KiDS SBI

Simulation-Based Inference of KiDS-1000 Cosmic Shear

Maximilian von Wietersheim-Kramsta
maximilian.von-wietersheim-kramsta@durham.ac.uk

21/05/2023 – COSMO21, Chania

Kilo-Degree Survey

Durham University

UCL

Kiyam Lin

Nicolas Tessore

Benjamin Joachimi

Arthur Loureiro

Robert Reischke

Angus Wright

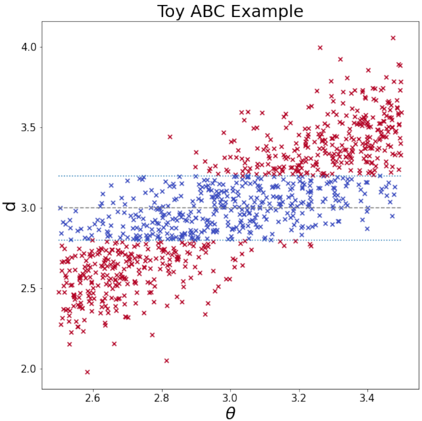
Simulation-Based Inference (SBI)

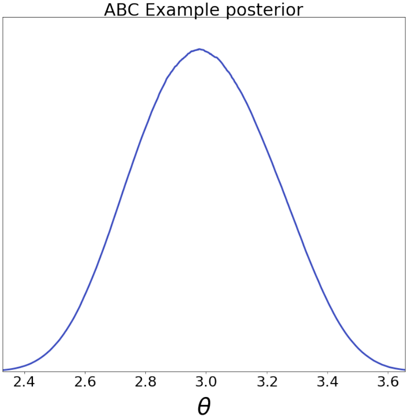
A.k.a. likelihood-free or implicit likelihood inference

- Signal and noise modelling as complex as simulations can be.
- Likelihood can take an arbitrary form (non-Gaussian).
- Full Bayesian uncertainty propagation from measurements to parameters.
- Number of simulations required similar to the number needed for numerical covariances [Lin et al. 2022; arxiv:2212.04521].

SBI: Approximate Bayesian Computation

$$P(\boldsymbol{\theta}|\boldsymbol{d}) = \frac{P(\boldsymbol{d}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})}{P(\boldsymbol{d})} \propto P(\boldsymbol{\theta}, \boldsymbol{d}) \cdot P(\boldsymbol{\theta}) \tag{1}$$





SBI: Density Estimation

$$P(\boldsymbol{\theta}|\boldsymbol{d}) = \frac{P(\boldsymbol{d}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})}{P(\boldsymbol{d})} \propto P(\boldsymbol{\theta}, \boldsymbol{d}) \cdot P(\boldsymbol{\theta}) \tag{2}$$

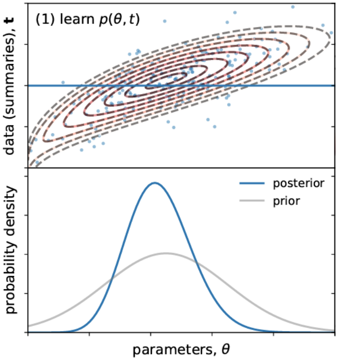
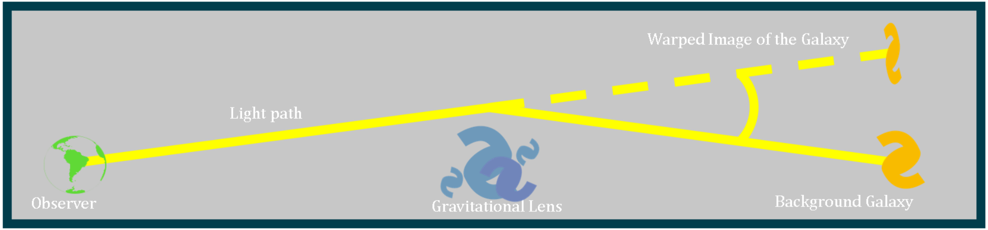
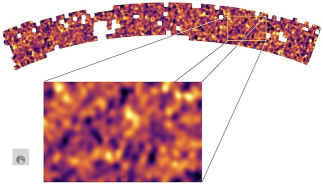


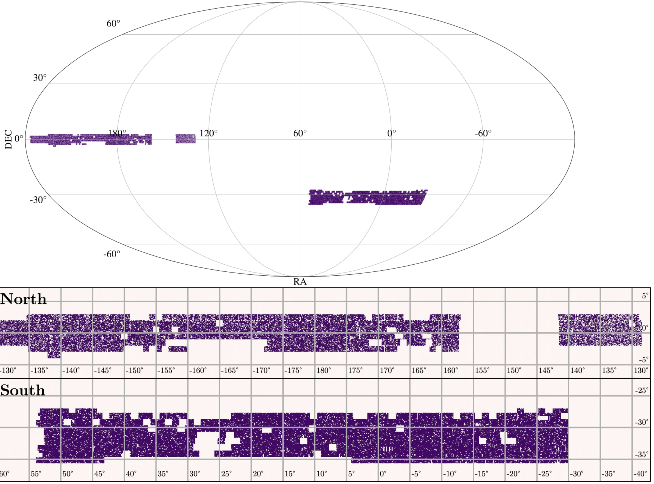
Figure: [Alsing et al. 2019]

Weak Gravitational Lensing

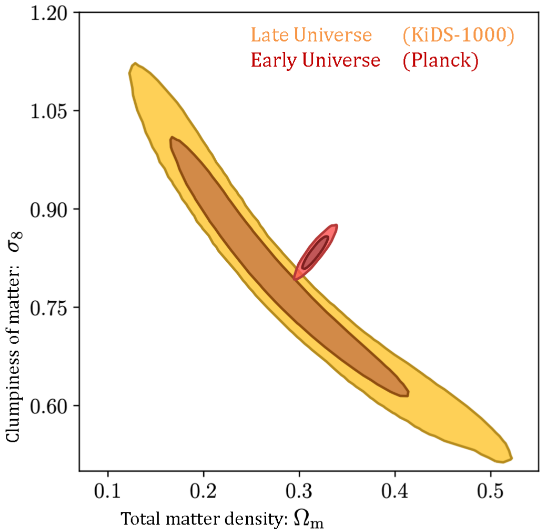




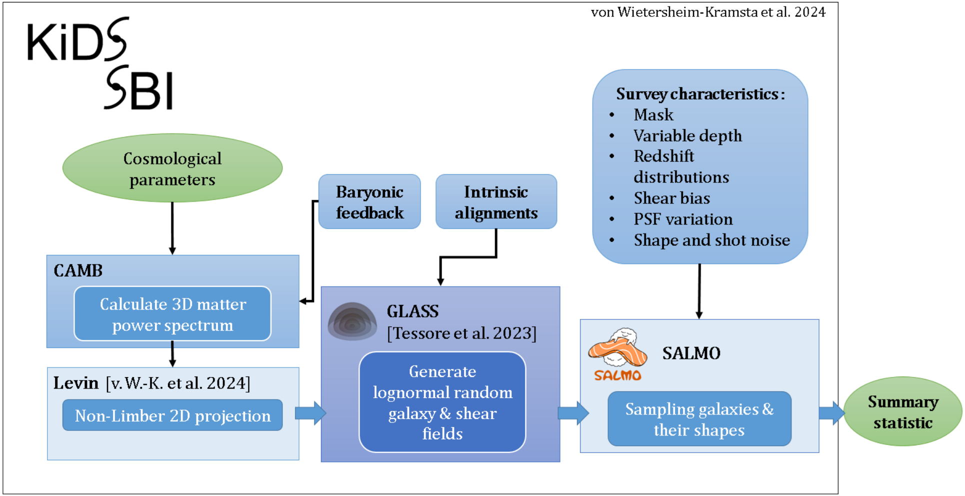
Kilo-Degree Survey: KiDS-1000



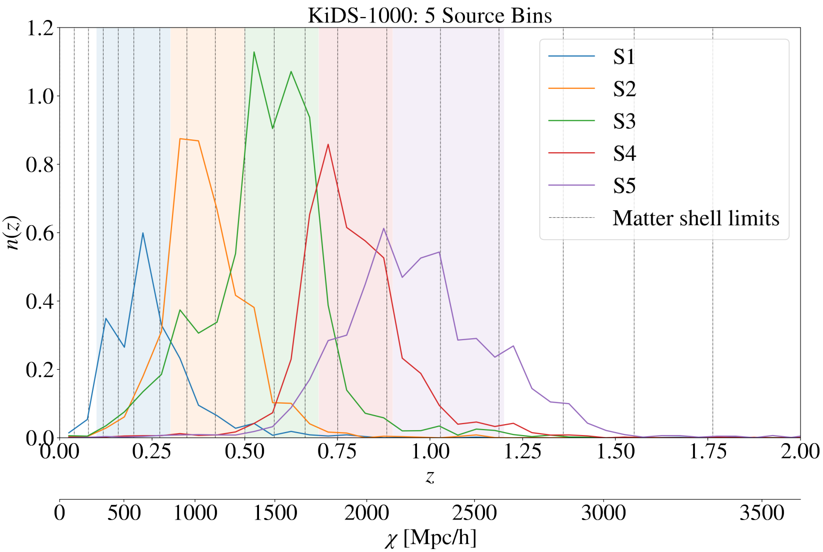
Previous KiDS-1000 Cosmic Shear Results



Forward-Simulation Pipeline



Geometry of the simulations & tomography



Geometry of the simulations

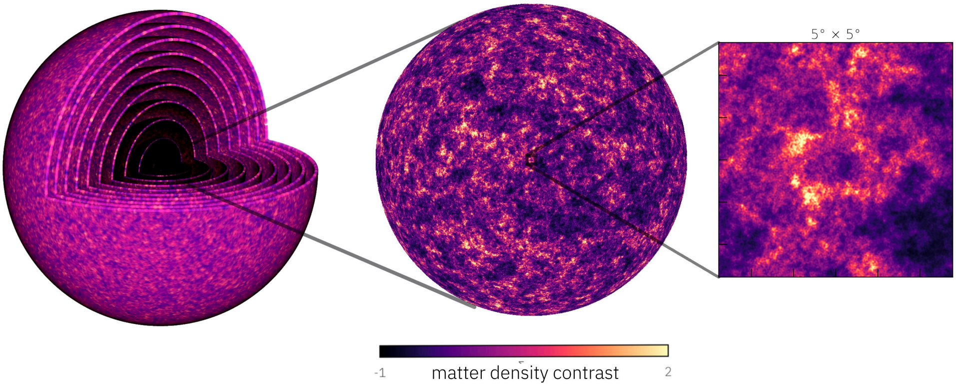
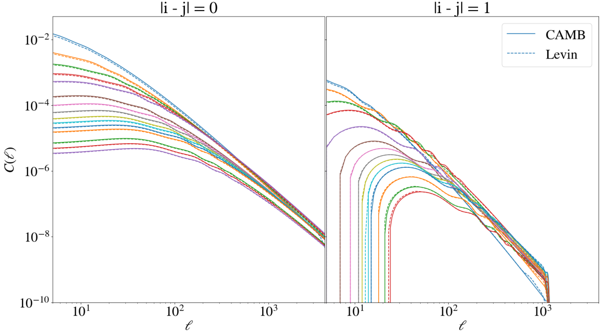
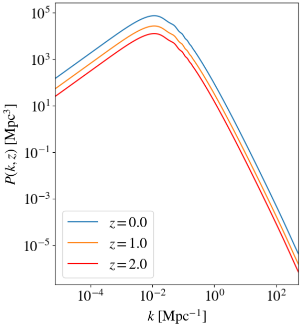


Figure: [Tessore et al. 2023; arxiv:2302.01942]

Non-Limber Projection: Levin

$$C^{(ij)}_{\delta\delta}(\ell) = \frac{2}{\pi} \int d\chi\, W^{(i)}(z[\chi]) \int d\chi'\, W^{(j)}(z'[\chi]) \int dk\, k^2\, P(k, z[\chi], z'[\chi])\, j_\ell(k\chi)\, j_\ell(k\chi') \tag{3}$$





[von Wietersheim-Kramsta et al. 2024; arXiv:2404.15402]

GLASS: Generator of Large Scale Structure

$$C_{\delta\delta}^{(ij)}(\theta) = \langle \delta^{(i)}(\theta)\, \delta^{(j)*}(\theta) \rangle \tag{4}$$

$$\kappa(\theta, z) = \frac{3\Omega_m}{2} \int_0^z \mathrm{d}z' \, \frac{f_k(z')\,[f_k(z') - f_k(z)]}{f_k(z)} \frac{1+z'}{E(z')}\, \delta(\theta, z') \tag{5}$$

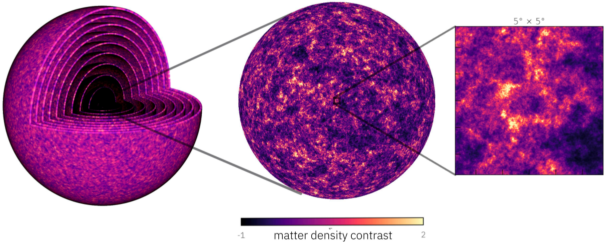


Figure: [Tessore et al. 2023; arxiv:2302.01942]

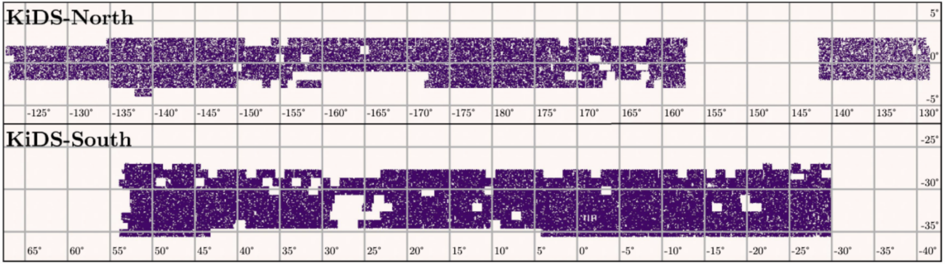
Galaxy Positions & Survey Characteristics

Shell Tomographic bin

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

Pixel

Mask



[von Wietersheim-Kramsta et al. 2024; arXiv:2404.15402]

Galaxy Positions & Survey Characteristics

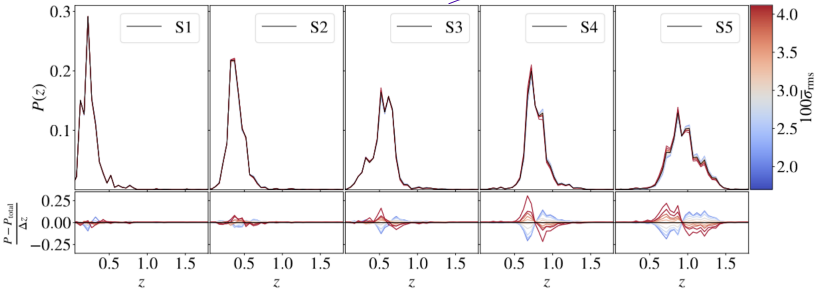
Shell

Tomographic bin

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

Pixel

Redshift distributions



[von Wietersheim-Kramsta et al. 2024; arXiv:2404.15402]

Galaxy Positions & Survey Characteristics

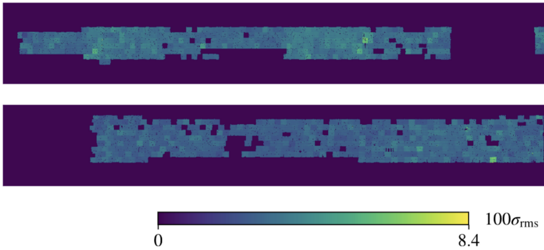
Shell

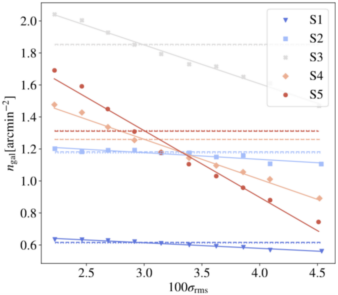
Tomographic bin

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

Pixel

Galaxy density (variable depth)





[von Wietersheim-Kramsta et al. 2024; arXiv:2404.15402]

Variable Depth

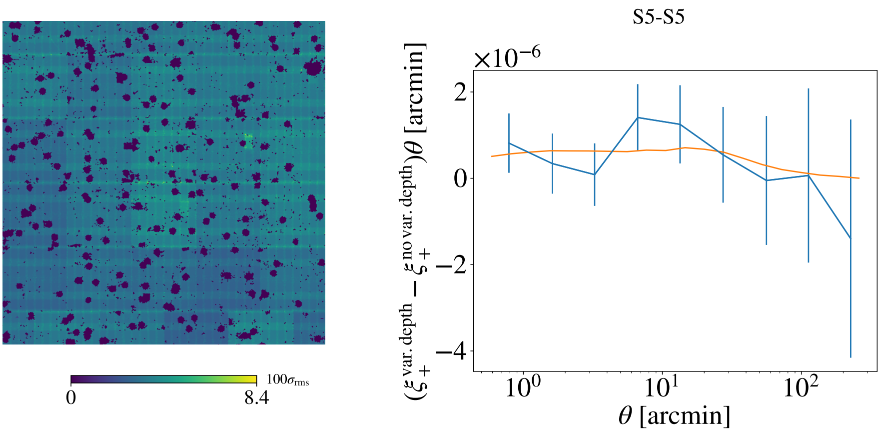
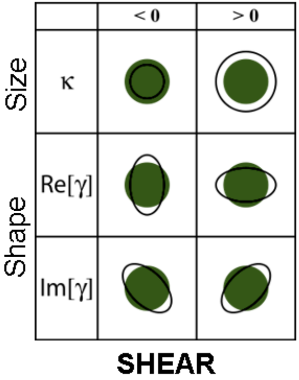


Figure: Orange: theory [Heydenreich et al. 2020; arXiv:1910.11327]. Blue: forward-simulations.

Cosmic Shear & Galaxy Shapes

$$g(\boldsymbol{\Theta}) = \frac{\gamma(\boldsymbol{\Theta})}{1 - \kappa(\boldsymbol{\Theta})} \tag{6}$$

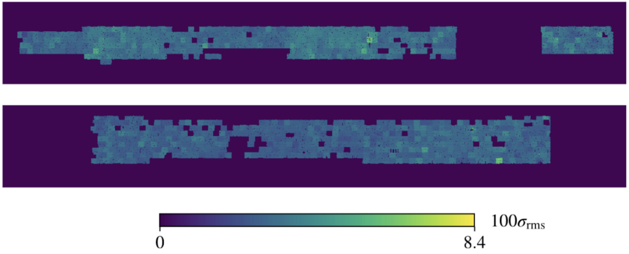


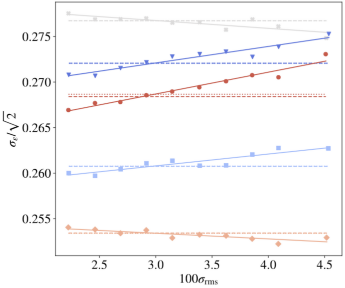
Galaxy Shapes & Survey Characteristics

Reduced shear

$$\epsilon_{\text{lensed}}(\mathbf{\Theta}) = \frac{\epsilon_{\text{int}} + g(\mathbf{\Theta})}{1 + g^*(\mathbf{\Theta})\epsilon_{\text{int}}}$$

Intrinsic shapes
(variable depth)





[von Wietersheim-Kramsta et al. 2024; arXiv:2404.15402]

Galaxy Shapes & Survey Characteristics

Pixel

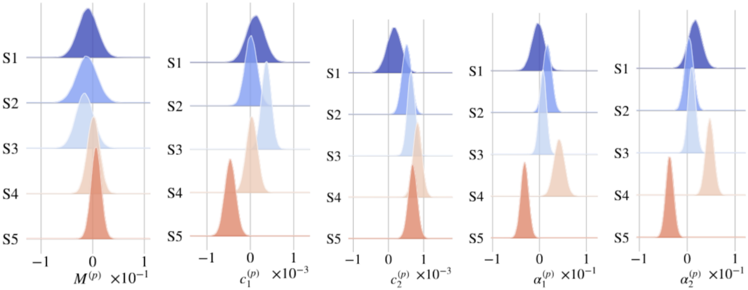
$$\epsilon_{\rm obs}(p, m; \boldsymbol{\Theta}) = \left(1 + M^{(p)}\right)\epsilon_{\rm lensed}(\boldsymbol{\Theta}) + \alpha^{(p)}\epsilon_{\rm PSF}(m) + \underbrace{\cancel{\beta^{(p)}\delta\epsilon_{\rm PSF}}}_{0} + c^{(p)}$$

Tomographic bin

Multiplicative shear bias

PSF shear bias

Additive shear bias



[von Wietersheim-Kramsta et al. 2024; arXiv:2404.15402]

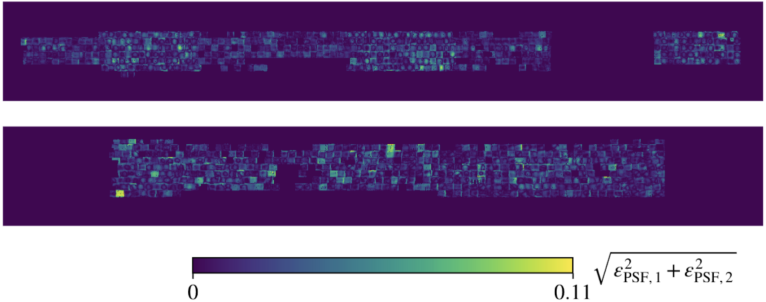
Galaxy Shapes & Survey Characteristics

Pixel

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

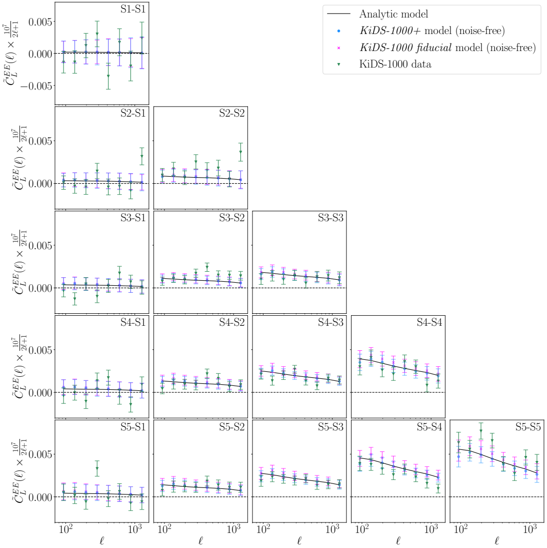
Tomographic bin

PSF shear bias

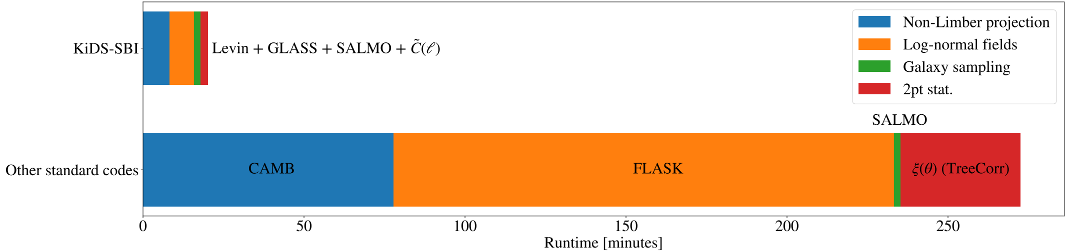


[von Wietersheim-Kramsta et al. 2024; arXiv:2404.15402]

Measurement: Pseudo Angular Power Spectra

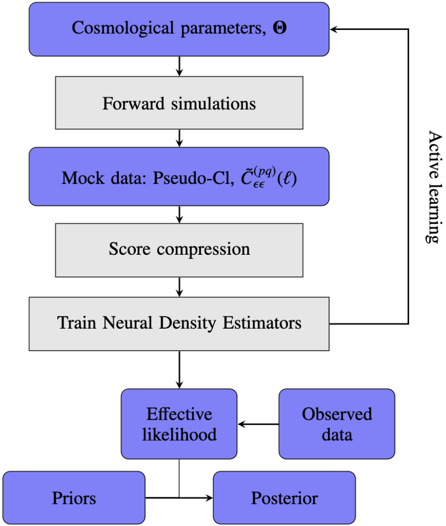


Forward-simulations Speed



[von Wietersheim-Kramsta et al. 2024; arXiv:2404.15402]

KiDS-SBI



Cosmological parameters, $\boldsymbol{\Theta}$
Forward simulations
Mock data: Pseudo-Cl, $\tilde{C}_{\epsilon\epsilon}^{(pq)}(\ell)$
Score compression
Train Neural Density Estimators
Active learning
Effective likelihood
Observed data
Priors
Posterior

KiDS-SBI: Parameters and Priors

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	$\omega_{\rm c}$	Flat	[0.051, 0.255]	0.118
Baryonic matter density	$\omega_{\rm b}$	Flat	[0.019, 0.026]	0.026
Scalar spectral index	$n_{\rm s}$	Flat	[0.84, 1.1]	0.901
Intrinsic alignment amp.	$A_{\rm IA}$	Flat	[-6, 6]	0.264
Baryon feedback amp.	$A_{\rm bary}$	Flat	[2, 3.13]	3.1
Redshift displacement	$\boldsymbol{\delta}_z$	Gaussian	$\mathcal{N}(\mathbf{0}, \mathbf{C}_z)$	$\mathbf{0}$
Multiplicative shear bias	$\boldsymbol{M}^{(p)}$	Gaussian	$\mathcal{N}(\overline{\boldsymbol{M}}^{(p)}, \boldsymbol{\sigma}_{\boldsymbol{M}}^{(p)})$	$\overline{\boldsymbol{M}}^{(p)}$
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	$\mathcal{N}(\overline{\alpha}_{1,2}^{(p)}, \sigma_{\alpha_{1,2}}^{(p)})$	$\overline{\alpha}_{1,2}^{(p)}$

[von Wietersheim-Kramsta et al. 2024; arXiv:2404.15402]

KiDS-SBI: Score Compression

Further massive compression from summary statistics to reduce dimensionality via score compression [Alsing et al. 2018]

$$\mathcal{L} = \mathcal{L}_* + \delta\boldsymbol{\theta}^T \nabla \mathcal{L}_* - \frac{1}{2}\delta\boldsymbol{\theta}^T \mathbf{J}_* \delta\boldsymbol{\theta}, \tag{7}$$

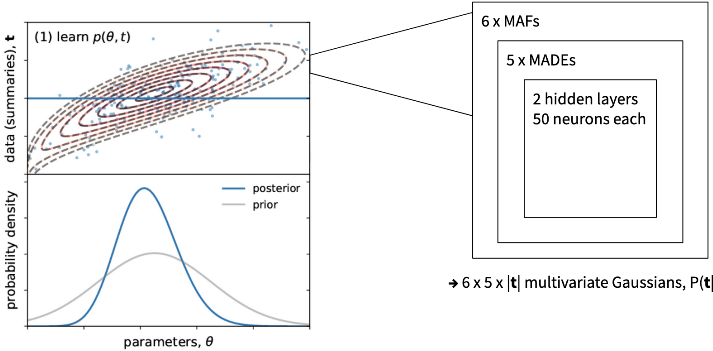
$$\boldsymbol{t} = \nabla \boldsymbol{\mu}^T \boldsymbol{C}^{-\mathbf{1}}(\boldsymbol{d} - \boldsymbol{\mu}). \tag{8}$$

The score compression is repeated after an initial SBI which refines the fiducial parameter estimate.

KiDS-SBI: Neural Likelihood Estimation (DELFI)

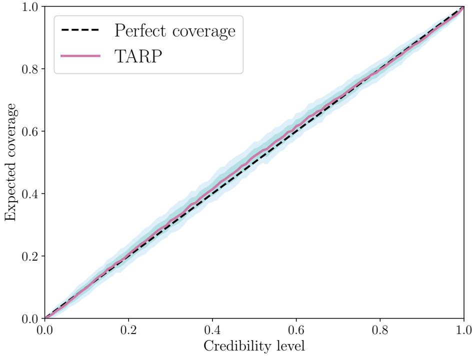
An ensemble of six independent Masked Autoregressive Flows (MAFs) is combined to characterise the likelihood

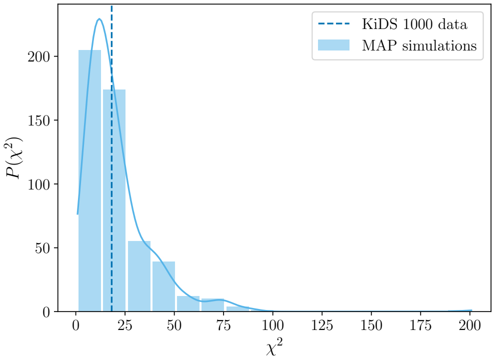
$$p(\mathbf{t}|\theta;\mathbf{w}) = \prod_{\alpha=1}^{N_{\mathrm{NDEs}}} \beta_\alpha p_\alpha(\mathbf{t}|\theta;\mathbf{w}) \tag{9}$$



→ 6 x 5 x |**t**| multivariate Gaussians, P(**t**|***Θ***)

Internal Consistency & Goodness-of-Fit





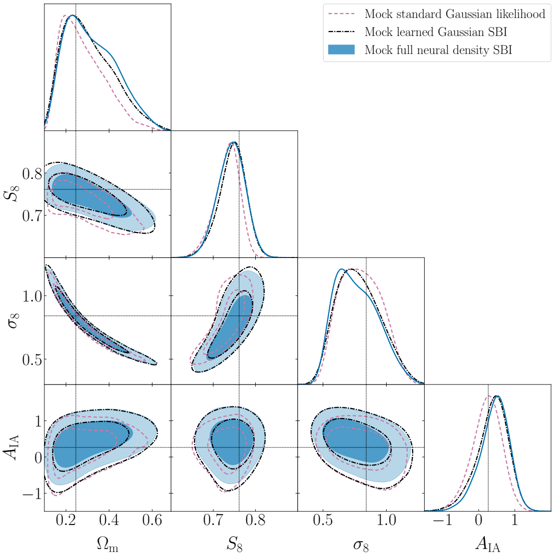
[von Wietersheim-Kramsta et al. 2024; arXiv:2404.15402]

KiDS-SBI: Mock Analysis

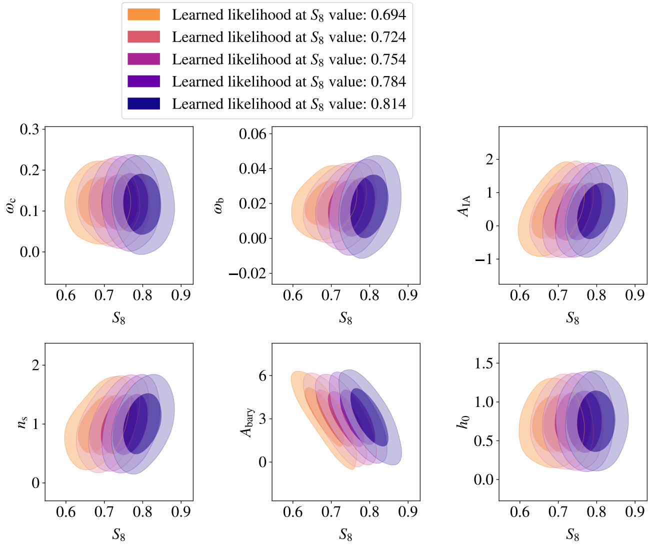
Compare different types of inferences:

- "Standard" Gaussian likelihood assumption with a fixed covariance matrix
- SBI with a learned multivariate Gaussian (MDN)
  $\Rightarrow$ Gaussian likelihood with a parameter-dependent covariance matrix
- **Full neural density SBI** (ensemble of MAFs)
  $\Rightarrow$ non-Gaussian likelihood with a parameter-dependent covariance matrix

KiDS-SBI: Mock Analysis

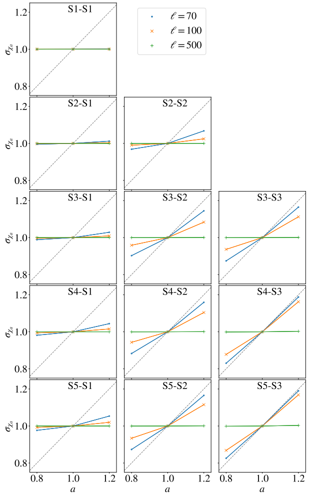


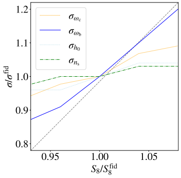
S_8-dependent Uncertainty



S_8-dependent Uncertainty from Theory

Scaling of the analytical Gaussian likelihood in KiDS-1000 if $C^{(ij)}_{\mathrm{scaled}}(\ell) = a\, C^{(ij)}(\ell)$





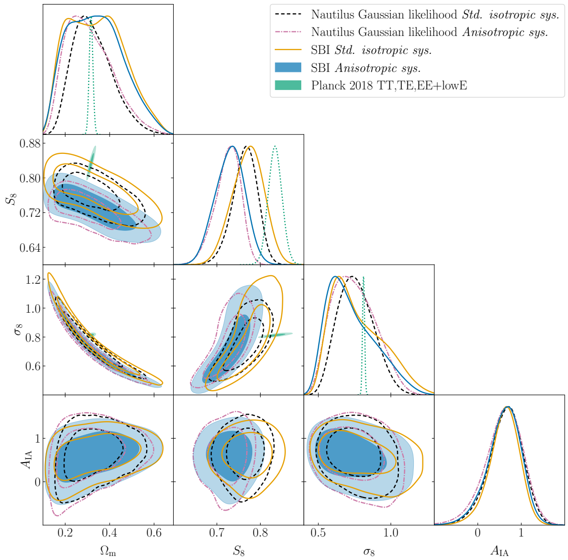
KiDS-SBI: Forward Models

Two main models:

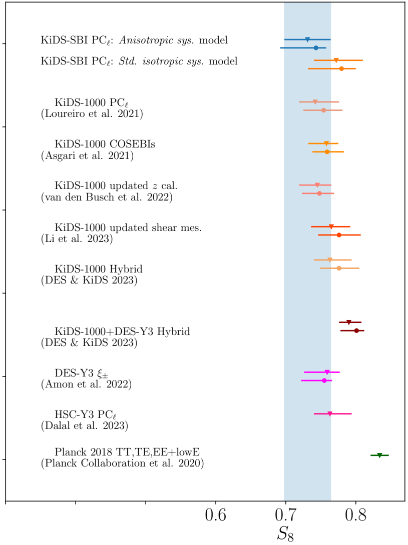
- **"Standard" isotropic systematics model**:
  galaxy selection function isotropic in each tomographic bin

- **Anisotropic systematics model**:
  anistropic galaxy selection (position, shapes and redshift)
  + anisotropic PSF distortions

Cosmological Results for Flat ΛCDM



Cosmological Results for Flat ΛCDM



KiDS-SBI: Conclusions

- SBI is a powerful tool to rigorously conduct cosmological inference and model testing
- We report $\mathbf{S_8 = 0.731 \pm 0.033}$ which is in 2.9σ tension with Planck 2020
- Neglecting variable depth and shear biases can bias S_8 by approx. 5%
- Cosmic shear likelihood is consistent with a Gaussian
- Its covariance is measurably cosmology-dependent (cosmic variance)
  $\Rightarrow$ uncertainty on S_8 is 10% higher

[von Wietersheim-Kramsta et al. 2024; arXiv:2404.15402]

Questions?

References

Niall Jeffrey, Justin Alsing, François Lanusse (2020)
Likelihood-free inference with neural compression of DES SV weak lensing map statistics
MNRAS Volume 501, Issue 1, February 2021, Pages 954–969.

Arthur Loureiro, et al. (2021)
KiDS & Euclid: Cosmological implications of a pseudo angular power spectrum analysis of KiDS-1000 cosmic shear tomography
arxiv 2110.06947v1 .

David Levin (1994)
Fast integration of rapidly oscillatory functions
JCAM Volume 67, Issue 1, 20 February 1996, Pages 95-101.

References

Justin Alsing et al. (2019)
Fast likelihood-free cosmology with neural density estimators and active learning
Monthly Notices of the Royal Astronomical Society Volume 488, Issue 3, September 2019, Pages 4440-4458.

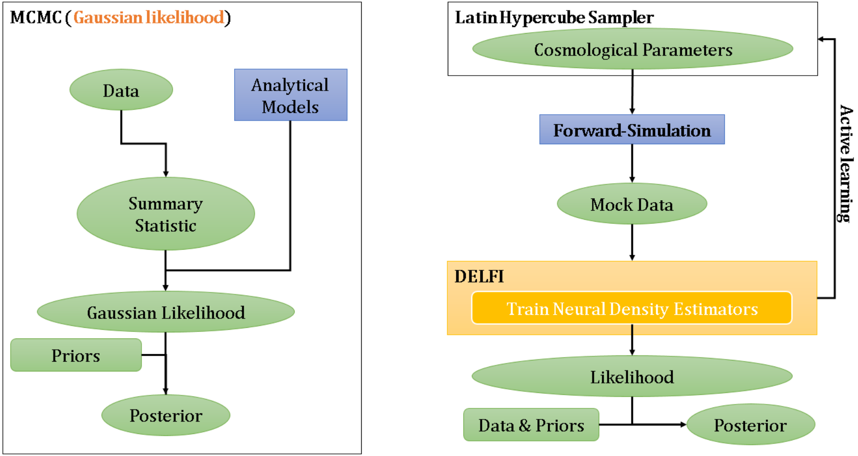
Justin Alsing et al. (2018)
Massive optimal data compression and density estimation for scalable, likelihood-free inference in cosmology
Monthly Notices of the Royal Astronomical Society Volume 477, Issue 3, July 2018, Pages 2874–2885.

Appendices A - Pipeline Setup

- Shear-shear weak lensing from KiDS-1000 using pseudo-Cls
- Latin hypercube of cosmology values as simulation input

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	$\mathcal{N}(\overline{\alpha}_{1,2}^{(p)}, \sigma_{\alpha_{1,2}}^{(p)})$	$\overline{\alpha}_{1,2}^{(p)}$

Appendices A - Pipeline Setup



Appendices B - Compression and DELFI

- A stack of NDEs are used in DELFI

$$p(\mathbf{t}|\theta;\mathbf{w}) = \prod_{\alpha=1}^{N_{\mathrm{NDEs}}} \beta_\alpha p_\alpha(\mathbf{t}|\theta;\mathbf{w}), \tag{10}$$

- A mixture of Gaussian Mixture Density Networks (MDNs) and Masked Autoregressive Flows (MAFs) are employed in this ensemble
- Further massive compression from summary statistics to reduce dimensionality via score compression

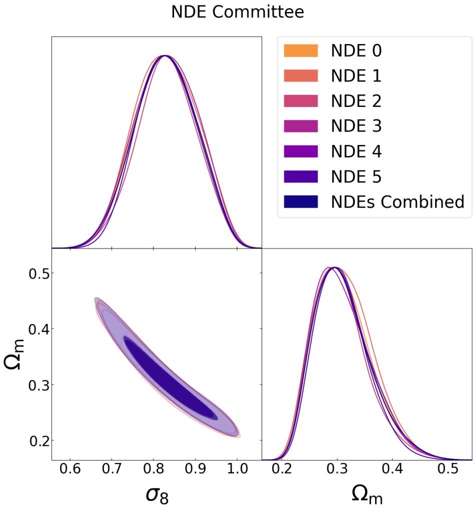
$$\mathcal{L} = \mathcal{L}_* + \delta\boldsymbol{\theta}^T \nabla \mathcal{L}_* - \frac{1}{2}\delta\boldsymbol{\theta}^T \mathbf{J}_* \delta\boldsymbol{\theta}, \tag{11}$$

$$\boldsymbol{t} = \nabla\boldsymbol{\mu}^T \boldsymbol{C}^{-1}(\boldsymbol{d} - \boldsymbol{\mu}), \tag{12}$$

- Remap this to a MLE estimate via the Fisher matrix

$$\hat{\theta} = \theta_* + \boldsymbol{F}_*^{-1}\nabla\mathcal{L}_* = \theta_* + \boldsymbol{F}_*^{-1}\boldsymbol{t}_*, \tag{13}$$

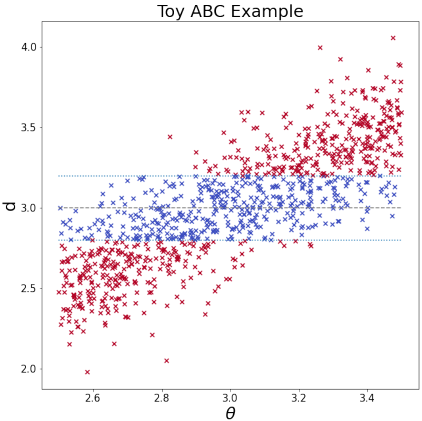
Appendices C - PyDELFI

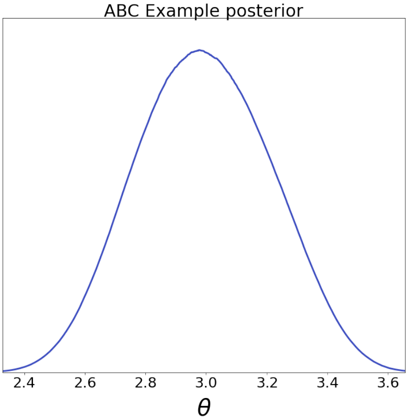


[Alsing et al. 2019]

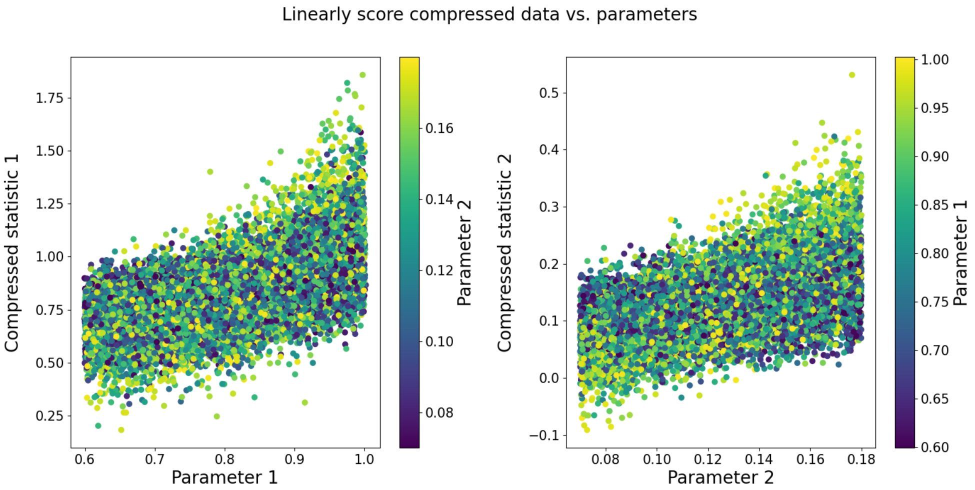
Appendices D - SBI - ABC

$$P(\boldsymbol{\theta}|\boldsymbol{d}) = \frac{P(\boldsymbol{d}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})}{P(\boldsymbol{d})} \tag{14}$$





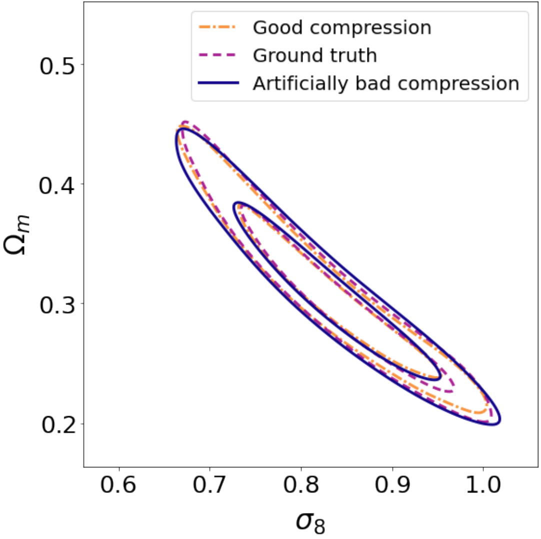
Appendices E - Score Compression



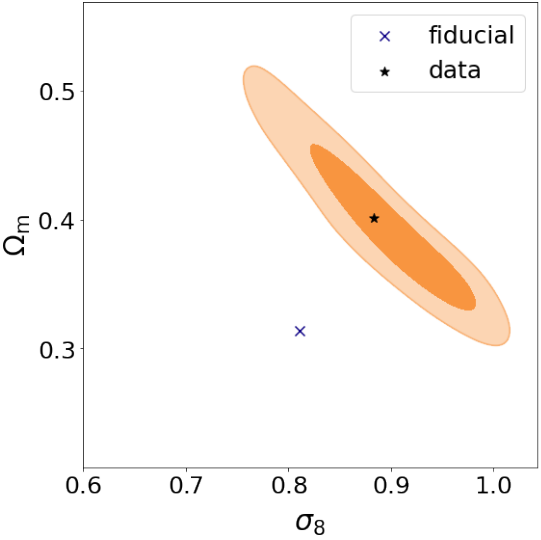
[Alsing et al. 2018]

$$\mathbf{t} = \mathbf{M}(\mathbf{d})$$

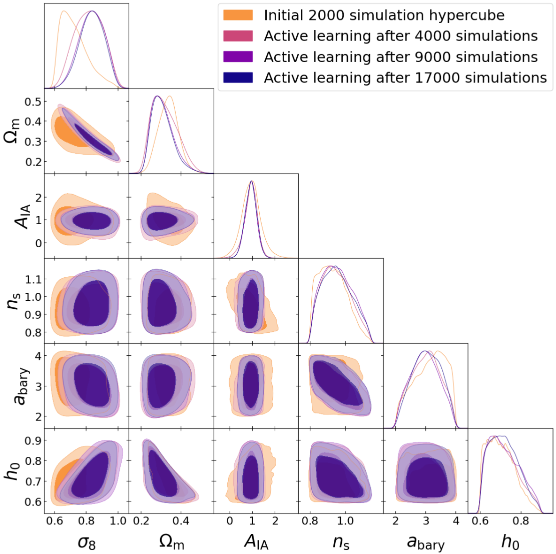
Appendices F - Sensitivity to Compression



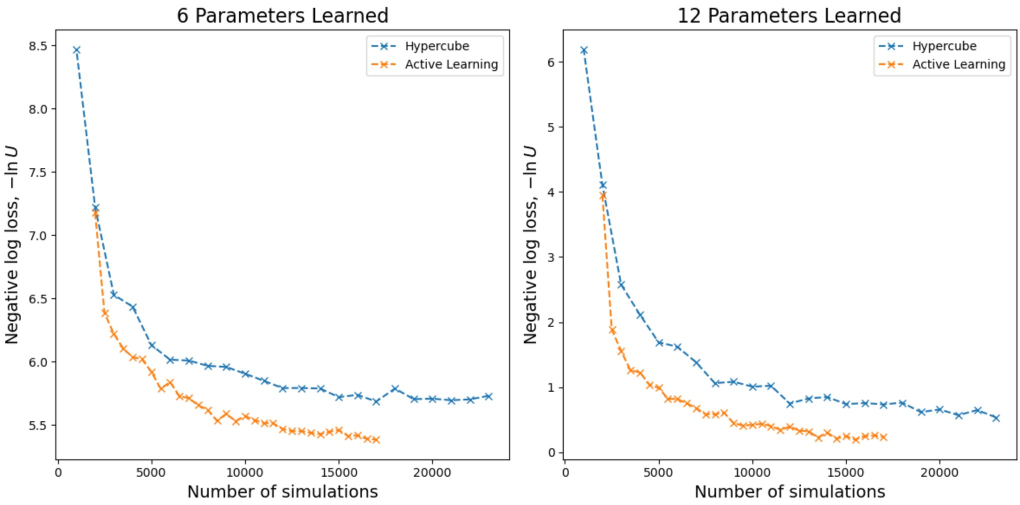
Appendix G - Sensitivity to Fiducial Cosmology



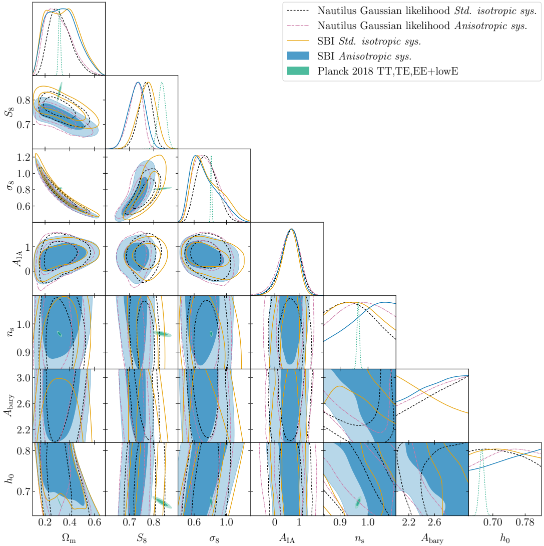
Appendix H - Active Learning



Appendix I - Simulation Number Sufficiency



Appendix J - Full Posterior



Appendix K - Signal Test