Probing Dark Matter with Strong Gravitational Lensing through an Effective Density Slope

Atınç Çağan Şengül1★ and Cora Dvorkin1†

1 Harvard University, Department of Physics, Cambridge, Massachusetts, 02138, U.S.A.

9 August 2022

ABSTRACT

Many dark matter (DM) models that are consistent with current cosmological data show differences in the predicted (sub)halo mass function, especially at sub-galactic scales, where observations are challenging due to the inefficiency of star formation. Strong gravitational lensing has been shown to be a useful tool for detecting dark low-mass (sub)halos through perturbations in lensing arcs, therefore allowing the testing of different DM scenarios. However, measuring the total mass of a perturber from strong lensing data is challenging. Over or underestimating perturber masses can lead to incorrect inferences about the nature of DM. In this paper, we argue that inferring an effective slope of the dark matter density profile, which is the power-law slope of perturbers at intermediate radii, where we expect the perturber to have the largest observable effect, is a promising way to circumvent these challenges. Using N-body simulations, we show that (sub)halo populations under different DM scenarios differ in their effective density slope distributions. Using realistic mocks of Hubble Space Telescope observations of strong lensing images, we show that the effective density slope of perturbers can be robustly measured with high enough accuracy to discern between different models. We also present our measurement of the effective density slope $\gamma = 1.96^{+0.12}_{-0.12}$ for the perturber in JVAS B1938+666, which is a $2\sigma$ outlier of the cold dark matter scenario. More measurements of this kind are needed to draw robust conclusions about the nature of dark matter.

Key words: gravitational lensing: strong, cosmology: dark matter

1 INTRODUCTION

Cosmological observations and galaxy dynamics have shown us that about 84% of all matter in the universe is composed of dark matter (DM), whose presence can be inferred via its gravitational influence. However, the Standard Model of particle physics lacks a description of it. The nature and interactions of DM remain one of the greatest puzzles of fundamental physics in modern times.

The standard Cold Dark Matter (CDM) paradigm predicts that dark matter is clustered into gravitationally bound systems called halos. These halos are expected to roughly follow a power-law spectrum from galaxy clusters to ultra-faint dwarf galaxies (Davis et al. 1985; Navarro et al. 1996b, 1997; Moore et al. 1999; Green et al. 2004; Wang et al. 2020). While the CDM model provides a good description of the universe at large cosmological scales (Planck Collaboration et al. 2020; Gil-Marín et al. 2016; Abbott et al. 2018; Troxel et al. 2018; Abolfathi et al. 2018; Springel et al. 2006), smaller scales remain largely untested. Moreover, many theoretical models describing the microphysics of dark matter disagree with CDM at small scales (Bode et al. 2001; Kaplinghat 2005; Anderhalden et al. 2013; Lovell et al. 2014; Bullock & Boylan-Kolchin 2017; Buckley & Peter 2018; Tulin & Yu 2018).

Strong gravitational lensing is a promising observable that can be used as a probe of dark matter fluctuations at all scales, in particular those below that of the typical dwarf galaxies. Furthermore, in the coming years, we expect to have an increase in observed lensed systems, with tens of thousands expected to be discovered with upcoming optical imaging surveys (Collett 2015; Bechtol et al. 2019; Jacobs et al. 2019; Huang et al. 2021).

The (sub)halo mass function, which is the number of dark matter (sub)halos as a function of mass, is commonly used to probe dark matter scenarios, in particular its low-mass end, where many theoretical DM models predict a suppression in the number of expected subhalos or halos along the line of sight (Angulo et al. 2013; Schneider 2015). So far, there have been a handful of claimed detections of low-mass substructure in strong lens systems (Vegetti et al. 2010, 2012; Hezaveh et al. 2016b) through perturbations in lensed arcs, as well as line-of-sight halos (Şengül et al. 2021). The lack of detection of (sub)halos together with the few detections have been used to constrain the properties of DM (Vegetti et al. 2014; Ritondale et al. 2019). At smaller physical scales where direct detection is not possible, the power spectrum of the collective perturbations of low-mass perturbers in the convergence field, either from subhalos (Hezaveh et al. 2016a; Cyr-Racine et al. 2016b; Diaz Rivero et al. 2018a; Cyr-Racine et al. 2019; Díaz Rivero et al. 2018b; Brennan et al. 2019) or line-of-sight halos (Şengül et al. 2020), has been proposed as a powerful probe of CDM. The magnification ratio of multiple images has also been proposed as a probe of small-scale structure in the lens (Mao & Schneider 1998; Metcalf & Madau 2001) and has been used

★ sengul@g.harvard.edu
† cdvorkin@g.harvard.edu

© 2022 The Authors

arXiv:2206.10635v2 [astro-ph.CO] 8 Aug 2022
to detect subhalos (Dalal & Kochanek 2002; Moustakas & Metcalf 2003; Amara et al. 2006; Fadely & Keeton 2011; Xu et al. 2012; MacLeod et al. 2013; Nierenberg et al. 2014; Nierenberg et al. 2017) and line-of-sight halos (Gilman et al. 2019). Additionally, the collective effects of the large population of low-mass (sub)halos have been harnessed by machine learning techniques that attempt to extract information that may be buried in the observations (Breuer et al. 2016; Daylan et al. 2018; Brehmer et al. 2019; Ostdiek et al. 2022h;a; Wagner-Carena et al. 2022) as well as other types of techniques (Birrer et al. 2017).

When working with gravitational lensing as a probe, the total extended mass of a perturber is not an optimal parameter, since it relies on assumptions about the extended mass distribution (Vegetti & Vogelsberger 2014; Vegetti et al. 2014; Minor et al. 2017). Unusually dense perturbers will have a high lensing effect (despite their possible low mass), and vice versa. Hence, predicting the total mass from strong lensing measurements can be misleading. The substructure detected in the SDSSJ0946+1006 lens system (Vegetti et al. 2010) is a good example of these challenges. Its total mass ($3.5 \times 10^9 M_\odot$) was inferred by assuming a tidal truncation due to the host halo. Considering only its mass, this detection is consistent with predictions of CDM. However, a recent analysis (Minor et al. 2021b) of the subhalo has shown that this substructure is inconsistent with CDM with an unusually steep density slope. A robust way of measuring perturber masses has been proposed in Refs. Minor et al. (2017); Minor et al. (2021a) where they define a “perturbation radius”, which is the distance from the perturber to the critical curve in the direction in which the perturbation to the magnification is largest. The effective perturber lensing mass (this is, the projected mass within the perturbation radius, divided by the log-slope of the main lens’ 2D density profile) has been shown to be a robust estimator of the mass of the perturber under changes in the density profile (Minor et al. 2017). In addition to measuring the perturber mass, we can probe the density of the perturber as it falls with increasing radius. More specifically, the effect of perturbers in strong lensing systems is quite sensitive to the slope of the average convergence of the perturber at a characteristic scale, where we expect the effect of the perturber to have the largest observable effect. We call this the effective density slope. This effective slope is expected to vary under different dark matter scenarios, so a measurement of the effective slope would provide fertile grounds for discerning between different dark matter models.

Measuring the effective density slope may also shed light on a discrepancy claimed between the observations of dwarf galaxies and CDM simulations, known as the “cusp-core” problem (Moore 1994; Flores & Primack 1994; de Blok et al. 2001; Gentile et al. 2004; de Blok 2010). The cusp-core problem refers to the fact that the observed central regions of galaxies (inferred from rotation curves) tend to have a lower amplitude and shallower slope of the density profile than the ones predicted from CDM simulations. When the effects of baryons are neglected, the CDM halos show a self-similarity over a large mass range in simulations. These halos can be approximated by the Navarro-Frenk-White (NFW) profile (Navarro et al. 1996, 1997), with a “cusp” at its center, where the density ($\rho$) diverges with the distance ($r$) to the center as $\rho \propto r^{-1}$. In contrast, the rotational speed of the stars in dwarf galaxies is observed to be roughly linear with radius, which implies a shallow inner slope, a “core”. The origin of this problem is yet unclear, and a number of solutions have been proposed to this problem, ranging from new physics in the form of warm dark matter (WDM) (Bode et al. 2001; Lovell et al. 2012), self-interacting dark matter (SIDM) (Spergel & Steinhardt 2000; Vogelsberger et al. 2012; Rocha et al. 2013; Elbert et al. 2015), ultra-light dark matter (Hu et al. 2000) or fluid dark matter (Peebles 2000) to the effect of baryons (Navarro et al. 1996a; Gnedin & Zhao 2002; Read & Gilmore 2005; Mashchenko et al. 2006; Governato et al. 2012; Brook et al. 2012; Pontzen & Governato 2012; Chan et al. 2015). A measurement of the effective density slope through gravitational lensing in dwarf galaxies should provide a powerful probe of CDM, since these systems are largely devoid of stars (Fitts et al. 2017; Read et al. 2017).

An observed diversity in rotation curves of dwarf galaxies has also been claimed to be poorly reproduced in simulations (Oman et al. 2015). This could be due to a limitation of numerical simulations in accounting for supernovae feedback or star formation mechanisms, or it could also be due to non-circular motions in the gas not accounted for (Oman et al. 2018; Marasco et al. 2018). Ref. Oman et al. (2015) shows that systems for which the rotation curves do not agree with simulations also present a deficit in the enclosed mass compared to CDM simulations. They also show that for many of these systems the inner slopes are similar while the rotation curves are very different. Different solutions in the form of dark matter self interactions and baryonic flows have been suggested to solve this problem (Kamada et al. 2017; Kaplinghat et al. 2016, 2020; Santos-Santos et al. 2020). However, these solutions each have shortcomings, hence it remains an open question.

In this work, we suggest the use of gravitational lensing to measure the power-law slope of the average convergence in an intermediate regime to discern between possible dark matter scenarios. This could help shed light on the problems mentioned above and give us new insights into the nature of dark matter. We will restrict our analysis to low-mass systems ($< 10^8 M_\odot$), where the effect of baryons is expected to be small (Fry et al. 2015; Pereira Wilson et al. 2022).

This paper is organized as follows. In §2 we discuss how we determine the characteristic scale at which the power-law slope of a perturber is measured with lensing. In §3 we analyze the density profiles obtained from N-body simulations of different DM scenarios and extract their effective power-law slopes. We test the accuracy and robustness of the measurements of the effective power-law slopes of perturbers using realistic mocks of strong lensing images in §4. In §5 we present a measurement of the effective slope of the density profile of a perturber previously found in the strong lens system JVAS B1938+666. In §6 we summarize the results and the implications of our analysis. We include more details of our analysis in Appendices A, B, and C.

2 REGION OF MAXIMUM OBSERVABILITY

In general, our ability to probe the density profile of a perturber at very small radii is limited by resolution and noise. On the other hand, at much larger radii, the effect of the perturber is weaker and can be absorbed into the smooth model of the lens. Between these inner and outer radii, there is a region where we are most sensitive to the changes in the average convergence, which we will call the region of maximum observability (RMO). The effect on a lensing image of two perturbers with different density profiles is indistinguishable if they have overlapping average convergences within the RMO.

When modeling a gravitationally lensed image with a perturber, one has to simultaneously model the main lens, the external shear, and the source light. In general, it is not possible to measure the average convergence of a perturber without assuming it has some functional form. Only then, one can measure the parameters of that functional form. Determining the sensitivity of the images to the average convergence at various radii becomes tractable if we consider the much simpler problem of measuring a perturber from a noisy
image when we know the main lens mass distribution, the perturber position, and the source-light distribution. Let \( I_\ell (\vec{y}) \), where \( \vec{y} \) is a 2D vector representing the angular position on the lens plane. All the uncertainties we derive will therefore be lower bounds on the true uncertainties from modeling where all the lens and source parameters are varied.

Our goal is to write a function that maps the average convergence of a perturber at various radii to the pixel brightness changes that it causes. We start by calling the surface brightness distribution on the image plane \( I(\vec{r}) \), where \( \vec{r} \) is a 2D vector representing an angular position on the image plane (we choose the center of the perturber to be \( \vec{r} = 0 \) without loss of generality). The unlensed and lensed brightness distributions are related by \( I(\vec{r}) \approx I_\ell (\vec{y}) \), where the source plane position is given by the lens equation \( \vec{y} = \vec{r} - \delta \vec{y} \), with \( \alpha(\vec{r}) \) being the angular deflection due to the lens. In gravitationally lensed systems, \( I(\vec{r}) \) is not directly measured. Instead, pixel values \( p_{i,j} \) (with some noise which we will not explicitly write for clarity) of an \( N_x \times N_y \) grid with \( i \in \{1, 2, ..., N_x \} \) and \( j \in \{1, 2, ..., N_y \} \) are measured. Each pixel value is given by

\[
p_{i,j} = \int_{i,j} d^2 \vec{x} I(\vec{x}),
\]

where \( \int_{i,j} \) represents the integral over the the area of the pixel \( (i, j) \) and \( I \) is the surface brightness convolved with the point spread function (PSF) \( \delta \):

\[
I(\vec{x}) = \int_{\mathbb{R}^2} d^2 \vec{u} I(\vec{u}) \delta(\vec{x} - \vec{u})
\]

\[
= \int_{\mathbb{R}^2} d^2 \vec{u} I_\ell (\vec{y}) \delta(\vec{x} - \vec{u}),
\]

We can substitute \( I_\ell (\vec{y}) \) by \( I(\vec{u}) \) in Eq. (2) to obtain Eq. (3), where \( \vec{y} = \vec{u} - \delta \vec{y} \).

In general, the maximum angular deflection due to typical perturbers considered in galaxy-galaxy lensing studies is much smaller than the average pixel size. For example, the angular deflection of an NFW perturber with mass \( M_{200} = 10^9 M_0 \) \( (M_{200} \approx 200 \text{ the critical density of the universe), concentration \( c_{200} = 12, \text{ lens redshift } z_L = 0.5 \text{ and source redshift } z_s = 1.5 \text{ is } |\alpha| \approx 0.003 \text{"}}.\) With that in mind, we split the angular deflection into two terms: \( \delta \vec{y} = \delta \vec{y}_m + \delta \vec{y}_p \), where \( \delta \vec{y}_m \) is the angular deflection due to the main lens and \( \delta \vec{y}_p \) is the angular deflection due to the perturber. Since \( |\alpha| \ll |\alpha_m| \), we expand \( I_\ell (\vec{y}) \) around \( \vec{y}_m = \vec{u} - \delta \vec{y}_m \), the source plane position that gets lensed to \( \vec{u} \) by the main lens, and only keep the terms up to linear order in \( \delta \vec{y}_p \):

\[
I(\vec{x}) = \int_{\mathbb{R}^2} d^2 \vec{u} I_\ell (\vec{v}_m) \cdot [\nabla I_\ell (\vec{v}_m)] \delta(\vec{x} - \vec{u}).
\]

Integrating the first term gives the unperturbed brightness, which we denote \( I_0(\vec{x}) \). We can write the effect of the perturber as differences of pixel values \( \Delta p_{i,j} \equiv p_{i,j} - p_{i,j}^0 \), where \( p_{i,j}^0 \) are the pixel values that we would observe if the perturber did not exist. These differences can be obtained by substituting the second term of Eq. (4) into Eq. (1),

\[
\Delta p_{i,j} = \int_{i,j} d^2 \vec{x} \int_{\mathbb{R}^2} d^2 \vec{u} [\nabla p_{i,j}(\vec{u}) \cdot \nabla I_\ell (\vec{v}_m)] \delta(\vec{x} - \vec{u}).
\]

We see that pixel value differences caused by a perturber are larger if the source has a large directional derivative along the direction of the angular deflection caused by the perturber. We can approximate the integral over the pixel area in Eq. (5) as the integrand evaluated at the center of the pixel multiplied by the pixel area. We can also approximate the convolution integral as a discrete sum, which gives

\[
\Delta p_{i,j} \approx \Delta x^2 \sum_{k,l} [-\bar{\alpha}_p(\vec{x}_{k,l}) \cdot \nabla I_\ell (\vec{y}_{m,k,l})] \delta_{i-k,j-l},
\]

where \( \vec{x}_{k,l} \) is the center of the pixel \( (k,l) \), \( \vec{y}_{m,k,l} = \vec{x}_{k,l} - \vec{\alpha}_m(\vec{x}_{k,l}) \) is the position on the source plane that maps onto \( \vec{x}_{k,l} \) after being lensed by only the main lens, and \( \delta_{i-k,j-l} \) is the discretized PSF.

For an axially symmetric perturber, the angular deflection is given by \( \bar{\alpha}_p(\vec{x}) = k(\vec{x}) \bar{\alpha}(\vec{x}) \), where \( k(\vec{x}) \) is the average convergence of the perturber within the angular radius \( |\vec{x}| \). We finally obtain a linear mapping from the average convergence values of the perturber at various radii to the pixel differences it causes:

\[
\Delta p_{i,j} \approx \Delta x^2 \sum_{k,l} [-\bar{\alpha}_p(\vec{x}_{k,l}) \cdot \nabla I_\ell (\vec{y}_{m,k,l})] \delta_{i-k,j-l} k(\vec{x}_{k,l}).
\]

Let \( K \) be the vector of model parameters where each element \( K_i \) is the average convergence at some radii \( r_i \). Let \( D \) be the data vector where each element is a pixel difference \( \Delta p_{i,j} \) of some pixel \((i,j)\). We now have a linear map \( L \) that takes a set of average convergences to their pixel differences, \( D = L K \). The elements of the matrix \( L \) can be obtained from the coefficients that multiply \( k(\vec{x}_{i,j}) \) in Eq. (7). Assuming Gaussian pixel errors given by a covariance matrix \( C \), we obtain uncertainties for each average convergence element, \( K_i \), as

\[
\sigma_i = \sqrt{A^{-1}}_{i,i}, \text{ where } A = L^T C^{-1} L.
\]

We will refer to these uncertainties as linear errors, and they can be used as lower bounds on the true uncertainties. We define the RMO as the region where the relative linear errors in average convergence are less than twice the lowest relative error. We will show in §4 that with this definition of the RMO we can consistently predict the effective power-law slope of a given perturber profile. We see from Eq. 8 that the linear errors, and consequently the RMO, is a function of the main lens, the perturber position, and the source-light distribution. Sharper source-light distributions with steeper gradients and higher image resolution broaden the RMO to include the more inner regions of the perturber profile.

In this section, we will analyze the density profile of high-resolution N-body simulations under different dark matter microphysics. The simulations used are the ETHOS (Effective Theory of Structure Formation) suite, originally presented in Refs. Vogelsberger et al. (2016);...
Cyr-Racine et al. (2016a). In particular, five different dark matter scenarios are analyzed: a cold dark matter scenario and four other non-standard scenarios that include dark matter-dark radiation interactions, that produce a cutoff in the primordial power spectrum, and dark matter self-interactions, that affect the non-linear structure formation (under the label of ETHOS1-4). The parameters used in each simulation are described in Table 1. In particular, the parameters of the ETHOS4 simulation have been found to give a reasonable fit to observations of the Milky Way satellite population (Vogelsberger et al. 2016).

For each dark matter scenario, we will extract the radial profiles of the perturbers in the simulation at different redshifts. In Fig. 2, we show an example of the normalized density profile of a CDM subhalo at redshift $z = 1.47$. We then calculate the average convergence of each profile as a function of radii, which is what the angular deflections of a perturber depend on. The power-law slope of the average convergence profile in the RMO (shown as the blue-shaded region in Fig. 2) is what we will call the effective power-law slope for each perturber. We should note that the RMO will be different for each perturber detected via strong lensing, depending on the resolution, main lens, source light, and the perturber position. Therefore, the RMO calculation needs to be done on a case-by-case basis. For the analysis in this section, we will assume that the main lens is at $z = 0.88$ and the source is at $z = 2.079$, as in the JVAS B1938+666 system. The rest of the properties of the main lens and the source are given by the best-fit parameters of the JVAS B1938+666 lens system, as reported in Ref. Şengül et al. (2021).

We sort the particles of each perturber into radial bins and fit a power-law density profile ($\rho \propto r^{-\gamma}$) around the perturber radius to obtain the effective power-law slope, $\gamma$, for each perturber. We assume that the number of particles in each radial bin is a Poissonian realization of an underlying power-law density profile. The details of this analysis can be found in Appendix A. We only considered perturbers that are more massive than $5 \times 10^7 M_\odot$ to ensure that the radial bins are sufficiently populated to have an accurate effective power-law slope estimate. The particle mass of the simulations we use is $2.8 \times 10^4 M_\odot$, which ensures that the smallest perturbers consist of more than $10^5$ particles. We observe significant differences between the distribution of effective slopes in different DM scenarios. In Fig. 3 we show the effective density slope distributions of the different DM models at redshift $z = 1.47$. The CDM population has a mean of $\gamma_{\text{CDM}} = 1.64$ with a standard deviation of $\sigma(\gamma_{\text{CDM}}) = 0.07$. In the ETHOS1-4 models, the dark matter interactions with dark radiation and with itself result in a shallower profile. The ETHOS4 distribution has the largest overlap with CDM with a mean of $\gamma_{\text{ETHOS4}} = 1.51$ and a standard deviation of $\sigma(\gamma_{\text{ETHOS4}}) = 0.06$.

We find a redshift dependence in the effective slope distributions of the different DM models, with a shift towards lower values of the slope as we increase the redshift. In Fig. 4, we show the distribution of $\gamma$ for CDM, ETHOS3, and ETHOS4, at two redshifts. The mean $\gamma$ of the CDM population shifts by $0.4\sigma$ (from $\gamma = 1.67$ to $\gamma = 1.64$ when going from $z = 0.88$ to $z = 1.47$, while the mean of the ETHOS4 distribution shifts by $0.6\sigma$ (from $\gamma = 1.55$ to $\gamma = 1.51$). The redshift dependence for ETHOS3 is stronger, which is expected since it has a higher DM self-interaction cross section. The mean $\gamma$ in the ETHOS3 model, goes from $\gamma = 1.29$ at $z = 0.88$ to $\gamma = 1.23$ at $z = 1.47$, shifting by $1\sigma$. This is likely due to the effect of tidal stripping (see the discussion below). This effect needs to be accounted for when using the effective density slope to probe different dark matter scenarios.

In Fig. 5 we show the distribution of the masses and the effec-
Probing Dark Matter with Strong Gravitational Lensing through an Effective Density Slope

Table 1. Relevant parameters of the ETHOS (Effective Theory of Structure Formation) models considered in this work. Here $\alpha_\chi$ is the coupling between the mediator and dark matter, $\alpha_\nu$ is the coupling between the mediator and a massless neutrinos-like fermion, $m_\phi$ is the mass of the mediator and $m_\chi$ is the mass of the DM particle. The parameter $a_4$ is the coefficient of the power-law expansion in redshift of the DM drag opacity caused by the DM-DR interaction, which determines the linear initial matter power spectrum (see Ref. Cyr-Racine et al. (2016a)). The parameters $h\sigma_{i30}$, $h\sigma_{i200}$ and $h\sigma_{i1000}$ are the DM self-interaction cross-sections at different velocities: 30, 200 and 1000 km/s and describe the velocity-dependent cross sections. We refer the reader to Ref. Vogelsberger et al. (2016) for more details on these models.

| Name   | $\alpha_\chi$ | $\alpha_\nu$ | $m_\phi$ [GeV/c²] | $m_\chi$ [GeV/c²] | $a_4$ [h/Mpc] | $(\sigma)_{30}/m_\chi$ [cm²/g] | $(\sigma)_{200}/m_\chi$ [cm²/g] | $(\sigma)_{1000}/m_\chi$ [cm²/g] |
|--------|---------------|---------------|-------------------|-------------------|--------------|-------------------------------|-------------------------------|-------------------------------|
| ETHOS1 | 0.071         | 0.123         | 0.723             | 2000              | 14095.65     | 4.98                          | 0.072                         | 0.0030                        |
| ETHOS2 | 0.016         | 0.03          | 0.83              | 500               | 1784.05      | 9.0                           | 0.197                         | 0.00097                       |
| ETHOS3 | 0.006         | 0.018         | 1.15              | 178               | 305.94       | 16.9                          | 0.48                          | 0.0028                        |
| ETHOS4 | 0.5           | 1.5           | 5.0               | 3700              | 286.09       | 0.16                          | 0.022                         | 0.00075                       |

Figure 3. Normalized distributions of the effective density profile slopes ($\gamma$) of (sub)halos with masses $5 \times 10^8 M_\odot < M < 2 \times 10^9 M_\odot$ for all 5 DM scenarios (CDM, ETHOS1-4) analyzed in this work, at $z = 1.47$. The y-axis shows the fractional abundance of (sub)halos. We see that the CDM population has the steepest mean effective slope.

Figure 4. Redshift evolution of the normalized distributions of the effective density slopes ($\gamma$) of (sub)halos with masses $5 \times 10^8 M_\odot < M < 2 \times 10^9 M_\odot$ for CDM, ETHOS3, and ETHOS4. We see that for all scenarios the slope distributions shift towards lower values for higher redshifts.

Figure 5. Distribution of mass vs. the effective density profile slope of CDM (in red) and ETHOS4 (in blue) perturbers at redshift $z = 1.47$. Dashed contours enclose 68% (darker color) and 95% (lighter color) (which correspond to 1$\sigma$ and 2$\sigma$) of the perturbers around the mean, respectively.

We further explore the distribution of the slopes and distances to the host center for the subhalos of the most massive halo in each set of simulations (in Fig. 6) and find a weak correlation between the two, most likely caused by the tidal forces of the host galaxy (Syer & White 1998). We find for all the models studied that the subhalos that are closer to the center of their host have lower slopes on average. This could be understood as follows: Tidal stripping can cause subhalos to lose over 90% of their total mass. During this mass loss, the shape of the inner profile of the subhalo is preserved while the mass in the outer radii gets tidally stripped (Penarrubbia et al. 2010; Ogiya et al. 2019). A subhalo that ends up with a certain mass $M$ after tidal stripping will have a larger scale radius compared to a subhalo.
which gives a field of view of with a full width at half maximum (FWHM) of 10 HST orbits (33,000 seconds) of exposure and a Gaussian PSF of the lens, and

where \( \gamma \) is the ratio of the major and minor axes. For the tidally stripped subhalo, we will be probing the inner region of an originally more massive halo that lost its mass. These inner regions have a shallower density slope compared to the slope at larger radii. This correlation needs to be taken into account when using gravitational lensing measurements to discern between different dark matter scenarios.

**4 MEASURING THE EFFECTIVE DENSITY SLOPE FROM HST-LIKE NOISY MOCKS**

We made realistic mock images of gravitational lenses at the resolution and sensitivity of Hubble Space Telescope (HST) observations to demonstrate the feasibility of measuring the effective density slope of perturbers. The images are 60×60 pixels with a pixel size of 0.025″, which gives a field of view of 1.5″ × 1.5″. We assume a noise level of 10 HST orbits (33,000 seconds) of exposure and a Gaussian PSF with a full width at half maximum (FWHM) of 0.049″. The sky brightness is 22.3 (AB magnitude) and the zero-point magnitude is 25.96.

In each image, we placed a perturber at a randomly chosen position near the critical curve, no dimmer than 50% of the brightest pixel. We model the main lens with an Elliptical Power-Law (EPL) mass profile given by Ref. Barkana (1999):

\[
\kappa(x, y) = \frac{3 - \gamma}{2} \left( \frac{\theta_E}{\sqrt{x^2 + y^2/q}} \right)^{\gamma - 1}, \tag{9}
\]

where \( \gamma \) is the negative power-law slope of the 3D mass distribution (see Appendix A for more details), \( \theta_E \) is the Einstein radius, \( x \) and \( y \) are the angular coordinates aligned with the major and minor axes of the lens, and \( q \) is the ratio of the major and minor axes. For the perturber, we assume an NFW profile, and we will fit it with an EPL mass profile with cylindrical symmetry, \( q = 1 \). In addition, we include an external shear to capture the effect of weak lensing along the line of sight.

We produced 6 mock images capturing possible variations in the models, which we list in Table 2 and show in Fig. 7. We set the lens redshift to \( z = 0.5 \) and the source redshift to \( z = 1.5 \) for all mock images. We analyze the mocks in this work using the publicly available code lenstract (Birrer et al. 2015; Birrer et al. 2021).

We will now explain each of the cases analyzed and refer the reader to Appendix B for more details on these models. We use dynasty (Speagle 2020), a publicly available nested sampling algorithm package, to sample the posterior probability distribution of the model parameters and to obtain the best fits from the sampled points. The main lens, external shear, and the perturber parameters are varied simultaneously during the posterior sampling. The means and standard deviations from the sampling as well as the true values for each case are shown in Table B1.

### Table 2. The properties of each of the models analyzed in this work (shown in Fig. 7). For the non-fiducial models we only list the parameters that differ from those of the fiducial. The remaining model parameters are chosen from uniform random distributions (see Appendix B for more details on these models). The effective perturber lensing mass is kept fixed (at \( 1.5 \times 10^{12} M_\odot \)) in all the models (this is why the *Steeper lens, High-concentration perturber* and *Smaller lens* models have a different value of \( M_{200} \)). The effective perturber mass is the projected mass within the perturbation radius, divided by the log-slope of the main lens’ 2-dimensional density profile (Minor et al. 2017).

| Model                  | Properties                      |
|------------------------|---------------------------------|
| Fiducial               | \( \theta_{E,\text{main}} = 0.55'' \) |
|                        | \( \gamma_{\text{main}} = 2.0 \) |
|                        | \( M_{200,\text{perturber}} = 10^9 M_\odot \) |
|                        | \( c_{200} = 12 \) |
|                        | \( m_{\text{AB}} = 22 \) |
|                        | \( N_{\text{sersic}} = 5 \) |
| Dimmer source          | \( m_{\text{AB}} = 24 \) |
| Simpler source         | \( N_{\text{sersic}} = 1 \) |
| Steeper lens           | \( \gamma_{\text{main}} = 2.3 \) |
|                        | \( M_{200,\text{perturber}} = 1.27 \times 10^9 M_\odot \) |
|                        | \( c_{200} = 30 \) |
| High-concentration perturber | \( M_{200,\text{perturber}} = 1.93 \times 10^8 M_\odot \) |
| Smaller lens           | \( \theta_{E,\text{main}} = 0.35'' \) |
|                        | \( M_{200,\text{perturber}} = 2.78 \times 10^8 M_\odot \) |
Figure 7. Mock images used in our analysis. The parameters of each image are described in Table 2.

chosen ellipticities, radii, and Sersic indices. The Sersic light profile is given by

$$I(x) = I_0 \exp \left[ -k \left( \frac{\sqrt{x_1^2 + x_2^2/q^2}}{R_s} \right)^{1/n_s} - 1 \right],$$

where $R_s$ is the Sersic radius, $I_0$ is the surface brightness at $R_s$, $x_1$ and $x_2$ are the coordinates of the centroid, $q$ is the ratio of the major and minor axes, and $n_s$ is the Sersic index. The normalizing constant $k$ ensures that $R_s$ is the half-light radius, which is where the surface brightness is half of the maximum value at the center. To test how much the source-light complexity affects our ability to measure the effective density slope, we also have created a mock lens that has a single Sersic profile as a source (this is the Simpler source model).

We model the source light with shapelets (Refregier 2003; Birrer et al. 2015; Birrer et al. 2019; Shajib et al. 2020; Şengü et al. 2021), which are orthonormal sets of weighted Hermite polynomials. The source complexity is controlled by the shapelet-order parameter, $n_{\text{max}}$, which determines the number of degrees of freedom $N$ via the relation $N = (n_{\text{max}} + 1)(n_{\text{max}} + 2)/2$. The size of the shapelet reconstruction is determined by the scaling parameter $\delta$. Given a set of lens-model and shapelet parameters, the brightness of each pixel in the image is a linear function of the shapelet coefficients. If one knows the covariance matrix of the noise of the pixels, the best fit shapelet coefficients can be obtained by a simple matrix inversion. The centroid and the width of the shapelet basis are freely varied like any other model parameter. The optimal $n_{\text{max}}$ changes depending on the complexity of the light distribution of the source galaxy and the resolution of the image. Therefore, we need to find the best fit $n_{\text{max}}$ for each image on a case-by-case basis. We do this by starting from a low value of $n_{\text{max}} = 5$ and iterating by 1 until the Bayesian information criterion (BIC) of the best fit is larger than the previous iteration. The BIC is given by $\text{BIC} = k \ln(n) + \chi^2$, where $k$ is the number of parameters in the model (this includes the number of shapelet coefficients) and $n$ is the number of unmasked pixels.

We find that in the Fiducial model, we can measure a slope of $\gamma = 1.559^{+0.054}_{-0.033}$. The posteriors of the perturber parameters are shown in Fig. 8a. The average convergence profile of the power-law best fit is compared with the true NFW profile in Fig. 8b where we see that the best-fit slope agrees with the average slope within the RMO. The red error bars are the 1$\sigma$ linear errors we get on the average convergences at each radius when all other model parameters are fixed. We note that this fiducial case is encouraging since we can measure $\gamma$ with an uncertainty of 3%, which is smaller than the differences in the means of the different DM populations (which is at least 7%, as shown in Fig. 3).

**Dimmer source:**

We analyze a case where the source is dimmer than the fiducial, with an AB magnitude of 24 instead of 22 (the amount of light we receive is reduced roughly by a factor of 6). We find, as expected, that the uncertainties of the model parameters are larger. The posteriors are shown in Fig. 9a. We see in 9b that the RMO has shifted to larger radii compared to the fiducial case, due to the change in source brightness, which agrees with the steeper power-law measurement from modeling. Larger linear errors within the RMO mean that the pixel differences are consistent with a wider range of power-law slopes, which limits our capacity to measure the effective slope of the system. Source brightness is the most important factor that limits our capacity to probe the effective density slope of the perturbers.

**Simpler source:**

If instead the source consists of a single Sersic profile instead of the 5 Sersics, we can measure the slope $\gamma = 1.502^{+0.05}_{-0.034}$. We see that lowering source complexity does not change our ability to probe the effective power-law slope of the perturber. The posteriors are shown in Fig. 10a. As we can see in Fig. 10b, the RMO shifts towards smaller radii, which is consistent with our measurement of a shallower slope of the perturber.

**Steeper lens:**

When the power-law slope of the matter distribution of the main lens is steeper (we increased it to 2.3), we measure a slope $\gamma = 1.556^{+0.091}_{-0.041}$ for the perturber. The uncertainty is slightly higher than that of the fiducial model, but the measurement is still precise enough to distinguish between the different DM scenarios analyzed in this work. The posteriors are shown in Fig. 11a. We find that the RMO does not get significantly affected by the change in the slope of the main lens, which can be seen by comparing Figs. 8b and 11b. Note that the perturber in this case has a larger $M_{200,\text{perturber}}$ compared to the fiducial one to keep the effective perturber lensing mass fixed.

**High-concentration perturber:**

If the concentration of the perturber is raised to $c_{200} = 30$, we find that we can measure a slope of $\gamma = 1.958^{+0.074}_{-0.076}$. The posteriors are shown in Fig. 12a. Our slope measurement is consistent with the average slope within the RMO, which is not significantly affected by the
Figure 8. Fiducial model: (a) Posterior probability distribution of the perturber parameters. The true position values are shown in solid green lines. The average power-law slope in the RMO is shown as the dashed green line. (b) Average convergence as a function of radius. The power-law best fit for the Fiducial model is shown in the cyan line. The true NFW profile is shown with solid green lines. The red error bars show the 1σ linear errors of the average convergence within each radii. The blue-shaded region corresponds to the RMO.

Figure 9. Dimmer source model: (a) Posterior probability distribution of the perturber parameters. The true position values are shown in solid green lines. The average power-law slope in the RMO is shown as the dashed green line. (b) Average convergence as a function of radius. The power-law best fit for the Dimmer source model is shown in the cyan line. The true NFW profile is shown with solid green lines. The red error bars show the 1σ linear errors of the average convergence within each radii. The blue-shaded region corresponds to the RMO.
Figure 10. **Simpler source** model: (a) Posterior probability distribution of the perturber parameters. The true position values are shown in solid green lines. The average power-law slope in the RMO is shown as the dashed green line. (b) Average convergence as a function of radius. The power-law best fit for the **Simpler source** model is shown in the cyan line. The true NFW profile is shown with solid green lines. The red error bars show the $1\sigma$ linear errors of the average convergence within each radii. The blue-shaded region corresponds to the RMO.

Figure 11. **Steeper lens** model: (a) Posterior probability distribution of the perturber parameters. The true position values are shown in solid green lines. The average power-law slope in the RMO is shown as the dashed green line. (b) Average convergence as a function of radius. The power-law best fit for the **Steeper lens** model is shown in the cyan line. The true NFW profile is shown with solid green lines. The red error bars show the $1\sigma$ linear errors of the average convergence within each radii. The blue-shaded region corresponds to the RMO.
change of the perturber profile, shown in Fig. 12b. An NFW perturber with a higher concentration has a lower scale radius. Around the scale radius, the effective slope is roughly $\gamma = 2$. Since the RMO for this case (shown in Fig. 12b) includes the scale radius, which is $r_s \approx 0.1''$, we measure an effective slope close to 2. The best fit power-law is compared with the true NFW profile in Fig. 12b. In this case, the perturber also has a larger $M_{200,\text{perturber}}$ compared to the fiducial one to keep the effective perturber lensing mass fixed.

**Smaller lens:**

Furthermore, we tested making the Einstein radius of the main lens smaller (0.3''), and we found that in this case we can measure the slope $\gamma = 1.465\pm 0.041\pm 0.032$. The posteriors are shown in Fig. 13a. The slope is consistent with the expected one from the RMO, shown in Fig. 13b. This case outperforms even the *Fiducial* model because, with a smaller Einstein radius, the source light (which is kept at the same magnitude) is now concentrated in a smaller number of pixels. Since we enforce that the perturber has to be close to those pixels, the deflections caused by the perturber are now measured more accurately. Another contributing factor is the perturber being more massive relative to the main lens (since now the latter has a smaller mass). The same perturber mass is now more dominant in the overall shape of the image. However, we note that in lenses with smaller Einstein radii, the probability of finding a perturber at bright enough positions is also smaller. Here again, the perturber has a larger $M_{200,\text{perturber}}$ value compared to the fiducial one to keep the effective lensing mass of the perturber fixed.

### 4.1 Robustness tests

To confirm that we are indeed measuring the power-law slope of the average convergence within the RMO, we performed robustness tests where we create a set of mock lenses with an NFW perturber where the concentration of the perturber varies from $c = 9$ to $c = 78$, while the mass within the RMO is kept fixed at $1.5 \times 10^7 \, M_\odot$. The choices we made for the rest of the lens parameters are the same as in the *Fiducial* case in §4. We expect to measure a steeper slope for the higher concentration case because an increase in concentration results in a decrease in the scale radius for an NFW halo. At the same scale, an NFW halo with a smaller scale radius has a steeper power-law slope than an NFW halo with a larger scale radius. In Fig. 14, we show the slope measurements as the green error bars obtained following the same modeling procedure described in §4, which involves sampling the posteriors of a model that has a power-law perturber. The cyan lines show the average power-law slope of the true NFW profile of each perturber within the RMO. We see that the measured slope agrees with the expected slope from the RMO for all the cases.

We have also tested for changes in the truncation radius, $r_t$, of an NFW with a fixed mass and scale radius. We measure (again by modeling a mock lens with a power-law perturber) the effective slope of a truncated NFW perturber (*Baltz et al. 2009*) with $\tau = r_t/r_s = 1, 2, 3, 4, 5$. In Fig. 14, we show with green error bars the 1$\sigma$ confidence interval of the slope measurements from mock lenses by fitting a power-law. The cyan line shows the average power-law slope of the true truncated NFW profile of each perturber within the RMO. For each case, we find that the measured slope matches the expected slope from the RMO.

### 5 MEASURING THE EFFECTIVE DENSITY SLOPE OF THE PERTURBER IN JVAS B1938+666

We will focus here on a perturber that has been previously detected in the strong lens system JVAS B1938+666 (*Lagattuta et al. 2012; Vegetti et al. 2012*), using the gravitational imaging technique (*Koopmans 2005*). The main lens in this system is at redshift $z = 0.881$ (*Tonry & Kochanek 2000*) and the source is at redshift $z = 2.059$ (*Riechers 2011*). The main lens has an Einstein radius of roughly $\theta_E \approx 0.45''$. We have re-analyzed this system and found that the perturber that was previously assumed to be a subhalo is likely a line-of-sight halo. We refer to Ref. *Şengül et al. (2021)* for a detailed discussion of this analysis. In this section, we will present our measurement of the effective density slope of the perturber in JVAS B1938+666. In Ref. *Vegetti et al. (2012)*, this perturber was shown to be consistent with a singular isothermal sphere (SIS) profile, which is a power-law with a slope of $\gamma = 2$. A subsequent analysis (*Vegetti & Vogelsberger 2014*) found that for this system, a family of density profiles is consistent with an SIS profile over a large region in parameter space. Considering the sensitivity and the source-light distribution, they show that the data are not sensitive to the size of the inner density core. Motivated by these findings and our results in the previous section, we model the main lens and the perturber as an Elliptical Power-Law, shown in Eq. (9), and an external shear. The perturber is modeled with a cylindrically symmetric power-law mass profile (by fixing $q = 1$ in Eq. 9). Although the true density profile of the perturber might deviate from this power-law at smaller and larger radii, the data are not currently sensitive to such deviations.

The source light is modeled with shapelets with an optimal shapelet index of $n_{max} = 10$. The shapelet index optimization is done by minimizing the BIC over models with different $n_{max}$, as in the previous section. The background and Poisson noise are set to match the noise level of the image used, following our previous analysis (*Şengül et al. 2021*).

We perform a likelihood analysis using *dynesty* (*Speagle 2020*), and show the posteriors of the perturber parameters in Fig. 16. The data, reconstruction, and residuals are shown in Appendix C. We find an effective density slope of $\gamma = 1.96^{+0.12}_{-0.12}$. Assuming that the underlying slope distributions are Gaussian, our measurement is a $> 2.1 \sigma$ outlier for a CDM slope distribution from simulations at $z = 1.47$ shown in Fig. 4. This result is also in $3.9 \sigma, 4.5 \sigma, 5.1 \sigma, \text{ and } 3.2 \sigma$ tension with ETHOS1-4, respectively. If the perturber were a subhalo, instead of a line-of-sight halo as Ref. *Şengül et al. (2021)* shows, its slope would have to be compared to the $\gamma = 0.88$ simulation snapshots. In this case the slope measurement would be a $1.9 \sigma$ outlier for CDM and a $3.4 \sigma, 4.1 \sigma, 4.6 \sigma, \text{ and } 2.7 \sigma$ outlier for the ETHOS1-4 models, respectively.

From our measurement of the power-law slope within the RMO and assuming that the mass distribution of the perturber follows an NFW profile, we can estimate the scale radius and the concentration of the perturber. Around the scale radius, the NFW profile slowly transitions from a slope of $\gamma = 3$ to $\gamma = 1$, as we move to smaller radii, as can be seen in Fig. 17. Therefore, the density profile of an NFW perturber that has a scale radius that is much smaller than the RMO will effectively behave like a power-law with a slope of $\gamma = 3$. Conversely, when the scale radius is much larger than the RMO, the density profile of the perturber will behave like a power-law with a slope of $\gamma = 1$. In our case, we are measuring a power-law slope of $\gamma = 2$, which means that we are probing the density profile at roughly the NFW scale radius. In contrast to our previous study with mocks, for the real JVAS B1938+666 system, we do not know the true source-light distribution, main lens profile, and the perturber...
**Figure 12.** High-concentration perturber model: (a) Posterior probability distribution of the perturber parameters. The true position values are shown in solid green lines. The average power-law slope in the RMO is shown as the dashed green line. (b) Average convergence as a function of radius. The power-law best fit for the High-concentration perturber model is shown in the cyan line. The true NFW profile is shown with solid green lines. The red error bars show the 1σ linear errors of the average convergence within each radii. The blue-shaded region corresponds to the RMO.

**Figure 13.** Smaller lens model: (a) Posterior probability distribution of the perturber parameters. The true position values are shown in solid green lines. The average power-law slope in the RMO is shown as the dashed green line. (b) Average convergence as a function of radius. The power-law best fit for the Smaller lens model is shown in the cyan line. The true NFW profile is shown with solid green lines. The red error bars show the 1σ linear errors of the average convergence within each radii. The blue-shaded region corresponds to the RMO.
Figure 14. The expected and measured power-law slope for perturbers with varying concentration. The power-law slope of the average convergence within the RMO is shown in cyan line. The green error bars correspond to the $1\sigma$ confidence interval of the slope measurement obtained modeling the mock lens with a power-law perturber.

Figure 15. The expected and measured power-law slope for perturbers with varying truncation radius. The power-law slope of the average convergence within the RMO is shown in cyan line. The green error bars correspond to the $1\sigma$ confidence interval of the slope measurement obtained modeling the mock lens with a power-law perturber.

Figure 16. Posterior probability distribution of the perturber parameters in the system JVAS B1938+666. The vertical lines correspond to the best fit values of the parameters.

Figure 17. The effective slope of an NFW profile as a function of radius in units of the scale radius $r_s$. The redshift of the perturber. The scale radius of the perturber needs to be within the RMO so that the power-law slope of the perturber at our measurement scale is close to $2$. If $r_s$ were any bigger (smaller), we would expect to measure a slope that is closer to $\gamma = 1(3)$. Considering the measured mass of the perturber, a scale radius of $r_s = 0.25 \text{kpc}$ implies a concentration of roughly $c_{200} = 60$, which is much higher than the expected value of $c_{200} = 10$ given by the mass-concentration relation found in CDM simulations (Dutton & Macciò 2014). This is consistent with our measurement of $c_{200} = 54^{+21}_{-15}$ in Ref. Şengül et al. (2021) and also matches the expectation in Ref. Minor et al. (2021a) that only high-concentration perturbers with masses $\lesssim 10^9 \text{M}_\odot$ are expected to be detected at the HST resolution.

6 CONCLUSIONS

It is standard practice, in strong gravitational lensing studies, to provide a measurement of a system by making assumptions about its extended profile. In some cases, these assumptions can change the position. Instead, we use the best fits from our analysis of the system outlined in Ref. Şengül et al. (2021) to calculate the RMO. In Fig. 18, we show the linear errors for the average convergence of the perturber.

The RMO is shown in the blue-shaded region. We see that we are most sensitive to the average convergence between roughly $0.04'' - 0.13''$. This corresponds to a physical scale of roughly $0.17 - 0.55 \text{kpc}$ at scale radius $r_s$.
Probing Dark Matter with Strong Gravitational Lensing through an Effective Density Slope

We demonstrate the feasibility of measuring the effective density slope by making mock measurements of Hubble Space Telescope-like data. We find that if the source galaxy is sufficiently bright (an AB magnitude of 22 or brighter), measuring the slope of a perturber with a mass of $1 \times 10^9 \, M_\odot$ is possible with a 3% uncertainty. This measurement also depends on accurate modeling of the main lens profile. Deviations of the main lens mass distribution from an elliptical power-law can result in a bias in the inferred effective power-law slopes for perturbers. The differences in the means of the distributions observed in different DM populations are generally larger than our measurement accuracy, suggesting that we should be able to distinguish between different DM scenarios as we accrue more measurements of the slopes of the perturbers in strong lensing systems.

As a step towards this effort, we measure in this work the effective slope of the density profile of the perturber in the strong lens system JVAS B1938+666. We find that its slope is steeper than what is expected from simulations, being in strong tension with ETHOS1-3. It is also a $3\sigma$ outlier for ETHOS4 and a $2\sigma$ outlier for CDM. This tension can also be seen as a tension in concentration values. For an NFW halo, a steep slope indicates a high concentration and vice versa. More measurements of the slopes of perturbers are needed to tell if this indicates new physics, as perturbers with steeper effective slopes (higher concentrations) have a stronger lensing signal and are more likely to be detected.

In the upcoming years, we expect to find tens of thousands of new strong lensing systems with experiments such as the Vera C. Rubin Observatory, Euclid, the Dark Energy Survey (DES), the Dark Energy Spectroscopic Instrument (DESI), the Square Kilometre Array (SKA), and the Nancy Grace Roman Telescope (Oguri & Marshall 2010; Collett 2015; McKeen et al. 2015; Huang et al. 2021). As the capabilities to observe strong lens systems increase, we expect to find a large number of perturbers. Follow-up analyses of these perturbers with high-resolution imaging from the Hubble Space Telescope, the James Webb Space Telescope, and the Extreme Large Telescope will provide us with a distribution of effective density slopes that can be then compared with simulations.

Acknowledgments

CD and ACS are partially supported by the Department of Energy (DOE) Grant No. DE-SC0020223. We thank Mark Vogelsberger and Jesús Zavala for providing us with the CDM and ETHOS simulations analyzed in this work. We would also like to thank Simon Birrer, Adam Coogan, Akaxia Cruz, Arthur Tsang, and Sebastian Wagner-Carena for valuable discussions.

Data Availability

The system analyzed in this work is JVAS B1938+666. The data of this system are publicly available on https://mast.stsci.edu/portal/Mashup/ Clients/Mast/Portal.html. The code used for our analysis is available at https://github.com/acagansengul/interlopers_with_lenstronomy. The data from the particle simulations will be shared on request to the corresponding author with permission of Mark Vogelsberger and Jesús Zavala.

APPENDIX A: EXTRACTING THE AMPLITUDE AND SLOPE OF THE AVERAGE CONVERGENCE FROM SIMULATIONS

This appendix shows the details behind the extraction of the effective density slopes from N-body simulations.

We sort the particles of each (sub)halo in the simulations in radial bins of equal width $\delta r$, proportional to the $r_{200}$ of that (sub)halo, around the center. The expected number of particles $\mu_i$ in each bin
i will be proportional to the density of the (sub)halo $\rho(r)$ at that radius,

$$
\mu_i = \frac{1}{m_p} \int_{r_i}^{r_i + \Delta r} dr \, 4\pi r^2 \rho(r) \approx \frac{4\pi r_i^2 \Delta r \rho(r_i)}{m_p},
$$

(A1)

where $m_p$ is the mass of a particle in the simulation. Using Eq. (A1), we obtain the density of each halo at various radii. The uncertainties of these densities are calculated assuming that the number of particles in each bin follows a Poisson distribution with expectation value $\mu_i$. The 2D surface mass density for each (sub)halo can be written in terms of the 3D density as

$$
\Sigma(r) = \int_0^\infty dz \rho \sqrt{r^2 + z^2}.
$$

(A2)

We are interested in the average convergence, denoted by $\bar{\kappa}(r)$, which is proportional to the mass within the 2D radius $r$ divided by $r^2$:

$$
\bar{\kappa}(r) \propto \frac{1}{r^2} \int_0^r dr' 2\pi r' \Sigma(r').
$$

(A3)

A 3D power-law density profile $\rho \propto r^{-\gamma}$ implies $\bar{\kappa} \propto r^{-\gamma+1}$. So we can extract $\gamma$ by calculating,

$$
\gamma = 1 - \frac{r(d\bar{\kappa}/dr)}{\bar{\kappa}(r)}
$$

(A4)

which we do numerically from the bin populations $n_i$ in the simulations.

APPENDIX B: ANALYSIS OF THE MOCK IMAGES

In this appendix, we show a complete list of the model parameters of the mock images analyzed in §4. In Table B1, we show the mean and 1σ uncertainties of the lens-model parameters of the different cases listed in Table 2. We see that the means and the uncertainties of the main lens and shear parameters are consistent with the true values for all six cases considered here.

Degeneracy Between Main Lens Model and the Perturber

We present in Fig. B1 the posterior probability distributions of the model parameters for the Fiducial case. We see a strong correlation between the Einstein radius of the main lens model ($\theta_{E,\text{main}}$) and the effective power-law slope of the perturber ($\gamma$). This indicates that failing to correctly model the main lens mass profile can result in a biased measurement of the power-law slope for perturbers.

Source Reconstruction

We present in Figs. B2, B3, B4, B5, B6, and B7 the source used for the models analyzed in §4, along with its reconstructions.

APPENDIX C: DATA, RECONSTRUCTION, AND RESIDUALS IN JVAS B1938+666

The Hubble Space Telescope image of JVAS B1938+666 used in this study, its reconstruction and residuals are shown in Fig. C1. The posterior probability distributions of the model parameters obtained by nested sampling are shown in Fig. C2.
### Table B1. Lens-model properties of each of the models shown in Fig. 7. The first row of each model shows the true values, while the second row shows the means and standard deviations of the measurements coming from a likelihood analysis, as described in §4. The parameters \( c_x \) and \( c_y \) are the \( x \) and \( y \) coordinates of the centroid of the main lens, and \( e_x \) and \( e_y \) are the \( x \) and \( y \) component of the ellipticity of the main lens profile. The true \( \theta_E \) and \( \gamma \) for the perturbers are not applicable because they follow NFW profiles instead of EPL profiles.
Figure B1. Posterior probability distribution of the model parameters of the Fiducial case described in Table 2. Model parameters are described in Table B1. We see that the Einstein radius of the main lens ($\theta_{E,\text{main}}$) is correlated with the effective power-law slope of the perturber ($\gamma$).
Probing Dark Matter with Strong Gravitational Lensing through an Effective Density Slope

Figure B2. Best fit of the Fiducial mock in Table 2. The top left panel shows the mock image with noise. The top middle panel shows the model reconstruction from the best fit, where the perturber position is marked with the white cross. The top right panel shows the residuals divided by the noise in each pixel. The bottom left panel shows the true source distribution that is made up of 5 Sersic components. The bottom middle panel shows the shapelet reconstruction of the source. The bottom right panel shows the convergence of the best fit model in log scale, where the perturber can be seen as a small peak.

Figure B3. Same as in Fig. B2 but for the Dimmer source mock (see Table 2).
Figure B4. Same as in Fig. B2 but for the Simpler source mock (see Table 2).

Figure B5. Same as in Fig. B2 but for the Steeper lens mock (see Table 2).
Figure B6. Same as in Fig. B2 but for the High-concentration perturber (see Table 2).

Figure B7. Same as in Fig. B2 but for the Smaller lens (see Table 2).
Figure C1. Lens-light subtracted Hubble Space Telescope image of the lens JVAS B1938+666, the model prediction (where the mass profile of the perturber is modeled with a free power-law slope, as in §5), the residuals of the model with a perturber, and the reconstructed source (left to right panels).
Figure C2. Posterior probability distribution of the model parameters in the system JVAS B1938+666. The parameters $\theta_{E,\text{main}}$, $\gamma_{\text{main}}$, $x_1$, $x_2$, $e_1$, and $e_2$ are the Einstein radius, power-law slope, the x and y coordinate of the center, and the Cartesian ellipticity components of the main lens profile, respectively. The parameters $\Gamma_1$ and $\Gamma_2$ are the Cartesian components of the external shear. The parameters $\theta_{E}$, $\gamma$, $x$, $y$, and $z$ are the Einstein radius, the power-law slope, the x and y coordinate of the center and the redshift of the perturber, respectively. The vertical lines correspond to the best fit values of the parameters.
Probing Dark Matter with Strong Gravitational Lensing through an Effective Density Slope

Shajib A. J., et al., 2020, MNRAS, 494, 6072
Speagle J. S., 2020, MNRAS, 493, 3132
Spergel D. N., Steinhardt P. J., 2000, Phys. Rev. Lett., 84, 3760
Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137
Syer D., White S. D. M., 1998, Monthly Notices of the Royal Astronomical Society, 293, 337
Tonry J. L., Kochanek C. S., 2000, AJ, 119, 1078
Troxel M. A., et al., 2018, Phys. Rev. D, 98, 043528
Tulin S., Yu H.-B., 2018, Phys. Rept., 730, 1
Vegetti S., Vogelsberger M., 2014, MNRAS, 442, 3598
Vegetti S., Koopmans L. V. E., Bolton A., Treu T., Gavazzi R., 2010, MNRAS, 408, 1969
Vegetti S., Lagattuta D. J., McKeen J. P., Auger M. W., Fassnacht C. D., Koopmans L. V. E., 2012, Nature, 481, 341
Vegetti S., Koopmans L. V. E., Auger M. W., Treu T., Bolton A. S., 2014, MNRAS, 442, 2017
Vogelsberger M., Zavala J., Loeb A., 2012, MNRAS, 423, 3740
Vogelsberger M., Zavala J., Cyr-Racine F.-Y., Pfrommer C., Bringmann T., Sigurdson K., 2016, MNRAS, 460, 1399
Wagner-Carena S., Aalbers J., Birrer S., Nadler E. O., Darragh-Ford E., Marshall P. J., Wechsler R. H., 2022, arXiv e-prints, p. arXiv:2203.00690
Wang J., Bose S., Frenk C. S., Gao L., Jenkins A., Springel V., White S. D. M., 2020, Nature, 585, 39
Xu D. D., Mao S., Cooper A. P., Gao L., Frenk C. S., Angulo R. E., Helly J., 2012, Monthly Notices of the Royal Astronomical Society, 421, 2553
Şengül A. c., Tsang A., Diaz Rivero A., Dvorkin C., Zhu H.-M., Seljak U., 2020, Phys. Rev. D, 102, 063502
Şengül A. c., Dvorkin C., Ostdiek B., Tsang A., 2021, arXiv e-prints, p. arXiv:2112.00749
de Blok W. J. G., 2010, Advances in Astronomy, 2010, 789293
de Blok W. J. G., McGaugh S. S., Bosma A., Rubin V. C., 2001, Astrophys. J. Lett., 552, L23

MNRAS 000, 1–23 (2022)