
Targeted Likelihood-Free Inference of Dark Matter Substructure in Strongly-Lensed Galaxies

Adam Coogan¹
a.m.coogan@uva.nl

Konstantin Karchev^{1,2}
kkarchev@sissa.it

Christoph Weniger¹
c.weniger@uva.nl

¹Gravitation Astroparticle Physics Amsterdam (GRAPPA)
Institute for Theoretical Physics (ITFA)
University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

²SISSA (Scuola Internazionale Superiore di Studi Avanzati)
via Bonomea 265, I-34136 Trieste, Italy.

Abstract

The analysis of optical images of galaxy-galaxy strong gravitational lensing systems can provide important information about the distribution of dark matter at small scales. However, the modeling and statistical analysis of these images is extraordinarily complex, bringing together source image and main lens reconstruction, hyper-parameter optimization, and the marginalization over small-scale structure realizations. We present here a new analysis pipeline that tackles these diverse challenges by bringing together many recent machine learning developments in one coherent approach, including variational inference, Gaussian processes, differentiable probabilistic programming, and neural likelihood-to-evidence ratio estimation. Our pipeline enables: (a) fast reconstruction of the source image and lens mass distribution, (b) variational estimation of uncertainties, (c) efficient optimization of source regularization and other hyperparameters, and (d) marginalization over stochastic model components like the distribution of substructure. We present here preliminary results that demonstrate the validity of our approach.

1 Introduction

The existence of dark matter, its overall abundance and its distribution at galactic scales are reasonably well-understood thanks to its gravitational interactions. However, determining the properties of the fundamental constituents of dark matter remains one of the major unresolved issues in physics [Bertone et al., 2005].

Detection of the smallest dark matter structures on subgalactic scales is a potential inroad. In the standard Λ CDM model of cosmology, dark matter is cold and noninteracting. As a result, n -body simulations predict that galactic dark matter halos should contain exponentially-many lower-mass substructures called *subhalos*. The subhalo mass function dn/dm describing the number of subhalos at a given mass can be different if dark matter is instead warm or has nongravitational interactions. In these theories, overdensities below a particular threshold are unable to collapse into subhalos, and the subhalo mass function drops to zero below some threshold (see e.g. Viel et al. [2013], Schneider et al. [2012]).

Since star formation is suppressed in subhalos with mass below $\sim 10^9 M_\odot$, it is difficult to detect light from these halos directly. Instead, searching for their gravitational effects is a promising approach. In this work we study strong galaxy-galaxy gravitational lenses, where a source galaxy’s light is dramatically distorted and multiply imaged by an intervening lens galaxy. The gravitational influence of subhalos imprints additional percent-level perturbations on the resulting observed image. To date, analyses of these systems have yielded two subhalo detections [Vegetti et al., 2010, 2012]

Fitting a lensing system and properly marginalizing over variations in the source, lens and subhalos to constrain properties of substructure is an extremely difficult problem. This is due to the dramatic range of possible source galaxy light distributions and the large number of parameters that must be marginalized over in realistic lensing models in order to perform substructure inference. On one hand, methods such the one in Vegetti and Koopmans [2009] use Bayesian lensing models to closely fit lensing observations and conventional sampling methods to compute posteriors for subhalo parameters. However, conventional sampling methods require sampling from the joint posterior over all model parameters. These methods are thus restricted to searching for effects from one or two subhalos. On the other hand, since many works have shown neural networks have the capacity to measure main lens parameters [Hezaveh et al., 2017, Morningstar et al., 2019] and detect subhalos [Diaz Rivero and Dvorkin, 2020, Ostdiek et al., 2020a,b], Brehmer et al. [2019] leveraged them through likelihood-free inference to marginalize over large numbers of subhalos and directly compute marginal posteriors for subhalo mass function parameters. However, this approach is extremely data-hungry when applied to even toy problems, since it requires amortizing over all possible variations in lensing systems.

In this work we merge the strengths of both approaches to perform targeted inference on individual lensing systems. In the first stage of our analysis, we introduce a new model for lensing systems that describes the source light distribution using an approximate Gaussian process [Rasmussen and Williams, 2006]. By implementing this in an automatically-differentiable manner and leveraging variational inference [Saul et al., 1996, Jordan et al., 1999, Hoffman et al., 2013, Bishop, 2006], we fit approximate posteriors for all $\mathcal{O}(10^5)$ source and lens parameters to an observation using gradient descent. By sampling from this approximate posterior, we generate training data that looks *very similar* to the observation. In the second analysis step, we use this data to train neural networks using the tools of likelihood-free inference to produce posterior distributions for substructure parameters. Applying these networks to the observation of interest yields inference results.

In the remainder of this paper, we describe our new model for lensing systems, our inference procedure, and present results from analyzing a mock lensing system. In upcoming work we will elaborate on our new source model [Karchev et al., 2020] and substructure inference approach [Coogan et al., 2020].

2 Modelling

Here we present a flexible Bayesian lensing model that has the capacity to model complex, high-resolution images and can be fit using gradient descent. The base model has three components: a lens, with parameters denoted z_l , a source, parametrized by z_s , and the instrument response, which here is simply pixel noise assumed to be normally distributed with known standard deviation. Ultimately, we consider all of the parameters of the base model as nuisance parameters that we will need to marginalise out in order to infer an additional set of lens dark matter substructure parameters (e.g. the position and mass of a single subhalo, as we present below), denoted z_d .

The lens models the presence of gravitating mass along the line of sight which bends the path of light (see fig. 3). In the thin lens formalism this process is described by the *lens equation* $\vec{\mathbf{p}} = \vec{\xi} - \vec{\alpha}(\vec{\xi})$, which relates image plane coordinates $\vec{\xi}$ to source plane coordinates $\vec{\mathbf{p}}$ via the displacement field $\vec{\alpha}$. These are all arrays of N 2-dimensional vectors, where N is the number of pixels in the analysed image. The displacement $\vec{\alpha}$ is a linear function of the overall projected mass distribution $\Sigma(\vec{\xi})$ in the image plane, which means that different mass components can be freely superposed. The displacement due to even a single component, however, can be a complex function of position, and in images of interest leads to the projection of disjoint sections of the image onto the same region in the source plane.

We include two smooth analytic components for the lens: a singular power-law ellipsoid (SPLE) for the main lens galaxy and external shear (see e.g. Chianese et al. [2020] for the displacement fields). This model has previously been used to model observational data and has been shown capable of modeling the combined distribution of dark and baryonic matter in the inner regions of galaxies at the percent level [Suyu et al., 2009]. It contains a total of eight parameters, which form \mathbf{z}_l .

We use a novel approach to modelling the sources in strongly lensed systems which has two main aims: to have a regularising effect so that the lens parameters can be constrained, and to treat source uncertainties so that they can be disentangled from the effect of lensing substructure. This is achieved with a generative model inspired by Gaussian processes, briefly described as:

$$\mathbf{f} = \mathbf{T}(\vec{\mathbf{p}}, \sigma) \mathbf{y}, \quad \text{with } \mathbf{y} \sim \mathcal{N}(\mathbf{0}, \alpha^2). \quad (1)$$

Here \mathbf{y} is an array of N flux parameters associated to the image pixels and assumed to have a Gaussian prior, while \mathbf{T} is a *transfer matrix*, defined below such that the true source fluxes \mathbf{f} have a given covariance $\mathbf{K} = \alpha^2 \mathbf{T} \mathbf{T}^T$. We use a Gaussian radial basis function to model the covariance between fluxes at two *points* in the source plane:

$$\text{cov}(f(\vec{p}_1), f(\vec{p}_2)) = k(\vec{p}_1, \vec{p}_2) = \alpha^2 \exp\left(-\frac{|\vec{p}_1 - \vec{p}_2|^2}{2\sigma^2}\right). \quad (2)$$

Its hyperparameters are α^2 and σ , describing respectively the prior variance and the correlation scale in the source plane.

However, we are interested in the covariance matrix of the light received in *pixels*, not at points. This depends on the intrinsic covariance between different points in the source (captured by eq. (2)) as well as the covariance due to the overlap between pixels projected back into the source plane. Accounting for this requires integrating eq. (2) over the overlap between each pair of back-projected pixels. Since this is extremely slow to compute exactly, we approximate each back-projected pixel using a 2D Gaussian (see fig. 4): $g_i(\vec{p}) = \mathcal{G}(\vec{p} - \vec{p}_i, \Sigma_i)$, with covariance matrix Σ_i derived to match the shape and total area of the pixel. Then the covariance matrix becomes

$$\text{cov}(\mathbf{f})_{ij} = \mathbf{K}_{ij} = \iint k(\vec{p}_1, \vec{p}_2) g_i(\vec{p}_1) g_j(\vec{p}_2) d^2\vec{p}_1 d^2\vec{p}_2 = 2\pi\alpha^2\sigma^2 \mathcal{G}(\vec{p}_i - \vec{p}_j, \Sigma_i + \sigma^2 \mathbf{I} + \Sigma_j). \quad (3)$$

Usually, one needs to calculate the transfer matrix \mathbf{T} from \mathbf{K} by e.g. matrix square root or Cholesky decomposition, but due to the high dimensionality of the problem these operations are infeasible. Instead, we realise that the outer product $\mathbf{T} \mathbf{T}^T$ is akin to a spatial convolution, and that the convolution of a Gaussian with itself is a Gaussian with twice the variance, which allows us to approximate $\mathbf{T}_{ik} \sim \mathcal{G}(\vec{p}_i - \vec{p}_k, \Sigma_i + \frac{\sigma^2}{2} \mathbf{I})$. The proper normalisation depends on the number density of pixels in the source plane; we give a full derivation in Karchev et al. [2020].

Finally, instead of allowing the kernel size to vary, we consider a number of independent *GP layers* with fixed $\sigma_{(k)}$ ranging from galaxy- to pixel-size in order to model details on various scales. Each layer thus has an independently-optimised array of flux parameters $\mathbf{y}_{(k)}$ and overall variance $\alpha_{(k)}^2$. The collection of all $\mathbf{y}_{(k)}$ forms the source parameter array \mathbf{z}_s , which together with the hyperparameters $\alpha_{(k)}^2$ are optimized variationally.

We implement our lensing model using the `pytorch` and `pykeops` libraries, which ensures the model output is automatically-differentiable with respect to all (hyper)parameters: \mathbf{z}_l , \mathbf{z}_s , $\alpha_{(k)}^2$, and enables us to perform all calculations on graphical processing units.

3 Statistical analysis

The overall analysis strategy splits in two steps.

1. Fit an approximate posterior $q_\phi(\mathbf{z}_s, \mathbf{z}_l | \mathbf{x}_0)$ for the lens and source parameters \mathbf{z}_l and \mathbf{z}_s to an observation \mathbf{x}_0 using variational inference, simultaneously optimizing the GP hyperparameters $\theta = \{\alpha_{(k)}^2\}$.
2. Sample data from this constrained model to train an inference network to predict $p(\mathbf{z}_d | \mathbf{x}_0)$, the marginal posterior for substructure parameters \mathbf{z}_d for the observation.

Variational inference We approximate the posterior over the source and lens parameters with a multivariate normal distribution over the lens parameters and a diagonal normal distribution over the source parameters:

$$q_\phi(\mathbf{z}_l, \mathbf{z}_s | \mathbf{x}_0) = \mathcal{G}_{\phi_l}(\mathbf{z}_l) \mathcal{N}_{\phi_s}(\mathbf{z}_s), \quad (4)$$

where $\phi = \{\phi_l, \phi_s\}$ denote the distribution’s mean and covariance parameters. This approximation is justified since all parameters are reasonably well-constrained. While it neglects source parameter correlations, the diagonal normal guide for the $\mathcal{O}(10^5)$ source parameters is necessary since the matrix inversion required for standard GP inference [Rasmussen and Williams, 2006] is prohibitively expensive. Since close-by source parameters are in general anti-correlated, we find that the neglect of correlations increases the variance of the posterior predictive distribution, which is consistent with the goal of using samples from the fitted model for training targeted neural inference networks. We leave a quantitative study of these effects to future work.

In order to optimize the above approximate posterior with respect to ϕ and also optimize the GP hyperparameters, we use gradient descent to maximize the evidence lower bound (ELBO) [Saul et al., 1996, Jordan et al., 1999, Bishop, 2006, Hoffman et al., 2013] for the observation, $\mathbb{E}_{\mathbf{z} \sim q_\phi(\mathbf{z} | \mathbf{x}_0)} [\ln p_\theta(\mathbf{x}_0, \mathbf{z}) - \ln q_\phi(\mathbf{z} | \mathbf{x}_0)]$. To this end, we use the probabilistic programming language pyro, and the auto-differentiation capabilities of pytorch.

Likelihood-free inference Next we seek the posterior $p(\mathbf{z}_d | \mathbf{x}_0)$ for the substructure parameters, marginalized over the lens and source parameters $\eta \equiv \{\mathbf{z}_l, \mathbf{z}_s\}$. The integral over η is intractable since it is very high-dimensional and the integrand is a complex nonlinear function. Instead, we approximate $p(\mathbf{z}_d | \mathbf{x}_0)$ using neural likelihood-to-evidence ratio estimation techniques [Hermans et al., 2019].

The strategy is to estimate the ratio $r(\mathbf{x}, \mathbf{z}_d) \equiv p(\mathbf{z} | \mathbf{x}) / p(\mathbf{z})$, from which the marginal posterior is easily recovered. To do this we train a classification network $d(\mathbf{x}, \mathbf{z}_d) \in [0, 1]$ to discriminate the hypotheses. In the first, the data and substructure parameters are drawn jointly from the parameter priors and model: $\mathbf{x}, \mathbf{z}_d \sim p(\mathbf{x}, \mathbf{z}_d)$. In the second, the data and substructure parameters are sampled marginally: $\mathbf{x}, \mathbf{z}_d \sim p(\mathbf{x})p(\mathbf{z}_d)$. We adopt the binary cross-entropy loss function and optimize the parameters of the classification network using gradient descent. After this we can compute $r(\mathbf{x}_0, \mathbf{z}_d) = d(\mathbf{x}_0, \mathbf{z}_d) / (1 - d(\mathbf{x}_0, \mathbf{z}_d))$ and obtain the substructure parameter posteriors for the observation of interest.

Formally, marginalizing over the source and lens requires training the classification network on all observationally possible lensing images. Framing this in terms of our lensing model, generating this training dataset requires sampling the lens, source and substructure parameters from their priors, computing the lensed image μ , and generating an observation by adding Gaussian pixel noise with standard deviation σ_n :

$$\mathbf{x} \sim \mathcal{N}(\mathbf{x} | \mu(\mathbf{z}_l, \mathbf{z}_s, \mathbf{z}_d), \sigma_n) p(\mathbf{z}_l) p(\mathbf{z}_s) p(\mathbf{z}_d). \quad (5)$$

Since we eventually are interested in sub-percent variations in the images, this is an extraordinarily complex endeavor, which would require a very large amount of training data as well as a network with very high capacity.

Instead, we propose a much simpler and more tractable approach where we *target* the analysis on a particular observation by constraining the sampling of the lens and source parameters. In particular, we sample $\mathbf{z}_l, \mathbf{z}_s \sim q_\phi(\mathbf{z}_l, \mathbf{z}_s | \mathbf{x}_0)$ rather than from their priors. This training data generation procedure excludes lensing systems incompatible with the observation being analyzed, dramatically decreasing the amount of data required for the classification network, allowing us to use simple architectures and shorter training times.

4 Results and discussion

As a preliminary test of our analysis procedure, we construct a 400×400 mock observation based on the Extremely Large Telescope’s [Observatory] capabilities that contains a single $\sim 10^9 M_\odot$ subhalo. The mock observation and source are shown in fig. 1, along with the mean observation and source reconstructed with a five-layer GP source model in our first analysis step. The observation reconstruction closely matches the mock observation, and captures very fine details in the source. The inferred values of the lens parameters are within $\sim 0.1 - 5\%$ of the true parameter values. For the

second analysis step, we train simple neural networks built from convolutional and fully-connected layers to give 2D and 1D posteriors for the subhalo’s position and mass, marginalized over the 174,458 source and lens parameters. The resulting posteriors, obtained with just 10,000 training examples, are shown in fig. 2. Our subhalo position reconstructions are accurate, with the true position of the subhalo in the mock observation (right column) lying within the 68% containment region. The mass posterior is centered close to the true subhalo mass, though it is reasonably broad. The whole analysis runs in a few hours on a single Nvidia GTX TitanX GPU.

Applying our pipeline to real data will require extending the lensing model. A more realistic version of our lensing system would contain ~ 320 subhalos above $10^5 M_\odot$ within the observation region. Dark matter halos along the line of sight are additionally expected to contribute 3 – 10 times more lensing distortions than subhalos [Despali et al., 2018]. Lastly, we have omitted the main lens’ light. All of these components are straightforward to incorporate in our framework. In upcoming work we will apply this extended analysis to search for subhalos and set constraints on the subhalo mass function using existing lensing observations in future work. More generally, we anticipate our approximate Gaussian process model and targeted likelihood-free inference strategy will find other exciting applications to difficult astrophysical imaging and data analysis problems.

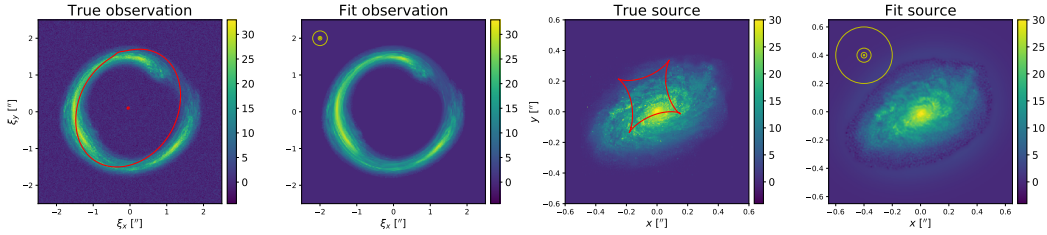


Figure 1: From left to right: the mock observation, fit observation, true source and fit source. The displayed fits are computed using the mean estimates of the source and lens parameters from the approximate posterior $q_\phi(\mathbf{z}|\mathbf{x})$. The red curves show the critical curve and caustic of the main lens. The yellow circles indicate the spatial scales of the five Gaussian process layers, the smallest of which is not visible here.

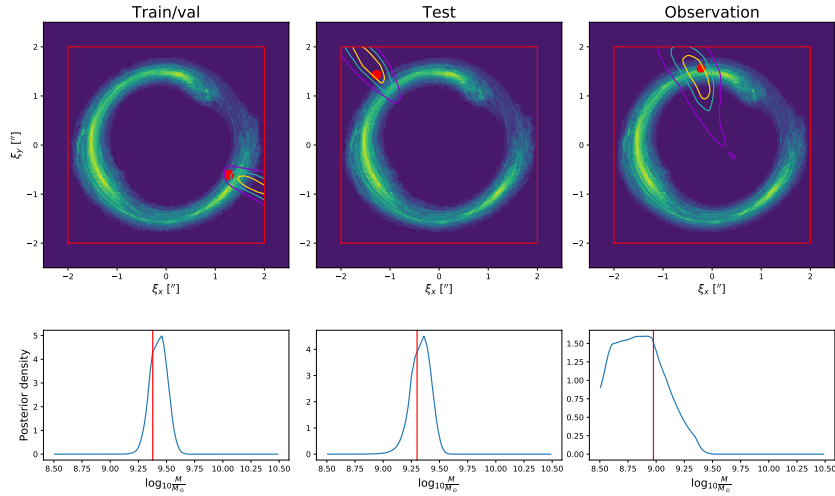


Figure 2: Preliminary results from our pipeline: posteriors for the subhalo position and mass for a systems from the training/validation set (left panel), a system from a held-out test set (middle), and the observation being analyzed (right). The purple, teal and yellow contours in the first row show the 99.7%, 95% and 68% containment regions. The red dots indicate the true subhalo position and the red box is the area from which the subhalos were uniformly sampled. The red lines in the second row mark the true subhalo masses. These were obtained with only 10,000 training examples.

Broader Impact

This work is focusing on the precision analysis of astronomical images based on forward models. Variants of the presented approach could be applicable to other areas of the physical sciences. Although we do not anticipate potential for misuse of the presented methods or the danger of grossly biased results, the usual care has to be exercised when drawing scientific conclusions based on a complex analysis machinery.

Acknowledgments and Disclosure of Funding

We thank Marco Chianese, Camila Correa, Gilles Louppe, Ben Miller and Simona Vegetti for useful discussions.

This work uses `numpy` [Harris et al., 2020], `scipy` [Virtanen et al., 2020], `matplotlib` [Hunter, 2007], `pyro` [Bingham et al., 2018], `pytorch` [Paszke et al., 2019], `pykeops` [Charlier et al., 2020], `jupyter` [Kluyver et al., 2016] and `tqdm` [da Costa-Luis et al., 2020]. Computations were carried out on the DAS-5 [Bal et al., 2016] cluster, which is funded by the Netherlands Organization for Scientific Research (NWO/NCF).

References

- H. Bal, D. Epema, C. de Laat, R. van Nieuwpoort, J. Romein, F. Seinstra, C. Snoek, and H. Wijshoff. A medium-scale distributed system for computer science research: Infrastructure for the long term. *Computer*, 49(05):54–63, may 2016. ISSN 1558-0814. doi: 10.1109/MC.2016.127.
- Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: Evidence, candidates and constraints. *Phys. Rept.*, 405:279–390, 2005. doi: 10.1016/j.physrep.2004.08.031.
- Eli Bingham, Jonathan P. Chen, Martin Jankowiak, Fritz Obermeyer, Neeraj Pradhan, Theofanis Karaletsos, Rohit Singh, Paul Szerlip, Paul Horsfall, and Noah D. Goodman. Pyro: Deep Universal Probabilistic Programming. *Journal of Machine Learning Research*, 2018.
- Christopher M. Bishop. *Pattern Recognition and Machine Learning*, chapter 10.1, pages 461–474. Springer Science and Business Media LLC, 2006. URL <https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>.
- Johann Brehmer, Siddharth Mishra-Sharma, Joeri Hermans, Gilles Louppe, and Kyle Cranmer. Mining for dark matter substructure: Inferring subhalo population properties from strong lenses with machine learning. *The Astrophysical Journal*, 886(1):49, nov 2019. doi: 10.3847/1538-4357/ab4c41. URL <https://doi.org/10.3847/1538-4357/2Fab4c41>.
- Benjamin Charlier, Jean Feydy, Joan Alexis Glaunès, François-David Collin, and Ghislain Durif. Kernel operations on the GPU, with autodiff, without memory overflows. *arXiv preprint arXiv:2004.11127*, 2020.
- Marco Chianese, Adam Coogan, Paul Hofma, Sydney Otten, and Christoph Weniger. Differentiable Strong Lensing: Uniting Gravity and Neural Nets through Differentiable Probabilistic Programming. *Mon. Not. Roy. Astron. Soc.*, 496(1):381–393, 2020. doi: 10.1093/mnras/staa1477.
- Adam Coogan, Konstantin Karchev, Benjamin Miller, and Christoph Weniger. Precision searches for subhalos in strong lensing images with targeted inference networks. 2020.
- Casper da Costa-Luis, Stephen Karl Larroque, Kyle Altendorf, Hadrien Mary, Mikhail Korobov, Noam Yorav-Raphael, Ivan Ivanov, Marcel Bargull, Nishant Rodrigues, Guangshuo CHEN, Charles Newey, James, Martin Zugnoni, Matthew D. Pagel, mjstevens777, Mikhail Dektyarev, Alex Rothberg, Alexander, Daniel Panteleit, Fabian Dill, FichteFoll, HeoHeo, Hugo van Kemenade, Jack McCracken, Max Nordlund, Orivej Desh, RedBug312, richardsheridan, Socialery, and Staffan Malmgren. `tqdm`: A fast, Extensible Progress Bar for Python and CLI, September 2020. URL <https://doi.org/10.5281/zenodo.4054194>.
- Giulia Despali, Simona Vegetti, Simon D. M. White, Carlo Giocoli, and Frank C. van den Bosch. Modelling the line-of-sight contribution in substructure lensing. *Mon. Not. Roy. Astron. Soc.*, 475(4):5424–5442, 2018. doi: 10.1093/mnras/sty159.
- Ana Diaz Rivero and Cora Dvorkin. Direct Detection of Dark Matter Substructure in Strong Lens Images with Convolutional Neural Networks. *Phys. Rev. D*, 101(2):023515, 2020. doi: 10.1103/PhysRevD.101.023515.

- Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, and et al. Array programming with numpy. *Nature*, 585(7825):357–362, Sep 2020. ISSN 1476-4687. doi: 10.1038/s41586-020-2649-2. URL <http://dx.doi.org/10.1038/s41586-020-2649-2>.
- Joeri Hermans, Volodimir Begy, and Gilles Louppe. Likelihood-free MCMC with Amortized Approximate Ratio Estimators. 3 2019.
- Yashar D. Hezaveh, Laurence Perreault Levasseur, and Philip J. Marshall. Fast Automated Analysis of Strong Gravitational Lenses with Convolutional Neural Networks. *Nature*, 548:555–557, 2017. doi: 10.1038/nature23463.
- Matthew D. Hoffman, David M. Blei, Chong Wang, and John Paisley. Stochastic variational inference. *Journal of Machine Learning Research*, 14(4):1303–1347, 2013. URL <http://jmlr.org/papers/v14/hoffman13a.html>.
- J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007. doi: 10.1109/MCSE.2007.55.
- Michael I. Jordan, Zoubin Ghahramani, and et al. An introduction to variational methods for graphical models. In *MACHINE LEARNING*, pages 183–233. MIT Press, 1999.
- Konstantin Karchev, Adam Coogan, and Christoph Weniger. Strong-lensing source reconstruction with differentiable probabilistic programming. 2020.
- Thomas Kluyver, Benjamin Ragan-Kelley, Fernando Pérez, Brian Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica Hamrick, Jason Grout, Sylvain Corlay, Paul Ivanov, Damián Avila, Safia Abdalla, and Carol Willing. Jupyter notebooks – a publishing format for reproducible computational workflows. In F. Loizides and B. Schmidt, editors, *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, pages 87 – 90. IOS Press, 2016.
- Warren R. Morningstar, Laurence Perreault Levasseur, Yashar D. Hezaveh, Roger Blandford, Phil Marshall, Patrick Putzky, Thomas D. Rueter, Risa Wechsler, and Max Welling. Data-Driven Reconstruction of Gravitationally Lensed Galaxies using Recurrent Inference Machines. 1 2019. doi: 10.3847/1538-4357/ab35d7.
- European Southern Observatory. Eso - e-elt instrumentation. <https://www.eso.org/sci/facilities/eelt/instrumentation/phaseA.html>. Accessed 2020-09-30.
- Bryan Ostdiek, Ana Diaz Rivero, and Cora Dvorkin. Detecting Subhalos in Strong Gravitational Lens Images with Image Segmentation. 9 2020a.
- Bryan Ostdiek, Ana Diaz Rivero, and Cora Dvorkin. Extracting the Subhalo Mass Function from Strong Lens Images with Image Segmentation. 9 2020b.
- Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019. URL <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.
- Lawrence K. Saul, Tommi S. Jaakkola, and Michael I. Jordan. Mean field theory for sigmoid belief networks. *CoRR*, cs.AI/9603102, 1996. URL <https://arxiv.org/abs/cs/9603102>.
- Aurel Schneider, Robert E. Smith, Andrea V. Maccio, and Ben Moore. Nonlinear Evolution of Cosmological Structures in Warm Dark Matter Models. *Mon. Not. Roy. Astron. Soc.*, 424:684, 2012. doi: 10.1111/j.1365-2966.2012.21252.x.
- 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.*, 691:277–298, 2009. doi: 10.1088/0004-637X/691/1/277.

- 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.*, 392:945, 2009. doi: 10.1111/j.1365-2966.2008.14005.x.
- 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.*, 408:1969, 2010. doi: 10.1111/j.1365-2966.2010.16865.x.
- 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*, 481:341, 2012. doi: 10.1038/nature10669.
- Matteo Viel, George D. Becker, James S. Bolton, and Martin G. Haehnelt. Warm dark matter as a solution to the small scale crisis: New constraints from high redshift Lyman- α forest data. *Phys. Rev. D*, 88:043502, 2013. doi: 10.1103/PhysRevD.88.043502.
- Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

Appendix

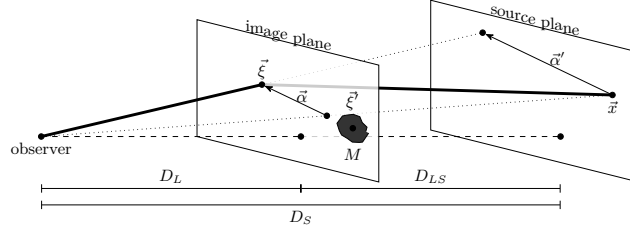


Figure 3: Geometry of strong lensing systems. The lensing mass M located at $\vec{\xi}'$ bends the light ray (thick line) emanating from the point \vec{x} , so that to the observer it looks like it is coming from the direction of $\vec{\xi}$. General relativity predicts the deflection angle $\vec{\alpha}'$ as viewed from the image plane based on the mass M and the distance $\vec{\xi} - \vec{\xi}'$. It then has to be rescaled by D_{LS}/D_S to obtain the displacement field $\vec{\alpha}$ (as viewed by the observer). 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.

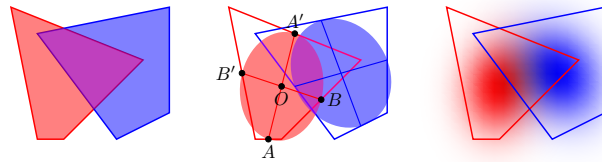


Figure 4: Our Gaussian approximation of pixels projected into the source plane. The projection of two pixels onto the source plane (a). A full treatment should consider their exact shapes and use an indicator function that is nonzero only inside the shaded areas. Instead, the projections are approximated to ellipses defined by the projections of their centers and the vectors AA' and BB' (b). The indicator functions are Gaussians with covariances derived from those ellipses (c).