UNIVERSITY OF AMSTERDAM

GRAPPA

GRavitation AstroParticle Physics Amsterdam

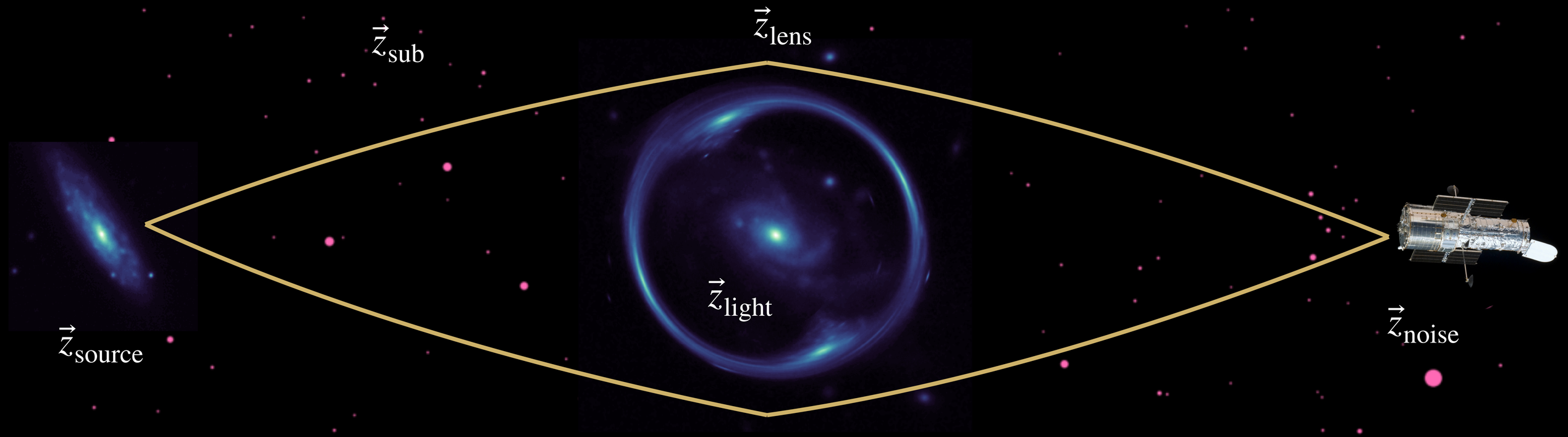
Sequential simulation-based inference for strong gravitational lensing

The importance of combining different inference strategies

Noemi Anau Montel
University of Amsterdam, GRAPPA Institute
n.anaumontel@uva.nl

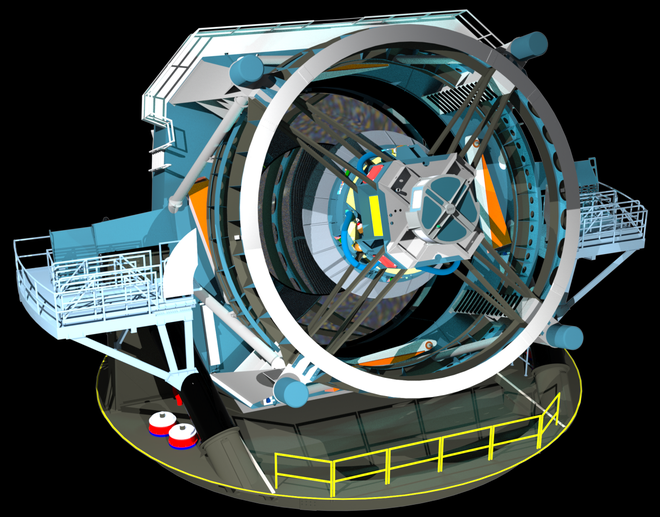
Strong gravitational lensing

Diverse and *complex* observations

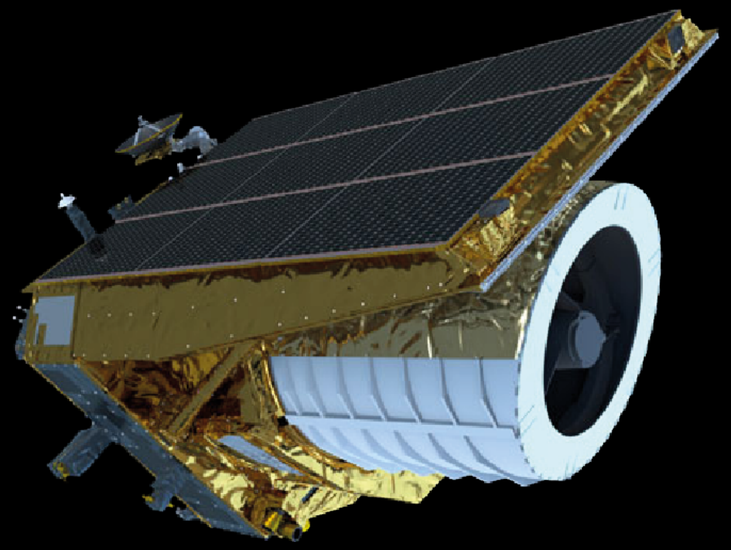


Strong gravitational lensing

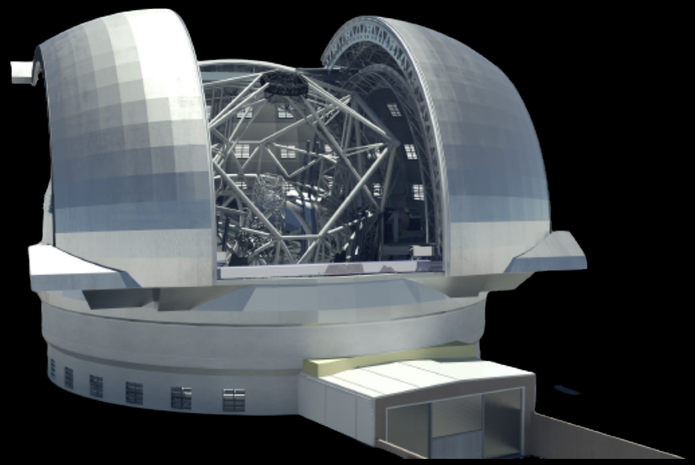
$\sim O(10^5)$ system in the near future (Collet 2015)



Vera C. Rubin Observatory
Large Synoptic Sky Survey



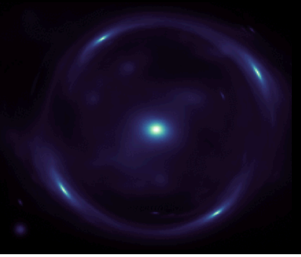
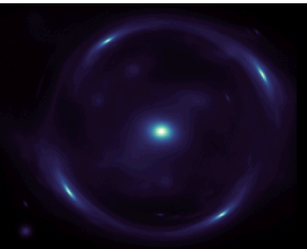
Euclid



Extremely Large Telescope

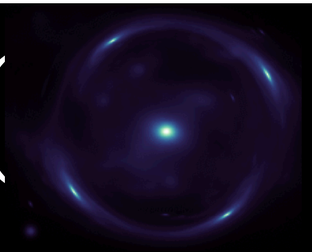
Strong gravitational lensing

A great inference challenge

$p(z \mid$  $) = \dfrac{p($  $\mid z)}{p($  $)} p(z)$

Strong gravitational lensing

A great inference challenge

$p(z \mid$  $) = \dfrac{p($  $\mid z)}{p($  $)} p(z)$

$\mathbf{z}_{\text{light}}$ $\mathbf{z}_{\text{source}}$ $\mathbf{z}_{\text{lens}}$ $\mathbf{z}_{\text{sub}} = \{(x, y, z, M)_1, \ldots, (x, y, z, M)_N\}$

$\boldsymbol{\theta}_{\text{sub}}$

Truncated Marginal Neural Ratio Estimation

A sequential simulation-based inference technique

For more information:

- Hermans et al. (2019)
- Miller et al. (2020)
- Miller et al. (2021)
- ***swyft*** package

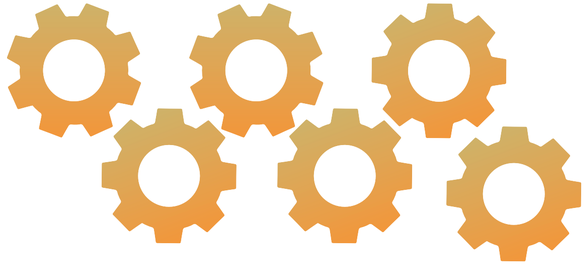
Truncated Marginal Neural Ratio Estimation

A sequential simulation-based inference technique

SBI: Can handle complex forward models

It is possible to improve the realism of the model without dealing with an *intractable* likelihood, only the ability to sample is needed.

$$\mathbf{x}, \mathbf{z} \sim p(\mathbf{x} \,|\, \mathbf{z}) p(\mathbf{z})$$



Truncated Marginal Neural Ratio Estimation

A sequential simulation-based inference technique

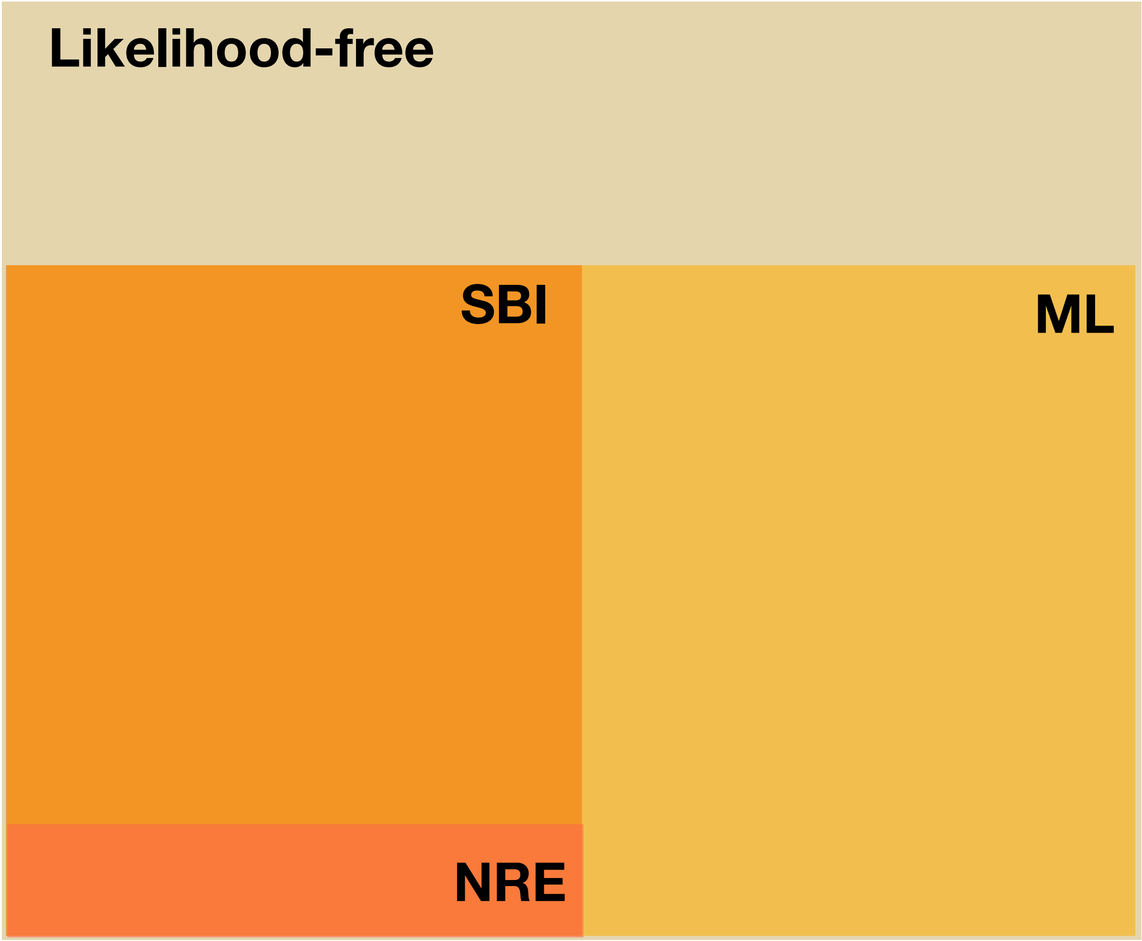
Neural Ratio Estimation

$$r(\mathbf{x};\mathbf{z}) = \frac{p(\mathbf{z}\,|\,\mathbf{x})}{p(\mathbf{z})} = \frac{p(\mathbf{x}\,|\,\mathbf{z})}{p(\mathbf{x})} = \boxed{\frac{p(\mathbf{x},\mathbf{z})}{p(\mathbf{x})p(\mathbf{z})}}$$

NRE rephrases posterior inference into a **binary classification problem** and then solves it by training a neural network on simulated data.

H0: $(\mathbf{x},\mathbf{z}) \sim p(\mathbf{x},\mathbf{z})$

H1: $(\mathbf{x},\mathbf{z}) \sim p(\mathbf{x})p(\mathbf{z})$

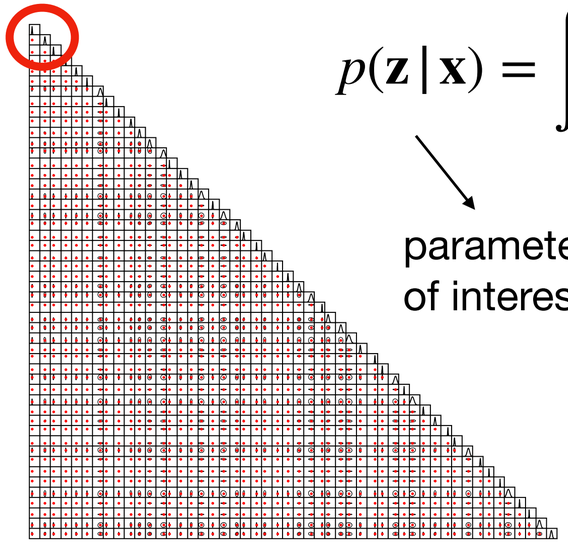


Truncated Marginal Neural Ratio Estimation

A sequential simulation-based inference technique

Marginal inference

Cherry-picking what we are interested in

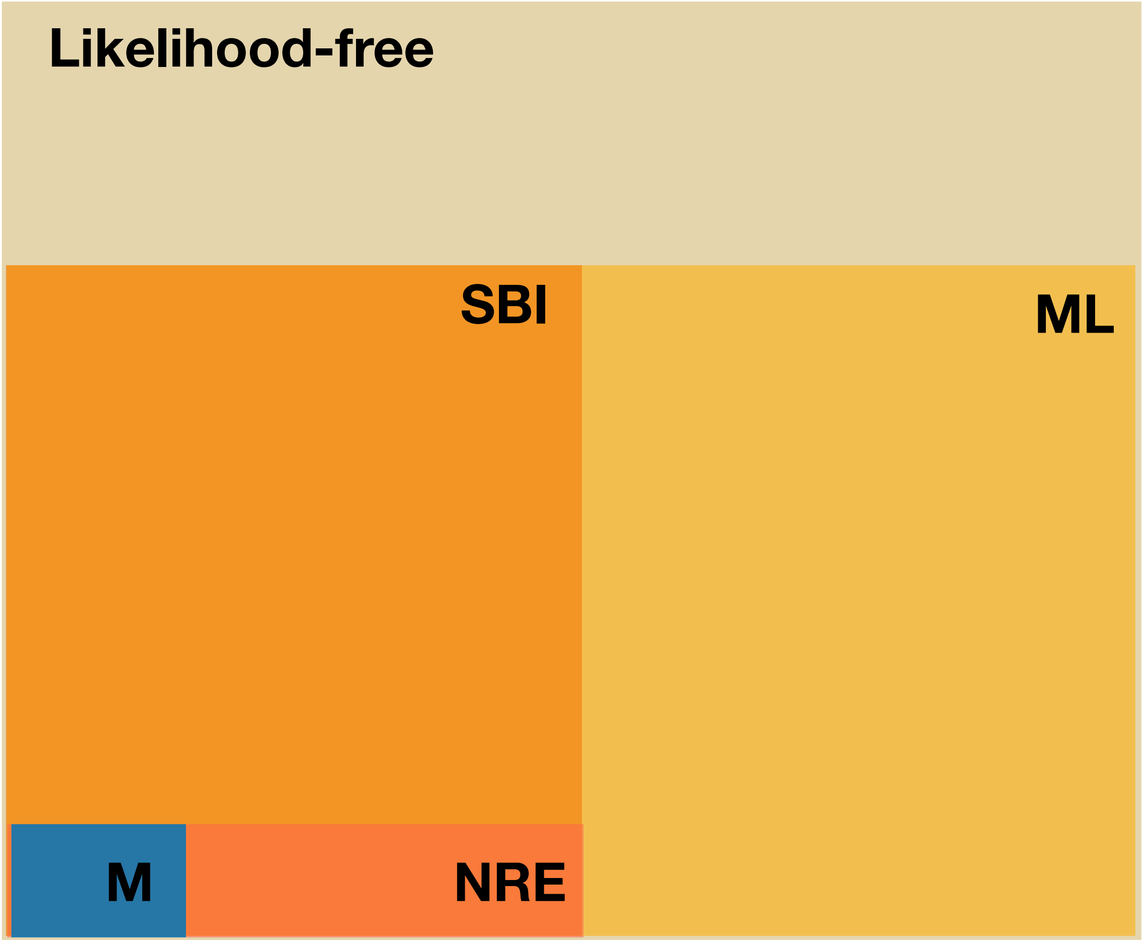


$$p(\mathbf{z}\,|\,\mathbf{x}) = \int \mathrm{d}\boldsymbol{\eta}\, p(\mathbf{z}, \boldsymbol{\eta}\,|\,\mathbf{x})$$

parameters of interest

nuisance parameters

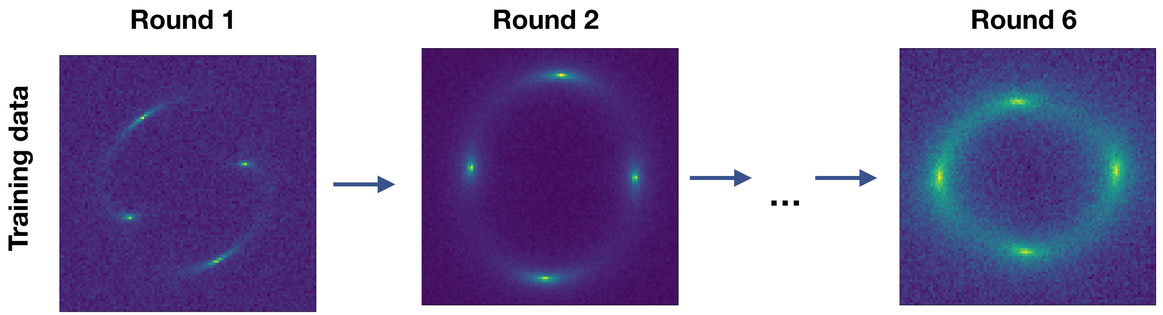
→ **Scalability** with dimensionality



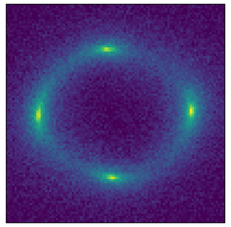
Truncated Marginal Neural Ratio Estimation

A sequential simulation-based inference technique

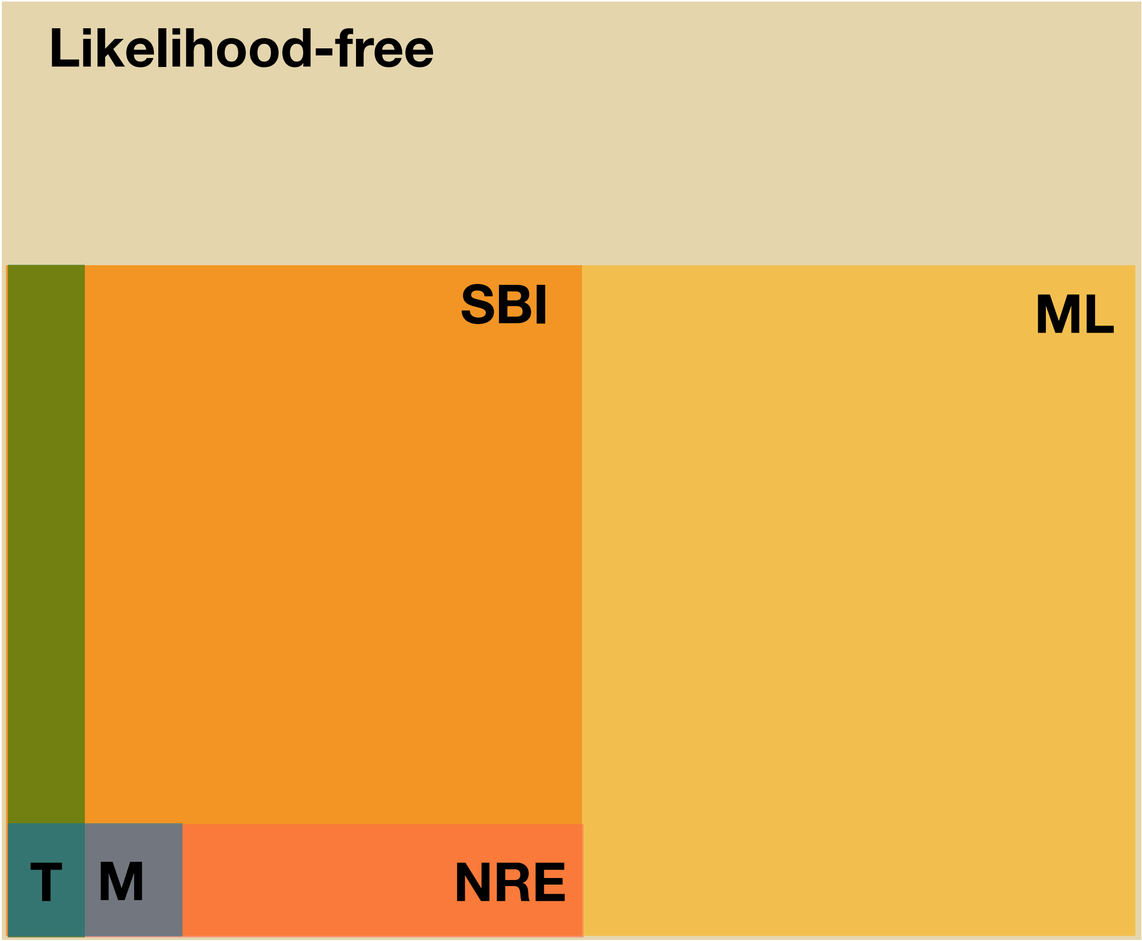
Sequential inference through prior truncation



Target observation

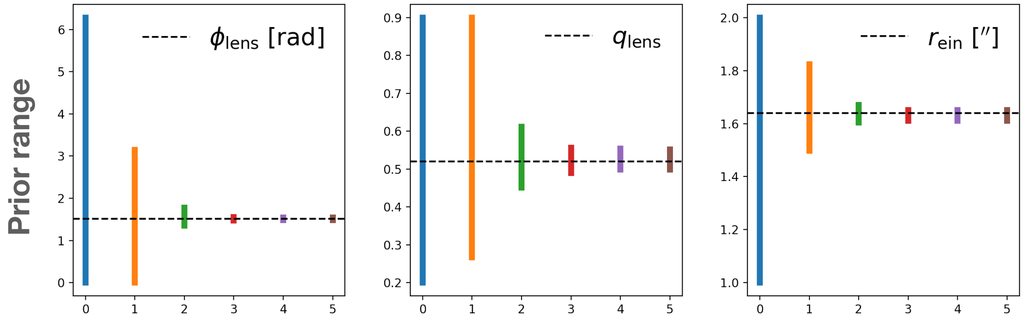


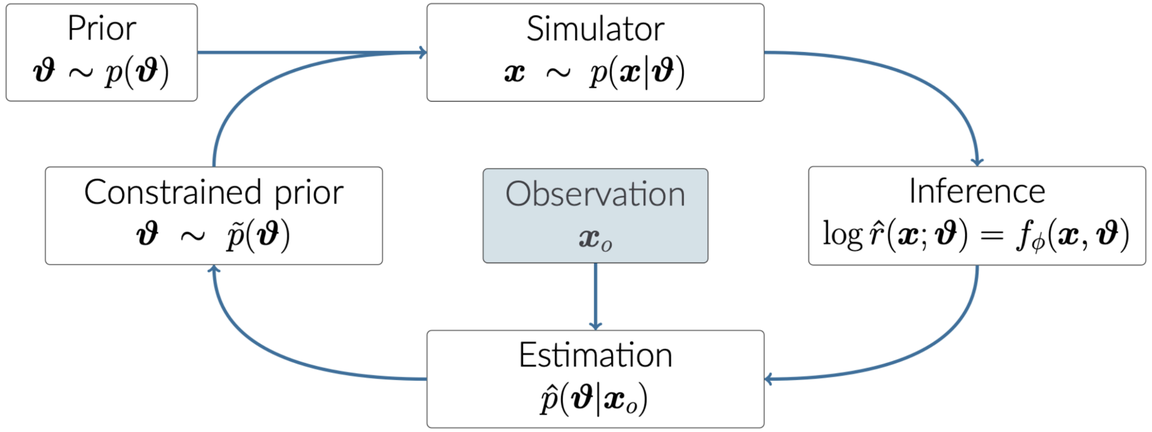
Learning with **high-precision**



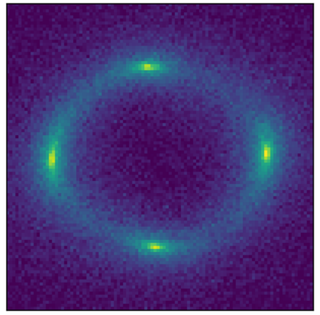
TMNRE for strong gravitational lensing

Reducing data variance



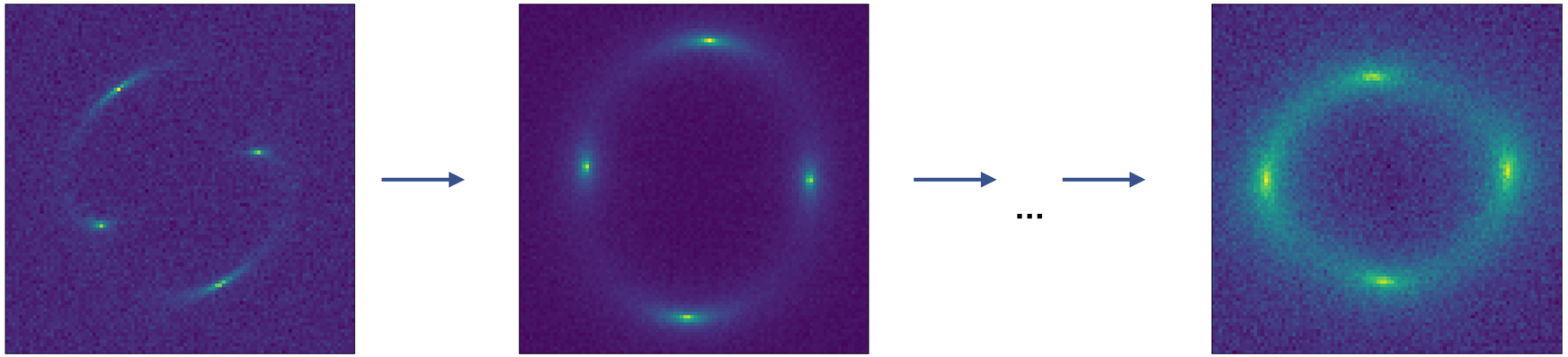


Target mock observation



Round 1 **Round 2** **Round 6**

Training data



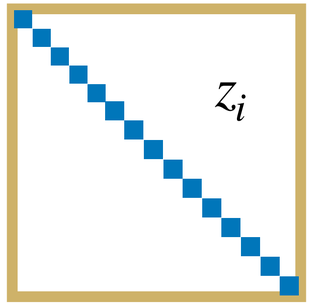
arxiv:2205.09126

x

Prior truncation schemes

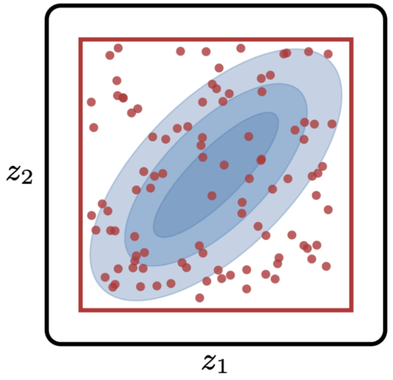
Parameter-wise truncation, based on the corresponding 1D marginal ratio.

$$\Gamma_i^{(R)} = \{z_i \in \mathbb{R} \mid r_R(z_i; \mathbf{x}) > \epsilon\}$$



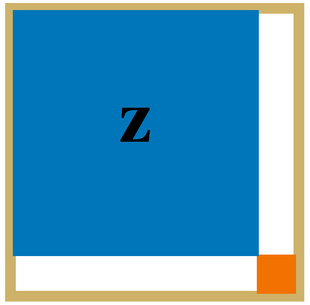
As a result, we obtain a truncation region that has the shape of a **hyper-rectangular box**.

$$\tilde{p}_R(\mathbf{z}) = \frac{1}{Z}\mathbb{I}(z_1 \in \Gamma_1^{(R-1)}) \times \cdots \times \mathbb{I}(z_N \in \Gamma_N^{(R-1)})p(\mathbf{z})$$



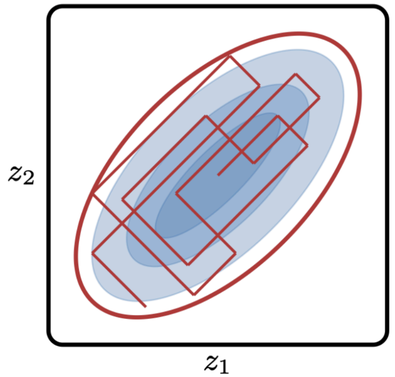
Block-wise truncation, to separate the parameters in blocks depending on what dominates data variance.

$$\Gamma^{(R)} = \{\mathbf{z} \in \mathbb{R}^N \mid r_R(\mathbf{z}; \mathbf{x}) > \epsilon\}$$



The **complex truncation region** is defined through a **hard likelihood constraint**. Sampling from this region is possible with nested sampling techniques (e.g. slice-sampling).

$$\tilde{p}_R(\mathbf{z}) = \frac{1}{Z}\mathbb{I}(\mathbf{z} \in \Gamma^{(R-1)})p(\mathbf{z})$$



TMNRE applications to strong gravitational lensing

1. **Needle in the haystack problem**

   Subhalo's parameters inference

2. **Hierarchical inference: distilling information from a dataset**

   Substructure population parameters inference from lensing images

3. **Block-wise truncation scheme for macromodel correlations**

   Deblending lens light and source light

4. **Image reconstruction task**

   Background source flux variations reconstruction

1. Subhalo's parameters inference

Model (~ $O(10^3)$ parameters)

$\mathbf{z}_{\text{macro}}$

$\mathbf{z}_{\text{source}} = \{x, y, \phi, q, r_{\text{e}}, I_e, n\}$

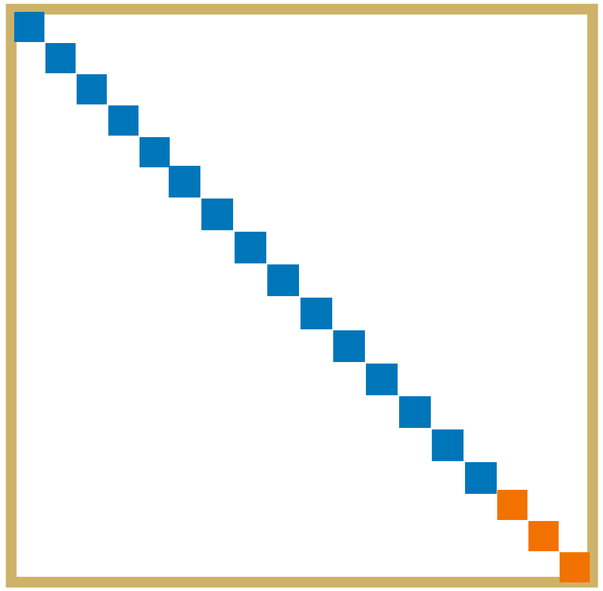
$\mathbf{z}_{\text{lens}} = \{x, y, \phi, q, r_{\text{Ein}}, \gamma, \gamma_1, \gamma_2\}$

$\boldsymbol{\theta}_{\text{sub}}$

$\mathbf{z}_{\text{sub}} = \{(x, y, z, M)_1, \ldots, (x, y, z, M)_N\}$

$\mathbf{z}_{\text{sub,heavy}} = (x, y, z, M_{>M_i})_{\text{heavy}}$

Truncation strategy

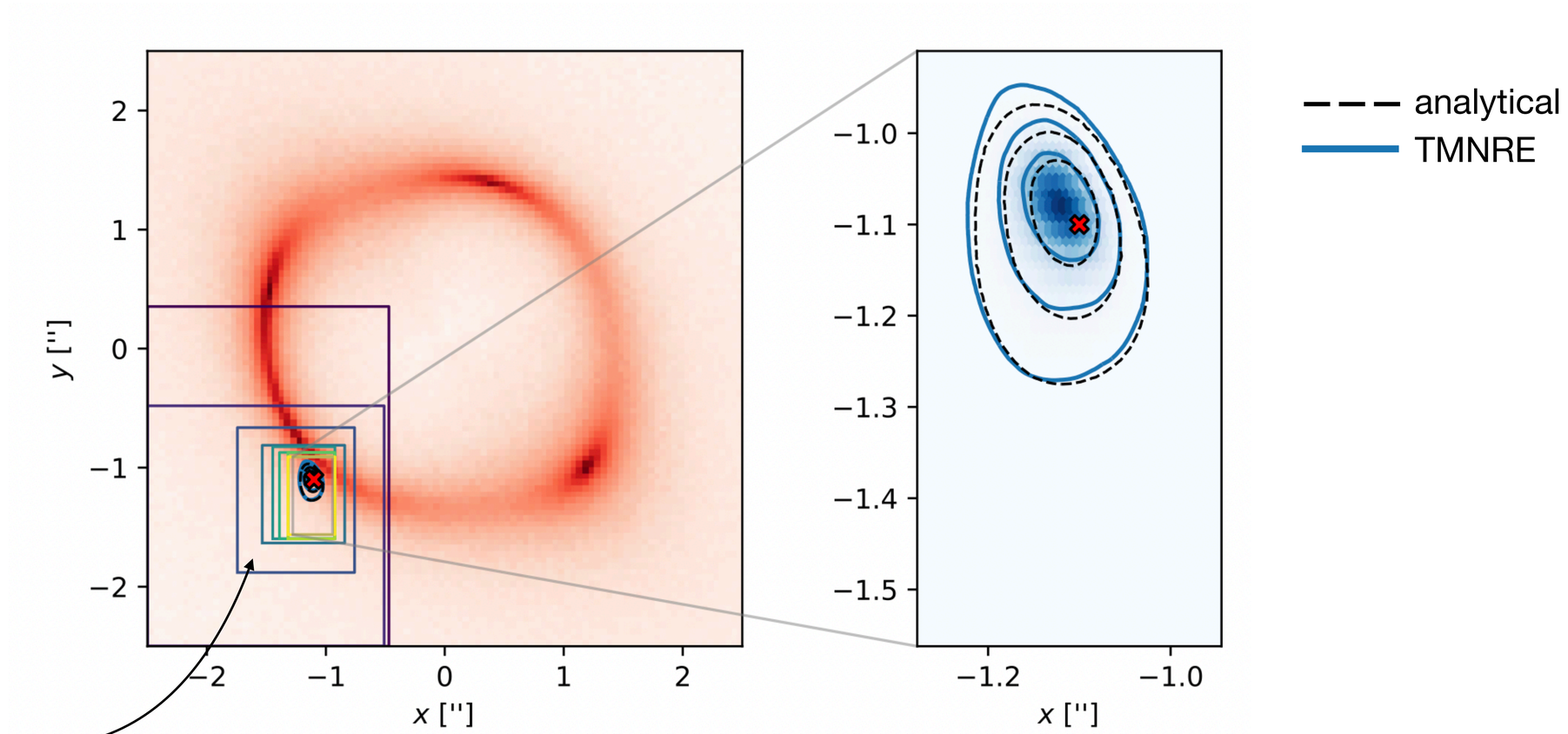


Box truncation on $\mathbf{z}_{\text{macro}}$ and $\mathbf{z}_{\text{sub,heavy}}$

arxiv:2209.09918

1. Subhalo's parameters inference

TMNRE reproduces analytically-calculable posteriors

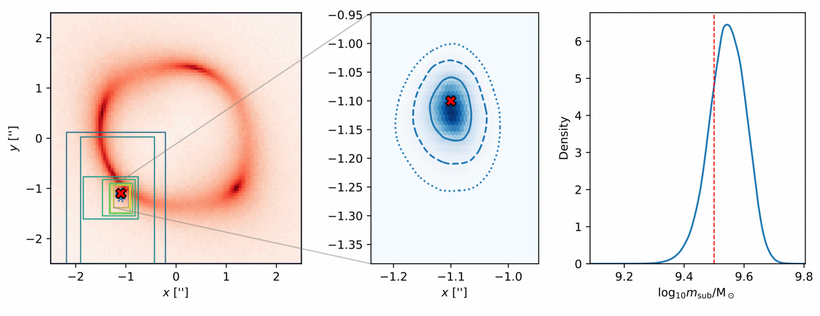


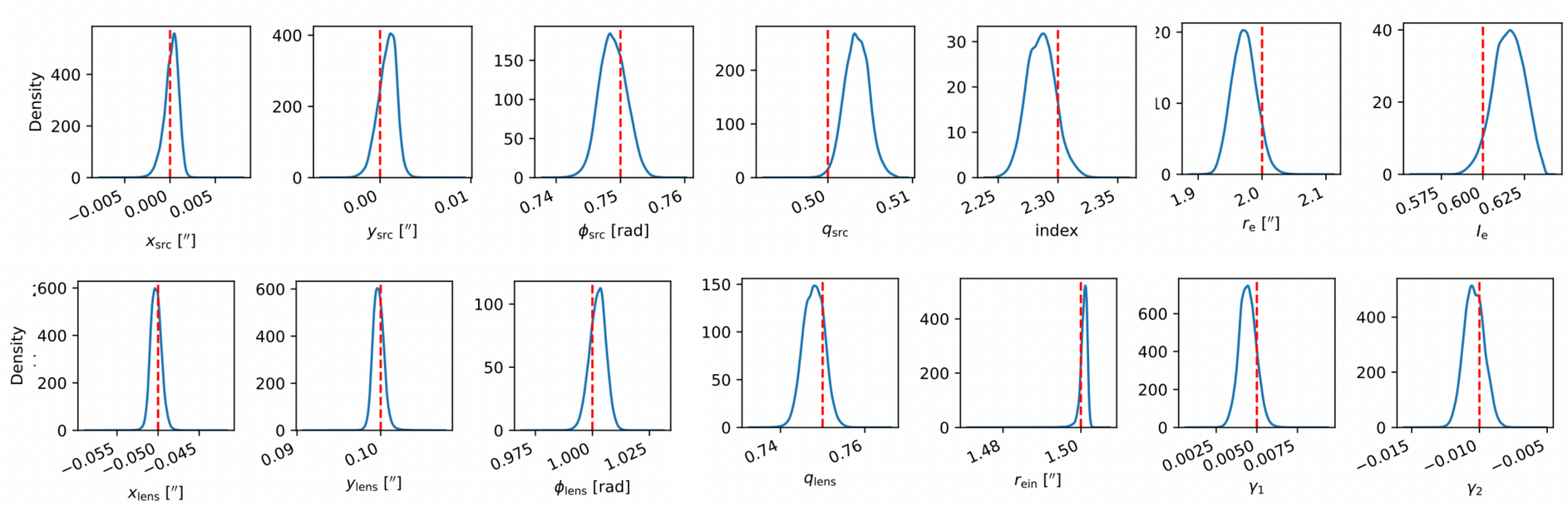
The boxes show the ranges of the truncated prior.

arxiv:2209.09918

1. Subhalo's parameters inference

Effect of the macromodel on subhalo measurements

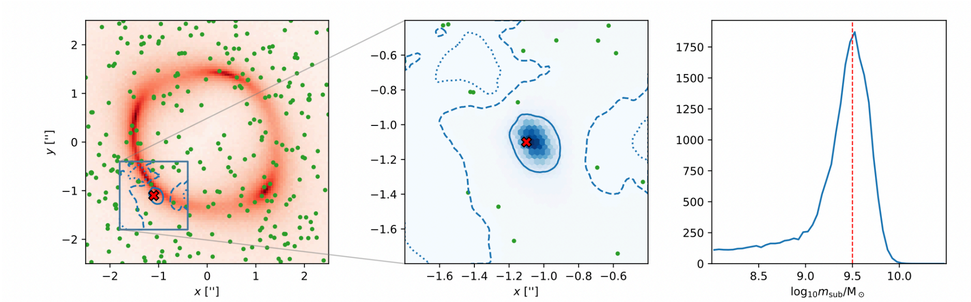


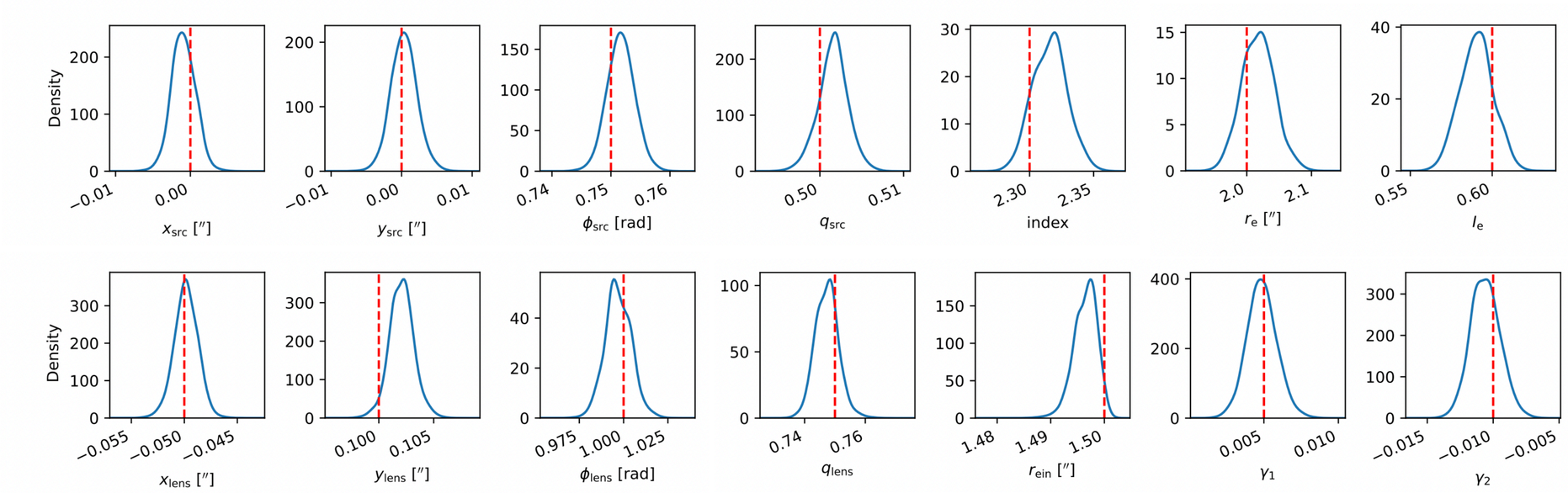


arxiv:2209.09918

1. Subhalo's parameters inference

Effect of the perturber population on subhalo measurements





arxiv:2209.09918

2. Hierarchical inference for substructure population parameter

Model (~ $O(10^3)$ parameters)

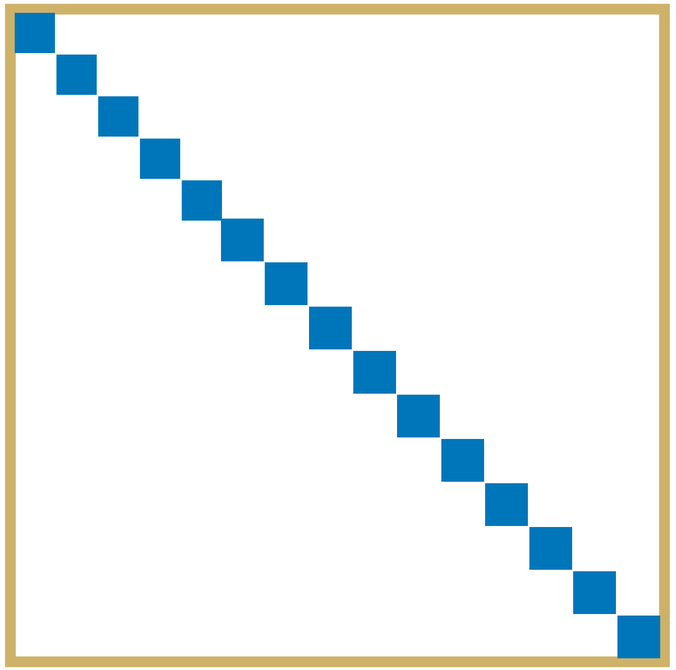
$\mathbf{z}_{\text{macro}}$

$\mathbf{z}_{\text{source}} = \{x, y, \phi, q, r_{\text{e}}, I_e, n\}$

$\mathbf{z}_{\text{lens}} = \{x, y, \phi, q, r_{\text{Ein}}, \gamma, \gamma_1, \gamma_2\}$

$\boldsymbol{\theta}_{\text{sub}}$ — $\mathbf{z}_{\text{sub}} = \{(x, y, z, M)_1, \ldots, (x, y, z, M)_N\}$

Truncation strategy

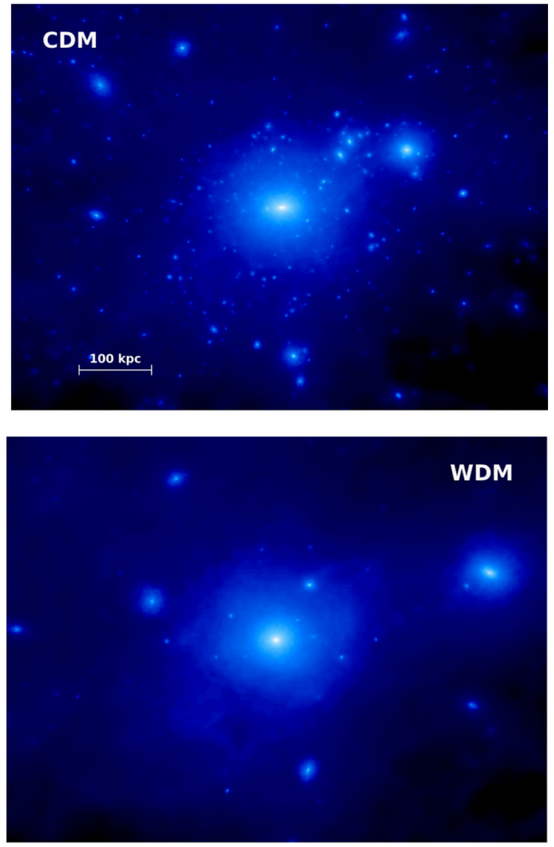


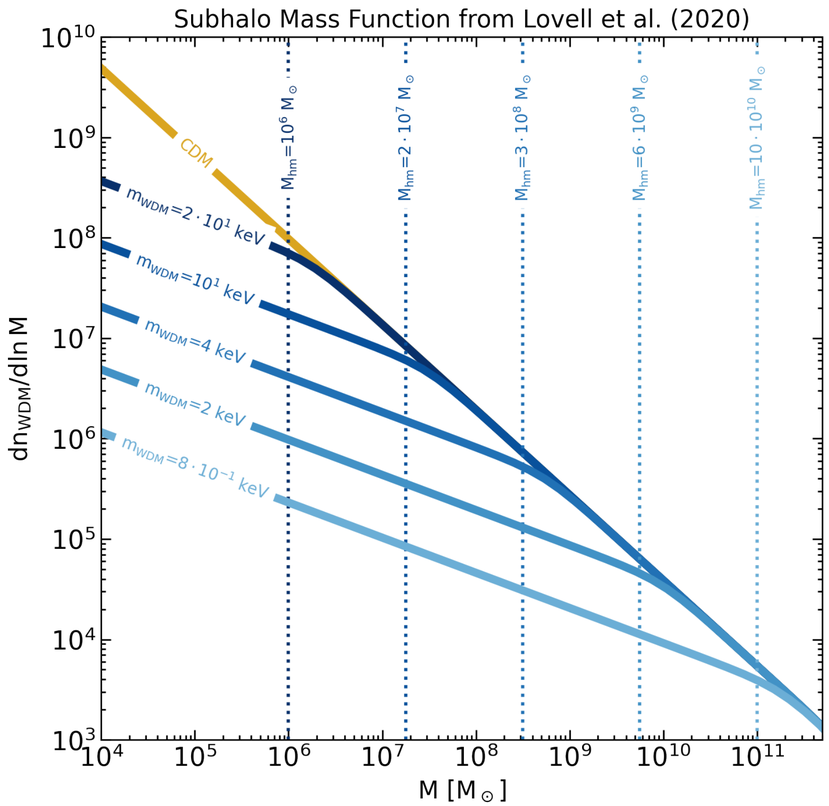
Box truncation on $\mathbf{z}_{\text{macro}}$

arxiv:2205.09126

2. Hierarchical inference

Dark matter is encoded in the properties of small-scale halos

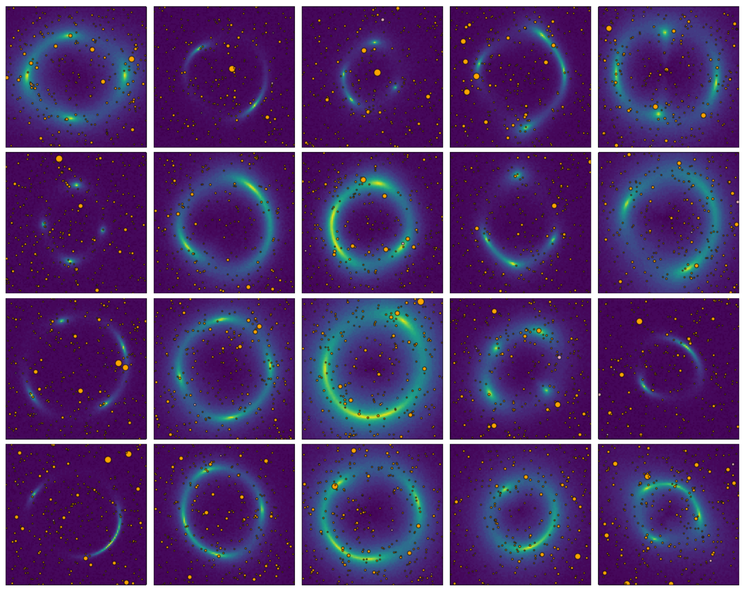


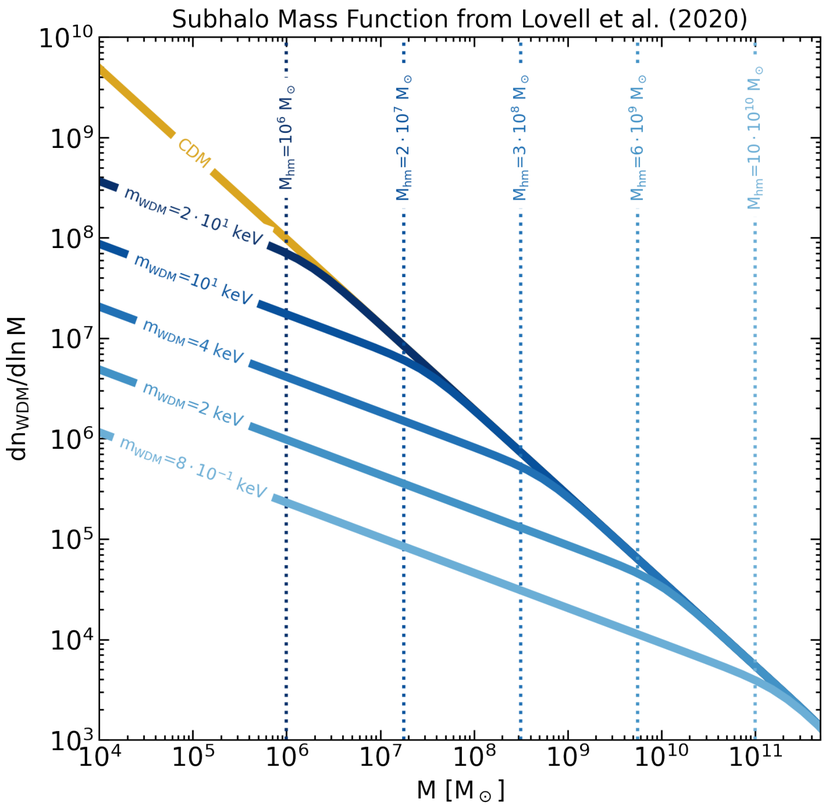


Bœhm et al. (2014)

2. Hierarchical inference

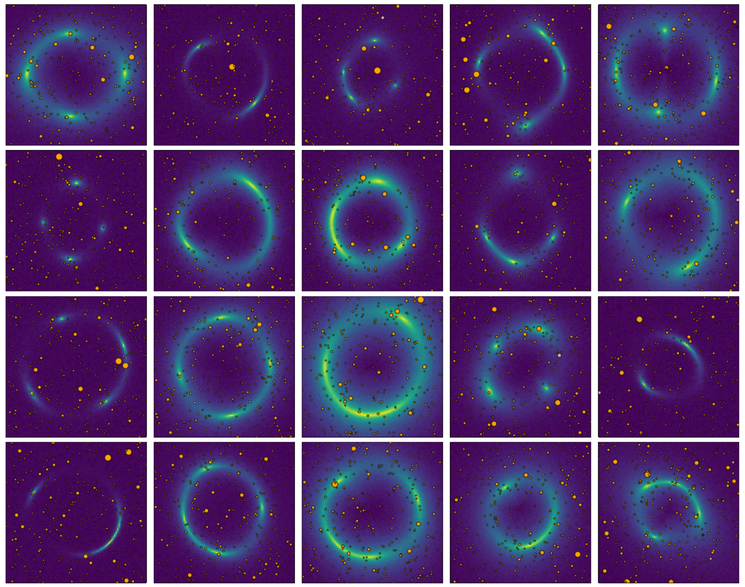
Dark matter is encoded in the properties of small-scale halos

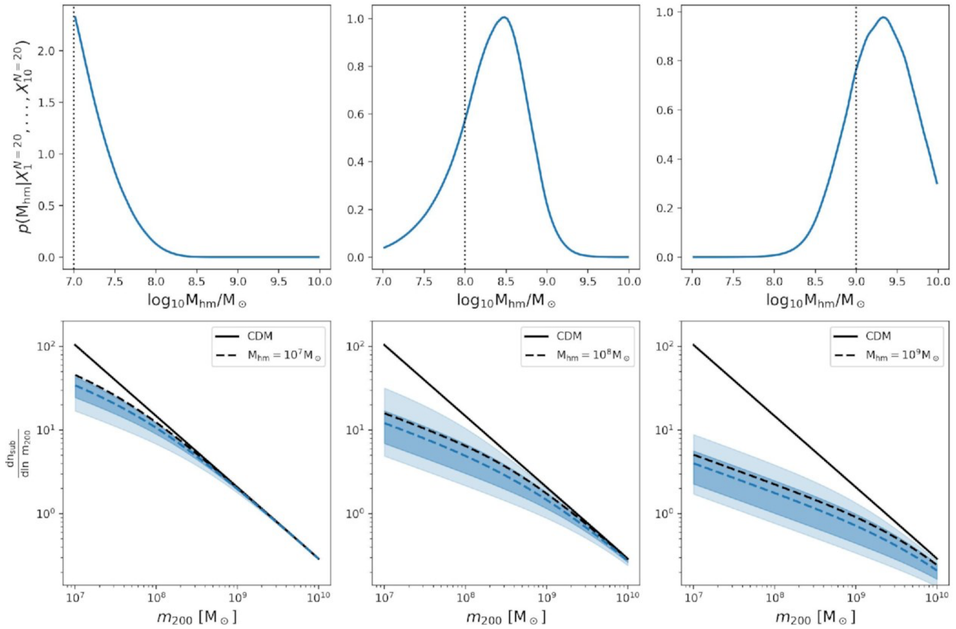




2. Hierarchical inference

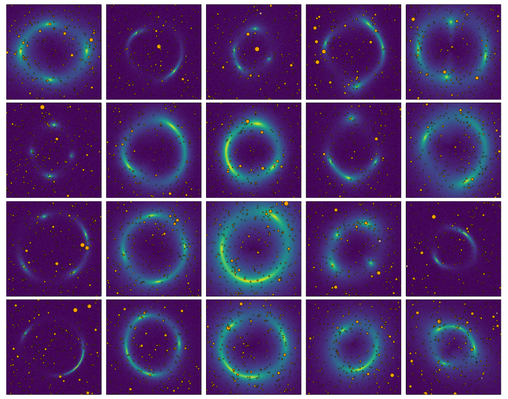
Dark matter is encoded in the properties of small-scale halos

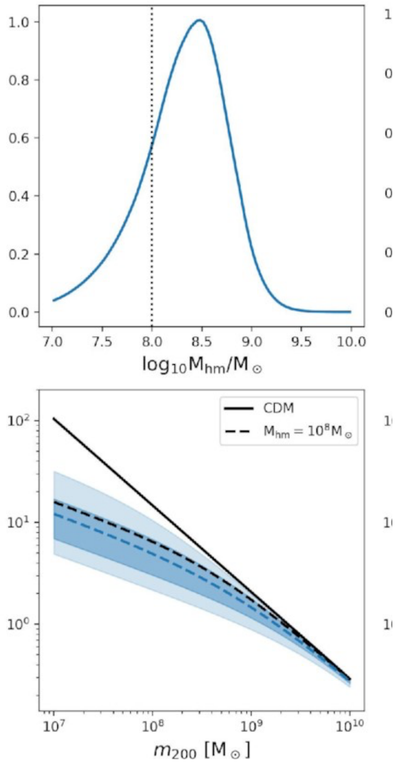




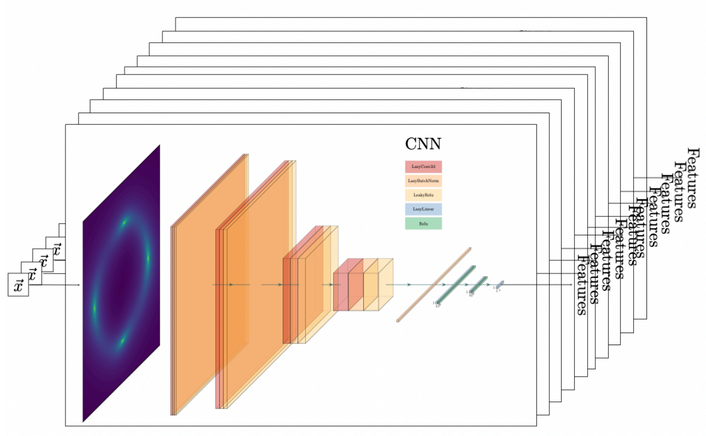
2. Hierarchical inference with NRE

Constraining DM with an *ensamble* of lenses





Distilling information from a dataset through neural embedding



Features from N lenses

$$r(z, \{x_1, \ldots, x_N\})$$

arxiv:2205.09126

3. Macromodel correlations

Model (25 parameters)

$\mathbf{z}_{\text{macro}}$

$$\mathbf{z}_{\text{light}} = \{x, y, \phi, q, r_{\text{e}}, I_e, n\}$$

$$\mathbf{z}_{\text{source}} = \{x, y, \phi, q, r_{\text{e}}, I_e, n\}$$

$$\mathbf{z}_{\text{lens}} = \{x, y, \phi, q, r_{\text{Ein}}, \gamma, \gamma_1, \gamma_2\}$$

$$\mathbf{z}_{\text{sub}} = (x, y, z, M_{>10^8 M_\odot})$$

Truncation strategy

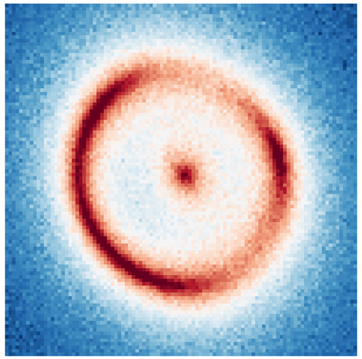
Correlated truncation on $\mathbf{z}_{\text{macro}}$ and $\mathbf{z}_{\text{sub}}$

3. Macromodel correlations

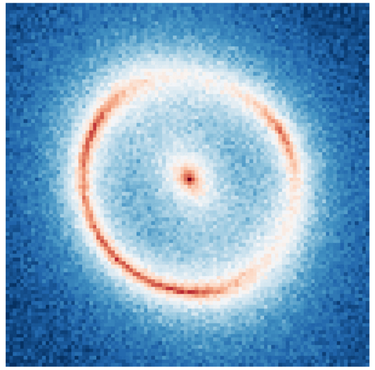
How to reduce data variance when parameters are correlated?

Targeted simulations

Box truncation

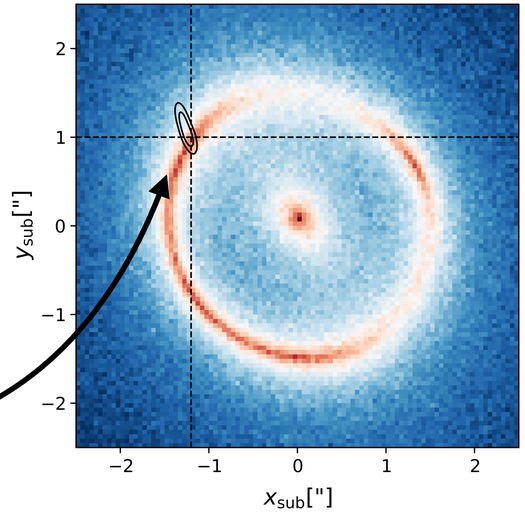


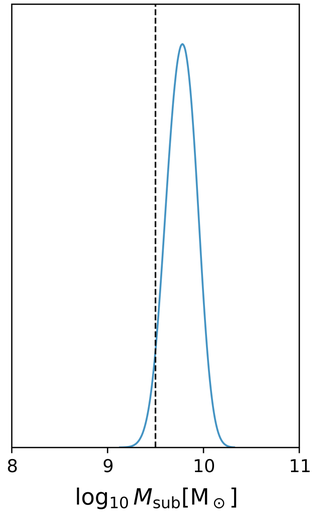
Correlated truncation



Substructure inference

Target observation





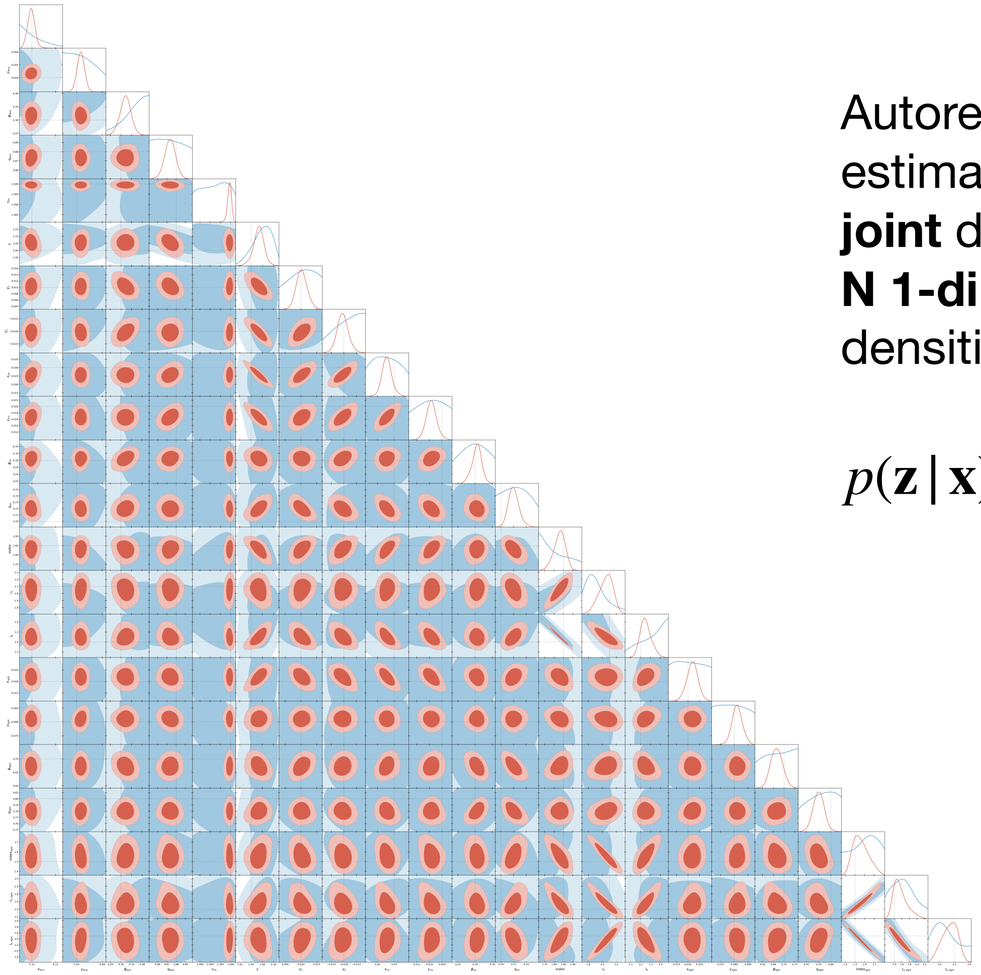
3. Macromodel correlations

Autoregressive NRE vs “vanilla” NRE

ANRE

NRE

Fixed simulation budget and number of weights



Autoregressive models turns the estimation of a **N-dimensional joint** density into the estimation of **N 1-dimensional conditional** densities:

$$p(\mathbf{z}\,|\,\mathbf{x}) = p(z_1\,|\,\mathbf{x}) \prod_{i=2}^{N} p(z_i\,|\,\mathbf{x}, z_{1:i-1})$$

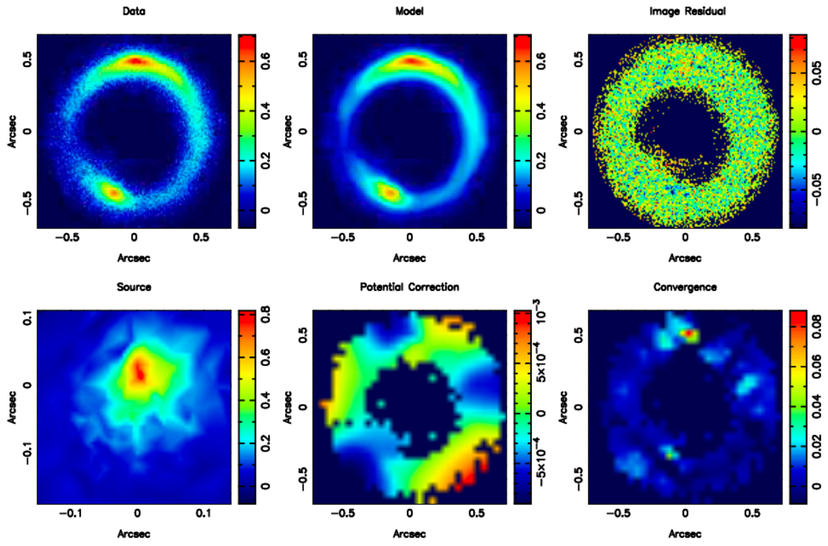
arxiv:2307.xxxx

Towards analysing lensing data with ML

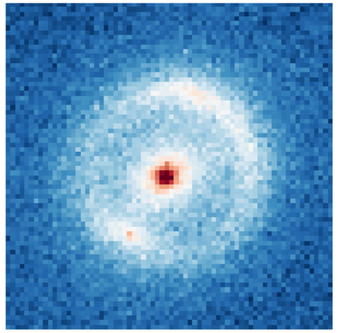
JVAS B1938+666: a case study

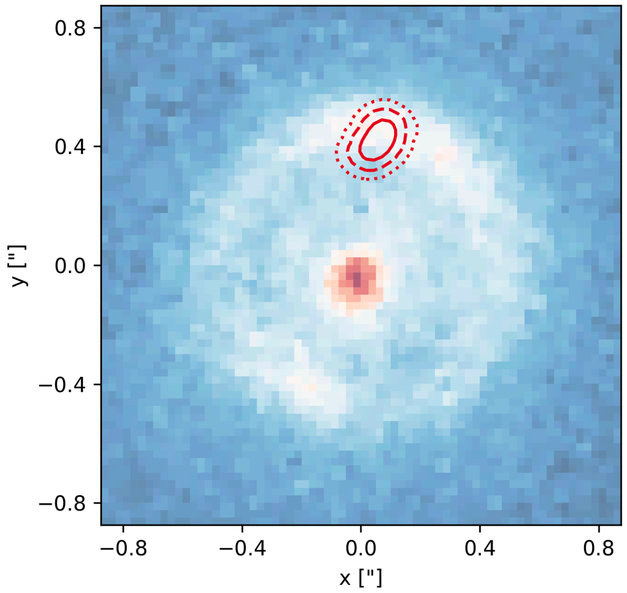
PRELIMINARY

Vegetti et al. (2012) → **detection**



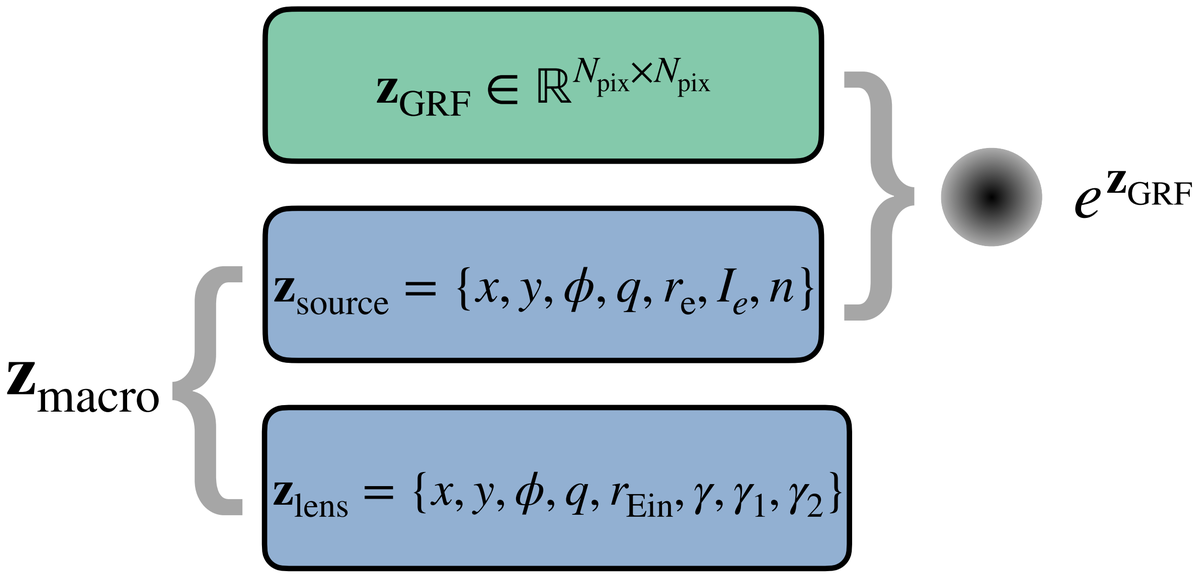
Targeted training data



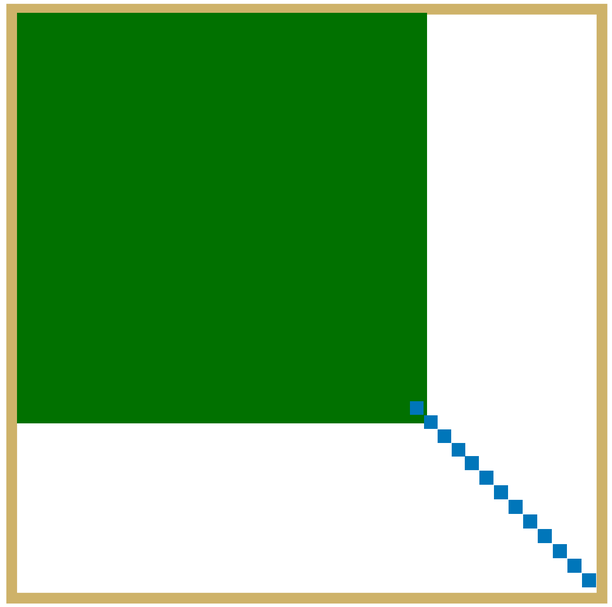


4. Image reconstruction with TMNRE

Model (~ $O(10^3)$ parameters)



Truncation strategy

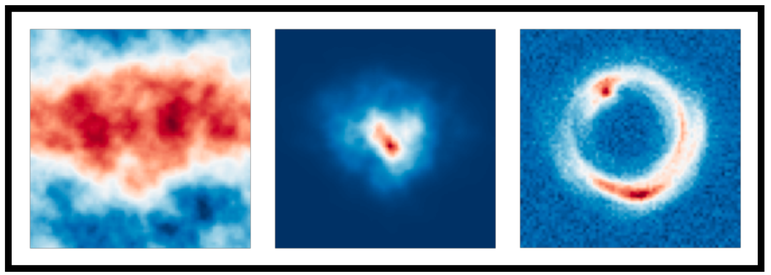


- **Full correlation on $\mathbf{z}_{\mathrm{GRF}}$ (Gaussian approximation)**
- **Box truncation on $\mathbf{z}_{\mathrm{macro}}$**

4. Image reconstruction with TMNRE

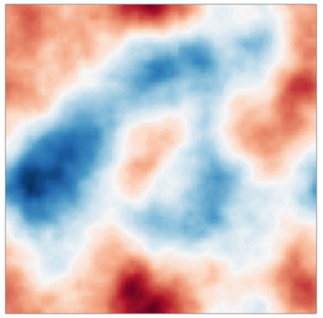
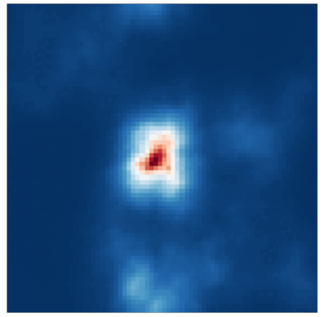
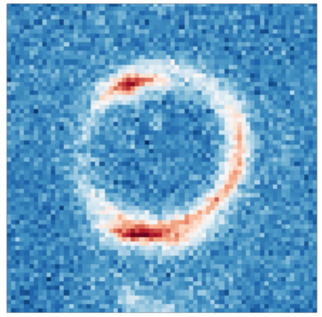
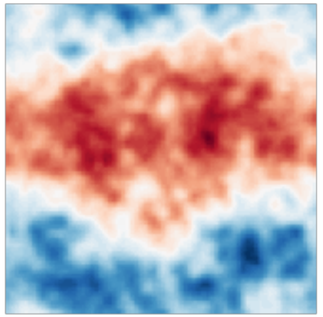
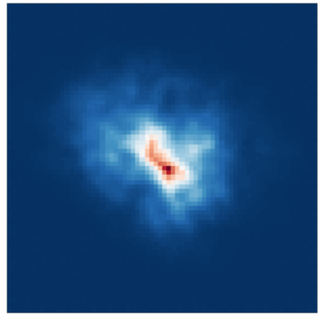
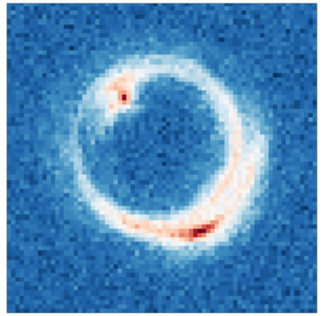
Towards including source variations

Target observation



We train the joined likelihood:

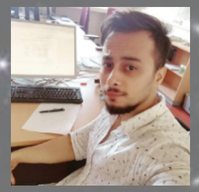
$$\frac{p(\mathbf{x} \mid \mathbf{z})}{p(\mathbf{x})}$$

	GRF	Source	Observation
Prior samples			
Posterior samples after 1 round			

Other TMNRE applications

Gravitational Waves

arXiv:2304.02035



Uddipta Bhardwaj

James Alvey

Stellar Streams

arXiv:2304.02032

James Alvey

Mathis Gerdes

Cosmology

Oleg Savchenko

Guillermo Franco Abellan

Supernovae Ia

arXiv:2209.06733

Konstantin Karchev

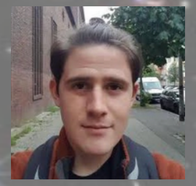
Point sources

arXiv:2211.04291

Noemi Anau Montel

Algorithms development

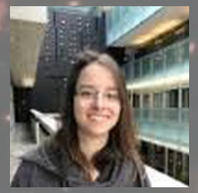
arXiv:2210.06170



Benjamin Miller

Large-scale structure

arXiv:2206.11312



Androniki Dimitriou

Camila Correa

Conclusions

- Strong lensing images analyses **results,** in general, **depend** on **data cuts**, model **approximations**, analysis **methods.**

- With **TMNRE** we hope to combine all known sources of **uncertainties** in the analysis and draw coherent conclusions based on the ***full* model** and ***all* data**.

- This is possible by **assembling different inference strategies** (neural networks and truncation schemes) to **coherently** perform **distinct analysis tasks on the same data**.

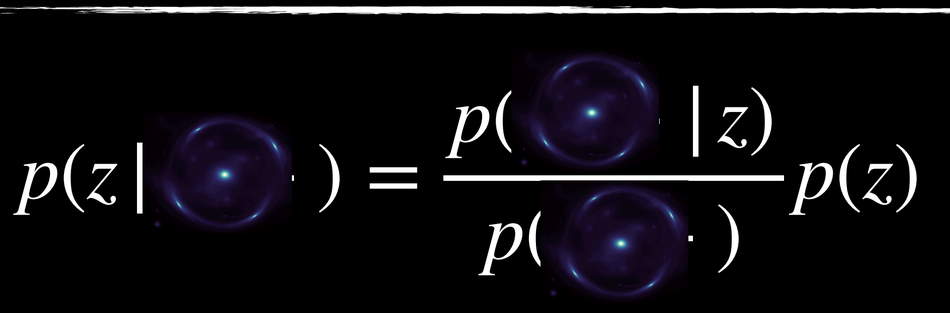


Image reconstruction

Density field reconstruction

Object detection

Population level parameter inference

Light emissions deblending

Hierarchical inference

All marginalizing over other components uncertainties with uncertainties

Conclusions

- Strong lensing images analyses **results,** in general, **depend** on **data cuts**, model **approximations**, analysis **methods.**

- With **TMNRE** we hope to combine all known sources of **uncertainties** in the analysis and draw coherent conclusions based on the ***full* model** and ***all* data**.

- This is possible by **assembling different inference strategies** (neural networks and truncation schemes) to **coherently** perform **distinct analysis tasks on the same data**.