Simulation-Based Inference:
But what is my model even doing?

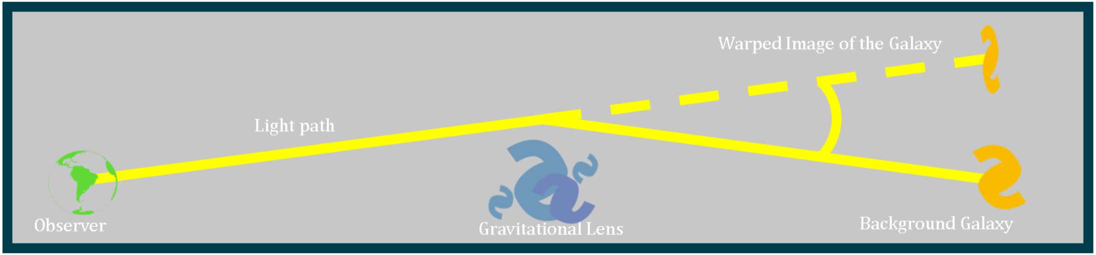
Maximilian von Wietersheim-Kramsta
maximilian.von-wietersheim-kramsta@durham.ac.uk

10/04/2024 – SBI for Galaxy Evolution, University of Bristol

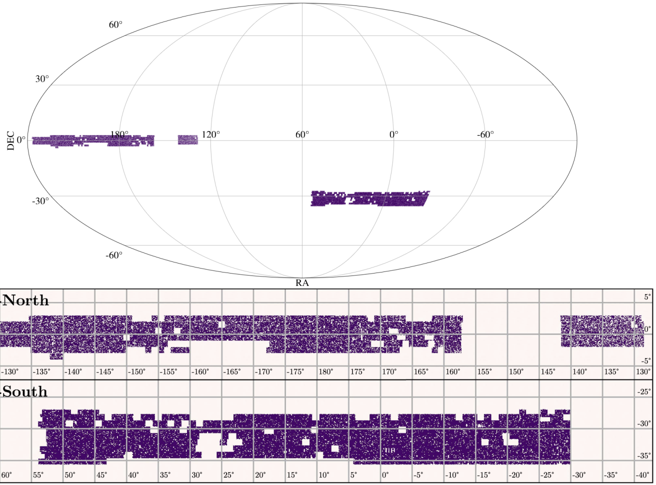
Simulation-Based Inference (SBI) - FAQs

- How do I pick a forward model?
- Can I do a blind analysis?
- My posterior is robust and reliable, but is my model accurate?
- How do I get the most out of the results of my SBI?

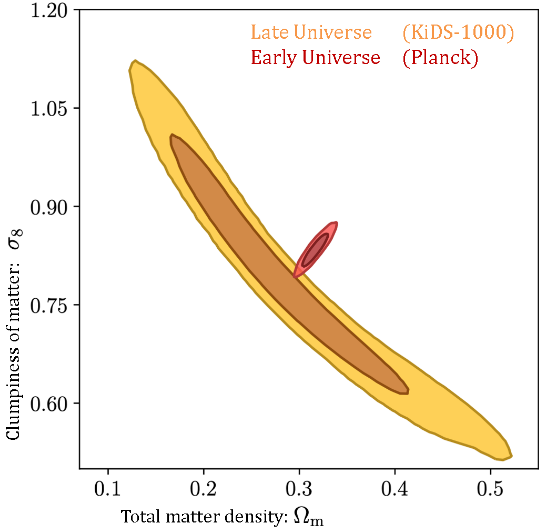
Example: Weak Gravitational Lensing



Kilo-Degree Survey: KiDS-1000



Previous KiDS-1000 Cosmic Shear Results



What forward model should I use?

What can we get away with?

1. In what are we interested? - Parameters, model comparison...

2. What is our measurement? - Summary statistics, images...

3. To which systematics will the measure be sensitive?

4. What type of SBI suits our problem? (see Matt Ho's and Lucas Makinen's talks) - NPE, NLE, NRE, active learning...

Example: KiDS-1000 Cosmic Shear

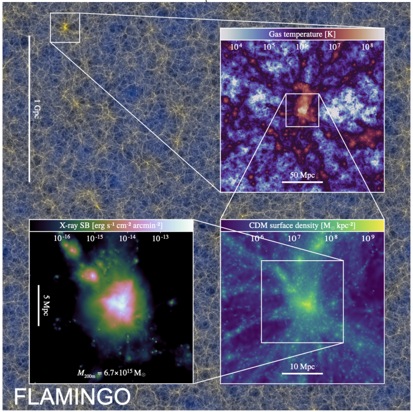
1. Interested in parameter inference
2. Measure 2-point statistics (pseudo-Cls)
3. A range of systematics:
   - Physical: intrinsic alignments, baryonic feedback, (galaxy bias)
   - Observational: photometric redshift uncertainty, masking, galaxy shape calibration uncertainties, anisotropies in the selection function
4. Suited for Neural Likelihood Estimation

Complexity

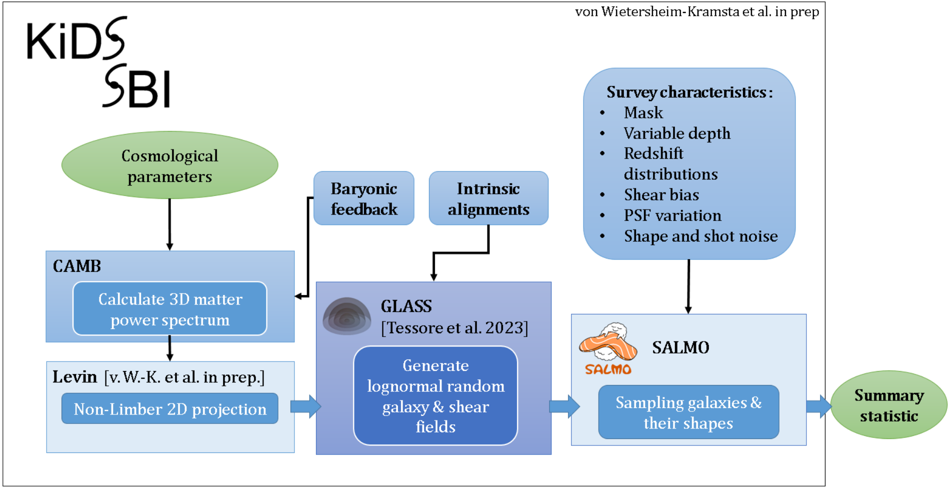
$$C_{\epsilon\epsilon}^{(pq)}(\ell) = C_{\mathrm{gg}}^{(pq)}(\ell) + C_{\mathrm{gI}}^{(pq)}(\ell) + C_{\mathrm{Ig}}^{(pq)}(\ell) + C_{\mathrm{II}}^{(pq)}(\ell)$$

$$C_{ab}^{(pq)}(\ell) = \int_0^\infty \frac{d\chi}{f_k^2(\chi)}\, W_a^{(p)}(\chi)\, W_b^{(q)}(\chi)\, P_\delta\!\left(\frac{\ell + 1/2}{f_k(\chi)}, \chi\right)$$

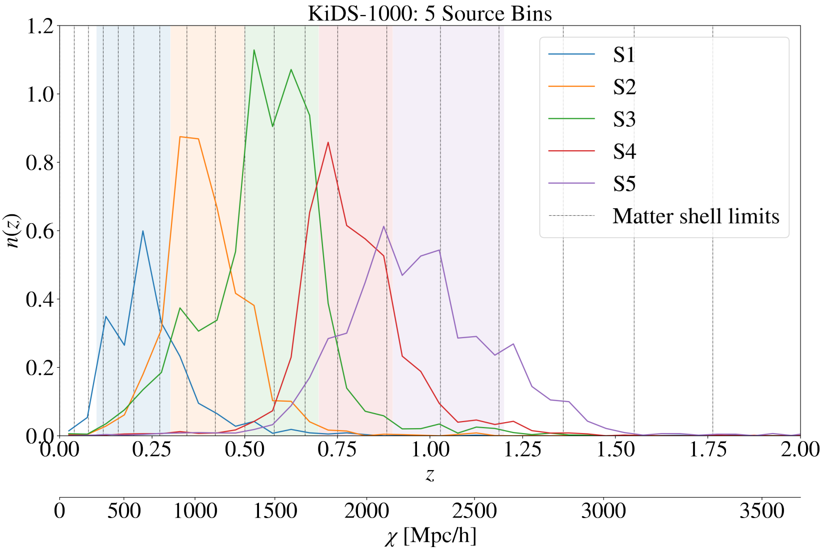
$$\tilde{C}_{\epsilon\epsilon,\mu}^{(pq)}(\ell;\mathbf{\Theta}) = \sum_{\ell'=0}^{\ell'_{\max}} \sum_{\ell''=0}^{\ell''_{\max}} \sum_{\nu=1}^{3} \sum_{\nu'=1}^{3} M_{\mu\nu',\ell\ell''}\, M'^{(pq)}_{\nu'\nu,\ell''\ell'}\, C_{\epsilon\epsilon,\nu}^{(pq)}(\ell';\mathbf{\Theta})$$



Statistical Forward Simulations



Photometric Redshift Uncertainties



Geometry of the simulations

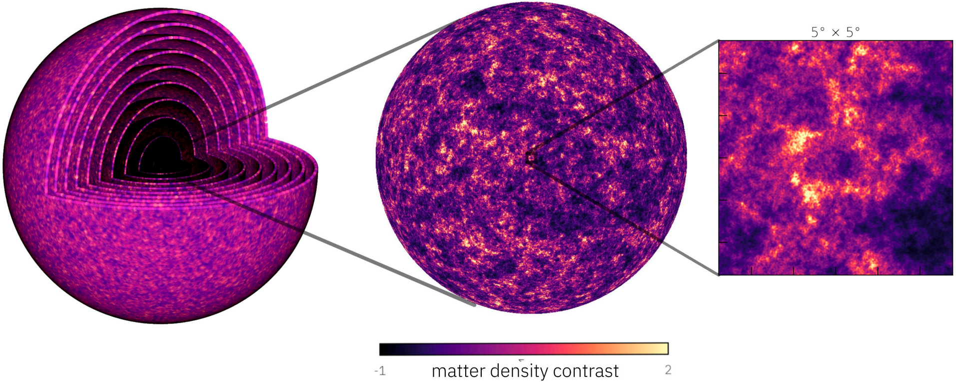
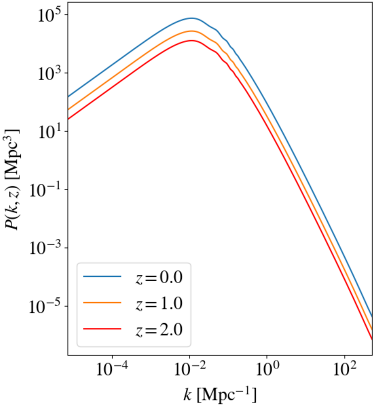
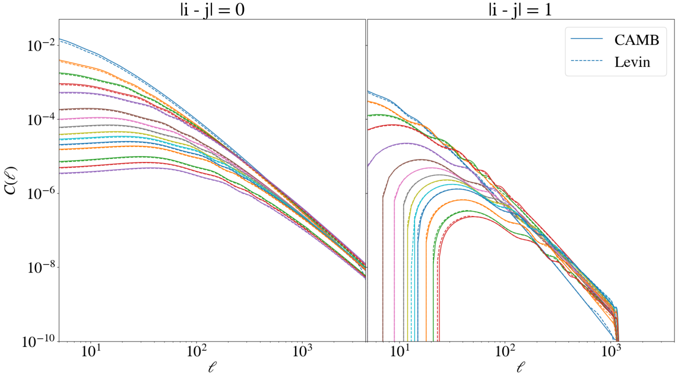


Figure: [Tessore et al. 2023; arxiv:2302.01942]

Non-Limber Projection: Levin





GLASS: Generator of Large Scale Structure

$$C_{\delta\delta}^{(ij)}(\theta; \mathbf{\Theta}) = \langle \delta^{(i)}(\theta; \mathbf{\Theta})\, \delta^{(j)*}(\theta; \mathbf{\Theta}) \rangle \tag{1}$$

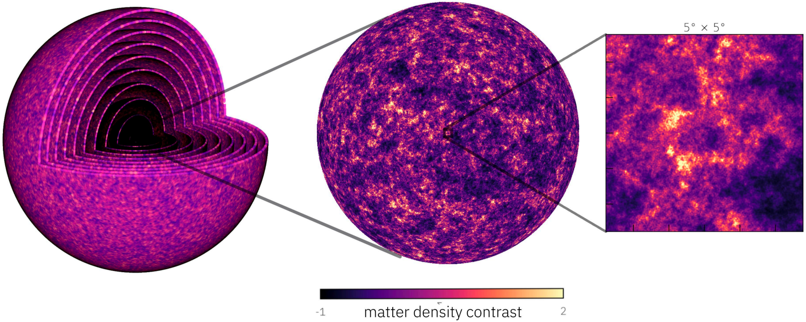


Figure: [Tessore et al. 2023; arxiv:2302.01942]

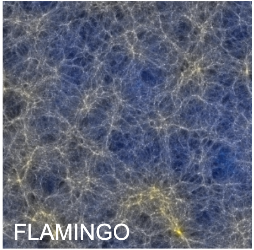
Galaxy Sampling

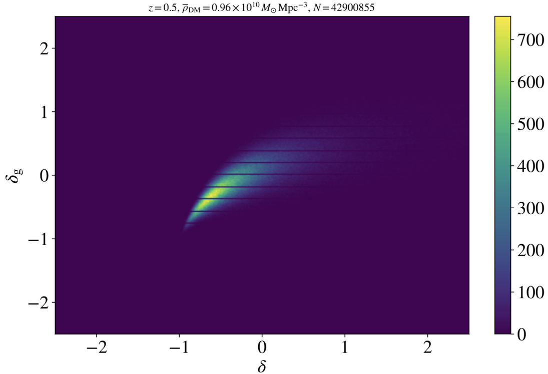
Shell Tomographic bin

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

Pixel

Galaxy bias





Survey Characteristics

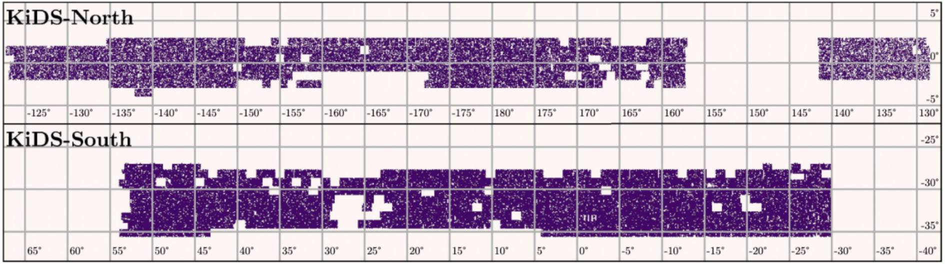
Shell

Tomographic bin

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

Pixel

Mask



Survey Characteristics

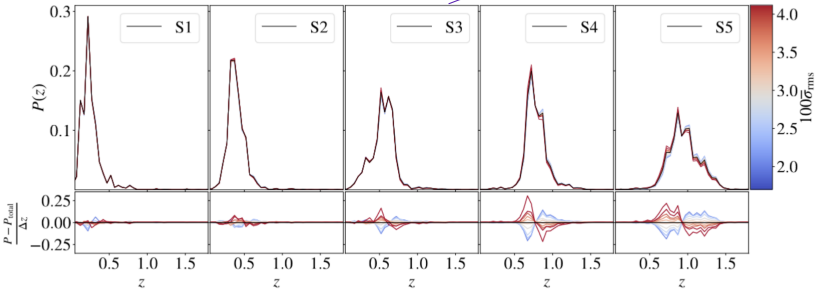
Shell

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



Survey Characteristics

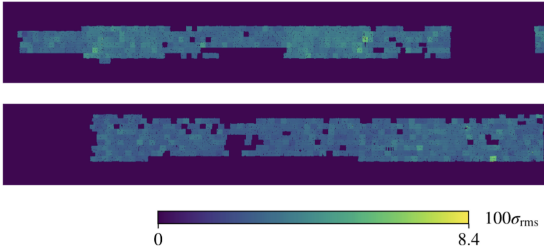
Shell

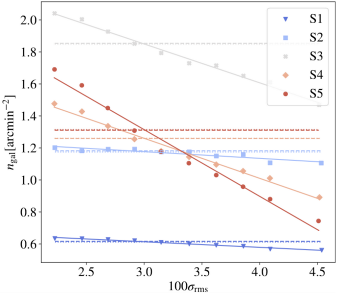
Tomographic bin

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

Pixel

Galaxy density (variable depth)



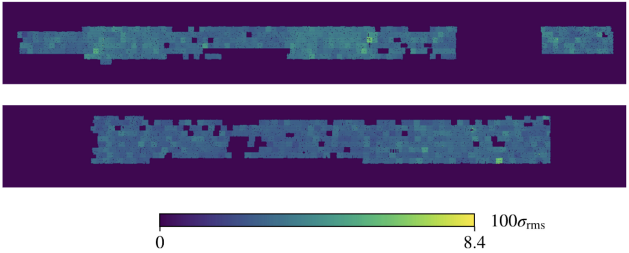


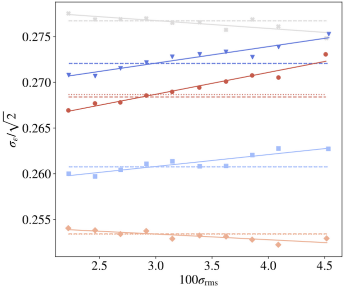
Survey Characteristics

Reduced shear

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

Intrinsic shapes
(variable depth)





Survey Characteristics and Pre-marginalisation

Pixel

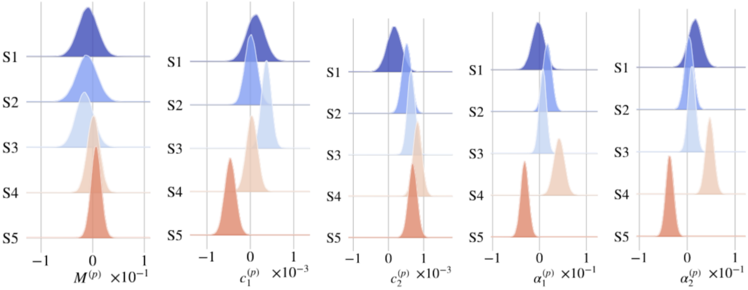
$$\epsilon_{\rm obs}(p, m; \boldsymbol{\Theta}) = \left(1 + M^{(p)}\right)\epsilon_{\rm lensed}(\boldsymbol{\Theta}) + \alpha^{(p)}\epsilon_{\rm PSF}(m) + \cancelto{0}{\beta^{(p)}\delta\epsilon_{\rm PSF}} + c^{(p)}$$

Tomographic bin

Multiplicative shear bias

PSF shear bias

Additive shear bias



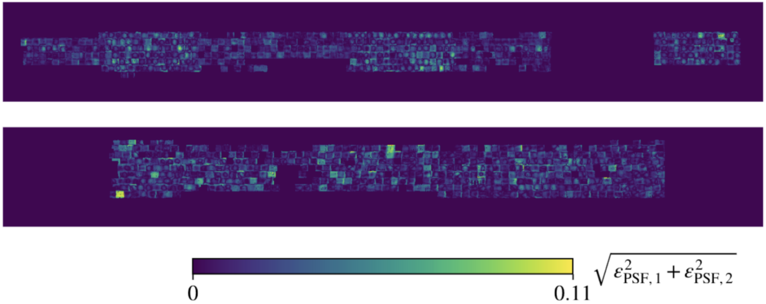
Survey Characteristics

Pixel

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

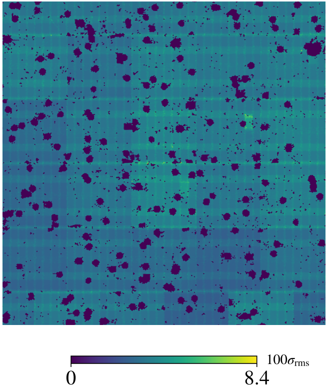
Tomographic bin

PSF shear bias



Is my model accurate?

Test Individual Contributions



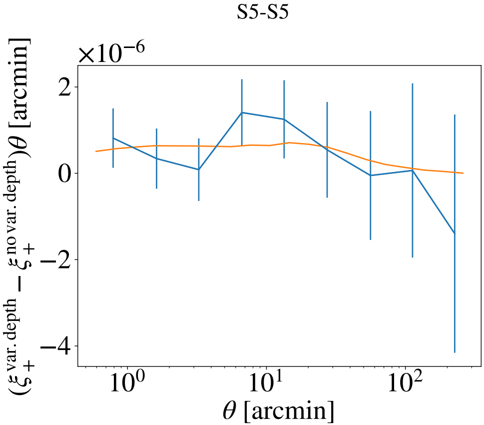


Figure: Varible Depth. Orange: theory [Heydenreich et al. 2020; arXiv:1910.11327]. Blue: forward-simulations.

Compare to Other Uncertainty Models

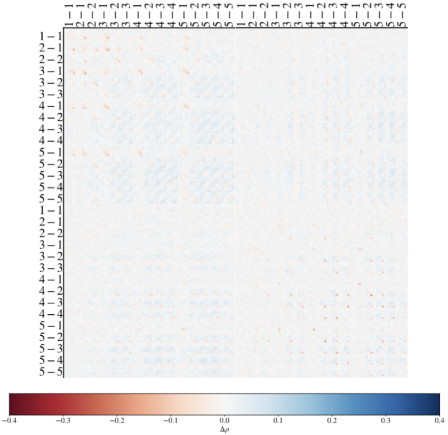
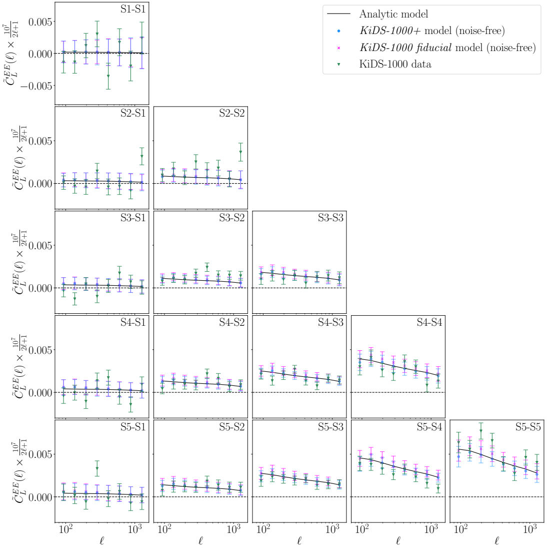


Figure: Comparison between sample and analytical covariance matrix [Joachimi et al. 2021; arXiv:2007.01844].

Can I do a blind analysis?

Measurement: Pseudo-Cls



Blinding in NLE

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

"Bayes said $p(\boldsymbol{\theta}, \boldsymbol{d}|M)$, and there was statistics."

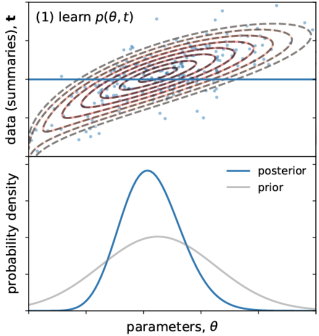
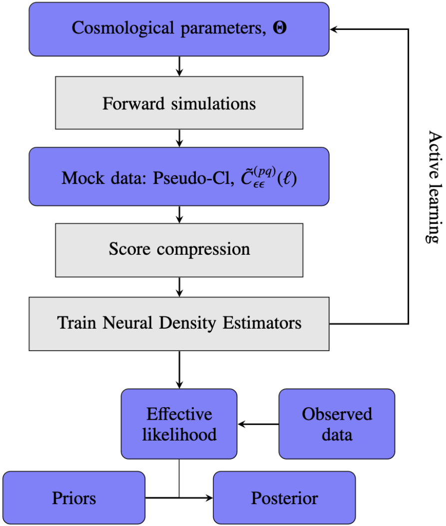


Figure: [Alsing et al. 2019]

How do I get the most out of the results of my SBI?

KiDS-SBI

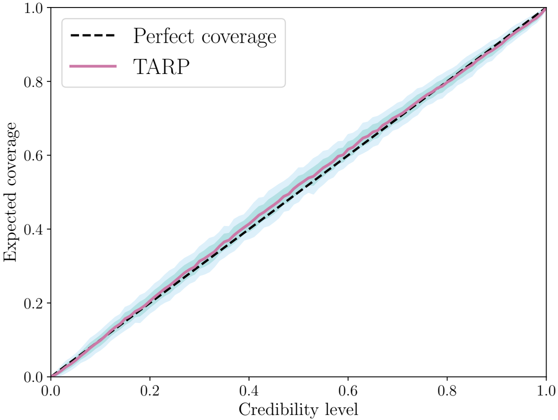


Cosmological parameters, $\boldsymbol{\Theta}$
Forward simulations
Mock data: Pseudo-Cl, $\tilde{C}_{\epsilon\epsilon}^{(pq)}(\ell)$
Score compression
Train Neural Density Estimators
Active learning
Effective likelihood
Observed data
Priors
Posterior

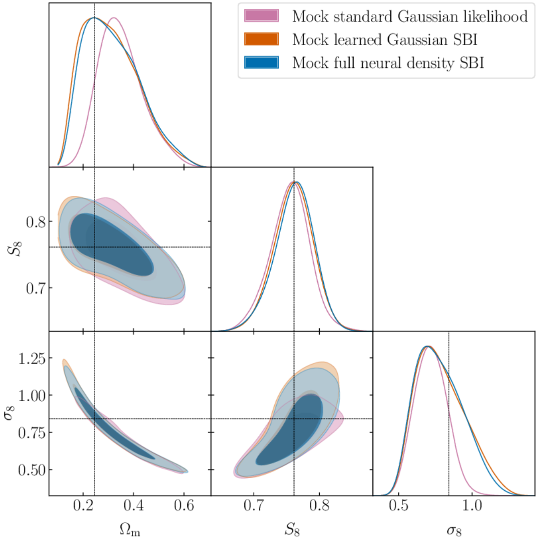
KiDS-SBI: Parameters and Priors

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

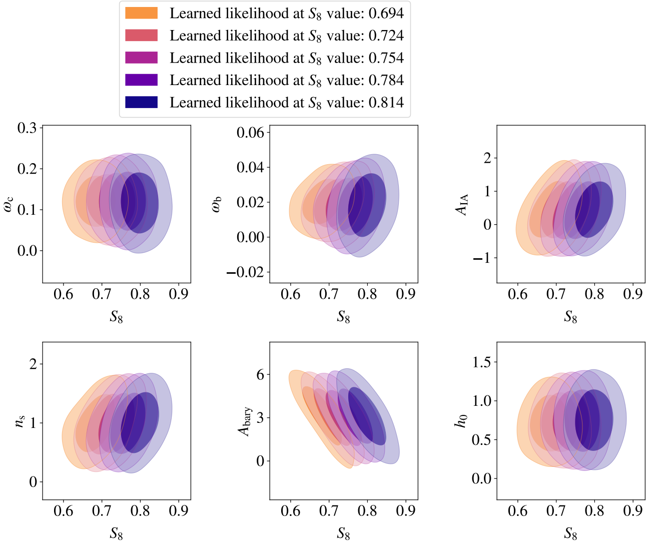
Check for Bias



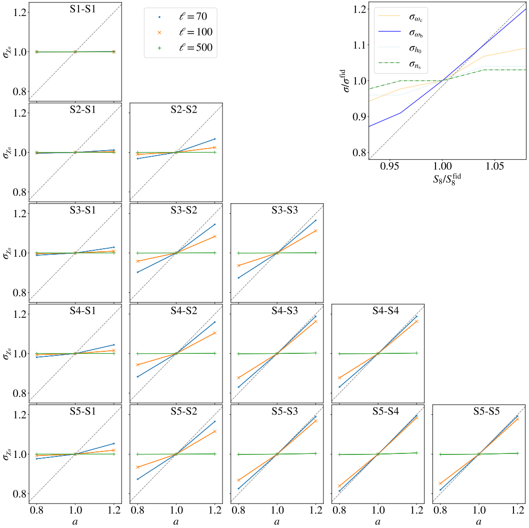
Test Different Likelihoods



Inspect the Likelihood: S_8-dependent Uncertainty



S_8-dependent Uncertainty from Theory



Play with Models and Mock Data

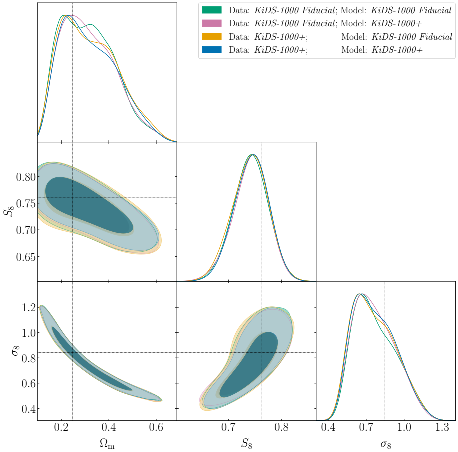


Figure: Noise-free mocks. *KiDS-1000+* model: includes anisotropic systematics. *KiDS-1000 fiducial* model: assumes isotropic systematics.

Play with Models and Mock Data

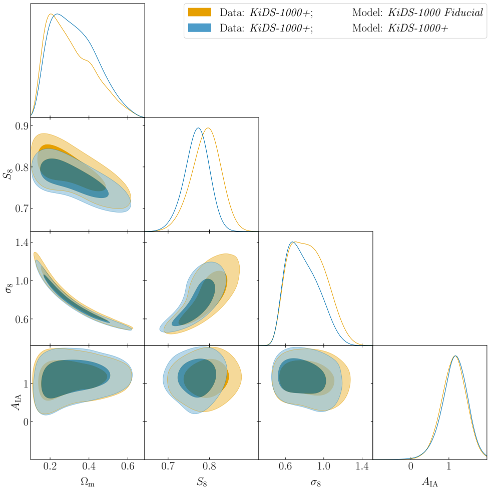


Figure: Noisy mocks.

KiDS-SBI: Conclusions

- Neglecting variable depth and shear biases can bias S_8 by approx. 5%

- Cosmic shear likelihood is consistent with a Gaussian, but its covariance is measurably cosmology-dependent (cosmic variance) $\rightarrow$ uncertainty on S_8 is higher

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 B - Compression and DELFI

- A stack of NDEs are used in DELFI

$$p(\mathbf{t}|\theta; \mathbf{w}) = \prod_{\alpha=1}^{N_{\mathrm{NDEs}}} \beta_\alpha p_\alpha(\mathbf{t}|\theta; \mathbf{w}), \tag{3}$$

- 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

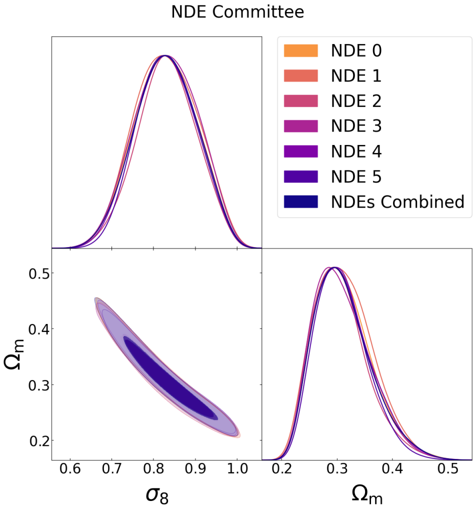
$$\mathcal{L} = \mathcal{L}_* + \delta\boldsymbol{\theta}^T \nabla \mathcal{L}_* - \frac{1}{2}\delta\boldsymbol{\theta}^T \mathbf{J}_* \delta\boldsymbol{\theta}, \tag{4}$$

$$\boldsymbol{t} = \nabla \boldsymbol{\mu}^T \boldsymbol{C}^{-1}(\boldsymbol{d} - \boldsymbol{\mu}), \tag{5}$$

- Remap this to a MLE estimate via the Fisher matrix

$$\hat{\theta} = \theta_* + \boldsymbol{F}_*^{-1} \nabla \mathcal{L}_* = \theta_* + \boldsymbol{F}_*^{-1} \boldsymbol{t}_*, \tag{6}$$

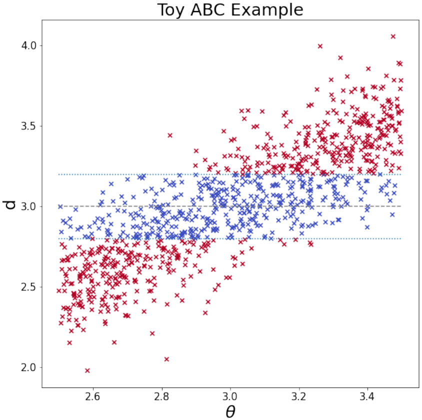
Appendices C - PyDELFI

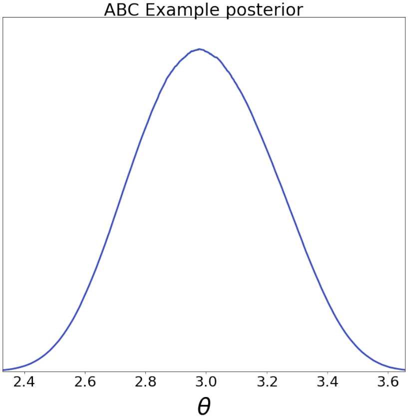


[Alsing et al. 2019]

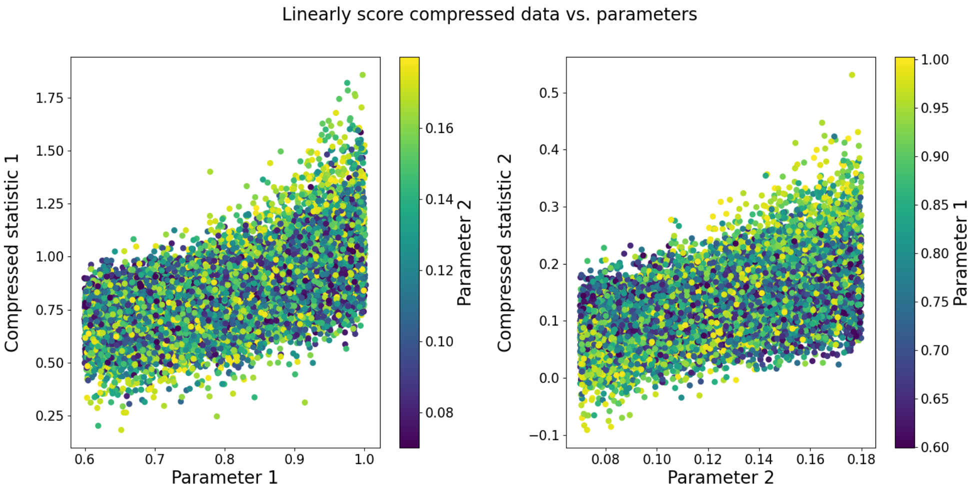
Appendices D - SBI - ABC

$$P(\boldsymbol{\theta}|\boldsymbol{d}) = \frac{P(\boldsymbol{d}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})}{P(\boldsymbol{d})} \tag{7}$$





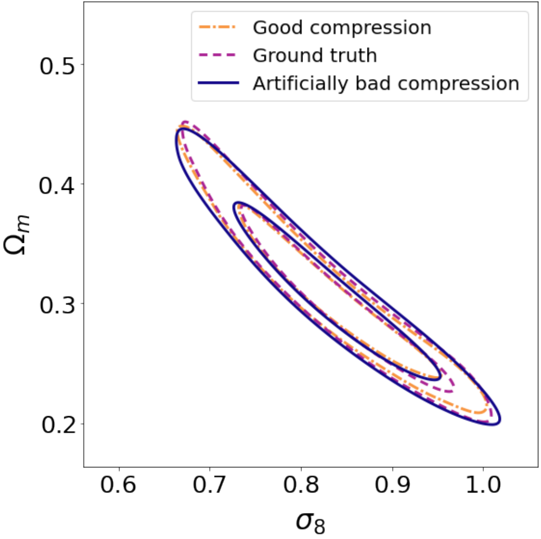
Appendices E - Score Compression



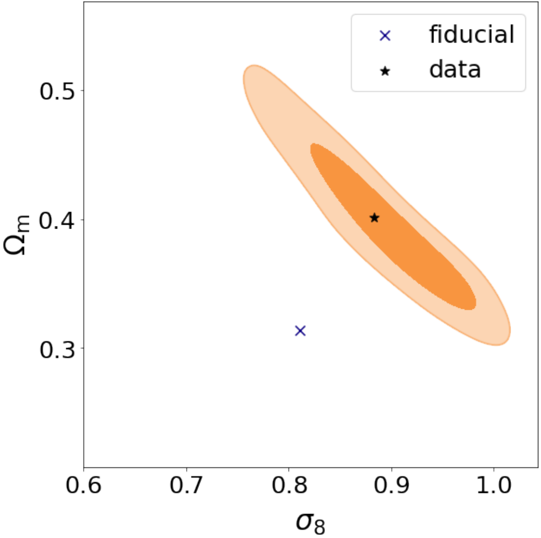
[Alsing et al. 2018]

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

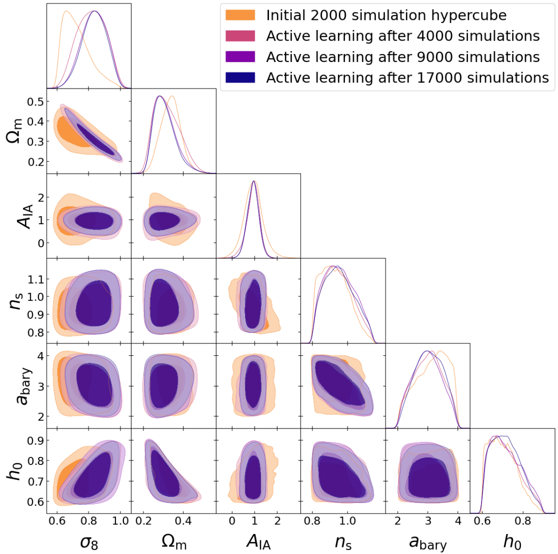
Appendices F - Sensitivity to Compression



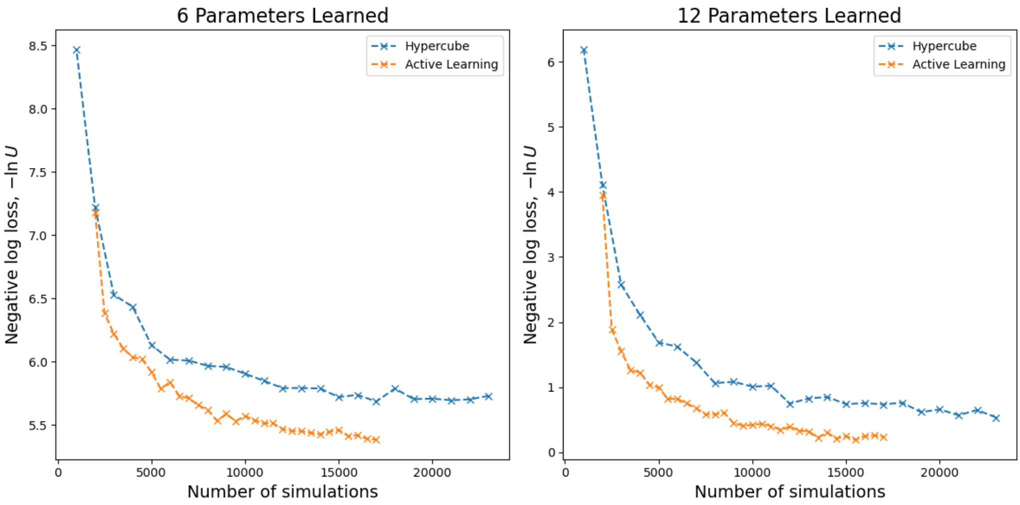
Appendix G - Sensitivity to Fiducial Cosmology



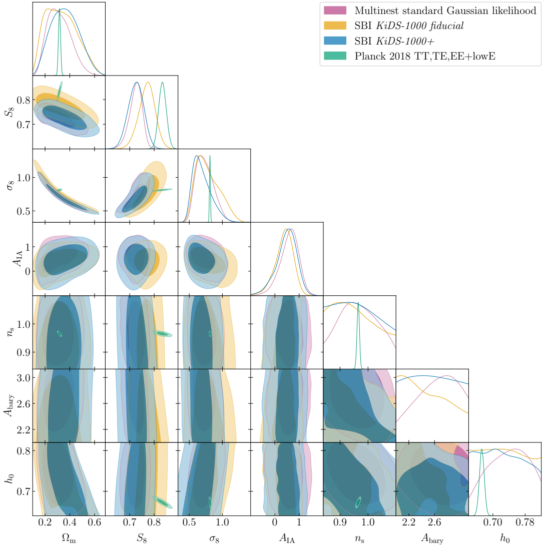
Appendix H - Active Learning



Appendix I - Simulation Number Sufficiency



Appendix J - Full Posterior



Appendix K - Signal Test