Cosmology from the integrated shear 3-point correlation function:

simulated likelihood analyses with machine-learning emulators

Zhengyangguang (Laurence) Gong (USM, LMU)

with: Anik Halder, Alex Barreira, Stella Seitz and Oliver Friedrich

https://arxiv.org/abs/2304.01187

Ecole thématique du CNRS
Future Cosmology
25.04.2023

Modelling the integrated shear 3PCF

$$\zeta_{\pm}(\alpha) = \left\langle \hat{M}_{\rm ap}(\boldsymbol{\theta}_C)\, \hat{\xi}_{\pm}(\alpha;\boldsymbol{\theta}_C) \right\rangle = \frac{1}{A_{\rm 2pt}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi} \mathcal{B}_{\pm}(\ell) J_{0/4}(\ell\alpha)$$

$$\mathcal{B}_{\pm}(\ell) = \int \mathrm{d}\chi\, \frac{q_{\kappa}^3(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{\delta}^{\rm 3D}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi\right) e^{2i(\phi_{\boldsymbol{\ell}_2} \mp \phi_{-\boldsymbol{\ell}_{12}})} U(\boldsymbol{\ell}_1) W(\boldsymbol{\ell}_2 + \boldsymbol{\ell}) W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell})$$

Compensated filter

Top-hat filter

Line-of-sight projection

3D matter bispectrum

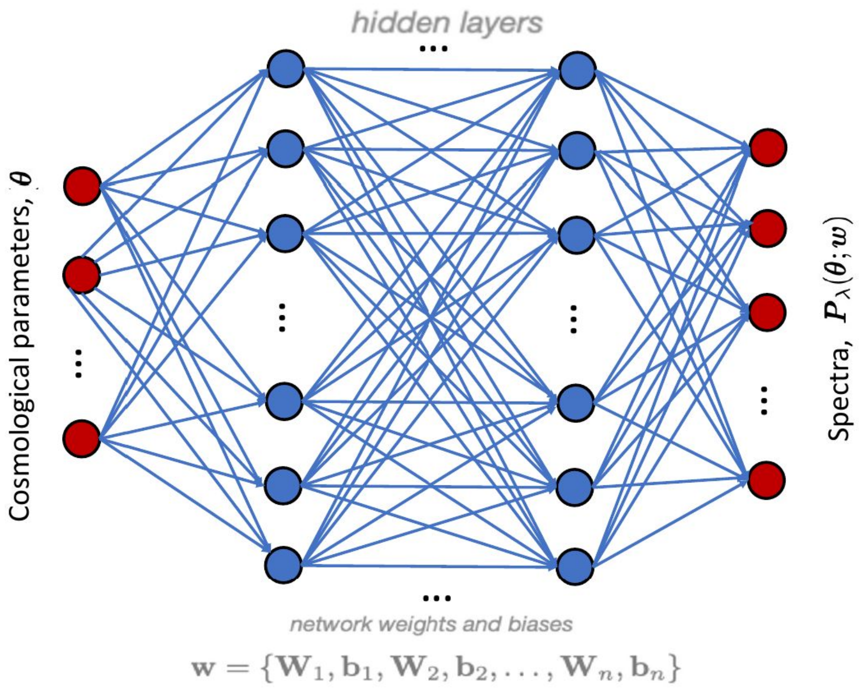
Window functions

Emulation

Halder et al. (2021)
[arxiv:2102.10177](https://arxiv.org/abs/2102.10177)
Halder and Barreira (2022)
[arxiv:2201.05607](https://arxiv.org/abs/2201.05607)

- Emulate the 4-dimensional integration which is computationally expensive
- Leaving the line-of-sight projection out of emulation: Preserving the flexibility in the systematic modelling

Emulate the integrated bispectrum using neural network (NN)



Mancini et al. (2021)
[arXiv:2106.03846]

- The emulator is constructed using the package containing a suite of fully connected NN: Cosmopower

- Cosmopower:

  training feature: $\{\Omega_m, \ln(10^{10} A_s), w_0, c_{\min}, z\}$

  training label: pre-computed spectra at 100 multipoles

  training on GPU

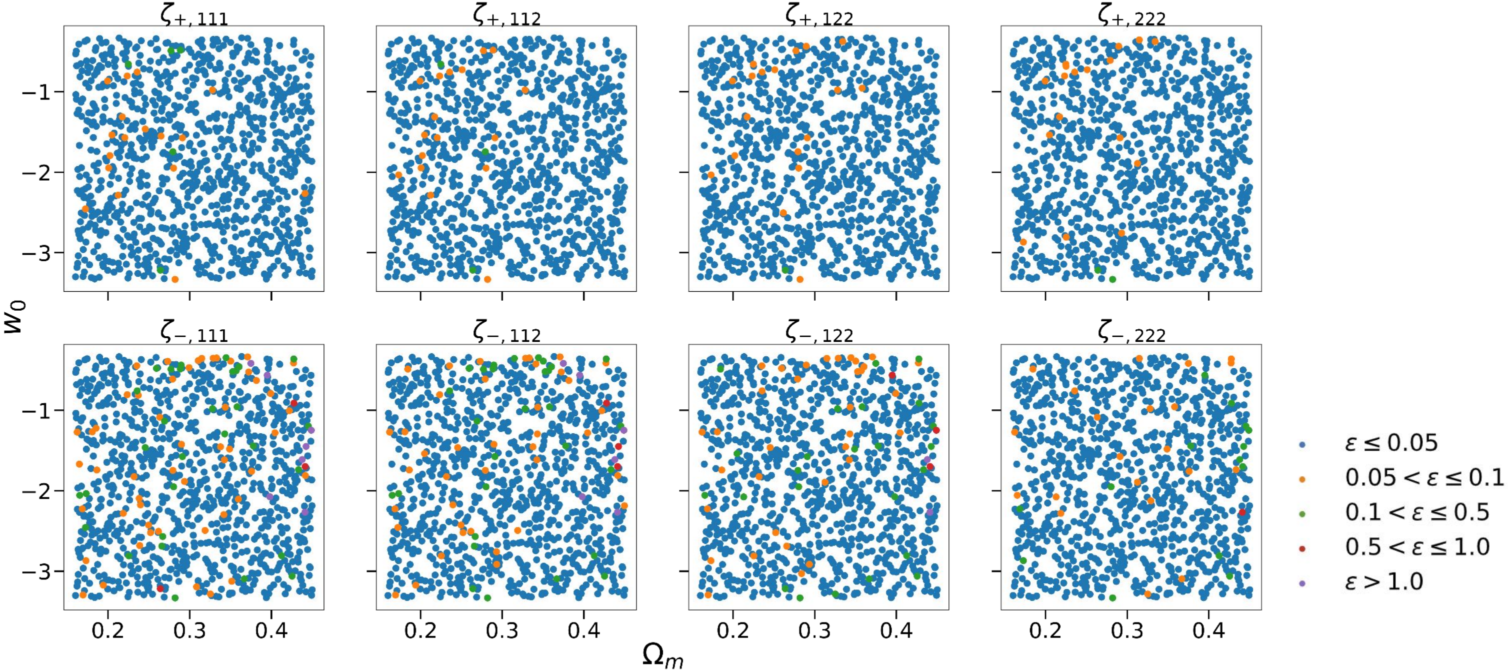
- Each emulator is for a specific filter size:

  100000 training nodes (10% validation)

  1000 testing nodes

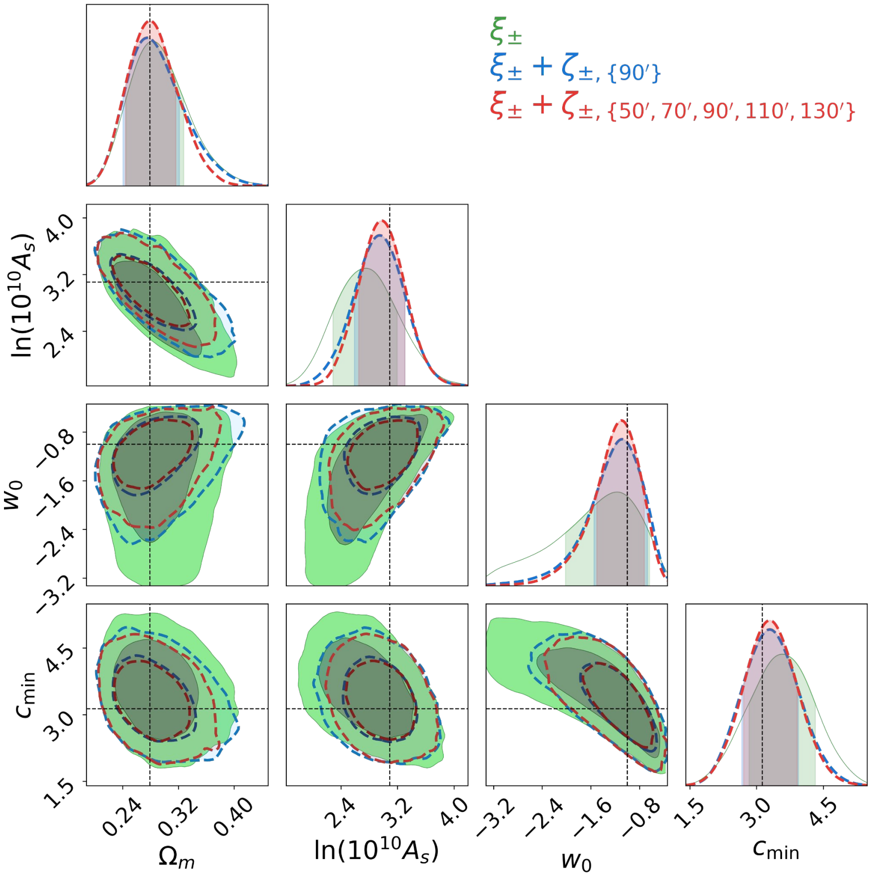
Emulation accuracy test

$$\epsilon \equiv \left| \frac{\chi^2_{\mathrm{emu},i}}{\chi^2_{\mathrm{test},i}} - 1 \right|$$



- All 8 non-redundant integrated shear 3PCF from 2 DES Y3-like tomographic bins
- Using chi2 fractional difference as the emulation accuracy metric: It describes how closely the emulators describe the log-likelihood surface w.r.t the theory model predictions

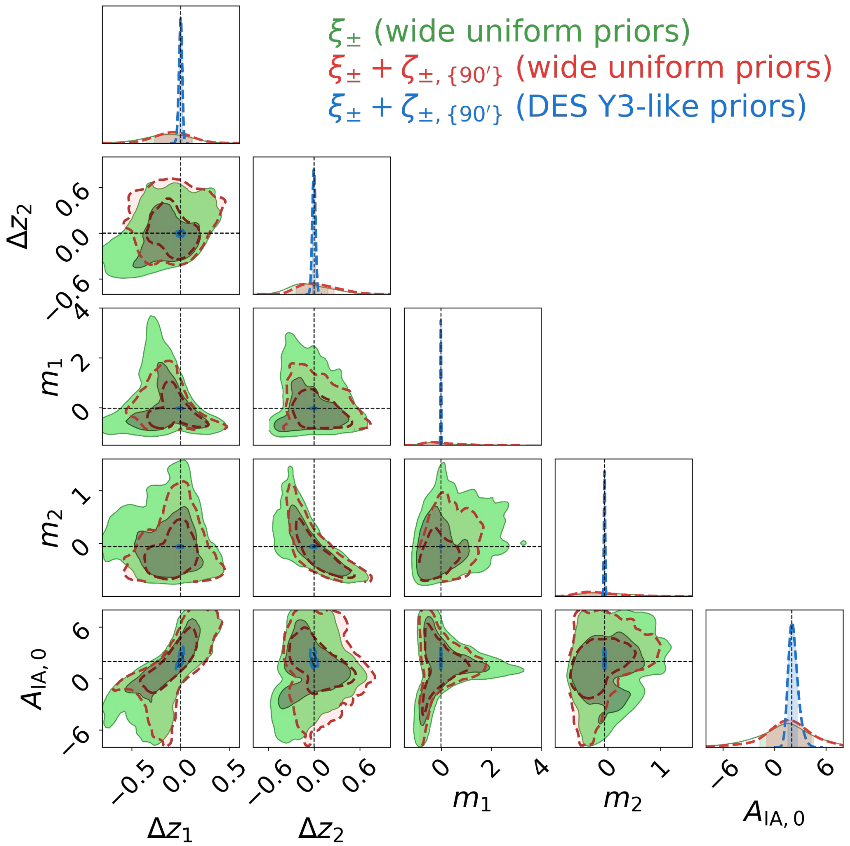
The impact of the aperture size



- We train 5 emulators for integrated shear 3PCF with different filter sizes: {50', 70', 90', 110', 130'}
- Marginalized over systematic parameters – photo-z, shear bias and intrinsic alignment (NLA) parameters

Aperture sizes (arcmin)	Ω_m	$\ln\left(10^{10}A_s\right)$	w_0	c_min
50	1.2%	9.0%	18.1%	4.8%
70	1.2%	16.9%	31.9%	11.6%
90	**3.7%**	**20.2%**	**38.4%**	**15.1%**
110	1.2%	19.1%	34.1%	11.0%
130	1.2%	16.9%	32.6%	12.3%
{50, 70, 90}	2.5%	24.7%	39.1%	15.8%
{50, 90, 130}	3.7%	23.6%	41.3%	16.4%
{70, 90, 110}	6.2%	25.8%	39.1%	15.1%
{90, 110, 130}	8.6%	25.9%	42.8%	15.8%
{50, 70, 90, 110, 130}	**12.4%**	**28.1%**	**44.9%**	**19.9%**

The impact of systematics and their modelling

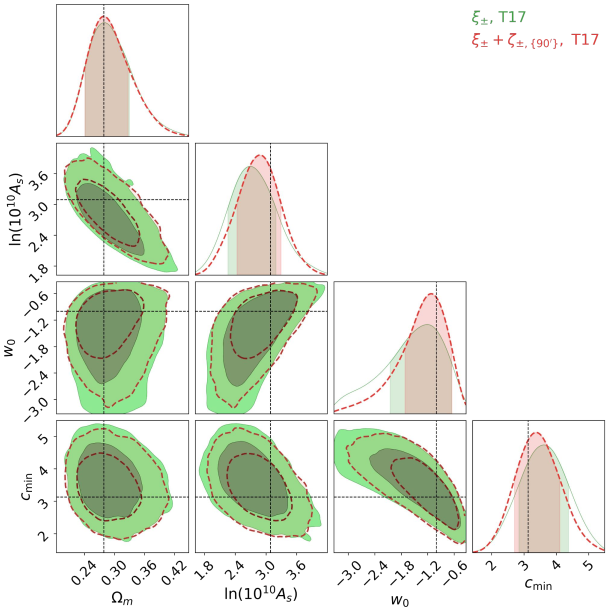


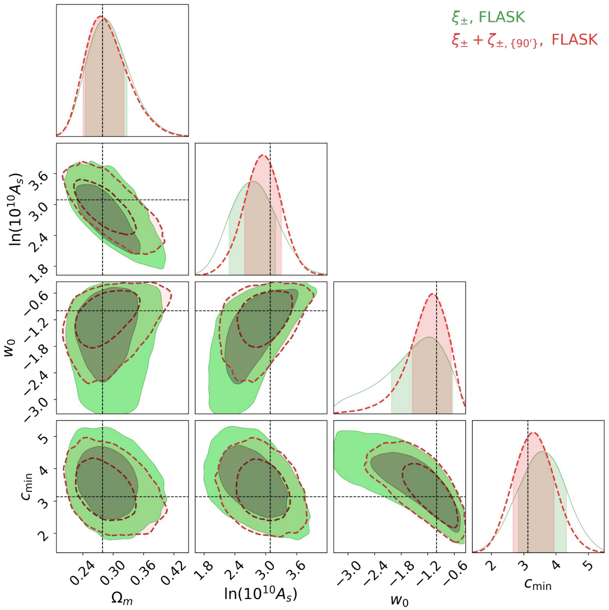
- Idea: 2-point and 3-point statistics depend differently on systematic parameters → Possible self-calibration of systematic parameters that can reduce the need for external calibration data sets

- There is indeed a visible level of systematics self-calibration from combining ξ± with ζ±,

- Quantitatively it is in contrast with the results reported in the work https://arxiv.org/abs/2010.00614

The impact of different covariance estimates





- FLASK-based covariance may not be suitable to cosmological constraints using 3-point cosmic shear information
- Real-data analyses may require using more expensive N -body simulations, or calculating the covariance matrix analytically

Covariance type	Ω_m	$\ln\left(10^{10}A_s\right)$	w_0	$c_{\min}$
FLASK (lognormal)	3.7%	20.2%	38.4%	15.1%
T17 (N-body simulations)	3.5%	8.8%	26.1%	8.7%

Summary

- Our analysis pipeline is accurate and able to yield unbiased parameter constraints from our N -body simulation DES Y3-like data vectors.

- Aperture size 90 arcmin is what results in the largest information gain from $\zeta_{\pm}$. The combination of several filter sizes can improve the constraints further but at the cost of dealing with a larger data vector and covariance matrix.

- We do not find significant improvements of systematic constraints in combined $\xi_{\pm} + \zeta_{\pm}$ analyses; i.e. the mitigation of systematic effects still requires prior calibration from external data

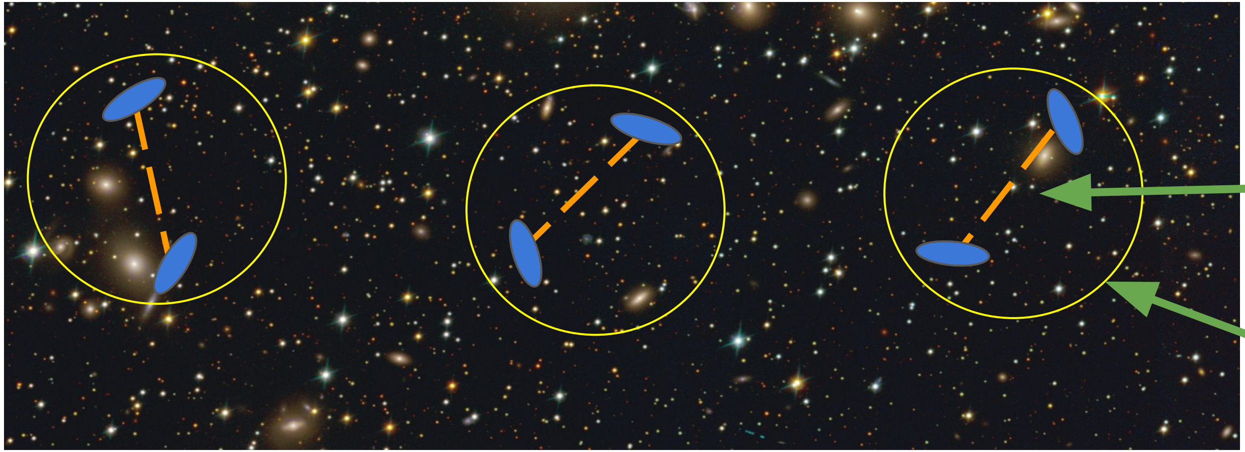
- Lognormal realizations might not provide reliable estimates of the $\zeta_{\pm}$ covariance matrix.

- The next step is to apply this higher-order statistic pipeline to DES Y3 data and extract cosmological information

Thank you!

Additional slides

Integrated shear 3-point correlation functions



Position-dependent shear 2-pt correlation in top-hat aperture

$\hat{\xi}_{\pm}(\theta; \boldsymbol{\theta_C})$

Aperture mass

$M_{\mathrm{ap}}(\boldsymbol{\theta_C})$

Weighted tangential shear inside the filter

$$\zeta_{\pm}(\theta) \equiv \langle M_{\mathrm{ap}}(\boldsymbol{\theta_C})\, \hat{\xi}_{\pm}(\theta; \boldsymbol{\theta_C}) \rangle$$

Halder et al. (2021)
arxiv:2102.10177

→ Directly observable higher-order statistic of the cosmic shear field

→ Probes the line-of-sight projection of the 3D matter bispectrum

Data preparation and pre-processing

	Prior range
Cosmological parameters (emulated)	
Ω_{m}	$U\,[0.16, 0.45]$
$\ln(10^{10} A_s)$	$U\,[1.61, 4.20]$
w_0	$U\,[-3.33, -0.33]$
Baryonic feedback parameter (emulated)	
$c_{\min}$	$U\,[1.0, 5.5]$
Systematic parameters (not emulated)	
Δz_1	$\mathcal{N}(0.0, 0.023)$
Δz_2	$\mathcal{N}(0.0, 0.020)$
m_1	$\mathcal{N}(0.0261, 0.012)$
m_2	$\mathcal{N}(-0.061, 0.011)$
$A_{\mathrm{IA},0}$	$U\,[-5.0, 5.0]$
α_{IA}	0 (fixed)

With the additional emulated parameter redshift z between 0.0 and 2.0

- Emulation prior:

  Too wide: 1. a waste of training data;
  2. Labels can experience numerical instability or give unusual predictions that form prominent outliers

  Too narrow: Parameter inference will be dominated by priors

- Scale primordial power spectrum amplitude logarithmically;
  Scale the training labels: integrated bispectrum and matter power spectrum with log10

Including other weak lensing systematics

- **Photometric redshift uncertainty**

$$n^i(z) \rightarrow n^i(z - \Delta_z^i)$$

- **Shear calibration** (bias from shear measurement pipeline)

$$\zeta_{\pm,\mathrm{obs}}^{ijk}(\alpha) = (1+m_i)(1+m_j)(1+m_k)\zeta_{\pm,\mathrm{true}}^{ijk}(\alpha)$$

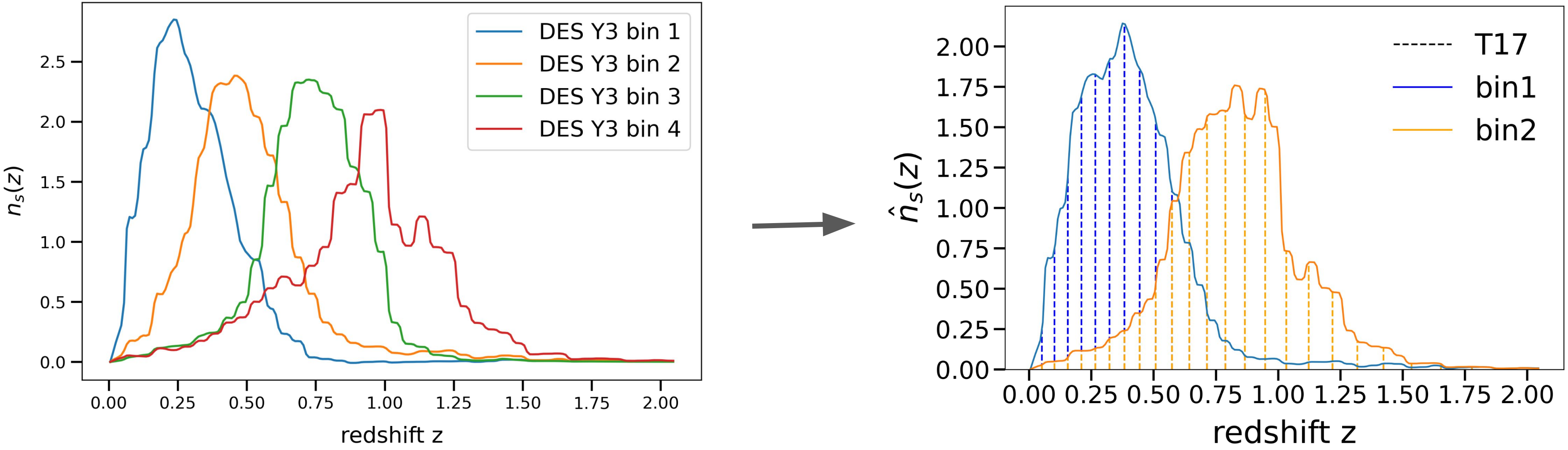
- **Intrinsic alignment** (non-linear linear alignment NLA model)

$$q_\kappa^i(\chi) \longrightarrow q_\kappa^i(\chi) - A\left(z\left(\chi\right)\right) \frac{n_\kappa^i(z(\chi))}{\bar{n}_\kappa^i} \frac{dz}{d\chi}$$

$$A(z) = A_{\mathrm{IA},0} \left(\frac{1+z}{1+z_0}\right)^{\alpha_{\mathrm{IA}}} \frac{C_1 \rho_{\mathrm{m},0}}{D(z)}$$

- We do not include these components in the emulation so that the flexibility enables others to adopt different models

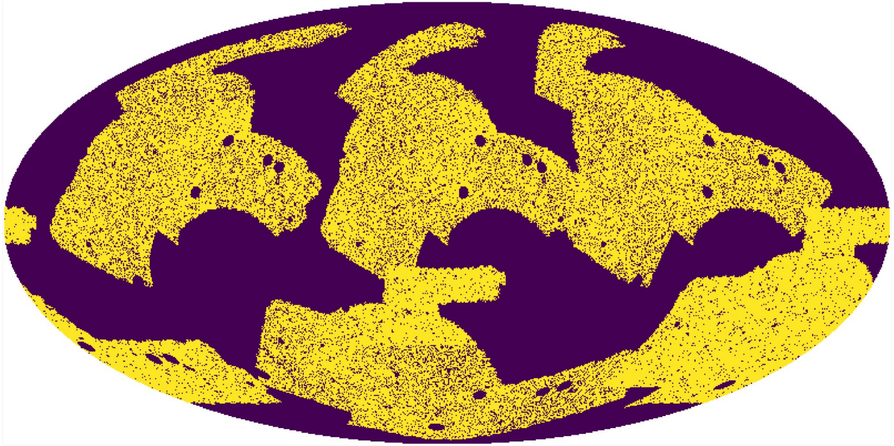
Covariance estimation I



- We merge 4 DES Y3 source redshift bins into 2 via a weighted summation
- Increase the signal-to-noise ratio of the integrated shear 3PCF

Covariance estimation II

5 rotated DES Y3-like footprints

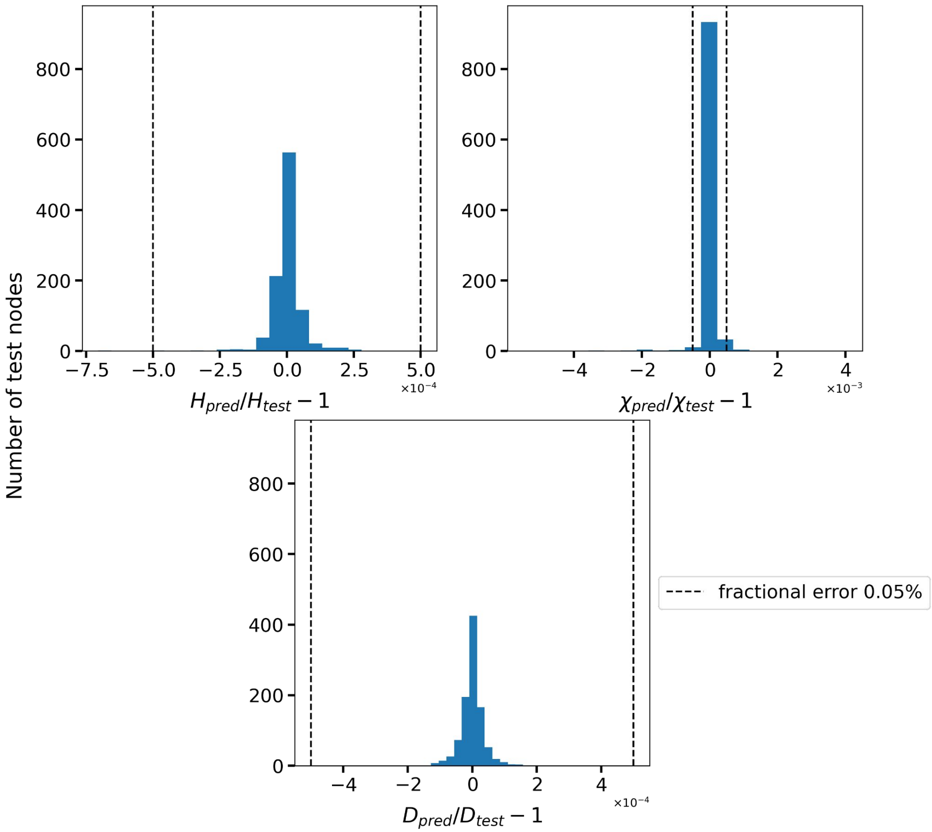


- Superimpose the DES Y3 footprint onto the full sky simulation map and rotate it to five non-overlapping locations
- Add shape noise following the equation:

$$\gamma_{\rm pix} = \gamma_{\rm noise} + \gamma_{\rm sim} = \frac{\sum_{j=1}^{N} \omega_j \gamma_{j,{\rm DES}} \exp(i\phi_j)}{\sum_{j=1}^{N} \omega_j} + \gamma_{\rm sim}$$

- Select mass apertures that have enough number of valid pixels
- Estimate data covariance from both N-body T17 simulation and FLASK log-normal maps

Including other weak lensing systematics



- The emulator is constructed using the package exploiting Gaussian Process: GPflow (de G. Matthews et al (2016) arXiv:1610.08733)
- GPflow:

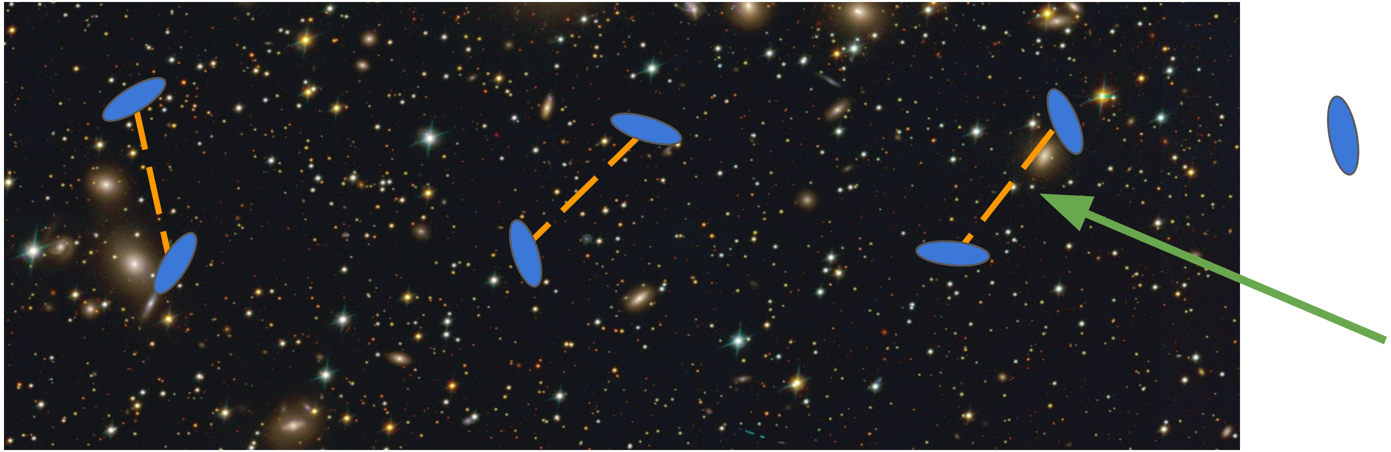
  training feature: $\{\Omega_{\rm m}, A_s, w_0, z\}$

  training label: H, chi or D (growth factor)

  training on GPU
- 10000 training nodes

  1000 testing nodes

Cosmic shear 2-point correlation functions



Background source galaxy ellipticity

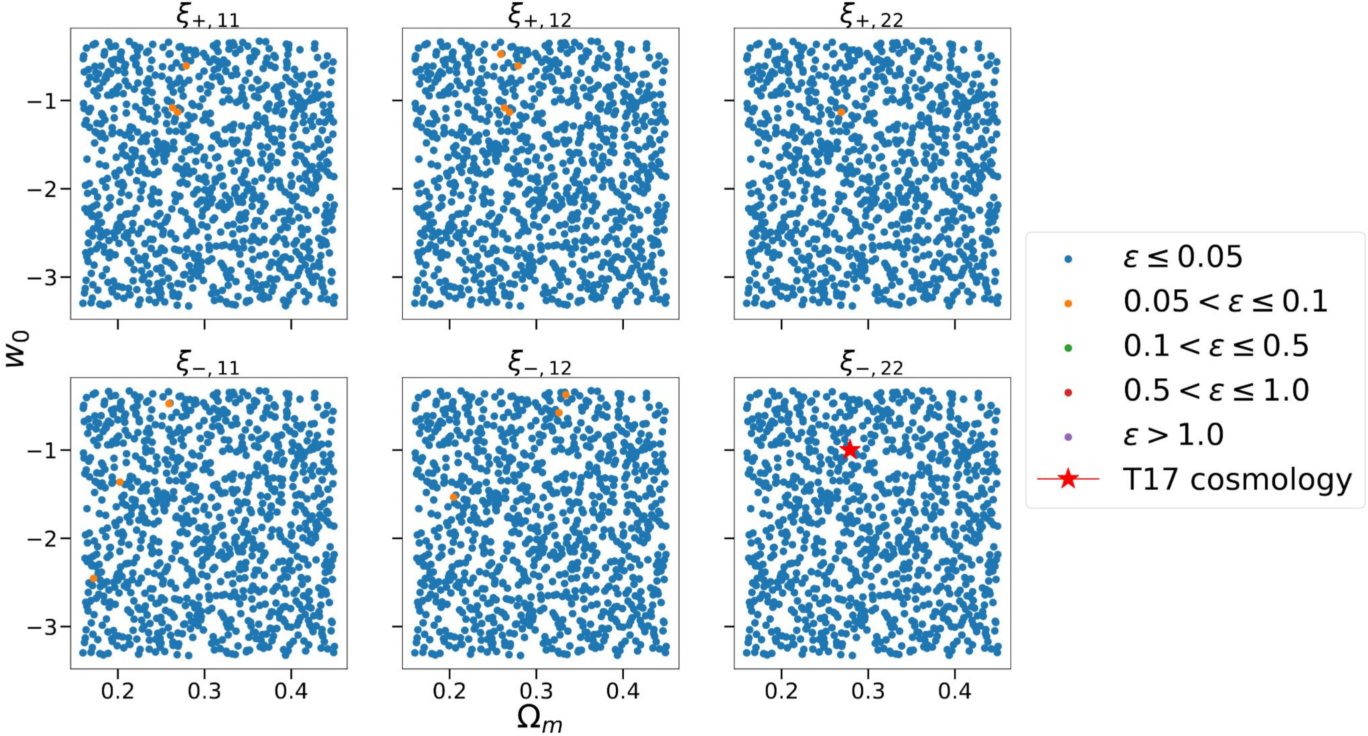
$$\hat{\xi}_{\pm}(\theta)$$

Shear 2-point correlation functions

→ Probes the line-of-sight projection of the 3D matter power spectrum

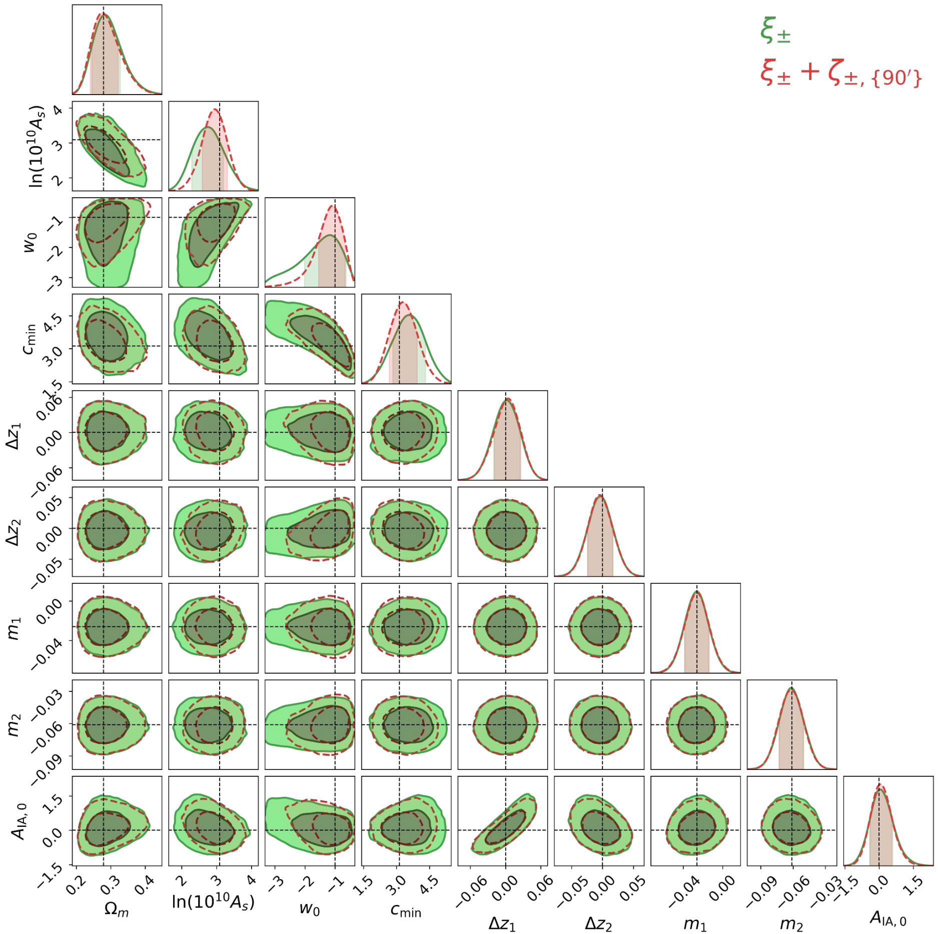
→ But cosmic shear is a non-Gaussian field with information beyond 2-point correlations!

Emulation for shear 2PCF



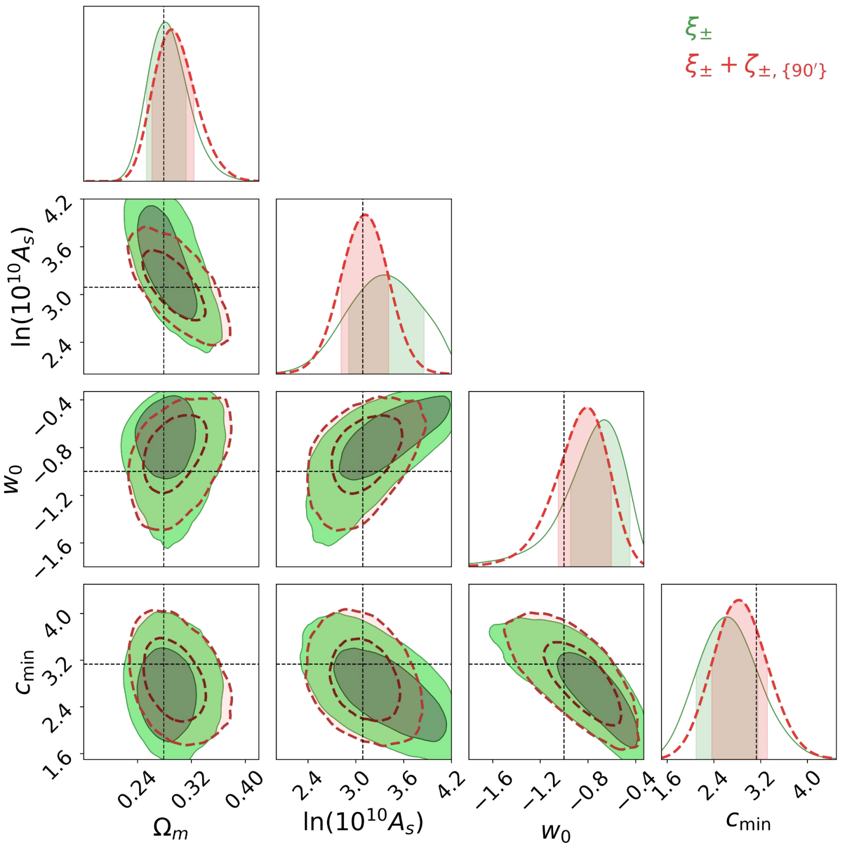
- All 6 shear 2PCFs from 2 DES Y3 tomographic bins

Simulated DESY3 MCMCs with 2PCF and integrated 3PCF



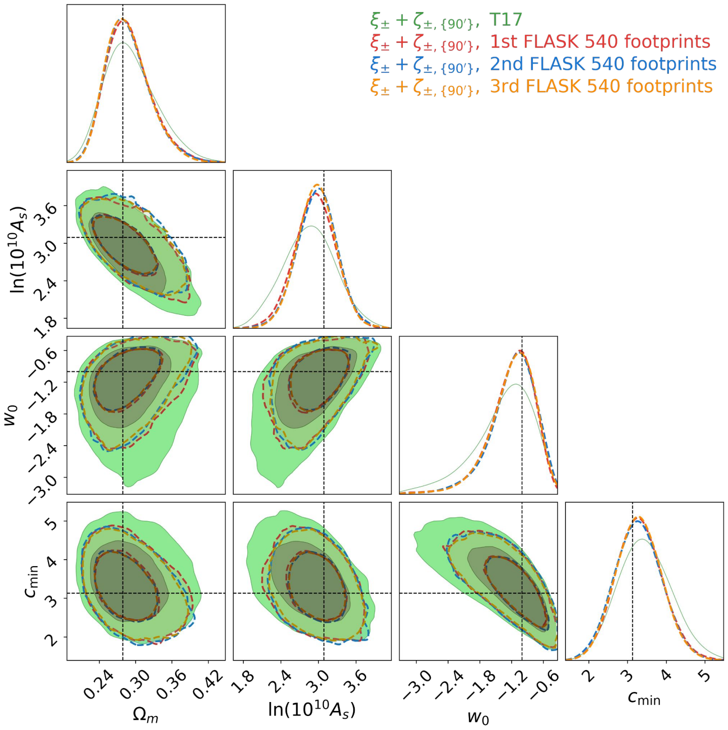
- Green: shear 2PCF only
  Red: shear 2PCF & integrated shear 3PCF (*for a single filter size*)
- MCMC on GPU: using emcee affine invariant sampler and sample **million points in ~ 1 hour**
- The inferred systematic parameter covariance is dominated by the corresponding DES Y3 gaussian priors

Validation on the T17 cosmic shear maps

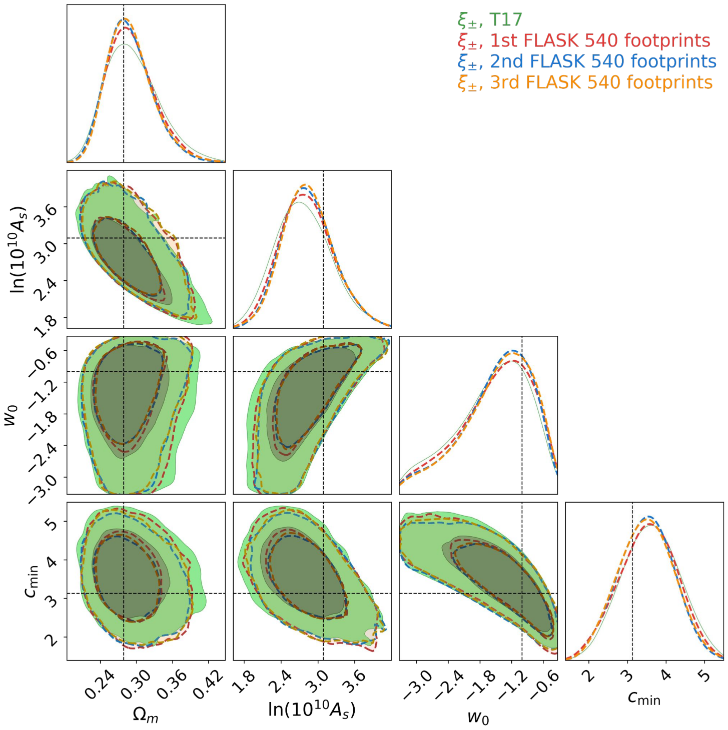


- The data vector comes from the average over 540 DES Y3-like footprints on T17 shear maps (Takahashi et al (2017), arXiv:1706.01472)
- The data covariance matrix is estimated from 1500 DES Y3-like footprints on FLASK log-normal shear maps
- The inferred parameter covariance is not biased from the fiducial T17 cosmological parameter values

T17 vs. FLASK



$\xi_\pm + \zeta_{\pm,\{90'\}}$, T17
$\xi_\pm + \zeta_{\pm,\{90'\}}$, 1st FLASK 540 footprints
$\xi_\pm + \zeta_{\pm,\{90'\}}$, 2nd FLASK 540 footprints
$\xi_\pm + \zeta_{\pm,\{90'\}}$, 3rd FLASK 540 footprints
$\ln(10^{10}A_s)$
w_0
$c_{\rm min}$
Ω_m



$\xi_\pm$, T17
$\xi_\pm$, 1st FLASK 540 footprints
$\xi_\pm$, 2nd FLASK 540 footprints
$\xi_\pm$, 3rd FLASK 540 footprints
$\ln(10^{10}A_s)$
w_0
$c_{\rm min}$
Ω_m

Emulation of integrated shear 3PCF and the modelling of systematic effects

Gong, Halder, Barreira, Seitz & Friedrich 2023 (arxiv:2304.01187)

Results of simulated likelihood analyses

- Validation on the T17 cosmic shear maps
- The impact of the aperture size
- The impact of systematics and their modelling
- The impact of different covariance estimates