Inferring substructure in strong lenses using simulation-based inference

Siddharth Mishra-Sharma

Based on work in progress with Johann Brehmer, Joeri Hermans, Gilles Louppe and Kyle Cranmer

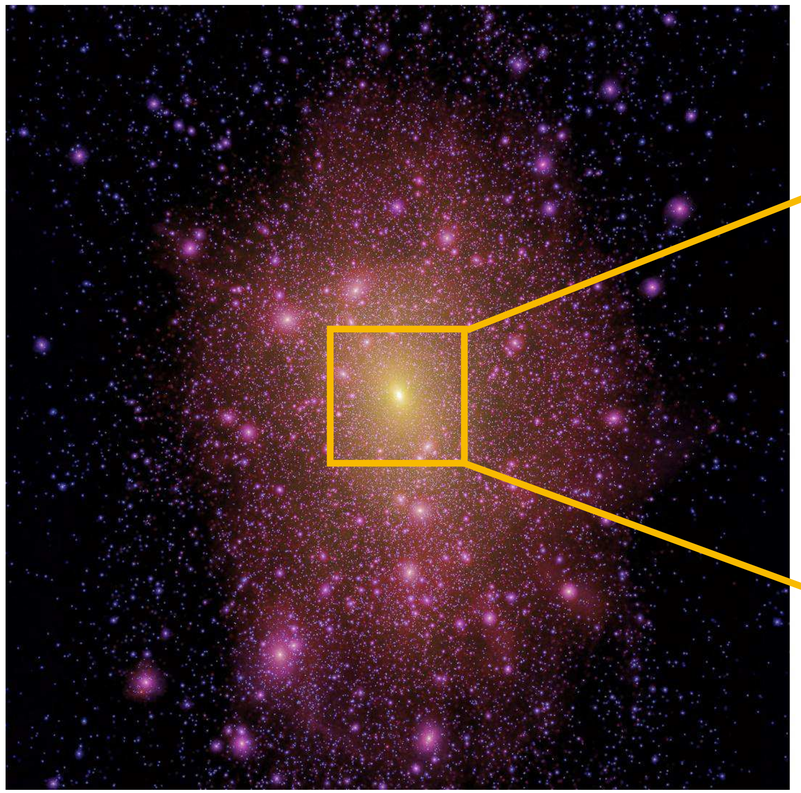
NYU | Center for Cosmology and Particle Physics

LSST Dark Matter Workshop, August 6 2019

KICP, Chicago, IL

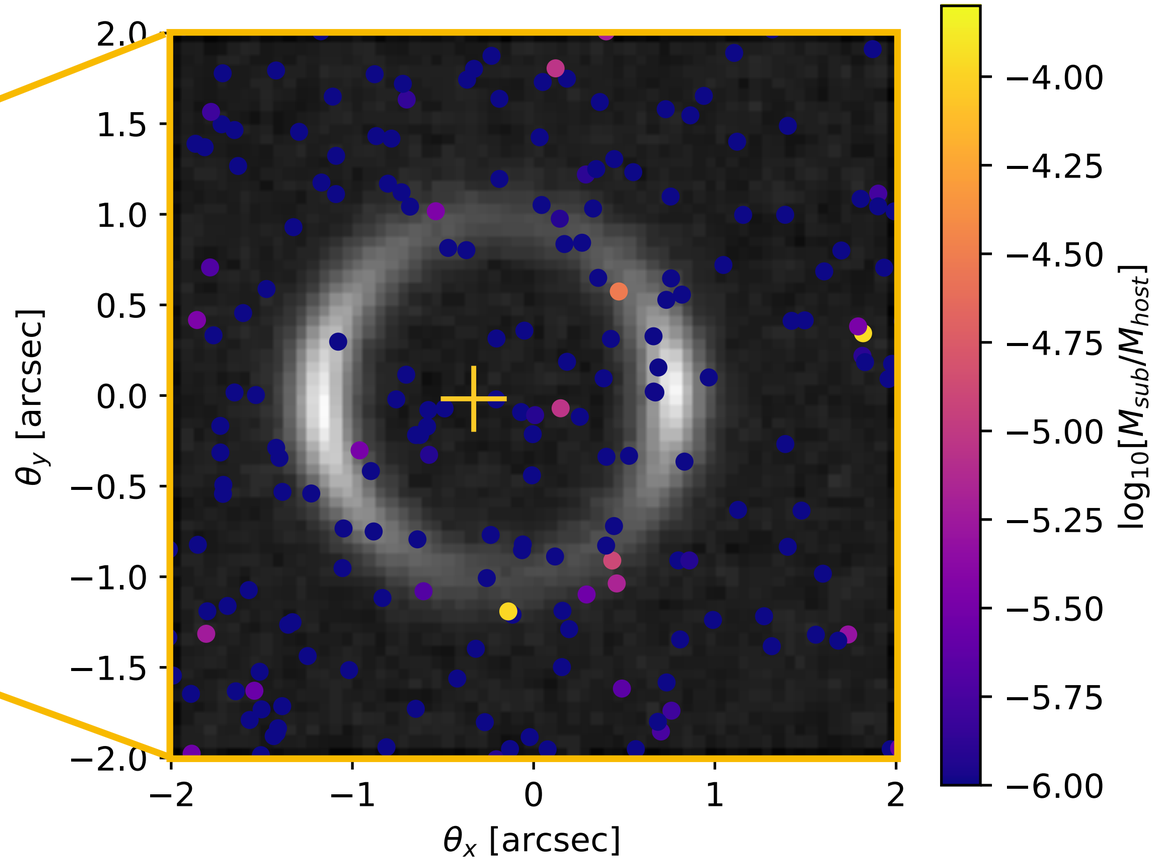
Gravitational lensing: effect of substructure

Presence of substructure down to small scales is one of the key predictions of CDM

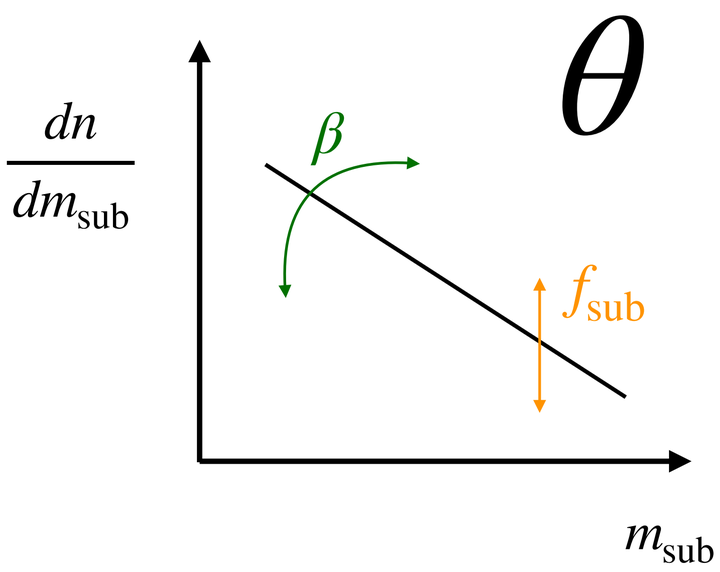


Lovell et al [1104.2929]

Substructure causes **percent-level shifts** in strongly lensed image



Substructure likelihood



Parameters of interest

Subhalo population parameters

$$\theta = \{f_{\rm sub}, \beta\}$$

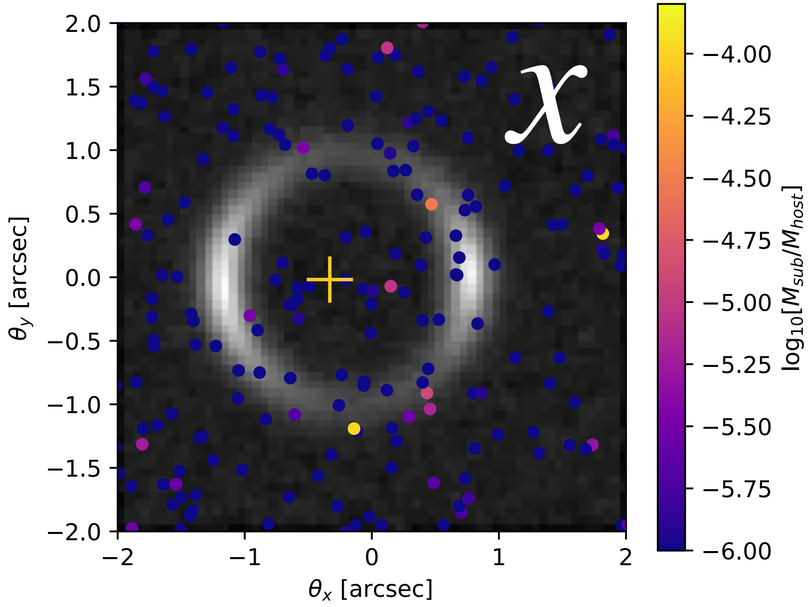
Prediction →

← Inference

$$p(x\,|\,\theta) = \int \mathrm{d}z\, p(x, z\,|\,\theta)$$

Latent variables

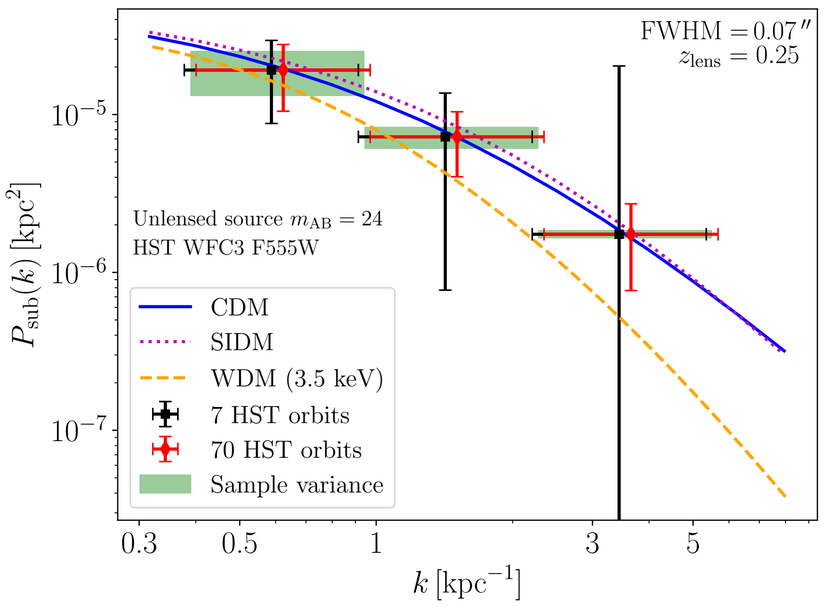
$$z = \{\vec{m}_{\rm sub}, \vec{r}_{\rm sub}\}$$



Huge latent space — full likelihood is intractable!

Inferring substructure through collective effects

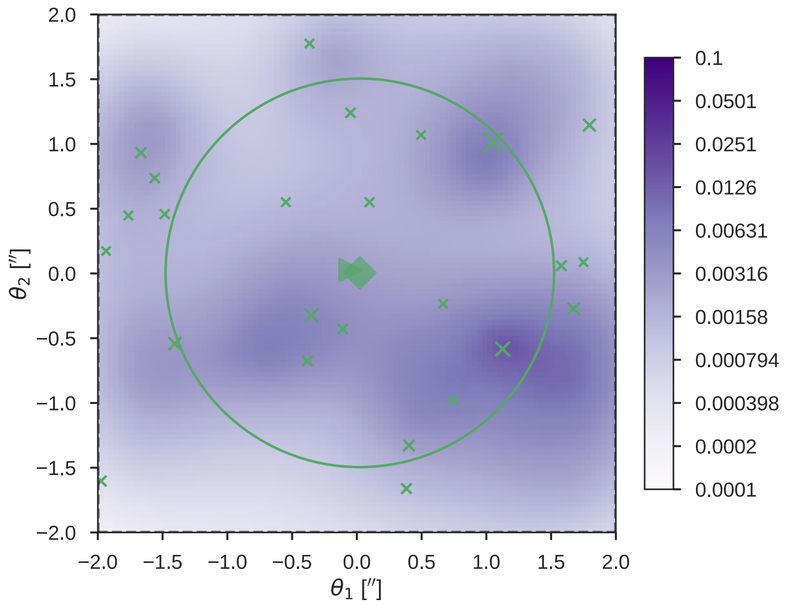
Power spectrum decomposition



Cyr-Racine et al [1806.07897]

See also
Rivero et al [1707.04590]
Rivero et al [1809.00004]
Brennan et al [1808.03501]
Hezaveh et al [1403.2720]

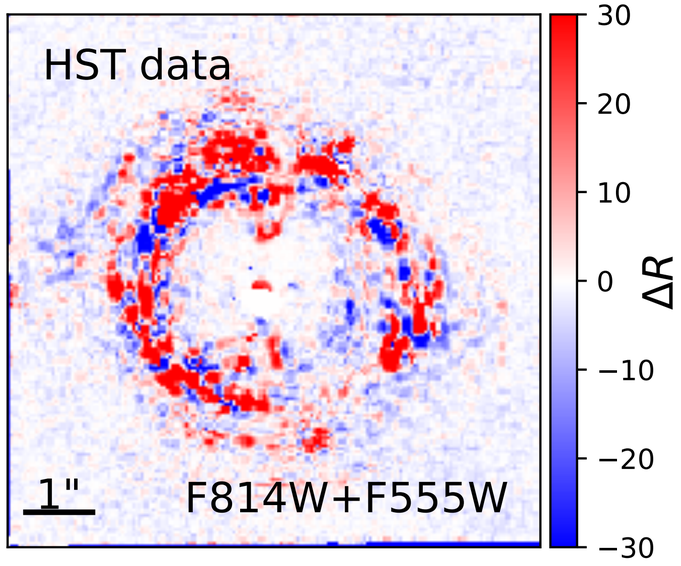
Trans-dimensional methods



Daylan et al [1706.06111]

See also
Brewer et al [1508.00662]

Summary statistics



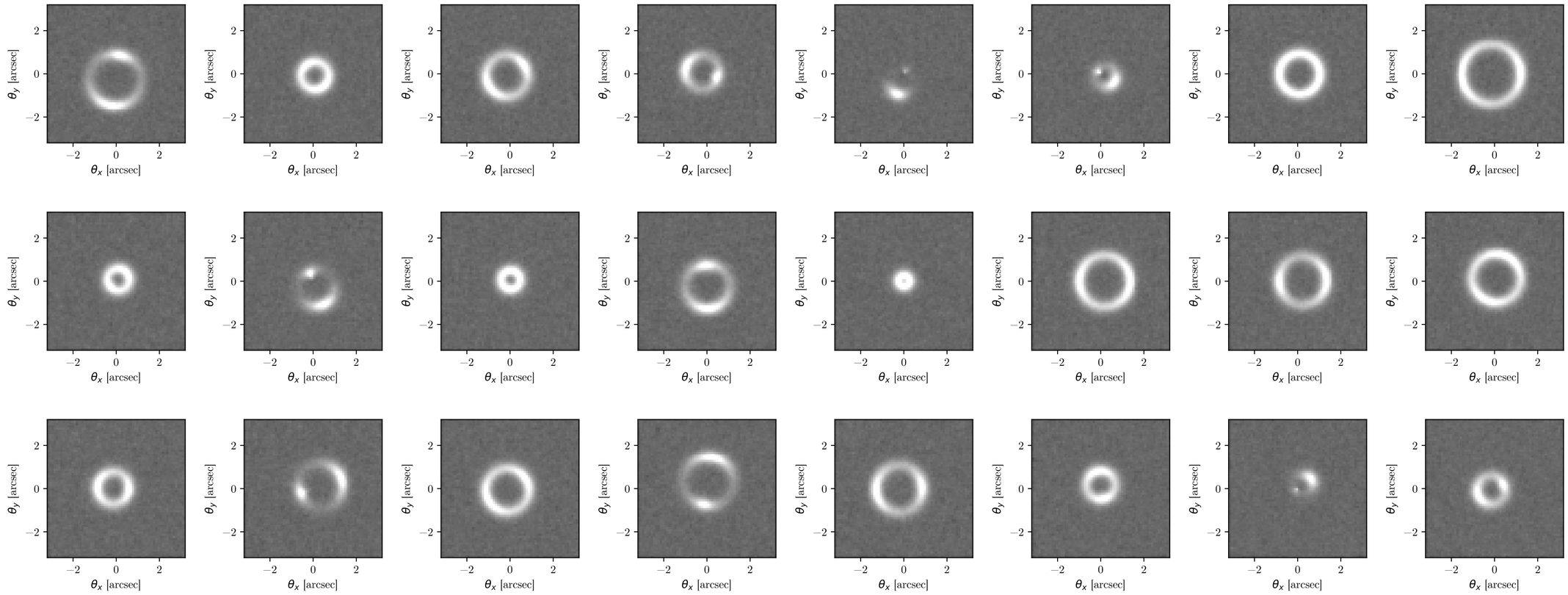
Birrer et al [1702.00009]

Goal

Future surveys like LSST, *Euclid* expected to deliver large samples of galaxy-galaxy strong lenses

$\mathcal{O}(10{,}000)$

Collett et al [1507.02657]

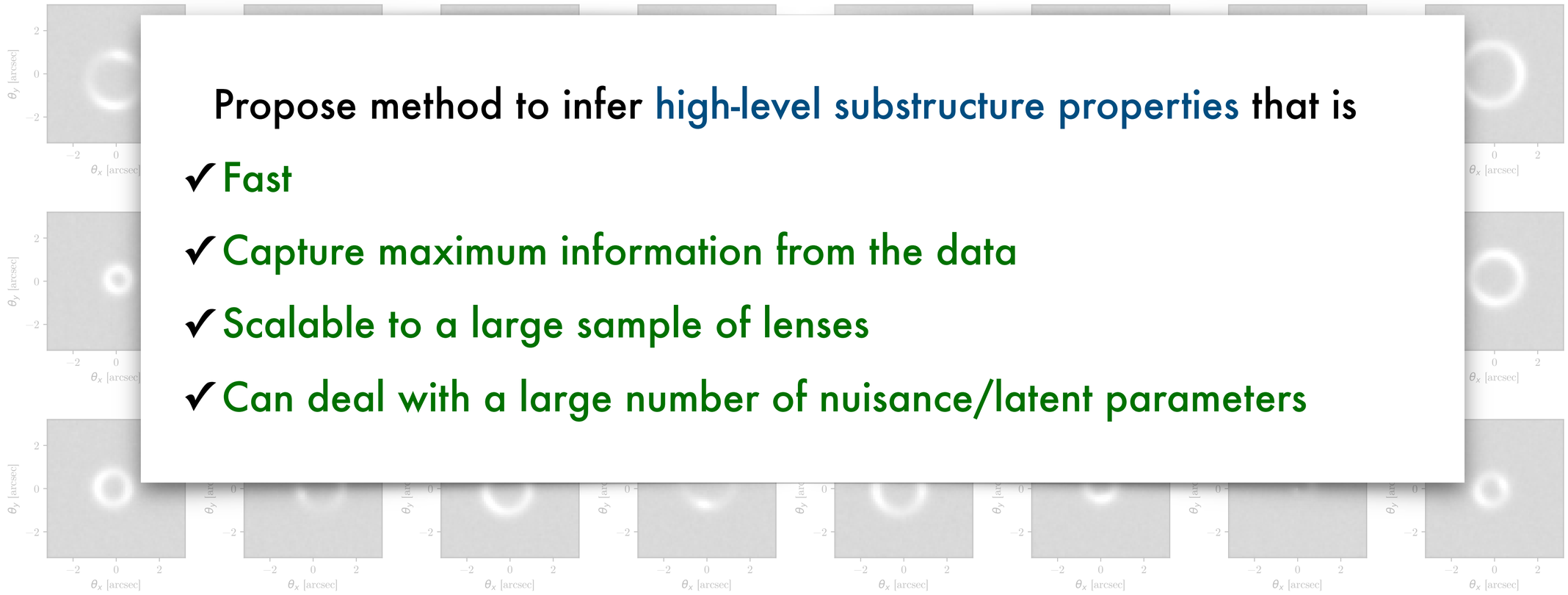


Goal

Future surveys like LSST, *Euclid* expected to deliver large samples of galaxy-galaxy strong lenses

$\mathcal{O}(10{,}000)$

Collett et al [1507.02657]

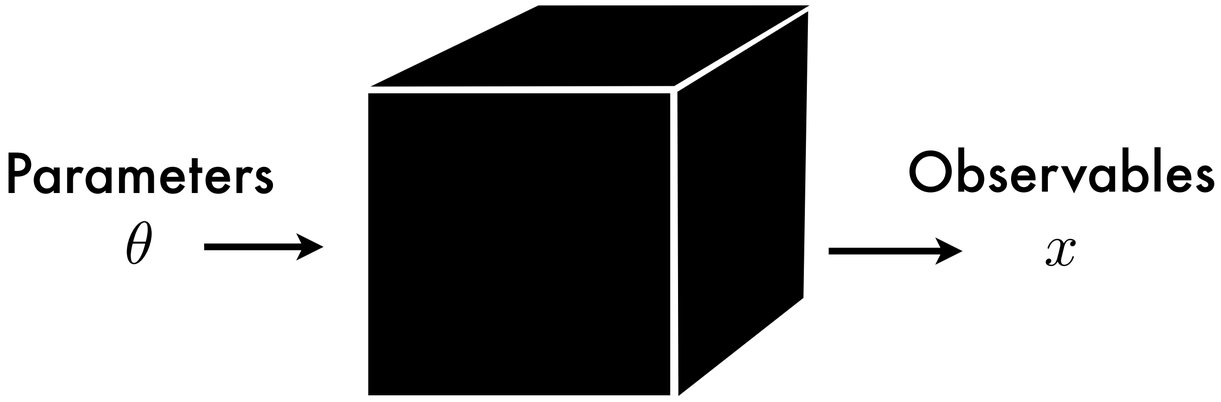


Propose method to infer high-level substructure properties that is

- ✓ Fast
- ✓ Capture maximum information from the data
- ✓ Scalable to a large sample of lenses
- ✓ Can deal with a large number of nuisance/latent parameters

Likelihood-free inference: opening the black box

Slides courtesy of Johann Brehmer

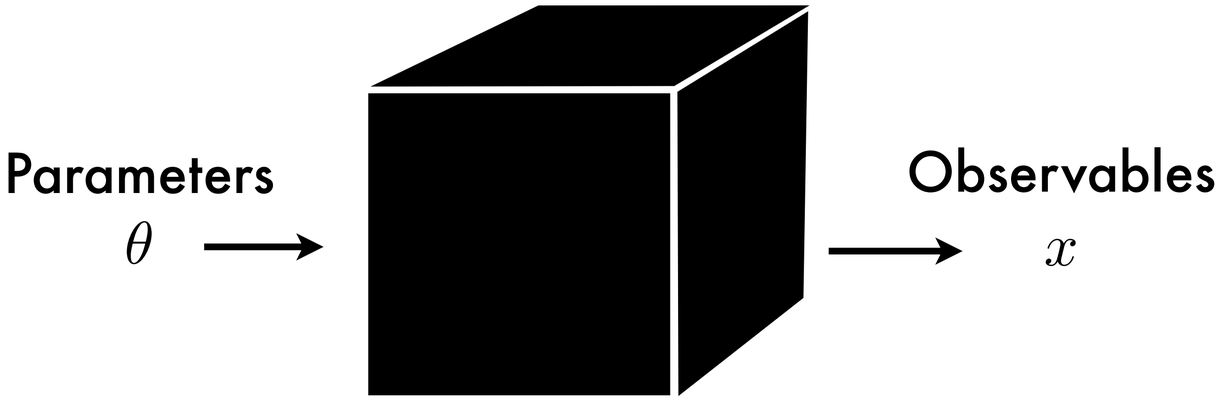


"Traditional" LFI treats the simulator as a generative black box: parameters in, samples out.

No one cares how the sausage is made.

Likelihood-free inference: opening the black box

Slides courtesy of Johann Brehmer



"Traditional" LFI treats the simulator as a generative black box: parameters in, samples out.

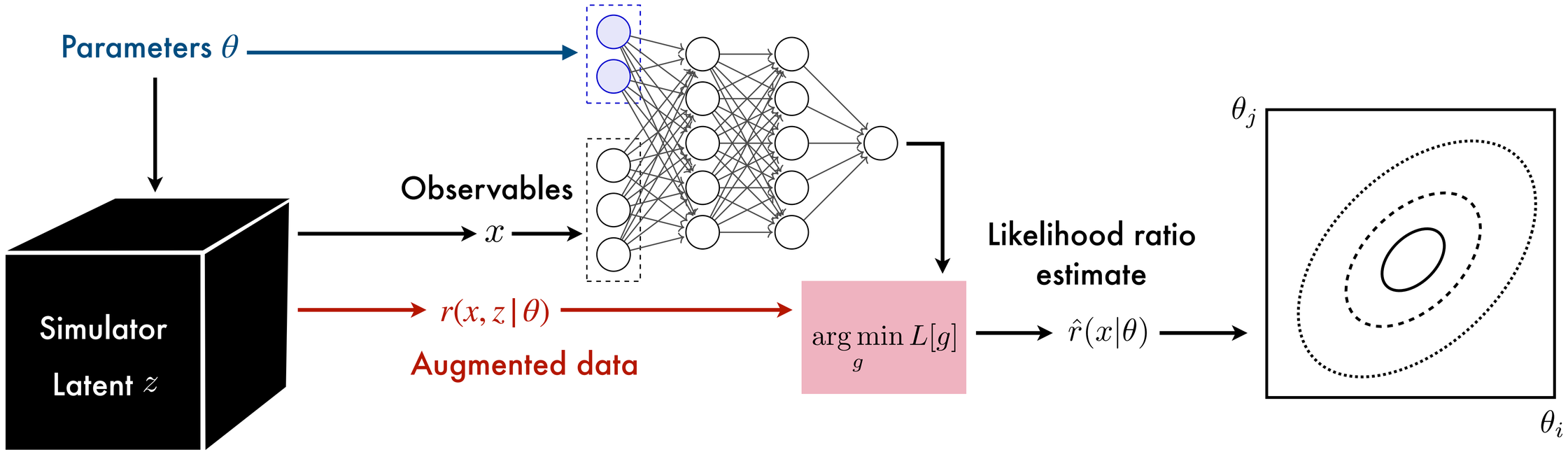
No one cares how the sausage is made.

But in real-life problems, we have access to the simulator code and some understanding of the microscopic processes.

We can extract more simulation from the simulator and use it to improve inference.

Overview

Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]



1. Simulation

Extract additional information from simulator

2. Machine Learning

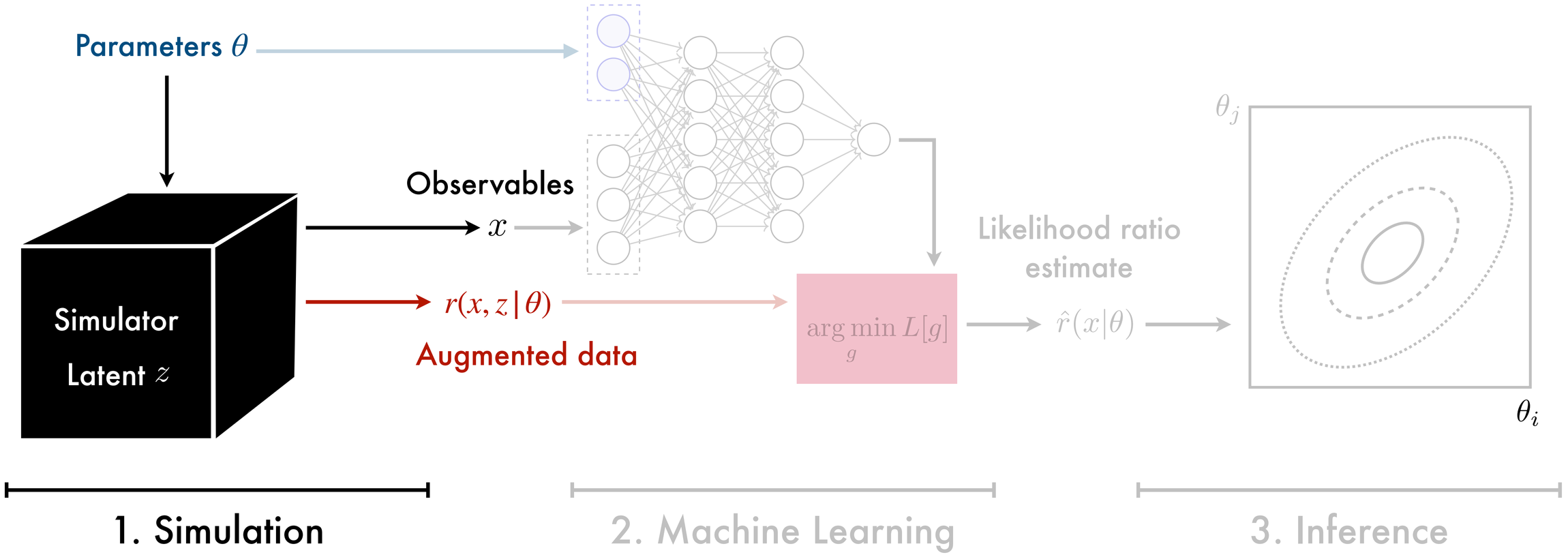
Use this information to train estimator for likelihood ratio

3. Inference

Limit setting with standard hypothesis tests

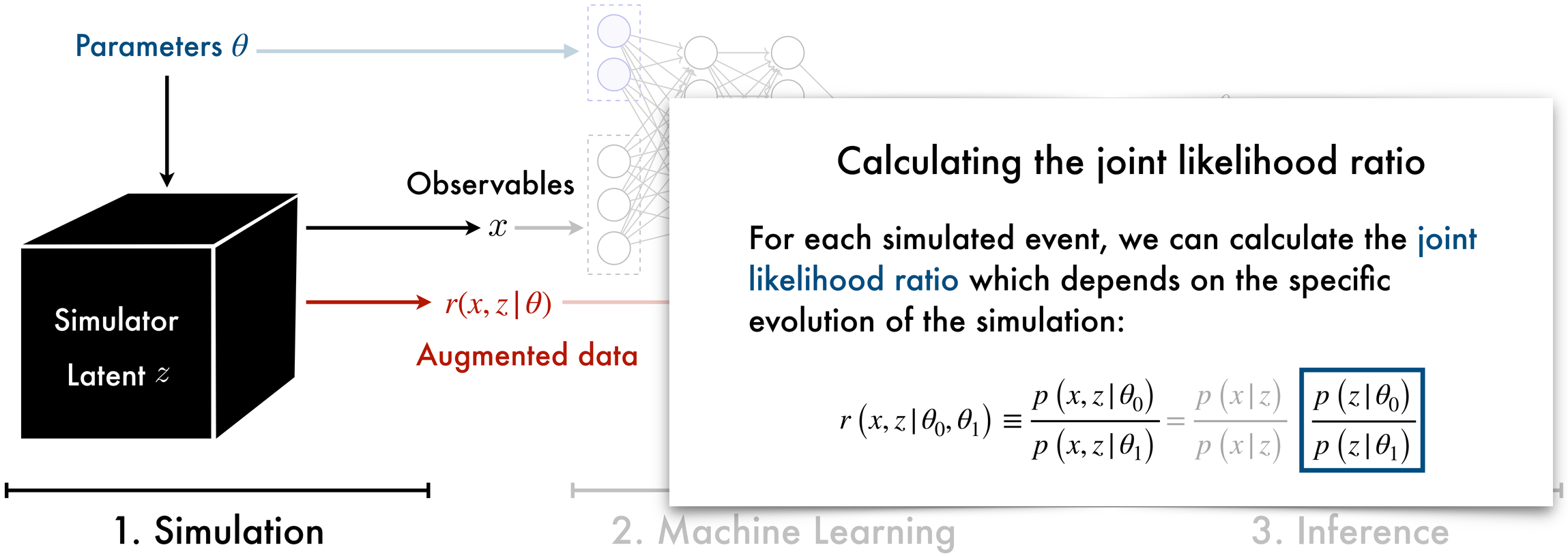
Simulation

Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]



Simulation

Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]



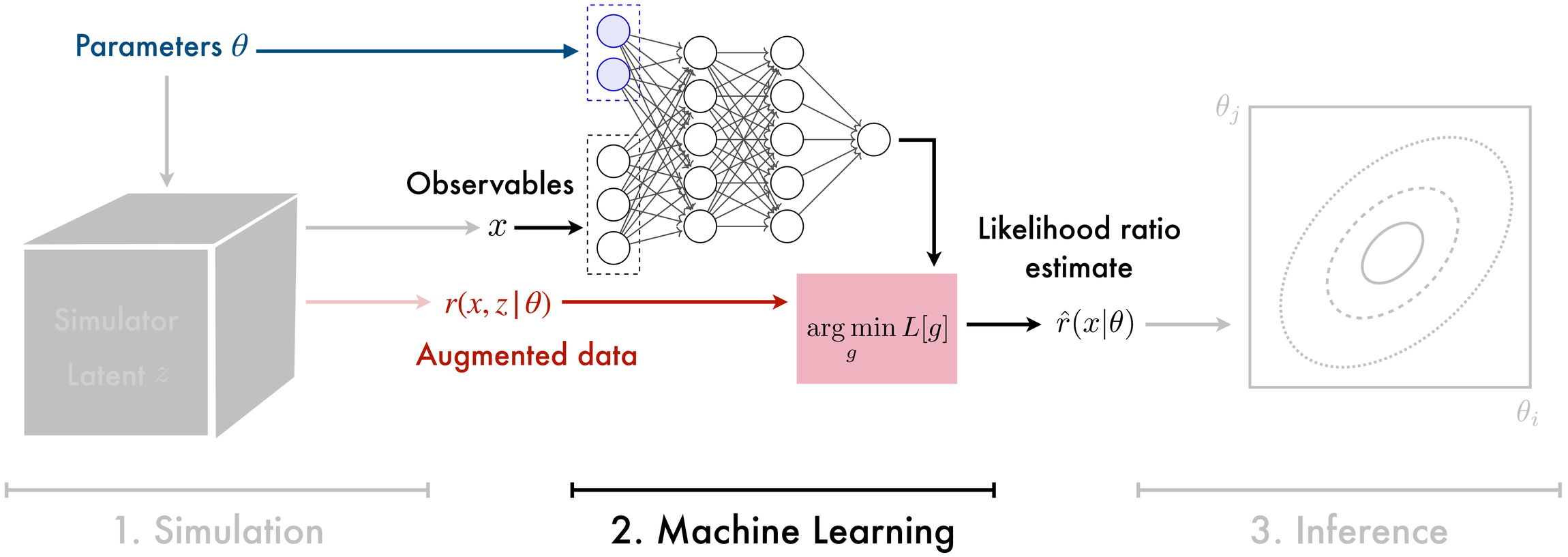
Calculating the joint likelihood ratio

For each simulated event, we can calculate the joint likelihood ratio which depends on the specific evolution of the simulation:

$$r\left(x, z|\theta_0, \theta_1\right) \equiv \frac{p\left(x, z|\theta_0\right)}{p\left(x, z|\theta_1\right)} = \frac{p\left(x|z\right)}{p\left(x|z\right)} \boxed{\frac{p\left(z|\theta_0\right)}{p\left(z|\theta_1\right)}}$$

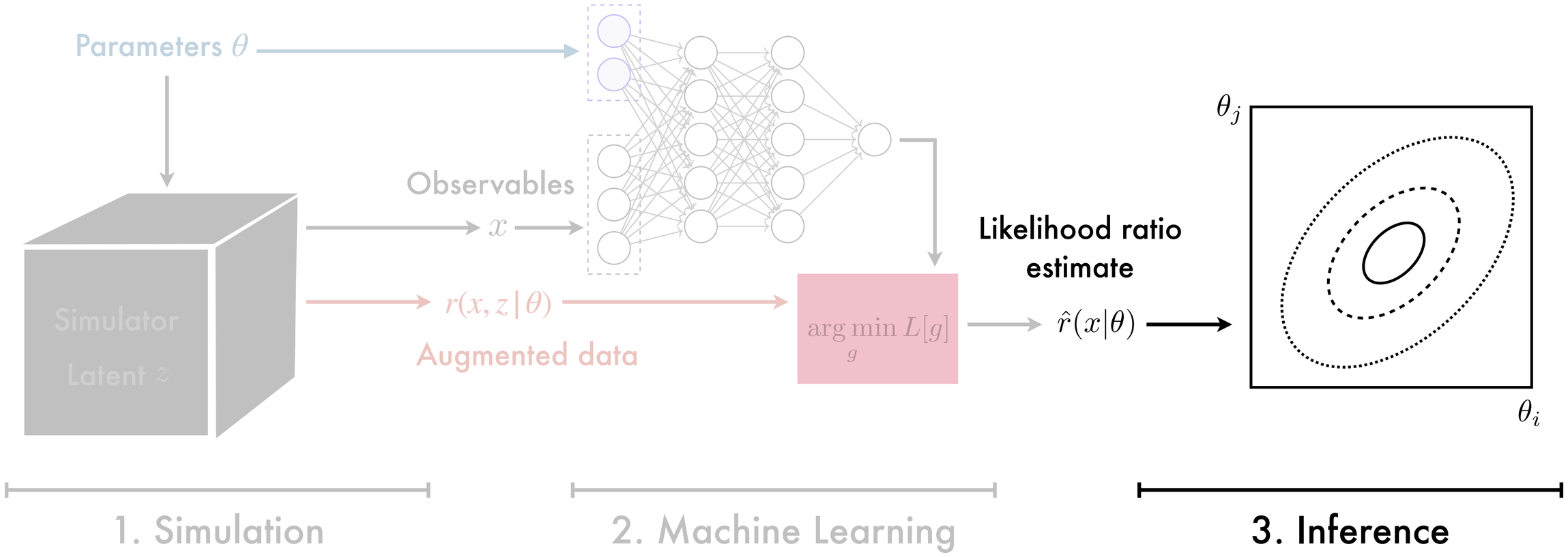
Machine learning

Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]



Inference

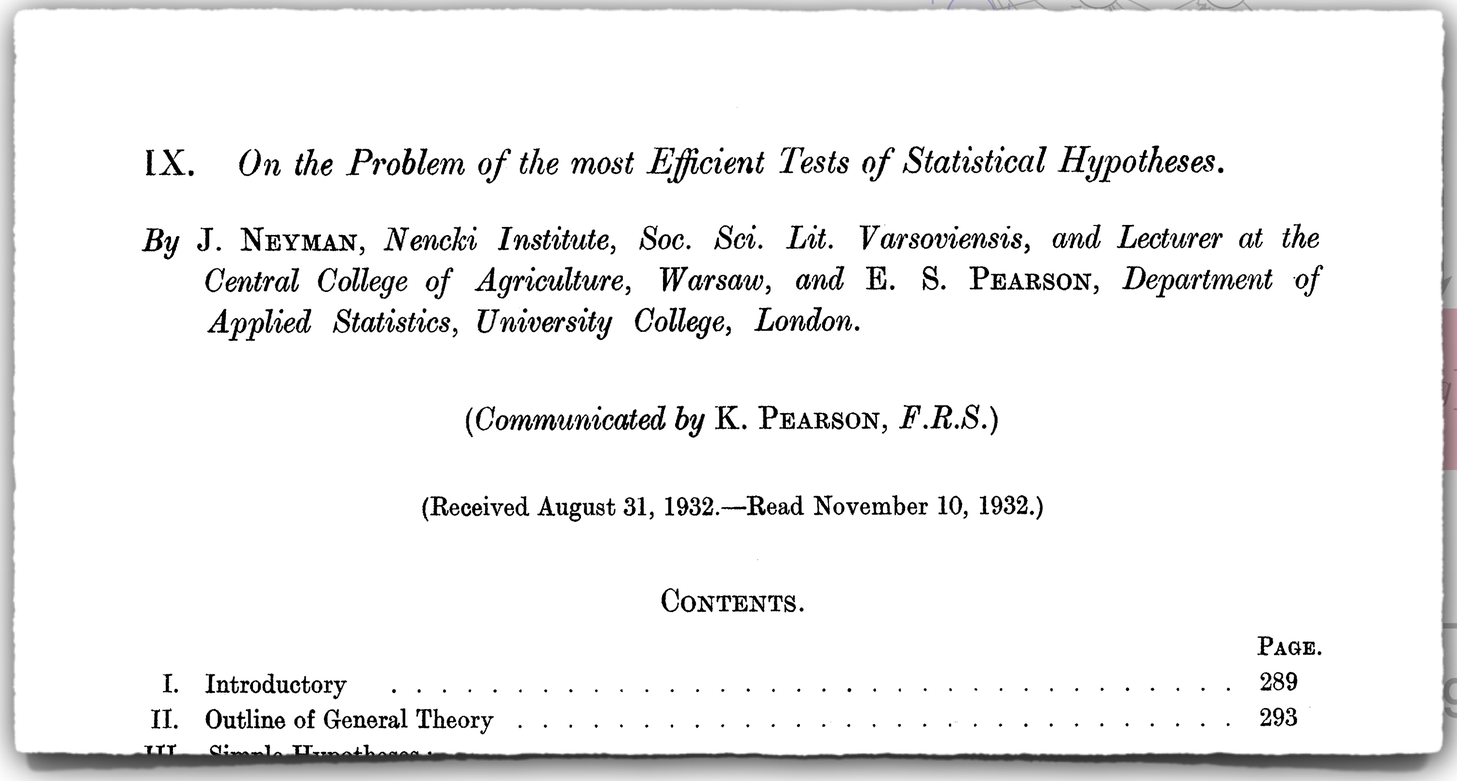
Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]



Inference

Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]

Parameters θ



IX. *On the Problem of the most Efficient Tests of Statistical Hypotheses.*

By J. Neyman, *Nencki Institute, Soc. Sci. Lit. Varsoviensis, and Lecturer at the Central College of Agriculture, Warsaw, and* E. S. Pearson, *Department of Applied Statistics, University College, London.*

(*Communicated by* K. Pearson, *F.R.S.*)

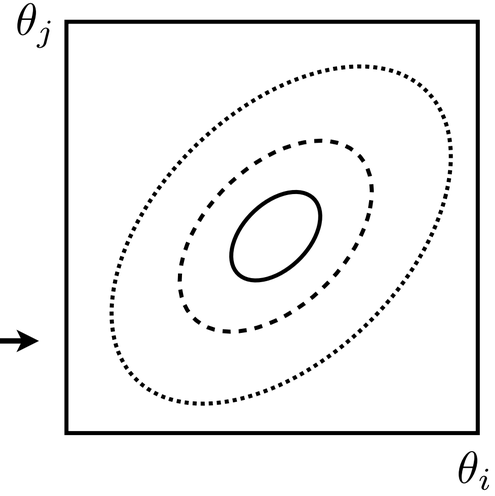
(Received August 31, 1932.—Read November 10, 1932.)

Contents.

		Page.
I.	Introductory	289
II.	Outline of General Theory	293

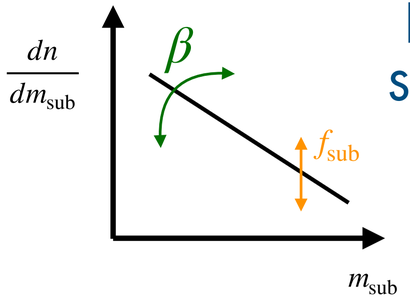
Likelihood ratio estimate

$\hat{r}(x|\theta)$



3. Inference

Application to substructure in strong lenses



Parameters of interest

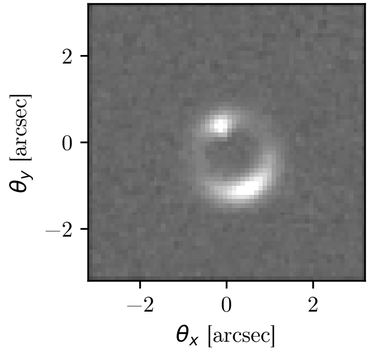
Subhalo population parameters

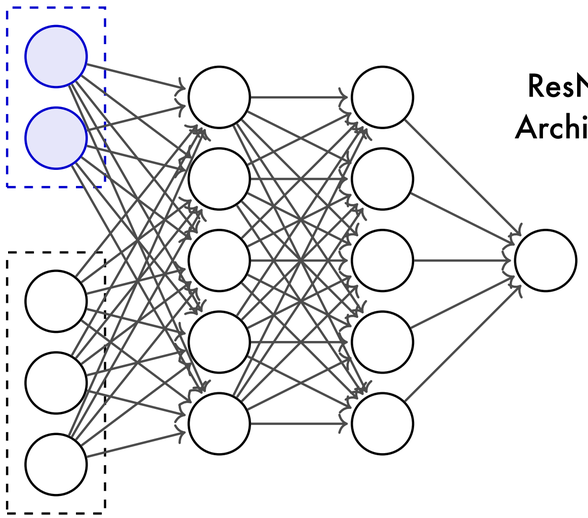
$\theta = \{f_{\rm sub}, \beta\}$

Observables

Lensed image

$x =$





ResNet-18 Architecture

Simulator

Latent

$z = \{\vec{m}_{\rm sub}, \vec{r}_{\rm sub}\}$

Nuisance

Host/source params, redshifts…

Joint likelihood ratio

$$r\left(x, z \,|\, \theta_0, \theta_1\right) = \frac{p(\vec{m}_{\rm sub}, \vec{r}_{\rm sub} \,|\, f_{\rm sub,0}, \beta_0)}{p(\vec{m}_{\rm sub}, \vec{r}_{\rm sub} \,|\, f_{\rm sub,1}, \beta_1)}$$

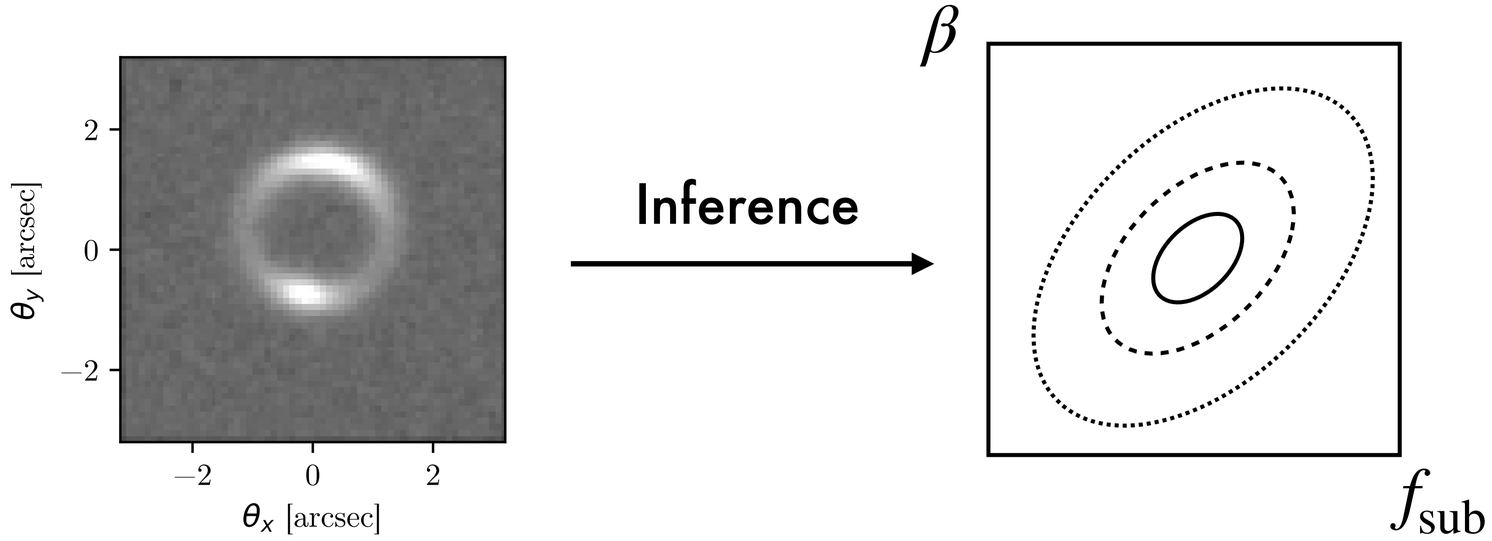
$$\arg\min_g L[g]$$

Likelihood ratio estimate

$\hat{r}(x|\theta)$

Application to substructure in strong lenses

After training the parameterized LR estimator…

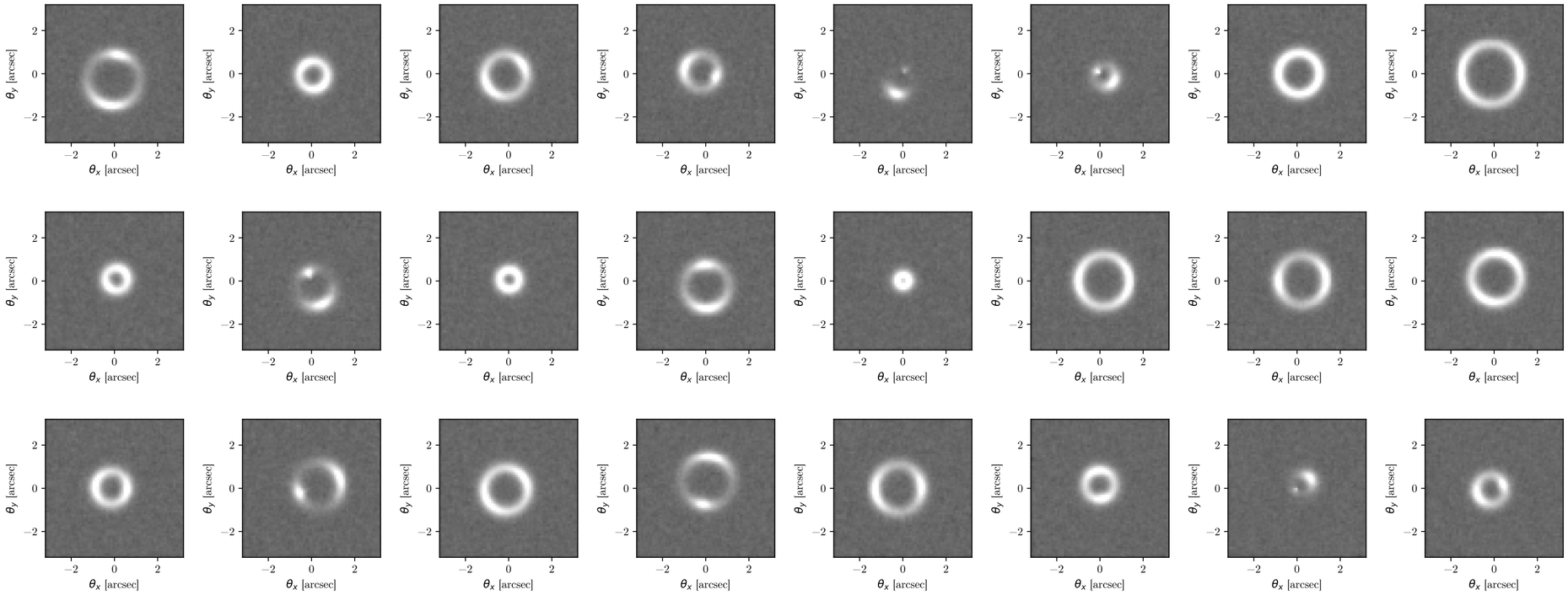


… can infer parameters of interest for given lensed image!

Proof of principle

Use simulated ensemble of galaxy-galaxy lenses observable by *Euclid*

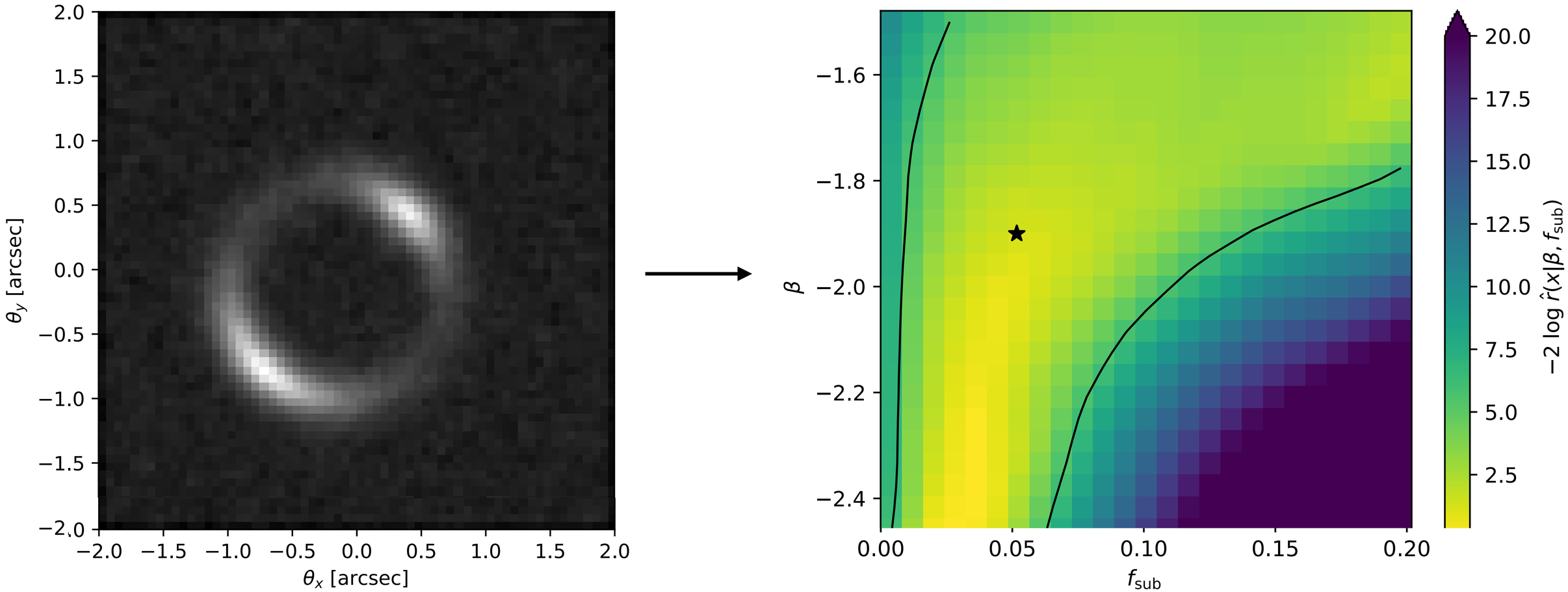
Collett et al [1507.02657]



1. Train likelihood ratio estimator with $f_{\rm sub} \sim [0,\, 0.2]$, $\beta \sim [-2.5, -1.5]$
2. Test on simulated data with $f_{\rm sub} = 0.05$, $\beta = -1.9$

Inferred likelihood ratios

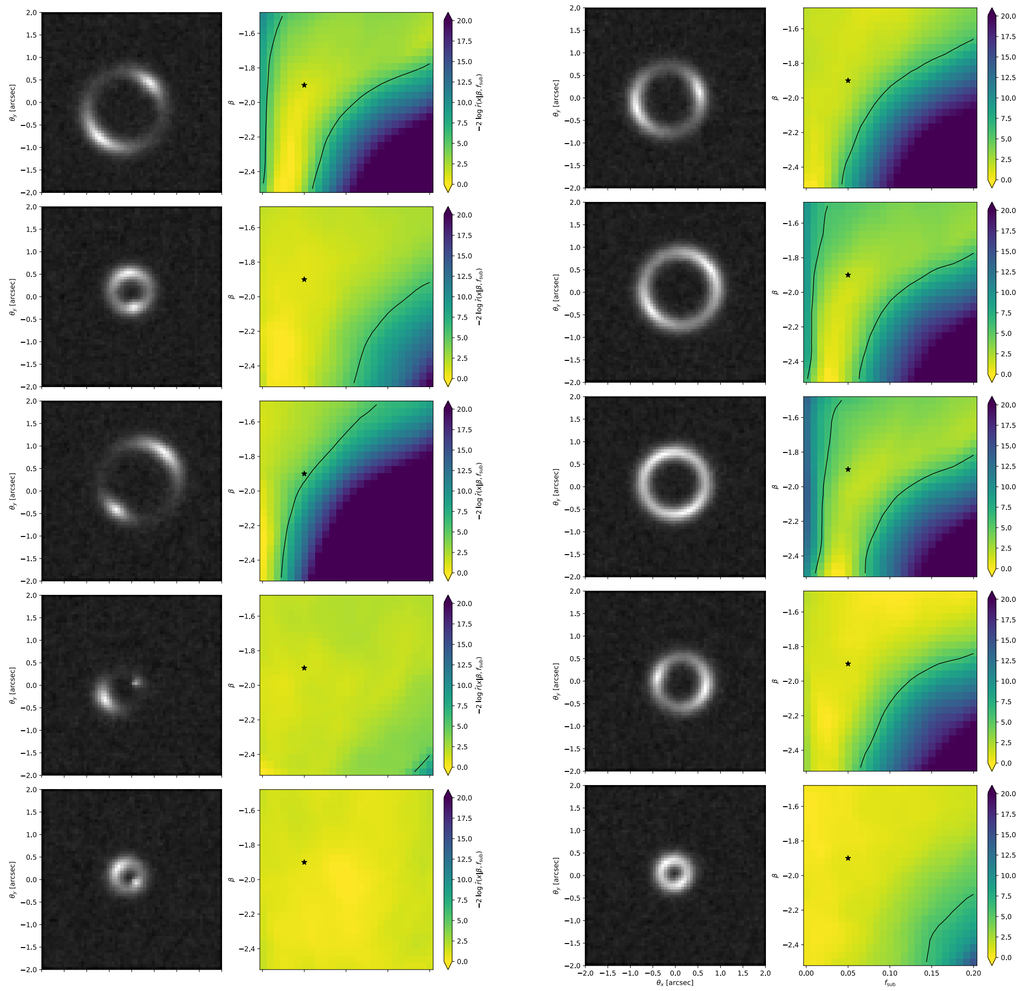
$f_{\text{sub}} = 0.05,\ \beta = -1.9$



Inferred likelihood ratios

$f_{\rm sub} = 0.05,\ \beta = -1.9$

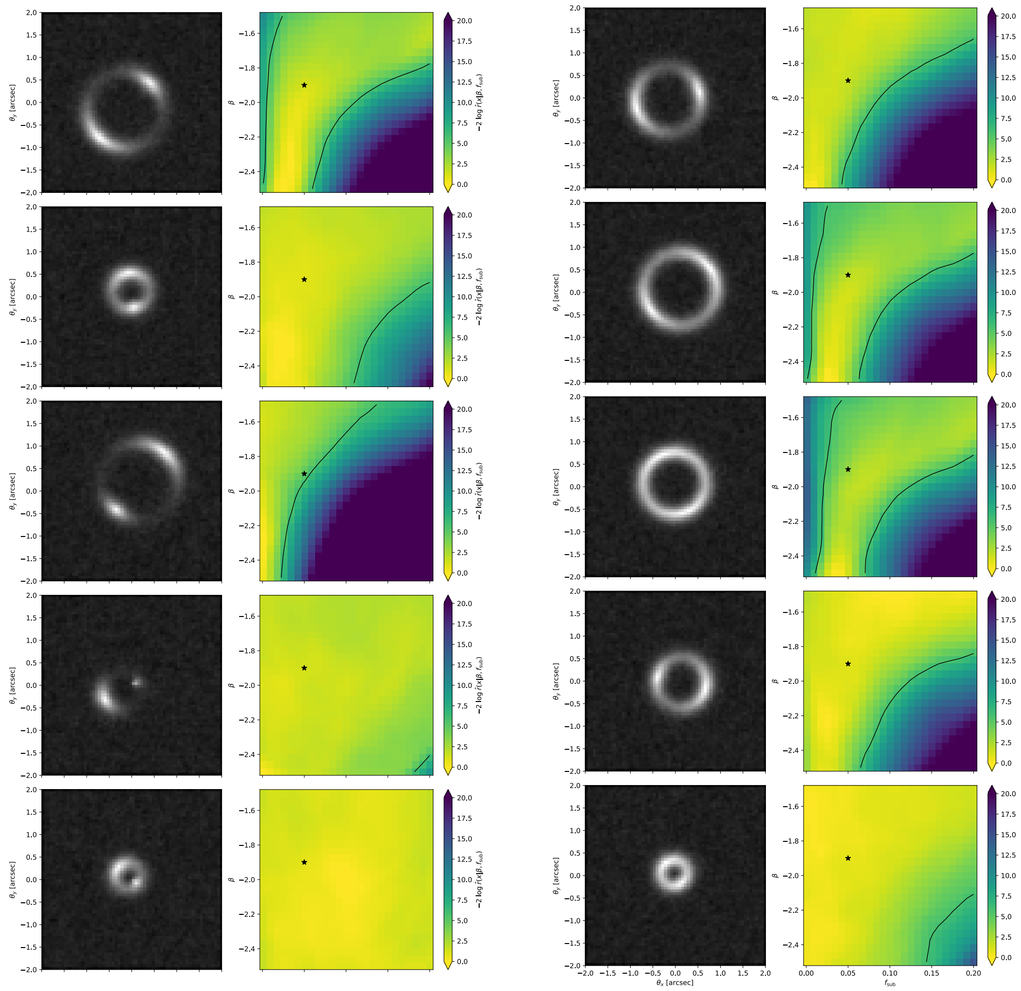
Analysis of Individual images



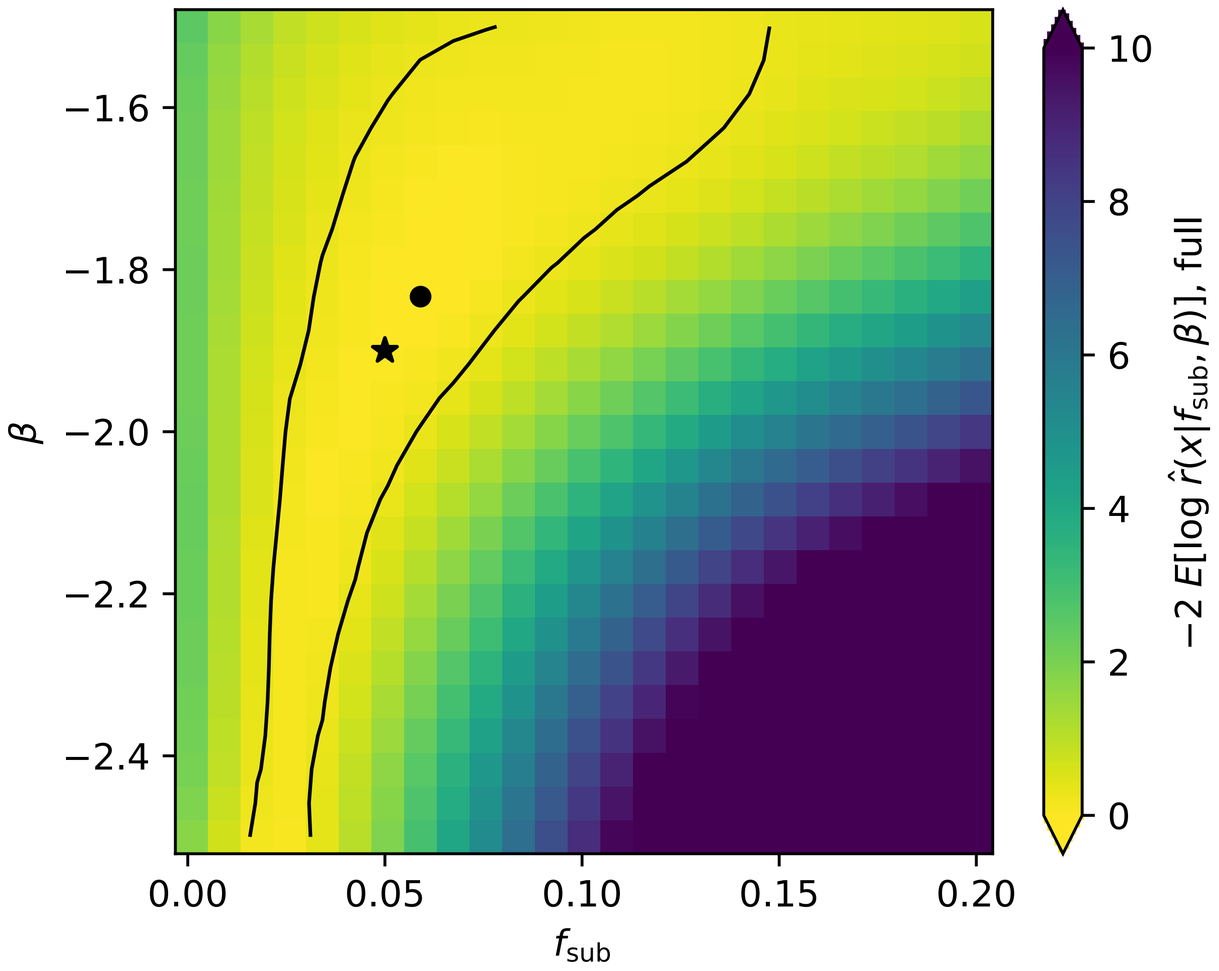
Inferred likelihood ratios

$f_{\rm sub} = 0.05,\ \beta = -1.9$

Analysis of Individual images

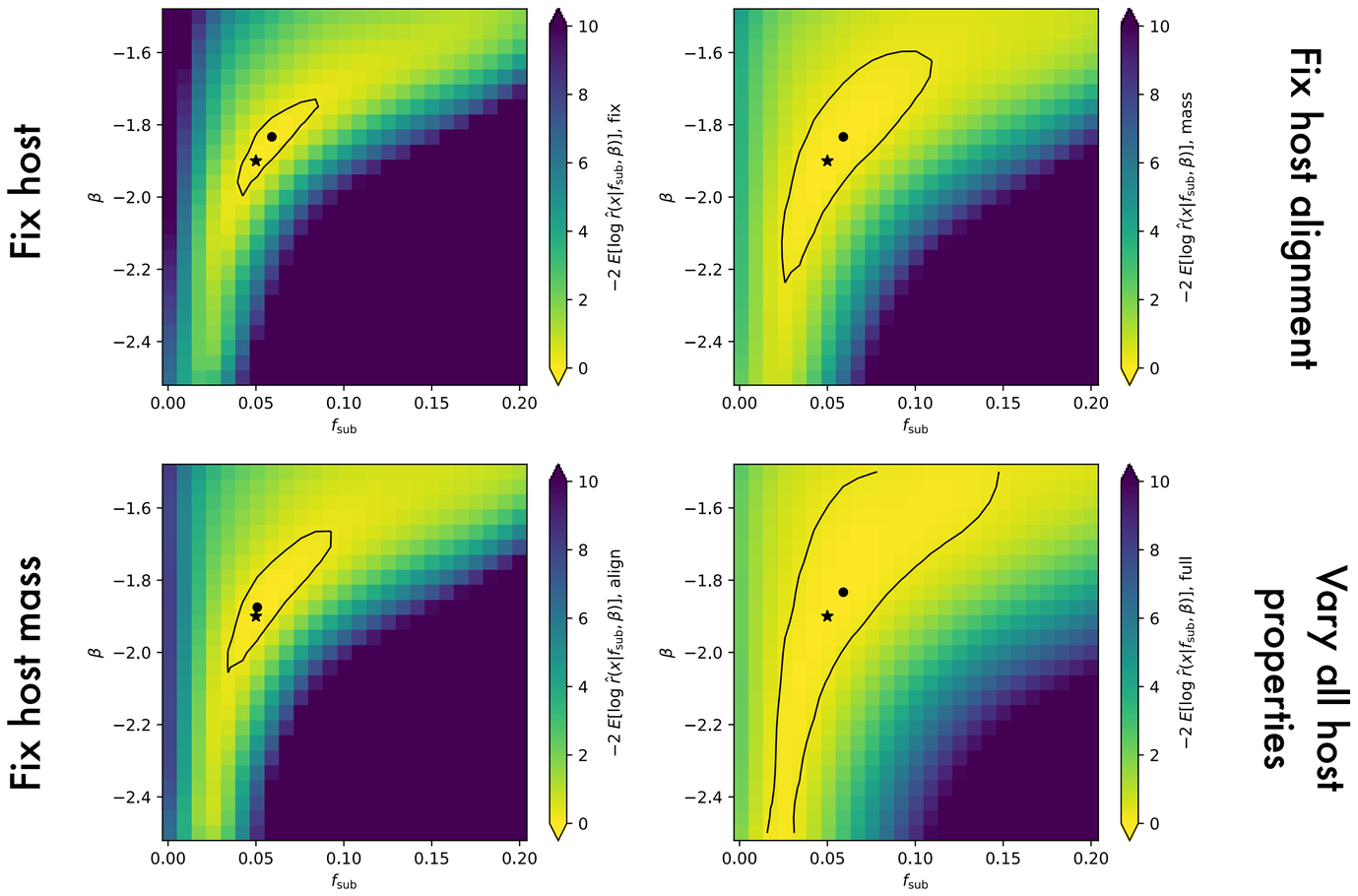


Combined analysis (20 images)

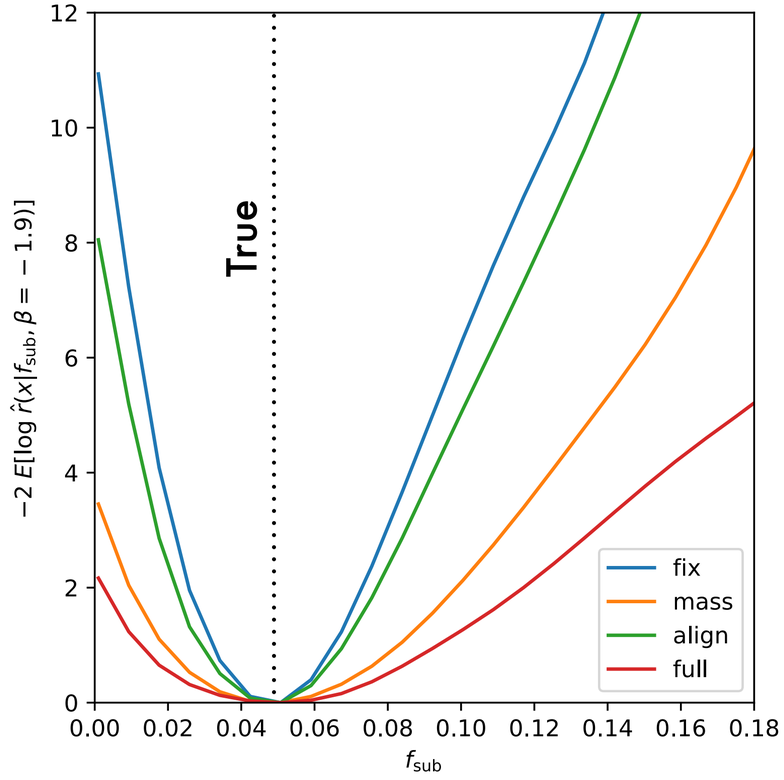


Inferred likelihood ratios

$f_{\rm sub} = 0.05,\ \beta = -1.9$



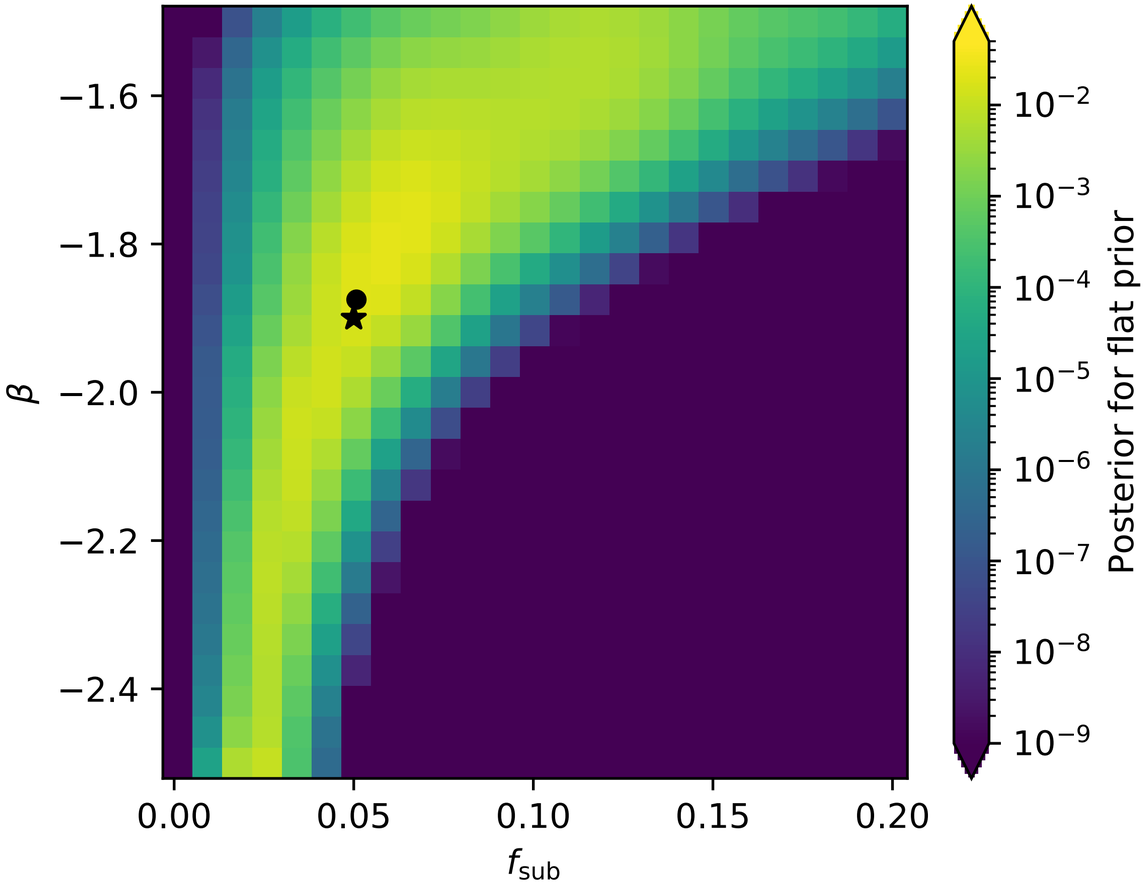
Profiles at $\beta = -1.9$



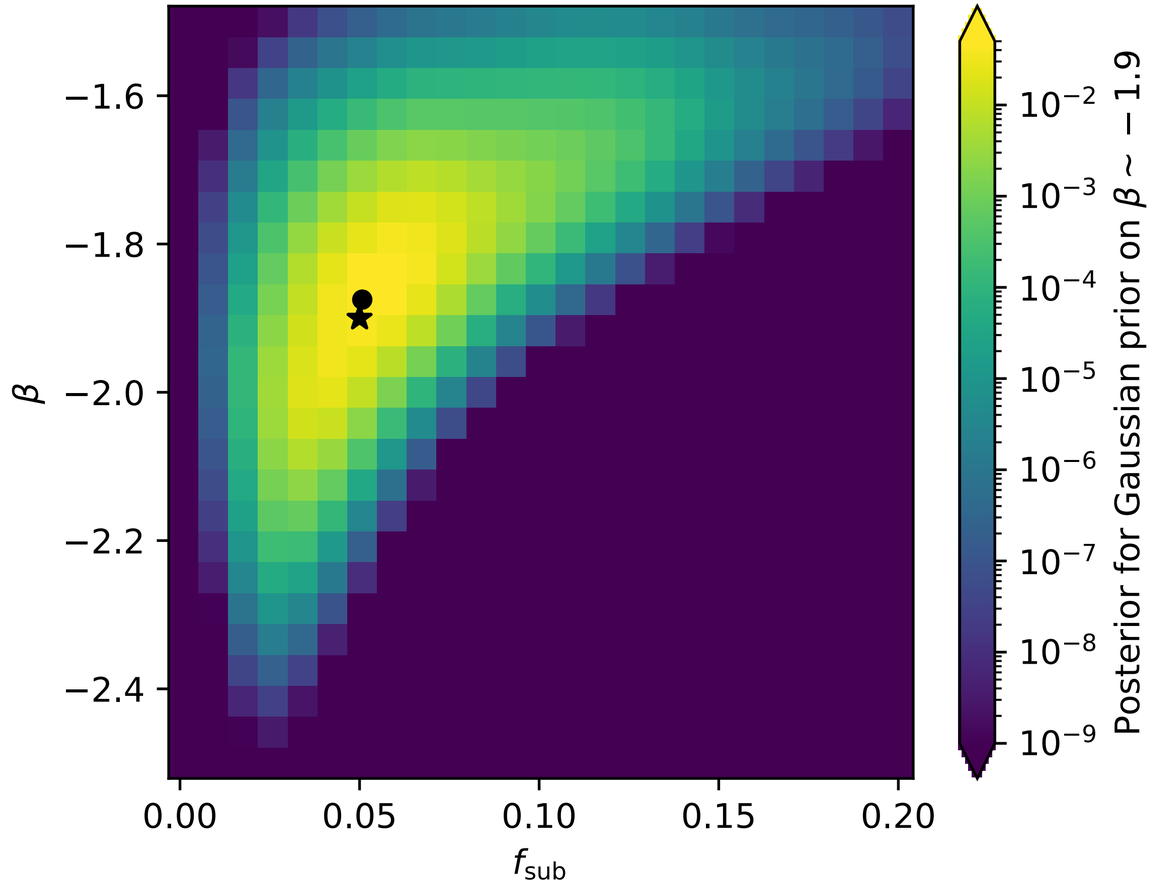
Bayesian interpretation

$f_{\rm sub} = 0.05,\ \beta = -1.9$

Uniform prior on $\{f_{\rm sub}, \beta\}$



Gaussian prior for $\beta \sim \mathcal{N}(-1.9,\ 0.1)$

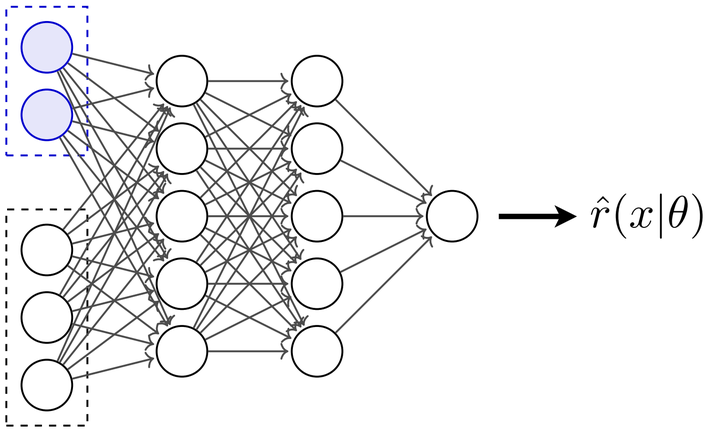


Summary

Estimating the likelihood ratio with machine learning

Powerful simulation-based estimators of the likelihood ratio provide a principled way to perform inference using additional information extracted from the simulator

Brehmer et al [1805.00013], Brehmer et al [1805.00020]



Inferring substructure in strong lenses

Fast, efficient, scalable way to analyze a population of galaxy-galaxy strong lenses to infer underlying substructure properties