Cosmology with SBI:

Forward Modelling Weak and Strong Lensing Observables

Maximilian von Wietersheim-Kramsta
mwiet.github.io

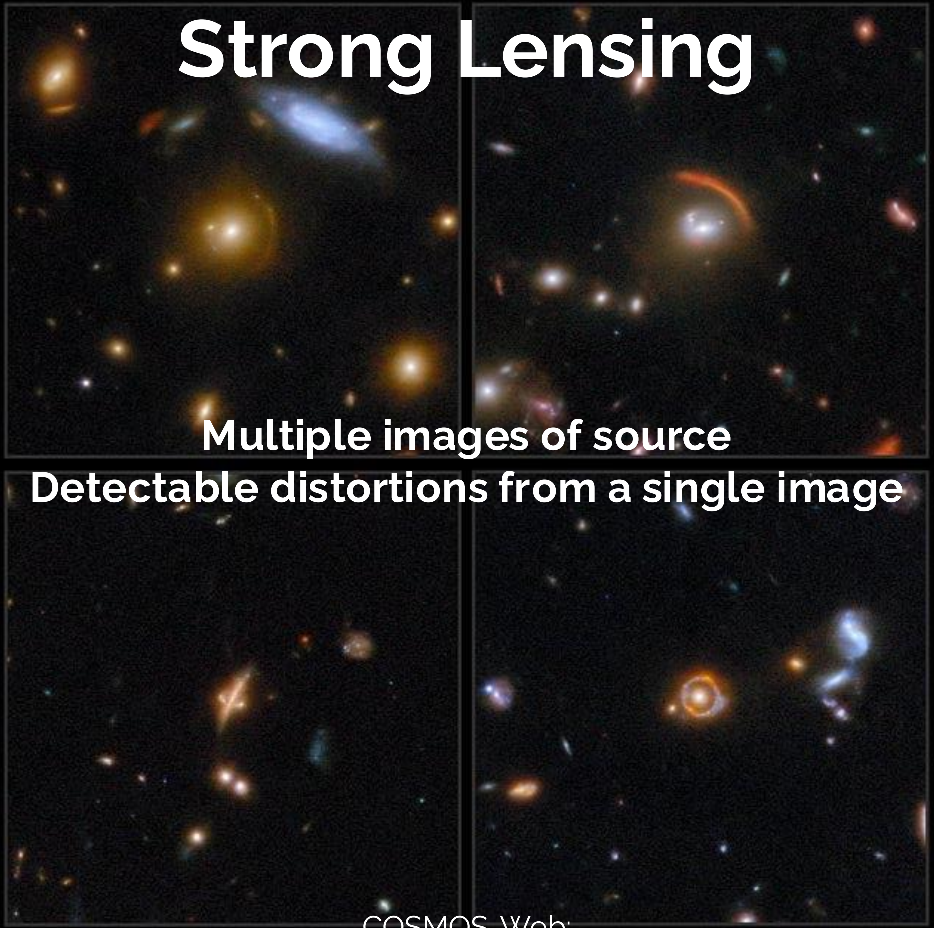
Oskar Klein Centre, Stockholm University

3 Oct. 2025

Gravitational Lensing

Credit: NASA's Goddard Space Flight Center Conceptual Image Lab

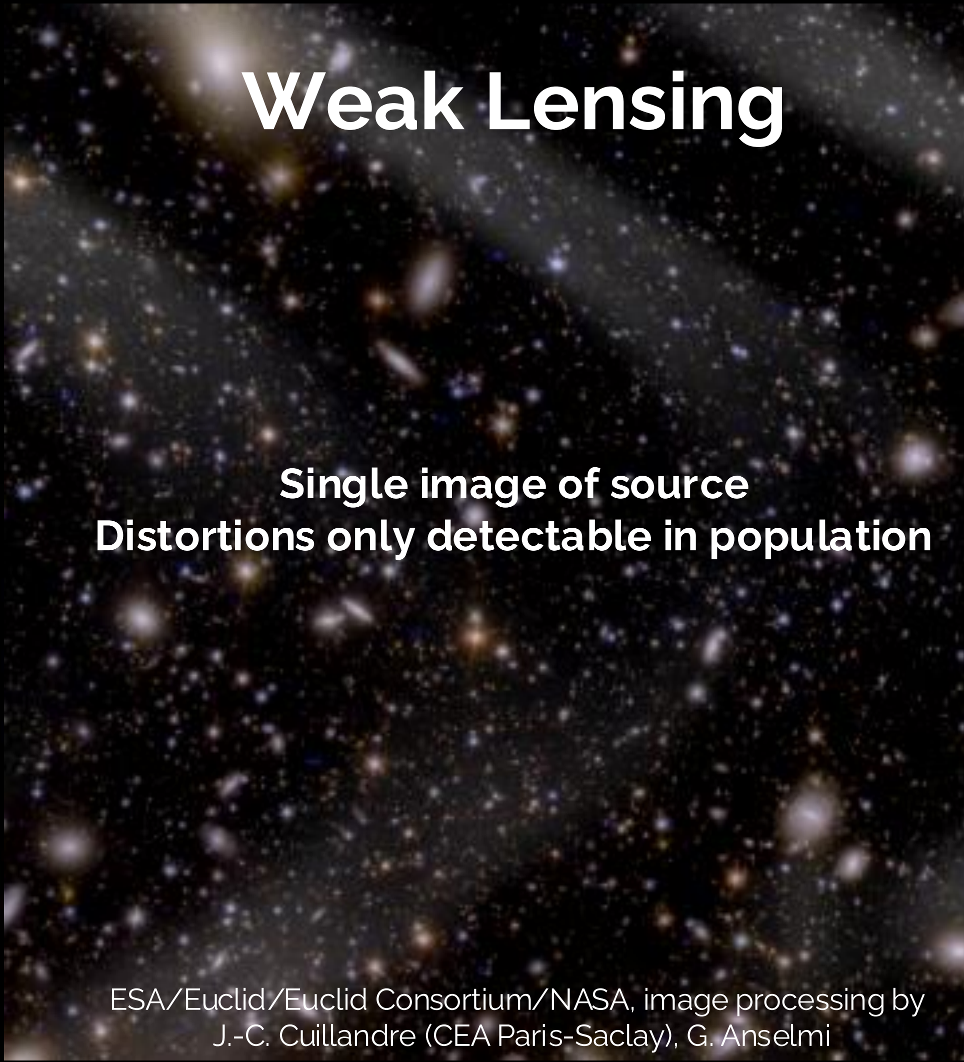
Strong Lensing



Multiple images of source
Detectable distortions from a single image

COSMOS-Web:
Nightingale, J. W., et al. (2025),
MNRAS, *543*(1), 203-222.

Weak Lensing

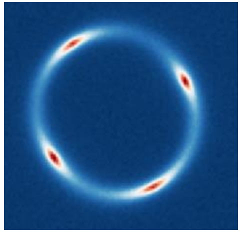


Single image of source
Distortions only detectable in population

ESA/Euclid/Euclid Consortium/NASA, image processing by
J.-C. Cuillandre (CEA Paris-Saclay), G. Anselmi

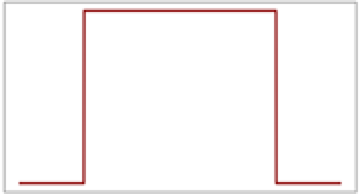
Recipe for Cosmological Inference

$\boldsymbol{d}$



Data

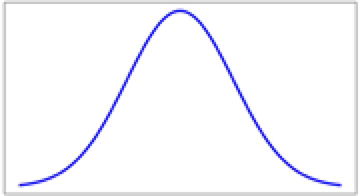
$P(\theta)$



$\boldsymbol{\theta}$

Prior

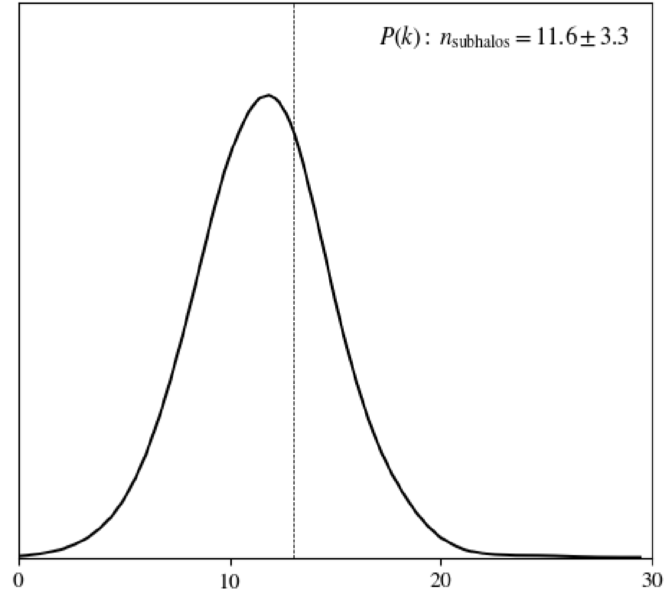
$P(\boldsymbol{d}|\theta)$



$\boldsymbol{d}$

Likelihood

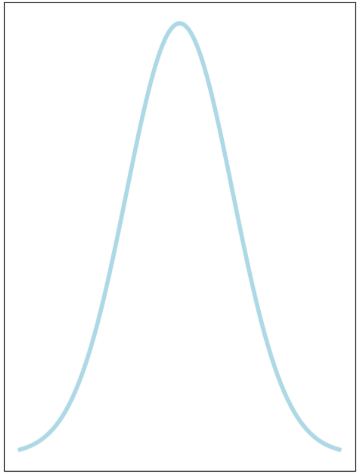
$P(\theta|\boldsymbol{d})$: Posterior



Modelling Likelihoods

Given a model!

$P(\boldsymbol{d}|\boldsymbol{\theta})$

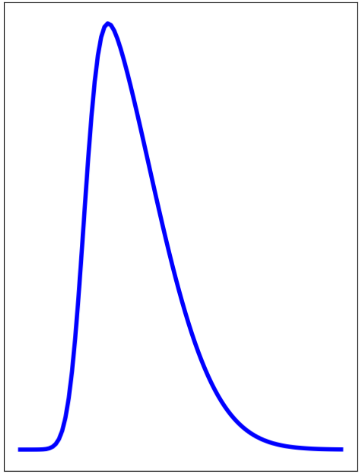


$\boldsymbol{d}$

Analytic

e.g.

$P(d|\theta) \propto e^{-(d-\mu)^2}$

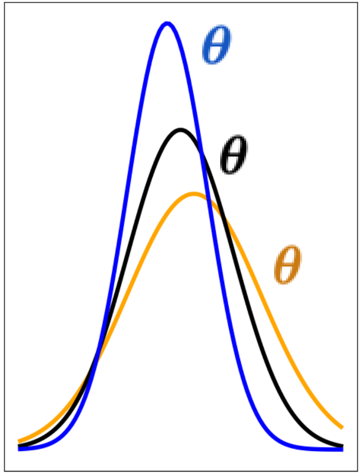


$\boldsymbol{d}$

Biased

e.g.

Instrumental systematics

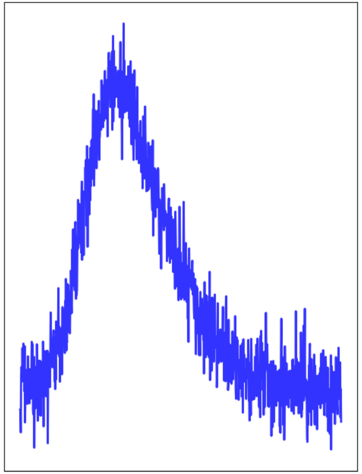


$\boldsymbol{d}$

Signal-dependent uncertainty

e.g.

Cosmic variance



$\boldsymbol{d}$

Intractable

e.g.

Non-trivial selection functions

Bayes' Theorem

Posterior

Likelihood

Prior

Joint probability

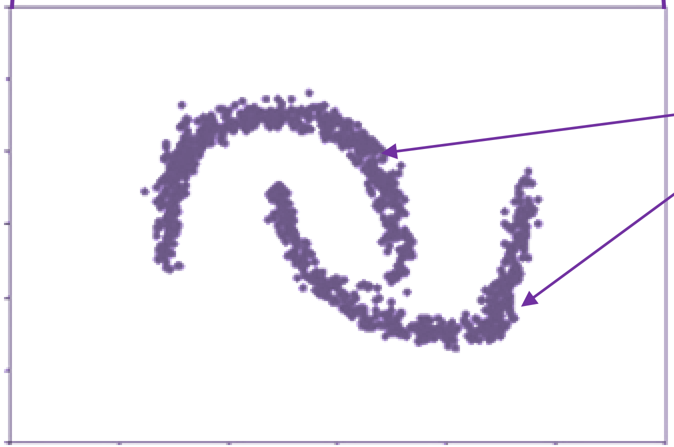
$$P(\boldsymbol{\theta}|\boldsymbol{d}) = \frac{P(\boldsymbol{d}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})}{P(\boldsymbol{d})} \propto P(\boldsymbol{\theta}, \boldsymbol{d}) \cdot P(\boldsymbol{\theta})$$

$\boldsymbol{\theta}$: Model parameters

$\boldsymbol{d}$: Data

Simulation-based
or
likelihood-free or
implicit likelihood
inference

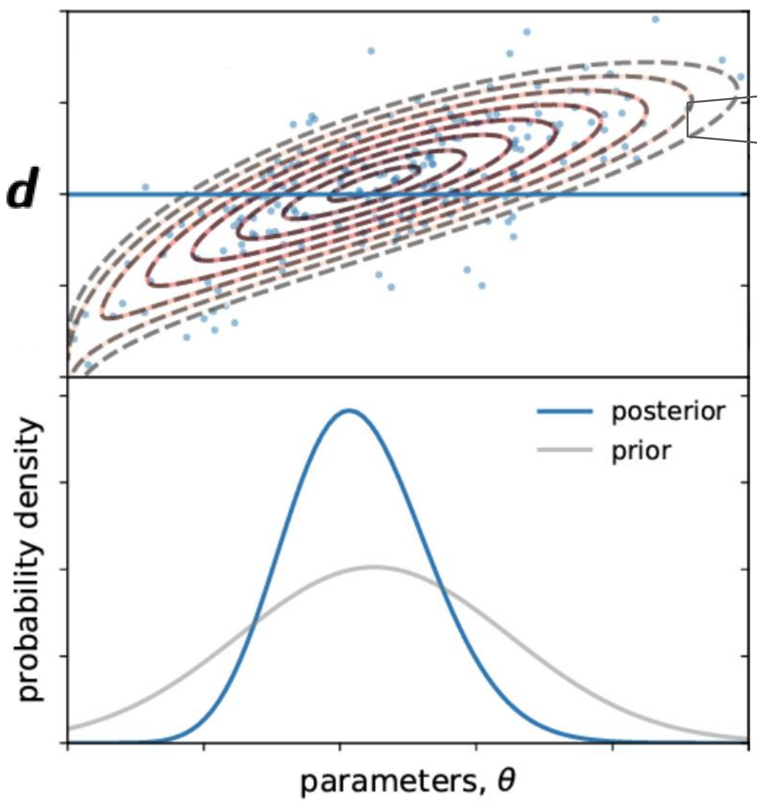
$\boldsymbol{d}$



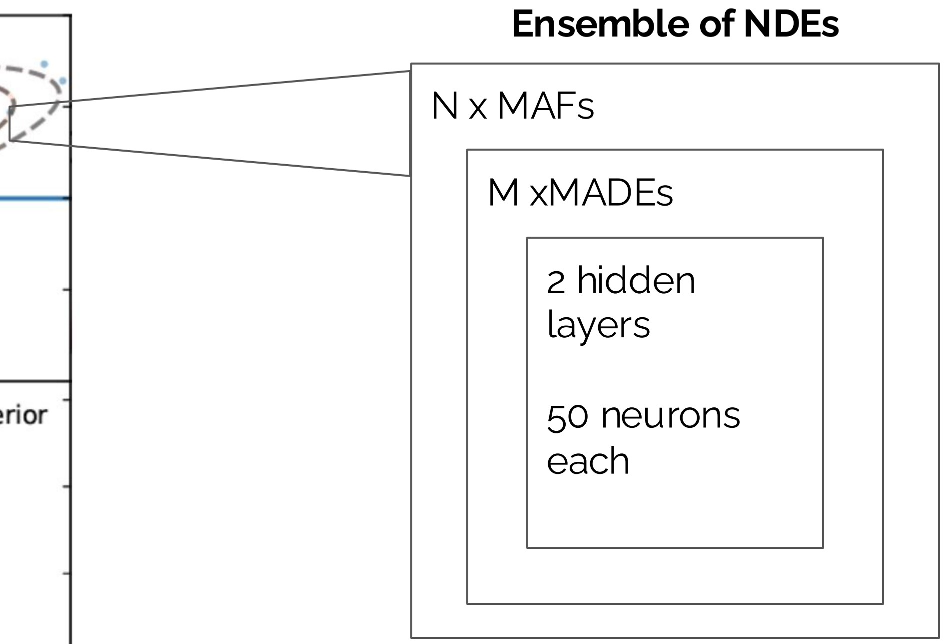
$\boldsymbol{\theta}$

Populate with
simulations
given a model

SBI: Neural Density Estimation



Alsing et al. (2019), MNRAS, 488(3), 4440-4458.



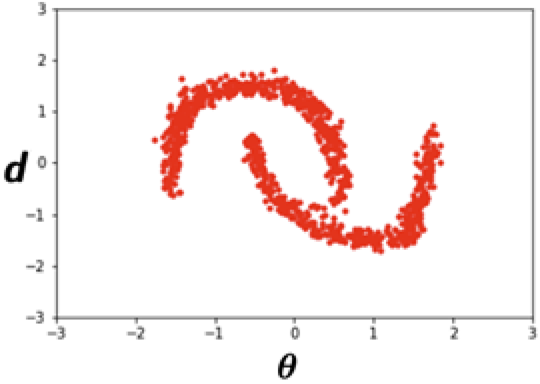
→ N x M x |***d***| multivariate Gaussians, P(***d***|***θ***, **w**)

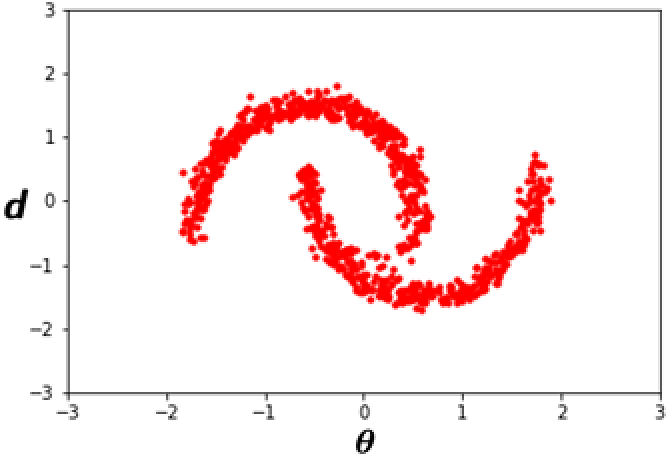
Masked Autoregressive Flows

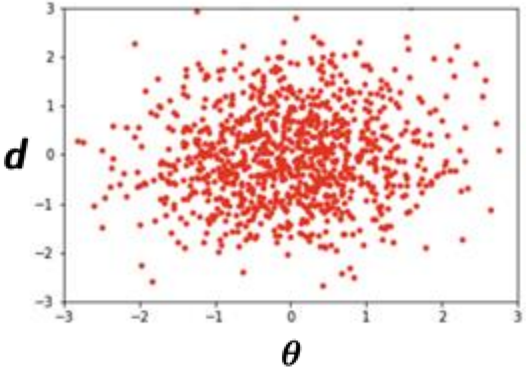
$P(\boldsymbol{\theta}, \boldsymbol{d})$
from simulations

Learn transformations
to

Gaussian
distribution



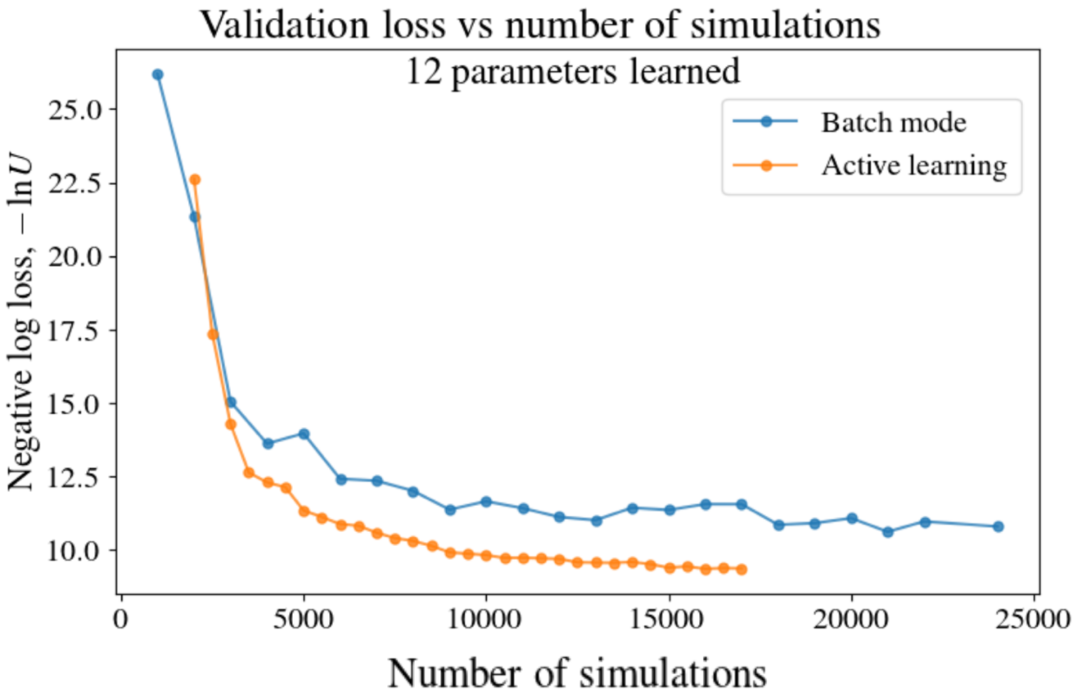




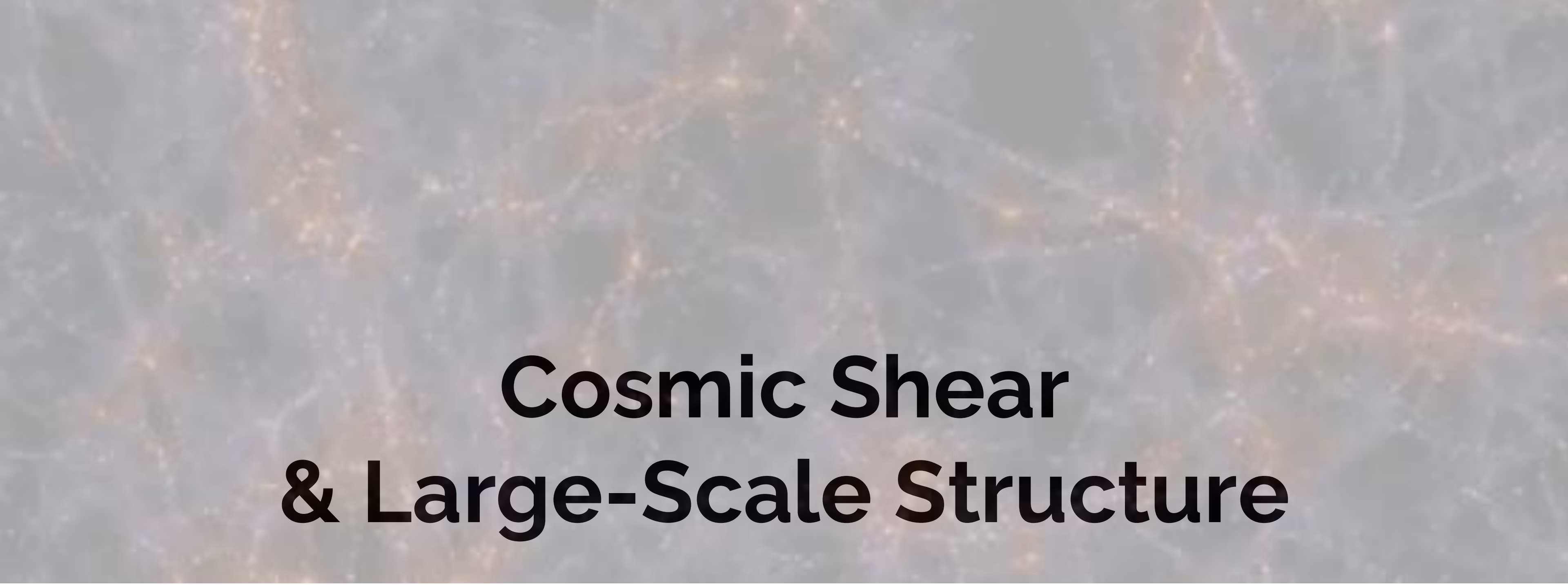
Credit: Eric Jang

Invertible

Masked Autoregressive Flows



Lin, von Wietersheim-Kramsta, et al. (2022), MNRAS 524(4), pp. 6167-6180



Cosmic Shear
& Large-Scale Structure

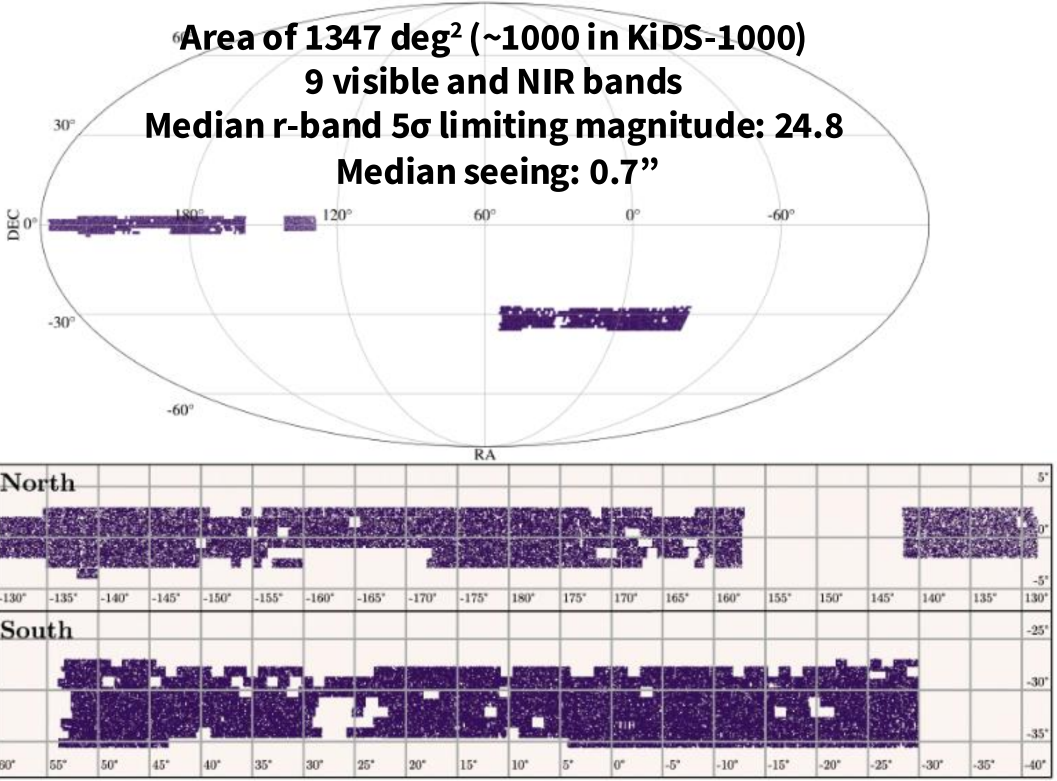
In collaboration with K. Lin, N. Tessore, B. Joachimi, A. Loureiro, R. Reischke, A.H. Wright

von Wietersheim-Kramsta, Lin et al. (2024), A&A 694, A223.

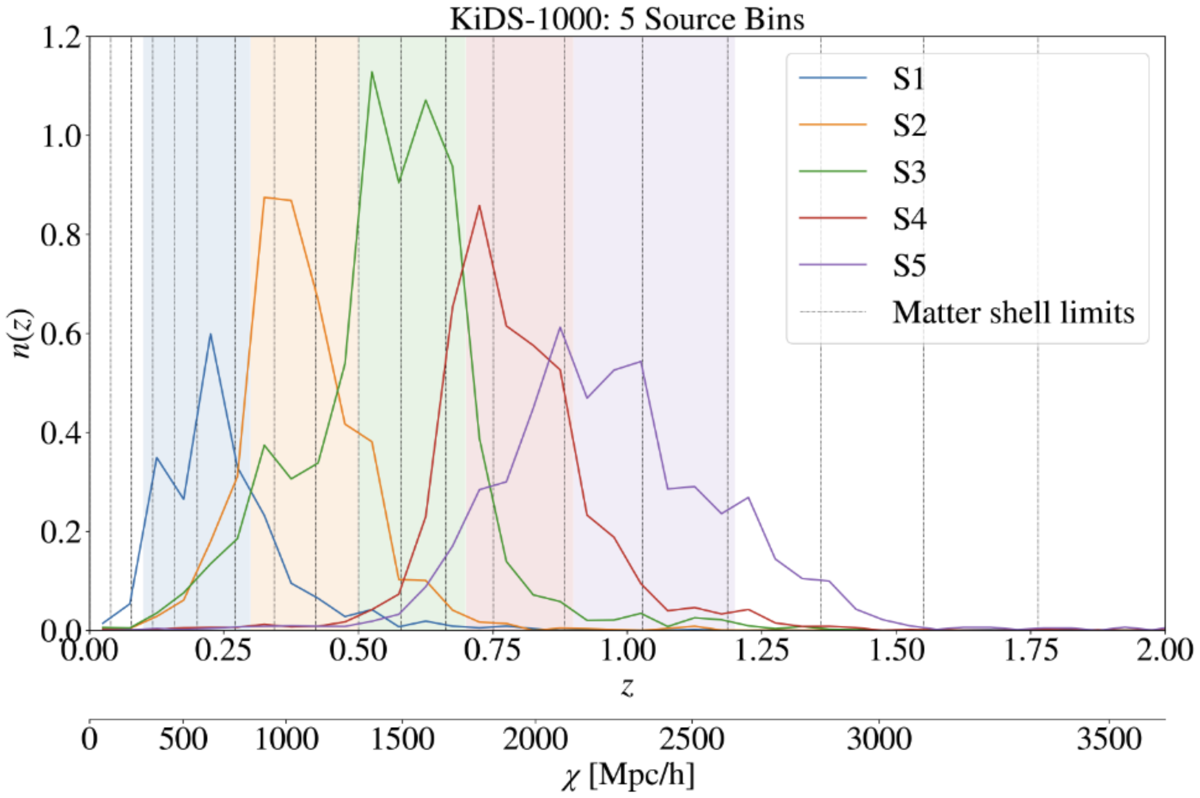
Kilo-Degree Survey

ESO VLT Survey Telescope

Area of 1347 deg^2 (~1000 in KiDS-1000)
9 visible and NIR bands
Median r-band 5σ limiting magnitude: 24.8
Median seeing: 0.7”

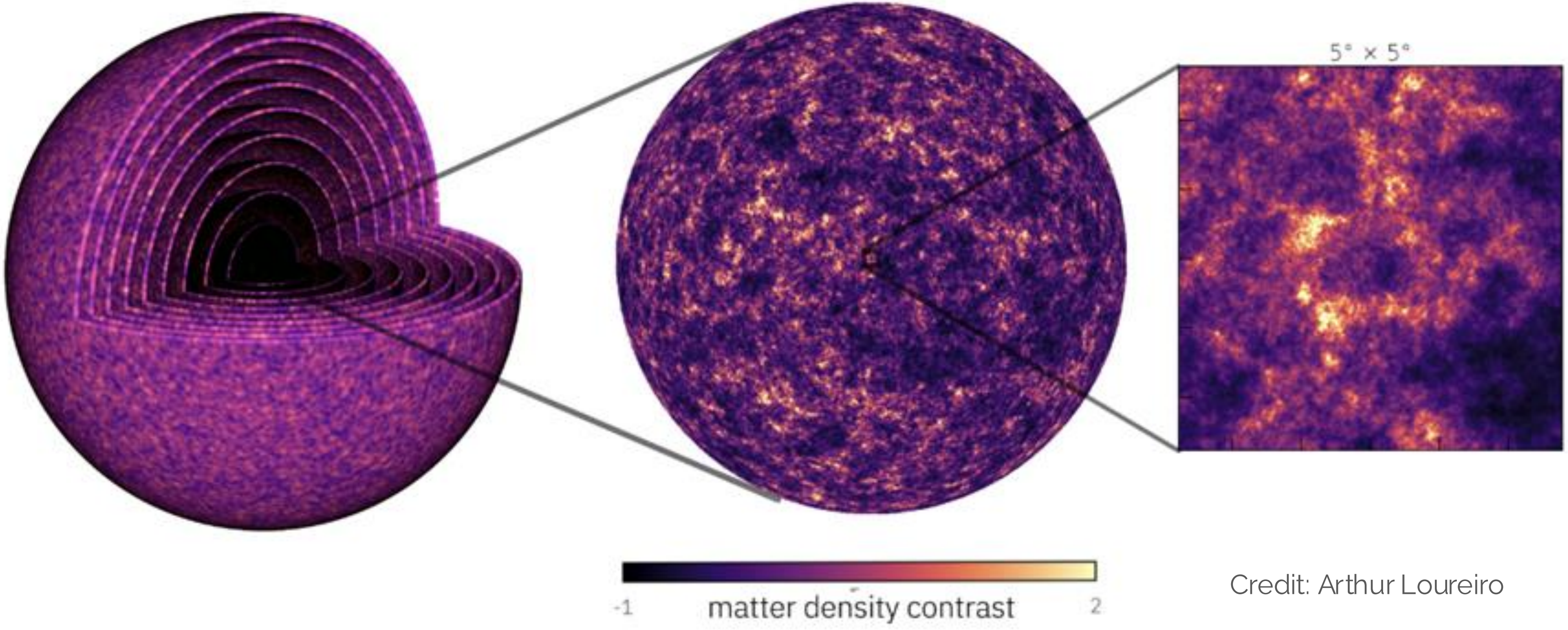


Harnessing the Photometric Uncertainties



Simulating Large-Scale Structure

GLASS: Generator for Large Scale Structure



Credit: Arthur Loureiro

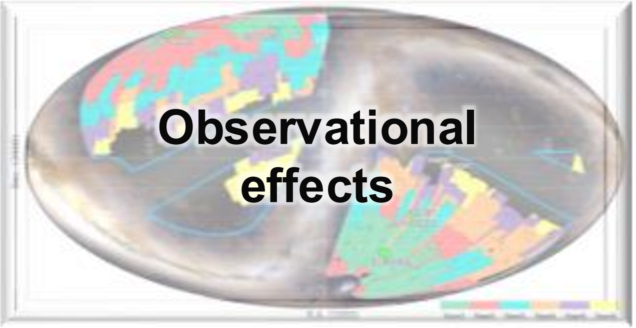
Tessore, et al. (2023), OJA 6 (March).

Cosmic Shear Systematics

Modelled at field level:

PSF Variations

Shear Measurement Bias

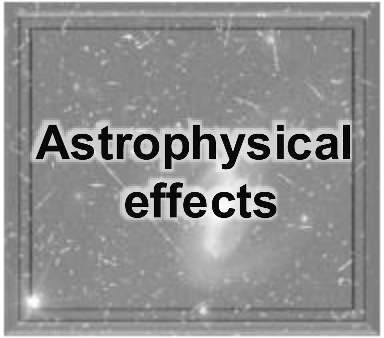


Spatial variability in selection

Redshift uncertainty + variations

Survey mask

Shape & shot noise

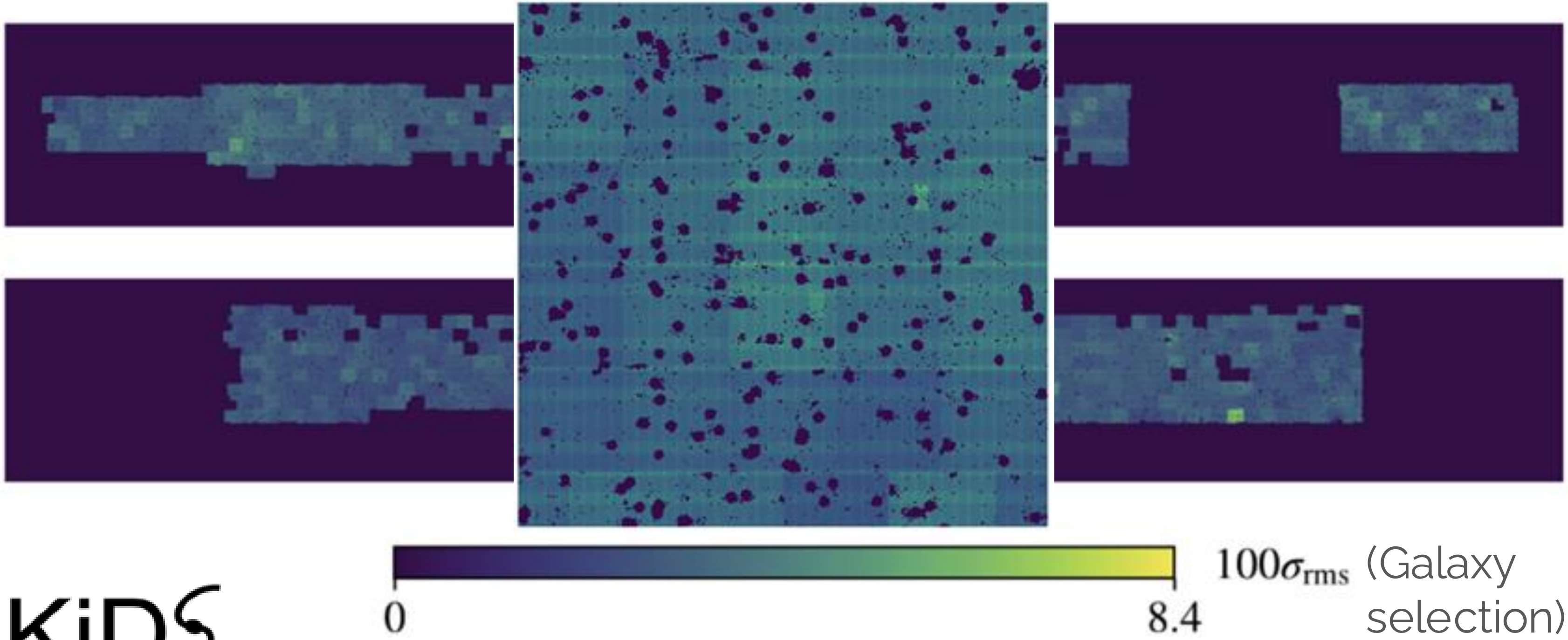


Source clustering

NLA intrinsic alignments

Baryon feedback (*2-point level*)

Realistic Selection and Systematics



(Galaxy selection)

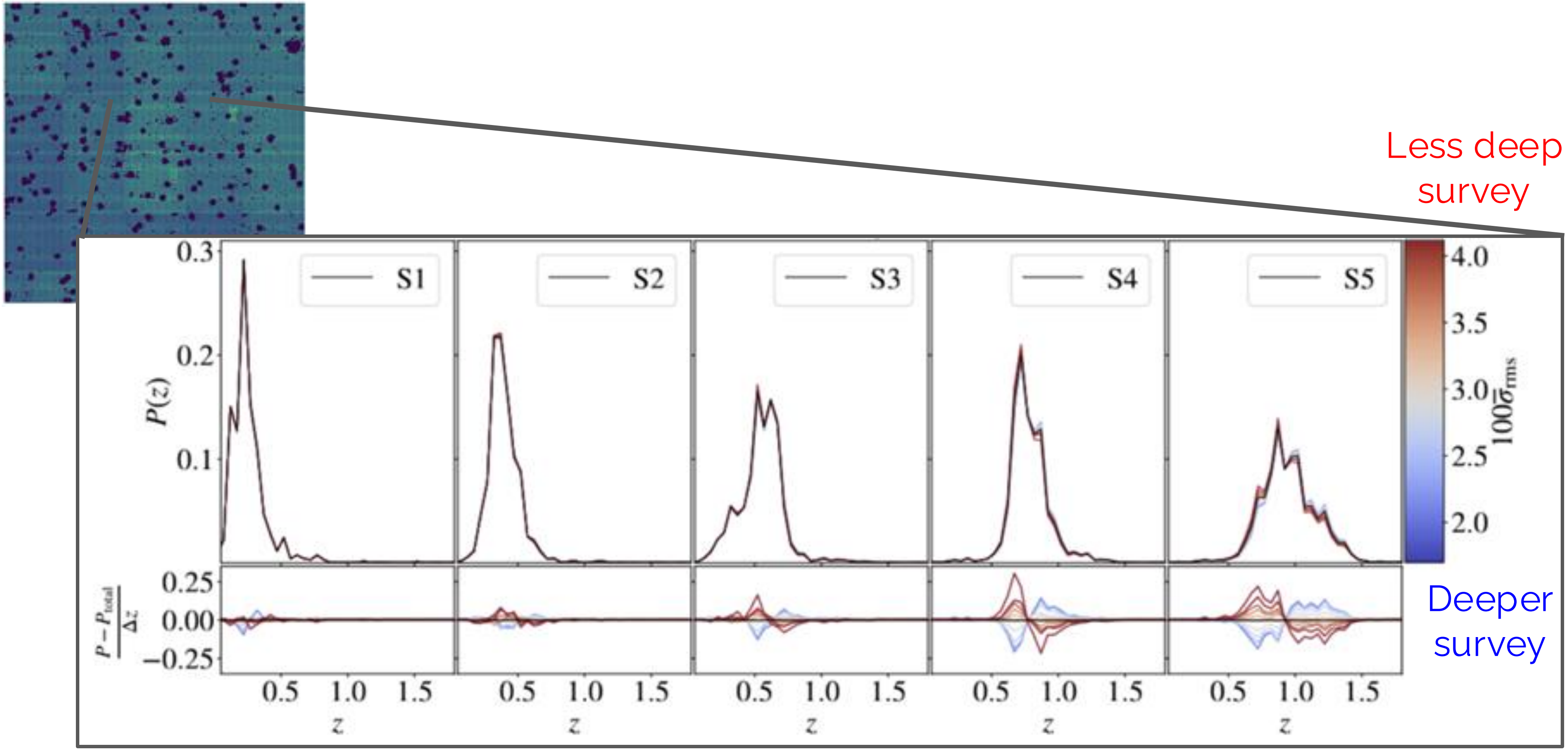
Kilo-Degree Survey

Deeper survey

Less deep survey

von Wietersheim-Kramsta, Lin et al. (2024), A&A 694, A223.

Depth and Galaxy Redshift



Less deep survey

Deeper survey

Sampling Galaxies

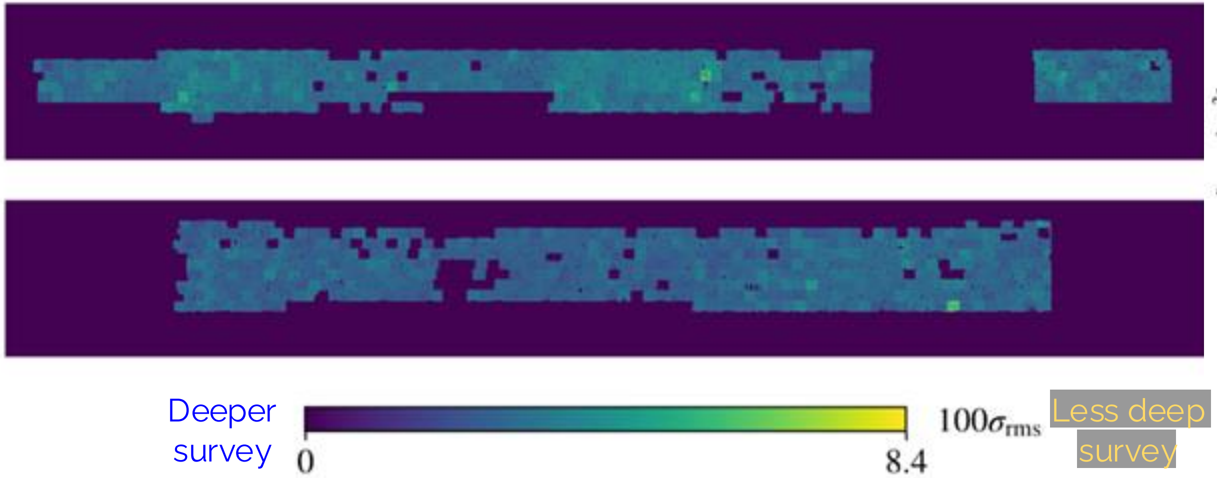
Shell

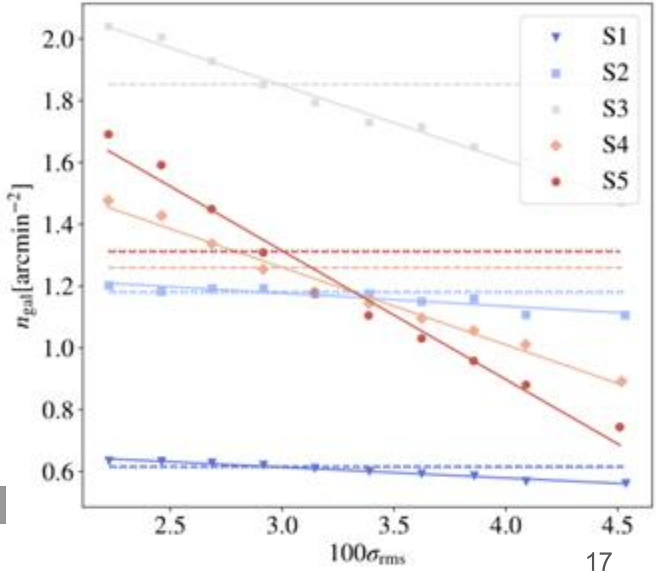
Tomographic bin

$$\langle N_m^{(i)(p)} \rangle(\Theta) = w_m(\Omega_{\rm survey}) \left[1 + b^{(i)} \delta^{(i)}(\theta_m; \Theta)\right] P_m(p|i)\, n^{(p)}_{{\rm gal},m}\, A_{{\rm pix},m}$$

Pixel

Galaxy density (variable depth)





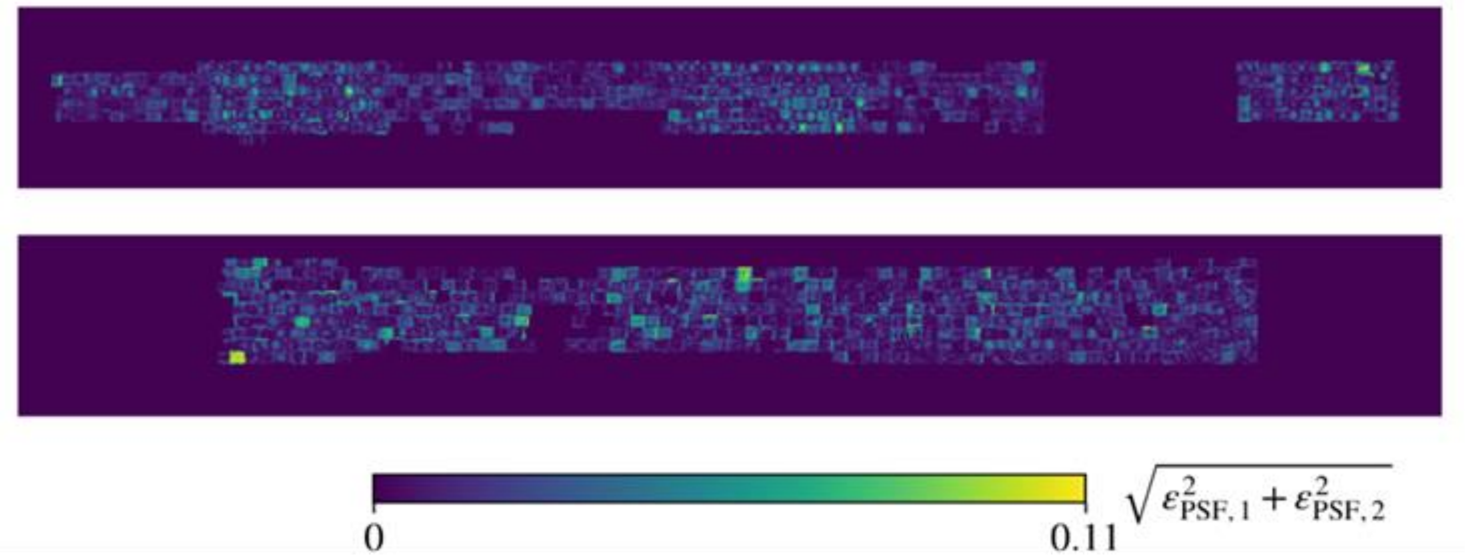
PSF Residuals

Pixel

$$\epsilon_{\rm obs}(p, m; \Theta) = \left(1 + M^{(p)}\right)\epsilon_{\rm lensed}(\Theta) + \alpha^{(p)}\epsilon_{\rm PSF}(m) + \underbrace{\cancel{\beta^{(p)}\delta\epsilon_{\rm PSF}}}_{0} + c^{(p)}$$

Tomographic bin

PSF shear bias



Shear Biases

Pixel

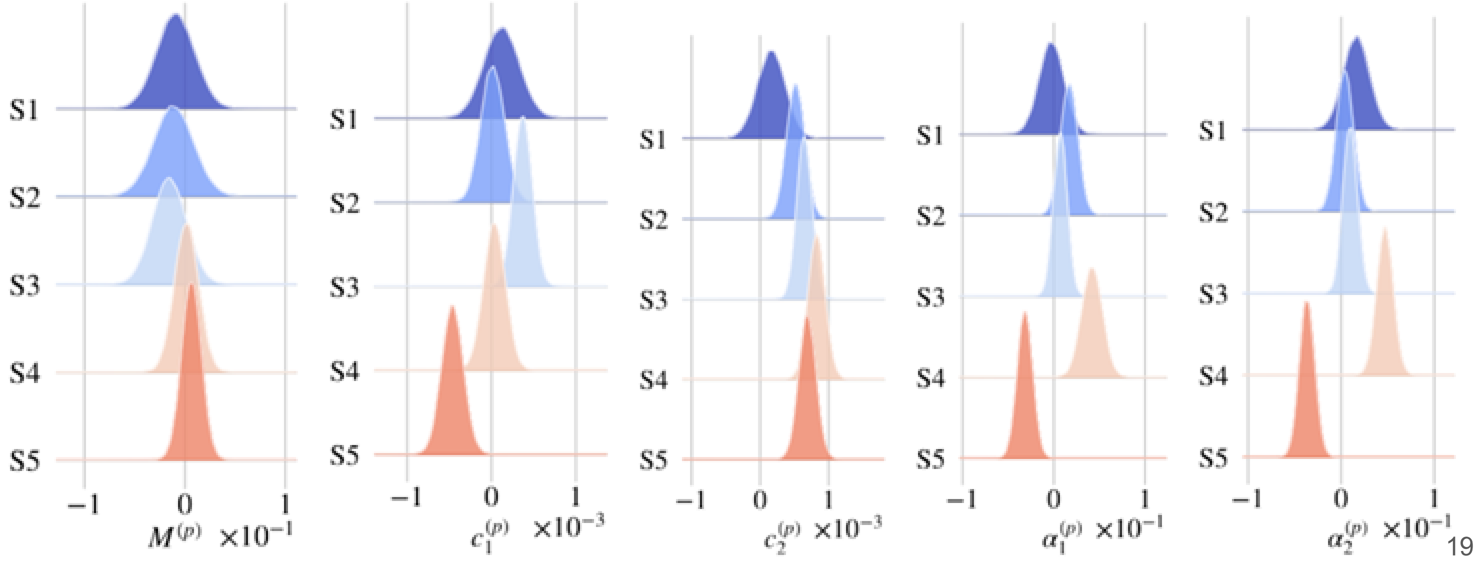
$$\epsilon_{\rm obs}(p, m; \Theta) = \left(1 + M^{(p)}\right)\epsilon_{\rm lensed}(\Theta) + \alpha^{(p)}\epsilon_{\rm PSF}(m) + \underbrace{\cancel{\beta^{(p)}\delta\epsilon_{\rm PSF}}}_{0} + c^{(p)}$$

Tomographic bin

Multiplicative shear bias

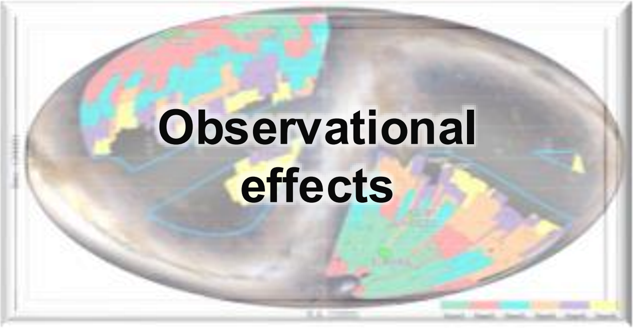
PSF shear bias

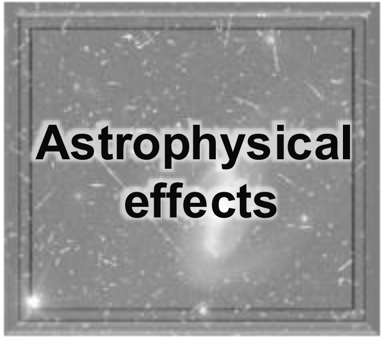
Additive shear bias



Cosmic Shear Systematics

Modelled at field level:





PSF Variations

Shear Measurement Bias

Spatial variability in selection

Redshift uncertainty + variations

Survey mask

Shape & shot noise

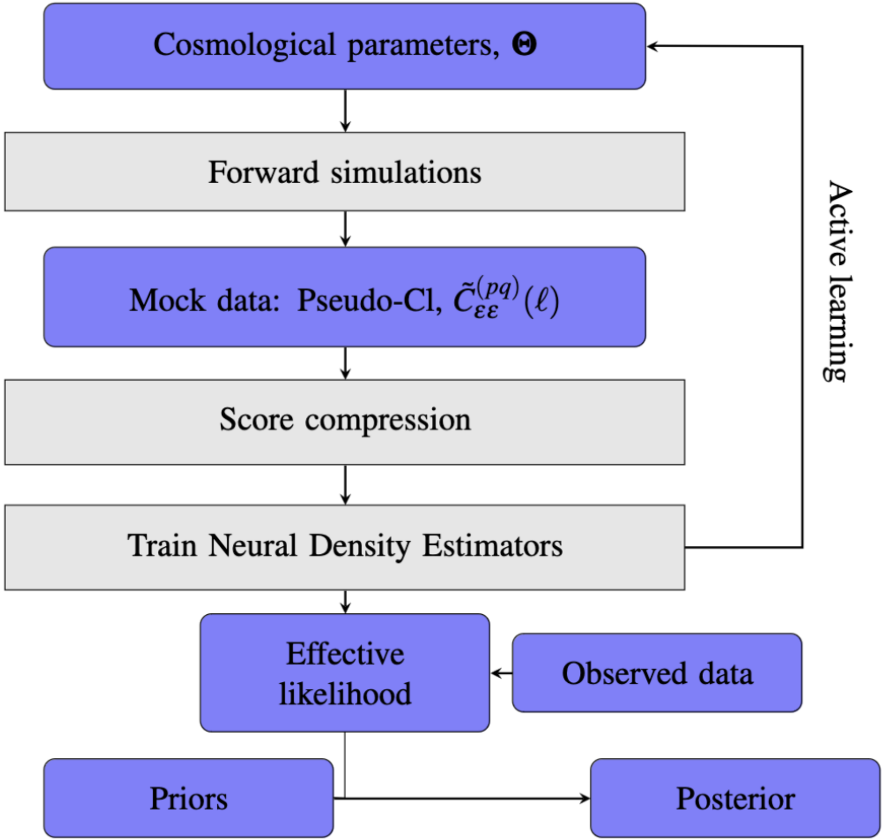
Source clustering

NLA intrinsic alignments

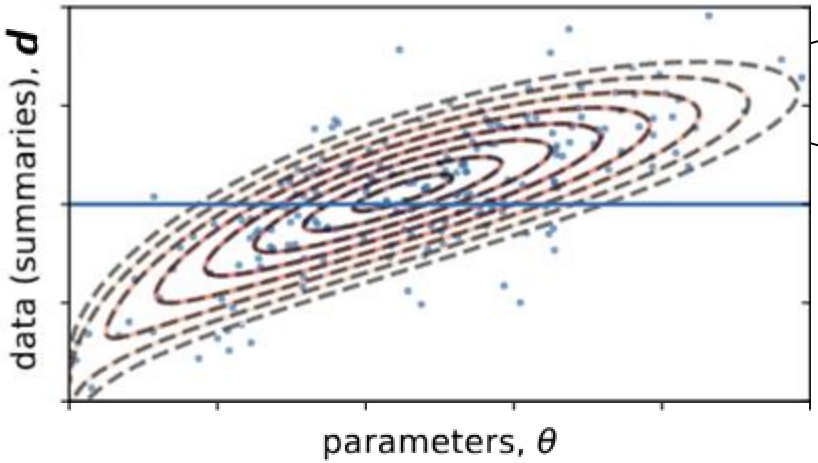
Baryon feedback (*2-point level*)

All correlated across 5 tomographic bins!

SBI: Sequential NDE



SBI: Neural Likelihood Estimation



Varying **32 parameters**
18,000 simulations

Ensemble of 5 NDEs

6 x MAFs

5 x MADEs

2 hidden layers
50 neurons each

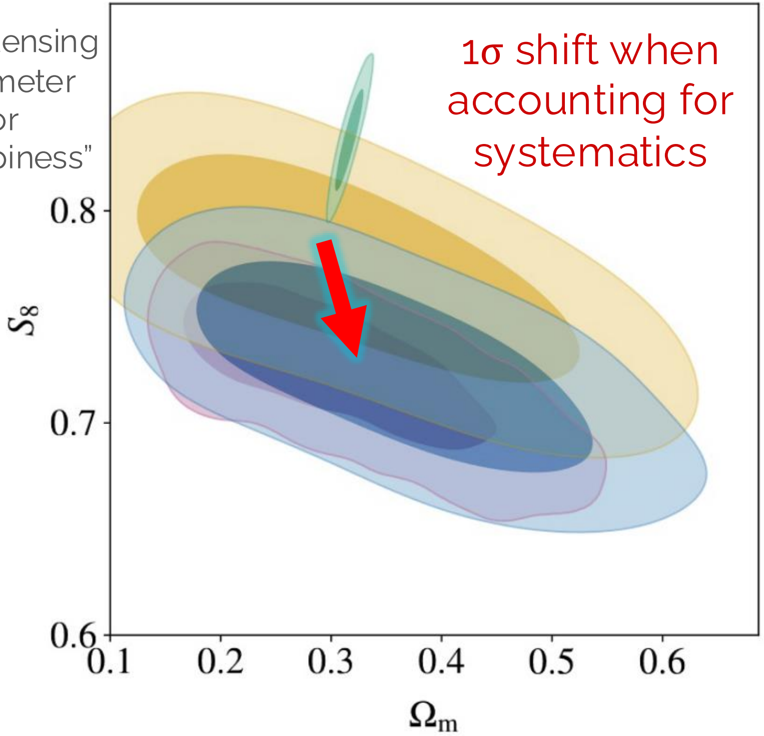
→ 6 x 5 x |**d**| multivariate Gaussians, $P(\mathbf{d}|\boldsymbol{\theta}, \mathbf{w})$

5 cosmological + 7 nuisance + 25 pre-marginalised parameters

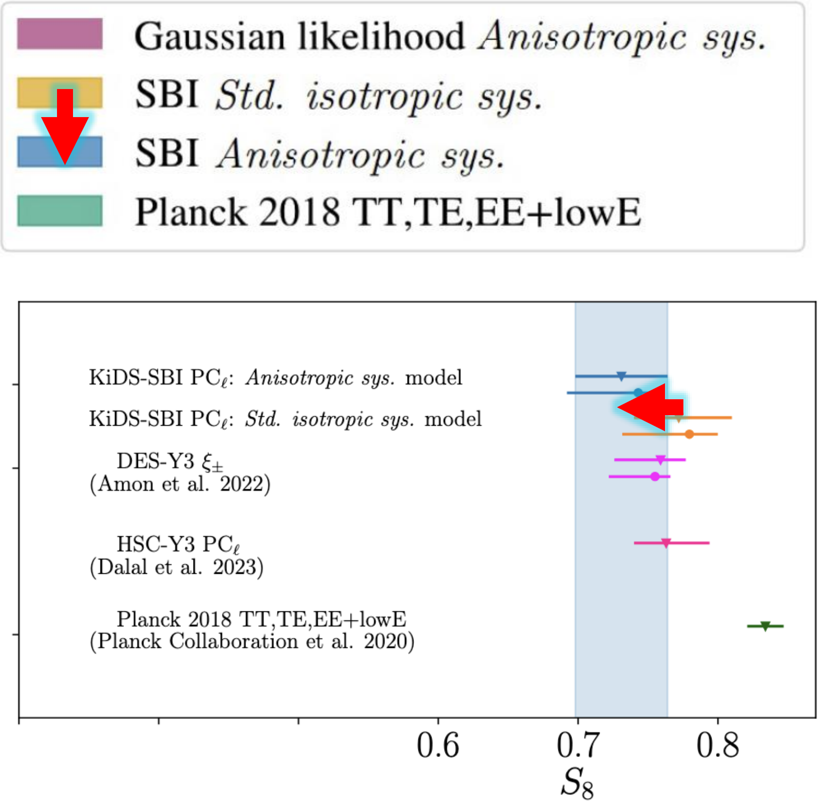
Parameter	Symbol	Prior type	Prior range	Fiducial
Density fluctuation amp.	S_8	Flat	[0.1, 1.3]	0.76
Hubble constant	h_0	Flat	[0.64, 0.82]	0.767
Cold dark matter density	ω_c	Flat	[0.051, 0.255]	0.118
Baryonic matter density	ω_b	Flat	[0.019, 0.026]	0.026
Scalar spectral index	n_s	Flat	[0.84, 1.1]	0.901
Intrinsic alignment amp.	A_{IA}	Flat	[-6, 6]	0.264
Baryon feedback amp.	A_{bary}	Flat	[2, 3.13]	3.1
Redshift displacement	$\boldsymbol{\delta}_z$	Gaussian	$\mathcal{N}(\mathbf{0}, \boldsymbol{C}_z)$	$\mathbf{0}$
Multiplicative shear bias	$M^{(p)}$	Gaussian	$\mathcal{N}(\overline{M}^{(p)}, \sigma_M^{(p)})$	$\overline{M}^{(p)}$
Additive shear bias	$c_{1,2}^{(p)}$	Gaussian	$\mathcal{N}(\overline{c}_{1,2}^{(p)}, \sigma_{c_{1,2}}^{(p)})$	$\overline{c}_{1,2}^{(p)}$
PSF variation shear bias	$\alpha_{1,2}^{(p)}$	Gaussian	$\mathcal{N}(\overline{\alpha}_{1,2}^{(p)}, \sigma_{\alpha_{1,2}}^{(p)})$	$\overline{\alpha}_{1,2}^{(p)}$

SBI in Cosmic Shear

Weak lensing parameter for "clumpiness"

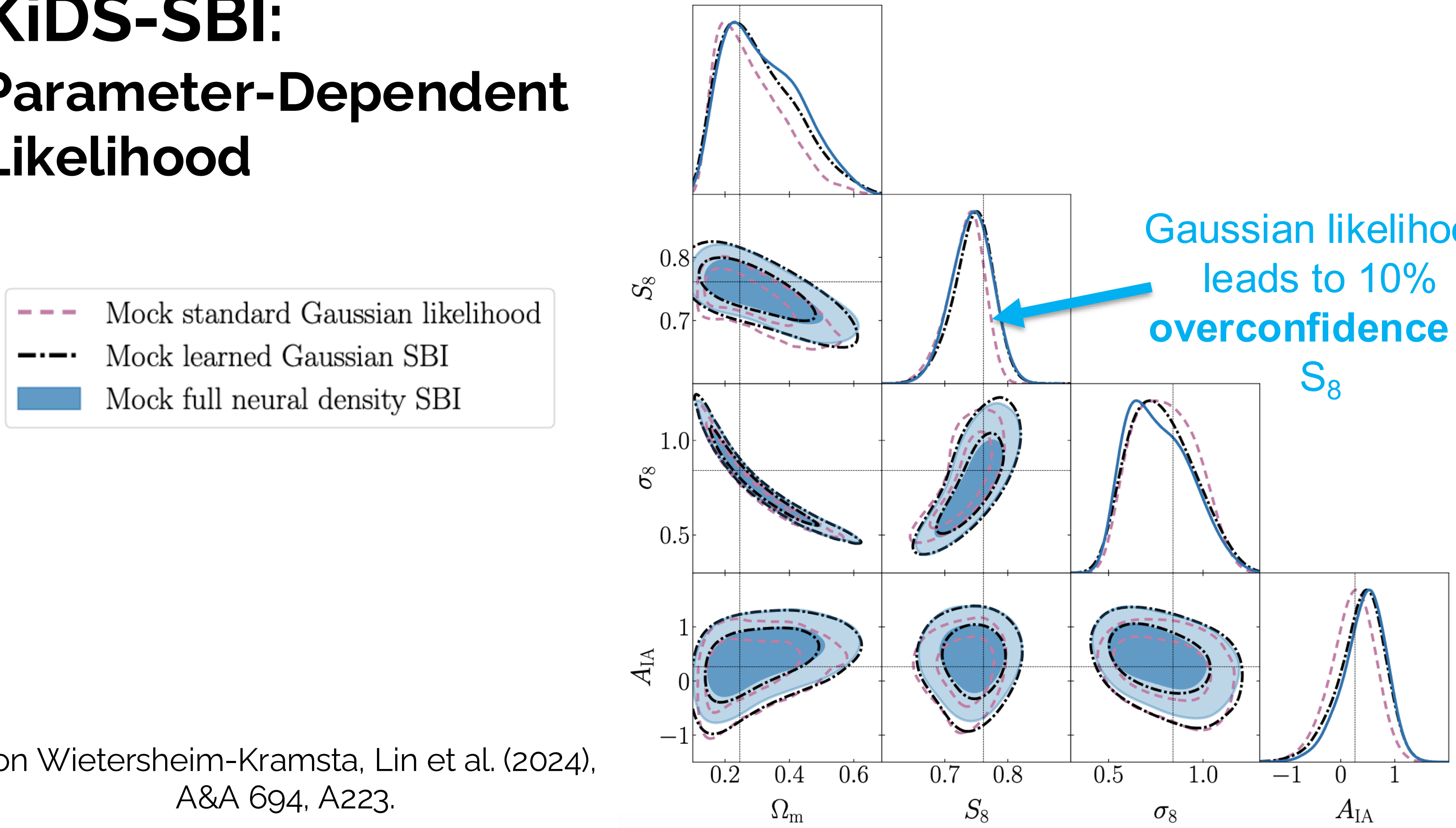


KiDS-1000 cosmic shear only



von Wietersheim-Kramsta, Lin et al. (2024), A&A 694, A223.

KiDS-SBI:
Parameter-Dependent Likelihood

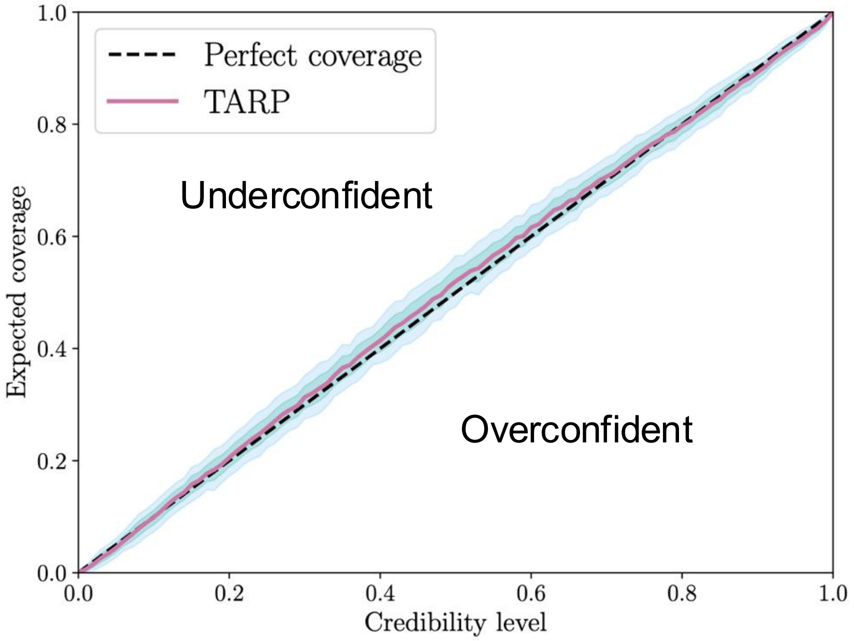


Gaussian likelihood leads to 10% **overconfidence** in S_8

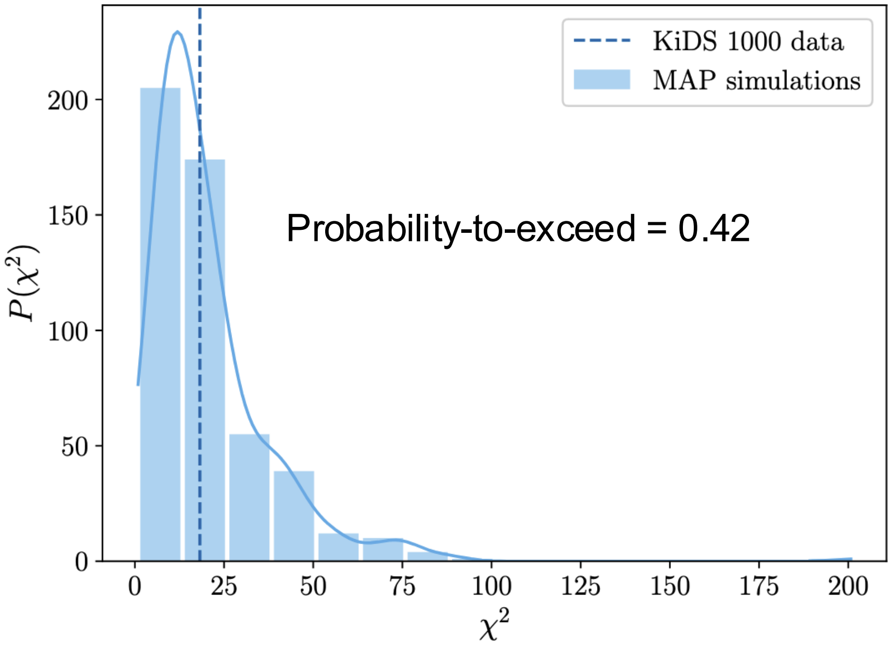
von Wietersheim-Kramsta, Lin et al. (2024), A&A 694, A223.

SBI: Accuracy Testing

Agreement between learnt posterior & forward model



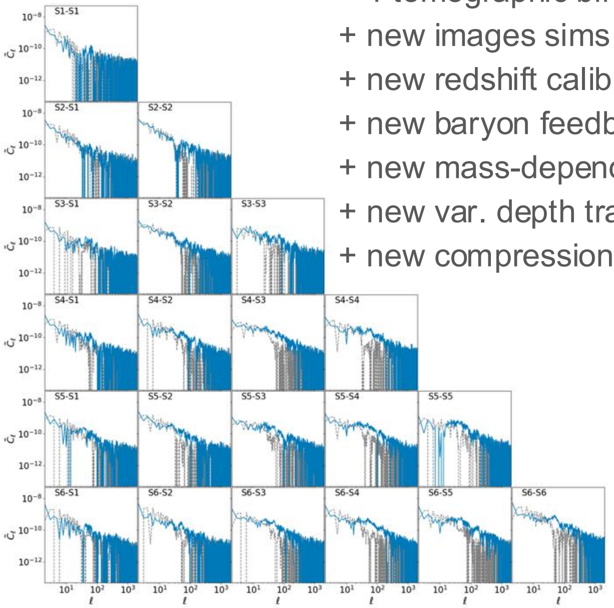
How likely is the real data given the model?



Extensions to KiDS-SBI

KiDS-SBI with KiDS-Legacy

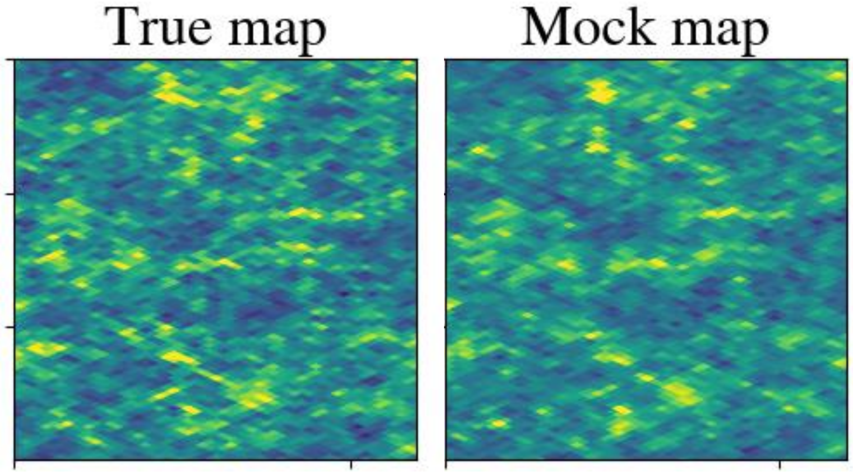
+ extra 350 deg^2
+ 1 extra i-band pass
+ 1 tomographic bin
+ new images sims for calib.
+ new redshift calibration
+ new baryon feedback model
+ new mass-dependent IAs model
+ new var. depth tracer
+ new compression



Lin, von Wietersheim-Kramsta, et al. (in prep.)

3x2pt analysis (shear x clustering)

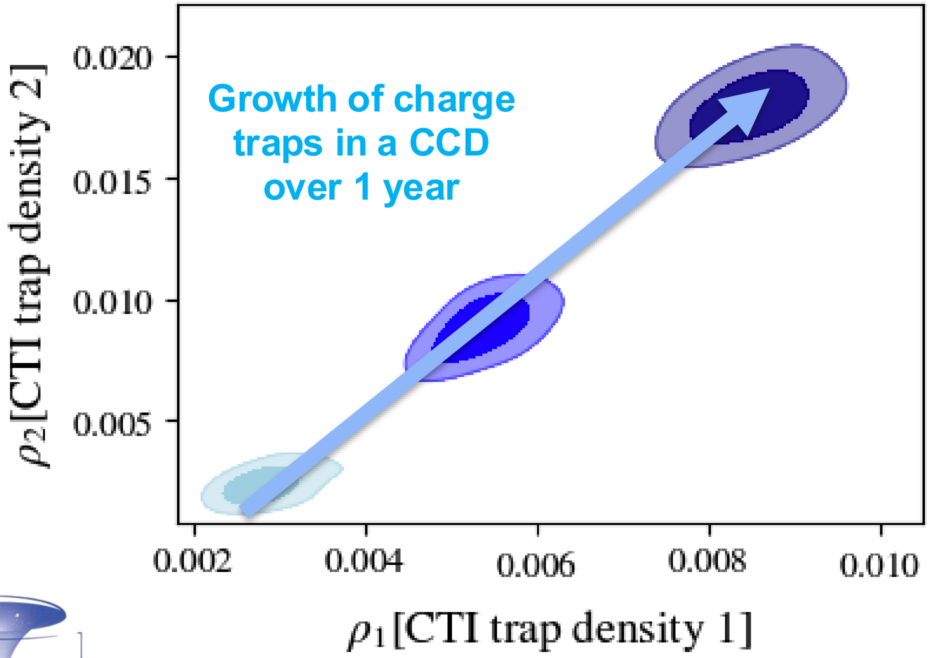
Forward simulating field-level galaxy bias on the sphere

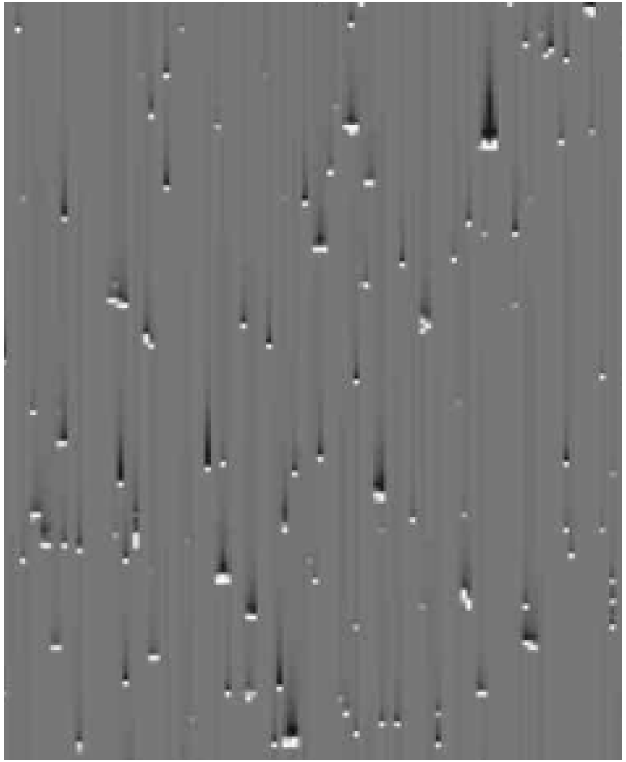


von Wietersheim-Kramsta, et al. (in prep.)

Considerations for Stage IV

e.g. Radiation Damage



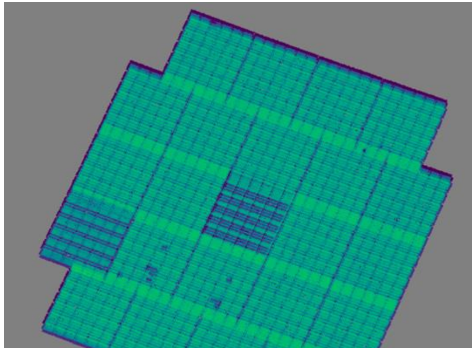


Massey et al. (2025)
McCracken et al. (2025)

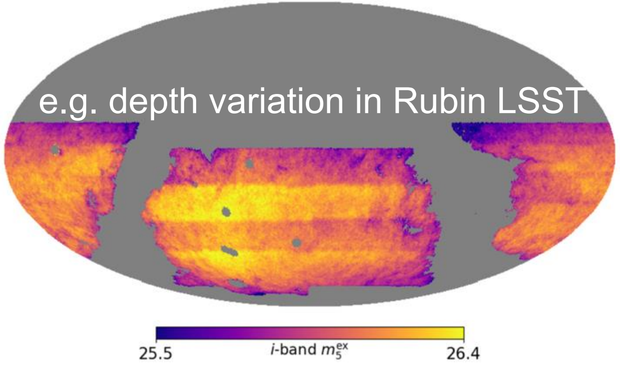
Considerations for Stage IV

e.g. Variable Depth

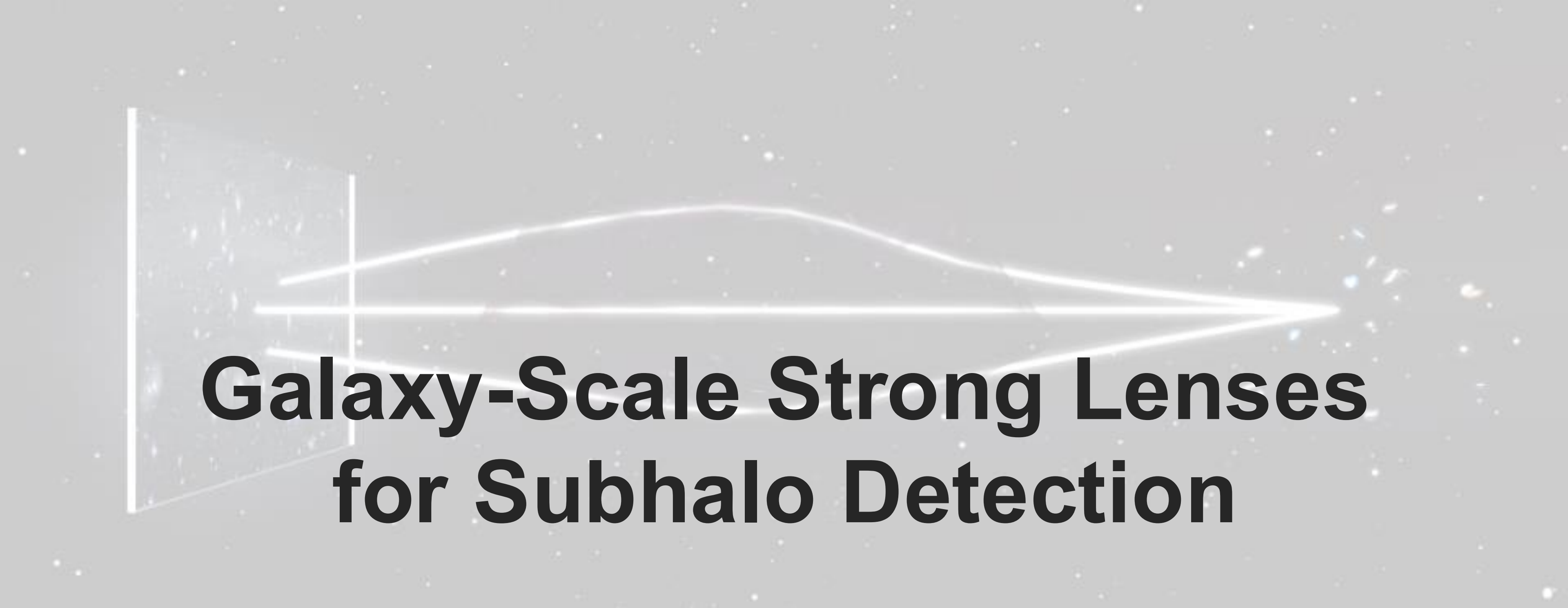
Angular Variation	vs.	Line-of-Sight Variation
• Shapes measured by VIS • Single-visit survey • Space-based		• Photometry by partner surveys • (Some) multi-epoch surveys • (Some) ground-based



Euclid Collaboration 2024 (Q1 release)



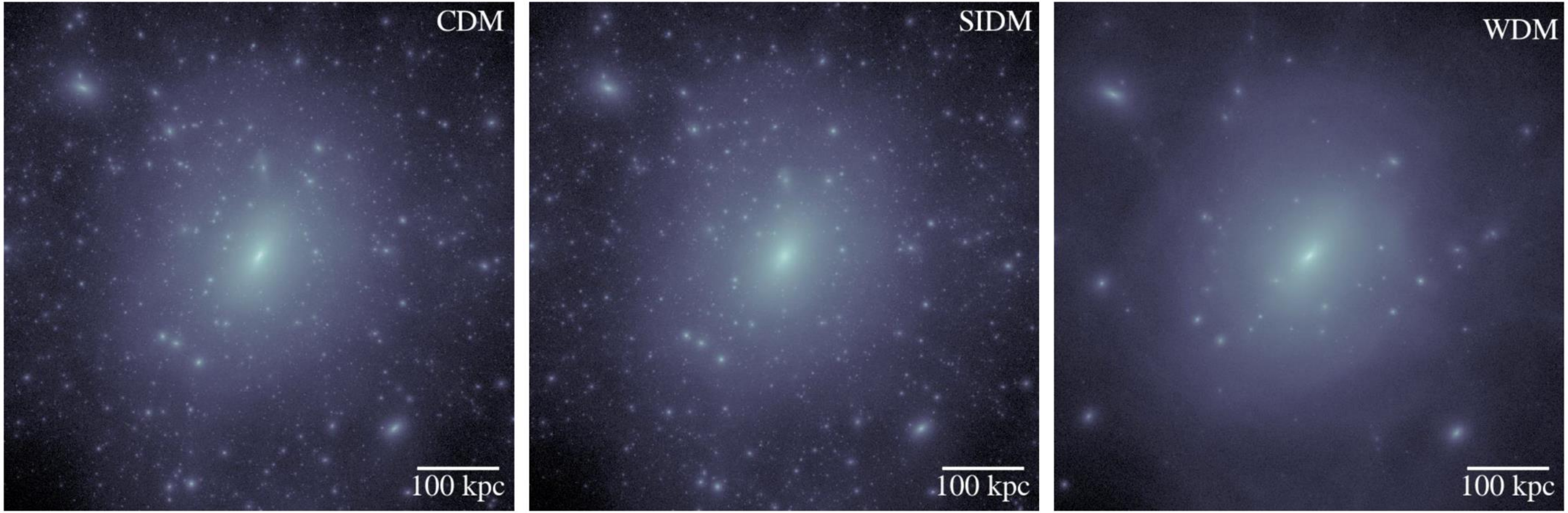
Hang et al. 2024



Galaxy-Scale Strong Lenses for Subhalo Detection

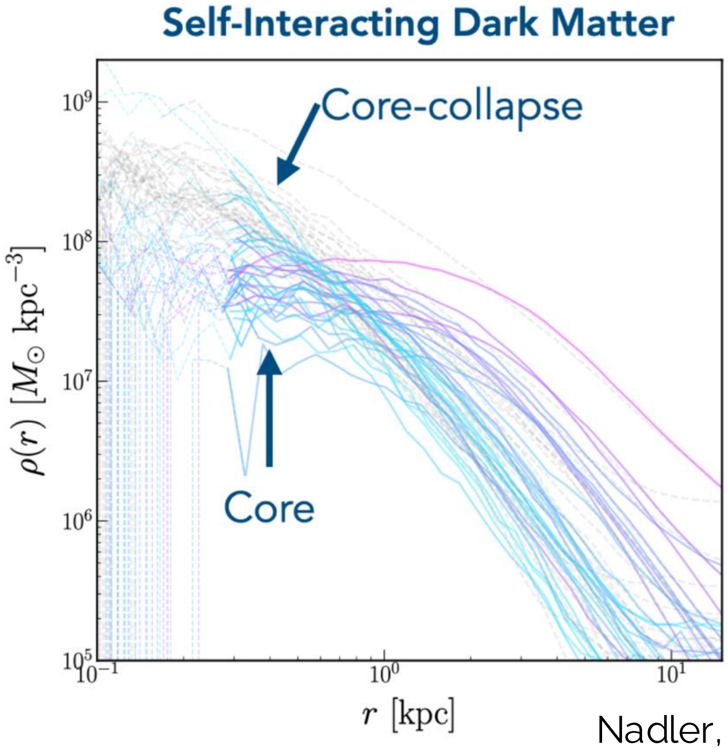
In collaboration with R. Massey, Q. He, J. Nightingale, L. Makinen, A. Robertson, A. Amvrosiadis, L. Fung, S. Lange, C. Frenk, S. Cole, R. Li, et al.

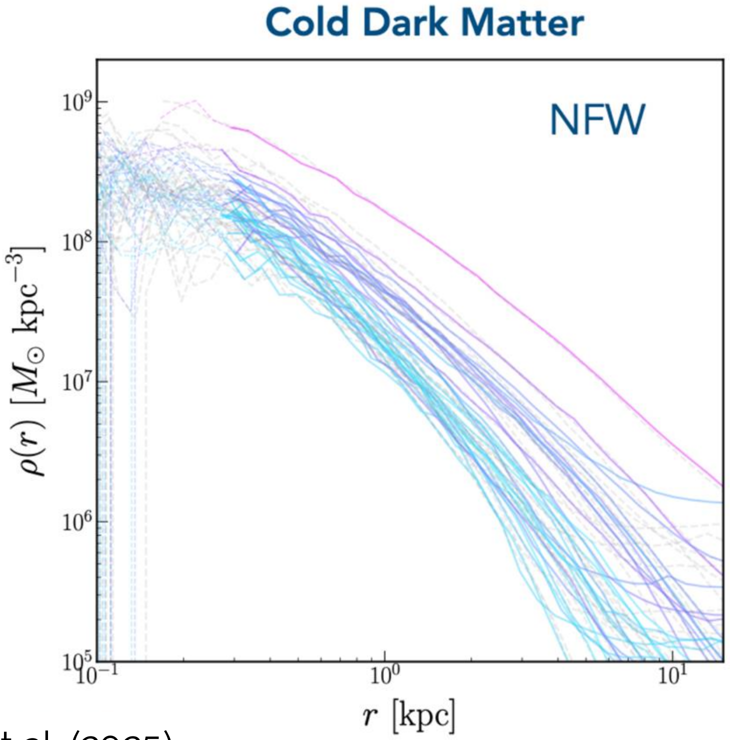
Substructure & the Nature of Dark Matter



Bullock & Boylan-Kolchin (2017), ARA&A, 55:343-387.

Substructure & the Nature of Dark Matter

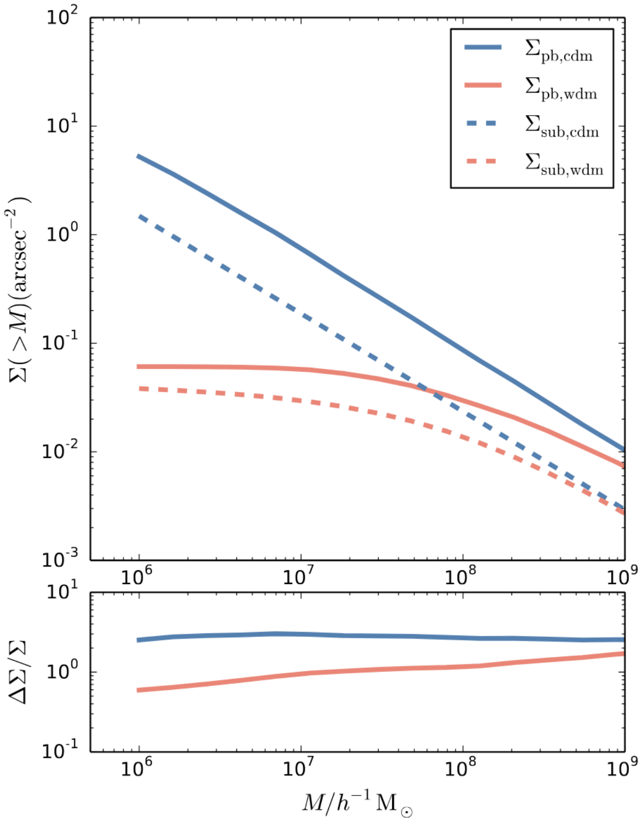




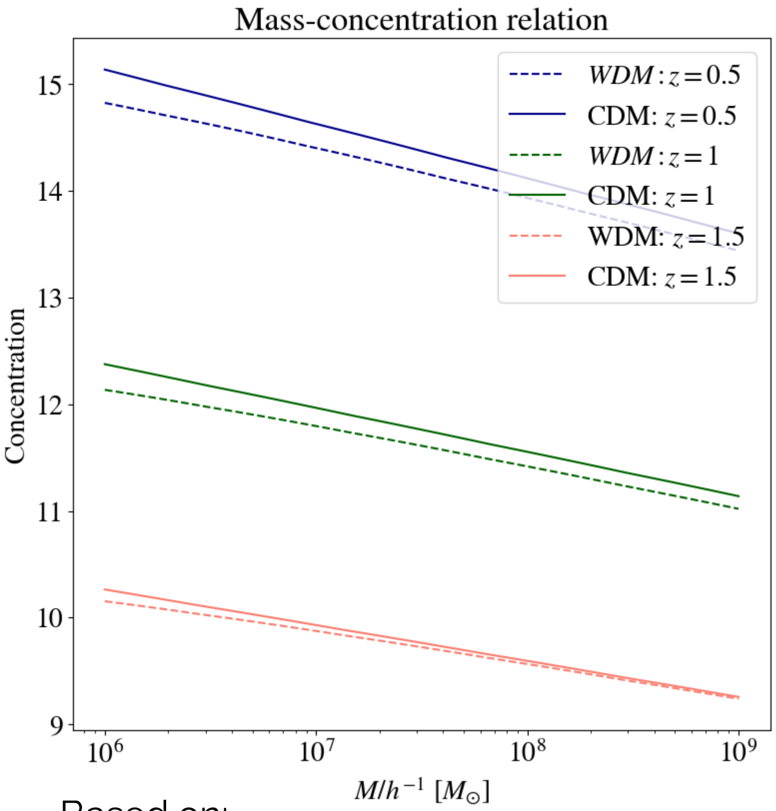
Nadler, E. O., et al. (2025), *arXiv:2503.10748*.

Forward Modelling: Substructure

Number densities of **perturbing interlopers and subhaloes**



Li et al. (2016), MNRAS, 468(2), 1426-1432.



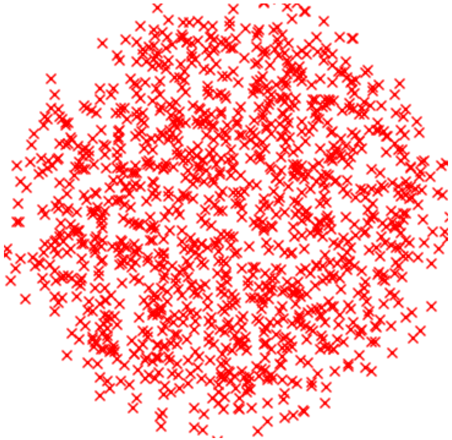
Based on:
Ludlow et al. (2016), MNRAS, 460(2), 1214-1232.

Forward Modelling: Realism

Source:
- Elliptical Core-Sersic
- $z = 1$
- **Axis ratio $\in [0.3, 0.85)$**
- **Axial tilt $\in [30, 70]°$**

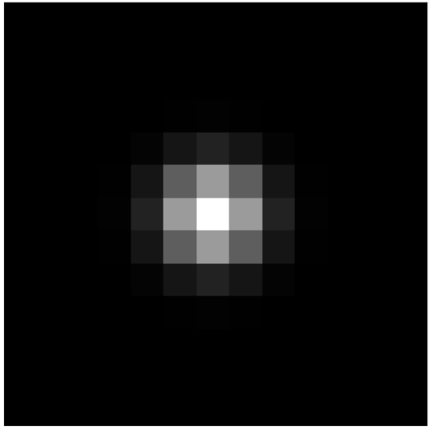
Lens:
- Power law mass
- $z = 0.5$
- No external shear
- **$R_E \in [1.0, 1.5]''$**

Perturbers:
- Warm Dark Matter
- Truncated NFW mass
- $M_{hf} = 10^7$
- **$n_{subhalos} \in [0, 30]$**
- + Perturbing interlopers

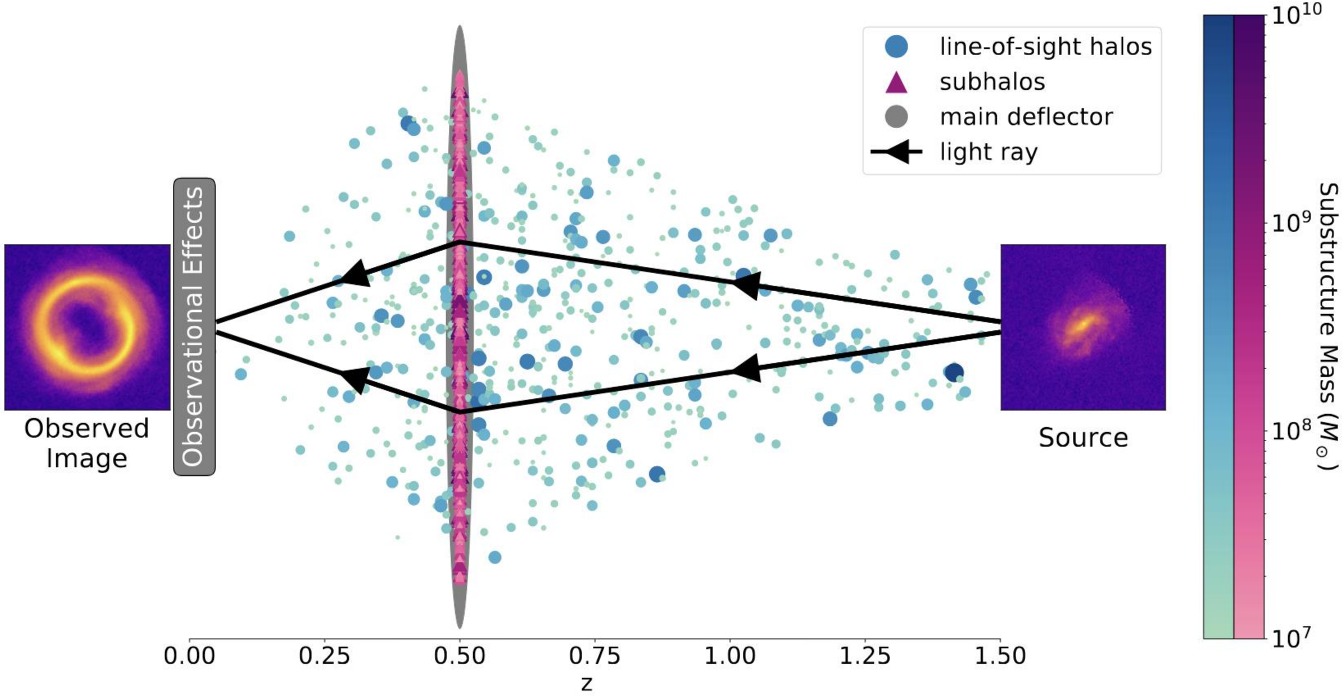


He et al. (2022), MNRAS, 511(2), 3046-3062.

Observational Effects
(HST-like)
- Exposure = 8000s
- Sky background = 0.1
- Pixel scale = 0.05”
- σ_{PSF} = 0.05”
- + Poisson noise



Forward Modelling

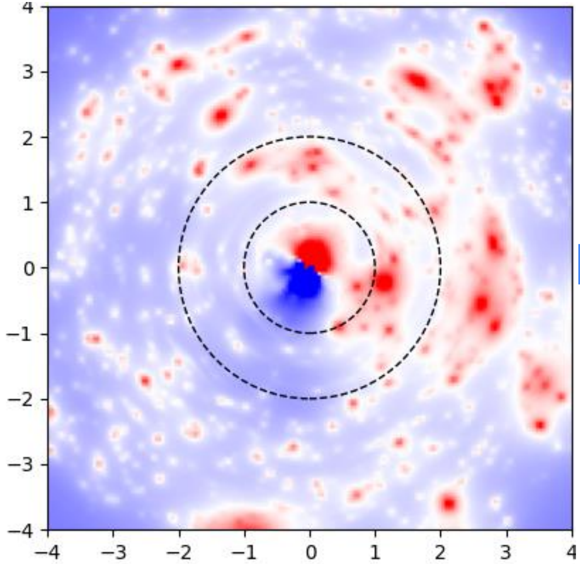


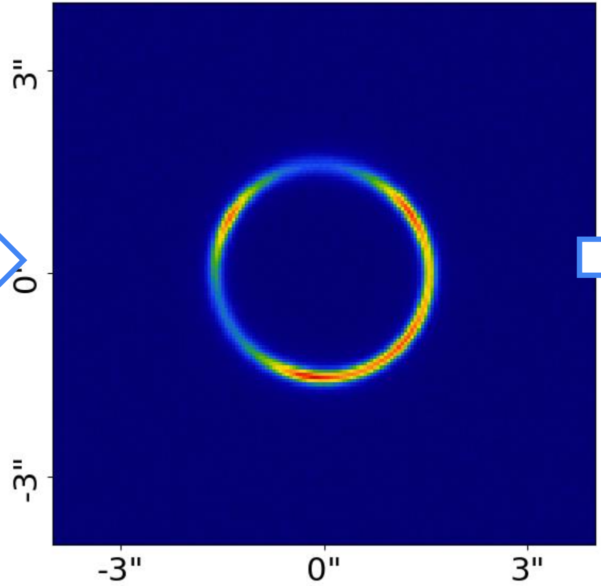
Wagner-Carena et al. (2023), AJ, *942*(2), 75.

Forward Modelling

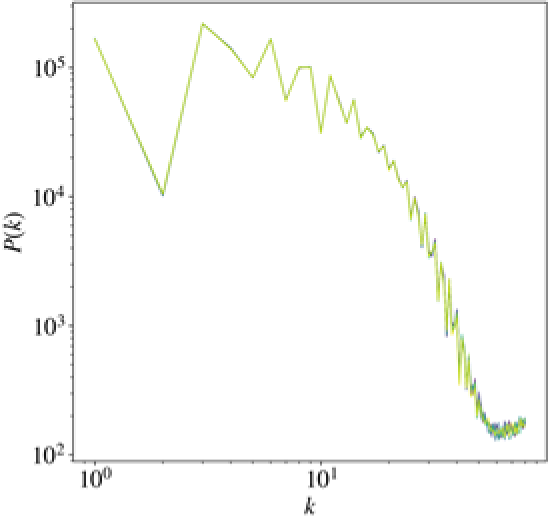
AutoLens: Mock Observation

Nightingale et al., (2021), JOSS, 6(58), 2825





Compression/summary statistic: P(k) + CNN



von Wietersheim-Kramsta, et al. (in prep.)

RINSE & REPEAT 1000 TIMES!

SBI: Neural Posterior Estimation

Ensemble of 2 NDEs

2 x MAFs

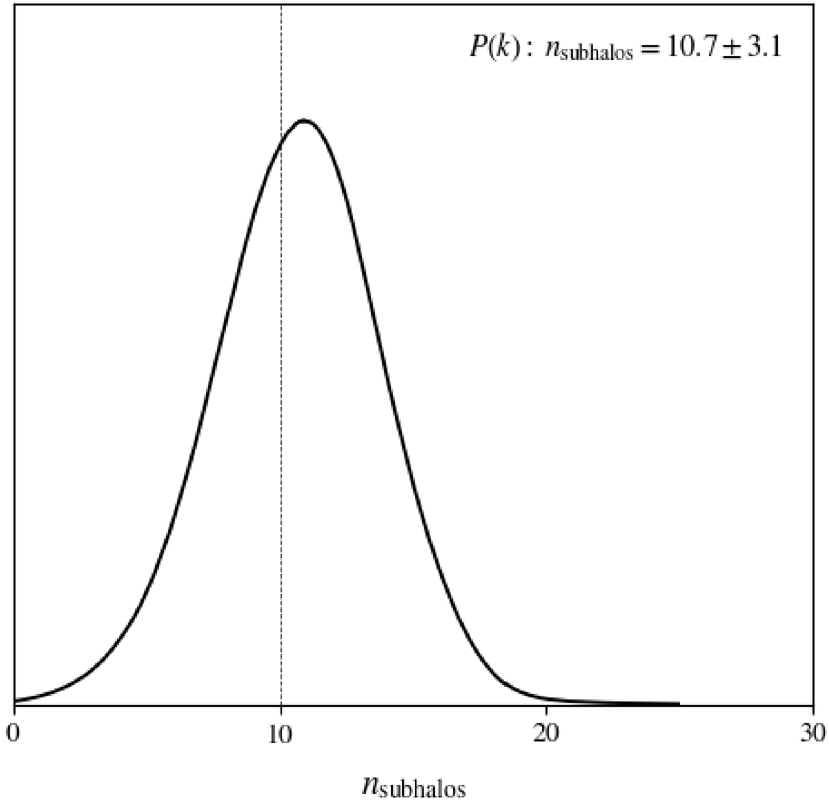
5 x MADEs

2 hidden layers

50 neurons each

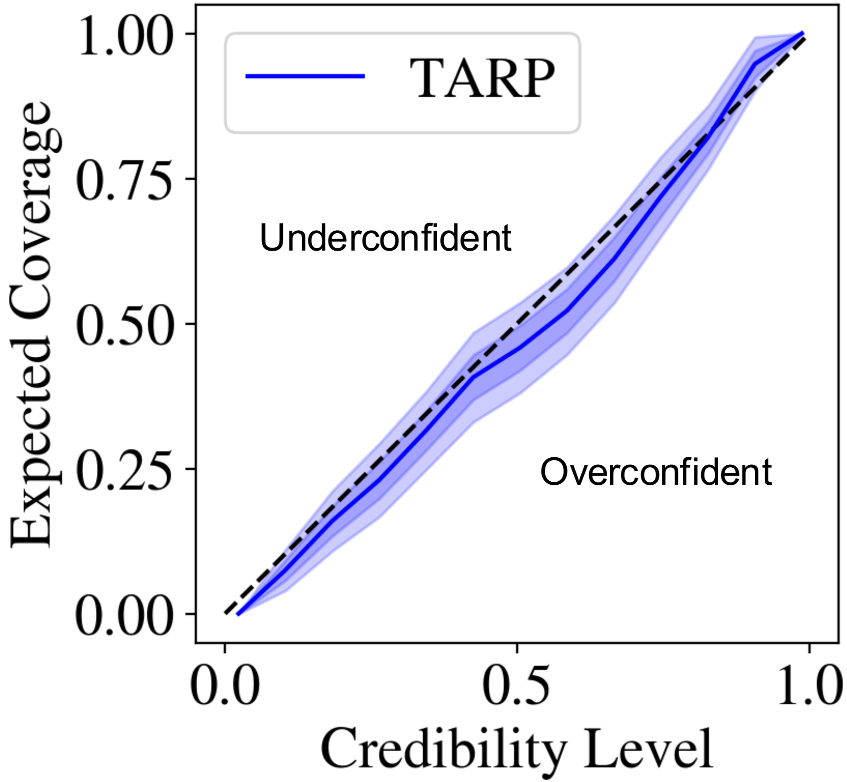
Varying **1 parameter**

1,000 simulations



von Wietersheim-Kramsta, et al. (in prep.)

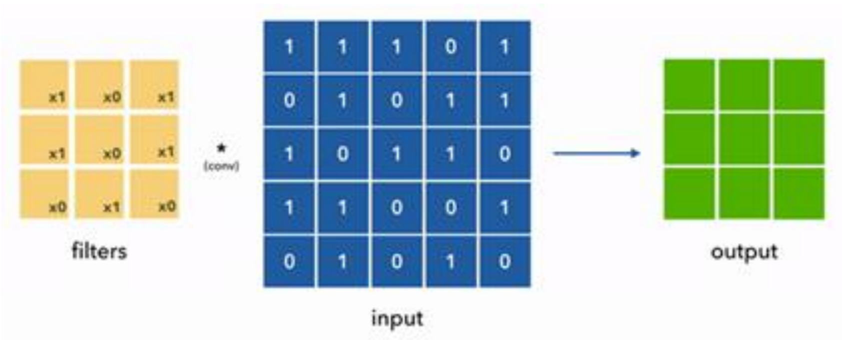
SBI: Coverage



SBI: Other Compressions

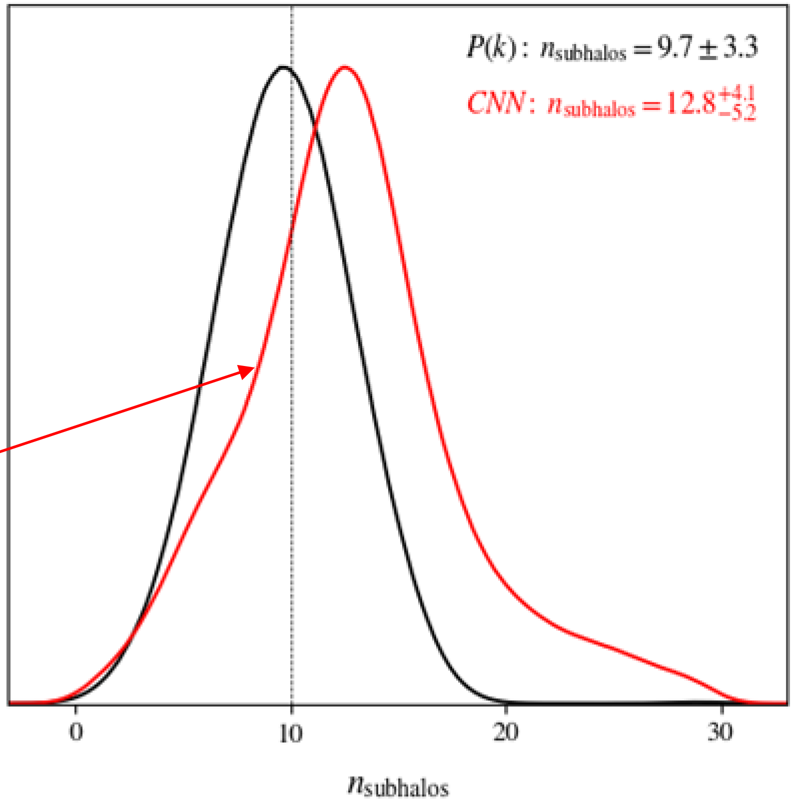
Other compression schemes/summary statistics:

Convolutional Neural Networks



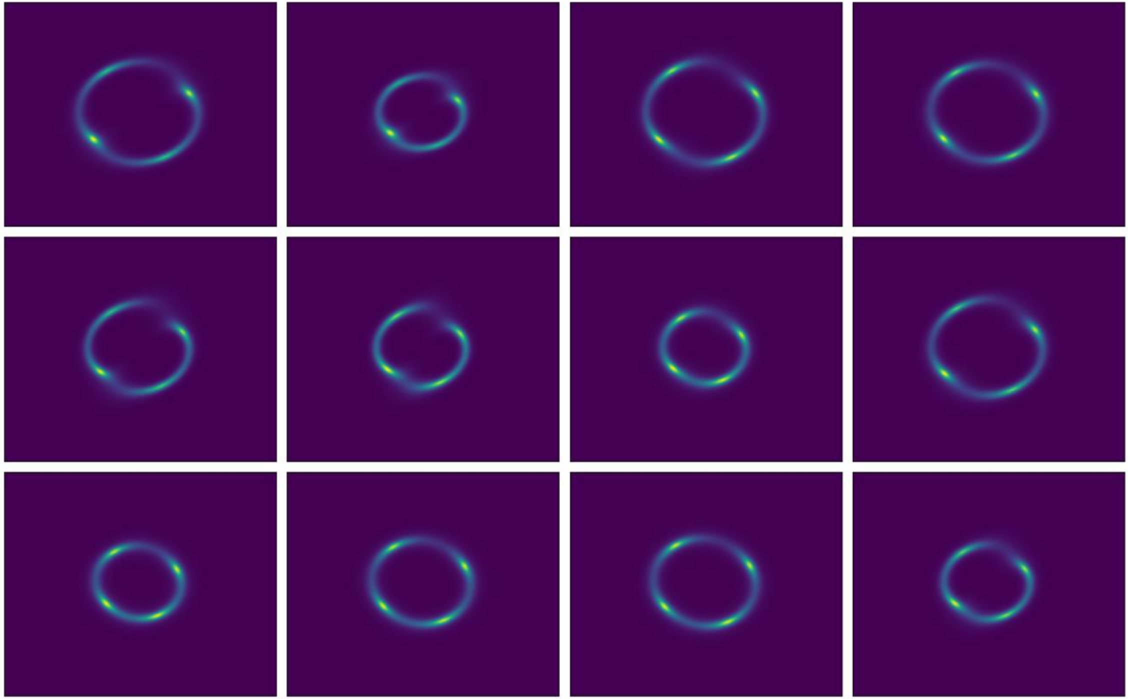
Learn weights based on all simulated images

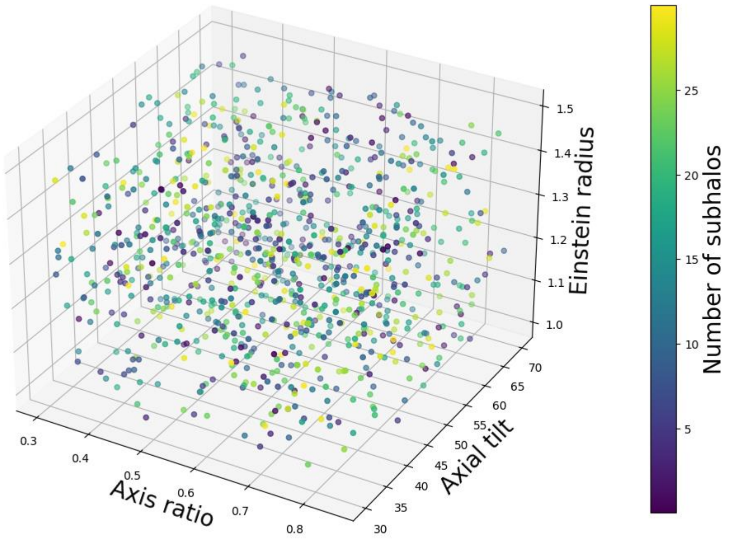
Convolutional layer



von Wietersheim-Kramsta, et al. (in prep.)

SBI: Other Compressions

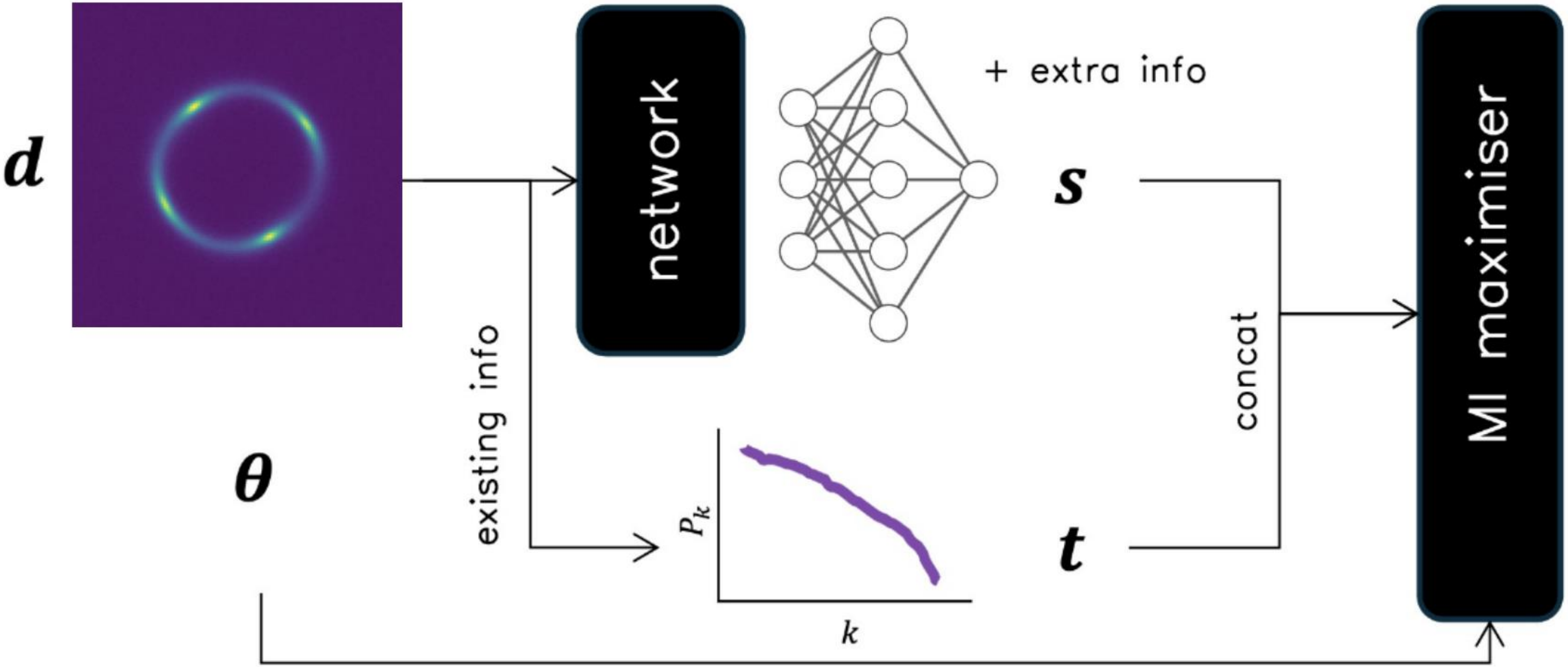




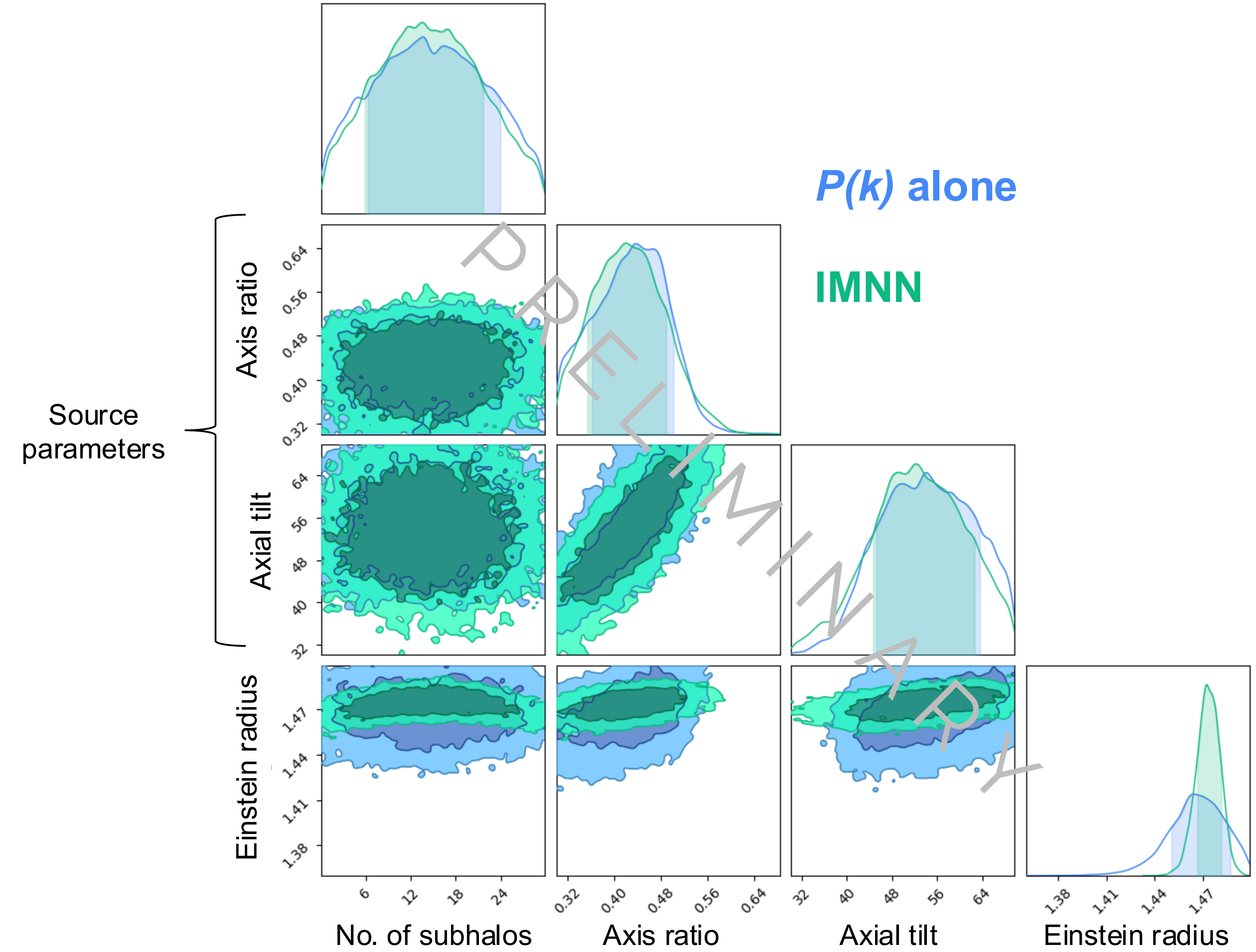
Varying **4 parameters**

von Wietersheim-Kramsta, et al. (in prep.)

SBI: IMNN Compression

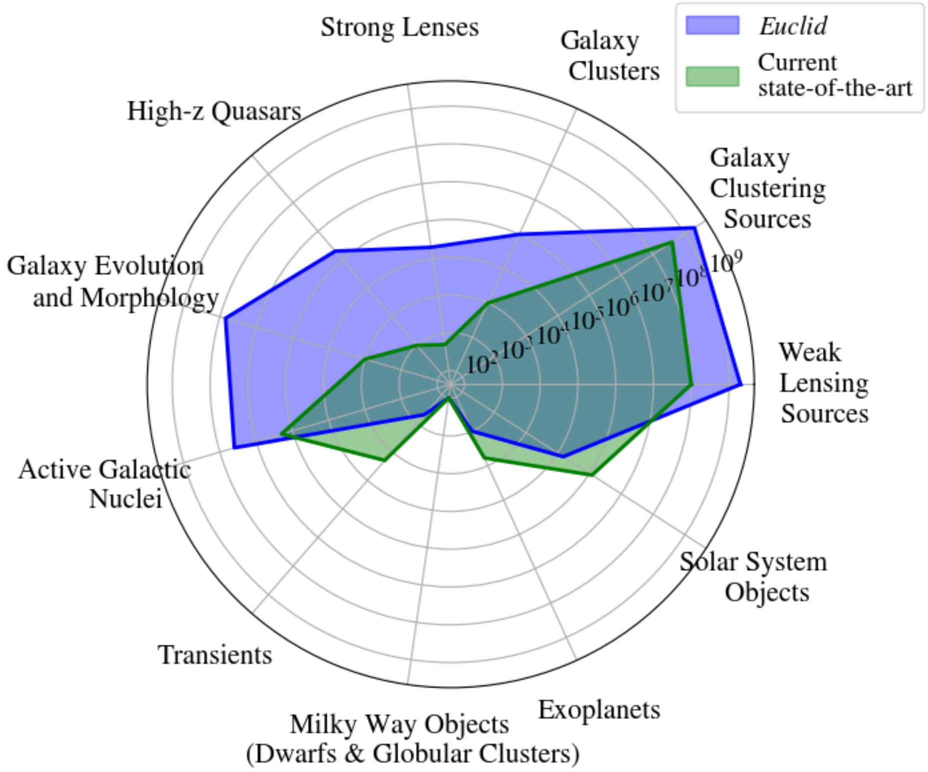


Makinen, T. L., et al. (2025), JCAP, *2025*(01), 095.



von Wietersheim-Kramsta, et al. (in prep.)

Future Considerations (Stage IV)

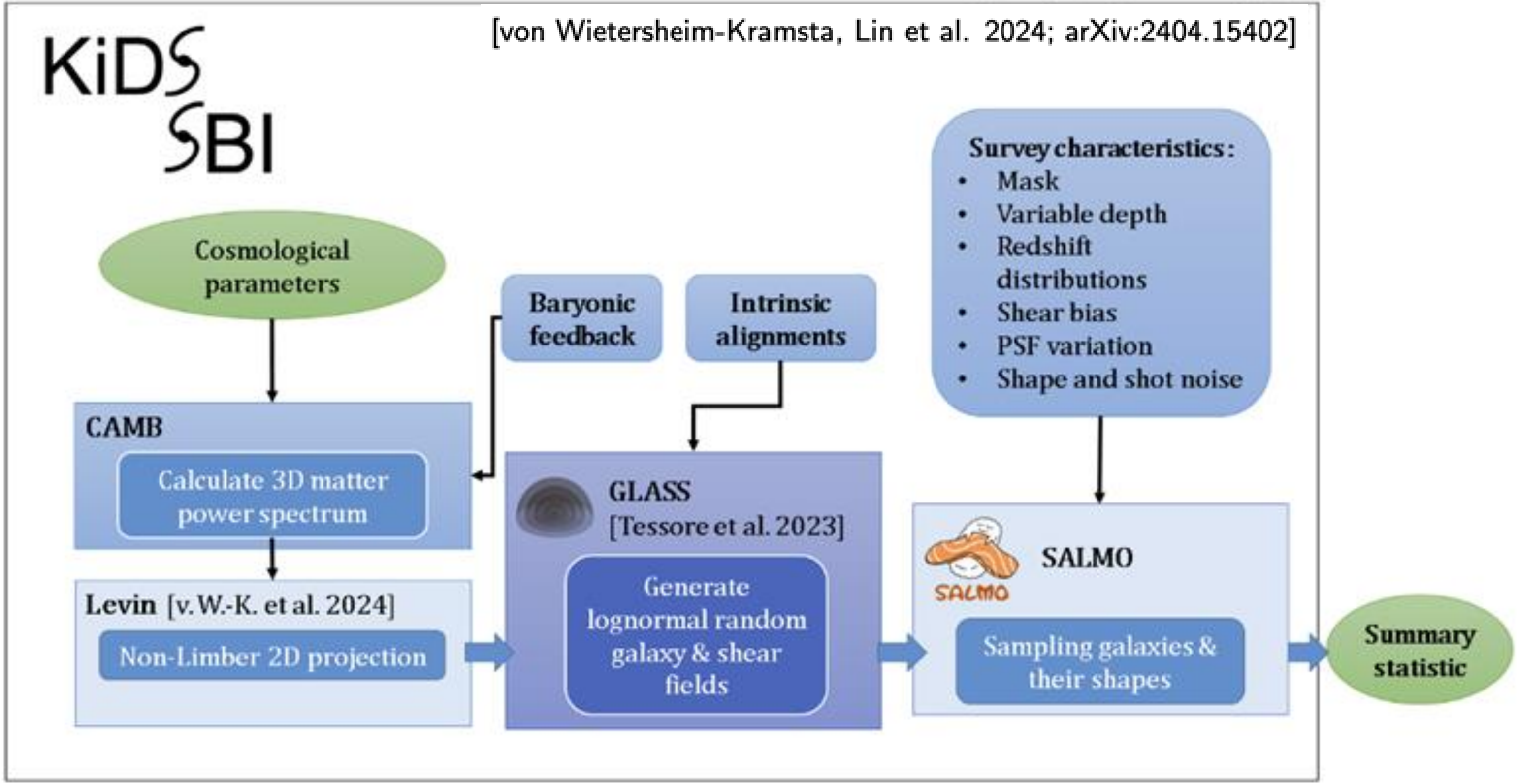


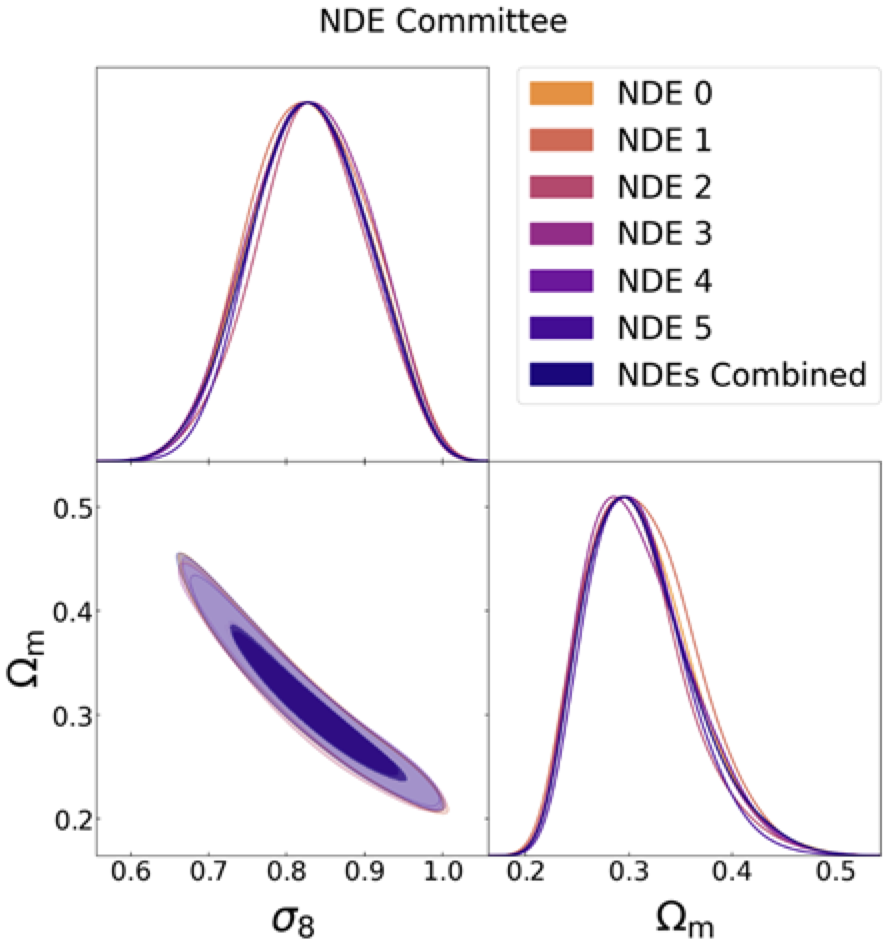
10^3 → 2×10^5
strong lenses in 5 years

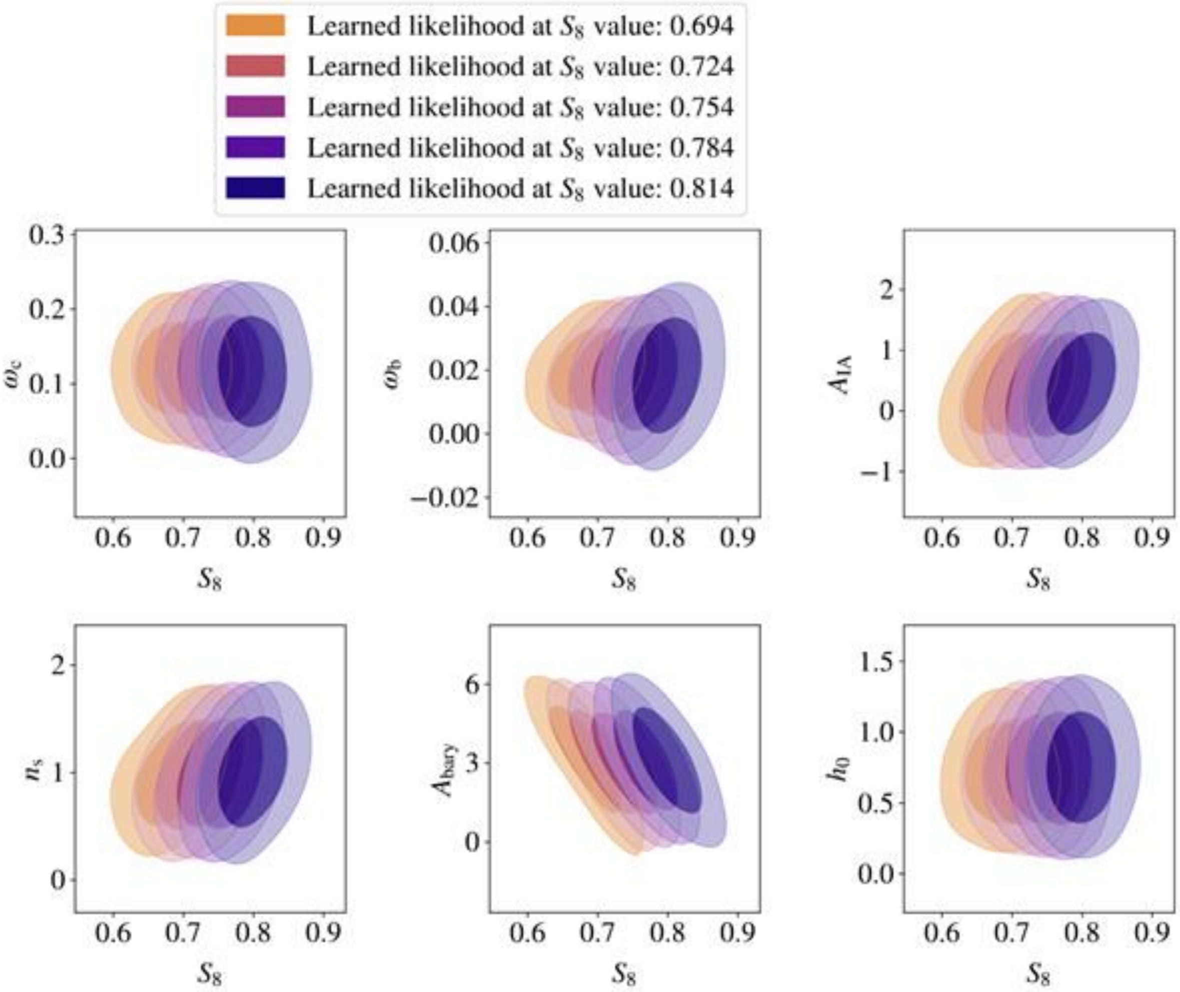
Questions?

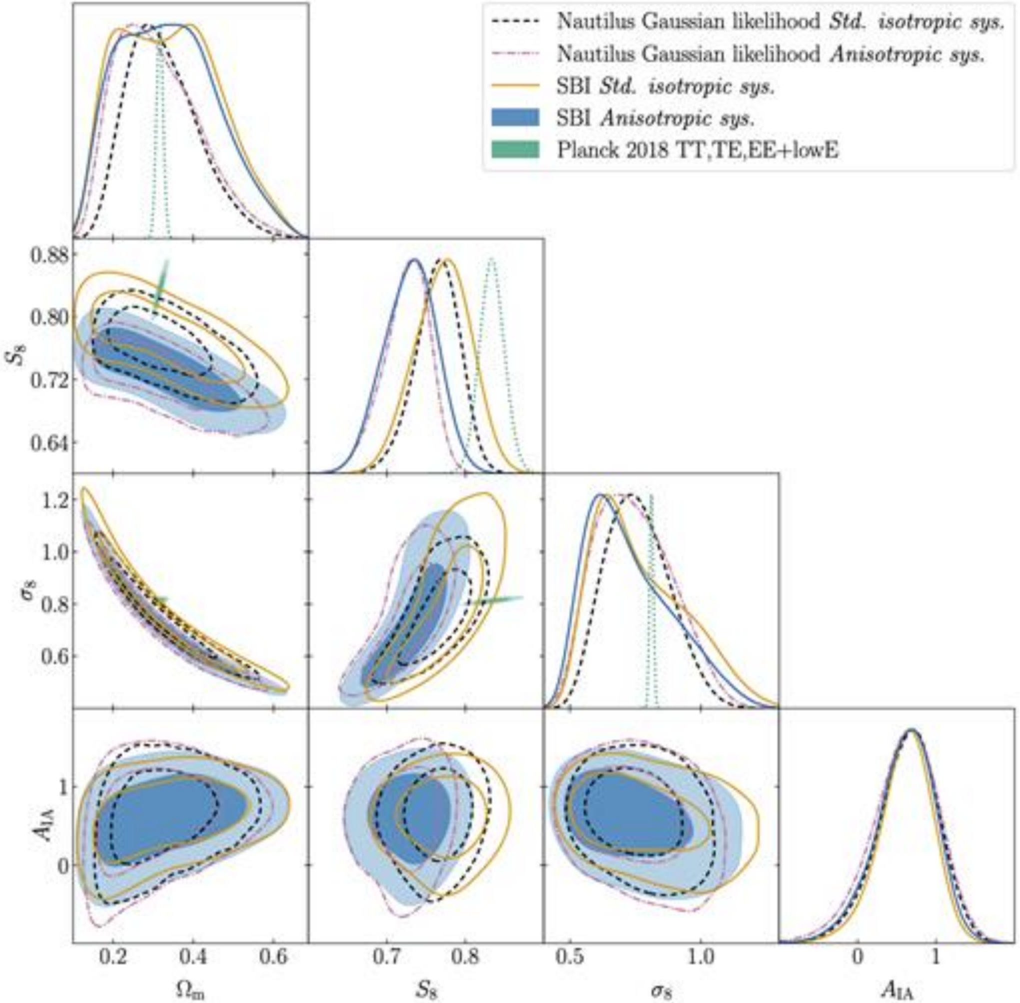
Appendix

Forward Simulations




NDE Committee
NDE 0
NDE 1
NDE 2
NDE 3
NDE 4
NDE 5
NDEs Combined
Ω_m
0.5
0.4
0.3
0.2
0.6
0.7
0.8
0.9
1.0
σ_8
0.2
0.3
0.4
0.5
Ω_m


Learned likelihood at S_8 value: 0.694
Learned likelihood at S_8 value: 0.724
Learned likelihood at S_8 value: 0.754
Learned likelihood at S_8 value: 0.784
Learned likelihood at S_8 value: 0.814
ω_c
ω_b
A_{IA}
n_s
A_{bary}
h_0
S_8


Nautilus Gaussian likelihood *Std. isotropic sys.*
Nautilus Gaussian likelihood *Anisotropic sys.*
SBI *Std. isotropic sys.*
SBI *Anisotropic sys.*
Planck 2018 TT,TE,EE+lowE
S_8
σ_8
A_{IA}
Ω_m
S_8
σ_8
A_{IA}