
1

SIMULATION-BASED INFERENCE FOR STRONG GRAVITATIONAL LENSING

[TO DO:

- Resolution of Fig. 3.13, 3.14, 3.16, 3.17, 3.18, 3.19
- Merge methodology + write tasks subdivision
- Polish pop results

]

Precision analysis of galaxy-galaxy strong gravitational lensing images provides a unique way of characterizing small-scale dark matter halos, and could allow us to uncover the fundamental properties of dark matter's constituents. Recently, gravitational imaging techniques made it possible to detect a few heavy subhalos. However, gravitational lenses contain numerous subhalos and line-of-sight halos, whose subtle imprint is extremely difficult to detect individually. Existing analyses are reliant on likelihood-based methods like Markov-chain Monte Carlo or nested sampling, and generally require various compromises to the realism of lensing models for the sake of computational tractability, such as ignoring the numerous other subhalos and line-of-sight halos in the system, assuming a particular form for the source model, compressing observations into hand-crafted summary statistics (e.g.a power spectrum of residuals), and requiring the noise to have a known likelihood function.

Here we show that a simulation-based inference method called truncated marginal neural ratio estimation (TMNRE) makes it possible to relax these requirements by training neural networks to directly compute marginal posteriors for parameters of interest from lensing images. By performing a set of proof-of-concept inference tasks on simulated data, first, we verify the accuracy of TMNRE and show it can compute posteriors for subhalo parameters marginalized over populations of hundreds of

substructures that would be undetectable on their own, as well as lens and source uncertainties. Second, we show how to perform hierarchical inference of the dark matter halo mass function cutoff mass from multiple lensing images. Furthermore, we show that since TMNRE learns a posterior function it enables direct statistical checks that would be extremely expensive with likelihood-based methods. Our results show that the full subhalo and line-of-sight halo population must be included when measuring the properties of individual dark matter substructures with this technique, and that TMNRE is able to put tight constraints on the mass of warm dark matter in the multi-keV regime. TMNRE is well-suited for analyzing complex lensing data. These first analysis demonstrate that our framework is in principle able to extract the wealth of information regarding dark matter’s nature contained in existing lensing data and in the large sample of lenses that will be delivered by near-future telescopes.

This chapter is based on work from [1, 2].

1.1 Introduction

Dark matter puzzle. Over the past several decades, numerous astrophysical probes including rotational curves of spiral galaxies [3], galaxy-cluster dynamics [4], cosmic microwave background [5], gravitational lensing observations [6], have established dark matter (DM) as one of the major components of the universe, comprising about 85% of its mass. However, up to the present time, the fundamental nature of DM is still one of the key unresolved puzzle in physics. For many years, the cold dark matter (CDM) paradigm [7] has been able to accurately reproduce vastly disparate large-scale observations across all epochs. In this model, DM is massive, neutral, non-relativistic, and collisionless.

Despite providing a stunning description of the observed distribution of matter on large scales ($> \mathcal{O}(\text{Mpc})$), the agreement between CDM predictions and observations at galactic and sub-galactic scales has been less clear. At present, there is continued debate over whether the known abundance of dwarf galaxies and the density profiles of low-mass galaxies are in tension with the predictions of Λ CDM (respectively dubbed the missing satellites problem [8, 9] and the cusp-core problem [10], reviewed in 11). Solutions to these tensions include the impact of baryonic processes, such as supernovae feedback and reionization processes [12], or alternative DM physics.

The latter approach requires an alteration of DM particle physics, such that large-scale predictions remain unaffected, but the number of small-scale substructures is suppressed. DM models which are warm instead of cold [13–15], collisional instead of collisionless [16], or quantum on macroscopic scales rather than classical [17] predict a diverse array of possible configurations of low-mass halos and could potentially resolve these tensions [18]. Unfortunately, light DM halos are difficult to probe as they are

not expected to accumulate enough baryonic matter to form stars and hence are truly *dark* [19, 20]. If DM has significant self-interactions, such halos might be detectable by searching for the self-annihilation or decay products of DM [21]. However, even if such interactions are not present, light halos can potentially be probed through their irreducible gravitational effects. In this chapter we study one such probe: galaxy-galaxy strong gravitational lensing.

Strong gravitational lensing as a dark matter probe. In strong gravitational lensing, the gravitational field of a mass distribution acts as a *lens* by magnifying and distorting the light flux coming from a background *source* [22]. This leads to multiple magnified and distorted images of the source, as explained by general relativity. The effect is sensitive only to how matter is distributed, regardless of its physical nature (baryonic/DM), and thus provides a direct way of probing the distribution of DM at small scales (see [23] for a recent review).

Indeed, a *perturber* (i.e., a subhalo or line-of-sight (LOS) halo lying somewhere between the observer and source) positioned near one of these images contributes additional, much more localized distortions, on top of the main lens mass distribution. By carefully analyzing the relationship between the multiple images of the source, the distortions from a perturber can be disentangled from possible variations in the source light and its properties can be measured. Moreover, a population of dark perturbers can collectively cause perturbations to images that can be detected statistically in order to constrain population-level parameters, such as the suppression scale in the low-mass end of the halo mass function (HMF), which are dictated by the fundamental properties of DM. Therefore, gravitational lensing provides a pristine probe of small-scale structures and can in principle distinguish between DM scenarios.

Strong lensing image analysis. Various different methods have been suggested to analyse the effects of small-scale structures on lensing images [24]. These methods usually target two different types of lensing systems that differ in the lensed source: quadruply-lensed quasars, and extended background galaxies that get lensed into extended arcs or complete Einstein rings.

In the former case, the source is a nearly point-like quasar that is lensed into four compact images (“*quads*”). These images’ positions and flux ratios comprise the summary statistics for these systems. The presence of a perturber near one of these images would cause anomalies in the ratios of their fluxes relative to what would be predicted assuming a smooth lens mass distribution. Evidence for flux ratio anomalies due to perturber was first found in [25]. Later, [26] derived a statistical constraint on the substructure fraction in the lensing galaxies using a small sample of seven lensed quasars. [27] showed that flux-ratio anomalies can also be used to detect individual

low-mass subhalos. Several studies derived upper limits on the subhalo mass function [28], also including perturbations due to LOS halos [29, 30]. Further investigations [31–33] pointed out the importance of correctly modeling baryonic structure in the main lens, in order to avoid systematic errors while constraining DM substructure abundance with flux-ratio anomalies.

Here we focus on *gravitational imaging*, which refers to the analysis of lenses with extended arcs [34–36]. The observation in this case consists of a whole image. On one hand, such images cover a larger area of the sky than the four point-like images in quads, potentially providing more sensitivity to detect perturbations due to perturbers, in the form of percent-level variations in the shape of the predicted lensed light based on a smooth lens model. On the other hand, extracting this information requires modeling the source galaxy’s light, which generally has a complex morphology. The gravitational imaging technique was first introduced in [37] and further developed in [35, 36]. Its application to real data has so far yielded to several detections of individual heavy ($> 10^8 M_\odot$) perturbers using deep, high-resolution observations in the optical from the Hubble Space Telescope (HST) and Keck as well as in radio data from Atacama Large Millimeter/submillimeter Array [38–42]. Moreover, measurements and non-detections of individual perturbers in samples of gravitational lens systems can be converted to constraints on the (sub)halo mass function and thus dark matter’s properties [43–45].

Established gravitational imaging analyses such as the method in [35, 41] use *likelihood-based* inference to infer the properties of perturbers. The central mathematical object in such approaches is the likelihood, a probabilistic model $p(\mathbf{x} | \Theta)$ for the data \mathbf{x} given some parameters Θ for the lens, source, perturbers and possibly other (hyper)parameters.¹ Likelihood-based tools, such as Markov-chain Monte Carlo (MCMC) or nested sampling [46], do not directly produce marginal posteriors but instead compute the *joint posterior* $p(\Theta | \mathbf{x}_0)$, which must then be marginalized over.

The computational expense of sampling from the high-dimensional joint posterior imposes restrictions on the realism of lensing models that can be analyzed. One such restriction common to most analyses is to assume a particular form of the noise and source model so that the source uncertainties can be excluded from the sampling and marginalized over analytically [35, 38–41]. This makes it difficult to explore more complex source models described by e.g. generative machine learning methods or noise artifacts like cosmic ray streaks that cannot be described by an analytic likelihood.

An additional difficulty with likelihood-based analyses is that each run of MCMC or nested sampling produces posterior samples for just a single observation. Directly

¹For example, the hyperparameters could include the pixel size for pixelated sources or strength of source regularization.

exploring the systematics, biases and other statistical properties of a particular lensing model is thus extremely time-consuming, necessitating rerunning posterior sampling many times for different input observations. This also makes analyses such as mapping perturber measurement sensitivity costly. It is noteworthy that recently [47] pushed to the limits how far one can feasibly go using likelihood-based analyses, fitting 54 images with 5 different mass models.

Likelihood-based analyses also typically assume no more than two perturbers are present in each image. Allowing for n perturbers would cause the joint posterior to become highly multimodal, with approximately $n!$ modes due exact invariance of the observation under relabeling of perturbers. In these frameworks, it is then challenging to perform statistical inference of quantities such as the posterior of mass and position of a single heavy subhalo in the presence of a population of lighter perturbers, or the posterior for substructure population-level parameter of interest, marginalizing over all source, lens, and substructures parameters to get the marginal of interest.

Transdimensional Bayesian inference [48, 49] partially overcomes this traditional likelihood-based methods' challenge, by using transdimensional MCMC to infer the probabilities of different possible populations of perturbers, albeit at substantial computational cost.

An alternative strategy involves linearising the gravitational potential via a *Taylor expansion* of the lens equation. By employing a Taylor expansion, it becomes feasible to capture all small-scales DM substructures without the need to parameterise them directly in the likelihood function [35, 37, 50]. This technique is therefore able to account for the full DM subhalo population (a part from the effects due to the curl-component induced by multi-plane lensing effects).

Another approach is to circumvent measuring individual perturbers by instead engineering *summary statistics* from first principles, such as the convergence power spectrum (PS) for different subhalo populations [51–53], and also for LOS populations [54]. However, this approach is not directly applicable to observations, because we do not have access to the true displacement field from the data. Building upon this, it is possible to relate the PS of the surface brightness fluctuations in strong lens images to the lens potential fluctuations arising from DM distribution that contribute to the convergence PS [55–57]. Another summary statistic that has been employed is the residuals between the image and best-fit reconstruction excluding substructure, that is related to the (sub)halo mass function parameters. In particular, this summary statistic has been successfully employed to constrain the HMF suppression scale using approximate Bayesian computation (ABC) [58, 59]. This approach reduces the dimensionality of the problem and enable inference of the collective effects of a large number of low-mass substructures at the statistical level. However, it is unknown how much information such approach discards.

Another class of methods that has developed in recent years uses *neural networks* to measure lens parameters [60–62], quantifying the structure of gravitational lens potential [63], detect individual subhalos [64], distinguish different types of DM substructure based on their lensing signatures [65], and classify whether each pixel in an image contained a subhalo in a given mass bin [66, 67]. Still, these methods need lots of data to amortize over all possible variations in lensing systems. In fact, amortized methods learn the posterior for any data, generated by any parameter over the whole range of the prior. But learning an amortized posterior is unnecessary if only a small range of parameters are consistent with a target observation.

This work. In this work, we demonstrate that a simulation-based inference (SBI) [68] method called truncated marginal neural ratio estimation [69, 70], from here on TMNRE, can circumvent these inference challenges to *measure the properties of individual perturbers* and to *measure the suppression scale of the subhalo HMF directly from images by combining multiple observations*.

In a nutshell, SBI, first presented in ??, refers to a class of statistical inference methods that use the output of a stochastic simulator that need not have a known likelihood. In particular, neural ratio estimation (NRE), first presented in [71], trains a neural network to map from observations directly to *marginal posteriors* for a specified subset of model parameters (e.g. the position and mass of a perturber). This bypasses the requirement of likelihood-based inference to sample the joint posterior. In contrast to methods like ABC, this also removes the need to engineer summary statistics [59] as they are in effect learned directly from the training data. Since NRE learns a marginal posterior *function*, it is straightforward to check the statistical properties of the inference results for different observations. TMNRE further extends NRE by focusing training data generation in the regions of parameter space most relevant for analyzing a particular observation over a sequence of inference rounds. This substantially reduces the number of simulations required to train the inference network as well as the required network complexity.

Several other works have applied SBI to substructure lensing. In [72] a likelihood-ratio estimation technique similar to TMNRE was employed to measure density profile parameters of subhalos from images. [73] recently applied neural posterior estimation to measure the subhalo HMF normalization in mock lensing images using real galaxy images as sources. [74] utilized a “likelihood-based” SBI method requiring the simulator’s score² to measure the slope and normalization of a subhalo HMF in simple mock images.

The present work complements these efforts in several ways. As far as measuring

²The score is the derivative of the log-likelihood for a given observation with respect to the model’s parameters.

individual perturbers complements is concerned, first, it offers a path towards cross-checking current substructure measurements under different modeling assumptions. Second, inference based on perturbers provides a level of interpretability beyond measuring subhalo HMF parameters directly from images, and moreover the opportunity to test different DM models through measuring the properties of individual subhalos. Third, measuring the heaviest subhalos in an observation enables modeling them explicitly in lensing simulations, which could reduce the training data requirements and improve inference accuracy for direct subhalo HMF measurements.

On the other hand, as far as measuring the suppression scale of the subhalo HMF directly from an ensemble images is concerned, first, this work demonstrates the sensitivity of the TMNRE approach to the HMF suppression scale. Second, it illustrates the importance of combining the information coming from different observations in the statistical analysis, especially in light of near-future data delivery. Currently, there are around a hundred strong lensing observations suitable for substructure inference, most of which come from the SLACS [75] and BELLS [76] surveys. In the near future, new and future telescopes like JWST [77], ELT [78], Euclid [79, 80], SKA [81], and LSST [82] will greatly increase the quality of data suitable for gravitational imaging analyses as well as its quantity, from $\mathcal{O}(100)$ to $\mathcal{O}(10^5)$ images [83, 84].

This chapter is organized as follows. In section 1.2 we describe our strong lensing model, which uses an analytic source and main lens in conjunction with well-motivated perturber models. We then carry out two distinct analysis tasks with TMNRE: inference of the parameters of a single subhalo in a lensing image section 1.4, and hierarchical inference of the halo mass function cutoff scale from an ensemble of lenses section 1.5. This work will help form the basis for SBI-based analysis of strong lensing images as a dark matter probe in existing and future lensing data.

1.2 Modelling strong lensing observations

Here we review how we model strong lensing images. We implement our lensing model in PyTorch [85] so that we can leverage GPUs to rapidly generate large numbers of observational data (a single-channel telescope image).

1.2.1 The physics of strong lensing

Before delving into modelling details, we briefly summarise the key points of the physics of gravitational lensing, referring the reader to e.g. [86] for a more detailed overview. In strong-lensing systems the mass distribution of a foreground galaxy gravitationally lenses the light rays coming from a background source, resulting in an arc-like image in the case of an extended galaxy source. We assume that mass

densities are low enough to treat the gravitational field of the matter in the image plane in the Newtonian approximation of general relativity (GR). In this case the metric is fully characterized by the lens' gravitational potential ψ . We also adopt the thin lens approximation, which assumes all the lens mass lies in a single *image plane* and all the source light is emitted from a *source plane*.

We will be using $\vec{\xi} \equiv (\xi_x, \xi_y)$ and $\vec{x} \equiv (x, y)$ as two-dimensional angular coordinates in the image and source planes respectively, and use z to indicate distances along the orthogonal dimension. Since the image plane covers a small angular patch of the sky and the lensing deflections are small in the Newtonian limit, the coordinate system can be treated as Cartesian.

The configuration, then, is determined by two fields: the distribution of surface brightness in the source plane, $\beta(\vec{x})$, and the distribution of mass in the image plane, described by the projected potential:

$$\Psi(\vec{\xi}) \equiv \int_{-\infty}^{\infty} \psi(\xi_x, \xi_y, z) dz \quad (1.1)$$

Under these assumptions, the source-plane coordinate to which a light ray through the image plane traces back is given by the simple *lens equation*

$$\vec{x} = \vec{\xi} - \vec{\alpha}(\vec{\xi}). \quad (1.2)$$

The displacement field $\vec{\alpha}$ is determined by the projected potential:

$$\vec{\alpha}(\vec{\xi}) = \frac{D_{LS}}{D_S} \frac{2}{c^2} \frac{\vec{\nabla}_{\vec{\xi}} \Psi(\vec{\xi})}{D_L}, \quad (1.3)$$

where the gradient is taken in the image plane and should have dimensions of inverse length, whence the introduction of the observer-lens angular diameter distance D_L . The expression also involves the angular diameter distances D_{LS} (from the lens to the source), and D_S (from the observer to the source).³ We illustrate in [Figure 1.1](#) the geometry of the system.

From a computational standpoint it is more convenient to represent [Equation 1.3](#) as an integral. Using the Poisson equation $\nabla_2 \Psi = 4\pi G \rho$ and setting appropriate boundary conditions:

$$\vec{\alpha}(\vec{\xi}) = \frac{4G}{c^2} \frac{D_{LS}}{D_L D_S} \int \Sigma(\vec{\xi}') \frac{\vec{\xi} - \vec{\xi}'}{|\vec{\xi} - \vec{\xi}'|^2} d(D_L \vec{\xi}'), \quad (1.4)$$

where the integral is performed over the image plane, and where c is the speed of light and G is the gravitational constant.

³We compute angular diameter distances with `astropy` [87, 88] using the flat cosmology from Planck [89].

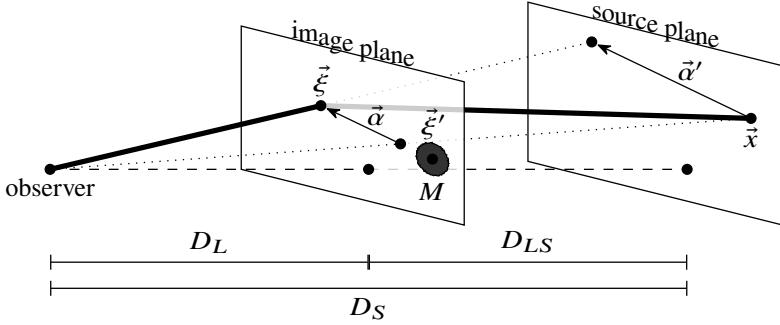


Figure 1.1: Thin lens geometry. The lensing mass M located at $\vec{\xi}'$ bends the light ray (thick line) emanating from the point \vec{x} , so that to the observer it looks like it is coming from the direction of $\vec{\xi}$. General relativity predicts the deflection angle $\vec{\alpha}'$ as viewed from the image plane based on the mass M and the impact parameter $\vec{b} = \vec{\xi} - \vec{\xi}'$. It then has to be rescaled by $\frac{D_{LS}}{D_S}$ to obtain the displacement field $\vec{\alpha}$ (as viewed by the observer) for use in [Equation 1.2](#). Angles are all assumed to be small enough that they can be used for Euclidean calculations. The dashed line is the optical axis perpendicular to the planes and connects the origins of the coordinate systems for each plane. The figure is a reproduction of Figure 2 in [90].

Here Σ is the projected (surface) mass density related to the 3D mass density ρ by an integration along the coordinate perpendicular to the lens plane z :

$$\Sigma(\vec{\xi}) \equiv \int_{-\infty}^{\infty} \rho(\xi_x, \xi_y, z) dz, \quad (1.5)$$

similarly to the expression for the projected potential.

The expression in [Equation 1.4](#) can be simplified by introducing the convergence κ in terms of the critical surface density Σ_{cr} :

$$\kappa(\vec{\xi}) = \frac{\Sigma(\vec{\xi})}{\Sigma_{\text{cr}}}, \quad \Sigma_{\text{cr}} \equiv \frac{c^2}{4\pi G D_L D_{LS}}. \quad (1.6)$$

If $\kappa > 1$, a lens can form multiple images [86, Section 2.6]. The threshold for this to happen is set by the critical surface density, and it is evident from its expression that for a fixed distance to the lens (D_L) further away sources (D_S) are lensed more easily, since they require smaller deflections.

Furthermore, it can be shown that the convergence κ is related to the isotropic part of the trace of the Jacobian of the lensing transformation through

$$\frac{d\vec{x}}{d\vec{\xi}} = \begin{pmatrix} 1 - \kappa & 0 \\ 0 & 1 - \kappa \end{pmatrix} - \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix}, \quad (1.7)$$

where the second part is the shear (discussed more in details in [subsection 1.2.3](#)). Another important quantity is the inverse of the Jacobian's determinant, the magnification

$$\mathcal{M} \equiv \left| \frac{d\vec{\xi}}{d\vec{x}} \right| = [(1 - \kappa)^2 - (\gamma_1^2 + \gamma_2^2)]^{-1}, \quad (1.8)$$

which is a measure of how much the solid angle spanned by the source is enlarged, or equivalently, of how gravitational focusing directs a larger fraction of the energy radiated by the source to the observer [91]. Thus, when the convergence κ and the shear γ are equal or greater than unity, sources are *strongly* magnified, and we can use the magnitude of the magnification to differentiate between *strong* from *weak* lensing regimes.

While lensing does change the apparent solid angle of a source, it is worth noting that it conserves energy, since it merely alters the trajectories of photons rather than creating or destroying them. As a result, the surface brightness $\beta(\vec{\xi})$ in the image plane is equal to the surface brightness at the point to which it traces back in the source plane (assuming one isolated source):

$$\beta(\vec{\xi}) = \beta(\vec{x}(\vec{\xi})). \quad (1.9)$$

Given the physics of strong lensing, in order to fully specify a strong-lensing model we then need two main ingredients: the lens model, which describes the total mass distribution of the lens, and the source model, which describes the surface brightness profile of the background source. It is common to split the lens model into a macroscopic smooth component (main lens and external shear) and a substructure⁴ component, due to subhalos and line-of-sight halos. Each lensing ingredient can be directly superimposed by summing their respective displacement fields in the lens plane:

$$\vec{\alpha} = \vec{\alpha}_{\text{lens}} + \vec{\alpha}_{\text{ext}} + \sum_{i=1}^{N_{\text{sub}}} \vec{\alpha}_{\text{sub},i} + \sum_{i=1}^{N_{\text{los}}} \vec{\alpha}_{\text{los},i}. \quad (1.10)$$

In the following sections, we will describe each component of the model we use to simulate mock images of gravitational lenses: the source, the main lens, the dark matter perturbers, and instrumental noise.

⁴Throughout this thesis, we use the terms ‘small-scale structures’, ‘substructures’, and ‘low-mass halos’ when considering both subhalos of the main lens and line-of-sight halos.

1.2.2 Source model

To model the surface brightness of the source galaxy, we adopt the widely-used Sérsic profile [92]. The surface brightness distribution is parameterized by

$$\beta(\vec{x}) = I_e \exp \left\{ -k_n \left[\left(\frac{R(\vec{x})}{r_e} \right)^{1/n} - 1 \right] \right\}, \quad (1.11)$$

where I_e is the surface intensity at the half-light radius r_e . The radial parameter $r(\vec{x}) = \sqrt{r_x^2 + r_y^2}$ is the length of the elliptical radial coordinate vector

$$\begin{pmatrix} r_x \\ r_y \end{pmatrix} = \begin{pmatrix} \sqrt{q_s} & 0 \\ 0 & 1/\sqrt{q_s} \end{pmatrix} \begin{pmatrix} \cos \phi_s & \sin \phi_s \\ -\sin \phi_s & \cos \phi_s \end{pmatrix} \begin{pmatrix} x - x_{0,s} \\ y - y_{0,s} \end{pmatrix}, \quad (1.12)$$

which depends on the source's position angle ϕ_s , axis ratio q_s , and center of light position $(x_{0,s}, y_{0,s})$.

The normalization k_n is related to the index n by an implicit transcendental equation in terms of the complete and lower incomplete gamma functions $2\gamma(2n, k_n) = \Gamma(2n)$. We use the expansion in series from [93],

$$k_n \approx 2n - \frac{1}{3} + \frac{4}{405n} + \frac{46}{25515n^2} + \frac{131}{1148175n^3} - \frac{2194697}{30690717750n^4}. \quad (1.13)$$

valid over a wide range of indices, $n > 0.36$. For typical galaxies $1/2 < n < 10$.

Therefore, our source model is parametrised by seven variables that we collect in the vector $\boldsymbol{\theta}_s \equiv \{I_e, r_e, x_{0,s}, y_{0,s}, q_s, \phi_s, n\}$. We fix the source's redshift to $z_{\text{source}} = 2$.

1.2.3 Main lens model

We adopt the singular power law ellipsoid (SPLE) model for the main lens galaxy, which is capable of modeling the gravitational potentials of strong lenses to near the percent level [94]. Furthermore, the SPLE model has been shown to be a more adequate representation of the combined DM and baryon mass distribution in the inner regions of galaxies, to which strong lensing is most sensitive [94]. The SPLE deflection field can be expressed in closed-form as a complex field $\alpha = \alpha_x + i\alpha_y$ [95, 96]:

$$\begin{aligned} \vec{\alpha}^{\text{SPLE}}(\vec{\xi}) = & \theta_E \frac{2q_l^{1/2}}{1+q_l} \left(\frac{\theta_E}{r} \right)^{\gamma-2} e^{i\varphi} \\ & \cdot {}_2F_1 \left(1, \frac{\gamma-1}{2}, \frac{5-\gamma}{2}, -\frac{1-q_l}{1+q_l} e^{2i\varphi} \right). \end{aligned} \quad (1.14)$$

Here (r, φ) are elliptical coordinates, related to the Cartesian coordinates $\vec{\xi}$ through a transformation parametrized by the lens' orientation ϕ_l , axis ratio q_l and position

$(\xi_{x,0}, \xi_{y,0})$:

$$\begin{pmatrix} r_x \\ r_y \end{pmatrix} = \begin{pmatrix} q_1^{1/2} & 0 \\ 0 & q_1^{-1/2} \end{pmatrix} \begin{pmatrix} \cos \phi_l & \sin \phi_l \\ -\sin \phi_l & \cos \phi_l \end{pmatrix} \begin{pmatrix} \xi_x - \xi_{x,0} \\ \xi_y - \xi_{y,0} \end{pmatrix}, \quad (1.15)$$

$$\tan \varphi = \frac{r_y}{r_x}. \quad (1.16)$$

In the circular ($q = 1$) isothermal ($\gamma = 2$) case this reduces to $\alpha = \theta_E e^{i\varphi}$, i.e. is an isotropic constant, as is well-known. Since the hypergeometric function ${}_2F_1$, however, is not implemented in PyTorch, we instead pretabulate its value as a function of ϕ_l , q_1 and γ and interpolate at runtime, as described in [97]. This results in very fast code with negligible output degradation.

The slope γ has a complicated degeneracy with the size of the source [98, 99]. Roughly, larger γ values cause the spatial scale of the source to increase [100, sec. 3.3]. For simplicity we fix $\gamma = 2.1$. In principle, inferring the slope is possible, but it requires more training data and leads to increased uncertainties in both lens and source parameters.

We also assume the lens galaxy's light has been perfectly subtracted, and fix its redshift to $z_{\text{lens}} = 0.5$.

To account for the weak lensing due to large-scale structure located along the line of sight to the source, we also include an external shear component, which is constant across the image plane:

$$\vec{\alpha}^{\text{shear}}(\vec{\xi}) = \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix} \vec{\xi}. \quad (1.17)$$

Our main lens model thus has seven parameters: the SPLE parameters $(\xi_{x,0}, \xi_{y,0}, \phi_l, q_1, \theta_E)$ and the external shear parameters (γ_1, γ_2) , which we denote collectively with $\boldsymbol{\theta}_l$.

1.2.4 Small-scale structures model

DM substructures can be divided into two categories: subhalos that orbit around the main halo at the lens redshift, and LOS halos distributed between the source and the observer. LOS halos are a more direct probe of free-streaming-induced small-scale structure suppression, because they are less affected by baryonic processes and environmental effects, such tidal stripping interactions with the main halo [101]. For this reason and the fact that they are expected to be more abundant than subhalos in a lensing system [101, 102], it is very important to model them as well, in order to correctly estimate the collective effects of all substructures on the lensing image.

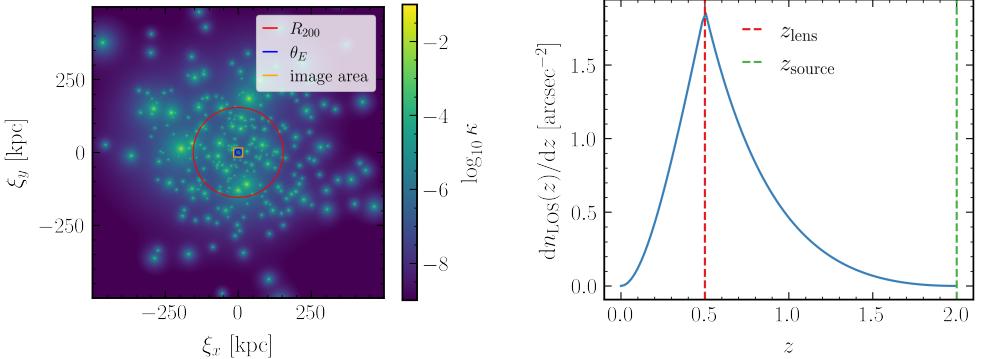


Figure 1.2: *Left panel:* Convergence map for a CDM subhalo population in the adopted mass range. The convergence map shows how the deflecting mass from all the subhalo lenses is distributed. The full map size is 1×1 Mpc. We mark in red the virial radius of the main lens halo, in blue its Einstein radius, and in orange the 5×5 arcsec lensing image area. *Right panel:* LOS halos distribution in redshift for our source and lens redshifts configuration, $z_{\text{lens}} = 0.5$ and $z_{\text{source}} = 2$.

Density profile

To model the density profiles of small-scale DM halos we adopt the smoothly truncated universal 3D mass density profile from [103]:

$$\rho_{\text{tNFW}}(r) = \frac{\rho_s}{r/r_s(1+r/r_s)^2} \frac{1}{1+(r/r_t)^2}. \quad (1.18)$$

Here r is the three-dimensional distance from the center of the halo, ρ_s and r_s are respectively the scale density and scale radius that specify a Navarro-Frenk-White (NFW) profile [104], and $r_t \equiv \tau r_s$ is the tidal truncation radius that depends on the history of the subhalo. Typical values of the truncation scale τ range from 4 – 10 for spherically symmetric lenses [30, 56]; we fix $\tau = 6$ for simplicity. Compared to the standard NFW form, which has an infinite total mass, the truncated NFW contains an additional truncation term that makes the profile decay as r^{-5} for large radii, resulting in a finite total mass given by:

$$m_\tau = 4\pi\rho_s r_s^3 \frac{\tau^2}{(\tau^2+1)^2} [(\tau^2-1)\ln\tau + \tau\pi - (\tau^2+1)]. \quad (1.19)$$

With a fixed truncation scale, the truncated NFW profile is fully determined by the same parameters that determine the NFW profile: the virial mass $m_{200,\text{sub}}$ ⁵ and the

⁵We parameterize subhalos by what would be their mass up to the virial radius r_{200} using the untruncated profile, with the same central density ρ_s and scale radius r_s as the truncated one. This

concentration $c = r_{200}/r_s$ of the halo. The latter measures how concentrated the mass of a halo is and fixes the density normalization; in principle, it varies from one subhalo to the next and shows dependencies on mass and redshift of the main halo. Here, instead of adopting a concentration-mass relation, we fix $c_{200} = 15$, which is roughly the average value for perturbers in the mass range $10^7 - 10^{10} M_\odot$ [105, Fig. 7]. We anticipate that accounting for scatter in the mass-concentration relation might actually improve our ability to measure subhalos' parameters as higher concentrations lead to substantially stronger lensing signals [106]. The equations for calculating the displacement field of a truncated NFW halo, given its mass and position, are fully elaborated by [103, Appendix A].

To generate perturber populations, we must choose values for their density normalizations and scale radii. Since simulation studies typically measure the halo mass $m_{200,\text{sub}}$ and the concentration c , it is more convenient to sample populations from distributions over these parameters. These variables can then be mapped to the parametrization above via

$$r_s = \frac{1}{c} \left[\frac{3m_{200,\text{sub}}}{4\pi 200\rho_{\text{cr}}(z_{\text{lens}})} \right]^{1/3}, \quad (1.20)$$

$$\rho_s = \rho_{\text{cr}}(z_{\text{lens}}) \frac{1}{3} \frac{c^3}{\log(1+c) - c/(1+c)}. \quad (1.21)$$

For simplicity, we model LOS halos using exactly the same profile even though they typically have not undergone tidal stripping.

The parameters of an individual subhalo which are not fixed are thus $\boldsymbol{\eta}_{\text{sub}} \equiv (\mathbf{x}_{\text{sub}}, \mathbf{y}_{\text{sub}}, m_{200,\text{sub}})$, where the second and third components are the projected position of the subhalo. In the case of LOS halos, the parameter set also includes the redshift z_{los} . In the following sections, we will use arrows to denote vectors, e.g. $\vec{m}_{200,\text{sub}}$ is an ordered set of masses, one for each simulated subhalo, and bold letters with an arrow to indicate arrays of vectors, e.g. $\vec{\eta}_{\text{sub}} \equiv (\mathbf{x}_{\text{sub}}, \mathbf{y}_{\text{sub}}, \mathbf{m}_{200,\text{sub}})$ is an ordered set of mass and positions in the lens plane, one for each simulated subhalo, and equivalently for LOS halos $\vec{\eta}_{\text{los}} \equiv (\mathbf{x}_{\text{los}}, \mathbf{y}_{\text{los}}, \mathbf{z}_{\text{los}}, \mathbf{m}_{200,\text{los}})$. We collectively define the parameters of the full population of perturbers, subhalos and LOS halos, as $\vec{\eta}_p$. In the next two subsections, we describe how we sample these parameters.

Generating subhalos

We sample subhalo masses from the CDM mass function of [107]:

$$\frac{1}{m_{200}} \frac{dn_{\text{sub}}(m_{200}, z)}{d \log m_{200}} = A_M (1+z)^{1/2} m_{200}^\alpha \exp \left[-\beta \left(\frac{m_{200}}{M} \right)^3 \right], \quad (1.22)$$

is defined as the mass of the halo enclosed in a sphere where the untruncated halo's average density is 200 times the critical density.

where M ⁶ is the mass of the main lensing halo and m_{200} the subhalo mass. The free parameters in this function were fit to hydrodynamical cosmological simulations that included baryons in [108]. In particular, we use the fits to EAGLE, which give $\alpha = 0.85$ (given in the text) and $(A_M, \beta) = (2.4 \times 10^{-4} M_\odot^{\alpha-1}, 300)$ (extracted from their figures). The expected number of subhalos in a given mass interval for the lens halo system can be computed by integrating the mass function over that interval.

The spatial distribution of subhalos has been shown to follow an Einasto profile [109]. However, since the virial radius of a typical main lens halo is much larger than its Einstein radius, and hence, than the image plane, we approximate the distribution to be uniform in the lensing image area. Still, we derive the total number of expected subhalos within the image, \bar{n}_{sub} , via the Einasto fit of [108]. We find that on average $\bar{n}_{\text{sub}} = 4$ subhalos fall within the lensing image area in our adopted lensing configuration and mass range. Thereafter we generate the subhalo population by sampling the number of subhalos from Poisson(\bar{n}_{sub}), drawing their masses from the subhalo mass function (Equation 1.22) and sampling their projected positions uniformly over the lens plane. Since the vast majority of subhalos fall outside the lens plane, we expect their lensing effect to be mostly degenerate with external shear, and thus do not simulate them. In the left panel of Figure 1.2 we show the convergence map for one realization of our subhalo population.

Generating line-of-sight halos

As described in [54, Figure 3], we first compute the average number of LOS halos in the double-pyramid geometry connecting the observer, lens-plane and source. For each simulation we sample the number of LOS halos from Poisson(\bar{n}_{los}). We then sample their redshifts and projected positions uniformly over the double-pyramid region and draw their masses from the mass function in [110], assuming an overdensity with respect to the critical density of the universe at the epoch of analysis of $\Delta = 200$.

More in details, we infer the number of detectable LOS halos by integrating their mass function in the mass range adopted for the analysis and within the double-cone volume

$$\bar{n}_{\text{los}} = \int_0^{z_{\text{src}}} \int_{m_{200,\min}}^{m_{200,\max}} n_{\text{los}}(m_{200}, z) dm_{200} \frac{dV}{dz} dz. \quad (1.23)$$

On average we get $\bar{n}_{\text{los}} = 260$ LOS halos projected in our lens plane. Similarly to what we do with the subhalo population, when generating simulated images, we draw the

⁶The total mass of the lens galaxy is described by the Einstein radius of the system, a very well-constrained parameter in lensing inference analyses. For the purpose of describing subhalos, we need to be able to map the measured properties of the lens (the Einstein radius θ_E) onto the properties of the host halo (the mass M). For simplicity, we compute the mass of the host halo transforming the Einstein radius distance measure into a mass measure. We would like to point out a similar approach from [74], where they relate the central velocity dispersion of a singular isothermal ellipsoid lens mass distribution profile to the virial mass of the host halo.

number of LOS halos from Poisson(\bar{n}_{los}), we then sample their masses and redshift from the [110] halo mass function and sample their projected positions uniformly over the lensing image area. As with subhalos, we ignore any LOS halos lying outside the double pyramid volume. In right panel of [Figure 1.2](#) we show the distribution of LOS halos in redshift for our lens and source redshifts configuration.

To avoid expensive iterative ray-tracing through the lens planes of each LOS halo, we project them as effective subhalos into the lens plane, using the relations derived in [54] to rescale their scale radii and masses. Following [54], LOS halos at comoving distance χ can be treated as subhalos on the main-lens plane with an effective projected mass density given by:

$$\Sigma_{\chi,\text{eff}}(D_L \vec{x}; m_{200}, r_s, \tau) = \Sigma(D_L \vec{x}; m_{200,\text{eff}}, r_{s,\text{eff}}, \tau). \quad (1.24)$$

The effective scale radius $r_{s,\text{eff}}$ and mass $m_{200,\text{eff}}$ are respectively

$$r_{s,\text{eff}} = \frac{D_L}{g(\chi) D_\chi} r_s, \quad (1.25)$$

and

$$m_{200,\text{eff}} = f(\chi) \frac{\Sigma_{\text{cr},l}}{\Sigma_{\text{cr},\chi}} \left(\frac{D_l}{g(\chi) D_\chi} \right)^2 m_{200}. \quad (1.26)$$

The piecewise functions $f(\chi)$ and $g(\chi)$ are:

$$f(\chi) = \begin{cases} 1 - \beta_{\chi l} & \chi \leq \chi_L \\ 1 - \beta_{l \chi} & \chi > \chi_L \end{cases}, \quad (1.27)$$

and

$$g(\chi) = \begin{cases} 1 & \chi \leq \chi_L \\ 1 - \beta_{l \chi} & \chi > \chi_L \end{cases}, \quad (1.28)$$

with $\beta_{ij} = \frac{D_{ij} D_s}{D_j D_{is}}$, where D_i is the angular diameter distance from the observer to plane i, and D_{ij} is the angular diameter distance from lens plane i to lens plane j, and χ_L is the comoving distance to the main-lens plane. We have also introduced the critical surface density at plane i

$$\Sigma_{\text{cr},i} \equiv \frac{c^2 D_s}{4\pi G D_i D_{is}}. \quad (1.29)$$

It should be noted that effective convergence methods, like the one we adopt [54], do not fully capture the subtleties and degeneracies of multi-plane lens analysis, by disregarding how, when a DM small-scale halo is not in the lens plane, the lens mass

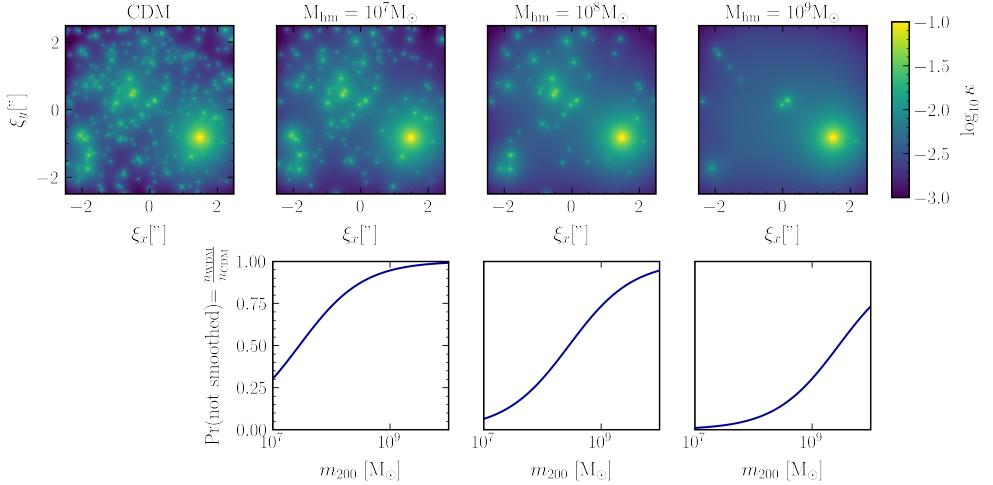


Figure 1.3: *Top:* Convergence maps for a population of LOS halos with masses sampled from the CDM HMF in the adopted mass range (??) and projected in the lens plane following the prescription from [54]. In the second to fourth columns, we imitate the effect of WDM with different cutoff masses (as labeled in the titles) via our smoothing scheme. *Bottom:* We show the probability with which LOS halos do not get smoothed, equal to the ratio between the WDM and the CDM HMF (Equation 1.33). The smoothing is stochastic, so for each realization of the smoothing different halos are smoothed.

model can absorb its lensing signal [102, 106, 111]. The omission of this effect may lead to an overestimate of the LOS halos contribution and needs to be addressed before this technique can be safely used for the analysis of real data.

For LOS halos we adopt the same concentration and truncation scale values used for subhalos.

1.2.5 Modelling free-streaming effects in WDM

Among DM models, warm dark matter (WDM) particle candidates, sterile neutrinos [112] and gravitinos [113], are well-motivated from a particle physics perspective. In WDM models DM particles have non-negligible thermal velocities that allow them to free-stream out of density perturbations, effectively preventing small-scale structure formation. The suppression scale at which this happens depends on model parameters and is parametrised by the half-mode mass M_{hm} in the HMF. Therefore, one of the viable way to discriminate between CDM and alternative WDM models is to constrain the low-mass end of the HMF by measuring the half-mode mass M_{hm} .

The free-streaming effects of WDM are well described in terms of the half-mode wavelength λ_{hm} , which corresponds to the scale at which the DM transfer function falls to half the CDM transfer function. We can define the half-mode mass as the mass contained within a radius of the half-mode wavelength:

$$M_{\text{hm}} = \frac{4\pi\Omega_M\rho_{\text{crit}}}{3} \left(\frac{\lambda_{\text{hm}}}{2} \right)^3, \quad (1.30)$$

where Ω_M is the matter density parameter and ρ_{crit} is the critical density of the universe. Following [114], the half-mode wavelength,

$$\lambda_{\text{hm}} = 2\pi\alpha_{\text{hm}} \left(2^{\nu/5} - 1 \right)^{-1/(2\nu)}, \quad (1.31)$$

is the scale below which the initial density perturbations are completely erased, with $\nu = 1.12$ and, assuming that all DM is warm,

$$\alpha_{\text{hm}} = 0.049 \left(\frac{m_{\text{WDM}}}{\text{keV}} \right)^{-1.11} \left(\frac{\Omega_{\text{DM}}}{0.025} \right)^{0.11} \left(\frac{h}{0.7} \right)^{1.22} h^{-1}\text{Mpc}. \quad (1.32)$$

We then have a one-to-one mapping between the mass of the WDM particle and the half-mode mass. For strong lensing, the half-mode mass can be thought of as an effective cutoff mass below which the DM mass function is strongly suppressed. To model this suppression in the WDM mass function we adopt for both subhalos and LOS halos the functional form from [115]:

$$\frac{n_{\text{WDM}}}{n_{\text{CDM}}} = \left(1 + \left(\alpha \frac{M_{\text{hm}}}{m_{200}} \right)^\beta \right)^\gamma, \quad (1.33)$$

with best fit parameters $\alpha = 4.2$, $\beta = 2.5$, and $\gamma = 0.2$ for subhalos, $\alpha = 2.3$, $\beta = 0.8$, and $\gamma = 1$ for central halos.

Smoothing substructures

The observational signature of WDM is, thus, the absence of small-scale structures. However, in the current parameterization, this is accompanied by the removal of the corresponding mass enclosed in them, whereas in reality the mass will still be present but will be diffused throughout the smooth main halo. This effect is manifested in a correlation between the half-mode mass and the main-halo Einstein radius: suppressing more substructure leads to an increase in the inferred Einstein radius since the total mass of the system (within the image) is tightly constrained by the size of the observed ring (or arcs).

We introduce a prescription for dealing with this degeneracy, which well captures the physical reality of structure suppression due to free streaming. Halos that should

be suppressed are not present because the DM particles that should make them up are freely streaming, and their mass is therefore more diluted throughout the main halo. Therefore, rather than removing or adding substructures as a response to a changing cutoff, we still sample substructures from the CDM mass function, but we smooth the displacement field generated by halos that should be suppressed based on the aforementioned prescription by [115] to hide their lensing signature. In other words, each sampled small-scale halo has a probability equal to the ratio between the WDM and the CDM HMF (Equation 1.33) of not being smoothed.

We then effect the smoothing by convolving the deflection field of each individual sub-/LOS halo with a radially-symmetric filter

$$f \propto 1 - \exp\left(-\left(\frac{r}{r_{\text{smooth}}}\right)^{n_{\text{smooth}}}\right). \quad (1.34)$$

This filtering preserves the far-field lensing signature of the halo, which is only determined by its total mass. By default, we choose the smoothing scale to be equal to the halo virial radius: $r_{\text{smooth}} = r_{200}$, and the smoothing exponent $n_{\text{smooth}} = 2$.

In the top row of Figure 1.3 we visualize the convergence maps in the lens plane for the same realization of LOS halos drawn from CDM distributions (panel 1), and with different cutoff masses implemented with our smoothing scheme (panels 2-4). In the bottom row, we show how we decide to smooth the lensing signature of certain halos based on the ratio between the WDM and the CDM HMF (Equation 1.33).

1.2.6 Instrumental effects

We generate mock data with comparable quality to HST observations. All images are $5'' \times 5''$ with $0.05''$ resolution (100×100 pixels). In our simulations, we do not include a point-spread function (PSF) for simplicity, but this component cannot be disregarded in real data analysis. To account for the fact that each pixel in the image corresponds to a finite collecting region in the sky, we generate our images at a resolution eight times higher than the target resolution and downsample. In experiments we found that neglecting this effect can have a significant impact on inference results. Lastly, we add Gaussian and uncorrelated pixel noise to our observations such that the brightest pixels are approximately 30 times the noise level (after downsampling), representative of HST data.

1.3 Methodology

1.4 Measuring single subhalo parameters

1.4.1 Mock data generation

Table 1.1

1.4.2 Inference strategy

[TO DO: Rephrase] We now apply TMNRE to three different substructure lensing problems of increasing complexity. For all tasks we use the same general ratio estimator architecture. It consists of an initial compression network that maps the 100×100 pixel images into a feature vector. This feature vector is concatenated to $\vec{\eta}_{\text{sub}}$ (and separately to η_{src} and η_{lens} for tasks where they are also inferred). The vector is then passed to a multilayer perceptron (MLP) which outputs an estimate of the 2D and 1D marginal likelihood-to-evidence ratios for $(x_{\text{sub}}, y_{\text{sub}})$ and $\log_{10} m_{\text{sub}}/\text{M}_\odot$ respectively (with separate MLPs used to estimate the 1D ratios for η_{src} and η_{lens}).

For each ratio estimator we begin the first training round with 10 000 training examples. We then truncate each parameter’s prior. If none of the truncated priors shrank by at least 20%, we increase the number of training examples by a factor of 1.5 for the next inference round. A fresh network is then trained using simulations drawn from the truncated prior. Convergence of the ratio estimator is declared after five such consecutive increases in the training set size. For tasks in which we must infer the macromodel parameters we first train the macromodel ratio estimator using this procedure and use the resulting truncated priors to generate training data for the subhalo ratio estimators using the same training procedure. We use the implementation of TMNRE in `swyft`⁷ [116], which is built on PyTorch and `pytorch-lightning`⁸.

The training data for our ratio estimators differs in important ways from typical datasets studied by machine learning researchers, making the choice of a good compression network an interesting challenge. Consider, for example, the machine learning problem of classifying the content of natural images. Natural images are distinguished by a hierarchy of visual features at different scales (for example, small-scale features such as textures and edges which comprise large-scale features like the head of an animal or part of an object). A good image classifier should be translation-invariant, producing the same output regardless of the position of an image’s contents.

⁷<https://github.com/undark-lab/swyft/>

⁸<https://www.pytorchlightning.ai/>

Since the deep convolutional neural network (CNN) architecture has an inductive bias towards learning a hierarchy of features and are translation invariant, CNNs are widely used in computer vision.

The training data for our ratio estimators does not share these features. Different perturber configurations produce images with slightly different relationships between the multiple images of the source galaxy. The variations between images lie near the Einstein ring, and do not show the same rich hierarchical structure of natural images. This means that inductive biases of CNNs are not necessarily beneficial in the context of substructure lensing.

In our experiments, we used CNNs in the ratio estimators for the macromodel parameters, finding their performance to be adequate. However, we found they produced much too wide 2D marginals for the position of a subhalo. Instead, we found the MLP Mixer [117] to work well.⁹ Roughly, the MLP Mixer splits the image into patches, stacks the patches and passes each pixel in the stack through an MLP, acting as a dilated convolution. Another MLP is then applied along the channel dimension of the mixed patches, and the process is iterated. The MLP Mixer thus directly examines the relationships between pixels in disparate parts of the image, which is exactly how the properties of subhalos are imprinted. We expect that other architectures that split the image into patches such as Vision Transformer (ViT) [118] could work well for the compression network, though ViT is known to require large amounts of training data.

The architectures of our macromodel and subhalo compression networks are given in [section 1.A](#). While we did not perform a full hyperparameter exploration, we found the batchnorm layers to be crucial for stable training of the CNN used for the macromodel ratio estimator. Since our images are roughly one-quarter the area of the images studied in the paper introducing MLP Mixer, we use a substantially smaller model than they suggest. Using dropout in the MLP Mixer and classifier MLPs improved performance. Varying the number of hidden layers and their size in the classifiers had little impact.

We used the Adam optimizer with an initial learning rate of 6×10^{-3} for the macromodel ratio estimator and 4×10^{-4} for the subhalo ratio estimator (found through a learning rate test) and a batch size of 64. The learning rate was reduced by a factor of 0.1 whenever the validation loss plateaued for 3 epochs. Training was run for no longer than 30 epochs.

⁹The MLP Mixer implementation we use can be found at <https://github.com/lucidrains/mlp-mixer-pytorch>.

	Parameter	True value	Initial prior	First inferred in
Subhalo	$x_{\text{sub}} [\text{''}]$	-1.1	$\mathcal{U}(-2.5, 2.5)$	subsection 1.4.3
	$y_{\text{sub}} [\text{''}]$	-1.1	$\mathcal{U}(-2.5, 2.5)$	subsection 1.4.3
	$\log_{10} m_{\text{sub}} / M_{\odot}$	9.5	$\mathcal{U}(8, 10.5)$	subsection 1.4.4
SPLE	$x_{\text{lens}} [\text{''}]$	-0.05	$\mathcal{U}(-0.2, 0.2)$	
	$y_{\text{lens}} [\text{''}]$	0.1	$\mathcal{U}(-0.2, 0.2)$	
	$\varphi_{\text{lens}} [{}^{\circ}]$	1	$\mathcal{U}(0.5, 1.5)$	
	q_{lens}	0.75	$\mathcal{U}(0.5, 1)$	subsection 1.4.4
	γ	2.1	—	
	$r_{\text{ein}} [\text{''}]$	1.5	$\mathcal{U}(1, 2)$	
Shear	γ_1	0.005	$\mathcal{U}(-0.5, 0.5)$	
	γ_2	-0.010	$\mathcal{U}(-0.5, 0.5)$	subsection 1.4.4
Source	$x_{\text{src}} [\text{''}]$	0	$\mathcal{U}(-0.2, 0.2)$	
	$y_{\text{src}} [\text{''}]$	0	$\mathcal{U}(-0.2, 0.2)$	
	$\varphi_{\text{src}} [{}^{\circ}]$	0.75	$\mathcal{U}(0.5, 1.25)$	
	q_{src}	0.5	$\mathcal{U}(0.2, 0.8)$	subsection 1.4.4
	n	2.3	$\mathcal{U}(1.5, 3)$	
	$r_e [\text{''}]$	2.0	$\mathcal{U}(0.5, 3)$	
	I_e	0.6	$\mathcal{U}(0.1, 2)$	

Table 1.1: True subhalo and macromodel parameter values and priors used in the first TMNRE inference round in our three inference tasks. The last column references the first section in which the indicated parameter is inferred rather than being fixed to its true value. The slope of the main lens is fixed to 2.1, as explained in subsection 1.2.3. The main lens and source redshifts are set to $z_{\text{lens}} = 0.5$ and $z_{\text{src}} = 2$ respectively. For the analysis in subsection 1.4.5 involving a population of light perturbers, we sample the number of LOS and subhalos from Poisson distributions with means $\bar{n}_{\text{los}} = 265.6$ and $\bar{n}_{\text{sub}} = 3.1$ respectively, and restrict their masses to the range $10^7 - 10^8 M_{\odot}$. The halo mass functions and redshift distributions are described in detail in subsection 1.2.4. For all perturbers we fix the concentration to $c = 15$ and truncation scale $\tau = r_t/r_s = 6$.

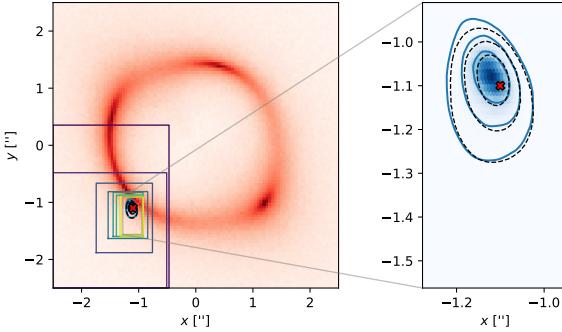


Figure 1.4: Validation of TMNRE through inference of the position of a subhalo, with macromodel parameters fixed to their true values and the subhalo’s mass fixed to $10^9 M_\odot$. The observation is shown in the left panel. The blue and dashed black contours correspond to the posterior inferred with TMNRE and computed analytically respectively, indicating the 68%, 95% and 99.7% credible regions. The red \times shows the subhalo’s true position. The blue through yellow boxes in the left panel show the ranges of the truncated prior based on the 1D marginals for the subhalo’s coordinates. The zoom-in on the right encompasses the range of the final truncated prior. The distorted blue hex-bin histogram shows the magnitude of the inferred posterior.

1.4.3 Subhalo position inference with fixed mass, source and lens

We first consider the case where the only free parameters in the lens are the position of a single $10^9 M_\odot$ subhalo, $\boldsymbol{\eta}_{\text{sub}} = (x_{\text{sub}}, y_{\text{sub}})$. The prior is taken to be uniform over the image plane (i.e., $\mathcal{U}(-2.5, 2.5)$ for both coordinates). The posterior for $\boldsymbol{\eta}_{\text{sub}}$ can then be computed analytically. Adopting a uniform prior over $\boldsymbol{\eta}_{\text{sub}}$ covering the image plane and using the fact the posterior is much narrower, we have

$$\log p(\boldsymbol{\eta}_{\text{sub}} | \mathbf{x}) \sim -\frac{1}{2} \sum_{i,j} \left(\frac{x_{ij} - f_{ij}(\boldsymbol{\eta}_{\text{sub}})}{\sigma_n} \right)^2, \quad (1.35)$$

where the sum runs over pixels and we dropped terms independent of $\boldsymbol{\eta}_{\text{sub}}$.

[Figure 1.4](#) shows the truncation regions for each round and compares the analytically-computed posterior with the posterior inferred using TMNRE. While the truncation regions and posterior estimates in early rounds are extremely broad compared to the analytically-computed posterior, TMNRE successfully identifies the region of the image containing the subhalo. After 10 inference rounds the truncation region stabilizes and the inferred posteriors agree well with the true ones for each

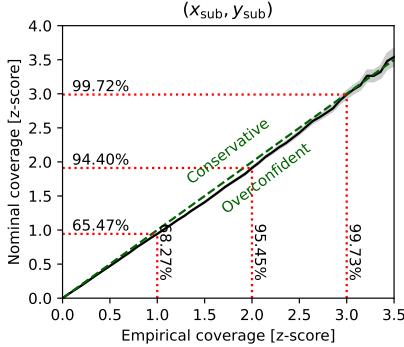


Figure 1.5: Coverage plot for inference task where only the subhalo’s position is free (see Figure 1.4), showing our ratio estimator produces posteriors of the correct size on average. In detail, the black curve shows the empirical versus nominal coverage, estimated by computing posteriors for 10 000 observations drawn from the final truncated prior. The statistical uncertainty of this estimate is plotted in grey; its derivation is explained in detail in [119]. For a perfectly-calibrated ratio estimator, the black curve would lie along the diagonal green dashed line. The red dashed lines indicate the empirical and nominal coverage of the $1 - 3\sigma$ credible regions.

coordinate. To complement this visual check we also check the coverage for samples from the final round of TMNRE in Figure 1.5. We find the empirical and nominal coverage to be in good agreement, with our ratio estimator very slightly underestimating the width of the posterior beyond the 95% confidence level.

Having validated TMNRE in this simple scenario, we now turn to more complex inference tasks where the posteriors of interest cannot be derived analytically.

1.4.4 Subhalo mass and position inference

Next we aim to infer the position and mass of a single subhalo, $\eta_{\text{sub}} = (m_{\text{sub}}, x_{\text{sub}}, y_{\text{sub}})$, in a system where the source and main lens parameters are also unknown. The priors for the 17 parameters of the model are given in Table 1.1. Due to the relatively low dimensionality, inference on this model is within the reach of likelihood-based tools such as MCMC or nested sampling. In addition, it can be implemented in a differentiable manner, making the application of methods such as Hamiltonian Monte Carlo (HMC) possible [97, 120]. Running such expensive scans is beyond the scope of this paper.

The final posteriors for the subhalo parameters are shown in Figure 1.6. The true values of all parameters fall within the $\sim 68\%$ credible intervals of the inferred

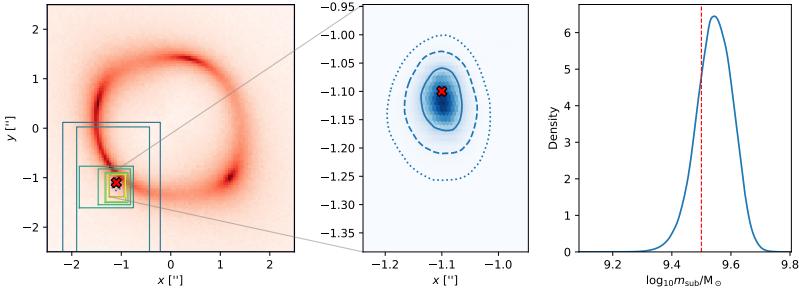


Figure 1.6: Marginal posteriors inferred with TMNRE for a subhalo’s 2D position (left and center) and mass (right) in a lens with unknown macromodel parameters. See the caption of Figure 1.4 for further details, though note we have instead used solid, dashed and dotted lines respectively to mark the 68%, 95% and 99.7% credible regions of the position posterior. The range of the x -axis in the right panel shows the final-round truncated prior for the subhalo’s \log_{10} -mass.

posteriors. We find the effect of the uncertain macromodel is not too strong (at least for this noise realization), with the size of the subhalo position posterior being comparable to what we found in the previous inference task. Figure 1.7 demonstrates that our ratio estimator has good coverage with respect to the constrained prior. In Figure 1.8 and Figure 1.9 we display the marginal posteriors and coverage plots for all 14 source and main lens parameters, which demonstrate they are well-calibrated.

1.4.5 Subhalo mass and position inference with a population of perturbers

For our final inference task we extend the previous one by aiming to infer the position and mass of a relatively heavy target subhalo while marginalizing over a population of lighter perturbers of unknown size. The priors for the perturber population are summarized in Table 1.1 and subsection 1.2.4. Our lensing images contain on average about 260 LOS halos and 3 subhalos in the lens plane. This means on average about 800 parameters are required to describe such images. Likelihood-based sampling of this high-dimensional, transdimensional posterior requires techniques such as reversible-jump MCMC [48, 49]. To marginalize over the perturber population with TMNRE, their parameters are sampled over during data generation but not passed to the ratio estimator.

Since the population of perturbers can contain a member with mass greater than our target subhalo, we need to make this inference task well-defined by “labeling” the

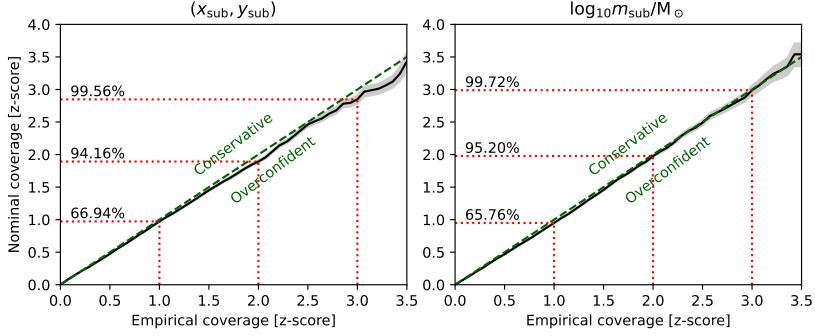


Figure 1.7: Coverage plots for subhalo position and mass ratio estimators learned from the observation in Figure 1.6. These again indicate the estimators’ credible regions are on average the correct size for observations drawn from the final-round truncated prior. See Figure 1.5 for an explanation of the format.

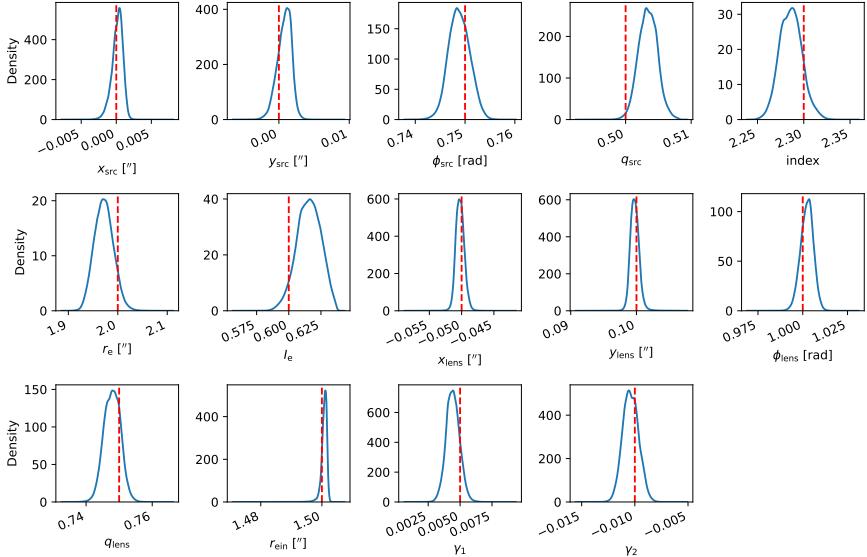


Figure 1.8: The 1D marginal posteriors for all macromodel parameters of the lensing system shown in Figure 1.6. The posteriors were computed using a CNN-based ratio estimator. The first seven panels correspond to the source parameters, the next five are for the main lens and the last two are for the external shear. All posteriors encompass the true parameter values (vertical red dashed lines) within the $\sim 2\sigma$ interval.

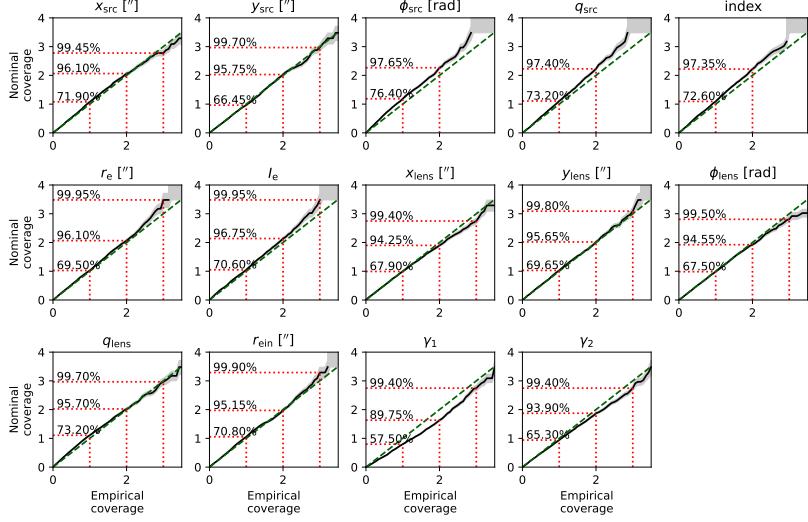


Figure 1.9: Coverage plots for the 1D marginal macromodel parameter posteriors of the lensing system from Figure 1.6, using the same format as in Figure 1.5. The posteriors generally have coverage, with a few being slightly conservative (ϕ_s , q_s and the source index) and the shear posteriors being slightly overconfident.

subhalos of interest. We accomplish this by making the perturber population lighter than the target subhalo, with mass restricted to the range $10^7 - 10^8 M_\odot$. We further assume the target subhalo has been localized to a $1.4'' \times 1.4''$ patch of the image around its true position.

The final-round inference results for η_{sub} plotted in Figure 1.10 show that inclusion of the perturber population has a substantial effect on the posteriors. The posterior for the subhalo's mass peaks around the true value, but has a long tail extending towards the lower boundary of the prior. This indicates we are only able to obtain an upper bound on the subhalo mass rather than a measurement, and cannot exclude the possibility its mass is the lowest value consistent with the prior. Having validated our analysis in simpler cases and checked our ratio estimator has good coverage, we conclude our marginal posteriors are in fact close to the true ones.

Our results are roughly in line with the image segmentation analysis of [66, 67], which found subhalos of mass above roughly $10^{8.5} M_\odot$ were resolvable in similar mock observations. In addition, while the 68% credible region for the subhalo's position contains its true position, the 95% and 99.7% credible regions cover nearly the whole prior region.

The posteriors for the source and lens parameters are shown in Figure 1.11. While

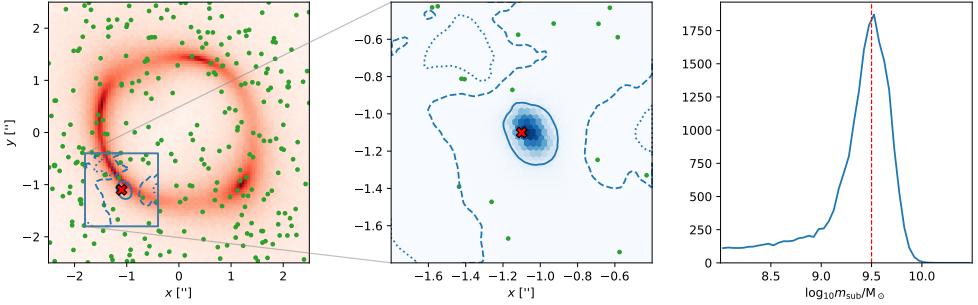


Figure 1.10: Subhalo position and mass posteriors obtained with TMNRE, now marginalizing over a population of $10^7 - 10^8 M_\odot$ LOS/subhalos (green dots) in addition to the unknown macromodel. See the caption of Figure 1.6 for details. The initial prior for the subhalo’s position is indicated by the blue box in the left panel. The range of the x -axis in the right panel shows the prior on the \log_{10} of its mass. For both the subhalo’s mass and position, the width of the inferred posteriors prevents TMNRE from truncating the priors.

some of the parameters’ posteriors have comparable widths to those found in the previous inference task (namely $\phi_{s/l}$, $q_{s/l}$, the source index, I_e , γ_1 and γ_2), others are measured much less precisely due to the stochastic perturber population ($x_{s/l}$, $y_{s/l}$, r_e and θ_E). We omit coverage plots for this analysis as they are of comparably-good quality to those in the previous subsection.

1.5 Hierarchical inference of dark matter mass

In this section, we show our results for hierarchical population parameters inference. First, we describe the simulated data in subsection 1.5.1 and the inference strategy in subsection 1.5.2. We then show how we constrain the lens and source parameters in subsection 1.5.3. Next, we show our results for the HMF cutoff mass and describe how we can combine the information from different strong lensing images in subsection 1.5.4. In the same subsection, we show our results on the DM mass. Finally, we directly assess the statistical behaviour of the trained neural networks in subsection 1.5.5.

1.5.1 Mock data generation

For this inference task we use our lensing simulator at its full capacity, by including lens, source, substructures and different DM models that depend on the half-mode-

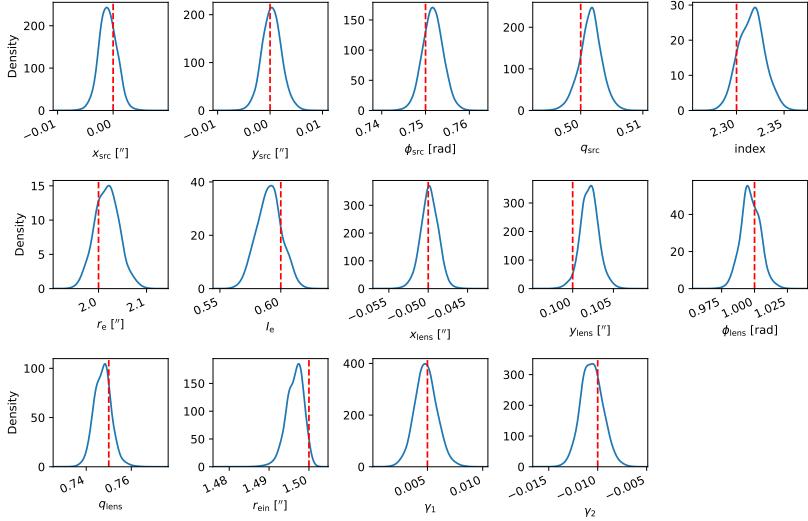


Figure 1.11: Macromodel 1D marginal posteriors as in [Figure 1.9](#), but for the inference task where a population of $10^7 - 10^8 M_\odot$ are present in the observation. This has the effect of broadening most of the posteriors.

mass M_{hm} :

$$p(\mathbf{x} \mid \boldsymbol{\eta}_{\text{src}}, \boldsymbol{\eta}_{\text{lens}}, \vec{\eta}_{\mathbf{P}}, \vartheta) = \mathcal{N}(\mathbf{x} \mid \text{obs}(\boldsymbol{\eta}_{\text{src}}, \boldsymbol{\eta}_{\text{lens}}, \vec{\eta}_{\mathbf{P}}, M_{\text{hm}}), \sigma^2). \quad (1.36)$$

In [Figure 1.12](#) we show a gallery of twenty mock strong-lensing images we use as target observations. These mock observations have been generated with arbitrary lens and source parameters drawn from the initial prior in [Table 1.2](#).

1.5.2 Inference strategy

The inference neural network used to perform TMNRE is split into two different components: an embedding network $C_\phi(\mathbf{x})$ and a binary classification network. The embedding network compresses data into a low-dimensional feature vector, estimating the best possible summary statistics from the full input image. The binary classification network is the marginal classifier that performs the actual ratio estimation. It passes the featurized observational data concatenated with the parameter of interest into a MLP to estimate the marginal likelihood-to-evidence ratios. The network architecture can be expressed as:

$$d_\phi(\mathbf{x}, \vartheta) = \text{MLP}_\phi(\text{features} = C_\phi(\mathbf{x}), \vartheta) = \sigma[\log \hat{r}(\mathbf{x}, \vartheta)]. \quad (1.37)$$

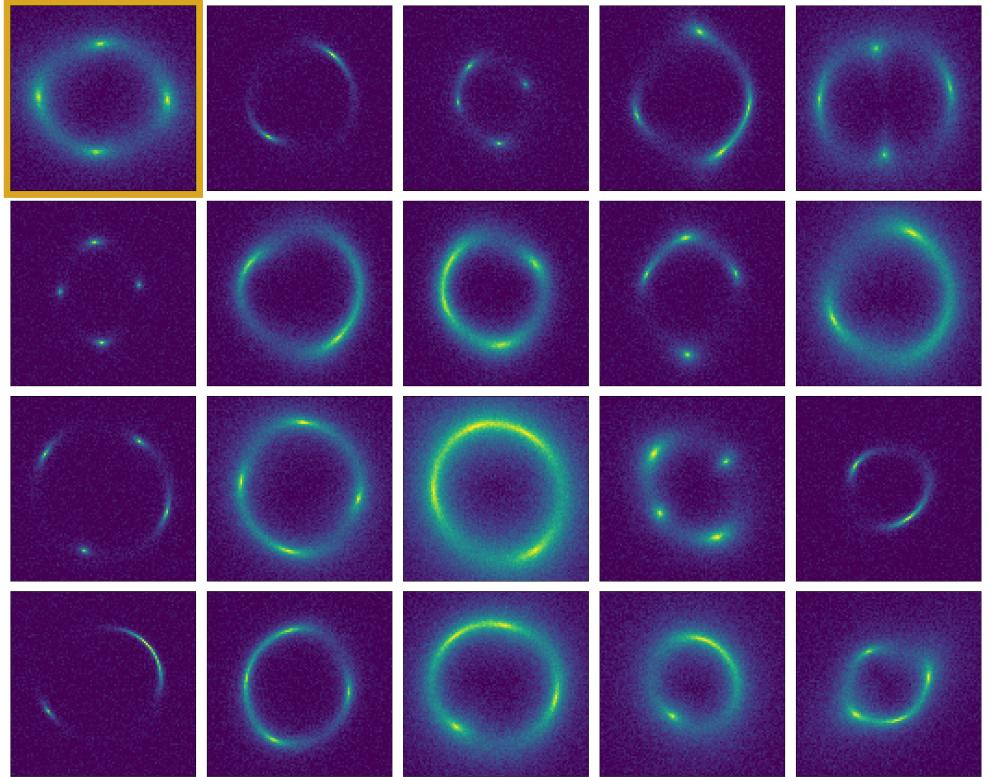


Figure 1.12: We present a gallery of twenty mock strong-lensing images we use as target observations. These mock observations have been generated with arbitrary lens and source parameters drawn from the initial prior in ???. Their peak signal-to-noise (SNR) is ~ 30 , representative of HST data. We analyse these images by first constraining their lens and source parameters proposal distribution in subsection 1.5.3. Then, we combine them in order to infer the cutoff mass scale in subsection 1.5.4. For the first one (upper left corner, framed in orange) of these images we show our results of the first part of the pipeline (subsection 1.5.3) in Figure 1.14, Figure 1.15, and Figure 1.16.

Parameter	True value	Prior	Description
Main lens			
θ_E ["]		$\mathcal{U}(1., 2.)$	SPLE
$\xi_{0,x}$ ["]		$\mathcal{U}(-0.2, 0.2)$	Einstein radius
$\xi_{0,y}$ ["]		$\mathcal{U}(-0.2, 0.2)$	lens center x axis
q_l		$\mathcal{U}(0.1, 1.)$	lens center y axis
ϕ_l [rad]		$\mathcal{U}(0, 2\pi)$	axis ratio
γ	2.1	-	rotation angle
z_{lens}	0.5	-	slope
			lens redshift
External shear			
γ_1		$\mathcal{U}(-0.05, 0.05)$	1 st component
γ_2		$\mathcal{U}(-0.05, 0.05)$	2 nd component
Source			
I_e		$\mathcal{U}(0., 4.)$	Sérsic
r_e ["]		$\mathcal{U}(0.1, 2.5)$	surface intensity
x_0 ["]		$\mathcal{U}(-0.1, 0.1)$	effective radius
y_0 ["]		$\mathcal{U}(-0.1, 0.1)$	source center x axis
q_s		$\mathcal{U}(0.1, 1.)$	source center y axis
ϕ_s [rad]		$\mathcal{U}(0, 2\pi)$	axis ratio
n		$\mathcal{U}(0.1, 4.)$	position angle
z_{src}	2	-	index
			source redshift
Subhalos			
\vec{p} ["]	$\in [-2.5, 2.5]$	$\mathcal{U}_{2D}(-2.5, 2.5)$	tNFW
m_{200} [M_\odot]	$\in [10^7, 10^{10}]$	[107]	position
c_{200}	15.	-	virial mass
τ	6.	-	concentration
			truncation
LOS halos			
\vec{p} ["]	$\in [-2.5, 2.5]$	$\mathcal{U}_{2D}(-2.5, 2.5)$	projected tNFW
m_{200} [M_\odot]	$\in [10^7, 10^{10}]$	[110]	position
z_{LOS}	$\in [0, z_{\text{src}}]$	[110]	virial mass
c_{200}	15.	-	LOS redshift
τ	6.	-	concentration
			truncation
WDM			
M_{hm} [M_\odot]		$\log \mathcal{U}(10^7, 10^{10})$	half-mode mass

Table 1.2: Summary of model parameters used for the simulated images in this work. When a prior distribution is not specified, the parameter is fixed to the true value.

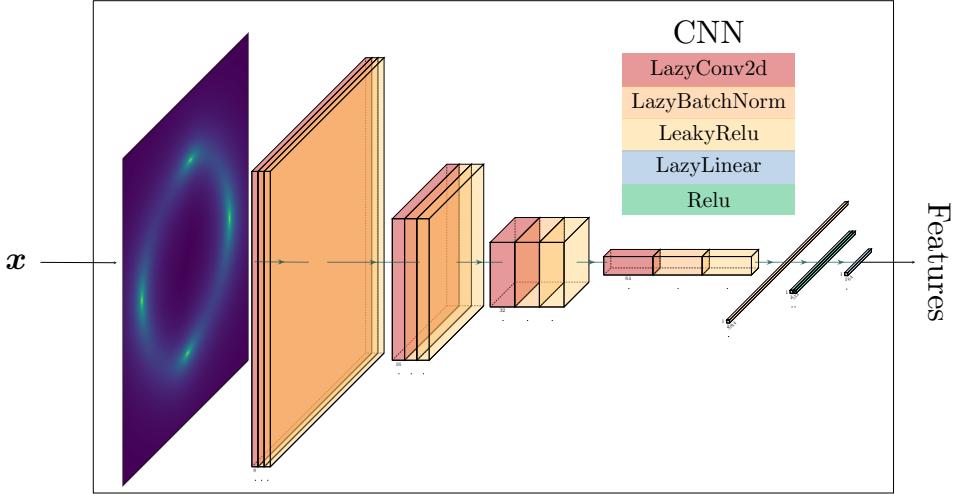


Figure 1.13: Illustration of the embedding CNN architecture used in the first part of the pipeline to constrain lens and source parameters. The observation x gets compressed into features: estimates of the best possible data summary statistic, by the CNN. In describing the CNN layers we follow PyTorch [85] convention. To create the illustration we have used [121].

For the embedding network, in both steps of the pipeline, we adopt a simple CNN. In Figure 1.13 we show the CNN architecture used to constrain lens and source parameters. The one used to estimate the cutoff mass has a similar structure.

1.5.3 Constraining lens and source parameters

We constrain lens and source parameters regions with TMNRE (??) with multiple sampling and training rounds.

In total, we perform six sampling and training rounds. In each round, we simulate 10^5 observations, of which 90% are used as the training dataset, and the remaining 10% as the validation dataset. Evaluations of the network on the mock target image are used to truncate the training data proposal distribution after each round, so that the region for lens and source parameters is targeted. The first training round is performed on the dataset generated from the initial source and lens parameters priors, shown in Table 1.2. In Figure 1.14 we show the initial prior and the following constrained proposal distributions. It can be seen that after the first round just a few of the parameters proposal distributions get truncated, e.g. the Einstein radius. By

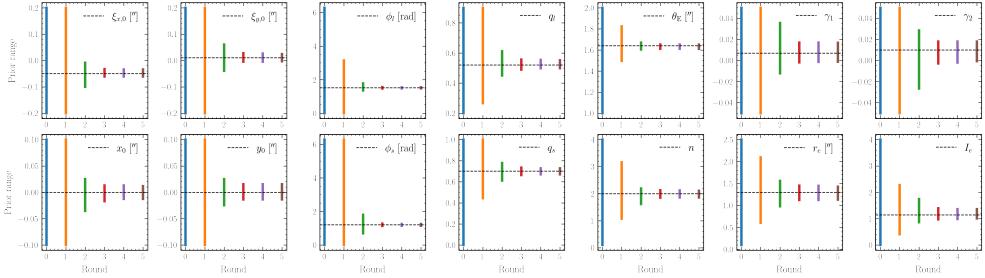


Figure 1.14: Constrained proposal distribution. Visualization of the sequential truncation of the lens and source proposal distributions over the six rounds of training. The particular target is the first mock image (framed in orange in Figure 1.12), whose parameters are depicted as black dashed horizontal lines.

having truncated these initial parameters, in the following rounds the other parameters can be better learned by the network and so constrained. In Figure 1.15, we show samples from the first five training datasets, which demonstrate that the constrained regions are indeed the ones that are likely to produce data similar to the targeted image \mathbf{x} . After the sixth round of training, it is not possible anymore to truncate the proposal distribution region based on the predetermined threshold, as seen in Figure 1.16. The truncation scheme has then efficiently identified the constrained region for lens and source parameters consistent with the targeted observation. Using the last constrained dataset, it is then easier in the second step of the pipeline to train a marginal neural ratio estimator to perform the final inference on the cutoff mass.

We would like to stress that these constrained proposal distributions correctly account for lens and source parameters uncertainties. In all our simulated data, the substructure parameters $\vec{\eta}_p$ are randomly sampled from their prior, in order to account for the presence of substructure. This has the desirable outcome of approximately accounting for the average effect that an additional mass component has on the main lens parameters (e.g. inferring an unbiased Einstein radius) and contributes to the source and lens uncertainties.

1.5.4 Dark matter inference

For the second step of the pipeline, we train an inference network to learn the cutoff mass on the last constrained dataset.

From initial tests, we have found that features from a single image are very hard to learn for the classifier, resulting in a very noisy ratio estimator. In order to reduce the estimator uncertainty, we then train the cutoff mass classifier on a dataset $X^N =$

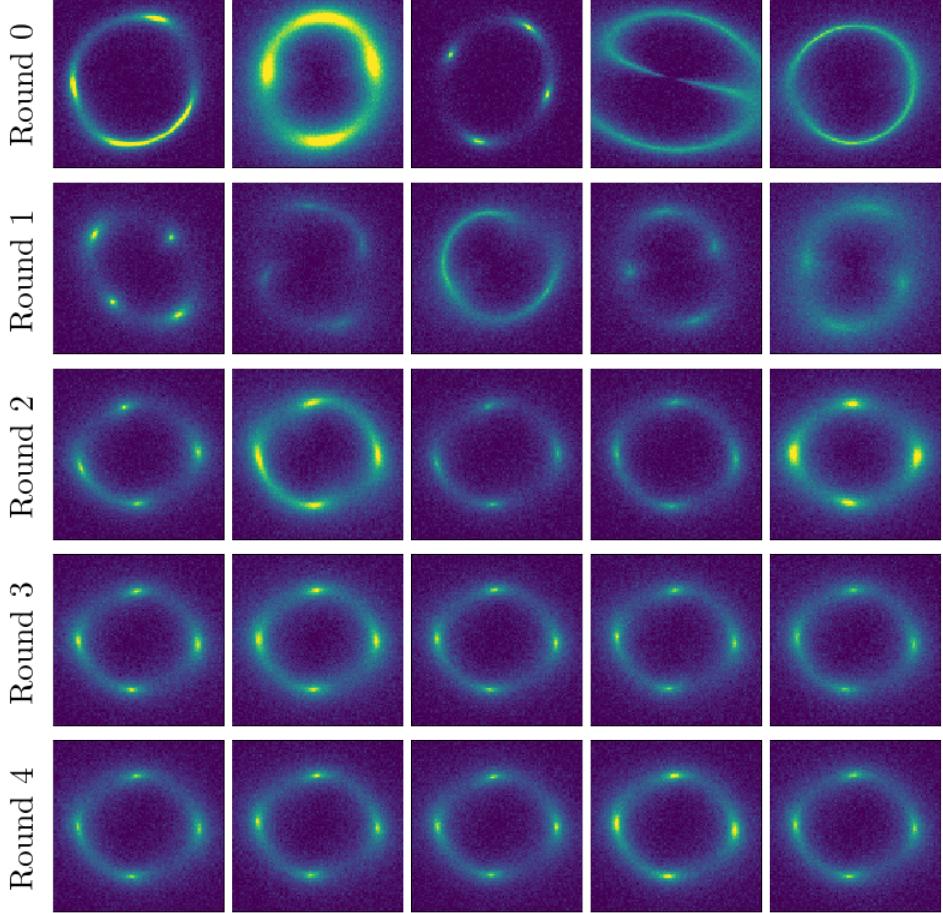


Figure 1.15: Training data targeting the first mock observation (framed in orange in [Figure 1.12](#)). In each row, we show five examples of training data for the first five rounds. In the first round, we sample our data from the initial prior shown in [Table 1.2](#). For the following rounds, the lens and source parameters are sampled from the constrained proposal distributions, obtained by evaluating the network trained with the previous round dataset on our target observation (see [subsection 1.5.3](#)). It is evident that with each round the training data more closely resembles the target image \mathbf{x}_0 .

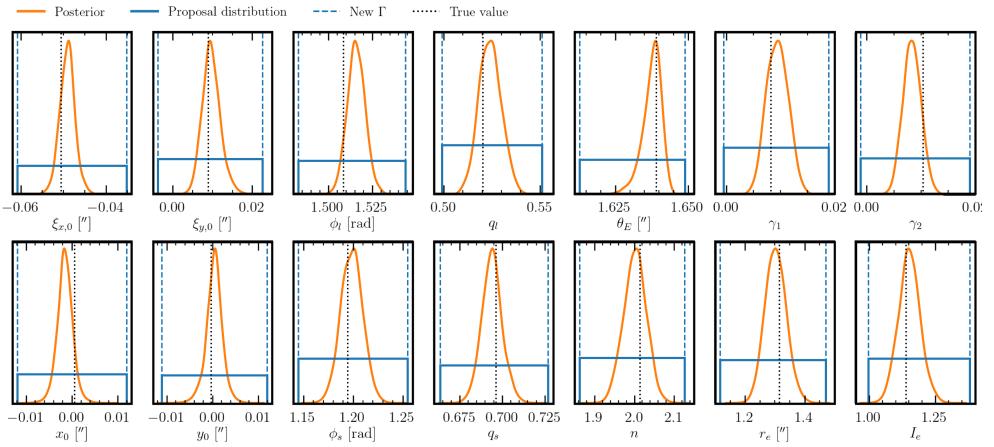


Figure 1.16: Lens and source parameters posteriors. In solid blue we show the last round of constrained proposal distributions for the first (upper left corner, framed in orange) target image in [Figure 1.12](#). The dotted black lines correspond to the true lens and source parameters values with which we have generated the target image. In orange, we show the estimated posteriors for lens and source parameters in the last training round. Based on the predetermined threshold, the new bounding limits Γ (dashed blue) do not change significantly from the previous constrained proposal distribution region, so it is not possible to constrain the proposal distribution more and we stop the truncation procedure.

$\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ of N different observations. For each observation, first, we constrain its lens and source parameters as explained in [subsection 1.5.3](#). Then, we train the cutoff classifier on the concatenation of the features coming from their embedding networks, effectively learning $r(M_{\text{hm}}; X^N)$. Note that the images in one dataset are sampled with the same cutoff mass M_{hm} , but different lens, source, and substructures realizations. In fact, our final goal is to apply the full pipeline to real data, which will all have different source, lens, and substructures configurations, but will have encoded the same DM properties.

In the first row of [Figure 1.17](#) we show the results from the inference network on ten test sets of lenses generated with a M_{hm} value of $10^7, 10^8, 10^9$ and, $10^{10} M_\odot$. Each curve is the posterior obtained for a set of $N = 20$ lenses. Each of the mock observations has lens and source parameters sampled from their own final constrained proposal distribution, and different substructure population.

Now that we have reduced the estimator noise, it is straightforward to perform inference on a group of sets of images by combining their ratios. Given a dataset $X^N = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ of images, the combined ratio for multiple M datasets is simply given by $r(M_{\text{hm}}; X_M^N) \propto \prod_{i=1}^M r(M_{\text{hm}}; X_i^N)$, where the proportionality is a ratio of evidences, independent of the parameter value, so it only accounts for a proper normalisation [71, 74]. In the second row of [Figure 1.17](#) we show the results for the combination of the $M = 10$ different posteriors shown in the first column.

In the third row we show a combined posterior for the WDM mass function from 200 images ($M = 10$ sets of $N = 20$ images). These plots show the uncertainty in the subhalo mass function under the assumption that it has the functional form in [Equation 1.33](#) with parameters from [115].

These first results show that our method is sensitive to the low-mass end of the HMF, and that we have unbiased results from combining just 10 sets of 20 observations, given that in the second panel of [Figure 1.17](#) the true input value for the half-mode mass M_{hm} is consistently contained within the estimated posterior. In [subsection 1.5.5](#) we will show a more sophisticated method to assess the statistical behaviour of our inference results.

Furthermore, we can translate the constraints we obtain on the cutoff mass to constraints on the WDM mass given the mapping between those two quantities defined in [subsection 1.2.5](#). In [Figure 1.18](#) we show our results for the WDM mass. Each column corresponds to a different cutoff mass input value, so a different WDM mass. In the first row, we plot five examples of the combined posterior density for $\log_{10} M_{\text{hm}}$ of $M = 10$ sets of $N = 20$ observations. In the second row, we show the corresponding color-coded five examples for m_{WDM} . In this case, we just transform the posterior from the first row using the parameterisation shown in [subsection 1.2.5](#), so we assume a flat prior on $\log_{10} M_{\text{hm}}$. Finally, in the last row, we show the WDM mass posterior

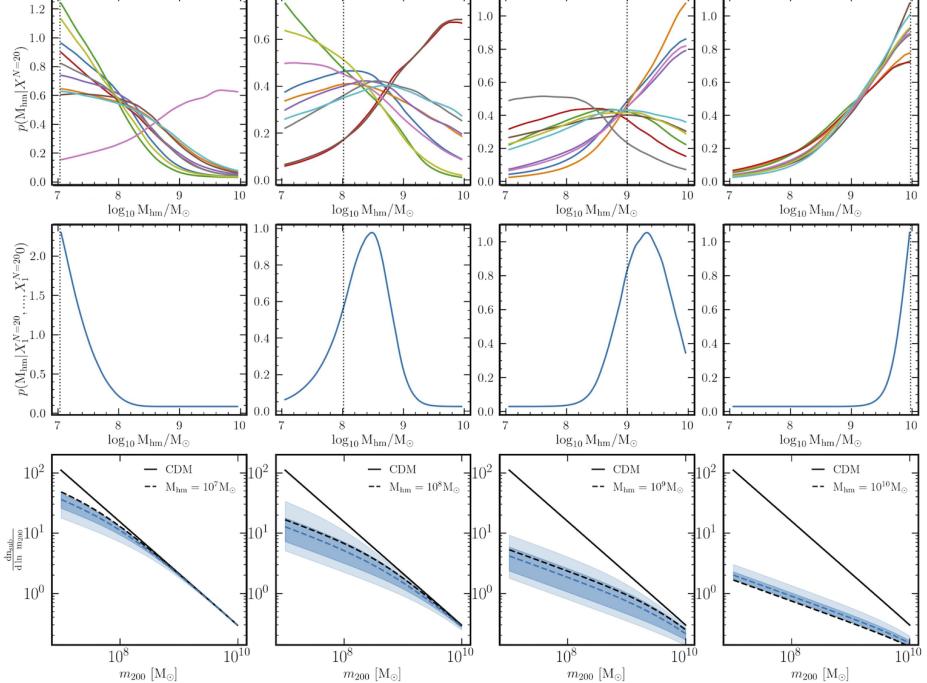


Figure 1.17: *Top:* Approximate posteriors for the half-mode mass derived from 10 different sets of 20 images. The dotted black line represents the true value of the half mode mass with which we have generated the images ($10^7, 10^8, 10^9, 10^{10} M_\odot$). *Middle:* We show the approximate posterior resulting from the combination of the $M = 10$ different posteriors shown in the first column, as explained in the text (subsection 1.5.4). *Bottom:* Subhalo mass function constraints derived from the cutoff mass posterior shown in the second column. The black solid line shows the CDM subhalo mass function according to Equation 1.22, whereas the black dashed one shows the WDM subhalo mass function according to Equation 1.33, given the true cutoff mass shown in the label. The blue dashed line shows the mean of the WDM subhalo mass function obtained by sampling 1000 samples from the cutoff mass posterior shown in the second panel and using this value in Equation 1.33. We also show the central 68 and 95 percentiles as shaded bands. These plots show how uncertain the subhalo mass function is under the assumption that it has the functional form in Equation 1.33 with parameters from [115].

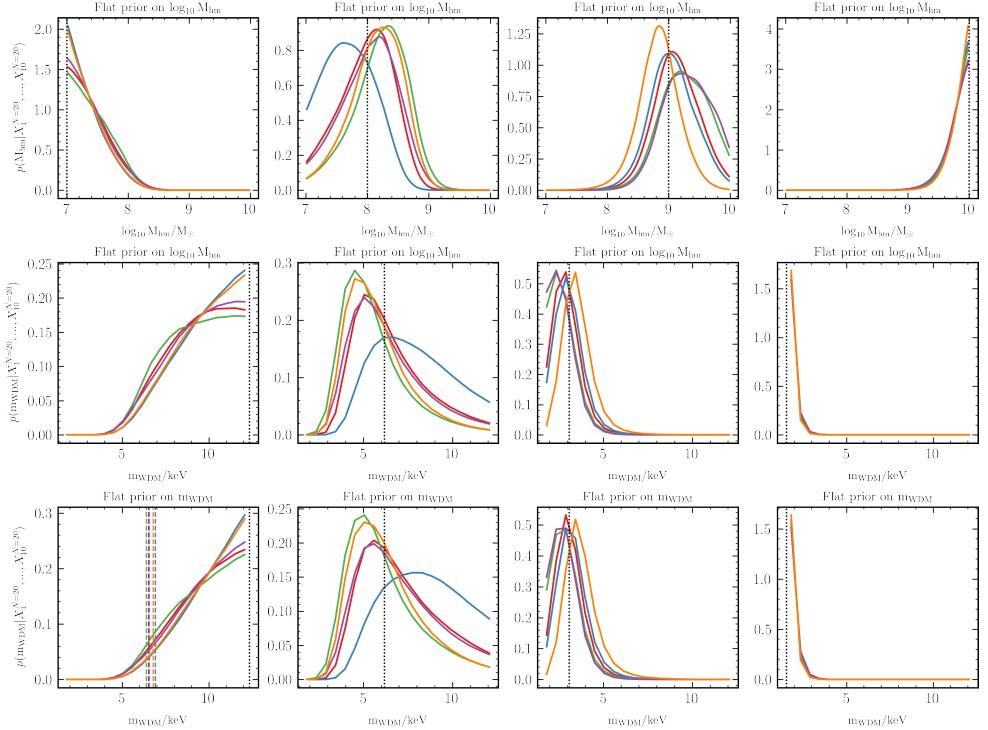


Figure 1.18: *Top:* We show five examples of combined posterior of $M = 10$ sets of $N = 20$ observations in terms of the cutoff mass (as the second row in Figure 1.17). The dotted black line represents the true input value of the half mode mass with which we have generated the analyzed mock observations ($10^7, 10^8, 10^9, 10^{10} \text{ M}_\odot$). *Middle:* Same results as shown in the first column but for the WDM mass. The dotted black line represents the true value of the WDM mass with which we have generated the analyzed mock observations, given the mapping between DM cutoff and DM mass in subsection 1.2.5. The WDM mass posteriors assume a flat prior on the cutoff mass. *Bottom:* Same results as shown in the first column but for the WDM mass and assuming a flat prior on the latter. In the first plot of the row, we show for the five examples the expected 95% credible lower limit on the WDM mass for the highest value of our prior distribution.

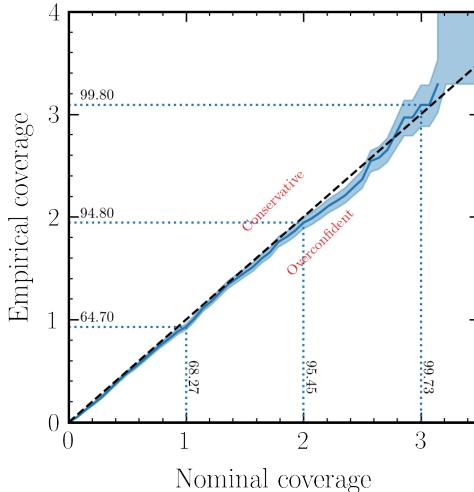


Figure 1.19: Empirical versus nominal expected coverage probabilities for the cutoff mass inference network. In case the line lies above (below) the black dashed diagonal line, the credible intervals are conservative (overconfident) and contain the true value with a frequency higher (lower) than nominally expected. We show the empirical (nominal) probabilities as horizontal (vertical) text.

densities assuming a flat prior on the latter. The posteriors in the second and third row are not actually the same because a flat prior $\log_{10} M_{\text{hm}}$ is different from a flat prior on m_{WDM} .

1.5.5 Credible interval testing

We would like to directly test and validate the statistical behaviour of our inference results by determining the expected coverage of the ratio estimator produced by the network. This can be easily done in *swyft* thanks to local amortization [116]. The goal is to compare the nominal and empirical expected coverage probabilities of estimated Bayesian credible intervals, which should coincide for a well-calibrated estimator. For the statistical formalism and definition of credible region and expected coverage probability, we refer the reader to [122]. In brief, an ideal estimator has matching empirical and nominal expected coverage, a conservative one predicts lower credibility than empirically obtained, and an overconfident one has higher nominal than empirical credibility. In plots like Figure 1.19, the line for an ideal ratio estimator should perfectly align with the diagonal, whereas for a conservative (overconfident) estimator, it will lie above (below) the diagonal. In combination with visually checking

the posteriors, this test supports the accuracy of the posterior estimator and is also particularly useful when one does not have access to the ground truth against which to compare the results. In Figure 1.19 we show the empirical versus nominal expected coverage probabilities for the cutoff mass inference network. We can see that the inference network for the half-mode mass has converged with good expected coverage.

1.6 Discussion

In this section, we discuss the improvements to the model and inference question which need to be addressed before we can safely apply our pipeline to the analysis of real data.

First, we have neglected effects such as inadequate *lens light* subtraction and assumed the lens light to be known. Regarding the *noise* model, we did not account for correlated pixel noise due to instrumental effects including the telescope’s PSF, multi-band observations, drizzling (e.g.see [73]), and even complex noise with an unknown likelihood function.

In this work, we have employed an analytic parameterisation (the Sérsic profile) as a lensed source light distribution model, which is adequate to analyze low-resolution images. However, to accurately model higher-fidelity lensing observations, such as those from on-going (e.g.HST) and future (e.g.JWST, ELT, SKA) telescopes, more *complex source models* need to be employed. Existing models, in order of complexity, are regularised pixellation of the source plane (see, e.g., [35, 90, 123]), source modelling through basis functions (e.g.shapelets [124] or wavelets [125]) attached to the source plane, and deep learning approaches (see, e.g., [62, 126]). The ability to accurately and precisely reconstruct the complex morphology of strong-lensing sources is of the utmost importance, as to disentangle the source surface brightness inhomogeneities from the percent-level fluctuations introduced by substructures in the lens. We anticipate that using sources with more complex morphologies will result in higher sensitivity to the DM cutoff mass, provided that it is possible to model these sources. In fact, the residuals between the image of an extended source lensed by the total lens potential (accounting for substructures), and that of the same source lensed only by the main lens component are proportional to the gradient of that source evaluated in the image plane [56, Equation 16].

Regarding WDM modelling, validation of our smoothing scheme (Equation 1.2.5) is required to accurately account for DM free-streaming effects. Moreover, we should account for uncertainties due to the assumed halo density profile by considering different DM distributions around galaxies (see, e.g., [127] for a review). We note that, thanks to its flexibility, our pipeline can incorporate any arbitrary DM model, as long as it specifies the form of the HMF and the density profiles of individual substructures.

Finally, we would like to draw the reader’s attention on the fact that in our modeling we assume that the halo mass of the lens is known exactly from its Einstein radius (see [Equation 1.2.4](#)). This is a strong assumption that has as a consequence the separation of substructure parameters $\vec{\eta}_p$ and lens parameter η_{lens} once we marginalize the posterior probability over the halo mass in [??](#). In particular, the second inference question we have addressed in this work in [section 1.5](#), constraining the cutoff mass of the subhalo mass distribution, is then a simplified version of the real one, which is to simultaneously determine the halo mass and subhalo mass distribution of the lenses from real data (see, e.g., [\[58\]](#)).

While this work used simple mock lenses, TMNRE makes it possible to add realism and parameters to a simulator without significantly altering the inference procedure, or necessarily increasing the simulation budget [\[119\]](#). It should therefore be straightforward to incorporate these various complexities we ignored in this work without fundamentally modifying the inference pipeline.

1.7 Conclusions

Measuring both the individual and collective properties of DM halos on sub-galactic scales by means of their gravitational effect is an important probe of the fundamental nature of DM. One of such probes, strong gravitational lensing, has sparked much interest over the last few years. Moreover, the development of fast and accurate techniques to extract information from strong lensing images is well motivated by the wealth of new high-resolution strong lensing observations that will become available in the near future.

In this work, we have presented the first step towards a new neural simulation based inference pipeline [\(??\)](#) to analyse present and future strong gravitational lensing systems in order to measure the properties of individual DM halos ([section 1.4](#)) and constrain the cutoff in the DM HMF, and so the DM mass ([section 1.5](#)). To this end, we have used a recent machine learning development, TMNRE, that makes it possible to *target* the analysis to a specific observation rather than amortize over all possible variations in lensing systems, making inference more efficient and precise. Thanks to TMNRE, we overcome the computational challenges of traditional MCMC, nested sampling and trans-dimensional MCMC methods, by directly learning the marginal posterior for the parameter of scientific-interest from the observation. Moreover, TMNRE leverages neural networks to directly learn the best summary statistic possible from the full input data, without having to compress the observation into hand-crafted summary statistics, like for ABC frameworks. The method is applicable to simulators with unknown likelihood functions and large or even variable numbers of input parameters. Lastly, the resulting inference networks can be poked

and prodded to confirm they are statistically well-behaved. This work is then a step forward towards making the analysis of strong lensing images for DM science faster, more efficient, and more accurate.

Our key results can be summarized as follows:

TMNRE can recover existing results. We verified the accuracy of TMNRE by confirming it reproduces analytically-calculable posteriors in a toy lensing scenario with known macromodel parameters and subhalo mass ([Figure 1.4](#)).

TMNRE enables direct marginal inference. Thanks to marginalized inference, the TMNRE based analysis is able to correctly propagate the lens and source parameters uncertainties, and account for the presence of a population of substructures, when estimating the marginal posterior of interest.

TMNRE enables statistical checks. Since the inference networks learned by TMNRE are locally amortized over a range of potential observations, we were able to test their statistical consistency. Our checks confirm that TMNRE on average produces posteriors with the correct width for the macromodel and subhalo parameters. Such tests would be extremely expensive with likelihood-based inference since they would require rerunning the sampling machinery on numerous mock observations.

The perturber population matters. We demonstrated that the sensitivity with which a subhalo’s parameters are measurable can be significantly degraded when marginalizing over a population of perturbers. While the 1σ regions of our position and mass posteriors were centered on the subhalo’s true parameters, they had heavy tails extending to the boundaries of our tight, manually-fixed priors. Given our validation, statistical checks and the fact TMNRE is so far the only method capable of performing the high-dimensional marginalization required for this analysis, *our results therefore suggest that the population of light perturbers should not be neglected*. However, it is important to highlight that we cannot strictly conclude from our results that the presence of a substructure population makes the inference of the properties of an individual subhalo unfeasible. What we find is that it makes the task more challenging for our particular SBI approach and network architectures. Whether better network architectures for the ratio estimator capable of modeling the posterior more accurately than MLP Mixer, or maybe the proper handling of the problem with a full trans-dimensional likelihood-based MCMC method dealing with the perturber population can resolve the issue remains open, and an important question to study in future work.

TMNRE enables hierarchical inference. We demonstrated that our framework is able to statistically extract the DM cutoff mass signal from a population of small-scale dark matter halos, by performing hierarchical inference on up to 200 observations

(subsection 1.5.4 and Figure 1.17). What we find is an expected 95% credible lower limits around 6.5 keV in the case of the scenario closest to CDM (see the bottom left panel in Figure 1.18), given the adopted prior and the various assumptions discussed in section 1.6.

A part from the possible framework improvements discussed in section 1.6, an interesting direction for further work is the use of TMNRE for model comparison. While here our ratio estimators were trained to compute the likelihood-to-evidence ratio, as pointed out in [71] it is possible to learn other ratios of densities. In particular ratio estimators can be used to learn the Bayes factor for assessing the strength of the evidence for different models. This could be used to determine whether an image contains a perturber or not, and to map the minimum-detectable perturber mass as a function of its position.

Overall, we believe using TMNRE to measure individual perturbers and perturbers' population parameters as described in this work provides a promising path towards uncovering the identity of dark matter.

Conv2d(1, 4, 8, 2, 1, bias=False)
BatchNorm2d(4)
LeakyReLU(0.2)
Conv2d(4, 8, 8, 2, 1, bias=False)
BatchNorm2d(8)
LeakyReLU(0.2)
Conv2d(8, 16, 8, 2, 1, bias=False)
BatchNorm2d(16)
LeakyReLU(0.2)
Conv2d(16, 32, 8, 2, 1, bias=False)
BatchNorm2d(32)
LeakyReLU(0.2)

Table 1.A.1: The convolutional compression network used in the macromodel parameter ratio estimator. The notation is taken from PyTorch: the arguments to Conv2d are the number of input channels, output channels, kernel size, stride and padding, respectively. The horizontal lines demarcate where the number of channels changes. The output of the network is flattened into a vector with 128 features.

Appendix

1.A Compression network architectures

The compressor architectures are given in Table ?? and Table ?? . Note that we standardize the images before providing them to the networks.

image_size	100
channels	1
patch_size	10
dim	256
depth	4
num_classes	32
dropout	0.1

Table 1.A.2: The details of the MLP Mixer compression network in the subhalo ratio estimator. We use the implementation from <https://github.com/lucidrains/mlp-mixer-pytorch>, with arguments given in the table.

BIBLIOGRAPHY

- [1] N. Anau Montel, A. Coogan, C. Correa, K. Karchev, and C. Weniger, “Estimating the warm dark matter mass from strong lensing images with truncated marginal neural ratio estimation,” *Mon. Not. Roy. Astron. Soc.* **518** no. 2, (2022) 2746–2760, [arXiv:2205.09126 \[astro-ph.CO\]](https://arxiv.org/abs/2205.09126).
- [2] A. Coogan, N. Anau Montel, K. Karchev, M. W. Grootes, F. Nattino, and C. Weniger, “The effect of the perturber population on subhalo measurements in strong gravitational lenses,” *Mon. Not. Roy. Astron. Soc.* **527** no. 1, (2024) 66–78, [arXiv:2209.09918 \[astro-ph.CO\]](https://arxiv.org/abs/2209.09918).
- [3] V. C. Rubin, N. Thonnard, and W. K. Ford, Jr., “Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/,” *Astrophys. J.* **238** (1980) 471.
- [4] F. Zwicky, “Die Rotverschiebung von extragalaktischen Nebeln,” *Helv. Phys. Acta* **6** (1933) 110–127.
- [5] Planck Collaboration, P. A. R. Ade *et al.*, “Planck 2015 results. XIII. Cosmological parameters,” *Astron. Astrophys.* **594** (2016) A13, [arXiv:1502.01589 \[astro-ph.CO\]](https://arxiv.org/abs/1502.01589).
- [6] A. N. Taylor, S. Dye, T. J. Broadhurst, N. Benitez, and E. van Kampen, “Gravitational lens magnification and the mass of Abell 1689,” *Astrophys. J.* **501** (1998) 539, [arXiv:astro-ph/9801158](https://arxiv.org/abs/astro-ph/9801158).
- [7] P. J. E. Peebles, “Large scale background temperature and mass fluctuations due to scale invariant primeval perturbations,” *Astrophys. J. Lett.* **263** (1982) L1–L5.
- [8] B. Moore, S. Ghigna, F. Governato, G. Lake, T. R. Quinn, J. Stadel, and P. Tozzi, “Dark matter substructure within galactic halos,” *Astrophys. J. Lett.* **524** (1999) L19–L22, [arXiv:astro-ph/9907411](https://arxiv.org/abs/astro-ph/9907411).
- [9] A. A. Klypin, A. V. Kravtsov, O. Valenzuela, and F. Prada, “Where are the missing Galactic satellites?,” *Astrophys. J.* **522** (1999) 82–92, [arXiv:astro-ph/9901240](https://arxiv.org/abs/astro-ph/9901240).
- [10] W. J. G. de Blok and S. S. McGaugh, “The Dark and visible matter content of low surface brightness disk galaxies,” *Mon. Not. Roy. Astron. Soc.* **290** (1997) 533–552, [arXiv:astro-ph/9704274](https://arxiv.org/abs/astro-ph/9704274).

Bibliography

- [11] J. S. Bullock and M. Boylan-Kolchin, “Small-Scale Challenges to the Λ CDM Paradigm,” *Ann. Rev. Astron. Astrophys.* **55** (2017) 343–387, [arXiv:1707.04256 \[astro-ph.CO\]](https://arxiv.org/abs/1707.04256).
- [12] J. S. Bullock, “Notes on the Missing Satellites Problem,” [arXiv:1009.4505 \[astro-ph.CO\]](https://arxiv.org/abs/1009.4505).
- [13] P. Colin, V. Avila-Reese, and O. Valenzuela, “Substructure and halo density profiles in a warm dark matter cosmology,” *Astrophys. J.* **542** (2000) 622–630, [arXiv:astro-ph/0004115](https://arxiv.org/abs/astro-ph/0004115).
- [14] C. J. Hogan and J. J. Dalcanton, “New dark matter physics: clues from halo structure,” *Phys. Rev. D* **62** (2000) 063511, [arXiv:astro-ph/0002330](https://arxiv.org/abs/astro-ph/0002330).
- [15] M. R. Lovell, C. S. Frenk, V. R. Eke, A. Jenkins, L. Gao, and T. Theuns, “The properties of warm dark matter haloes,” *Mon. Not. Roy. Astron. Soc.* **439** (2014) 300–317, [arXiv:1308.1399 \[astro-ph.CO\]](https://arxiv.org/abs/1308.1399).
- [16] D. N. Spergel and P. J. Steinhardt, “Observational evidence for selfinteracting cold dark matter,” *Phys. Rev. Lett.* **84** (2000) 3760–3763, [arXiv:astro-ph/9909386](https://arxiv.org/abs/astro-ph/9909386).
- [17] W. Hu, R. Barkana, and A. Gruzinov, “Cold and fuzzy dark matter,” *Phys. Rev. Lett.* **85** (2000) 1158–1161, [arXiv:astro-ph/0003365](https://arxiv.org/abs/astro-ph/0003365).
- [18] M. R. Buckley and A. H. G. Peter, “Gravitational probes of dark matter physics,” *Phys. Rept.* **761** (2018) 1–60, [arXiv:1712.06615 \[astro-ph.CO\]](https://arxiv.org/abs/1712.06615).
- [19] G. Efstathiou, “Suppressing the formation of dwarf galaxies via photoionization,” *Mon. Not. Roy. Astron. Soc.* **256** (1992) 43P–47P.
- [20] A. Fitts *et al.*, “FIRE in the Field: Simulating the Threshold of Galaxy Formation,” *Mon. Not. Roy. Astron. Soc.* **471** no. 3, (2017) 3547–3562, [arXiv:1611.02281 \[astro-ph.GA\]](https://arxiv.org/abs/1611.02281).
- [21] S. Adhikari *et al.*, “Astrophysical Tests of Dark Matter Self-Interactions,” [arXiv:2207.10638 \[astro-ph.CO\]](https://arxiv.org/abs/2207.10638).
- [22] C. S. Kochanek, “The saas fee lectures on strong gravitational lensing,” [astro-ph/0407232](https://arxiv.org/abs/0407232). <https://arxiv.org/pdf/astro-ph/0407232.pdf>.
- [23] S. Vegetti *et al.*, “Strong gravitational lensing as a probe of dark matter,” [arXiv:2306.11781 \[astro-ph.CO\]](https://arxiv.org/abs/2306.11781).
- [24] A. Drlica-Wagner, Y.-Y. Mao, S. Adhikari, R. Armstrong, A. Banerjee, N. Banik, K. Bechtol, S. Bird, K. K. Boddy, A. Bonaca, J. Bovy, M. R. Buckley, E. Bulbul, C. Chang, G. Chapline, J. Cohen-Tanugi, A. Cuoco, F.-Y. Cyr-Racine, W. A. Dawson, A. D. Rivero, C. Dvorkin, D. Erkal, C. D. Fassnacht, J. García-Bellido, and M. Giannotti, “Probing the fundamental nature of dark matter with the large synoptic survey telescope,” [1902.01055](https://arxiv.org/abs/1902.01055). <https://arxiv.org/pdf/1902.01055.pdf>.
- [25] S.-d. Mao and P. Schneider, “Evidence for substructure in lens galaxies?,” *Mon. Not. Roy. Astron. Soc.* **295** (1998) 587–594, [arXiv:astro-ph/9707187](https://arxiv.org/abs/astro-ph/9707187).

Bibliography

- [26] N. Dalal and C. S. Kochanek, “Direct detection of CDM substructure,” *Astrophys. J.* **572** (2002) 25–33, [arXiv:astro-ph/0111456](https://arxiv.org/abs/astro-ph/0111456).
- [27] A. M. Nierenberg, T. Treu, S. A. Wright, C. D. Fassnacht, and M. W. Auger, “Detection of substructure with adaptive optics integral field spectroscopy of the gravitational lens B1422+231,” *Mon. Not. Roy. Astron. Soc.* **442** no. 3, (2014) 2434–2445, [arXiv:1402.1496 \[astro-ph.GA\]](https://arxiv.org/abs/1402.1496).
- [28] A. M. Nierenberg, T. Treu, G. Brammer, A. H. G. Peter, C. D. Fassnacht, C. R. Keeton, C. S. Kochanek, K. B. Schmidt, D. Sluse, and S. A. Wright, “Probing dark matter substructure in the gravitational lens HE 0435–1223 with the WFC3 grism,” *Mon. Not. Roy. Astron. Soc.* **471** no. 2, (2017) 2224–2236, [arXiv:1701.05188 \[astro-ph.CO\]](https://arxiv.org/abs/1701.05188).
- [29] D. Gilman, S. Birrer, T. Treu, C. R. Keeton, and A. Nierenberg, “Probing the nature of dark matter by forward modelling flux ratios in strong gravitational lenses,” *Mon. Not. Roy. Astron. Soc.* **481** no. 1, (2018) 819–834, [arXiv:1712.04945 \[astro-ph.CO\]](https://arxiv.org/abs/1712.04945).
- [30] D. Gilman, S. Birrer, A. Nierenberg, T. Treu, X. Du, and A. Benson, “Warm dark matter chills out: constraints on the halo mass function and the free-streaming length of dark matter with eight quadruple-image strong gravitational lenses,” *Mon. Not. Roy. Astron. Soc.* **491** no. 4, (2020) 6077–6101, [arXiv:1908.06983 \[astro-ph.CO\]](https://arxiv.org/abs/1908.06983).
- [31] J. W. Hsueh, C. D. Fassnacht, S. Vegetti, J. P. McKean, C. Springola, M. W. Auger, L. V. E. Koopmans, and D. J. Lagattuta, “SHARP – II. Mass structure in strong lenses is not necessarily dark matter substructure: a flux ratio anomaly from an edge-on disc in B1555+375,” *Mon. Not. Roy. Astron. Soc.* **463** no. 1, (2016) L51–L55, [arXiv:1601.01671 \[astro-ph.CO\]](https://arxiv.org/abs/1601.01671).
- [32] J. W. Hsueh, L. Oldham, C. Springola, S. Vegetti, C. D. Fassnacht, M. W. Auger, L. V. E. Koopmans, J. P. McKean, and D. J. Lagattuta, “SHARP – IV. An apparent flux-ratio anomaly resolved by the edge-on disc in B0712+472,” *Mon. Not. Roy. Astron. Soc.* **469** no. 3, (2017) 3713–3721, [arXiv:1701.06575 \[astro-ph.GA\]](https://arxiv.org/abs/1701.06575).
- [33] J.-W. Hsueh, W. Enzi, S. Vegetti, M. Auger, C. D. Fassnacht, G. Despali, L. V. E. Koopmans, and J. P. McKean, “SHARP – VII. New constraints on the dark matter free-streaming properties and substructure abundance from gravitationally lensed quasars,” *Mon. Not. Roy. Astron. Soc.* **492** no. 2, (2020) 3047–3059, [arXiv:1905.04182 \[astro-ph.CO\]](https://arxiv.org/abs/1905.04182).
- [34] L. V. E. Koopmans, “Gravitational lensing & stellar dynamics,” *EAS Publ. Ser.* **20** (2006) 161, [arXiv:astro-ph/0511121](https://arxiv.org/abs/astro-ph/0511121).
- [35] S. Vegetti and L. V. E. Koopmans, “Bayesian Strong Gravitational-Lens Modelling on Adaptive Grids: Objective Detection of Mass Substructure in Galaxies,” *Mon. Not. Roy. Astron. Soc.* **392** (2009) 945, [arXiv:0805.0201 \[astro-ph\]](https://arxiv.org/abs/0805.0201).
- [36] S. Vegetti and L. V. E. Koopmans, “Statistics of CDM substructure from strong gravitational lensing: quantifying the mass fraction and mass function,” *Mon. Not.*

- Roy. Astron. Soc.* **400** (2009) 1583–1592, [arXiv:0903.4752 \[astro-ph.CO\]](#).
- [37] L. V. E. Koopmans, “Gravitational-mass imaging of CDM substructure,” *Mon. Not. Roy. Astron. Soc.* **363** (2005) 1136, [arXiv:astro-ph/0501324](#).
- [38] S. Vegetti, O. Czoske, and L. V. E. Koopmans., “Quantifying dwarf satellites through gravitational imaging: the case of SDSS J120602.09+514229.5,” *Mon. Not. Roy. Astron. Soc.* **407** (2010) 225, [arXiv:1002.4708 \[astro-ph.CO\]](#).
- [39] S. Vegetti, L. V. E. Koopmans, A. Bolton, T. Treu, and R. Gavazzi, “Detection of a Dark Substructure through Gravitational Imaging,” *Mon. Not. Roy. Astron. Soc.* **408** (2010) 1969, [arXiv:0910.0760 \[astro-ph.CO\]](#).
- [40] S. Vegetti, D. J. Lagattuta, J. P. McKean, M. W. Auger, C. D. Fassnacht, and L. V. E. Koopmans, “Gravitational detection of a low-mass dark satellite at cosmological distance,” *Nature* **481** (2012) 341, [arXiv:1201.3643 \[astro-ph.CO\]](#).
- [41] Y. D. Hezaveh *et al.*, “Detection of lensing substructure using ALMA observations of the dusty galaxy SDP.81,” *Astrophys. J.* **823** no. 1, (2016) 37, [arXiv:1601.01388 \[astro-ph.CO\]](#).
- [42] J. M. Diego, M. Pascale, B. J. Kavanagh, P. Kelly, L. Dai, B. Frye, and T. Broadhurst, “Godzilla, a monster lurks in the Sumburst galaxy,” *Astron. Astrophys.* **665** (2022) A134, [arXiv:2203.08158 \[astro-ph.GA\]](#).
- [43] S. Vegetti, L. V. E. Koopmans, M. W. Auger, T. Treu, and A. S. Bolton, “Inference of the cold dark matter substructure mass function at $z = 0.2$ using strong gravitational lenses,” *Mon. Not. Roy. Astron. Soc.* **442** no. 3, (2014) 2017–2035, [arXiv:1405.3666 \[astro-ph.GA\]](#).
- [44] S. Vegetti, G. Despali, M. R. Lovell, and W. Enzi, “Constraining sterile neutrino cosmologies with strong gravitational lensing observations at redshift $z \sim 0.2$,” *Mon. Not. Roy. Astron. Soc.* **481** no. 3, (2018) 3661–3669, [arXiv:1801.01505 \[astro-ph.CO\]](#).
- [45] E. Ritondale, S. Vegetti, G. Despali, M. W. Auger, L. V. E. Koopmans, and J. P. McKean, “Low-mass halo perturbations in strong gravitational lenses at redshift $z \sim 0.5$ are consistent with CDM,” *Mon. Not. Roy. Astron. Soc.* **485** no. 2, (2019) 2179–2193, [arXiv:1811.03627 \[astro-ph.CO\]](#).
- [46] J. Skilling, “Nested Sampling,” *AIP Conf. Proc.* **735** no. 1, (2004) 395.
- [47] J. W. Nightingale *et al.*, “Scanning for dark matter subhaloes in Hubble Space Telescope imaging of 54 strong lenses,” *Mon. Not. Roy. Astron. Soc.* **527** no. 4, (2023) 10480–10506, [arXiv:2209.10566 \[astro-ph.CO\]](#).
- [48] B. J. Brewer, D. Huijser, and G. F. Lewis, “Trans-Dimensional Bayesian Inference for Gravitational Lens Substructures,” *Mon. Not. Roy. Astron. Soc.* **455** no. 2, (2016) 1819–1829, [arXiv:1508.00662 \[astro-ph.IM\]](#).
- [49] T. Daylan, F.-Y. Cyr-Racine, A. Diaz Rivero, C. Dvorkin, and D. P. Finkbeiner, “Probing the small-scale structure in strongly lensed systems via transdimensional

- inference,” *Astrophys. J.* **854** no. 2, (2018) 141, [arXiv:1706.06111 \[astro-ph.CO\]](https://arxiv.org/abs/1706.06111).
- [50] A. Galan, G. Vernardos, A. Peel, F. Courbin, and J.-L. Starck, “Using wavelets to capture deviations from smoothness in galaxy-scale strong lenses,” *Astron. Astrophys.* **668** (2022) A155, [arXiv:2207.05763 \[astro-ph.IM\]](https://arxiv.org/abs/2207.05763).
- [51] A. Diaz Rivero, F.-Y. Cyr-Racine, and C. Dvorkin, “Power spectrum of dark matter substructure in strong gravitational lenses,” *Phys. Rev. D* **97** no. 2, (2018) 023001, [arXiv:1707.04590 \[astro-ph.CO\]](https://arxiv.org/abs/1707.04590).
- [52] A. D. Rivero, C. Dvorkin, F.-Y. Cyr-Racine, J. Zavala, and M. Vogelsberger, “Gravitational lensing and the power spectrum of dark matter substructure: Insights from the ethos n-body simulations,” *Phys. Rev. D* **98** (2018) 103517, [1809.00004](https://arxiv.org/abs/1809.00004).
<https://arxiv.org/pdf/1809.00004.pdf>.
- [53] S. Brennan, A. J. Benson, F.-Y. Cyr-Racine, C. R. Keeton, L. A. Moustakas, and A. R. Pullen, “Quantifying the power spectrum of small-scale structure in semi-analytic galaxies,” *Mon. Not. Roy. Astron. Soc.* **488** no. 4, (2019) 5085–5092, [arXiv:1808.03501 \[astro-ph.GA\]](https://arxiv.org/abs/1808.03501).
- [54] A. ÇaÇsan Şengül, A. Tsang, A. Diaz Rivero, C. Dvorkin, H.-M. Zhu, and U. Seljak, “Quantifying the line-of-sight halo contribution to the dark matter convergence power spectrum from strong gravitational lenses,” *Phys. Rev. D* **102** no. 6, (2020) 063502, [arXiv:2006.07383 \[astro-ph.CO\]](https://arxiv.org/abs/2006.07383).
- [55] S. Chatterjee and L. V. E. Koopmans, “The inner mass power spectrum of galaxies using strong gravitational lensing: beyond linear approximation,” *Mon. Not. Roy. Astron. Soc.* **474** no. 2, (2018) 1762–1772, [arXiv:1710.03075 \[astro-ph.CO\]](https://arxiv.org/abs/1710.03075).
- [56] F.-Y. Cyr-Racine, C. R. Keeton, and L. A. Moustakas, “Beyond subhalos: Probing the collective effect of the universe’s small-scale structure with gravitational lensing,” *Phys. Rev. D* **100** (2019) 023013, [1806.07897](https://arxiv.org/abs/1806.07897).
<https://arxiv.org/pdf/1806.07897.pdf>.
- [57] D. Bayer, S. Chatterjee, L. V. E. Koopmans, S. Vegetti, J. P. McKean, T. Treu, C. D. Fassnacht, and K. Glazebrook, “Probing sub-galactic mass structure with the power spectrum of surface-brightness anomalies in high-resolution observations of galaxy–galaxy strong gravitational lenses. II. Observational constraints on the subgalactic matter power spectrum,” *Mon. Not. Roy. Astron. Soc.* **523** no. 1, (2023) 1310–1325, [arXiv:1803.05952 \[astro-ph.GA\]](https://arxiv.org/abs/1803.05952).
- [58] S. Birrer, A. Amara, and A. Refregier, “Lensing substructure quantification in RXJ1131-1231: A 2 keV lower bound on dark matter thermal relic mass,” *JCAP* **05** (2017) 037, [arXiv:1702.00009 \[astro-ph.CO\]](https://arxiv.org/abs/1702.00009).
- [59] Q. He, A. Robertson, J. Nightingale, S. Cole, C. S. Frenk, R. Massey, A. Amvrosiadis, R. Li, X. Cao, and A. Etherington, “A forward-modelling method to infer the dark matter particle mass from strong gravitational lenses,” *Mon. Not. Roy. Astron. Soc.* **511** no. 2, (2022) 3046–3062, [arXiv:2010.13221 \[astro-ph.CO\]](https://arxiv.org/abs/2010.13221).
- [60] Y. D. Hezaveh, L. Perreault Levasseur, and P. J. Marshall, “Fast Automated

- Analysis of Strong Gravitational Lenses with Convolutional Neural Networks,” *Nature* **548** (2017) 555–557, [arXiv:1708.08842 \[astro-ph.IM\]](https://arxiv.org/abs/1708.08842).
- [61] L. Perreault Levasseur, Y. D. Hezaveh, and R. H. Wechsler, “Uncertainties in Parameters Estimated with Neural Networks: Application to Strong Gravitational Lensing,” *Astrophys. J. Lett.* **850** no. 1, (2017) L7, [arXiv:1708.08843 \[astro-ph.CO\]](https://arxiv.org/abs/1708.08843).
 - [62] W. R. Morningstar, L. Perreault Levasseur, Y. D. Hezaveh, R. Blandford, P. Marshall, P. Putzky, T. D. Rueter, R. Wechsler, and M. Welling, “Data-Driven Reconstruction of Gravitationally Lensed Galaxies using Recurrent Inference Machines,” [arXiv:1901.01359 \[astro-ph.IM\]](https://arxiv.org/abs/1901.01359).
 - [63] G. Vernardos, G. Tsagkatakis, and Y. Pantazis, “Quantifying the structure of strong gravitational lens potentials with uncertainty-aware deep neural networks,” [2010.07314. https://arxiv.org/pdf/2010.07314.pdf](https://arxiv.org/abs/2010.07314).
 - [64] A. D. Rivero and C. Dvorkin, “Direct detection of dark matter substructure in strong lens images with convolutional neural networks,” *Phys. Rev. D* **101** (2020) 023515, [1910.00015. https://arxiv.org/pdf/1910.00015.pdf](https://arxiv.org/abs/1910.00015).
 - [65] S. Alexander, S. Gleyzer, E. McDonough, M. W. Toomey, and E. Usai, “Deep Learning the Morphology of Dark Matter Substructure,” *Astrophys. J.* **893** (2020) 15, [arXiv:1909.07346 \[astro-ph.CO\]](https://arxiv.org/abs/1909.07346).
 - [66] B. Ostdiek, A. Diaz Rivero, and C. Dvorkin, “Image segmentation for analyzing galaxy-galaxy strong lensing systems,” *Astron. Astrophys.* **657** (2022) L14, [arXiv:2009.06663 \[astro-ph.CO\]](https://arxiv.org/abs/2009.06663).
 - [67] B. Ostdiek, A. Diaz Rivero, and C. Dvorkin, “Extracting the Subhalo Mass Function from Strong Lens Images with Image Segmentation,” *Astrophys. J.* **927** no. 1, (2022) 83, [arXiv:2009.06639 \[astro-ph.CO\]](https://arxiv.org/abs/2009.06639).
 - [68] K. Cranmer, J. Brehmer, and G. Louppe, “The frontier of simulation-based inference,” *Proc. Nat. Acad. Sci.* **117** no. 48, (2020) 30055–30062, [arXiv:1911.01429 \[stat.ML\]](https://arxiv.org/abs/1911.01429).
 - [69] B. K. Miller, A. Cole, G. Louppe, and C. Weniger, “Simulation-efficient marginal posterior estimation with swyft: stop wasting your precious time,” [arXiv:2011.13951 \[astro-ph.IM\]](https://arxiv.org/abs/2011.13951).
 - [70] B. K. Miller, A. Cole, P. Forré, G. Louppe, and C. Weniger, “Truncated marginal neural ratio estimation,” [2107.01214. https://arxiv.org/pdf/2107.01214.pdf](https://arxiv.org/abs/2107.01214).
 - [71] J. Hermans, V. Begy, and G. Louppe, “Likelihood-free MCMC with Amortized Approximate Ratio Estimators,” [arXiv:1903.04057 \[stat.ML\]](https://arxiv.org/abs/1903.04057).
 - [72] G. Zhang, S. Mishra-Sharma, and C. Dvorkin, “Inferring subhalo effective density slopes from strong lensing observations with neural likelihood-ratio estimation,” *Mon. Not. Roy. Astron. Soc.* **517** no. 3, (2022) 4317–4326, [arXiv:2208.13796 \[astro-ph.CO\]](https://arxiv.org/abs/2208.13796).

- [73] S. Wagner-Carena, J. Aalbers, S. Birrer, E. O. Nadler, E. Darragh-Ford, P. J. Marshall, and R. H. Wechsler, “From Images to Dark Matter: End-to-end Inference of Substructure from Hundreds of Strong Gravitational Lenses,” *Astrophys. J.* **942** no. 2, (2023) 75, [arXiv:2203.00690 \[astro-ph.CO\]](https://arxiv.org/abs/2203.00690).
- [74] J. Brehmer, S. Mishra-Sharma, J. Hermans, G. Louppe, and K. Cranmer, “Mining for Dark Matter Substructure: Inferring subhalo population properties from strong lenses with machine learning,” *Astrophys. J.* **886** no. 1, (2019) 49, [arXiv:1909.02005 \[astro-ph.CO\]](https://arxiv.org/abs/1909.02005).
- [75] A. S. Bolton, S. Burles, L. V. E. Koopmans, T. Treu, and L. A. Moustakas, “The sloan lens acs survey. 1. a large spectroscopically selected sample of massive early-type lens galaxies,” *Astrophys. J.* **638** (2006) 703–724, [arXiv:astro-ph/0511453](https://arxiv.org/abs/astro-ph/0511453).
- [76] **BOSS** Collaboration, J. R. Brownstein *et al.*, “The BOSS Emission-Line Lens Survey (BELLS). I. A large spectroscopically selected sample of Lens Galaxies at redshift ~ 0.5 ,” *Astrophys. J.* **744** (2012) 41–62, [arXiv:1112.3683 \[astro-ph.CO\]](https://arxiv.org/abs/1112.3683).
- [77] J. P. Gardner *et al.*, “The James Webb Space Telescope,” *Space Sci. Rev.* **123** (2006) 485, [arXiv:astro-ph/0606175](https://arxiv.org/abs/astro-ph/0606175).
- [78] J. D. Simon, S. Birrer, K. Bechtol, S. Chakrabarti, F.-Y. Cyr-Racine, I. Dell’Antonio, A. Drlica-Wagner, C. Fassnacht, M. Geha, D. Gilman, Y. D. Hezaveh, D. Kim, T. S. Li, L. Strigari, and T. Treu, “Testing the nature of dark matter with extremely large telescopes,” [1903.04742](https://arxiv.org/abs/1903.04742). <https://arxiv.org/pdf/1903.04742.pdf>.
- [79] **Euclid Imaging** Collaboration, A. Refregier, A. Amara, T. D. Kitching, A. Rassat, R. Scaramella, and J. Weller, “Euclid Imaging Consortium Science Book,” [arXiv:1001.0061 \[astro-ph.IM\]](https://arxiv.org/abs/1001.0061).
- [80] **EUCLID** Collaboration, R. Laureijs *et al.*, “Euclid Definition Study Report,” [arXiv:1110.3193 \[astro-ph.CO\]](https://arxiv.org/abs/1110.3193).
- [81] L. V. E. Koopmans, I. W. A. Browne, and N. J. Jackson, “Strong gravitational lensing with SKA,” *New Astron. Rev.* **48** (2004) 1085–1094, [arXiv:astro-ph/0409095](https://arxiv.org/abs/astro-ph/0409095).
- [82] **LSST Dark Energy Science** Collaboration, B. Abolfathi *et al.*, “The LSST DESC DC2 Simulated Sky Survey,” *Astrophys. J. Suppl.* **253** no. 1, (2021) 31, [arXiv:2010.05926 \[astro-ph.IM\]](https://arxiv.org/abs/2010.05926).
- [83] T. E. Collett, “The population of galaxy-galaxy strong lenses in forthcoming optical imaging surveys,” *Astrophys. J.* **811** no. 1, (2015) 20, [arXiv:1507.02657 \[astro-ph.CO\]](https://arxiv.org/abs/1507.02657).
- [84] J. P. McKean, N. Jackson, S. Vegetti, M. Rybak, S. Serjeant, L. V. E. Koopmans, R. B. Metcalf, C. D. Fassnacht, P. J. Marshall, and M. Pandey-Pommier, “Strong gravitational lensing with the ska,” [1502.03362](https://arxiv.org/abs/1502.03362). <https://arxiv.org/pdf/1502.03362.pdf>.

Bibliography

- [85] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. Curran Associates, Inc., 2019. <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- [86] M. Meneghetti, “Introduction to gravitational lensing,” 11, 2016.
https://www.ita.uni-heidelberg.de/~jmerten/misc/meneghetti_lensing.pdf.
- [87] **Astropy** Collaboration, T. P. Robitaille *et al.*, “Astropy: A Community Python Package for Astronomy,” *Astron. Astrophys.* **558** (2013) A33, [arXiv:1307.6212 \[astro-ph.IM\]](https://arxiv.org/abs/1307.6212).
- [88] **Astropy** Collaboration, A. M. Price-Whelan *et al.*, “The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package,” *Astron. J.* **156** no. 3, (2018) 123, [arXiv:1801.02634](https://arxiv.org/abs/1801.02634).
- [89] **Planck** Collaboration, N. Aghanim *et al.*, “Planck 2018 results. VI. Cosmological parameters,” *Astron. Astrophys.* **641** (2020) A6, [arXiv:1807.06209 \[astro-ph.CO\]](https://arxiv.org/abs/1807.06209). [Erratum: *Astron.Astrophys.* 652, C4 (2021)].
- [90] K. Karchev, A. Coogan, and C. Weniger, “Strong-lensing source reconstruction with variationally optimized Gaussian processes,” *Mon. Not. Roy. Astron. Soc.* **512** no. 1, (2022) 661–685, [arXiv:2105.09465 \[astro-ph.IM\]](https://arxiv.org/abs/2105.09465).
- [91] M. Bartelmann and P. Schneider, “Weak gravitational lensing,” *Phys. Rept.* **340** (2001) 291–472, [arXiv:astro-ph/9912508](https://arxiv.org/abs/astro-ph/9912508).
- [92] J. L. Sérsic, “Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy,” *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina* **6** (Feb., 1963) 41–43.
- [93] L. Ciotti and G. Bertin, “Analytical properties of the $r^{-(1/m)}$ luminosity law,” *Astron. Astrophys.* **352** (1999) 447, [arXiv:astro-ph/9911078](https://arxiv.org/abs/astro-ph/9911078).
- [94] S. H. Suyu, P. J. Marshall, R. D. Blandford, C. D. Fassnacht, L. V. E. Koopmans, J. P. McKean, and T. Treu, “Dissecting the Gravitational Lens B1608+656: Lens Potential Reconstruction,” *Astrophys. J.* **691** (2009) 277–298, [arXiv:0804.2827 \[astro-ph\]](https://arxiv.org/abs/0804.2827).
- [95] N. Tessore and R. B. Metcalf, “The elliptical power law profile lens,” *Astron. Astrophys.* **580** (2015) A79, [arXiv:1507.01819 \[astro-ph.CO\]](https://arxiv.org/abs/1507.01819).
- [96] C. M. O’Riordan, S. J. Warren, and D. J. Mortlock, “Galaxy mass profiles from strong lensing ii: The elliptical power-law model,” [2004.13435](https://arxiv.org/abs/2004.13435).
<https://arxiv.org/pdf/2004.13435.pdf>.
- [97] M. Chianese, A. Coogan, P. Hofma, S. Otten, and C. Weniger, “Differentiable Strong Lensing: Uniting Gravity and Neural Nets through Differentiable Probabilistic Programming,” *Mon. Not. Roy. Astron. Soc.* **496** no. 1, (2020) 381–393,

- [arXiv:1910.06157 \[astro-ph.CO\]](https://arxiv.org/abs/1910.06157).
- [98] P. Schneider and D. Sluse, “Mass-sheet degeneracy, power-law models and external convergence: Impact on the determination of the Hubble constant from gravitational lensing,” *Astron. Astrophys.* **559** (2013) A37, [arXiv:1306.0901 \[astro-ph.CO\]](https://arxiv.org/abs/1306.0901).
 - [99] P. Schneider and D. Sluse, “Source-position transformation – an approximate invariance in strong gravitational lensing,” *Astron. Astrophys.* **564** (2014) A103, [arXiv:1306.4675 \[astro-ph.CO\]](https://arxiv.org/abs/1306.4675).
 - [100] J. Nightingale and S. Dye, “Adaptive semi-linear inversion of strong gravitational lens imaging,” **1412.7436**. <https://arxiv.org/pdf/1412.7436.pdf>.
 - [101] G. Despali, S. Vegetti, S. D. M. White, C. Giocoli, and F. C. van den Bosch, “Modelling the line-of-sight contribution in substructure lensing,” *Mon. Not. Roy. Astron. Soc.* **475** no. 4, (2018) 5424–5442, [arXiv:1710.05029 \[astro-ph.CO\]](https://arxiv.org/abs/1710.05029).
 - [102] Q. He *et al.*, “Galaxy–galaxy strong lens perturbations: line-of-sight haloes versus lens subhaloes,” *Mon. Not. Roy. Astron. Soc.* **512** no. 4, (2022) 5862–5873, [arXiv:2110.04512 \[astro-ph.CO\]](https://arxiv.org/abs/2110.04512).
 - [103] E. A. Baltz, P. Marshall, and M. Oguri, “Analytic models of plausible gravitational lens potentials,” *JCAP* **01** (2009) 015, [arXiv:0705.0682 \[astro-ph\]](https://arxiv.org/abs/0705.0682).
 - [104] J. F. Navarro, C. S. Frenk, and S. D. M. White, “A Universal density profile from hierarchical clustering,” *Astrophys. J.* **490** (1997) 493–508, [arXiv:astro-ph/9611107](https://arxiv.org/abs/astro-ph/9611107).
 - [105] J. Richings, C. Frenk, A. Jenkins, A. Robertson, and M. Schaller, “A high-resolution cosmological simulation of a strong gravitational lens,” *Mon. Not. Roy. Astron. Soc.* **501** no. 3, (2021) 4657–4668, [arXiv:2005.14495 \[astro-ph.CO\]](https://arxiv.org/abs/2005.14495).
 - [106] N. C. Amorisco *et al.*, “Halo concentration strengthens dark matter constraints in galaxy–galaxy strong lensing analyses,” *Mon. Not. Roy. Astron. Soc.* **510** no. 2, (2022) 2464–2479, [arXiv:2109.00018 \[astro-ph.CO\]](https://arxiv.org/abs/2109.00018).
 - [107] C. Giocoli, G. Tormen, R. K. Sheth, and F. C. van den Bosch, “The Substructure Hierarchy in Dark Matter Haloes,” *Mon. Not. Roy. Astron. Soc.* **404** (2010) 502–517, [arXiv:0911.0436 \[astro-ph.CO\]](https://arxiv.org/abs/0911.0436).
 - [108] G. Despali and S. Vegetti, “The impact of baryonic physics on the subhalo mass function and implications for gravitational lensing,” *Mon. Not. Roy. Astron. Soc.* **469** no. 2, (2017) 1997–2010, [arXiv:1608.06938 \[astro-ph.GA\]](https://arxiv.org/abs/1608.06938).
 - [109] V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J. F. Navarro, C. S. Frenk, and S. D. M. White, “The Aquarius Project: the subhalos of galactic halos,” *Mon. Not. Roy. Astron. Soc.* **391** (2008) 1685–1711, [arXiv:0809.0898 \[astro-ph\]](https://arxiv.org/abs/0809.0898).
 - [110] J. L. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. S. Warren, G. Yepes, S. Gottlöber, and D. E. Holz, “Toward a halo mass function for precision cosmology: The Limits of universality,” *Astrophys. J.* **688** (2008) 709–728, [arXiv:0803.2706](https://arxiv.org/abs/0803.2706)

[[astro-ph](#)].

- [111] P. Fleury, J. Larena, and J.-P. Uzan, “Line-of-sight effects in strong gravitational lensing,” *JCAP* **08** (2021) 024, [arXiv:2104.08883 \[astro-ph.CO\]](#).
- [112] A. Boyarsky, M. Drewes, T. Lasserre, S. Mertens, and O. Ruchayskiy, “Sterile neutrino Dark Matter,” *Prog. Part. Nucl. Phys.* **104** (2019) 1–45, [arXiv:1807.07938 \[hep-ph\]](#).
- [113] J. R. Bond, A. S. Szalay, and M. S. Turner, “Formation of Galaxies in a Gravitino Dominated Universe,” *Phys. Rev. Lett.* **48** (1982) 1636.
- [114] A. Schneider, R. E. Smith, A. V. Maccio, and B. Moore, “Nonlinear Evolution of Cosmological Structures in Warm Dark Matter Models,” *Mon. Not. Roy. Astron. Soc.* **424** (2012) 684, [arXiv:1112.0330 \[astro-ph.CO\]](#).
- [115] M. R. Lovell, “Towards a general parametrization of the warm dark matter halo mass function,” [arXiv:2003.01125 \[astro-ph.CO\]](#).
- [116] B. K. Miller, A. Cole, C. Weniger, F. Nattino, O. Ku, and M. W. Grootes, “swyft: Truncated Marginal Neural Ratio Estimation in Python,” *J. Open Source Softw.* **7** no. 75, (2022) 4205.
- [117] I. Tolstikhin, N. Houlsby, A. Kolesnikov, L. Beyer, X. Zhai, T. Unterthiner, J. Yung, A. Steiner, D. Keysers, J. Uszkoreit, M. Lucic, and A. Dosovitskiy, “Mlp-mixer: An all-mlp architecture for vision,” [2105.01601](#).
<https://arxiv.org/pdf/2105.01601.pdf>.
- [118] A. Dosovitskiy *et al.*, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” [arXiv:2010.11929 \[cs.CV\]](#).
- [119] A. Cole, B. K. Miller, S. J. Witte, M. X. Cai, M. W. Grootes, F. Nattino, and C. Weniger, “Fast and credible likelihood-free cosmology with truncated marginal neural ratio estimation,” *JCAP* **09** (2022) 004, [arXiv:2111.08030 \[astro-ph.CO\]](#).
- [120] A. Gu *et al.*, “GIGA-Lens: Fast Bayesian Inference for Strong Gravitational Lens Modeling,” *Astrophys. J.* **935** no. 1, (2022) 49, [arXiv:2202.07663 \[astro-ph.IM\]](#).
- [121] H. Iqbal, “Harisiqbali88/plotneuralnet v1.0.0,” Dec., 2018.
<https://doi.org/10.5281/zenodo.2526396>.
- [122] J. Hermans, A. Delaunoy, F. Rozet, A. Wehenkel, V. Begy, and G. Louppe, “A Trust Crisis In Simulation-Based Inference? Your Posterior Approximations Can Be Unfaithful,” [arXiv:2110.06581 \[stat.ML\]](#).
- [123] S. H. Suyu, P. J. Marshall, M. P. Hobson, and R. D. Blandford, “A bayesian analysis of regularised source inversions in gravitational lensing,” *Mon. Not. Roy. Astron. Soc.* **371** (2006) 983–998, [arXiv:astro-ph/0601493](#).
- [124] S. Birrer and A. Amara, “Lenstronomy: multi-purpose gravitational lens modelling software package,” [arXiv:1803.09746 \[astro-ph.CO\]](#).
- [125] A. Galan, A. Peel, R. Joseph, F. Courbin, and J. L. Starck, “SLITronomy: towards a

Bibliography

- fully wavelet-based strong lensing inversion technique,” *Astron. Astrophys.* **647** (2021) A176, [arXiv:2012.02802 \[astro-ph.GA\]](https://arxiv.org/abs/2012.02802).
- [126] A. Adam, L. Perreault-Levasseur, and Y. Hezaveh, “Pixelated Reconstruction of Gravitational Lenses using Recurrent Inference Machines,” in *39th International Conference on Machine Learning Conference*. 7, 2022. [arXiv:2207.01073 \[astro-ph.IM\]](https://arxiv.org/abs/2207.01073).
- [127] P. Salucci, “The distribution of dark matter in galaxies,” *2019 Astronomy & Astrophysics Rv.* **27....2S** (11, 2018) , [1811.08843](https://arxiv.org/abs/1811.08843).
<https://arxiv.org/pdf/1811.08843.pdf>.