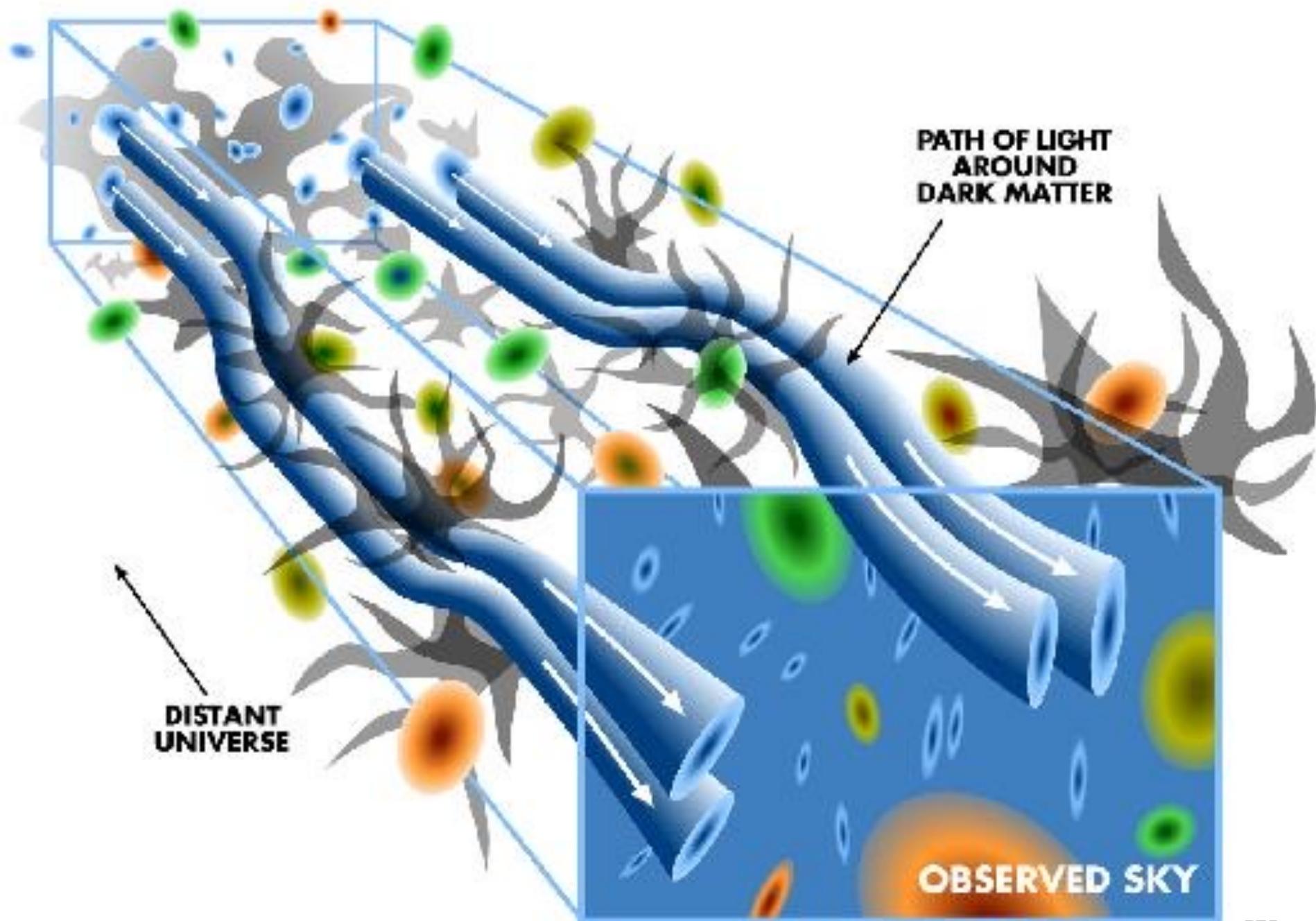


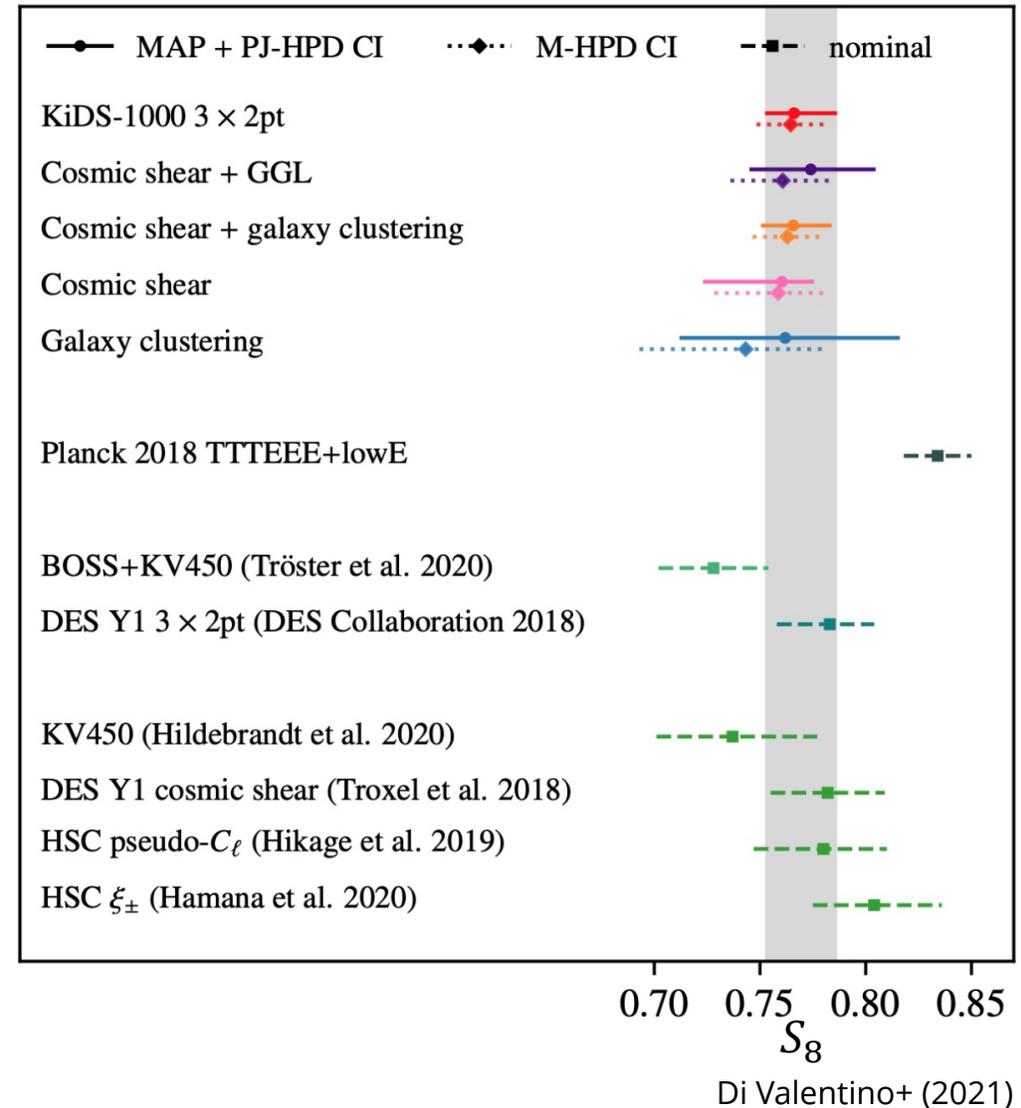
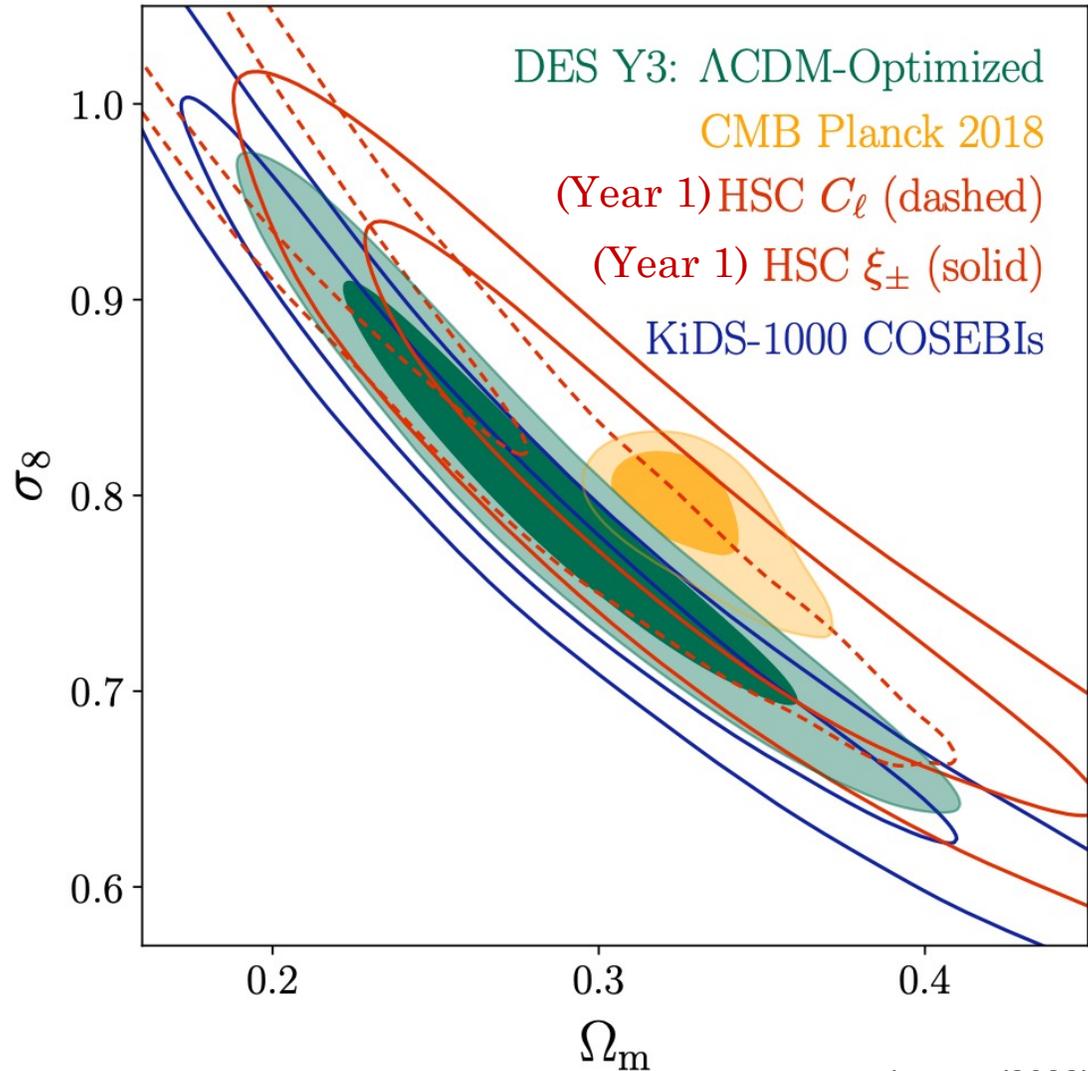
HSC Year 3: Cosmology Constraints and Systematics Challenges

Roohi Dalal, Andrina Nicola, Michael Strauss, and the HSC Weak
Lensing Working Group



S8 tension?

$$S_8 = \sigma_8 \sqrt{\Omega_m / 0.3}$$



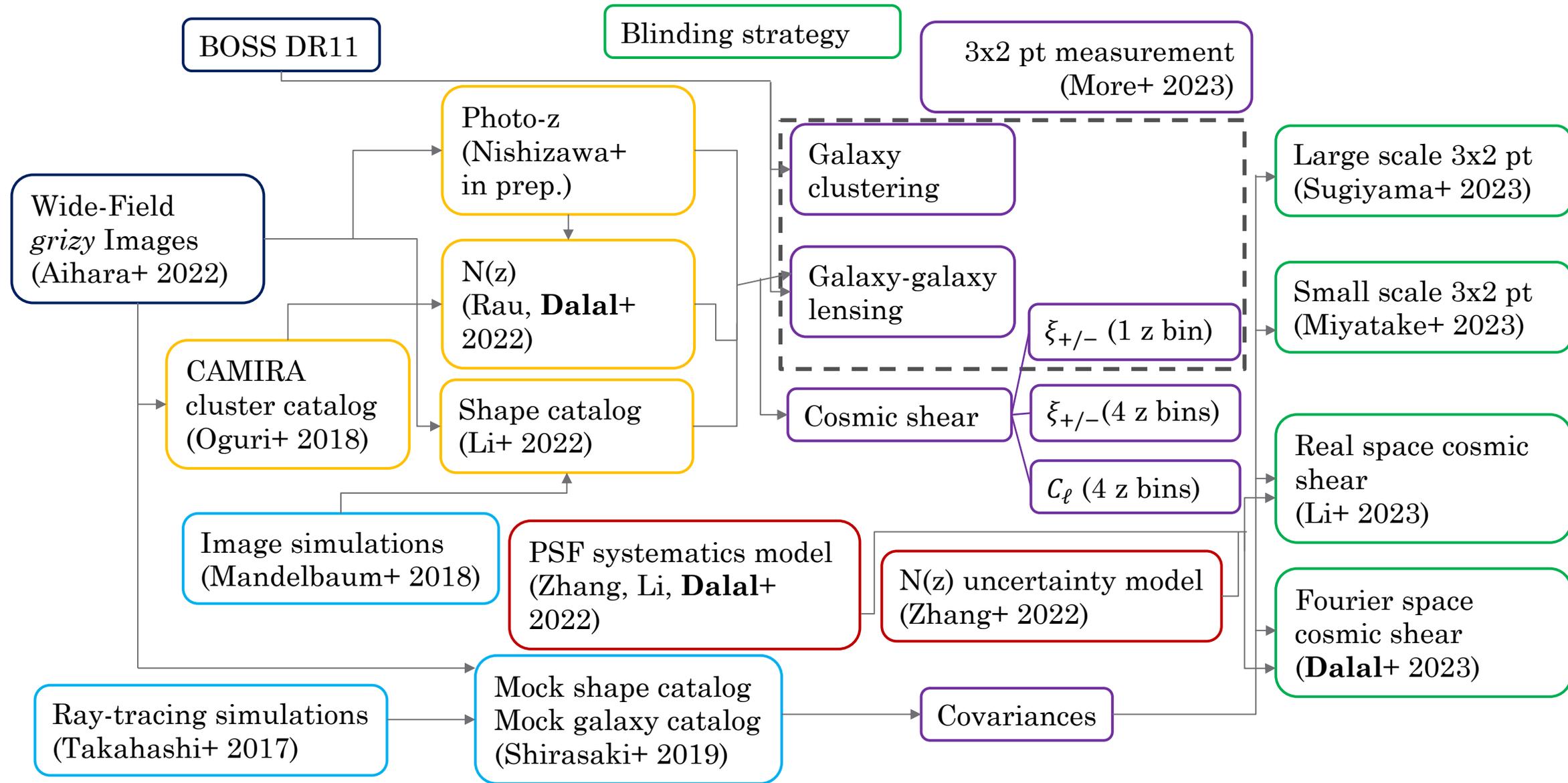
Hyper Suprime-Cam Survey

- Imaging survey using the 8.2 m Subaru telescope.
- 1.5 degree FOV
- 5 band survey (*grizy*)
- Depth ~ 26 mag
- *i*-band median seeing: 0.6 arcsec
- Y3 data release covers ~ 450 deg²



HSC COSMOS Ultra-Deep Field Aihara+ (2022)

Images → Measurement → Cosmology



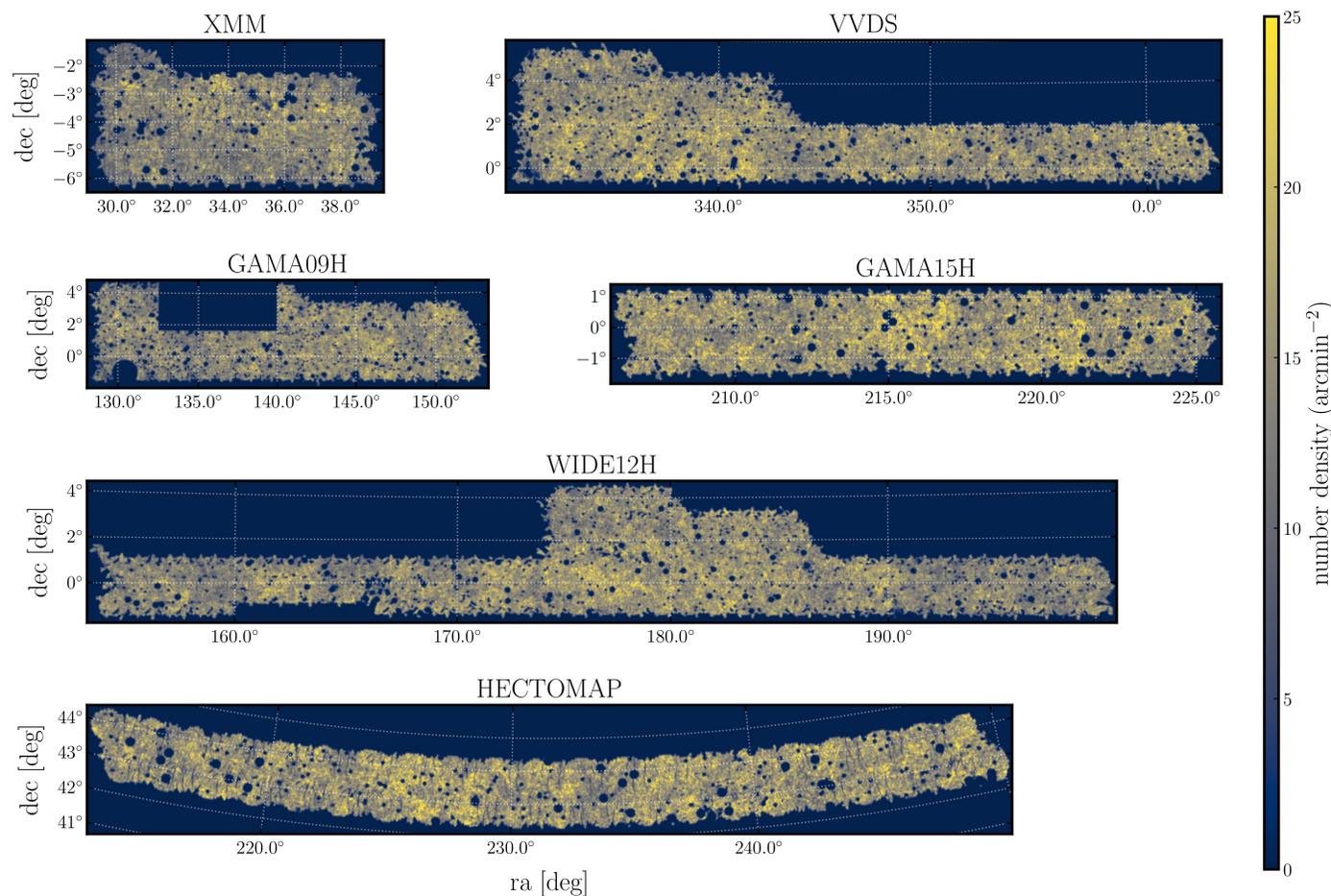
Data
Simulations
Intermediate data products
Data vectors
Modeling choices
Cosmological analyses

Y3 Shape Catalog (Li+ 2022)

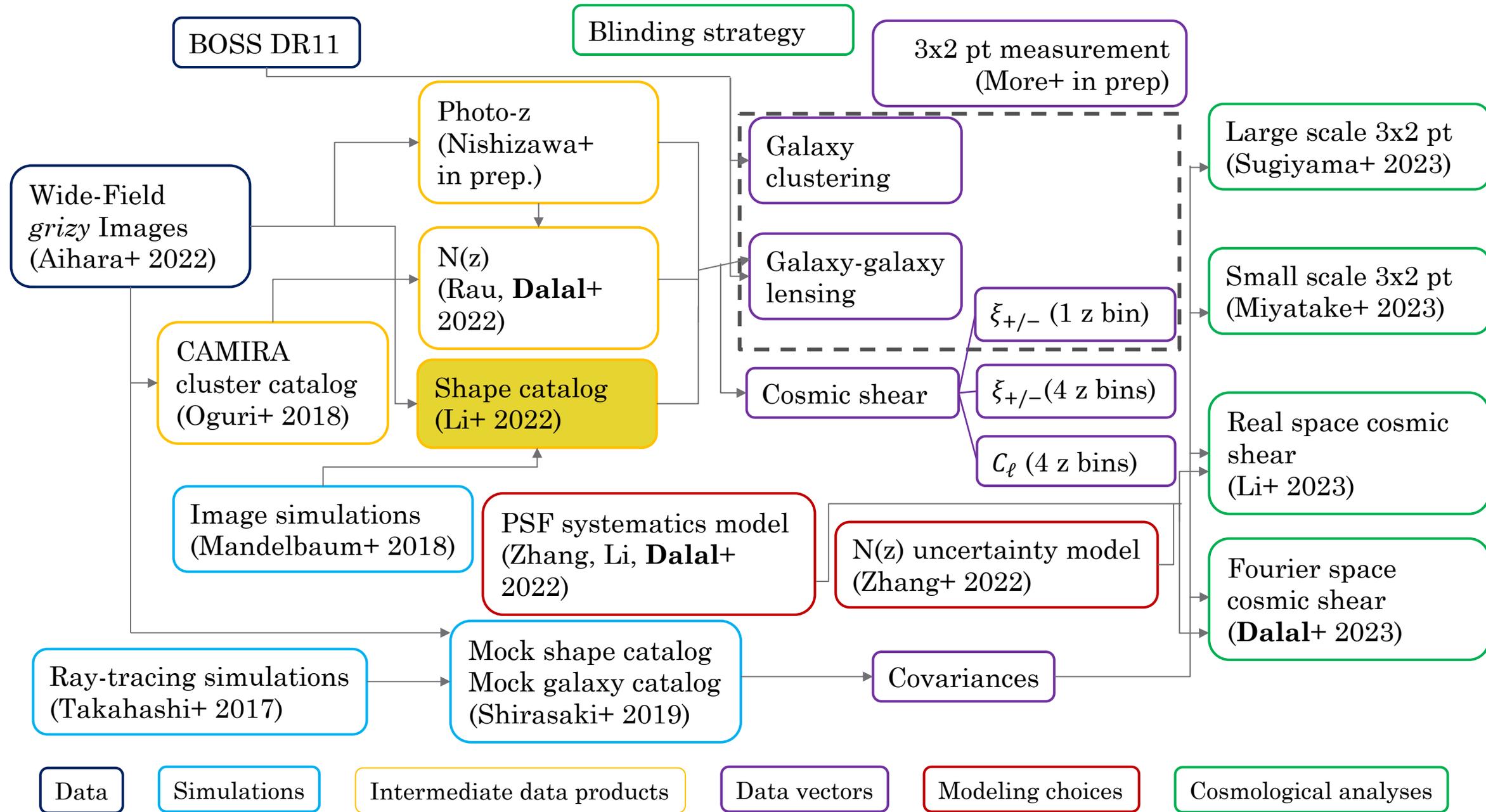


Led by Xiangchong Li (CMU)

- Area: 416 deg².
- Effective galaxy number density: 15 arcmin⁻².
- Magnitude-limited sample: $i < 24.5$ mag.
- Calibration of shape measurements using image simulations based on COSMOS HST images.
- Systematic uncertainties in shear estimation are below 1%.



Images → Measurement → Cosmology



N(z) inference (Rau, Dalal+ 2022)

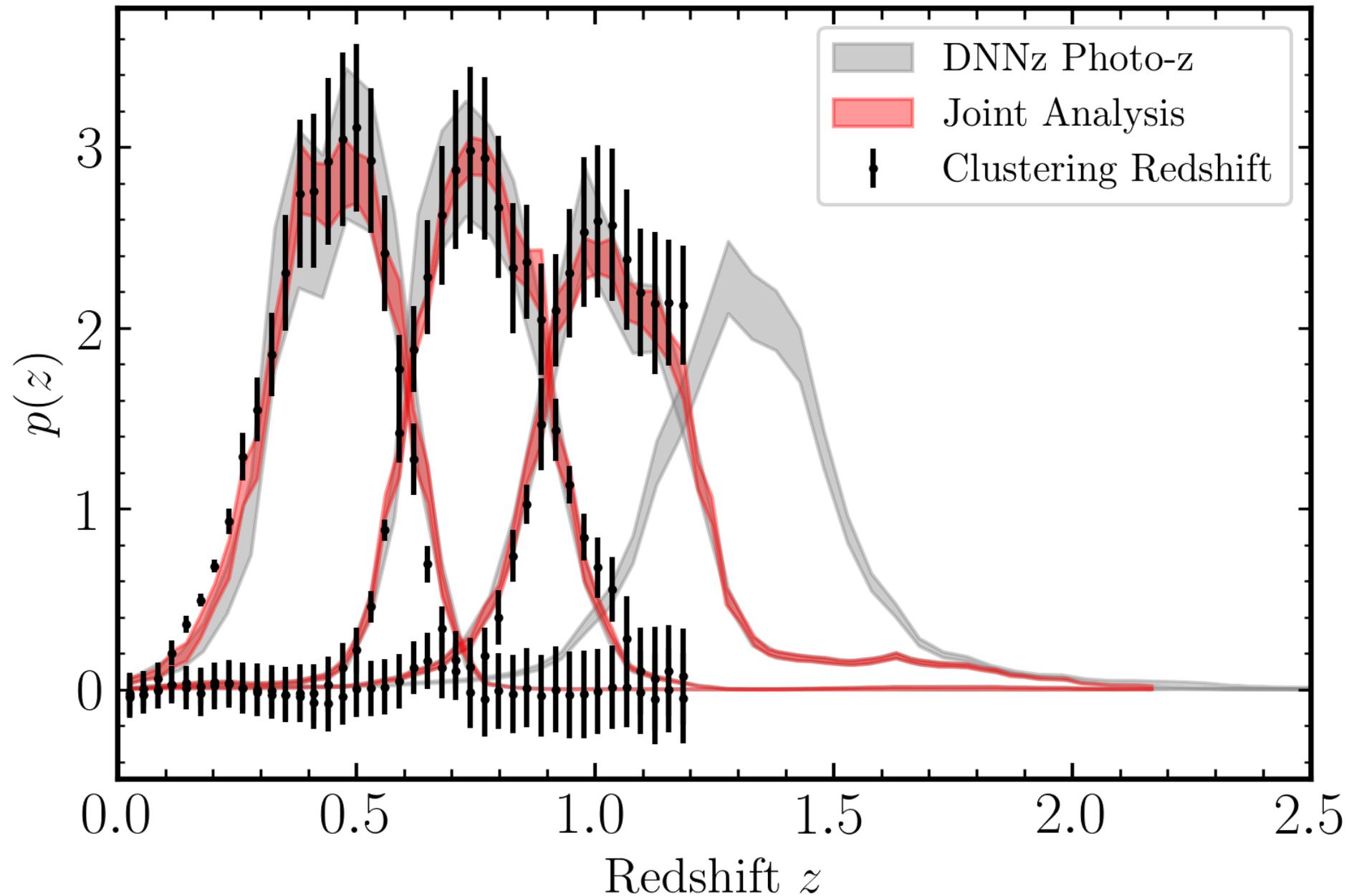


Led by Markus
Michael Rau
(Argonne)

3 photo-z codes for HSC Y3:

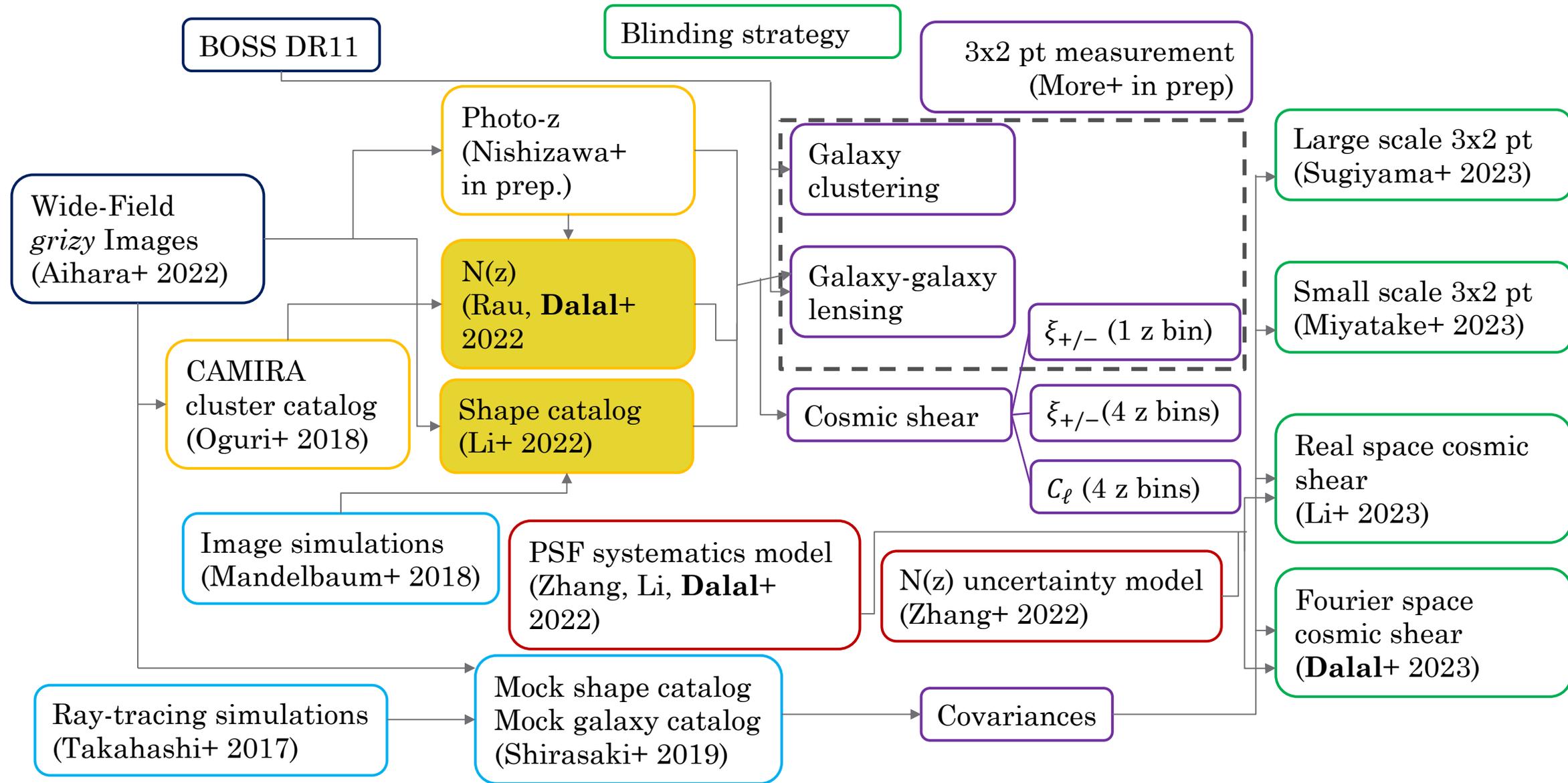
- Mizuki: Spectral Energy Distribution fitting (Tanaka 2015).
- DNNz: Neural network-based photo-z conditional density estimation (Nishizawa+ in prep).
- DEMPz: Machine learning code that uses conditional density estimation (Hsieh & Yee 2014).

Galaxies placed into 4 tomographic bins between $0.3 < z < 1.5$, based on best estimate of photo-z from DNNz.



Infer sample redshift distribution by combining photometric information with spatial clustering information from an LRG sample from the CAMIRA cluster finding algorithm.

Images → Measurement → Cosmology



Data
Simulations
Intermediate data products
Data vectors
Modeling choices
Cosmological analyses

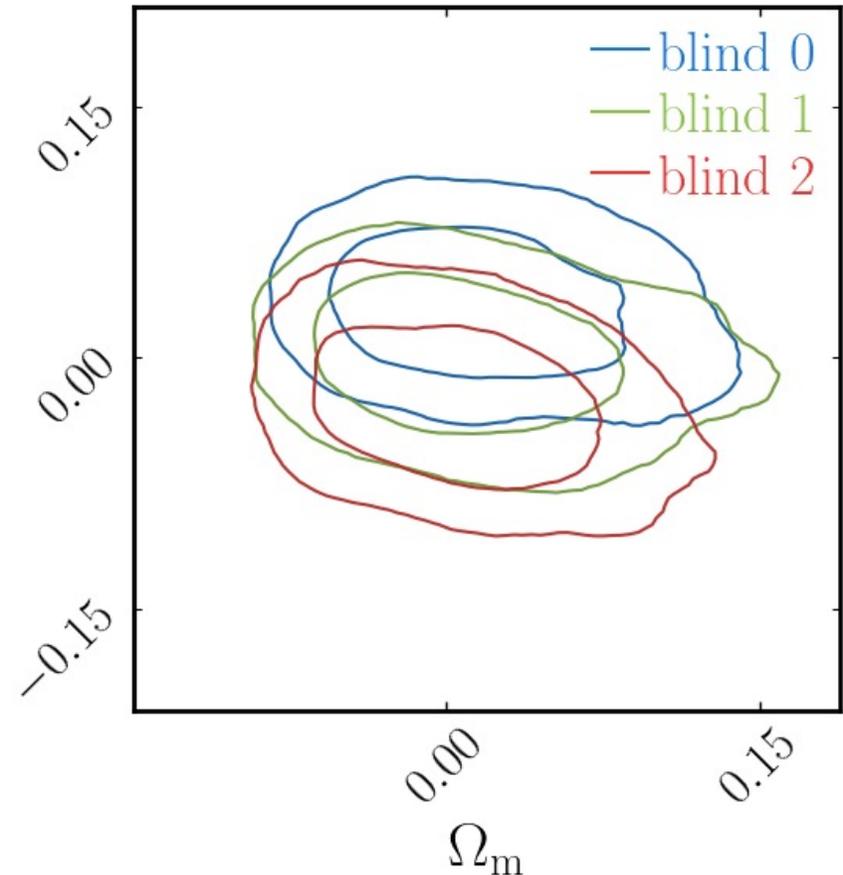
Blinding strategy

- Methodology to prevent confirmation bias.
- 2-tiered blinding strategy to ensure that unblinding one analysis does not automatically unblind the others.

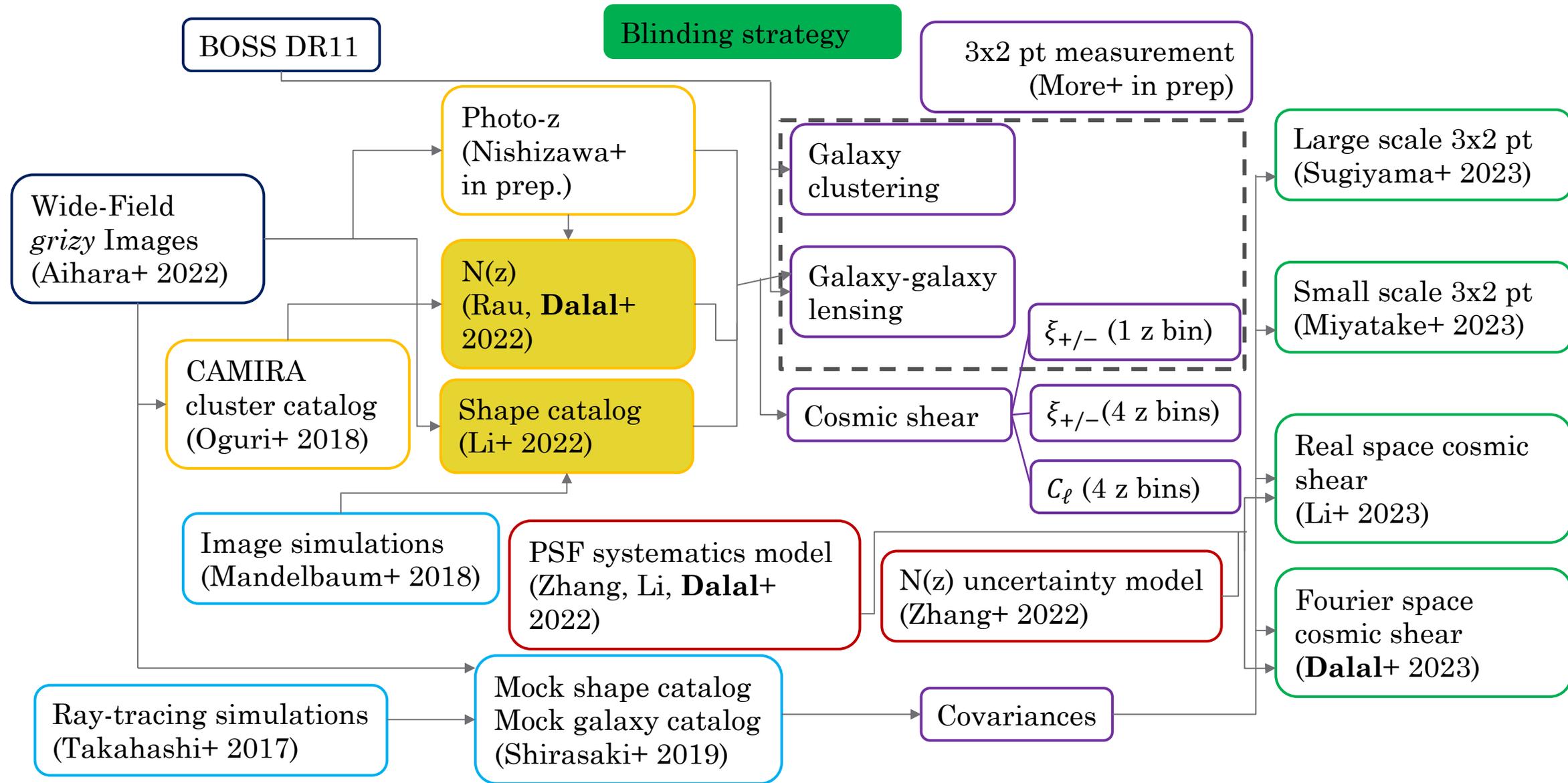
- Blinding on the level of multiplicative bias:

$$\mathbf{m}_{\text{cat}}^i = \mathbf{m}_{\text{true}} + \mathbf{dm}_1^i + \mathbf{dm}_2^i$$

- Each analysis team gets 3 different catalogs ($i = 0, 1, 2$). S_8
- \mathbf{dm}_1^i is removed by the analysis team lead before starting the analysis.
- The full analysis is done on all 3 blinded catalogs.
- One catalog has $\mathbf{dm}_2^i = 0$, i.e. is the true catalog, and this is only revealed after the analysis is complete, and all null tests and consistency checks have passed.



Images → Measurement → Cosmology



Data
Simulations
Intermediate data products
Data vectors
Modeling choices
Cosmological analyses

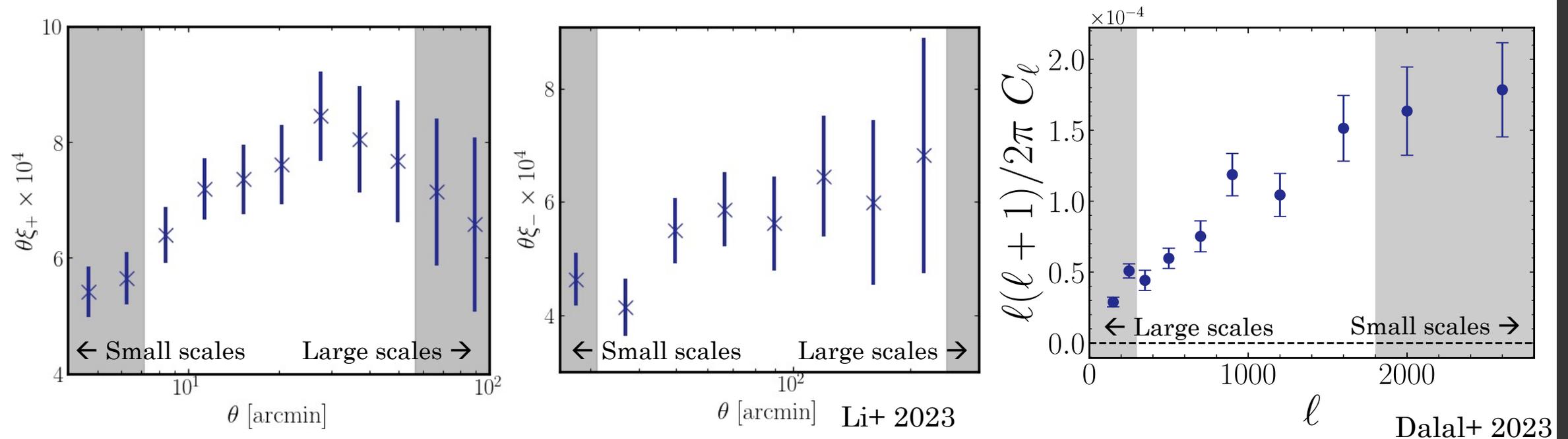
Cosmic shear measurements

When constructing cosmic shear data vectors, two methods we can use are:

- $\xi_{\pm}(\theta)$ (2 Point Correlation Function) – measures the correlation of shapes of galaxies with an angular separation of θ .

or

- C_{ℓ} (Angular Power Spectrum) – measures the second moment of the Fourier transform of the shear field (function of multipole, ℓ).

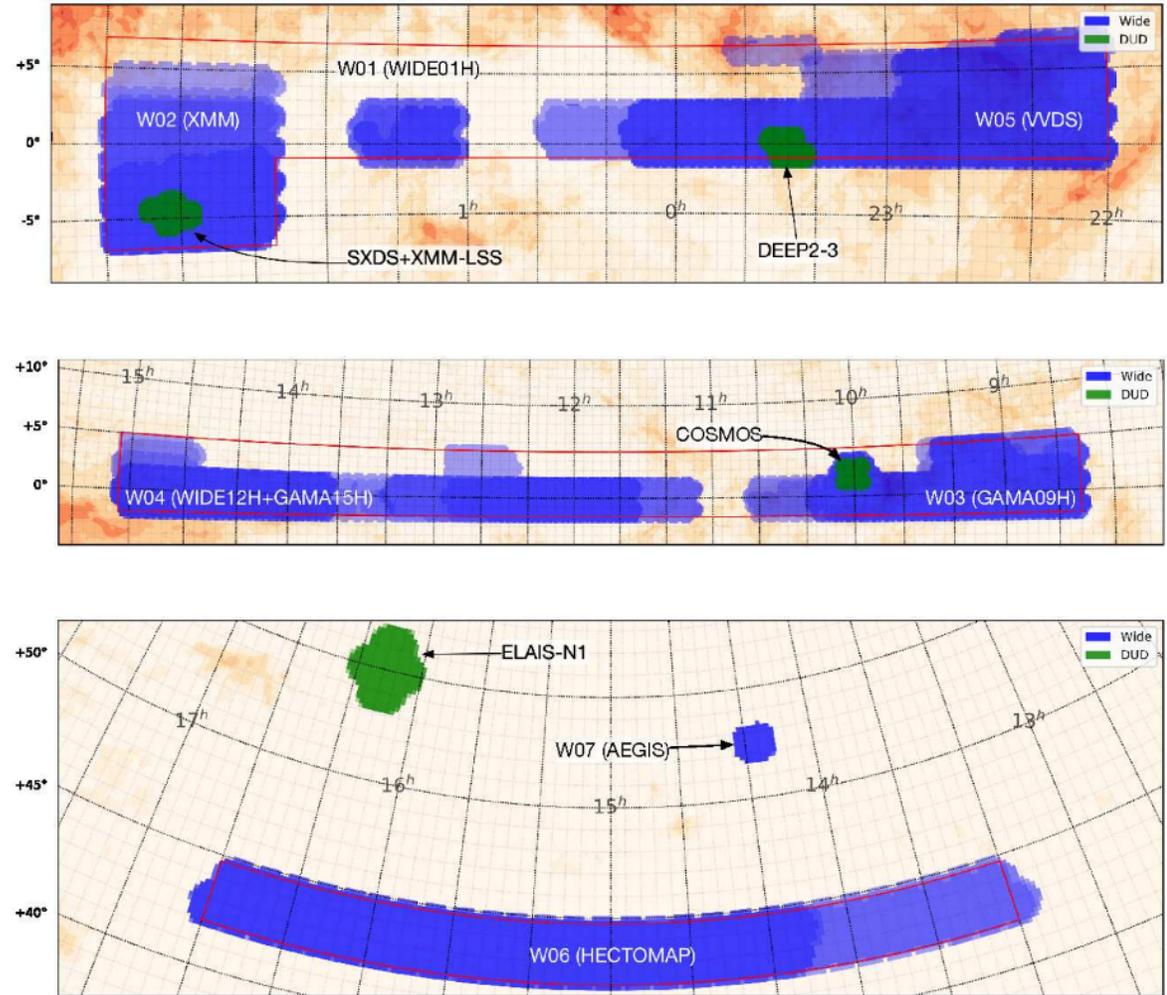


Cosmic shear power spectrum measurements

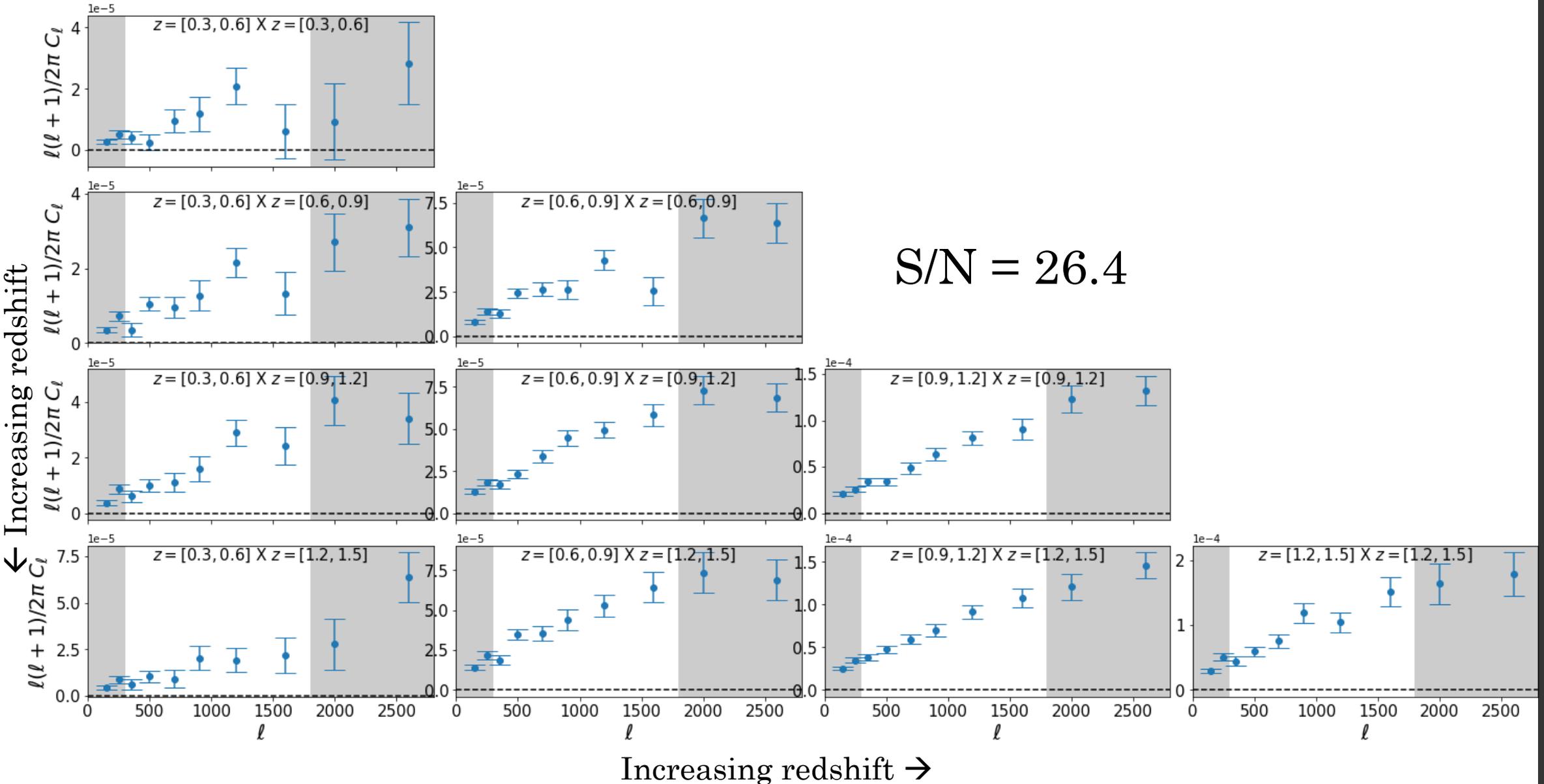
- Use NaMaster (Alonso+ 2019) to correct for biases due to partial sky coverage (Pseudo- C_ℓ).
- Measure 10 auto-and cross-correlation power spectra for 4 tomographic redshift bins between $z = 0.3$ and $z = 1.5$.
- Fiducial scale cuts: $300 \leq \ell \leq 1800$.

Due to evidence of systematics contamination at large scales.

Due to uncertainties in baryonic feedback and intrinsic alignment modeling.

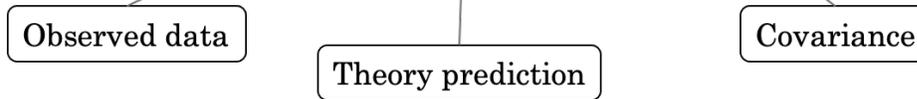


A high significance measurement



From data to cosmological constraints

$$-2 \ln \mathcal{L}(\hat{C}_\ell | \Theta) = \left(\hat{C}_\ell - C_\ell(\Theta) \right)^T \mathbf{C}^{-1} \left(\hat{C}_\ell - C_\ell(\Theta) \right)$$



Our power spectrum model is based on 23 parameters (5 cosmological, 6 astrophysical, 12 observational systematics). The model accounts for:

- Astrophysical effects
 - Baryonic feedback (HMCode2016, Mead+ 2016)
 - Intrinsic alignments (TATT, Blazek+ 2019)
- Systematics in the data
 - Point Spread Function systematics (Zhang, Li, Dalal+ 2022b)
 - Shear calibration biases
 - Redshift distribution uncertainties (Zhang+ 2022a)

We evaluate the likelihood throughout our parameter space using the PolyChord nested sampling algorithm (Handley+ 2019), implemented in CosmoSIS (Zuntz+ 2015).

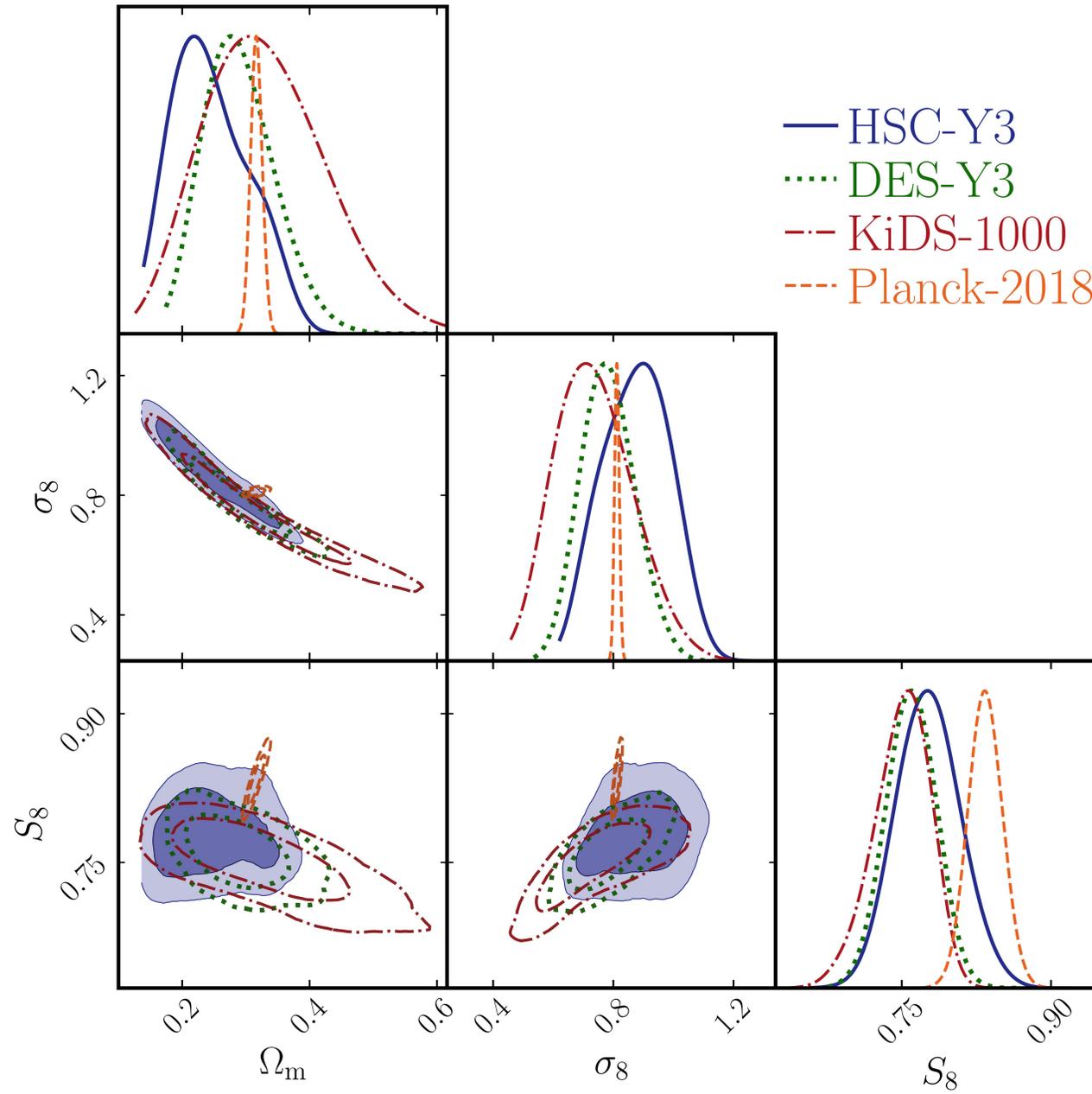
A 4% precision constraint:

$$S_8 = 0.776^{+0.032}_{-0.033}$$

p-value of best-fit model: 0.42

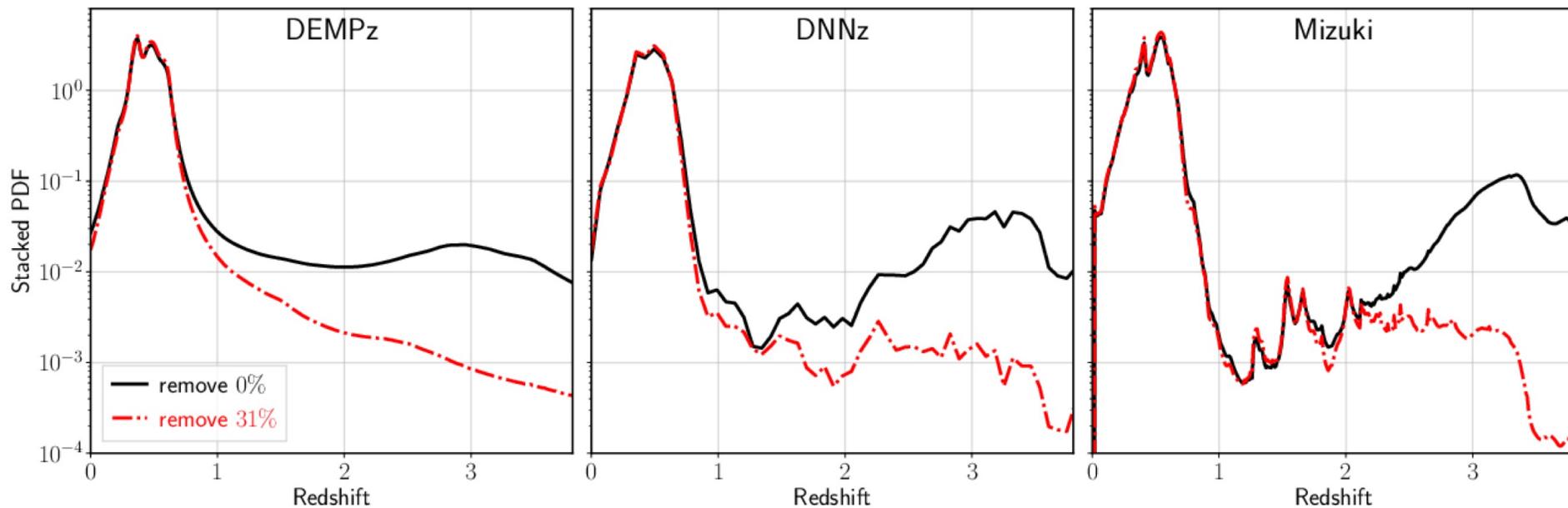
Consistency with Other Experiments

HSC-Y3 constraint from cosmic shear $C_{\ell s}$ is in 2σ tension with Planck.

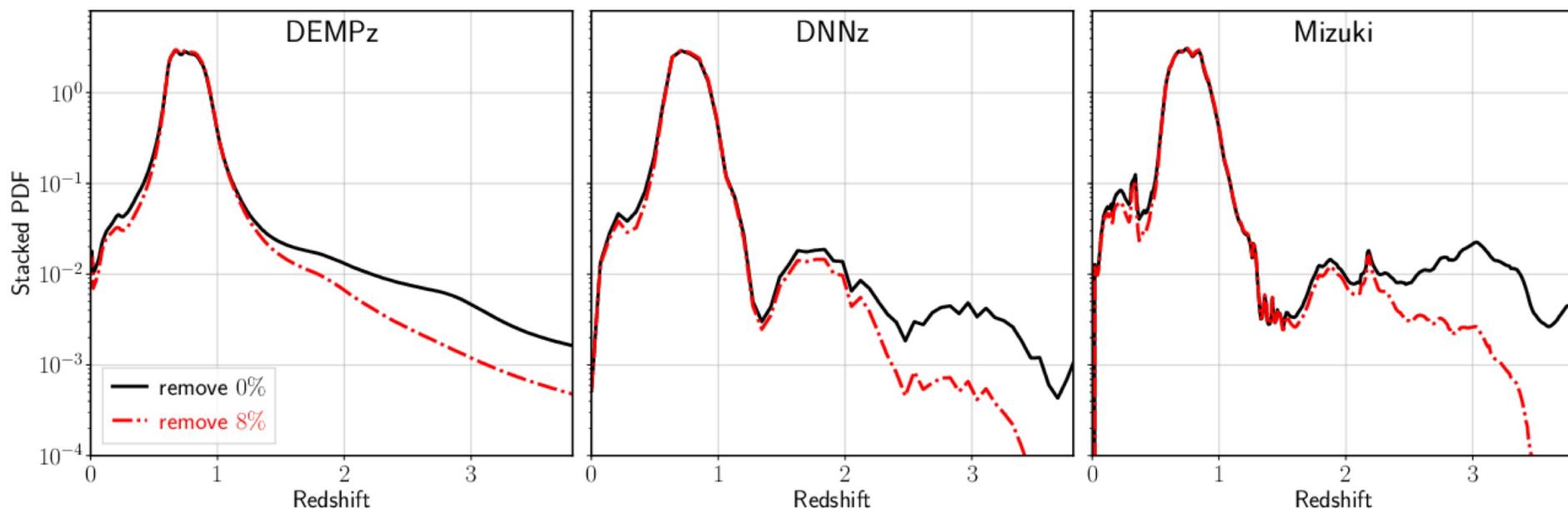


Redshift distribution uncertainties

First
tomographic
bin:
 $0.3 < z < 0.6$

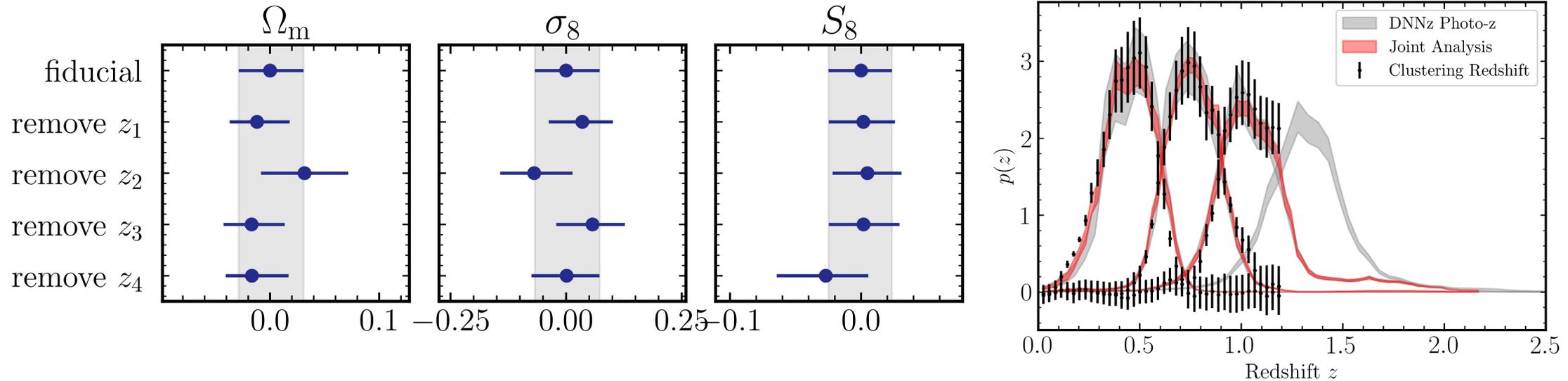


Second
tomographic
bin:
 $0.6 < z < 0.9$



DNNz and Mizuki photo-z estimates for a number individual galaxies have secondary peaks at high redshifts. We remove these galaxies from our sample.

Redshift Distribution Uncertainties



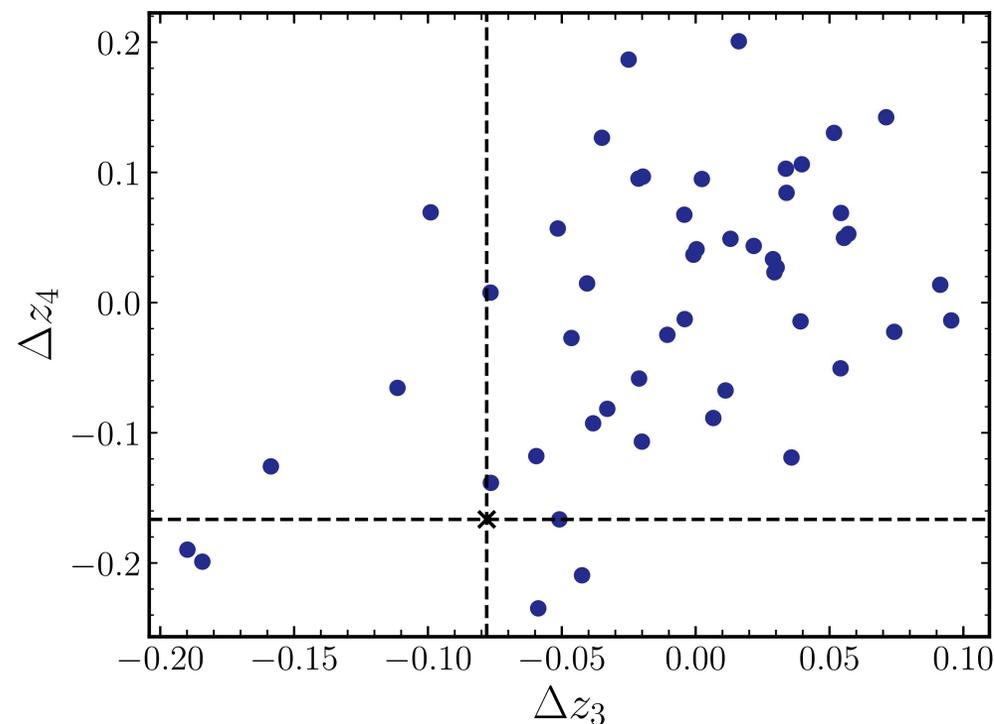
- We model uncertainties in the source redshift distribution using a shift model:

$$n^i(z) \rightarrow n^i(z + \Delta z_i)$$

- Redshift bin 3 ($0.9 < z \leq 1.2$) is only partially calibrated by clustering redshifts and bin 4 ($1.2 < z \leq 1.5$) is not at all calibrated.
- We adopt conservative, flat priors on Δz_3 and Δz_4 : $\mathcal{U}(-1, 1)$.

Residual Photo-z Errors

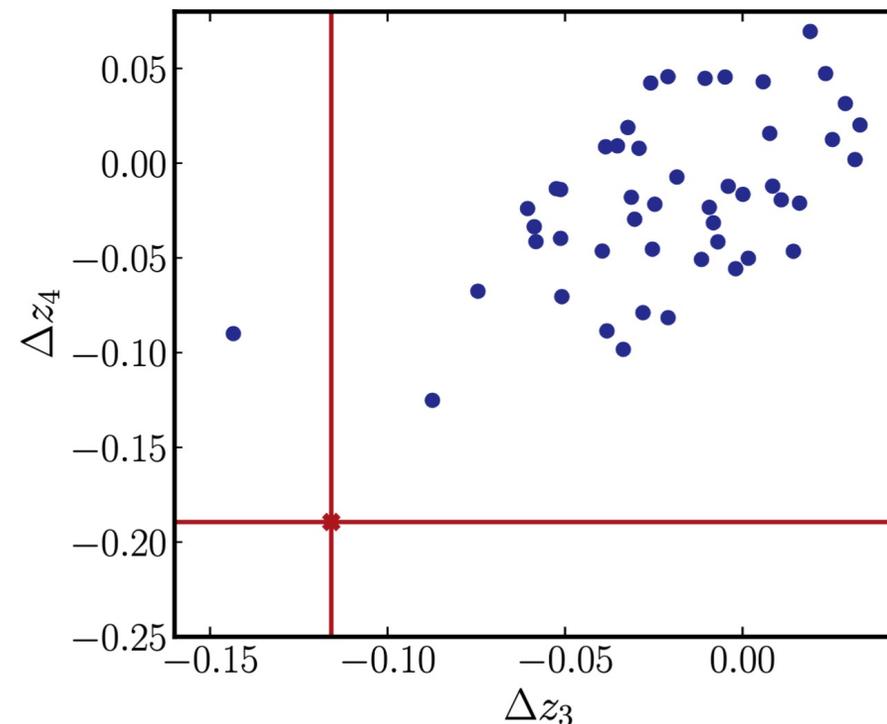
- Our constraining power is limited by the choice to use a conservative, wide, flat prior on the shifts in the third and fourth redshift bins.
- We find a 1.12σ shift in S_8 to a higher value when using informative priors.
- We find $\sim 2\sigma$ detections of significant shifts in these bins:
 - $\Delta z_3 = -0.076^{+0.056}_{-0.059}$
 - $\Delta z_4 = -0.157^{+0.094}_{-0.111}$
- Future work (e.g. with DESI and PFS) will be needed to calibrate these high redshift bins.



Power spectrum analysis

Residual Photo-z Errors

- Our constraining power is limited by the choice to use a conservative, wide, flat prior on the shifts in the third and fourth redshift bins.
- We find a 1.12σ shift in S_8 to a higher value when using informative priors.
- We find $\sim 2\sigma$ detections of significant shifts in these bins:
 - $\Delta z_3 = -0.115^{+0.052}_{-0.058}$
 - $\Delta z_4 = -0.192^{+0.088}_{-0.088}$
- Future work (e.g. with DESI and PFS) will be needed to calibrate these high redshift bins.



2PCF analysis

PSF Systematics

PSF Systematics Model (Zhang, Li, Dalal+ 2022)



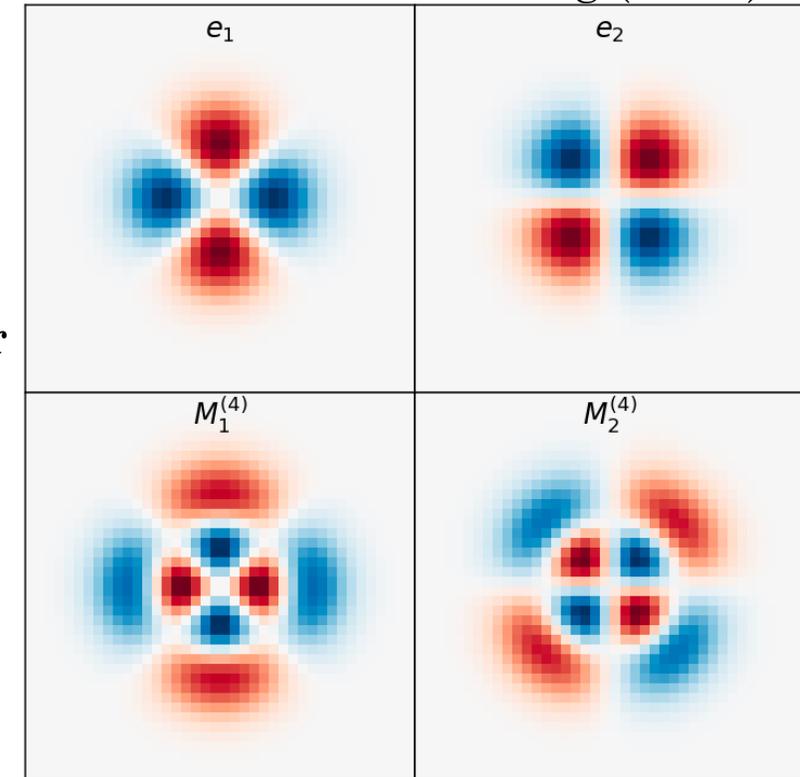
Led by Tianqing Zhang (CMU)

The point spread function can contaminate cosmic shear measurements in two ways:

1. **PSF Leakage** – the shape of the PSF coherently contaminates the inferred shear even when the PSF model is perfect (due to imperfect PSF deconvolution by the shear estimator).
2. **PSF Modeling Error** - when the PSF model inaccurately describes the actual PSF shape, the inferred shear can get an additive systematic term.

In the past, models of PSF systematics have been limited to second order moments of the PSF.

We extend the formalism to fourth order moments, and show that these have an important contribution to the bias in C_ℓ s.



The image response to the spin-2 quantity of the second moment e_1 and e_2 , and fourth moment $M_1^{(4)}$ and $M_2^{(4)}$.

PSF Systematics Model (cont.)

- Observed galaxy ellipticity: $\hat{g}_{\text{gal}} = g_{\text{gal}} + g + g_{\text{sys}}$
- $g_{\text{sys}} = \alpha^{(2)} e_{\text{PSF}} + \beta^{(2)} \Delta e_{\text{PSF}} + \alpha^{(4)} M_{\text{PSF}}^{(4)} + \beta^{(4)} \Delta M_{\text{PSF}}^{(4)}$
- The measured cosmic shear power spectrum becomes:

$$C_\ell \rightarrow C_\ell + \underbrace{\sum_{i=1}^4 \sum_{j=1}^4 \mathbf{p}_i \mathbf{p}_j C_\ell^{\mathbf{S}_i \mathbf{S}_j}}_{\Delta C_\ell}$$

with the parameter vector $\mathbf{p} = [\alpha^{(2)}, \beta^{(2)}, \alpha^{(4)}, \beta^{(4)}]$ and the PSF moments vector $\mathbf{S}_i = [e_{\text{PSF}}, \Delta e_{\text{PSF}}, M_{\text{PSF}}^{(4)}, \Delta M_{\text{PSF}}^{(4)}]$.

PSF Systematics Model (cont.)

- We compute the galaxy shear-PSF systematic cross-correlations, and compare them to the theory galaxy-PSF C_ℓ s based on the PSF-PSF correlations to fit the PSF systematic parameters $(\alpha^{(2)}, \beta^{(2)}, \alpha^{(4)}, \beta^{(4)})$

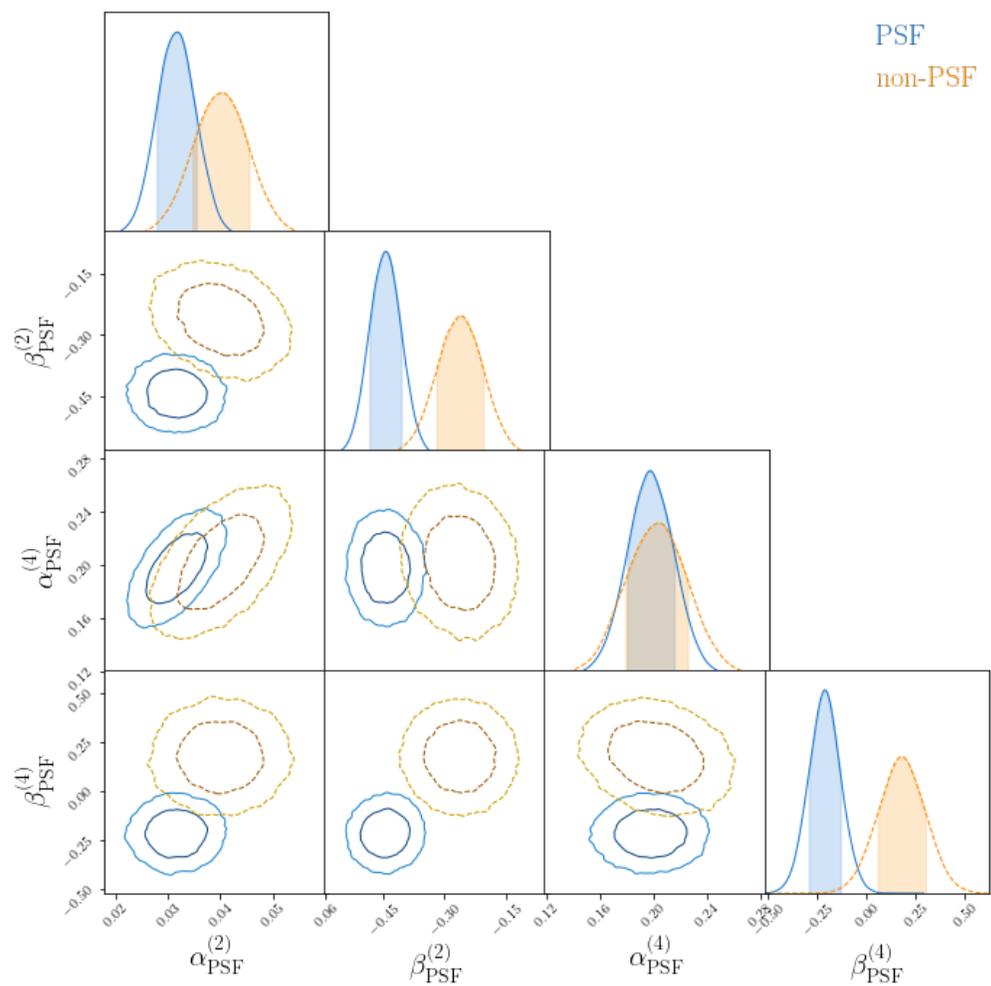
$$C_\ell^{\hat{g}_{\text{gal}} e_{\text{PSF}}} = \alpha^{(2)} C_\ell^{e_{\text{PSF}} e_{\text{PSF}}} + \beta^{(2)} C_\ell^{\Delta e_{\text{PSF}} e_{\text{PSF}}} + \alpha^{(4)} C_\ell^{M_{\text{PSF}}^{(4)} e_{\text{PSF}}} + \beta^{(4)} C_\ell^{\Delta M_{\text{PSF}}^{(4)} e_{\text{PSF}}}$$

$$C_\ell^{\hat{g}_{\text{gal}} \Delta e_{\text{PSF}}} = \alpha^{(2)} C_\ell^{e_{\text{PSF}} \Delta e_{\text{PSF}}} + \beta^{(2)} C_\ell^{\Delta e_{\text{PSF}} \Delta e_{\text{PSF}}} + \alpha^{(4)} C_\ell^{M_{\text{PSF}}^{(4)} \Delta e_{\text{PSF}}} + \beta^{(4)} C_\ell^{\Delta M_{\text{PSF}}^{(4)} \Delta e_{\text{PSF}}}$$

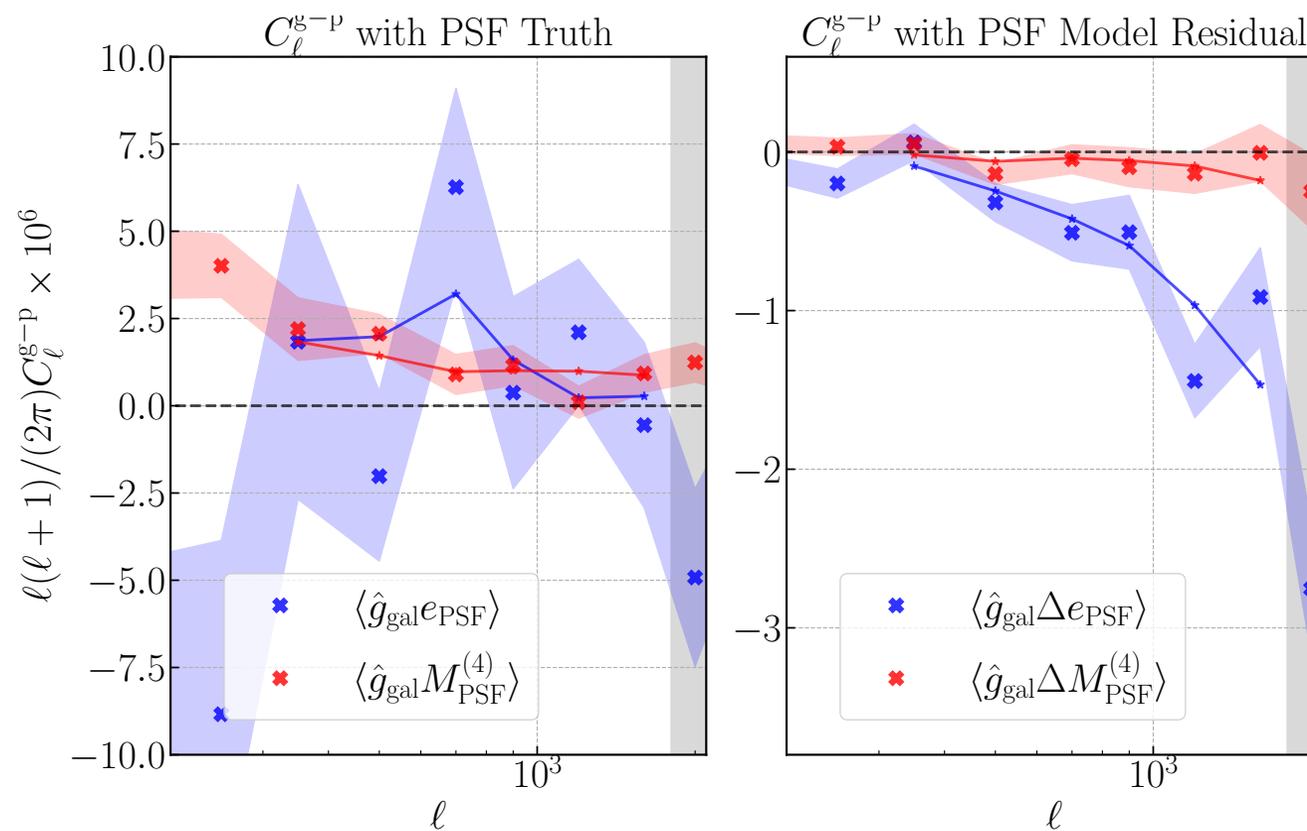
$$C_\ell^{\hat{g}_{\text{gal}} M_{\text{PSF}}^{(4)}} = \alpha^{(2)} C_\ell^{e_{\text{PSF}} M_{\text{PSF}}^{(4)}} + \beta^{(2)} C_\ell^{\Delta e_{\text{PSF}} M_{\text{PSF}}^{(4)}} + \alpha^{(4)} C_\ell^{M_{\text{PSF}}^{(4)} M_{\text{PSF}}^{(4)}} + \beta^{(4)} C_\ell^{\Delta M_{\text{PSF}}^{(4)} M_{\text{PSF}}^{(4)}}$$

$$C_\ell^{\hat{g}_{\text{gal}} \Delta M_{\text{PSF}}^{(4)}} = \alpha^{(2)} C_\ell^{e_{\text{PSF}} \Delta M_{\text{PSF}}^{(4)}} + \beta^{(2)} C_\ell^{\Delta e_{\text{PSF}} \Delta M_{\text{PSF}}^{(4)}} + \alpha^{(4)} C_\ell^{M_{\text{PSF}}^{(4)} \Delta M_{\text{PSF}}^{(4)}} + \beta^{(4)} C_\ell^{\Delta M_{\text{PSF}}^{(4)} \Delta M_{\text{PSF}}^{(4)}}$$

PSF Systematics Constraints

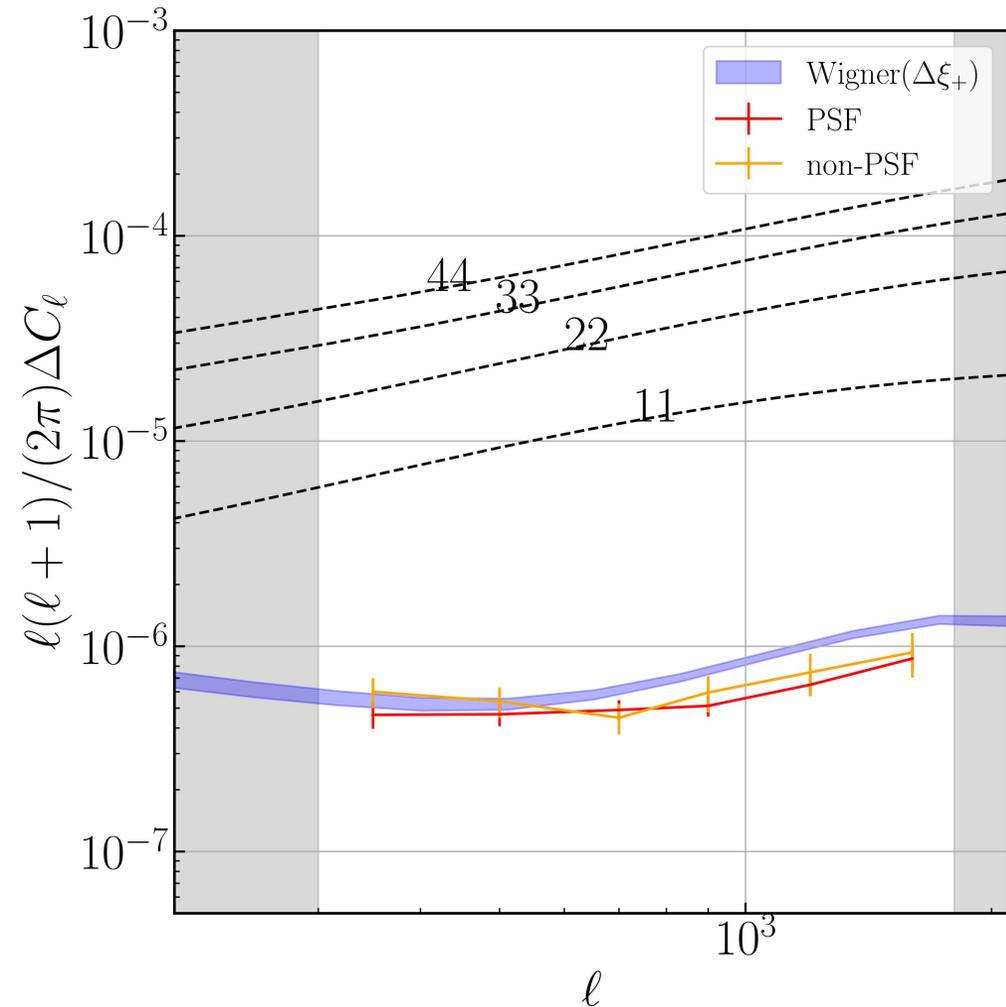


Constraints on the PSF parameters from fitting the shear-PSF cross-correlations.



Best-fit PSF model compared to measurements.

PSF Systematics Impact on Cosmology



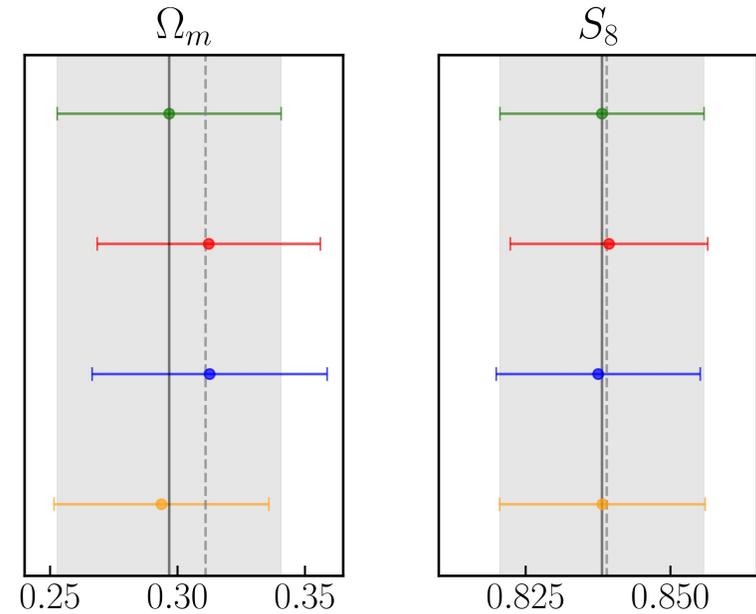
Impact of PSF model on cosmic shear C_ℓ s compared to prediction from real space analysis.

No Systematics

No Correction

Second Moments Correction

Fiducial Correction



Impact of PSF model on cosmological parameter constraints, compared to a simpler second moment model.

Baryonic feedback + Intrinsic Alignments

Power spectrum model

For the Λ CDM model, the linear matter power spectrum is a function of 5 cosmological parameters:

- Ω_m - Matter density
- H_0 - Hubble constant (expansion rate of the universe)
- A_s - Amplitude of the primordial power spectrum
- n_s - Tilt of the primordial power spectrum
- $\omega_b \equiv \Omega_b h^2$ - Baryon density ($h \equiv \frac{H_0}{100} \text{ km s}^{-1} \text{ Mpc}^{-1}$)

We use `baccoemu` to compute the linear power spectrum from these parameters.

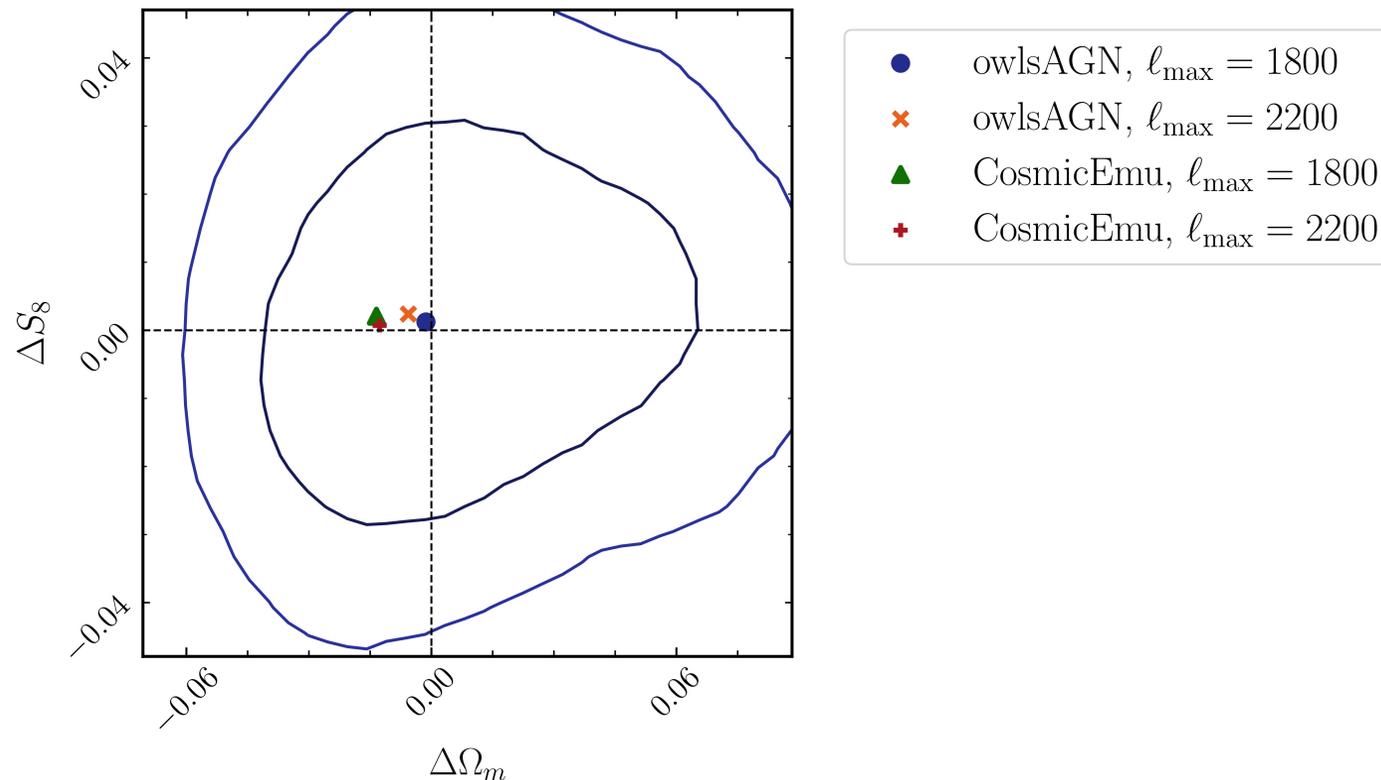
At small scales, the growth of structure is nonlinear, and baryonic feedback from AGN leads to suppression of the power spectrum.

We model the nonlinear power spectrum with `HMCode 2016`, adding one parameter to describe baryonic feedback:

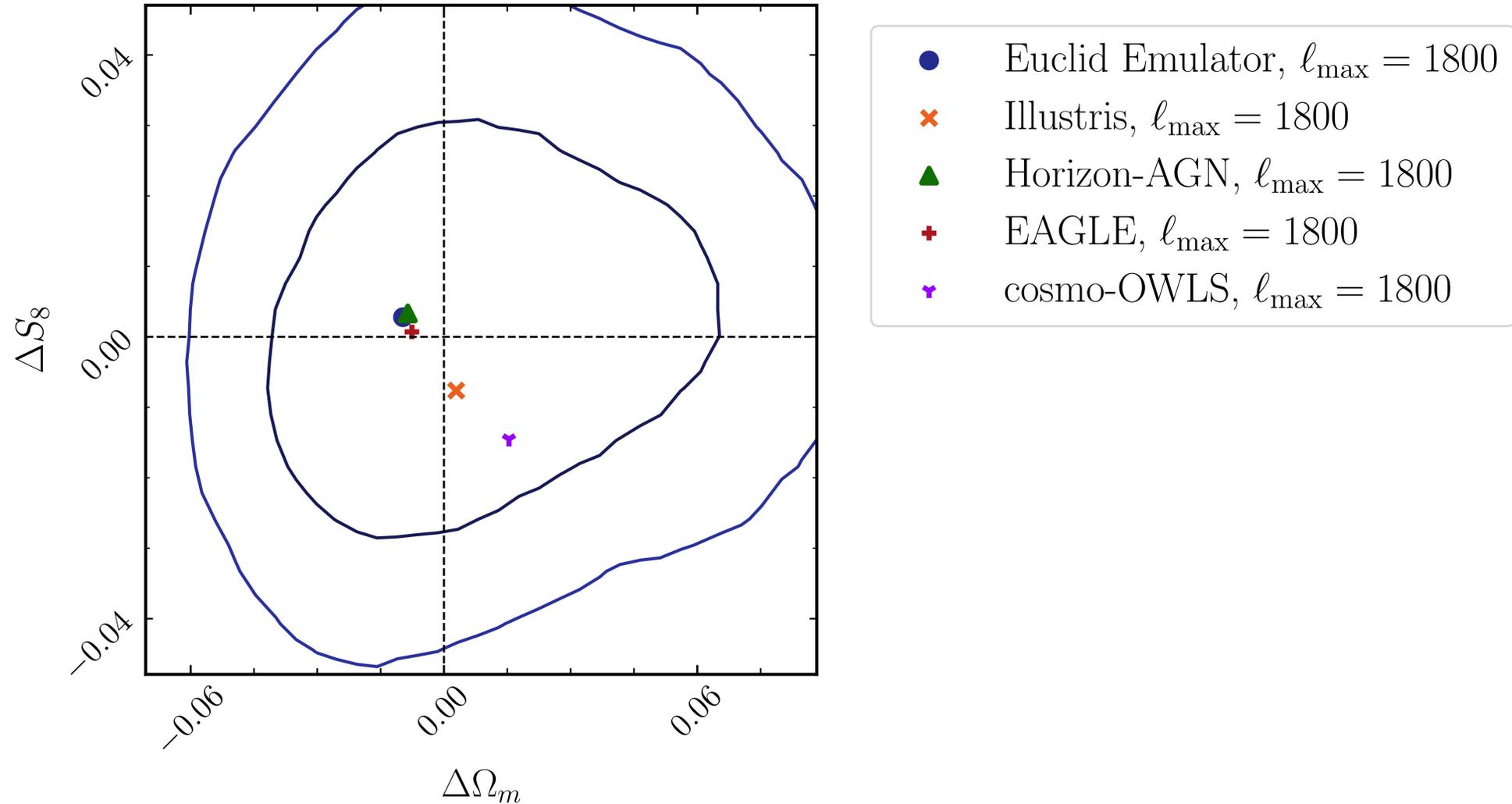
- A_{bary} - the amplitude of the halo mass concentration ($A_{\text{bary}} = 3.13$ for no baryons)

Model selection tests

We simulate “contaminated data vectors” with different models of baryonic feedback and nonlinear growth of structure, and analyze them with our fiducial model to understand how much model misspecification can bias our results.



Further model validation



Intrinsic Alignments

Power spectra get an additional contribution from the intrinsic shapes of galaxies being aligned with the tidal field of the gravitational potential.

2 different IA models:

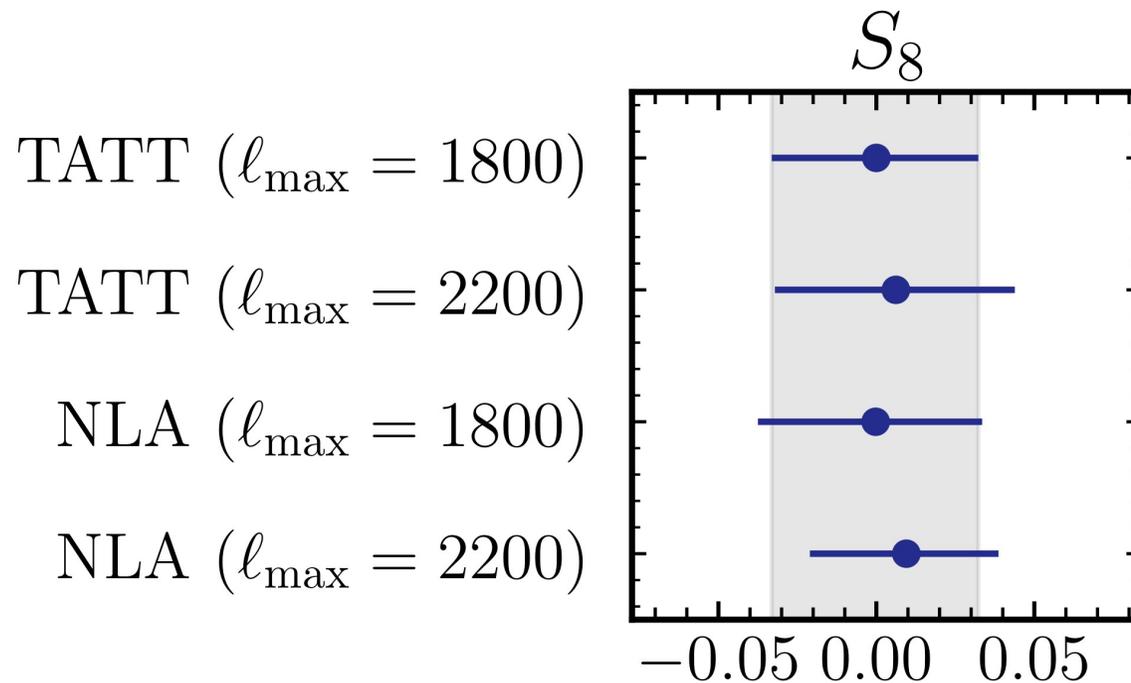
1. Tidal Alignments and Tidal Torquing (TATT)

- A_1 - Amplitude of IA power spectra scaling linearly with the tidal field.
- A_2 - Amplitude of IA power spectra scaling quadratically with the tidal field.
- η_1 - Redshift evolution of linear term.
- η_2 - Redshift evolution of quadratic term.
- b_{TA} - Galaxy bias parameter.

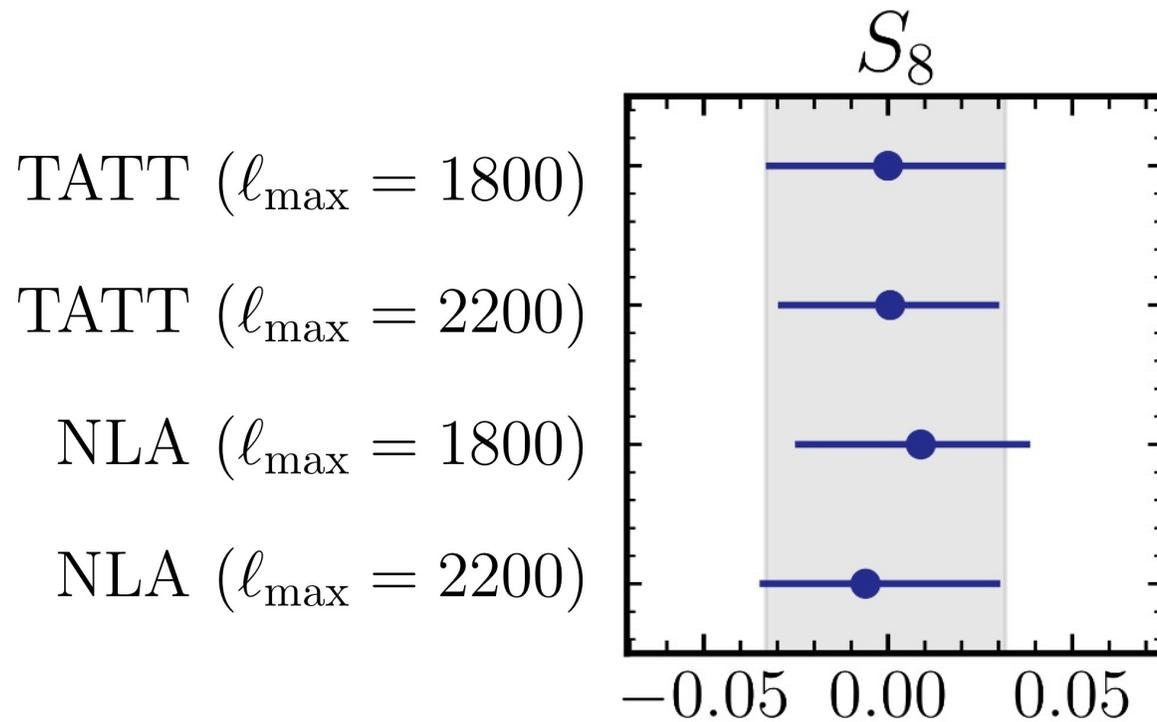
2. Non-linear alignments (NLA) – subset of TATT

- A_1 - Amplitude of IA power spectra scaling linearly with the tidal field.
- η_1 - Redshift evolution of linear term.

Blinded Catalog 0 (most constraining power)



Blinded Catalog 2 (true catalog)

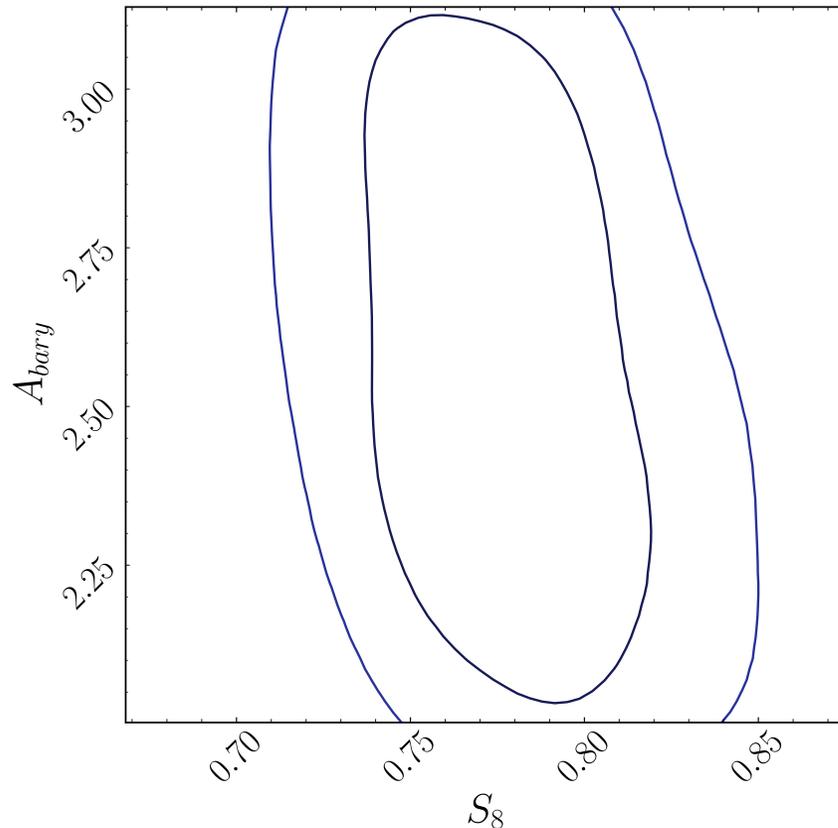


Baryonic Feedback Constraints

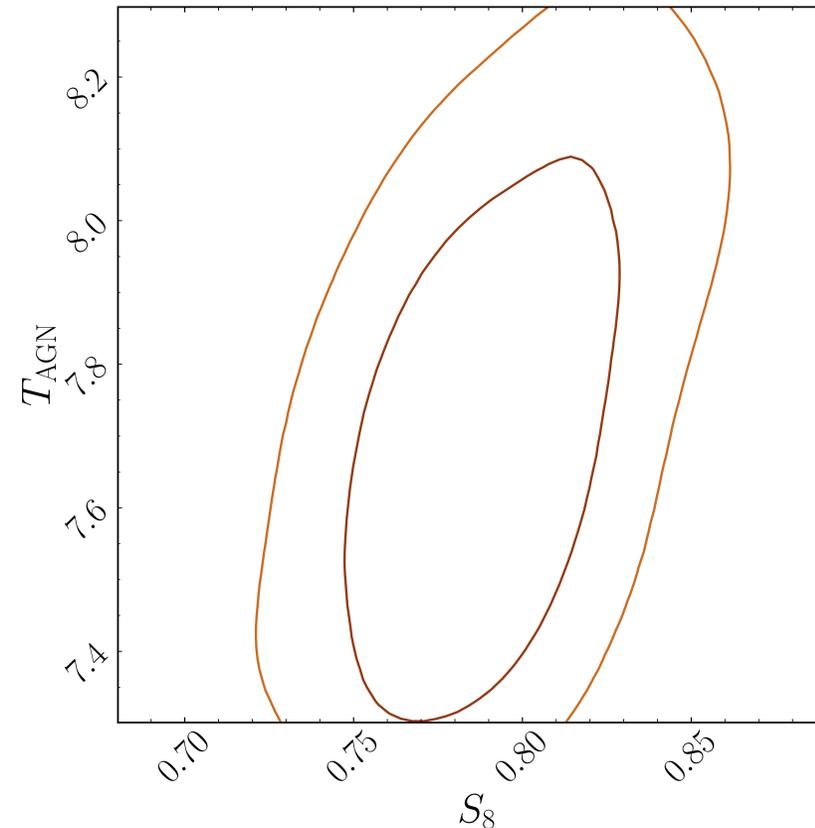
- 1.5 σ departure from no baryonic feedback in our fiducial analysis:
 $A_{bary} = 2.43_{-0.25}^{+0.46}$.
- When using HMCode2020: $T_{AGN} = 7.68_{-0.25}^{+0.27}$ (similar level of feedback as owlsAGN).
- Not modeling baryons leads to a 0.5 σ shift in S_8 to a lower value.

Baryonic Feedback and the S_8 Tension

Amon & Efstathiou (2022) suggest that the S_8 tension could be resolved by strong baryonic feedback.



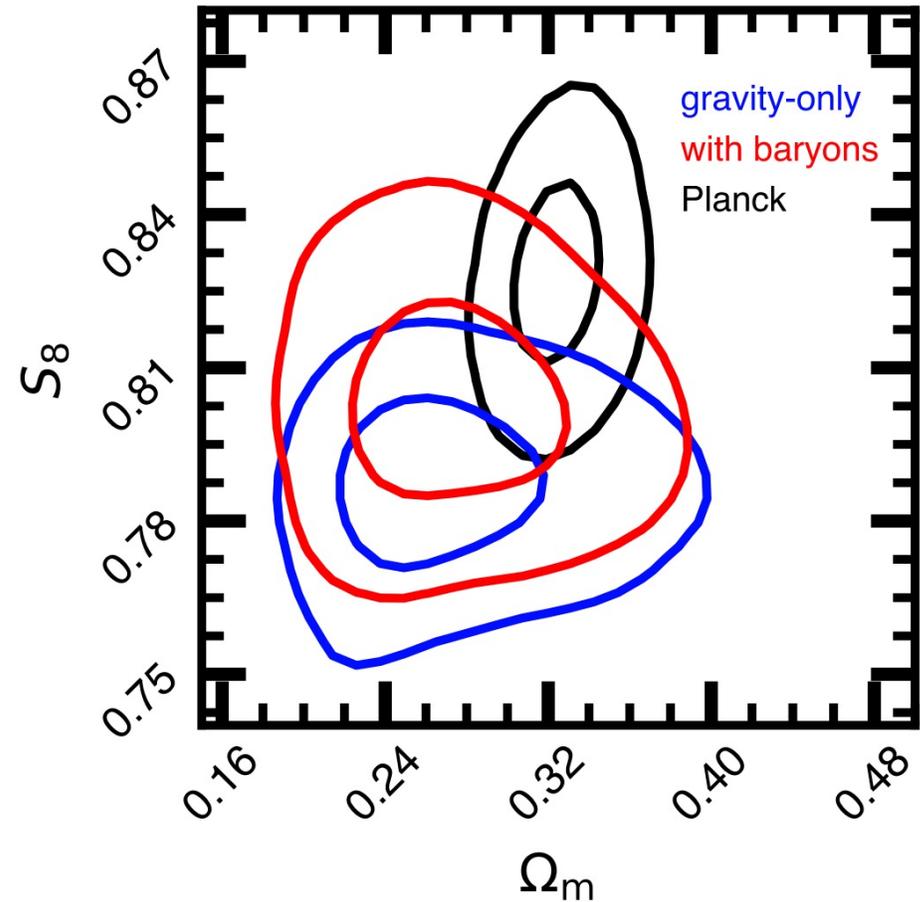
HMCCode 2016



HMCCode 2020

Baryonic Feedback and the S_8 Tension

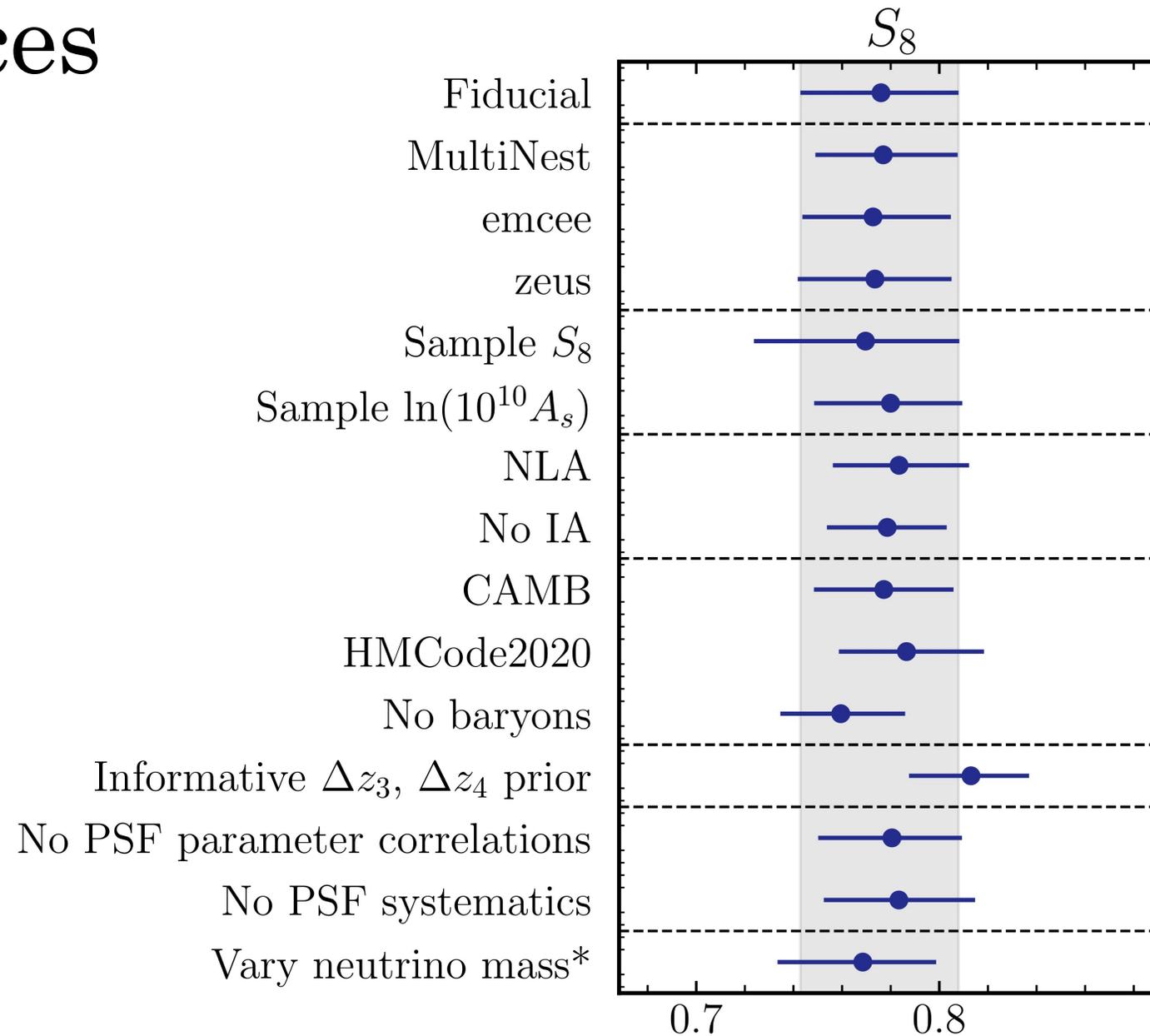
- Arico+ 2023 analyze all scales measured for the DES Y3 cosmic shear 2PCFs, using a 7 parameter “baryonification” model and find a higher value of S_8 .
- I’m currently working to develop a principled methodology to select the optimal scale cuts + baryon model to jointly constrain cosmology and feedback.
- We will then re-analyze the HSC power spectra with this optimized choice.
- Cross-correlations with CMB SZ measurements can also be used to directly measure baryons and place informative priors on feedback parameters for cosmology analyses.



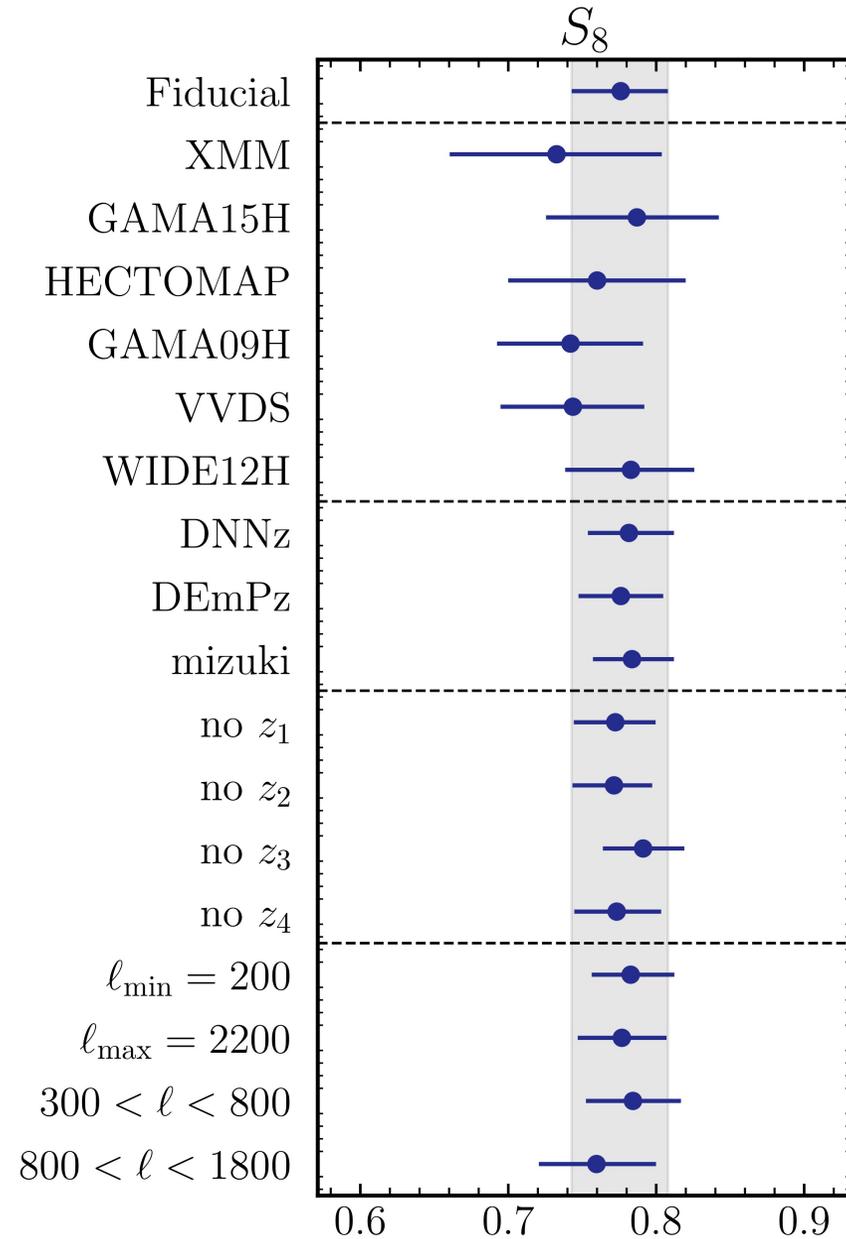
Arico+ 2023

Sensitivity to systematics

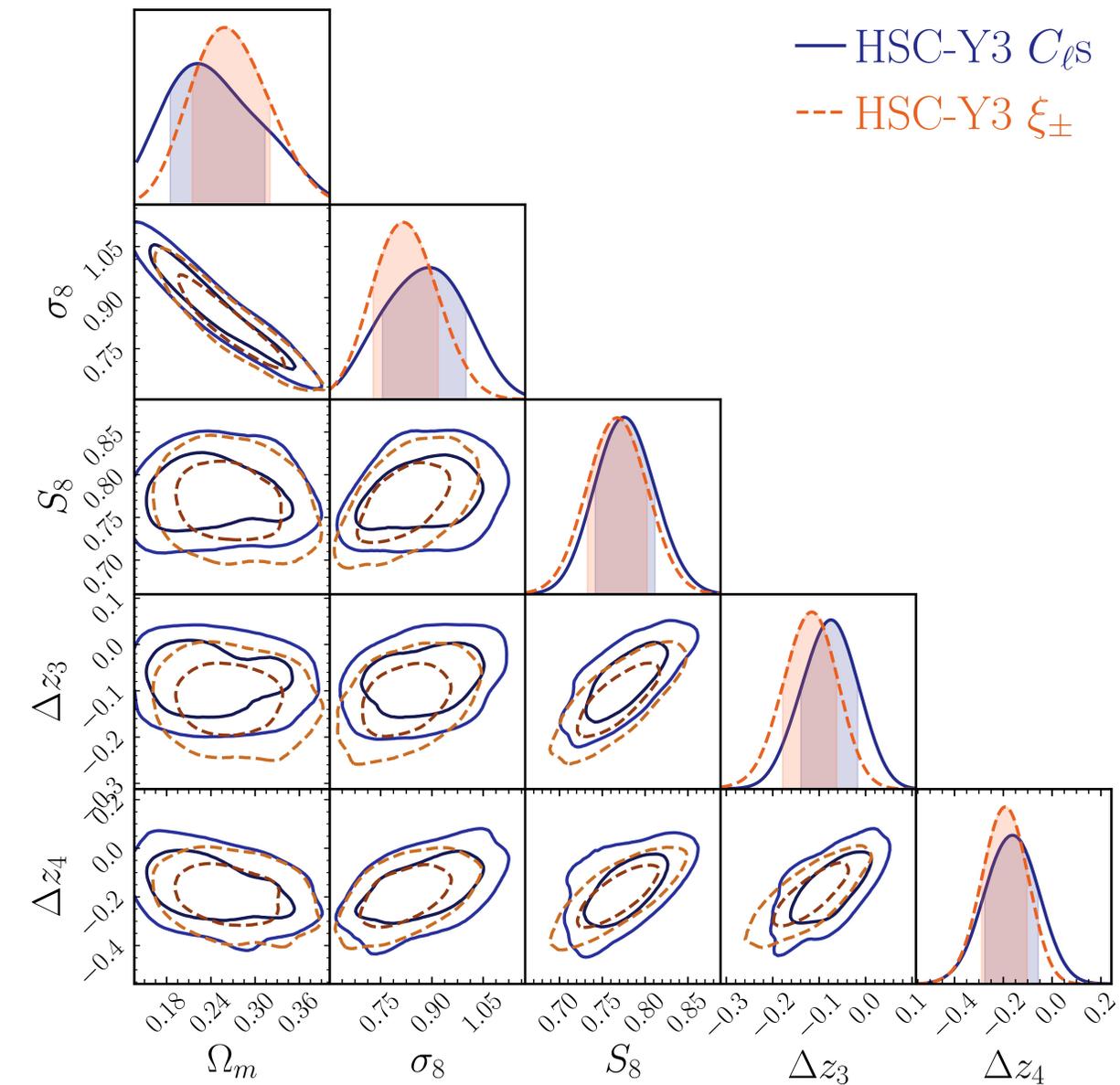
Robustness to Modeling and Analysis Choices



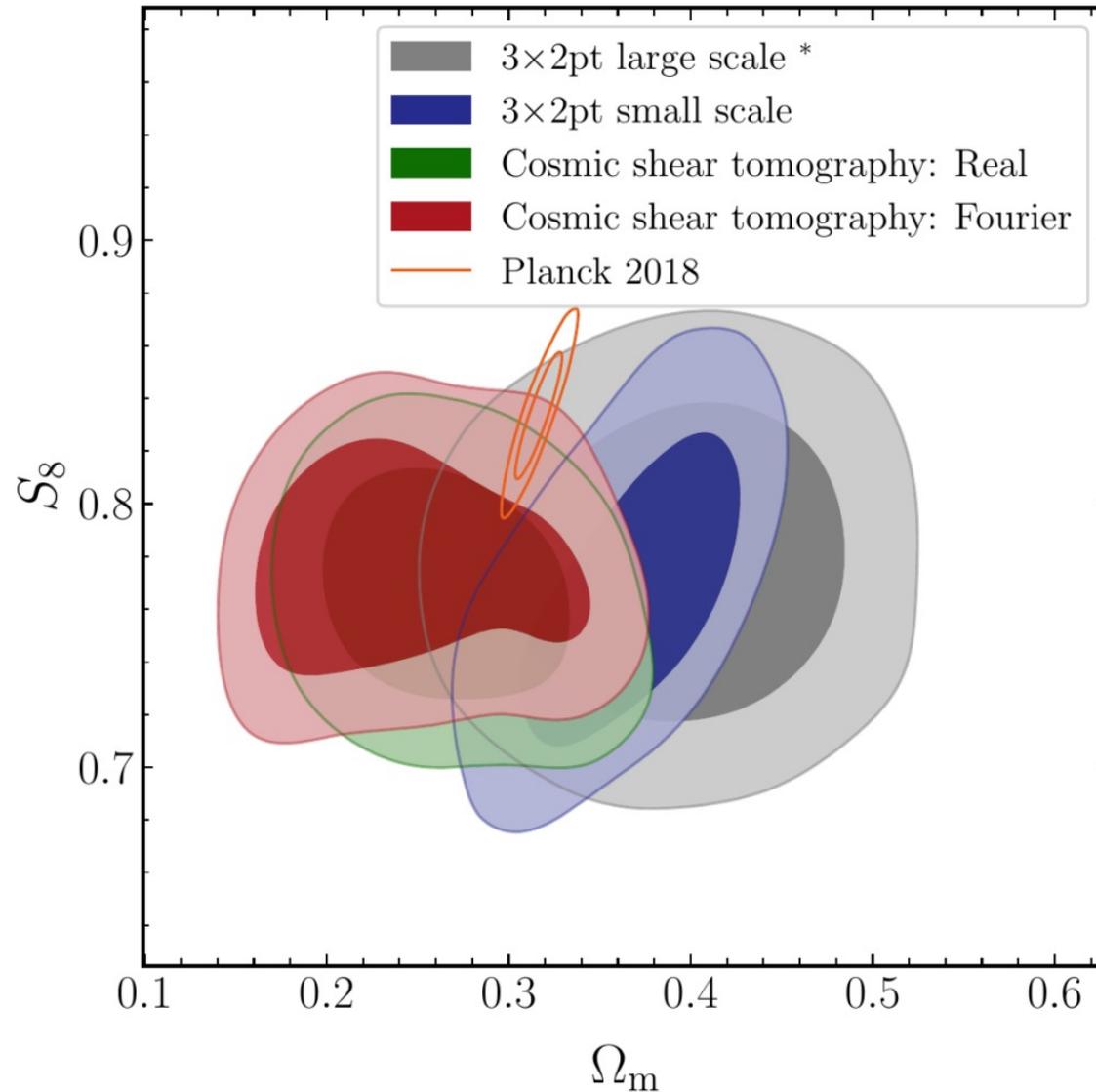
Internal Consistency Tests



Consistency with 2PCF Analysis



Consistency with Other HSC Analyses



Other Lessons Learned

- Shear estimation:
 - Multiband image simulations are necessary for calibrating redshift-dependent shear.
 - Self-calibrating shear estimators (e.g. metadetect, FPFS) will be needed in the future.
- PSF systematics:
 - Modeling of systematics from higher order PSF moments is crucial.
 - An expanded suite of null tests can help catch sources of systematic error.
 - Coordination between the pipeline team and the science analysis team is essential.
- Redshift inference:
 - Limited access to unbiased calibration data at the faint end of color space makes redshift calibration model-dependent.
 - High-redshift density tracers will be need for cross-correlation based calibration.
- Model selection:
 - Define model selection criteria in advance, including thresholds for acceptable levels of biases.
 - Use maximum a posteriori estimates for tests of model misspecification.

Improvements with Future Data

- We will need higher precision measurements to better understand the S_8 tension.
- The HSC final data release will cover $\sim 1000 \text{ deg}^2$ of the sky, also with extraordinary depth and seeing.
- The Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST) will cover $18,000 \text{ deg}^2$, going one magnitude deeper than HSC.
- Two upcoming space telescopes: Euclid and the Nancy Grace Roman Space Telescope.
- Ongoing work to better study modeling choices and develop analysis tools will be crucial.
- Lots of interesting work to be done at smaller scales, especially related to baryonic feedback!

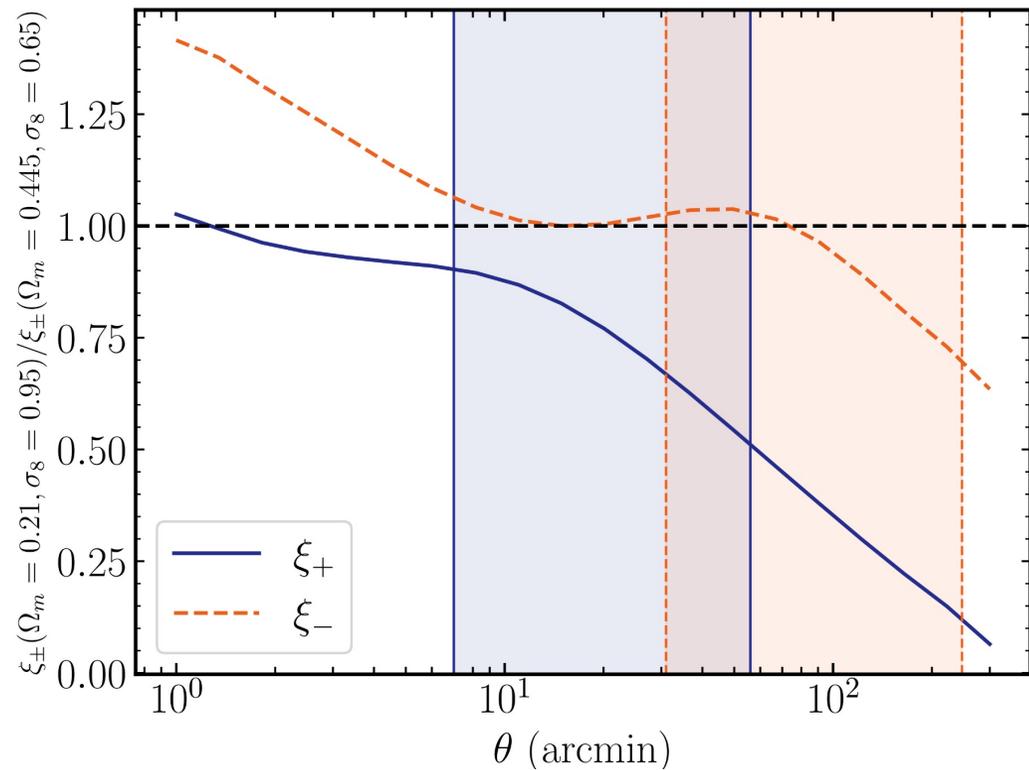
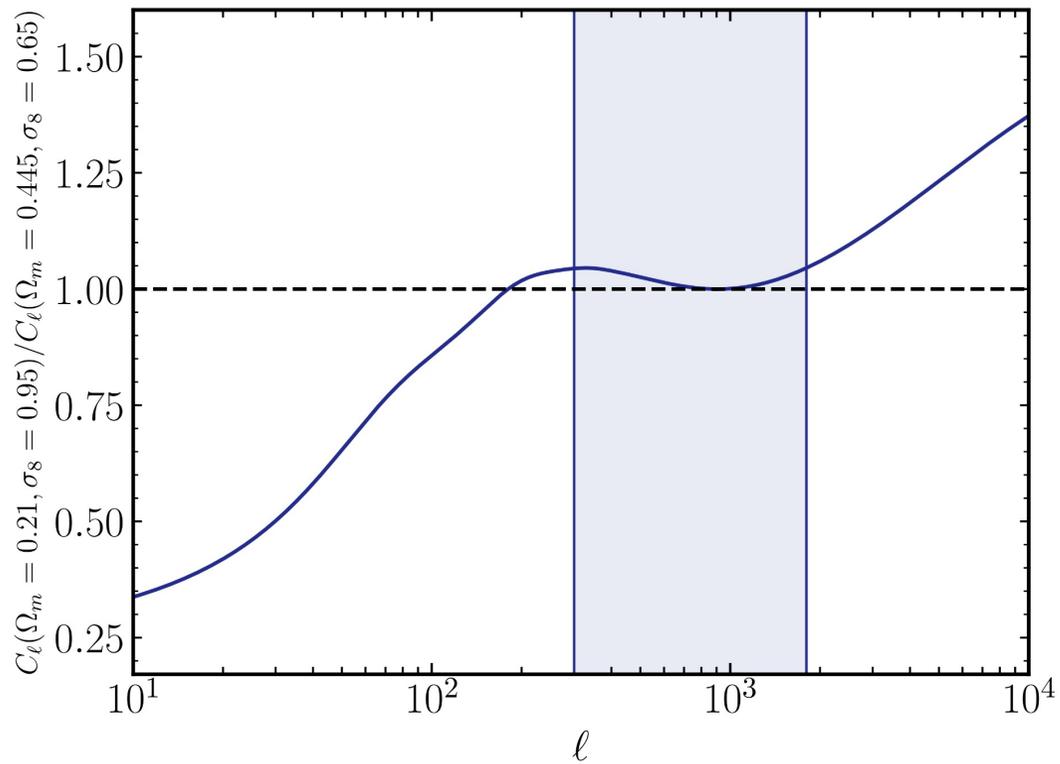
HSC Y3 Cosmology Papers

- Weak Lensing Tomographic Redshift Distribution Inference - Rau, Dalal+
- A General Framework for Removing Point Spread Function Additive Systematics in Cosmological Weak Lensing Analysis – Zhang, Li, Dalal+
- Cosmology from cosmic shear power spectra– Dalal+
- Cosmology from cosmic shear two-point correlation functions – Li+
- Measurements of the clustering of SDSS-BOSS galaxies, galaxy-galaxy lensing and cosmic shear – More+
- Cosmology from galaxy clustering and weak lensing with HSC and SDSS using the minimal bias model - Sugiyama+
- Cosmology from galaxy clustering and weak lensing with HSC and SDSS using the emulator based halo model– Miyatake+

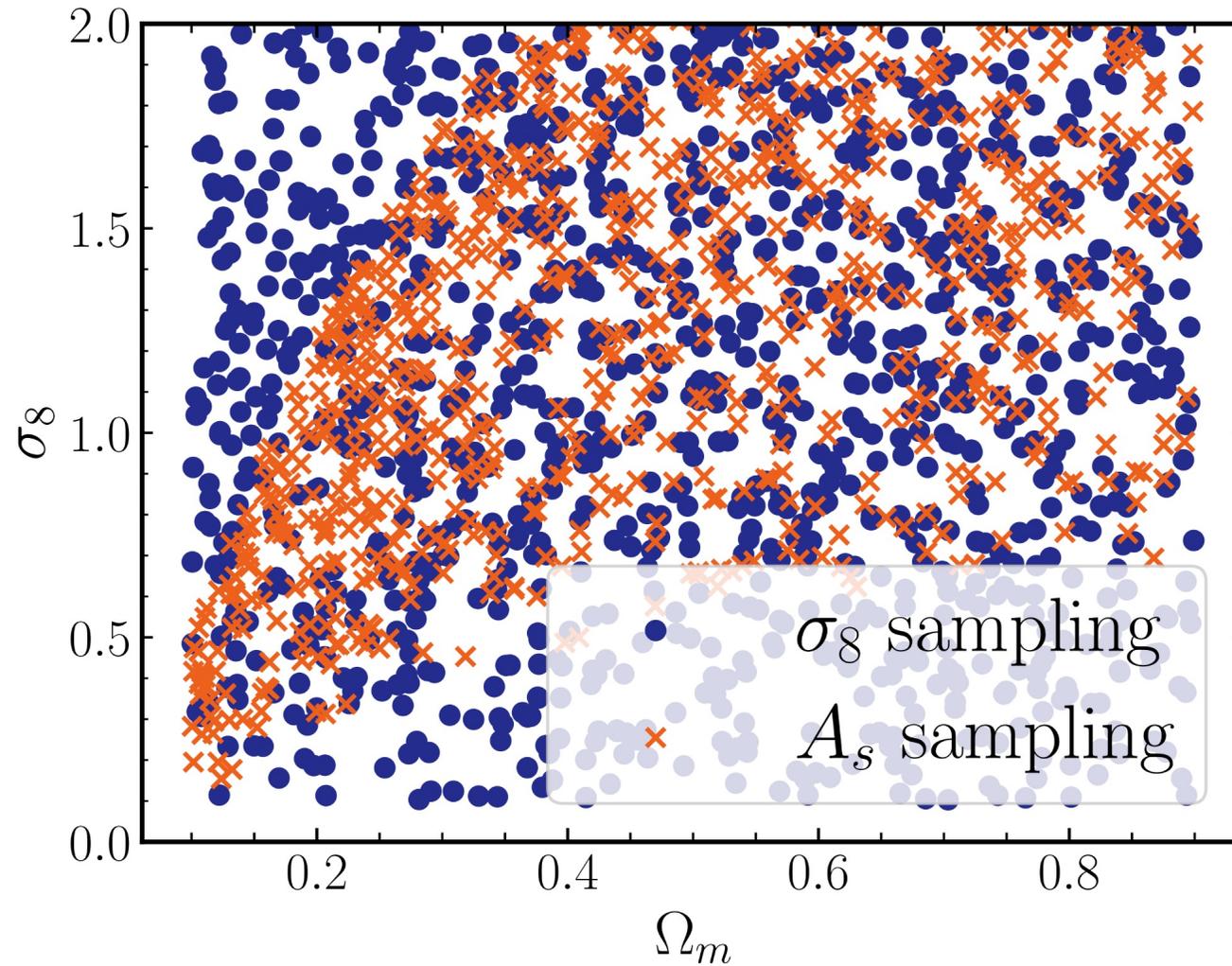
Many other upcoming HSC projects (including cluster cosmology led by Tomomi Sunayama, and tomographic galaxy clustering and cross-correlations with ACT).

Backup slides

Sampling A_s

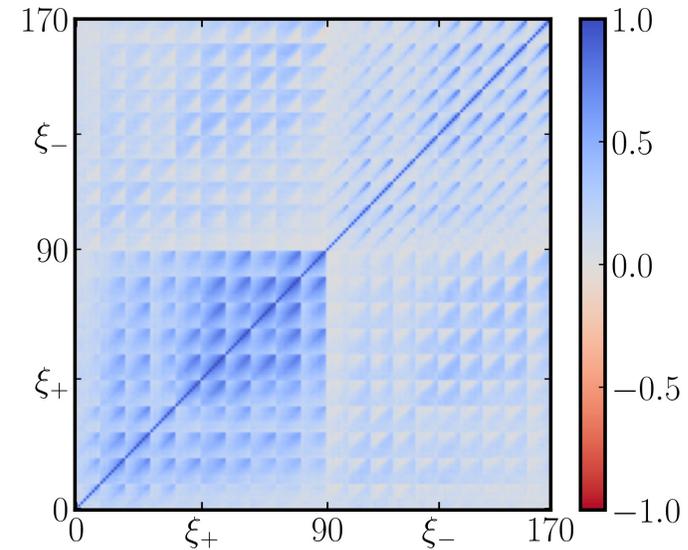


Sampling A_s

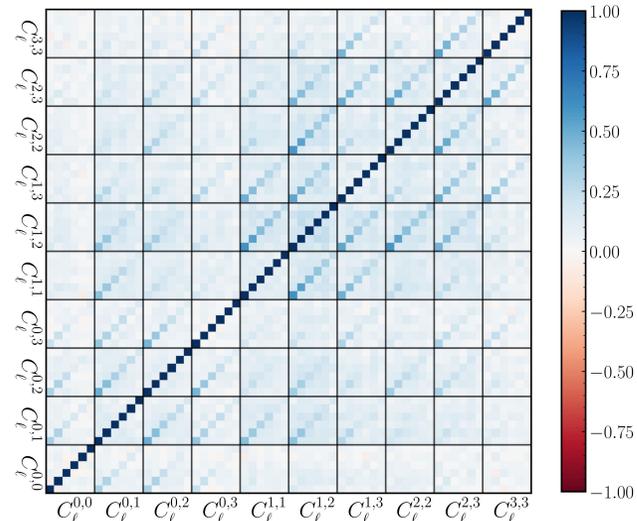


Cosmic shear power spectra

- Fourier space measurements of cosmic shear are complementary to real space measurements.
- ξ_{\pm} measurements are strongly correlated across scales (i.e. the covariance has large off-diagonal contributions). C_{ℓ} s are almost uncorrelated across multipoles.
- C_{ℓ} measurements are usually based on a pixelized map of the shear field (estimated with galaxy shapes), which can have a complicated window function.
- Partial sky coverage causes C_{ℓ} to be a biased estimator. We correct for this using the Pseudo- C_{ℓ} method.

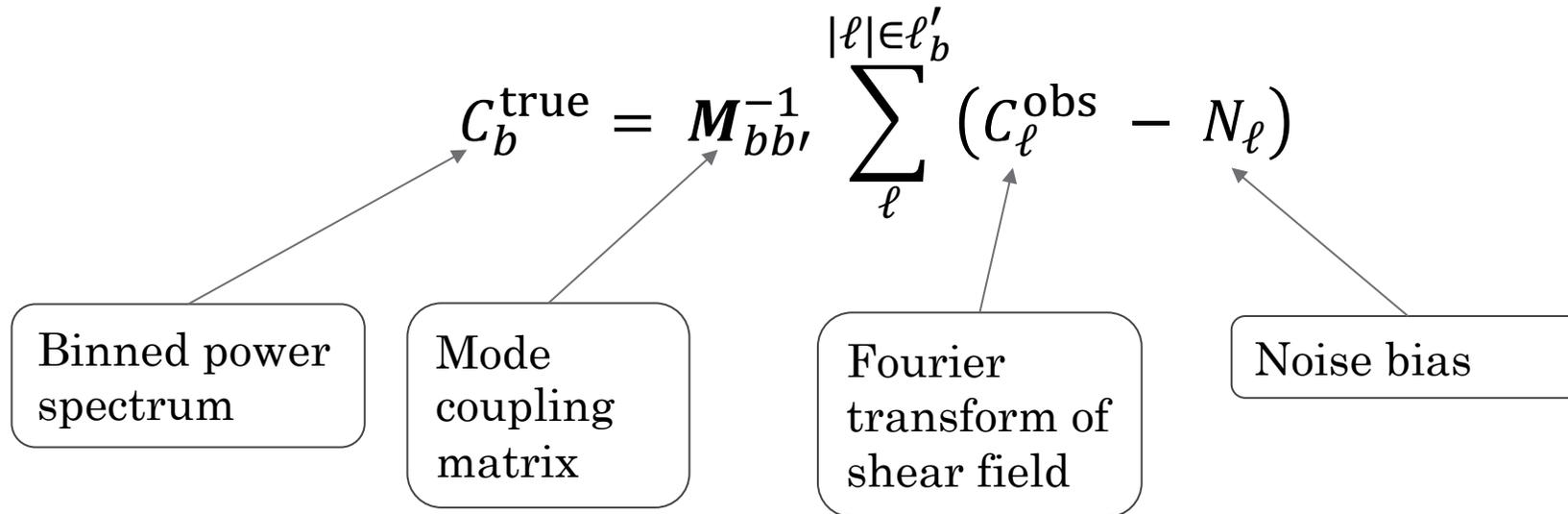


Real space correlation matrix



Fourier space correlation matrix

The Pseudo- C_ℓ method



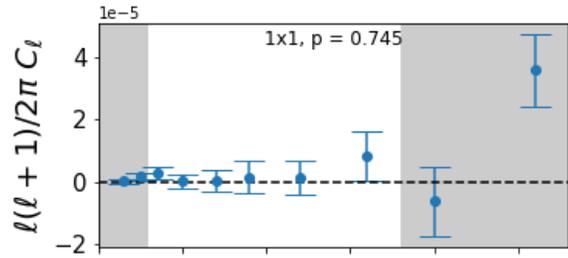
Implemented in NaMaster (Alonso+ 2019).

Computing spectra in 4 tomographic redshift bins of width $\Delta z = 0.3$ between $0.3 < z < 1.5$.

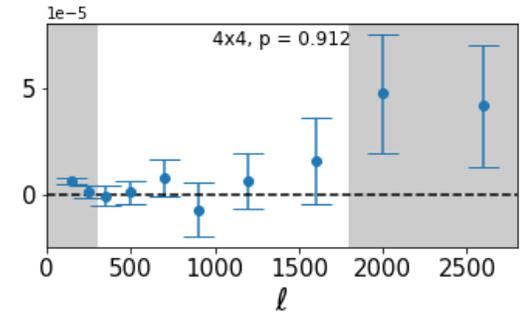
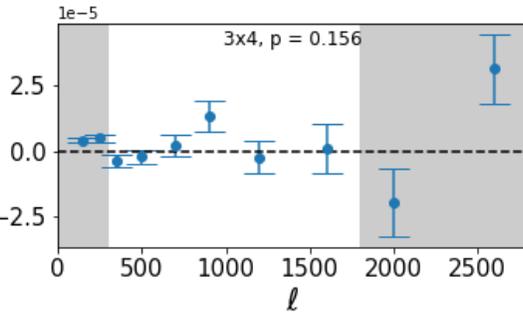
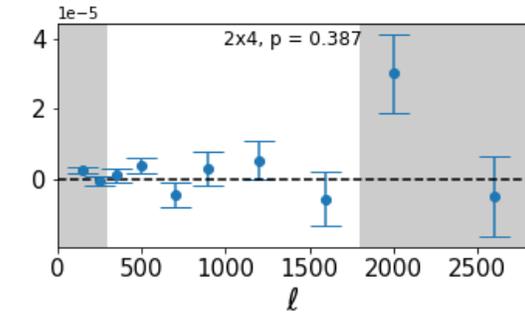
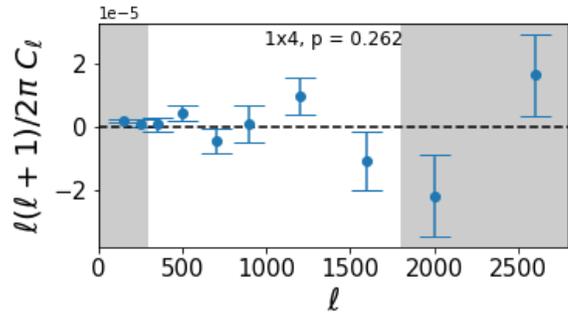
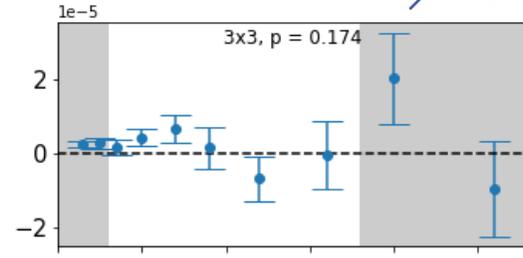
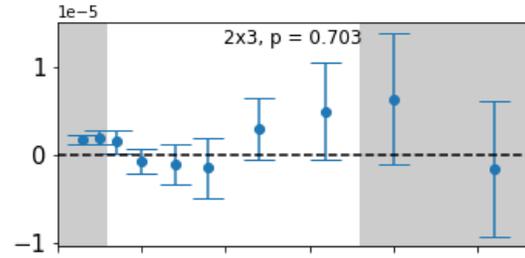
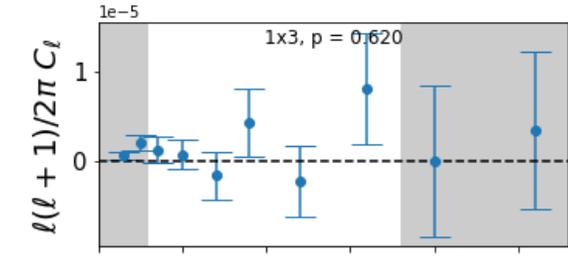
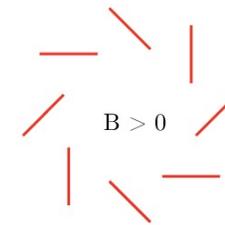
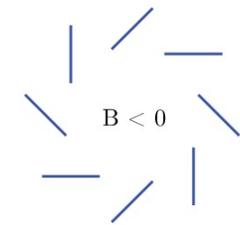
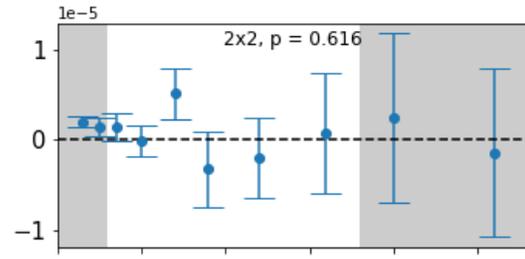
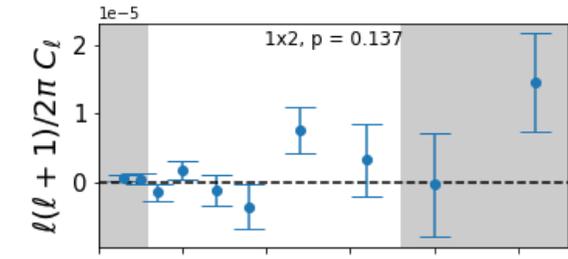
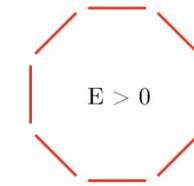
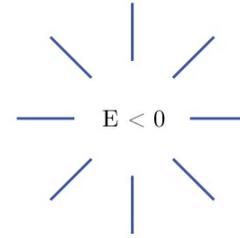
Covariance

- Covariance estimated from 1404 mock catalogs created following Shirasaki+ 2019:
- Use the full-sky lensing simulations of Takahashi+ 2017 combined with the observed photometric redshifts and angular positions of real galaxies:
 1. Set the RA and Dec of the survey window in the full-sky realization.
 2. Populate source galaxies on the light-cone using original angular positions and redshifts of the observed galaxies.
 3. Rotate the shape of each source galaxy at random to erase the real lensing signal.
 4. Add the lensing shear on each source galaxy using the lensing simulations.

No significant detection of B-modes



p-value = 0.50



← Increasing redshift

Increasing redshift →