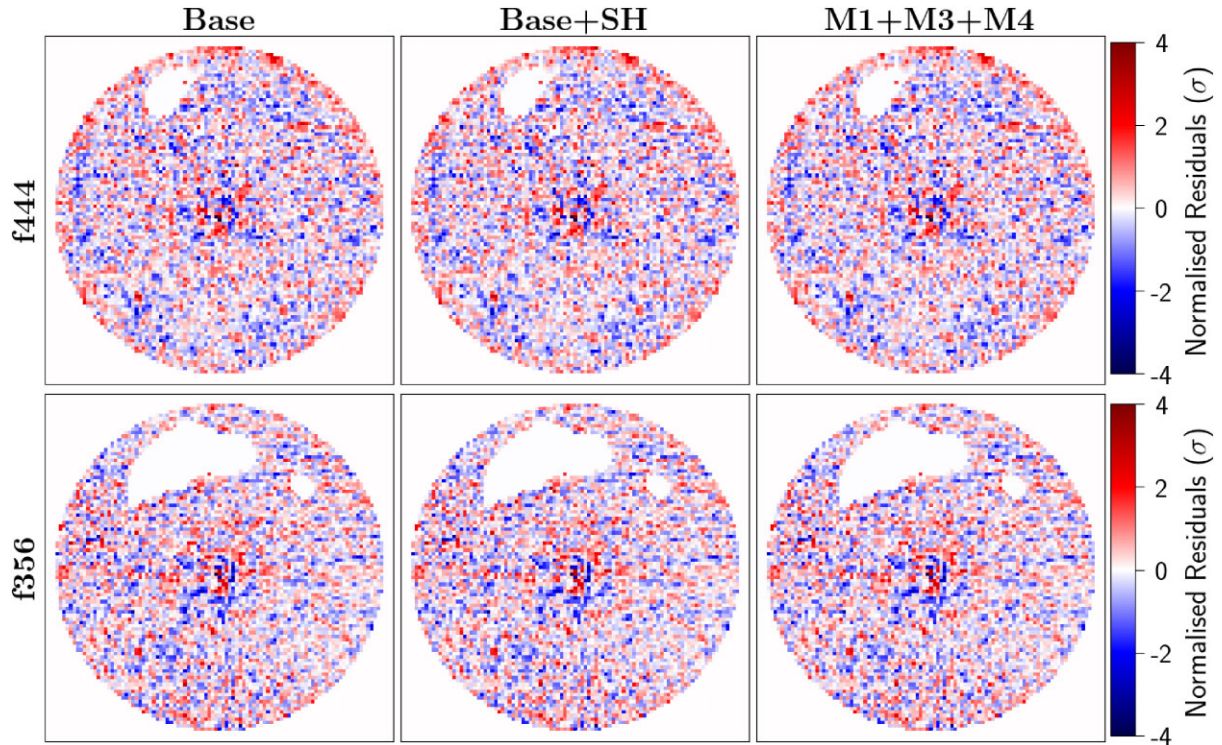


Figure 13. Normalized residuals (residuals divided by noise in each pixel) from the model fits to SPT0418 for the ‘base’ (EPL + shear) model, this model with subhalo (‘SH’), and the model containing $m = 1$, $m = 3$, and $m = 4$ multipoles (but no substructure). We see no areas where there are major correlated residual changes when extra complexity (angular or substructure) is added to the model.

availability of multiwavelength data for SPT2147 makes it an ideal candidate for future investigations into the subhalo’s density profile and its potential impact on mass estimates.

5.3 Substructure detection criteria

Previous strong lens studies that assume a power-law plus shear mass model have adopted different Bayesian evidence criteria for identifying DM subhalo candidates. Detection thresholds vary significantly: Vegetti et al. (2010, 2012) use $\Delta\log(\epsilon)$ of 50, Ritondale et al. (2019) use 100, and Nightingale et al. (2024) use 10. Similarly to this work, Hezaveh et al. (2016) claim a detection in a different strong lens with a value of $\Delta\log(\epsilon) = 47.3$ for fits using a power-law plus shear plus $m = 3$ and $m = 4$ multipoles. Many sensitivity mapping studies set thresholds between 50 and 10 (Amorisco et al. 2021; Despali et al. 2022; O’Riordan et al. 2023; O’Riordan & Vegetti 2024). An increase of 10 corresponds to roughly a 5σ detection, making this an appropriate benchmark for considering a detection as significant.

Our study shows that the significance of a DM subhalo detection is sensitive to the choice of mass model. For example, in SPT2147, a power-law plus shear model favoured a DM subhalo with evidence increases of $\Delta\log(\epsilon) = 60.35, 42.13, 20.54$ in the F444W, F356W, and F277W bands, respectively. In contrast, when including $m = 1$ and $m = 4$ multipoles, these values dropped to $\Delta\log(\epsilon) = 10.97, 9.96, 4.3$. This pattern is also observed in other analyses. For example, Nightingale et al. (2024) found that the evidence for a DM subhalo in SDSSJ0946 + 1006, a lens with a confirmed substructure, varied from $\Delta\log(\epsilon) = 72.36$ for a power-law plus shear model to 22.52 for a decomposed stars-plus-DM model. These results suggest that more refined mass models can reduce the

evidence threshold needed to detect subhaloes, as they better capture the underlying complexity of the lens.

This has important implications for constraints on the mass of the DM particle. As mentioned, sensitivity mapping studies often assume a threshold for detection of up to 50, which may be conservative in their estimates of DM subhalo detectability. If improved mass models, and cross-referencing across multipole filters, allow for a lower detection threshold, these forecasts could be adjusted accordingly. This could lead to a scenario where fewer lenses are needed to achieve comparable constraints on the DM particle mass, even though more complex mass models are now assumed. A more thorough evaluation of how to establish robust detection criteria is therefore essential for future studies, and whether our implementation, as described in Section 3.3.1, remains appropriate.

5.4 Contrasting SPT2147 with SPT0418

The contrast in results between SPT2147 and SPT0418 shows that we do not just detect multipoles simply because they are included in the model. SPT0418 is a prime example that actually not all lens mass models require high levels of complexity, and it is important to note that this improves the confidence that the detections in SPT2147 are not systematics where complexity is added to artificially inflate the evidence (i.e. the model does not just want all the parameters it can get), since that would then also be the case here. Further, as mentioned in Sections 4.1.4 and 4.2.4, there is a difference in the effects of the complexity additions on the model residuals – where SPT2147 shows evidence of a correlated improvement to the residuals along the lensing arcs, whereas SPT0418 shows no such improvements. This shows the physical meaning behind the changes (or lack thereof) to Bayesian evidences that are seen.

Given some of the recovered axis ratio parameters, the SPT0418 lens mass could likely be reasonably well described by a singular isothermal sphere mass model ('SIS': the simplest applicable model which is the EPL with $q = 1$, $\gamma = 2$). This dramatic mass simplicity could have numerous explanations: the level of source signal may be too weak to constrain any complexity in the lens mass; the source itself, which is a merger of two elliptical galaxies (Cathey et al. 2024), may, despite its large span over the source plane, lack suitably placed features to enable detection of complexity (i.e. reconstruction of a spiral has fewer lensing solutions than reconstruction of a blob); or alternatively, this lensing galaxy may genuinely have little to no angular structure. Regardless, the lack of multipole evidence in SPT0418 makes the indications shown of angular complexity in SPT2147 that much more compelling, and enhances the argument for further study of this lens. What would additionally be interesting is deeper imaging of SPT0418, in order to provide higher levels of source signal in order to re-test the application of multipoles to see if the evidences change.

6 SUMMARY AND CONCLUSIONS

We have investigated multiwavelength images from JWST of two strongly lensed systems to test the effects of angular (multipole) perturbations to the standard elliptical power-law mass model, and the impact of these perturbations on substructure detection. We find that (1) SPT2147-50 has evidence for angular complexity (of first- and fourth-order multipole combination), with Bayes factor $\Delta\log(\epsilon) \sim 47, 35, 16$ in f444, f356, f277 imaging respectively; (2) the inclusion of these multipoles reduces the Bayes factor for substructure by $\sim 46, 16, 18$ for f444, f356, f277; (3) irrespective of complexity, our highest SNR filter (f444) always gives a substructure detection above 4σ . We will refer to this as a DM *candidate*, as we still require the non-parametric potential corrections (Vegetti & Koopmans 2009) to definitively rule it out as a mass modelling systematic error. We do not, however, find any evidence for complexity in SPT0418 (angular or substructure), which indicates that the detections in SPT2147 are less likely to be systematics related to the modelling process.

A quantitative summary (see also Table C1 in the appendix) for SPT2147 is as follows:

(i) When assuming only a standard power law plus external shear model for the lens galaxy mass, we favour a DM subhalo inclusion with Bayes factor $\Delta\log(\epsilon) = 56.90, 25.70, 19.84$ in filters f444, f356, f277. These are all well above 5σ .

(ii) We find similarly strong evidence for angular mass complexity in the lens galaxy mass distribution. Inclusion of a combination of first-, third-, and fourth-order multipoles is favoured by Bayes factors $\Delta\log(\epsilon) = 45.19, 33.54, 12.05$ and inclusion of only first- and fourth-order multipoles is favoured by Bayes factors $\Delta\log(\epsilon) = 46.99, 35.04, 16.39$ (in filters f444, f356, f277, respectively; again all well above 5σ).

(iii) Based on the higher source-SNR filters f444 and f356 only, depending on the combinations of multipoles and substructure used, the first-order multipole has magnitude in the range 2.43–9.52 per cent; the third-order multipole has magnitude 0.29–1.34 per cent; and the fourth-order multipole has magnitude 0.30 – 1.48.

(iv) A dark substructure is still favoured in models that include multipoles, but to a lesser extent. By implementing a 'confidence region' of $\Delta\log(\epsilon) < 5$ from the highest evidence model in each filter, the best applicable model (within this region in all three filters) includes only first- and fourth-order multipoles

(plus substructure), and the Bayes factors for substructure additions are $\Delta\log(\epsilon) = 10.97, 9.96, 4.35$ in f444, f356, f277 (approximately 5.05, 4.85, 3.4σ).

(v) This best model substructure has mass $\log_{10} M_{200} = 10.87^{+0.53}_{-0.71} M_{\odot}$ in f444 and $\log_{10} M_{200} = 10.99^{+0.33}_{-0.37} M_{\odot}$ in f356. This is a substructure *candidate*, requiring more detailed future analysis to definitively confirm it is not a systematic resulting from an overly simplistic mass model.

Our research contributes to the expanding body of evidence that detecting DM subhaloes through strong lensing necessitates a meticulous separation of the DM subhalo signal from the radial and angular complexities inherent in the lens galaxy's mass distribution. This conclusion aligns with previous findings from cosmological simulations (He et al. 2023), observational data (Barone et al. 2024; Nightingale et al. 2024), and machine learning applied to large samples of simulated lenses (O'Riordan & Vegetti 2024), as well as studies explaining flux ratio anomalies in lensed quasars through mass model angular complexity (Gilman et al. 2024). Despite fitting one of the most complex parametric mass models currently possible – incorporating a power-law profile with external shear and first-, third-, and fourth-order multipoles – we still detect statistical evidence at a significance level exceeding 5σ in favour of the presence of a DM subhalo. Although we cautiously refer to this as a *candidate* and acknowledge the need for more advanced lens modelling including potential corrections to reach a definitive conclusion, our findings suggest that disentangling DM substructure from angular complexity is an achievable challenge. This progress will eventually enable us to critically assess the CDM paradigm.

SOFTWARE CITATIONS

This work uses the following software packages:

- (i) **ASTROPY** (Astropy Collaboration 2013; Price-Whelan et al. 2018)
- (ii) **BAGPIPES** (Carnall et al. 2018)
- (iii) **COLOSSUS** (Diemer 2018)
- (iv) **CORNER.PY** (Foreman-Mackey 2016)
- (v) **MATPLOTLIB** (Hunter 2007)
- (vi) **NAUTILUS** (Lange 2023)
- (vii) **NUMBA** (Lam, Pitrou & Seibert 2015)
- (viii) **NUMPY** (van der Walt, Colbert & Varoquaux 2011)
- (ix) **PYAUTOFIT** (Nightingale et al. 2021a)
- (x) **PYAUTOGALAXY** (Nightingale et al. 2018; Nightingale et al. 2023a)
- (xi) **PYAUTOLENS** (Nightingale & Dye 2015; Nightingale et al. 2018, 2021b)
- (xii) **PYSWARMS** (Miranda 2018)
- (xiii) **PYTHON** (Van Rossum & Drake 2009)
- (xiv) **SCIKIT-IMAGE** (Van der Walt et al. 2014)
- (xv) **SCIKIT-LEARN** (Pedregosa et al. 2011)
- (xvi) **SCIPY** (Virtanen et al. 2020)

ACKNOWLEDGEMENTS

This work was supported by Science and Technology Facilities Council (STFC) consolidated grants ST/X001075/1 and ST/T000244/1. Samuel Lange is supported by STFC studentship grant S0600186. Carlos Frenk acknowledges support by the European Research Council (ERC) through Advanced Investigator grant to C.S. Frenk, DMIDAS (GA 786910). This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology

on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). Most computing was performed on the COSMA supercomputers, where the equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1, and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. This work was also supported using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service, provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/T022159/1), and DiRAC STFC funding.

We would also like to thank the reviewer for their insightful comments which helped improve the paper.

DATA AVAILABILITY

Data for the observations used in this work are publicly available at the STScI MAST data base – mast.stsci.edu, with an overview of the observing project (*TEMPLATES*) available at (Rigby et al. 2024, doi: [10.3847/1538-4357/ad7501](https://doi.org/10.3847/1538-4357/ad7501)). The software used for lens modelling, PYAUTOLENS, is open source and available at github.com/Jammy2211/PyAutoLens. Model parameters and images, including lens and source light subtracted images, are available upon request; please contact the authors.

REFERENCES

- Aghanim N. et al., 2020, *A&A*, 641, A6
- Amorisco N. C. et al., 2021, *MNRAS*, 510, 2464
- Amvrosiadis A. et al., 2024, preprint ([arXiv:2407.12983](https://arxiv.org/abs/2407.12983))
- Amvrosiadis A. et al., 2025, *MNRAS*, 537, 1163
- Astropy Collaboration, 2013, *A&A*, 558, A33
- Ballard D. J., Enzi W. J. R., Collett T. E., Turner H. C., Smith R. J., 2024, *MNRAS*, 528, 7564
- Barone T. M. et al., 2024, *Communications Physics*, 7, 286
- Benacchio L., Galletta G., 1980, *MNRAS*, 193, 885
- Bender R., Surma P., Doebereiner S., Moellenhoff C., Madejsky R., 1989, *A&A*, 217, 35
- Bergamini P. et al., 2023, *ApJ*, 952, 84
- Bode P., Ostriker J. P., Turok N., 2001, *ApJ*, 556, 93
- Bro R., De Jong S., 1997, *Journal of Chemometrics*, 11, 393
- Cao X. et al., 2022, *Res. Astron. Astrophys.*, 22, 025014
- Cappellari M., 2002, *MNRAS*, 333, 400
- Cappellari M., 2016, *ARA&A*, 54, 597
- Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, *MNRAS*, 480, 4379
- Cathey J. et al., 2024, *ApJ*, 967, 11
- Chu Z., Lin W., Li G., Kang X., 2013, *ApJ*, 765, 134
- Cohen J. S., Fassnacht C. D., O’Riordan C. M., Vegetti S., 2024, *MNRAS*, 531, 3431
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- Despali G., Vegetti S., White S. D., Giocoli C., van den Bosch F. C., 2018, *MNRAS*, 475, 5424
- Despali G., Vegetti S., White S. D., Powell D. M., Stacey H. R., Fassnacht C. D., Rizzo F., Enzi W., 2022, *MNRAS*, 510, 2480
- Despali G., Heinze F. M., Fassnacht C. D., Vegetti S., Spingola C., Klessen R., 2024, preprint ([arXiv:2407.12910](https://arxiv.org/abs/2407.12910))
- Diemand J., Kuhlen M., Madau P., Zemp M., Moore B., Potter D., Stadel J., 2008, *Nature*, 454, 735
- Diemer B., 2018, *ApJ Supplement Series*, 239, 35
- van Dokkum P., Brammer G., Wang B., Leja J., Conroy C., 2024, *Nat. Astron.*, 8, 119
- Enzi W. et al., 2021, *MNRAS*, 506, 5848
- Enzi W. J. R., Krawczyk C. M., Ballard D. J., Collett T. E., 2024, preprint ([arXiv:2411.08565](https://arxiv.org/abs/2411.08565))
- Etherington A. et al., 2022, *MNRAS*, 517, 3275
- Etherington A. et al., 2023, *MNRAS*, 521, 6005
- Etherington A. et al., 2024, *MNRAS*, 531, 3684
- Foreman-Mackey D., 2016, *The J. Open Source Softw.*, 1, 24
- Frenk C., White S., 2012, *Annalen der Physik*, 524, 507
- Frenk C. S., White S. D., Efstathiou G., Davis M., 1985, *Nature*, 317, 595
- Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, *ApJ*, 327, 507
- Gao L., White S. D., Jenkins A., Stoehr F., Springel V., 2004, *MNRAS*, 355, 819
- Gilman D., Birrer S., Nierenberg A., Oh M. S. H., 2024, *MNRAS*, 533, 1687
- Girelli G., Pozzetti L., Bolzonella M., Giocoli C., Marulli F., Baldi M., 2020, *A&A*, 634, A135
- Hao C. N., Mao S., Deng Z. G., Xia X. Y., Wu H., 2006, *MNRAS*, 370, 1339
- He Q. et al., 2022, *MNRAS*, 512, 5862
- He Q. et al., 2023, *MNRAS*, 518, 220
- He Q. et al., 2024, *MNRAS*, 532, 2441
- Hezaveh Y. D. et al., 2016, *ApJ*, 823, 37
- Hunter J. D., 2007, *Comput Sci Eng*, 9, 90
- Johnson D., Fleury P., Larena J., Marchetti L., 2024, *J. Cosmol. Astropart. Phys.*, 2024, 055
- Keeton C., Kochanek C., Seljak U., 1997, *ApJ*, 482, 604
- Lam S. K., Pitrou A., Seibert S., 2015, Finkel H., ed., Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC – LLVM’15. Association for Computing Machinery, New York, USA, p. 1
- Lange J. U., 2023, *MNRAS*, 525, 3181
- Li R., Frenk C. S., Cole S., Gao L., Bose S., Hellwing W. A., 2016, *MNRAS*, 460, 363
- Lovell M. R., Frenk C. S., Eke V. R., Jenkins A., Gao L., Theuns T., 2014, *MNRAS*, 439, 300
- Ludlow A. D., Bose S., Angulo R. E., Wang L., Hellwing W. A., Navarro J. F., Cole S., Frenk C. S., 2016, *MNRAS*, 460, 1214
- Macciò A. V., Fontanot F., 2010, *MNRAS*, 404, L16
- Mercier W. et al., 2024, *A&A*, 687, A61
- Minor Q., Kaplinghat M., Chan T. H., Simon E., 2021a, *MNRAS*, 507, 1202
- Minor Q., Gad-Nasr S., Kaplinghat M., Vegetti S., 2021b, *MNRAS*, 507, 1662
- Miranda L. J. V., 2018, *J. Open Source Softw.*, 3, 433
- Moster B. P., Somerville R. S., Maulbetsch C., Van Den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, 710, 903
- Navarro J. F., Frenk C. S., White S. D., 1997, *ApJ*, 490, 493
- Newton O. et al., 2021, *J. Cosmology Astropart. Phys.*, 2021, 062
- Newton O., Lovell M. R., Frenk C. S., Jenkins A., Helly J. C., Cole S., Benson A. J., 2024, preprint ([arXiv:2408.16042](https://arxiv.org/abs/2408.16042))
- Nightingale J. W., Dye S., 2015, *MNRAS*, 452, 2940
- Nightingale J. W., Dye S., Massey R. J., 2018, *MNRAS*, 478, 4738
- Nightingale J. W., Massey R. J., Harvey D. R., Cooper A. P., Etherington A., Tam S.-I., Hayes R. G., 2019, *MNRAS*, 489, 2049
- Nightingale J. W., Hayes R. G., Griffiths M., 2021a, *J. Open Source Softw.*, 6, 2550
- Nightingale J. W. et al., 2021b, *J. Open Source Softw.*, 6, 2825
- Nightingale J. W. et al., 2023a, *J. Open Source Softw.*, 8, 4475
- Nightingale J. W. et al., 2023b, *MNRAS*, 521, 3298
- Nightingale J. W. et al., 2024, *MNRAS*, 527, 10480
- O’Riordan C. M., Vegetti S., 2024, *MNRAS*, 528, 1757
- O’Riordan C. M., Despali G., Vegetti S., Lovell M. R., Moliné A., 2023, *MNRAS*, 521, 2342
- Padilla N. D., Strauss M. A., 2008, *MNRAS*, 388, 1321
- Pedregosa F. et al., 2011, *J. Mach. Learn. Res.*, 12, 2825
- Peng B. et al., 2023, *ApJ*, 944, L36
- Planck Collaboration XIII, 2016, *A&A*, 594, A13
- Polletta M. et al., 2007, *ApJ*, 663, 81
- Pontoppidan K. M. et al., 2016, in Peck A. B., Seaman R. L., Benn C. R., eds, *Observatory Operations: Strategies, Processes, and Systems VI*, Vol. 9910, Proc. SPIE, United Kingdom, p. 991016

Powell D. M., Vegetti S., McKean J. P., Spingola C., Stacey H. R., Fassnacht C. D., 2022, *MNRAS*, 516, 1808

Price-Whelan A. M. et al., 2018, *AJ*, 156, 123

Reuter C. et al., 2020, *ApJ*, 902, 78

Rigby J. R. et al., 2024, *The Astrophysical Journal*, 978, 108

Ritondale E., Vegetti S., Despali G., Auger M., Koopmans L., McKean J., 2019, *MNRAS*, 485, 2179

Sibson R., 1981, Barnett V., ed. *Interpreting Multivariate Data*. John Wiley and Sons, New York, p. 21

Springel V. et al., 2005, *nature*, 435, 629

Springel V. et al., 2008, *MNRAS*, 391, 1685

Stacey H. R., Powell D. M., Vegetti S., McKean J. P., Fassnacht C. D., Wen D., O’Riordan C. M., 2024, *A&A*, 688, A110

Suyu S., Marshall P., Blandford R., Fassnacht C., Koopmans L., McKean J., Treu T., 2009, *ApJ*, 691, 277

Tessore N., Metcalf R. B., 2015, *A&A*, 580, A79

Tremblay B., Merritt D., 1995, *AJ*, 110, 1039

Tulin S., Yu H.-B., 2018, *Phys. Rep.*, 730, 1

Van Rossum G., Drake F. L., 2009, *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 1441412697

Vegetti S., Koopmans L. V. E., 2009, *MNRAS*, 392, 945

Vegetti S., Koopmans L., Bolton A., Treu T., Gavazzi R., 2010, *MNRAS*, 408, 1969

Vegetti S., Lagattuta D., McKean J., Auger M., Fassnacht C., Koopmans L., 2012, *nature*, 481, 341

Vegetti S. et al., 2024, *Space Sci. Rev.*, 220, 58

Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, *Phys. Rev. D*, 71, 063534

Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013, *Phys. Rev. D*, 88, 043502

Virtanen P. et al., 2020, *Nature Methods*, 17, 261

van der Walt S., Schönberger J. L., Nunez-Iglesias J., Boulogne F., Warner J. D., Yager N., Gouillart E., Yu T., 2014, *PeerJ*, 2, e453

Weiß A. et al., 2013, *ApJ*, 767, 88

Witt H. J., Mao S., 1997, *MNRAS*, 291, 211

Zavala J., Frenk C. S., 2019, *Galaxies*, 7, 81

van der Walt S., Colbert S. C., Varoquaux G., 2011, *Comput Sci Eng*, 13, 22

APPENDIX A: PHYSICAL INTERPRETATION OF MULTIPOLES

Fig. A1 shows a simple example of the effects on the lensing of a simple Sersic-light source of the first 4 multipole orders when added to a spherical power-law mass distribution with slope = 2.1. Note that we demonstrate a multipole strength of 0.1 for $m = 1$ and 0.05 for $m = 3$, $m = 4$ as the $m = 1$ multipole requires a greater strength to produce its effects (Amvrosiadis et al. 2024). Also note that in these perfect systems the multipole perturbation creates $2m$ multiple images, however, if we added noise, lens light and source complexity, some of these multiple images will blur back into the Einstein ring based on the system configuration.

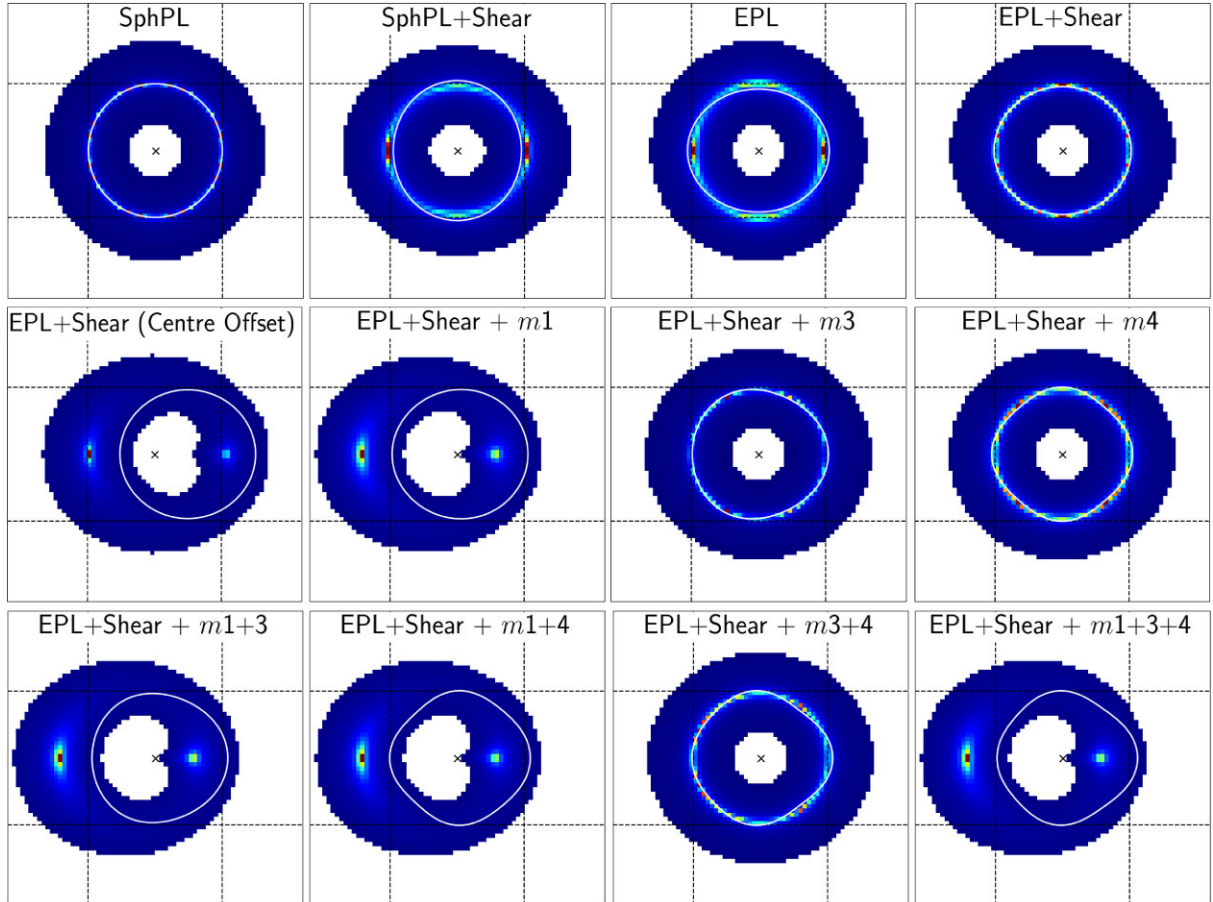


Figure A1. A simple system (spherical sersic source at $z = 4$ perfectly aligned with spherical power-law mass, slope = 2.1, $z = 0.5$), showing the effects of changes to centre ($\Delta x = +0.5$ arcsec), mass ellipticity ($q = 0.85$), addition of external shear ($k_y = 0.05$), and addition of multipole perturbations of order $m = 1, 3, 4$ with strengths $k_m = 0.05, 0.01, 0.01$. The image is a noiseless image, the white lines represent the lensing Critical Curves, and the dashed lines are drawn for reference at ± 1 arcsec, with the origin marked with the black cross. All angles in the simulation are set to zero.

At $\phi = 0$, the $m = 2$ multipole gives a similar effect to the power-law mass ellipticity and at $\phi = 90^\circ$, it is similar to the External Shear, and could explain therefore the indications in studies such as Etherington et al. (2024) that the External Shear is fitting to internal lens mass complexity. For a physical interpretation of the remaining multipoles, we visualize that the $m = 1$ multipole corresponds to a skew in the mass distribution, which could, for example, be due to the galaxy being mid- or post-merger. The $m = 3$ and $m = 4$ multipoles are more frequently used in strong lensing (e.g. Hezaveh et al. 2016; Stacey et al. 2024) and demonstrate triaxiality from $m = 3$, which has long been a component of galaxy surface brightness fits (e.g. Benacchio & Galletta 1980; Tremblay & Merritt 1995), and boxiness/disciness from the $m = 4$ (e.g. Bender et al. 1989). The majority of physical intuition on these multipoles comes from the light profiles of galaxies – see e.g. Hao et al. (2006); Padilla & Strauss (2008) for comprehensive applications of multipoles to Sloan Digital Sky Survey galaxies, or Cappellari (2016) for a review on galaxy structure using spectroscopy. Note that work is ongoing to achieve a quantitative rather than qualitative view on what multipole strengths and realizations are to be expected directly from mass modelling of strong lensing galaxies (Amvrosiadis et al. 2024; Stacey et al. 2024), but so far in the literature we see that the amplitudes of mass multipoles are in line with those from the light, but potentially mis-aligned (Stacey et al. 2024).

APPENDIX B: CALCULATION OF SUBSTRUCTURE VISIBILITY

To approximate if the theorized subhalo is detectable from its stellar emission (in the NIRCcam f444W filter, where we have the highest source signal), we first start with the best-fitting subhalo mass of $M_{200} = 10^{10.87} M_\odot$. We operate under some simplifying assumptions that there is no dust extinction and that the halo has retained its original mass (in reality dust would cause a lower SNR for the object, but if the halo is at the lens redshift, there would be tidal stripping which reduced its mass over time, meaning it could have a larger stellar component than calculated). Following the formalism of the stellar-to-halo mass relation proposed by Moster et al. (2010) in equation (B1), and taking the best-fitting parameters from table 1 of Girelli et al. (2020) at $z = 0.845$, we derive a rough estimate of $\log_{10}(M_*/M_h) \sim -2.22$ or $\log_{10} M_* \sim 8.65 M_\odot$. Taking a mass-to-light ratio of 1, this also gives a luminosity of $\log_{10}(L_*) \sim 8.65 L_\odot$.

$$\frac{M_*}{M_h}(z) = 2A(z) \left[\left(\frac{M_h}{M_{A(z)}} \right)^{-\beta(z)} + \left(\frac{M_h}{M_{A(z)}} \right)^{\gamma(z)} \right]^{-1} \quad (\text{B1})$$

We take an example SED from the SWIRE¹² template library (Polletta et al. 2007) for a 5-Gyr Elliptical Galaxy, with the SED shown in Fig. B1. This is only an approximation of the SED, but a reasonable one, particularly regarding age given the lookback time of ~ 7 Gyr. We estimate the Luminosity distance at this redshift using the COSMOTOOLS¹³ cosmological calculator, this gives $D_L = 5292$ Mpc for $H_0 = 71$, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$. Using this distance, we calculate the rest-frame wavelength for light at this redshift when observed in the centre of the f444 filter is $\lambda_{\text{emit}} = 2.41 \mu\text{m}$, and using the SED we get $L_{\lambda, \text{emit}} = 3.65 * 10^{32} \text{ergs}^{-1} \text{\AA}^{-1}$, which converts to an observed flux density at $\lambda = 4.44 \mu\text{m}$ of $F_{\lambda, \text{obs}} = 1.089 * 10^{-23} \text{ergs}^{-1} \text{cm}^{-2} \text{\AA}^{-1} =$

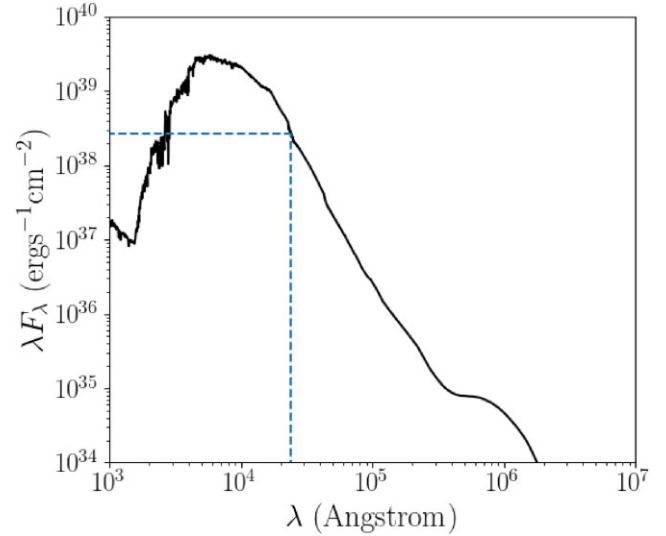


Figure B1. SED based on a 5-yr-old Elliptical Galaxy from the SWIRE library. The SED has been scaled using our calculated luminosity to represent the true non-normalized expected values of λF_λ . The blue dashed lines trace the values at our emitted light wavelength, $\lambda_{\text{emit}} = 2.41 \mu\text{m}$.

We use the JWST Exposure Time Calculator (ETC)¹⁴ (Pontopidan et al. 2016) to then calculate the upper-limit signal-to-noise ratio we would expect from an actual observation (we do not apply any extinction). We apply the same observing strategy from the *TEMPLATES* (Rigby et al. 2024) observing proposal,¹⁵ approximate the light as a point source, and use the same built-in SED of a 5-Gyr-old elliptical galaxy to get the observed filter flux. We also use the ETC ability to extract and use a noise map generated from the day of the observation (2022 September 9) and implement this to output the expected signal-to-noise ratio for this object to be $\text{SNR} \sim 0.16$.

APPENDIX C: ADDITIONAL FIT PARAMETER PLOTS

We show here first the parameters from the finalized fits in each filter for the subhalo models of SPT2147 in figure C1, with the prime M1 + M4 model highlighted. Most parameters are consistent between filters within a model, but there are some 3σ tensions, including importantly f277-f356 position angle tension in the primary M1 + M4 model. Additional tensions beyond 3σ are: f356-f444 tension in the Einstein radius for the base model; f277-f444 tension in the position angle for the base model, and f277-f356 position angle tension in M4, (M1 + M4), M3 + M4, and M1 + M3 + M4 models; f277-f356 tension in the shear angle for the M1 + M3 + M4 model.

Next, in Fig. C2 we show the change in the parameters of the fits from before adding a subhalo to after this addition, with the prime M1 + M4 model highlighted. We draw solid lines at 0 to show the lines of no change – errors are the quadrature sum of the parameter 3σ errors from the fits before and after and if errorbars cross the zero-line, the parameter change is consistent with 0. There are some non-zero-consistent parameters, but all M1 + M4 prime model parameters are consistent before and after the subhalo addition, excepting a small negative change in the f356 model for the M4 angle. Parameters with

¹²https://www.iasf-milano.inaf.it/~polletta/templates/swire_templates.html

¹³<http://www.bo.astro.it/~cappi/cosmotools>

¹⁴<https://jwst.etc.stsci.edu/>

¹⁵<https://www.stsci.edu/jwst/phase2-public/1355.pdf>

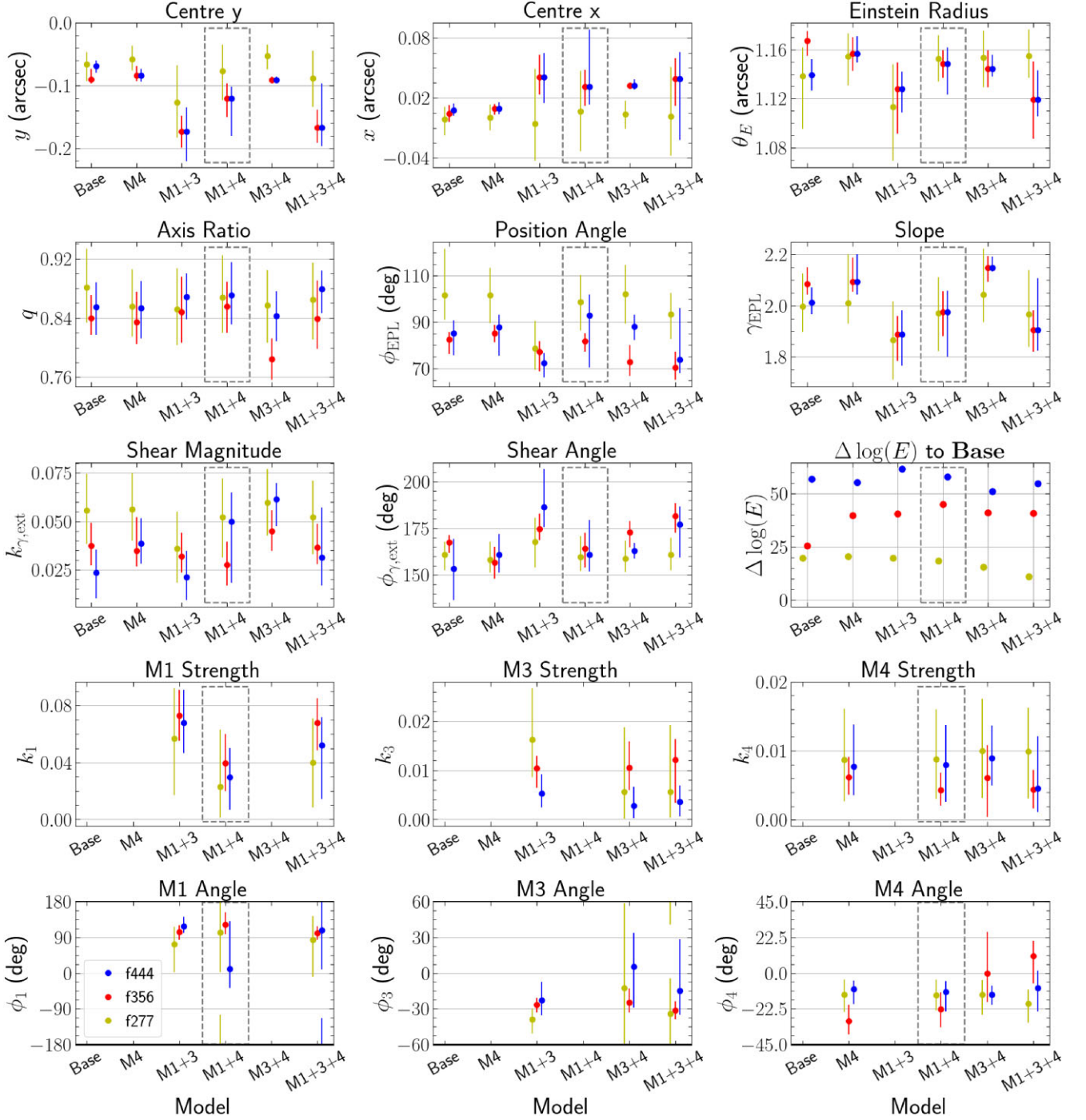


Figure C1. Reported parameters and associated errors for the fits after adding a subhalo to the highest evidence macro-models in SPT2147, as well as the base model. $\Delta \log(\epsilon)$ shown here is the total difference to the base no-subhalo model. Multipole angles have a symmetry through $360/n$ – i.e. a 90 deg error on the M4 multipole means it is unconstrained as this multipole has rotational symmetry through 90 deg, and so the errors are wrapped around $\pm(360/n)/2$. Evidence changes shown are in comparison to the base EPL + shear model. We highlight with the grey box our preferred final model of EPL + shear + M1 + M4.

changes outside of zero are: f444 base and f356 M3+M4 positive change of centre x position; f277, f356, f444 base, f356 M4, f356 M3 + M4, f444 M1 + M3 + M4 decrease in Einstein radius; f277 base increase and f356 M3 + M4 decrease in position angle; f444 base and f444 M1 + M3 + M4 decrease in power-law slope; f444 base decrease in shear magnitude; f356 M3 + M4 increase in shear angle; f444 M1 + M3 decrease in M1 strength, f356 M3 + M4 increase in M3 strength; f356 M4 and f356 M1 + M4 decrease in M4 angle.

We also show in Table C1 specific values for the $\Delta \log(\epsilon)$ evidence changes, maximal normalized residuals from the fits, and multipole strengths, for all the models fit both with and without the subhalo.

We finally show in Fig. C3 the parameters for the fits to SPT0418. We recall that no multipole models provide a high enough evidence increase to be worth using, so the prime model here is the base model. The position angles for the f356 fits are unconstrained (errors span the whole range), due to the high axis ratio – $q = 1$

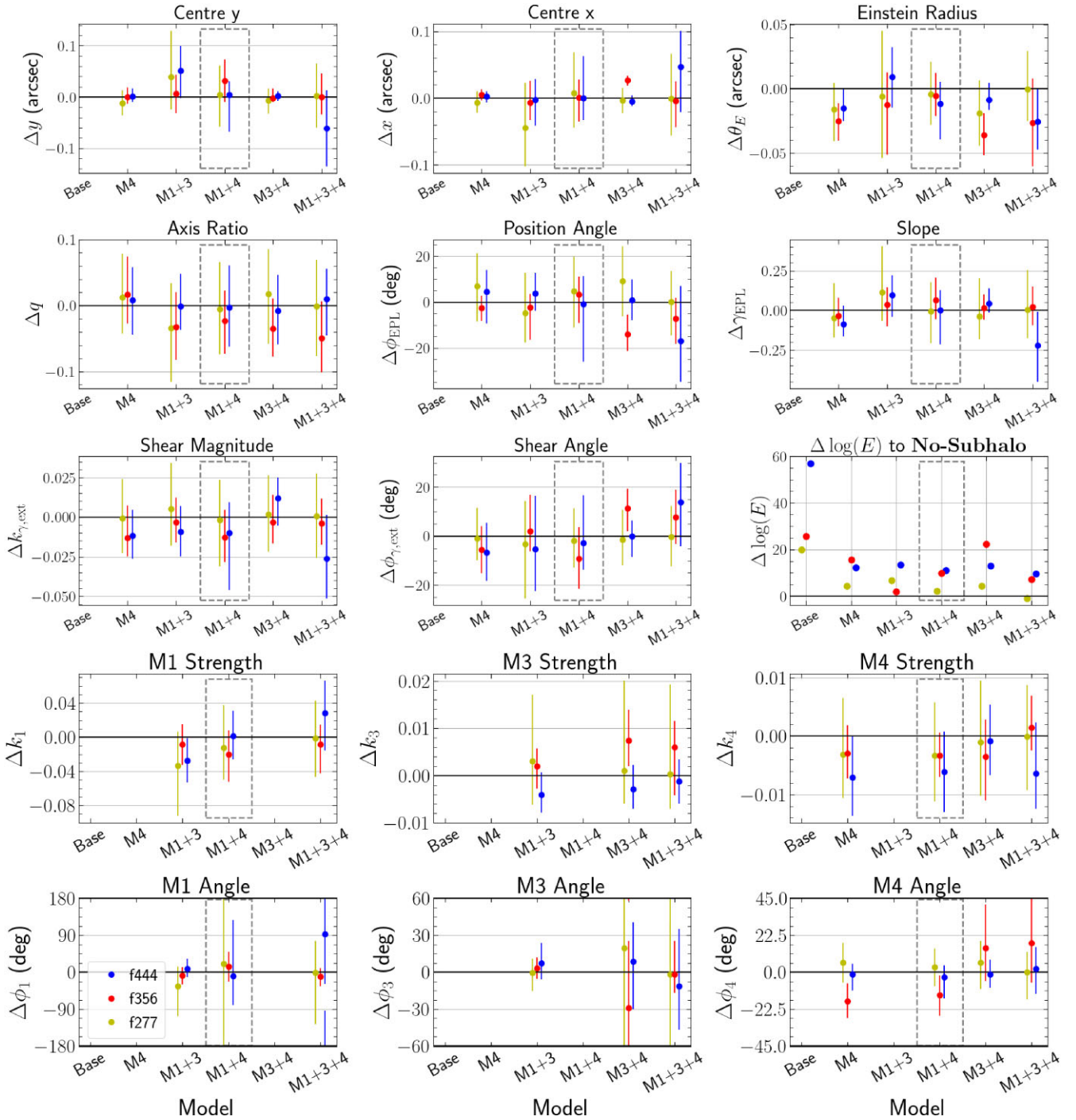


Figure C2. Changes in the mass macro-model parameters for SPT2147 once we add a substructure, as well as the associated error on the change. Solid lines are drawn at a change of 0. Multipole angles have a symmetry through $360/n$ – i.e. a 90 degree error on the M4 multipole means it is unconstrained as this multipole has rotational symmetry through 90 deg, and so the errors are wrapped around $\pm(360/n)/2$. Evidence changes shown are the total difference to the base no-subhalo model. We highlight with the grey box our preferred final model of EPL + shear + M1 + M4.

is spherical so the position angle has no physical meaning at this point. Shear angles are generally inconsistent between the f356 and f444 filters within a model, with f444 being higher. All other parameters are fully consistent between all models and filters, which

is expected when all models have a similar Bayesian evidence. All multipole strengths reach down very close to zero with their errorbars, again showing that the multipole models are not required for this lens.

Table C1. Evidence boosts, maximum normalized residuals, and multipole strengths for SPT2147 over all models both with and without the addition of a subhalo. All multipole ‘pre-sub’ evidences are compared to the base ‘pre-sub’ evidence. All ‘post-sub’ evidences are compared to the corresponding model without a subhalo.

| Filter | Model | PRE-sub | PRE-sub | PRE-sub | | | WITH-sub | WITH-sub | WITH-sub | | |
|--------|---------------|--------------------------------------|---|--|------|------|--|---|--|------|------|
| | | $\Delta \log(\epsilon)$ (to base) | Max normalized residual (σ) | Multipole strengths ($k_m * 10^{-2}$) | | | $\Delta \log(\epsilon)$ (to previous) | Max normalized residual (σ) | Multipole strengths ($k_m * 10^{-2}$) | | |
| | | | | m1 | m3 | m4 | | | m1 | m3 | m4 |
| f444W | Base | – | 3.58 | – | – | – | 56.90 | 3.59 | – | – | – |
| f444W | M1 | 2.36 | 3.67 | 4.98 | – | – | 55.83 | 3.71 | 3.94 | – | – |
| f444W | M3 | 21.81 | 3.63 | – | 0.94 | – | 28.98 | 3.60 | – | 0.29 | – |
| f444W | M4 | 43.13 | 3.66 | – | – | 1.48 | 12.15 | 3.59 | – | – | 1.11 |
| f444W | M1 + 3 | 48.16 | 3.73 | 9.52 | 0.94 | – | 13.36 | 3.74 | 6.79 | 0.54 | – |
| f444W | M1 + 4 | 46.99 | 3.59 | 2.81 | – | 1.41 | 10.97 | 3.50 | 2.96 | – | 0.80 |
| f444W | M3 + 4 | 38.01 | 3.58 | – | 0.57 | 0.98 | 13.03 | 3.54 | – | 0.29 | 0.90 |
| f444W | M1 + 3 + 4 | 45.19 | 3.59 | 2.43 | 0.49 | 1.10 | 9.68 | 3.75 | 5.23 | 0.36 | 0.46 |
| f356W | Base | – | 3.83 | – | – | – | 25.70 | 3.77 | – | – | – |
| f356W | M1 | 10.66 | 3.94 | 5.99 | – | – | 28.77 | 4.05 | 4.87 | – | – |
| f356W | M3 | 4.79 | 3.80 | – | 0.67 | – | 40.99 | 3.87 | – | 1.34 | – |
| f356W | M4 | 24.31 | 3.88 | – | – | 0.92 | 15.57 | 3.76 | – | – | 0.62 |
| f356W | M1 + 3 | 38.69 | 4.04 | 8.11 | 0.84 | – | 2.04 | 3.99 | 7.30 | 1.04 | – |
| f356W | M1 + 4 | 35.04 | 4.08 | 5.97 | – | 0.76 | 9.96 | 4.06 | 3.97 | – | 0.43 |
| f356W | M3 + 4 | 18.61 | 3.85 | – | 0.31 | 0.96 | 22.40 | 3.87 | – | 1.05 | 0.61 |
| f356W | M1 + 3 + 4 | 33.54 | 4.08 | 7.62 | 0.62 | 0.30 | 7.31 | 4.03 | 6.78 | 1.21 | 0.44 |
| f277W | Base | – | 3.47 | – | – | – | 19.84 | 3.49 | – | – | – |
| f277W | M1 | 4.10 | 3.48 | 5.30 | – | – | 13.82 | 3.49 | 3.51 | – | – |
| f277W | M3 | 5.68 | 3.44 | – | 1.72 | – | 10.58 | 3.47 | – | 1.37 | – |
| f277W | M4 | 16.29 | 3.72 | – | – | 1.18 | 4.25 | 3.46 | – | – | 0.87 |
| f277W | M1 + 3 | 12.99 | 3.45 | 9.00 | 1.33 | – | 6.78 | 3.48 | 5.67 | 1.63 | – |
| f277W | M1 + 4 | 16.39 | 3.56 | 3.58 | – | 1.20 | 2.07 | 3.47 | 2.31 | – | 0.88 |
| f277W | M3 + 4 | 11.23 | 3.51 | – | 0.46 | 1.10 | 4.35 | 3.72 | – | 0.57 | 1.00 |
| f277W | M1 + 3 + 4 | 12.05 | 3.55 | 4.12 | 0.54 | 1.00 | –0.99 | 3.49 | 3.99 | 0.57 | 0.99 |

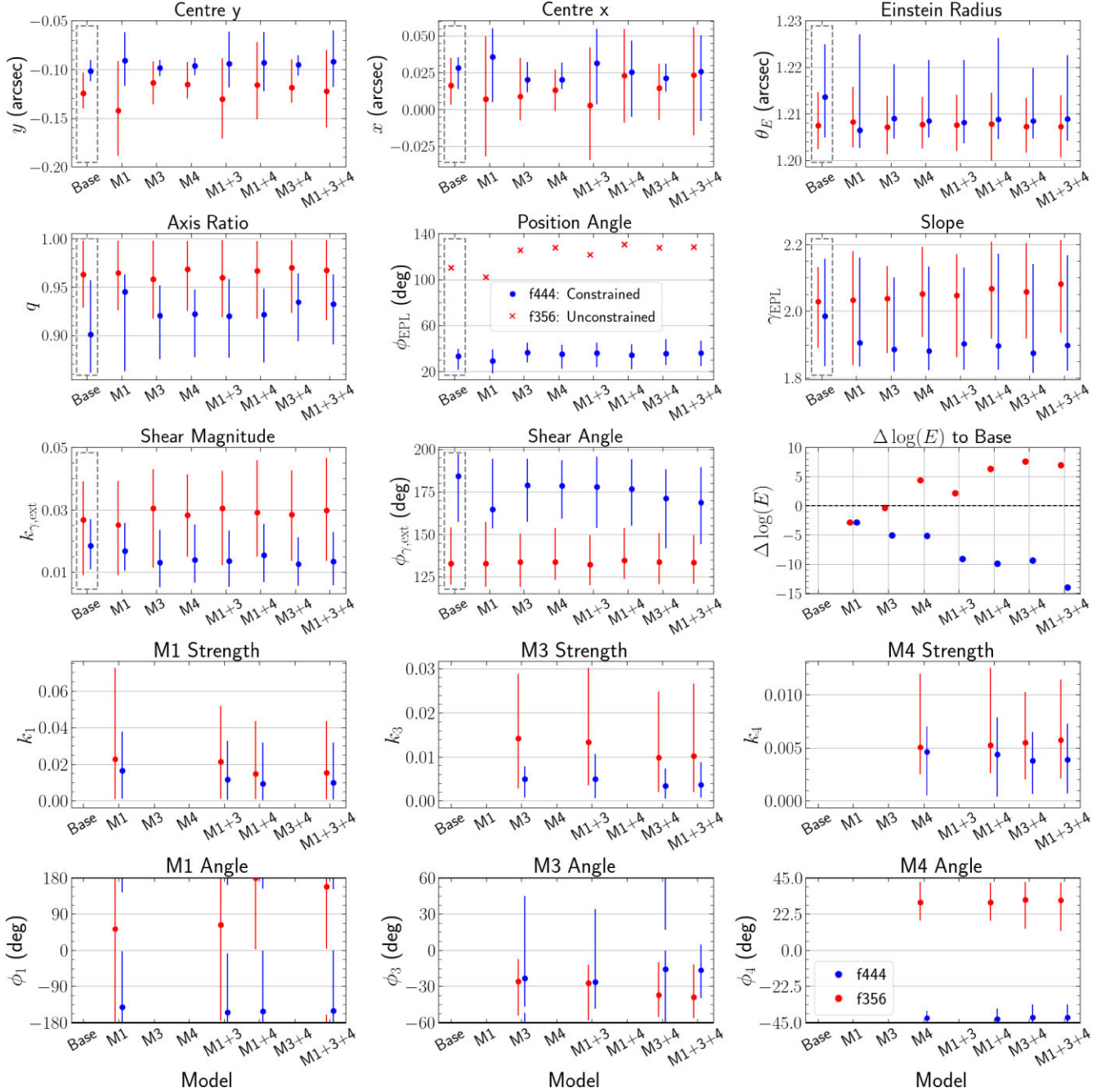


Figure C3. Reported parameters and associated errors for all fits (without a subhalo) to SPT0418. Multipole angles have a symmetry through $360/n$ – i.e. a 90 deg error on the M4 multipole means it is unconstrained as this multipole has rotational symmetry through 90 deg, and so the errors are wrapped around $\pm(360/n)/2$. Evidence changes shown are in comparison to the base EPL + shear model. We highlight with the grey box our preferred final model (base). The position angles in f356 show only the median value as the errors span the full parameter space leaving this angle unconstrained – this is due to the highly spherical axis ratio.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.