

Large Scale Structure of the Universe

Rogério Rosenfeld
IFT-UNESP/ICTP-SAIFR/LIneA

Dark Energy Survey & LSST



Five lectures on Cosmology and Large Scale Structure

Lecture I: The average Universe

Lecture II: Distances and thermal history

Lecture III: The perturbed Universe

Lecture IV: Theoretical challenges and surveys

→ Lecture V: Observational cosmology with LSS

Plan for Lecture V:

V.1 – How to find the best model of the Universe?

V.2 – DES results

V.1- How to find the best model of the Universe?

V.1.1 – Likelihood analysis

V.1.2 – Observables in photometric surveys

V.1.3 – Angular correlation function of clustering

V.1.4 – Angular correlation function of shapes

V.1.5 – Theoretical modelling of the observables

V.1.1 – Likelihood analysis

The main goal is to determine what is the best model that describes our Universe.

Easy steps:

- Pick a probe
- Pick a model
- Compute predictions from the model for a given set of parameters $\{p\}$
- Get some data
- Compare model predictions with data
- Find the best model with the corresponding values of parameters

Put all steps together in the so-called likelihood function:

$$\mathcal{L}(\{p\}) \propto \exp \left\{ -\frac{1}{2} \left(\mathcal{O}^{\text{th}} - \mathcal{O}^{\text{obs}} \right)_i^T \text{Cov}_{ij}^{-1} \left(\mathcal{O}^{\text{th}} - \mathcal{O}^{\text{obs}} \right)_j \right\}$$

Theoretical prediction depends on the model and its parameters.

Observations depend on the experiment.

The covariance matrix basically reflects the uncertainty in the experimental measurement.

Best model: maximize likelihood (actually, explore probability distributions of the parameters given data using MCMC – “posterior distributions” – more later)

Sounds pretty easy!

V.1.2 – Observables in photometric surveys

Imaging surveys have two types of information:

- Positions of galaxies (clustering)
- Shapes of galaxies (shear)

Each object comes with photometric redshift information.
There are several algorithms to estimate redshift.

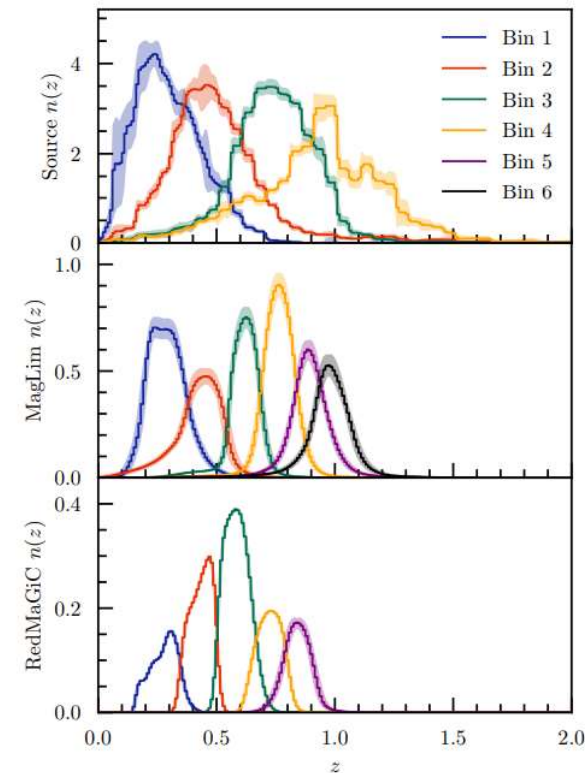
This is a major undertaking.

Some algorithms in DES: RedMagic, DNF, SOMPZ,...

Redshift information is contained in the redshift distribution:
 $n(z)$ for each redshift bin.

Observables: **angular correlation functions** for different redshift bins.

Giambiagi School 2023



V.1.3 – Angular correlation function of clustering

Angular correlation function of the galaxy clustering:

$$w^{ij}(\theta) = \langle \delta_g^i(\hat{n}) \delta_g^j(\hat{n} + \theta) \rangle$$

i,j= redshift bins

Measurements: one can “pixelize” the galaxy catalogue – way to divide the sphere into pixels (Healpix code)

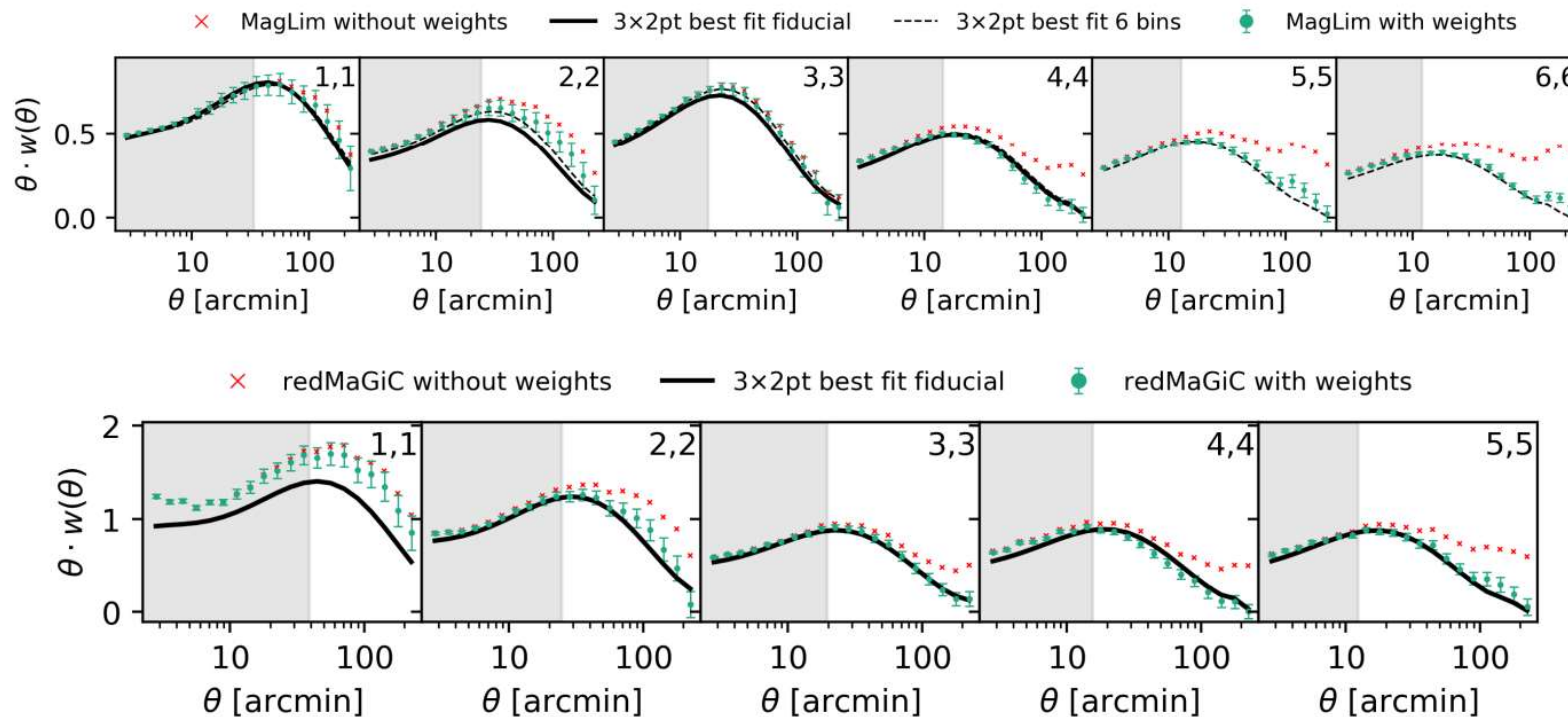
Pixel-based estimator for the angular correlation function:

$$\hat{w}^{ij}(\theta) = \sum_{\alpha=1}^{N_{Pix}} \sum_{\beta=1}^{N_{Pix}} \left(\frac{N_{\alpha}^i - \bar{N}^i}{\bar{N}^i} \right) \left(\frac{N_{\beta}^j - \bar{N}^j}{\bar{N}^j} \right) \Theta_{\alpha,\beta}$$

Number of galaxies
in pixel α .

=1 if pixels are
separated by an
angle within the
angular bin size
=0 otherwise

Measurements must be corrected by systematic effects including masking of pixels (due to bright sources), different exposures, different atmospheric conditions, etc.
 Example: results for DES-Y3 (there are 2 catalogues called MagLim and RedMagic)



2105.13540

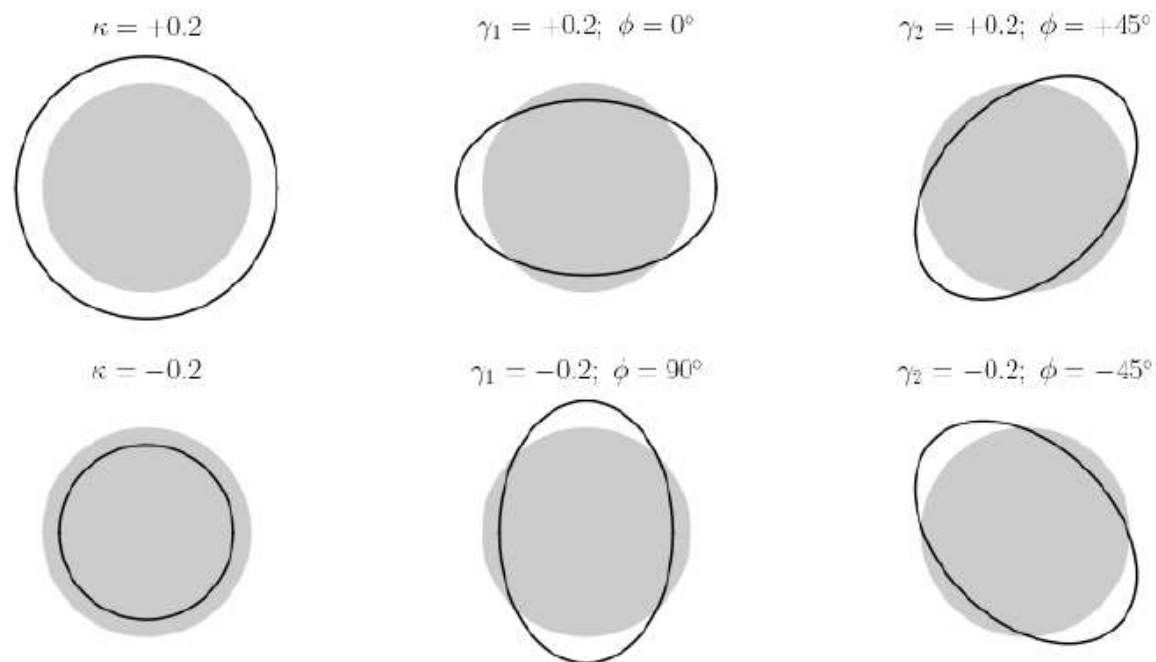
V.1.4 – Angular correlation function of shapes

Shapes of galaxies can be distorted by matter between the galaxy and us:
Usually the effect is small: weak gravitational lensing
(there are cases of strong distortions as well – strong gravitational lensing).



Gravitational lensing in the
JWST Deep Field

Weak gravitational lensing can be characterized by the magnification of the image (convergence κ) and the distortion (2-component shear $\gamma_{1,2}$).



Thesis of G. Giannini

Figure 1.8: Illustration of the effect of the convergence κ and shear field components $\gamma_{1,2}$ on a spherical source. For positive values of convergence, the image is magnified, vice versa for negative values. The shear components produce an anisotropic stretching of the image in different directions.

One can measure the ellipticity of a galaxy.

However, the shape before distortion was already elliptical.

Final shape is the sum of original ellipticity and the distortion (for small distortions).

$$\varepsilon = \varepsilon^0 + \gamma$$

Original shape is almost randomly oriented with some dispersion: shape noise

Original shape is not totally random due to gravitational interactions (“tidal torque”): intrinsic alignments have to be taken into account.

What one wants to measure is the 2-d shear field. DES uses an algorithm called METACALIBRATION to estimate the shear maps and parametrizes the uncertainties with a multiplicative (m) and additive (c) biases: $\gamma = m \gamma + c$

The cosmological signal is in the angular correlation of shapes:

$$\xi_{\pm} = \langle \gamma_{t,a} \gamma_{t,b} \rangle_{ab} \pm \langle \gamma_{\times,a} \gamma_{\times,b} \rangle_{ab}$$

In practice, the shear angular correlation function is estimated as:

$$\xi_{\pm}^{ij}(\theta) = \frac{\sum_{ab} w_a w_b \left(\hat{e}_{t,a}^i \hat{e}_{t,b}^j \pm \hat{e}_{\times,a}^i \hat{e}_{\times,b}^j \right)}{\sum_{ab} w_a w_b R_a R_b},$$

Sum is over galaxies in redshift bins (i,j) with weights w and a “response” R.
For details see DES papers 2105.13544 and 2105.13543

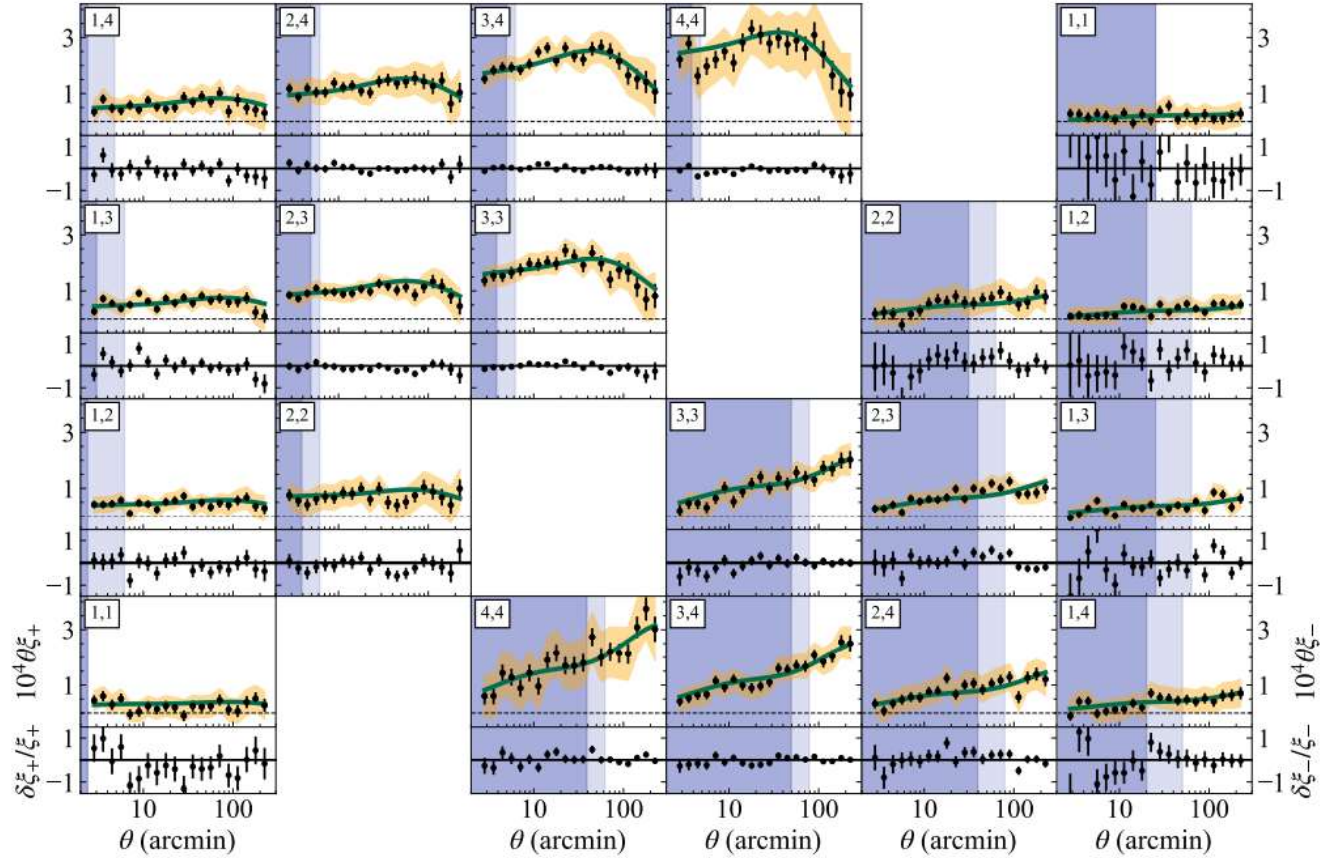


FIG. 3: Measured tomographic DES Y3 cosmic shear two-point correlation functions: $\xi_+(\theta)$ (left) and $\xi_-(\theta)$ (right), scaled by the angular separation, θ , to emphasize differences relative to the best-fit model (upper panels). The correlation functions are measured for each redshift bin pair, indicated by the label and the error bar represents the square root of the diagonal of the analytic covariance matrix. The best-fit Λ CDM theoretical prediction from the cosmic shear-only tomographic analysis is denoted by a green line. Scales excluded from the analysis, due to their sensitivity to small-scale systematics, are shaded in light blue for the *Fiducial* analysis and darker blue for the Λ CDM-*Optimized* analysis. The signal-to-noise of the measurement is 40 using all angular scales and 27 (31) using the *Fiducial* (Λ CDM-*Optimized*) scale-selection. For comparison, the yellow shaded region shows the Y1 uncertainty, with a factor of $\sim \sqrt{2}$ lower signal-to-noise. The lower panels plot the fractional difference between the measurements and best-fit, $\delta\xi_{\pm}/\xi_{\pm} = (\xi_{\pm} - \xi_{\pm}^{\text{theory}})/\xi_{\pm}^{\text{theory}}$. We find that the χ^2 per effective d.o.f of the Λ CDM model is $237.7/222.2 = 1.07$, and the p -value is 0.223.

V.1.5 – Theoretical modelling of the observables

V.1.5.a

Relation between angular correlation function $W(\theta)$
and angular power spectrum $C(l)$

• Projected 2-dimensional field $\delta_A^i(\hat{n})$
 \uparrow direction in sky
 i : redshift bin
 A : observable
 (assume $A=g$)

• Decompose 2d field in spherical harmonics

$$\delta(\hat{n}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} Y_{lm}(\hat{n})$$

• Angular correlation function $W(\theta) = \langle \delta(\hat{n}) \delta(\hat{n}+\theta) \rangle =$

$$= \sum_{lm} \sum_{l'm'} \underbrace{\langle a_{lm} a_{l'm'} \rangle}_{\substack{= \delta_{ll'} \delta_{mm'} \\ C_l}} Y_{lm}(\hat{n}) Y_{l'm'}(\hat{n}+\theta) = \sum_{lm} C_l Y_{lm}(\hat{n}) Y_{lm}(\hat{n}+\theta)$$

addition theorem: $\sum_{m=-l}^l Y_{lm}(\hat{n}) Y_{lm}(\hat{n}+\theta) = \frac{2l+1}{4\pi} P_l(\cos\theta)$

$$\cdot \quad \boxed{\frac{\delta_{ij}}{g_i g_j} W(\theta) = \sum_l \frac{2l+1}{4\pi} \frac{C_{ij}^{(l)}}{g_i g_j} P_l(\cos\theta)}$$

$C_{ij}^{(l)}$: galaxy clustering angular power spectrum

• there are similar equations for other observables (shear, galaxy shear)

• there is an approximation $W(\theta) = \int_0^\infty \frac{dl}{2\pi} l C(l) J_0(l\theta)$
(implemented in CCL) J_0 Bessel function

V.1.5.b

Modelling angular power spectra from 3D power spectrum

• Projected 2 dimensional field

$$\delta_A^i(\hat{m}) = \int d\chi \underbrace{g_A^i(\chi)}_{\substack{\text{window or selection function} \\ \cdot \text{ a redshift bin } i \text{ for probe A.}}} \underbrace{\delta_m(\hat{m}, \chi)}_{\substack{\text{3D matter density fluctuation}}}$$

\nearrow observable \uparrow direction in sky (angular variable)

$$\chi \equiv \chi(z)$$

comoving distance

- Fourier transform $\delta_m(\hat{n}, x)$:

$$\delta_m(\hat{n}, x) = \int \frac{d^3 k}{(2\pi)^3} \delta_m(\vec{k}, x) e^{i \vec{k} \cdot \hat{n} x / z}$$

- Use identity $e^{i \vec{k} \cdot \hat{n} x} = 4\pi \sum_{l, m} (i)^l j_l(kx) Y_{lm}(\hat{n}) Y_{lm}^*(\hat{k})$

to write:

$$\delta_m(\hat{n}, x) = 4\pi \sum_{l, m} Y_{lm}(\hat{n}) i^l \int \frac{d^3 k}{(2\pi)^3} j_l(kx) \delta_m(\vec{k}, x) Y_{lm}^*(\vec{k})$$

- Hence:

$$\delta_A^i(\hat{n}) \equiv \sum_{l, m} q_{lm}^i Y_{lm}(\hat{n}) = \sum_{l, m} \left[(4\pi)(i)^l \int dx q_A^i(x) \times \int \frac{d^3 k}{(2\pi)^3} j_l(kx) \delta_m(\vec{k}, x) Y_{lm}^*(\vec{k}) \right] Y_{lm}(\hat{n})$$

$$\Rightarrow q_{lm}^i = \int dx q_A^i(x) \int \frac{d^3 k}{(2\pi)^3} \delta_m (4\pi)(i)^l j_l(kx) Y_{lm}^*(\vec{k})$$

• Angular power spectrum C_ℓ :

$$\langle a_{\ell m}^i a_{\ell' m'}^j \rangle = \int d\chi_1 q^i(\chi_1) \int d\chi_2 q^j(\chi_2) \times$$

$$\times \int \frac{d^3 k_1}{(2\pi)^3} \frac{d^3 k_2}{(2\pi)^3} \underbrace{\langle \delta_m(k_1) \delta_m(k_2) \rangle}_{(2\pi)^3 \delta^3(\vec{k}_1 - \vec{k}_2) P_m(k)} \times (4\pi)^2 j_\ell(k_1 \chi_1) j_{\ell'}(k_2 \chi_2) \underbrace{Y_{\ell m}^*(\hat{k}_1)}_{Y_{\ell' m'}(\hat{k}_2)}$$

$$= \int d\chi_1 q^i(\chi_1) \int d\chi_2 q^j(\chi_2) \int dk k^2 \frac{(4\pi)^2}{(2\pi)^3} j_\ell(k\chi_1) j_{\ell'}(k\chi_2) \times$$

$$\times \underbrace{\int d\hat{k} Y_{\ell m}^*(\hat{k}) Y_{\ell' m'}(\hat{k})}_{\delta_{\ell\ell'} \delta_{mm'}}$$

$$= \frac{2}{\pi} \int dk k^2 P_m(k) Y_{\ell}^i(k) Y_{\ell}^j(k) \equiv C_\ell$$

$$Y_{\ell}^{i,j} \equiv \int d\chi q^{i,j}(\chi) j_{\ell}(k\chi)$$

- The famous Limber approximation (1953)

Integral of spherical Bessel functions j_ℓ are a nightmare to compute

Identity:
$$\frac{2}{\pi} \int dk k^2 j_\ell(kx_1) j_\ell(kx_2) = \frac{1}{x_1^2} \delta(x_1 - x_2)$$

If $P(k)$ is a slowly varying function of k :

$$\frac{2}{\pi} \int dk k^2 j_\ell(kx_1) j_\ell(kx_2) P(k) \simeq \frac{\delta(x_1 - x_2)}{x_1^2} P\left(k = \frac{l + 1/2}{x_1}\right)$$

\uparrow
 maximum of Bessel function

and

$$C(l)_{AB}^{ij} = \int d\chi \frac{q_A^i(\chi) q_B^j(\chi)}{\chi^2} P\left(k = \frac{l + 1/2}{\chi}\right)$$

actually not used in DES-Y3

V.1.5.c - Selection function for galaxies

$$q_g^i(x) = \underbrace{b^i(k, z(x))}_{\text{galaxy bias in bin } i} \underbrace{m_g^i(z(x))}_{\text{redshift distribution}} \underbrace{\frac{dz}{dx}}_{\text{transforms } dx \rightarrow dz}$$
$$\int dz m_g^i(z) = 1$$

note: we are neglecting an effect called redshift space distortions (RSD) that is small since redshift is not accurate.

See complete modelling in 1911.11947

Selection function for shear is different

In summary, the theoretical modelling for angular power spectra:

$$C_{\kappa\kappa}^{ij}(\ell) = \int d\chi \frac{q_{\kappa}^i(\chi) q_{\kappa}^j(\chi)}{\chi^2} P\left(\frac{\ell + \frac{1}{2}}{\chi}, z(\chi)\right), \quad (4)$$

$$C_{\delta_g\kappa}^{ij}(\ell) = \int d\chi \frac{q_{\delta}^i\left(\frac{\ell + \frac{1}{2}}{\chi}, \chi\right) q_{\kappa}^j(\chi)}{\chi^2} P\left(\frac{\ell + \frac{1}{2}}{\chi}, z(\chi)\right), \quad (5)$$

$$C_{\delta_g\delta_g}^{ij}(\ell) = \int d\chi \frac{q_{\delta}^i\left(\frac{\ell + \frac{1}{2}}{\chi}, \chi\right) q_{\delta}^j\left(\frac{\ell + \frac{1}{2}}{\chi}, \chi\right)}{\chi^2} P\left(\frac{\ell + \frac{1}{2}}{\chi}, z(\chi)\right), \quad (6)$$

$$\begin{aligned} q_{\kappa}^i(\chi) &= \frac{3H_0^2\Omega_m\chi}{2a(\chi)} \int_{\chi}^{\chi_h} d\chi' \left(\frac{\chi' - \chi}{\chi}\right) n_{\kappa}^i(z(\chi')) \frac{dz}{d\chi'}, \\ q_{\delta}^i(k, \chi) &= b^i(k, z(\chi)) n_{\delta}^i(z(\chi)) \frac{dz}{d\chi}. \end{aligned} \quad (7)$$

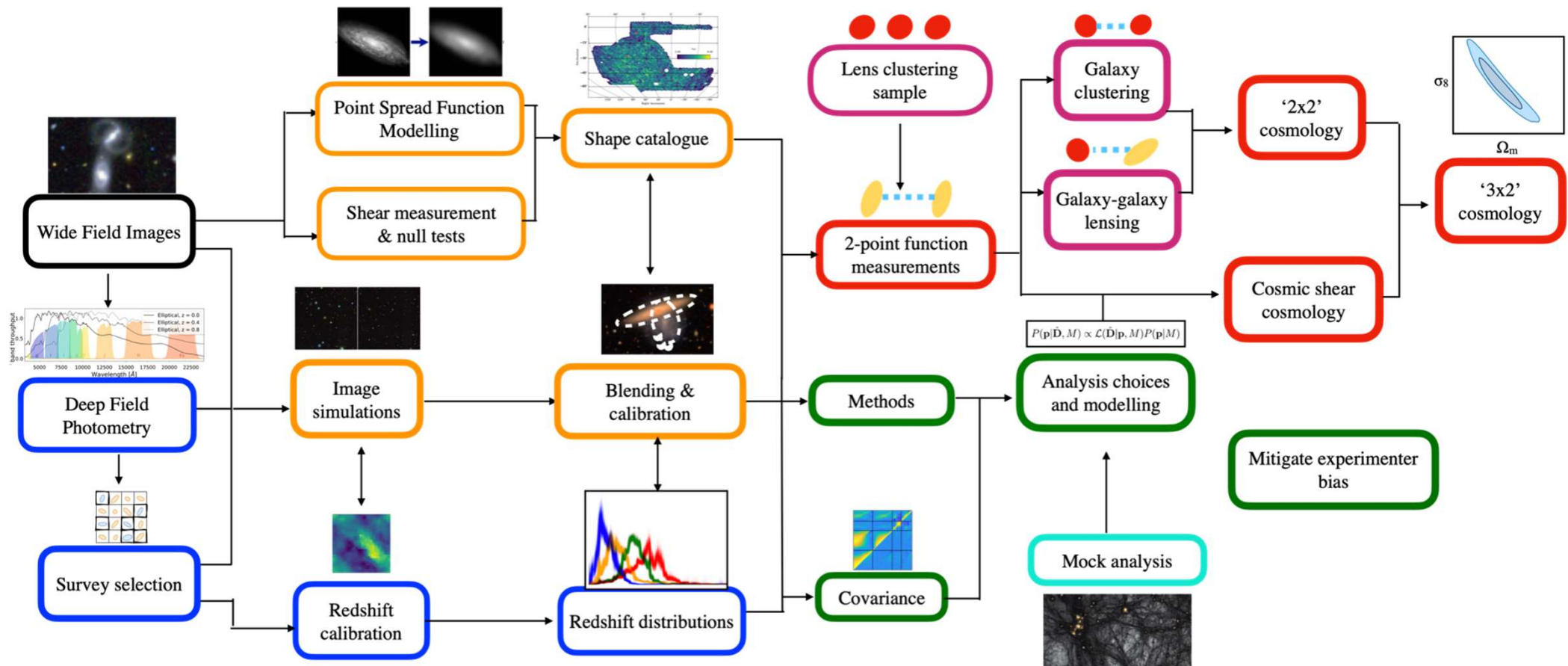
From DES-Y3 covariance paper
<https://arxiv.org/abs/2012.08568>

V.2- DES Y3 results

V.2.1 – From pixels to cosmology

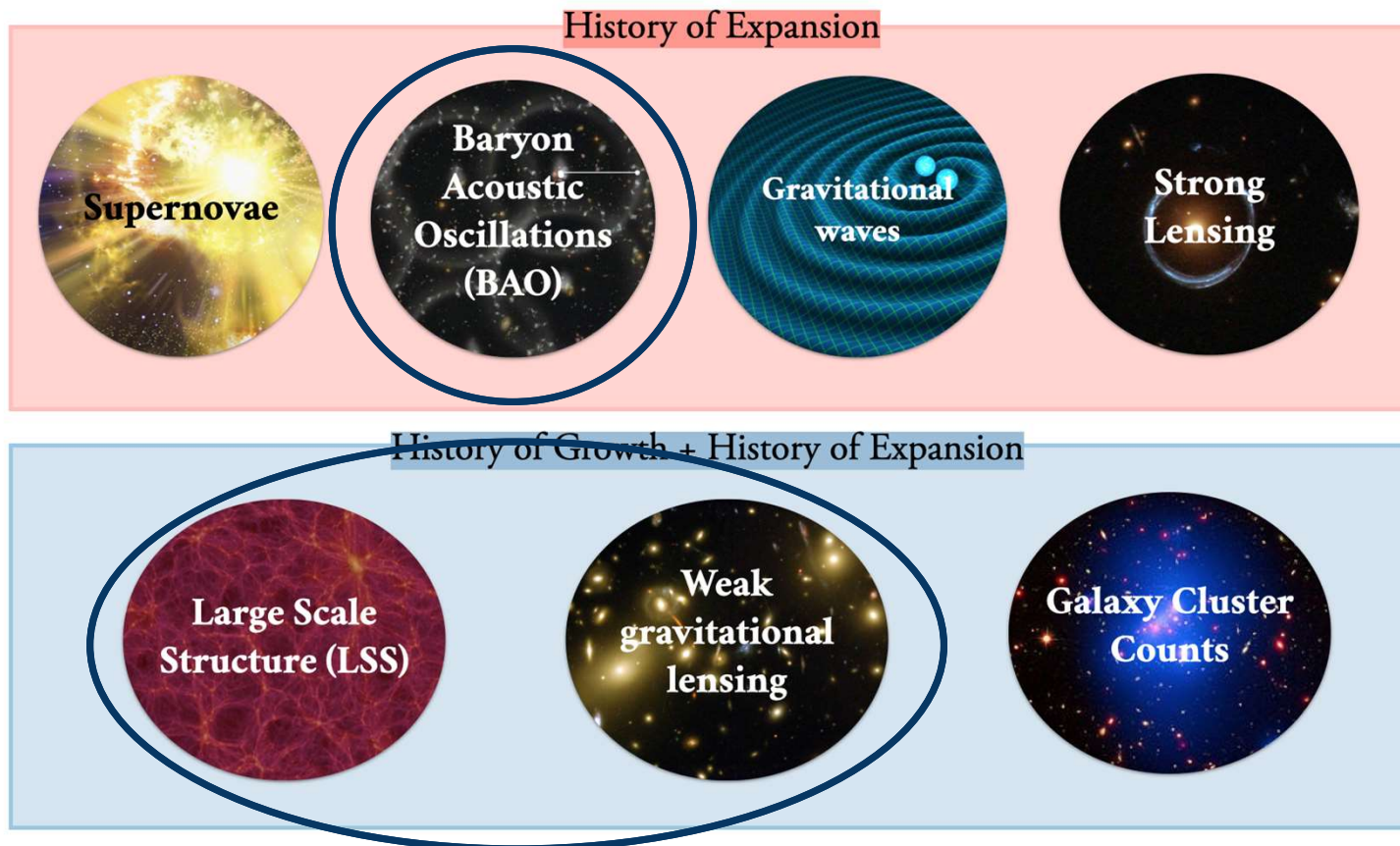
V.2.2 – 3x2pt analysis

V.2.1 – From pixels to cosmology



From A. Amon - DES-Y3 webinar

Cosmic probes within DES



DES-Y3 cosmology: 30+3 papers

Main results for:

1. 3x2pt correlation functions [Our group contributed with covariance matrix study]
2. Baryon acoustic oscillation [Our group contributed to the harmonic space analysis]
3. Harmonic space analysis [Led by our group + C. Doux]

V.2.2 – 3x2pt analysis

2105.13549

A self-consistent combined analysis maximises the cosmological information and robustly constrains astrophysical & observational systematic priors in the analysis! Most sensitive to Ω_m and S_8 .

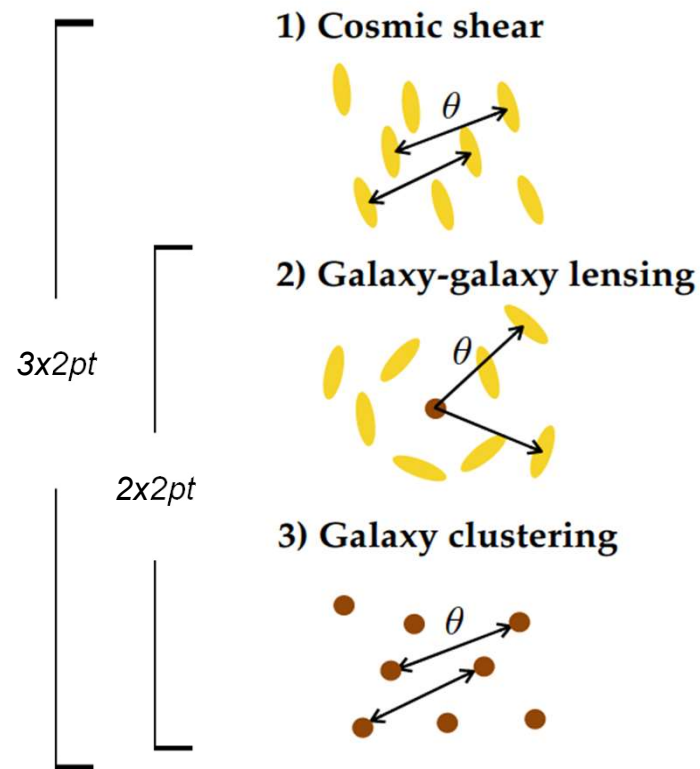
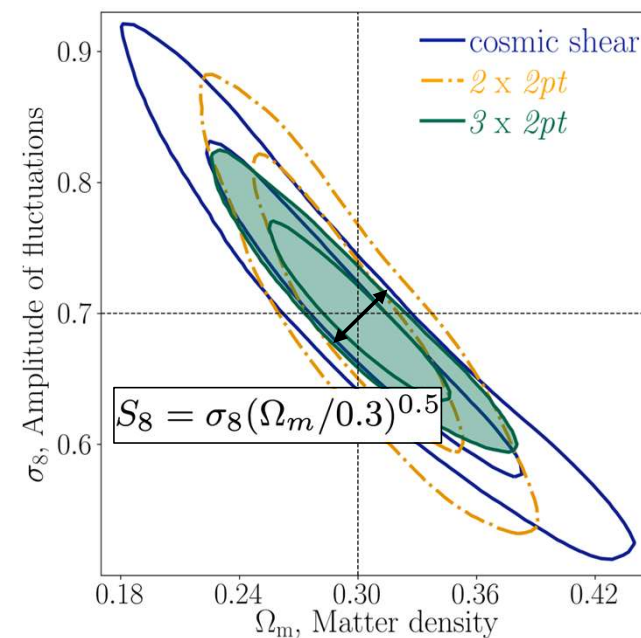


Image credit: Chihway Chang



How to estimate cosmological parameters?

Data vector:

$$\hat{\mathbf{D}} \equiv \{\hat{w}^i(\theta), \hat{\gamma}_t^{ij}(\theta), \hat{\xi}_{\pm}^{ij}(\theta)\}$$

Theoretical modelling
that depends on model M and
parameters \mathbf{p}

$$\mathbf{T}_M(\mathbf{p}) \equiv \{w^i(\theta, \mathbf{p}), \gamma_t^{ij}(\theta, \mathbf{p}), \xi_{\pm}^{ij}(\theta, \mathbf{p})\}$$

Gaussian likelihood that
depends on the covariance
matrix \mathbf{C}

$$\mathcal{L}(\hat{\mathbf{D}}|\mathbf{p}, M) \propto e^{-\frac{1}{2} [(\hat{\mathbf{D}} - \mathbf{T}_M(\mathbf{p}))^T \mathbf{C}^{-1} (\hat{\mathbf{D}} - \mathbf{T}_M(\mathbf{p}))]}$$

Posterior distribution of the
parameters that depend on
priors

$$P(\mathbf{p}|\hat{\mathbf{D}}, M) \propto \mathcal{L}(\hat{\mathbf{D}}|\mathbf{p}, M)P(\mathbf{p}|M)$$

Main issues in modeling correlation functions

- Photometric redshift uncertainties
- Galaxy bias relating galaxy with matter distributions
(does not affect shear)
- Intrinsic alignment of galaxies (does not affect clustering)
- Shear calibration (does not affect clustering)
- Baryonic effects in power spectrum of galaxies
- Nonlinear (gravity) clustering in the power spectrum

Issues in modeling are dealt with:

- Introduction of nuisance parameters to parametrize uncertainties
- Scale cuts designed to leave out small scales where nonlinear bias and baryonic effects are important.

These mitigation methods have to be implemented and tested against realistic simulations, with a requirement of accuracy in recovering cosmological parameters.

Most stringent cosmological constraints from a galaxy imaging survey.

Combined with external data: most stringent constraints overall

Test the six-parameter universe – the flat Λ CDM model:

$\{A_s, n_s, h, \Omega_m, \Omega_b \text{ and } \Omega_v\}$

σ_8 is a derived parameter (variance of linear fluctuations at scale of 8 Mpc/h)

Small extension- flat w CDM: + w (constant equation of state)

*Cosmological
parameters (7)*

*Nuisance
parameters
(30)*

Parameter	Prior	
Cosmology		
Ω_{m}	Flat	(0.1, 0.9)
$10^9 A_s$	Flat	(0.5, 5.0)
n_s	Flat	(0.87, 1.07)
Ω_{b}	Flat	(0.03, 0.07)
h	Flat	(0.55, 0.91)
$10^3 \Omega_{\nu} h^2$	Flat	(0.60, 6.44)
w	Flat	(−2.0, −0.33)
Lens Galaxy Bias		
$b_i (i \in [1, 4])$	Flat	(0.8, 3.0)
Lens magnification		
C_1^1	Fixed	1.21
C_1^2	Fixed	1.15
C_1^3	Fixed	1.88
C_1^4	Fixed	1.97
Lens photo-z		
$\Delta z_1^1 \times 10^2$	Gaussian	(−0.9, 0.7)
$\Delta z_1^2 \times 10^2$	Gaussian	(−3.5, 1.1)
$\Delta z_1^3 \times 10^2$	Gaussian	(−0.5, 0.6)
$\Delta z_1^4 \times 10^2$	Gaussian	(−0.7, 0.6)
$\sigma_{z,1}^1$	Gaussian	(0.98, 0.06)
$\sigma_{z,1}^2$	Gaussian	(1.31, 0.09)
$\sigma_{z,1}^3$	Gaussian	(0.87, 0.05)
$\sigma_{z,1}^4$	Gaussian	(0.92, 0.05)
Intrinsic Alignment		
$a_i (i \in [1, 2])$	Flat	(−5, 5)
$\eta_i (i \in [1, 2])$	Flat	(−5, 5)
b_{TA}	Flat	(0, 2)
z_0	Fixed	0.62
Source photo-z		
$\Delta z_s^1 \times 10^2$	Gaussian	(0.0, 1.8)
$\Delta z_s^2 \times 10^2$	Gaussian	(0.0, 1.5)
$\Delta z_s^3 \times 10^2$	Gaussian	(0.0, 1.1)
$\Delta z_s^4 \times 10^2$	Gaussian	(0.0, 1.7)
Shear calibration		
$m^1 \times 10^2$	Gaussian	(−0.6, 0.9)
$m^2 \times 10^2$	Gaussian	(−2.0, 0.8)
$m^3 \times 10^2$	Gaussian	(−2.4, 0.8)
$m^4 \times 10^2$	Gaussian	(−3.7, 0.8)



3x2pt results

We combine these into the **3x2pt** probe of large-scale structure.

A factor of 2.1 improvement in signal-to-noise from DES Year 1.

In Λ CDM:

$$S_8 = 0.776^{+0.017}_{-0.017} \quad (0.776)$$

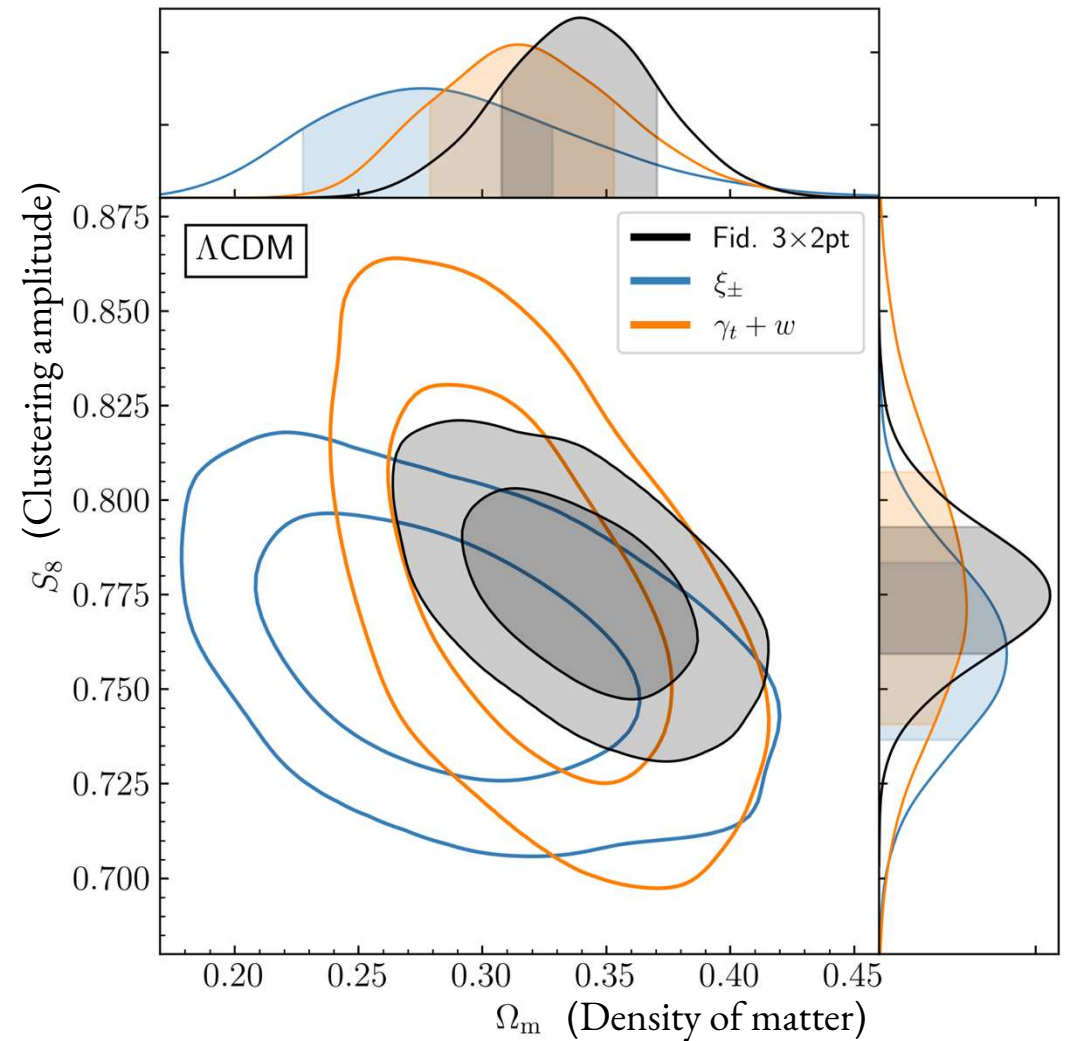
$$\Omega_m = 0.339^{+0.032}_{-0.031} \quad (0.372)$$

$$\sigma_8 = 0.733^{+0.039}_{-0.049} \quad (0.696)$$

In w CDM:

$$\Omega_m = 0.352^{+0.035}_{-0.041} \quad (0.339)$$

$$w = -0.98^{+0.32}_{-0.20} \quad (-1.03)$$



Consistency with Planck results

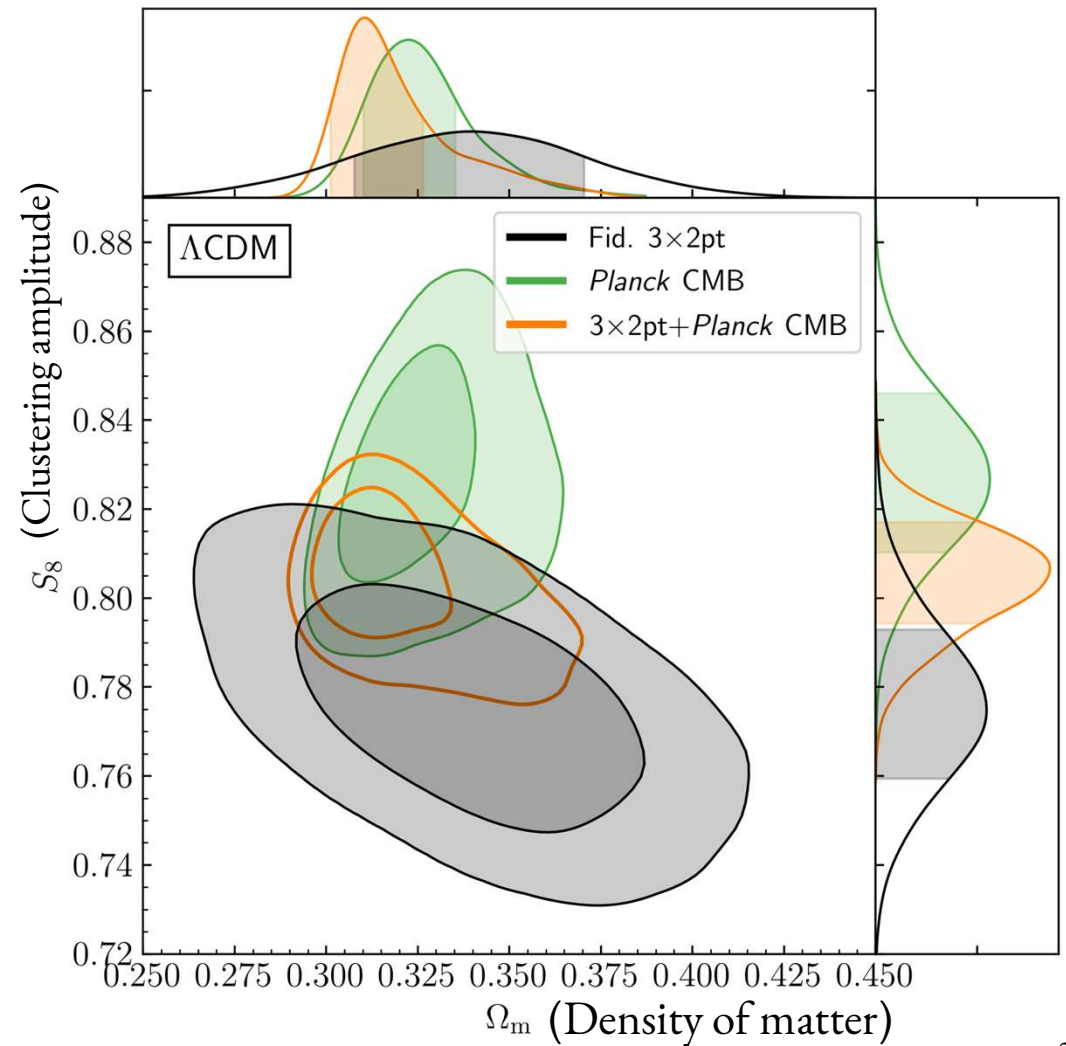
From DES-Y3 webinar

Low-z vs High-z in Λ CDM

We test the robustness of Λ CDM by comparing measurements of the clustering amplitude at low-redshift to the prediction from the cosmic microwave background (CMB) at high-redshift.

We find **no significant evidence of inconsistency** between **DES Y3 3 \times 2pt** and *Planck* CMB at 0.7-1.5 σ or $p=0.13$ -0.48.

Combining the results we find the orange contour.



Joint constraints

Combining all these data sets we find:

$$S_8 = 0.812^{+0.008}_{-0.008} \quad (0.815)$$

$$\text{In } \Lambda\text{CDM } \Omega_m = 0.306^{+0.004}_{-0.005} \quad (0.306)$$

$$\sigma_8 = 0.804^{+0.008}_{-0.008} \quad (0.807)$$

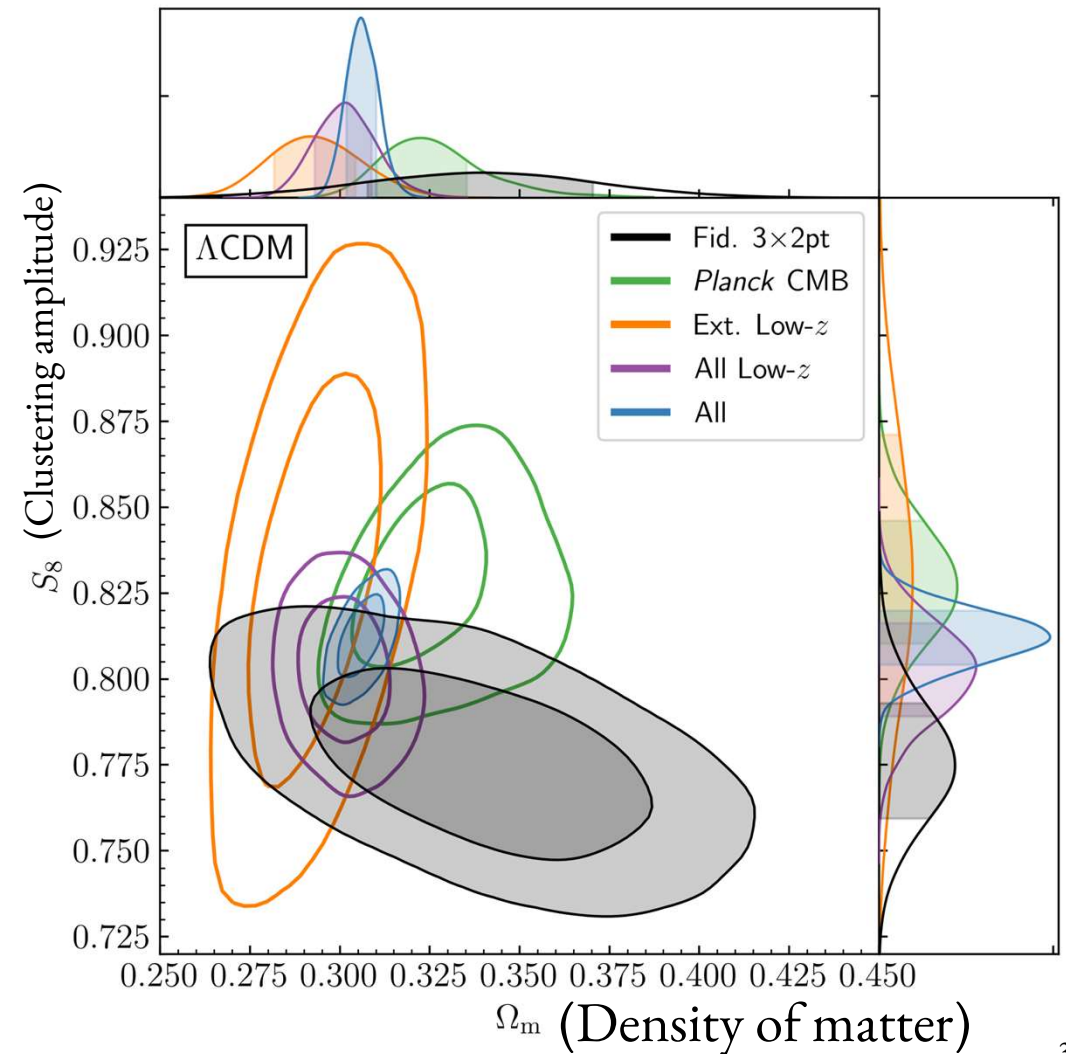
$$h = 0.680^{+0.004}_{-0.003} \quad (0.681)$$

$$\sum m_\nu < 0.13 \text{ eV (95\% CL)}$$

$$\text{In } w\text{CDM: } \sigma_8 = 0.810^{+0.010}_{-0.009} \quad (0.804),$$

$$\Omega_m = 0.302^{+0.006}_{-0.006} \quad (0.298),$$

$$w = -1.03^{+0.03}_{-0.03} \quad (-1.00)$$



Final remarks

Cosmological surveys are entering a new stage of precision!

Dark Energy Survey finishing the final data analysis.

Dark Energy Spectroscopic Instrument (DESI) is performing great:
Early Data Release in June 2023 with 1.2 million galaxies and quasars!

Euclid satellite launched July 1st 2023

Rubin Observatory's LSST will provide an unprecedented amount of data.
First light in 2024. Participants from Argentina, Brasil, Chile and Mexico!

More surveys: SPHEREx (2025), Roman Space Telescope (2027),

Amazing opportunities that come together with challenges:
Better understanding of nonlinear regime and galaxy bias for beyond Λ CDM models.

Joint studies with CMB: Planck, ACT, SO, S4, ...

Joint studies with Gravitational Waves:
LVK, Einstein Telescope, Cosmic Explorer, LISA, ...

LSST's eight Science Collaborations already in full action:
Data Challenges and Data Previews catalogues to test modelling, pipelines, etc.
Getting ready for real data!

Exciting times ahead! Come join the fun!

Gracias, Obrigado, Thank You