Aix*Marseille université
Socialement engagée

Euclid School 2023
Tomographic analysis of photometric galaxy clustering with Flagship 2.1

Vincent Duret

Supervisors : Stéphanie Escoffier, William Gillard

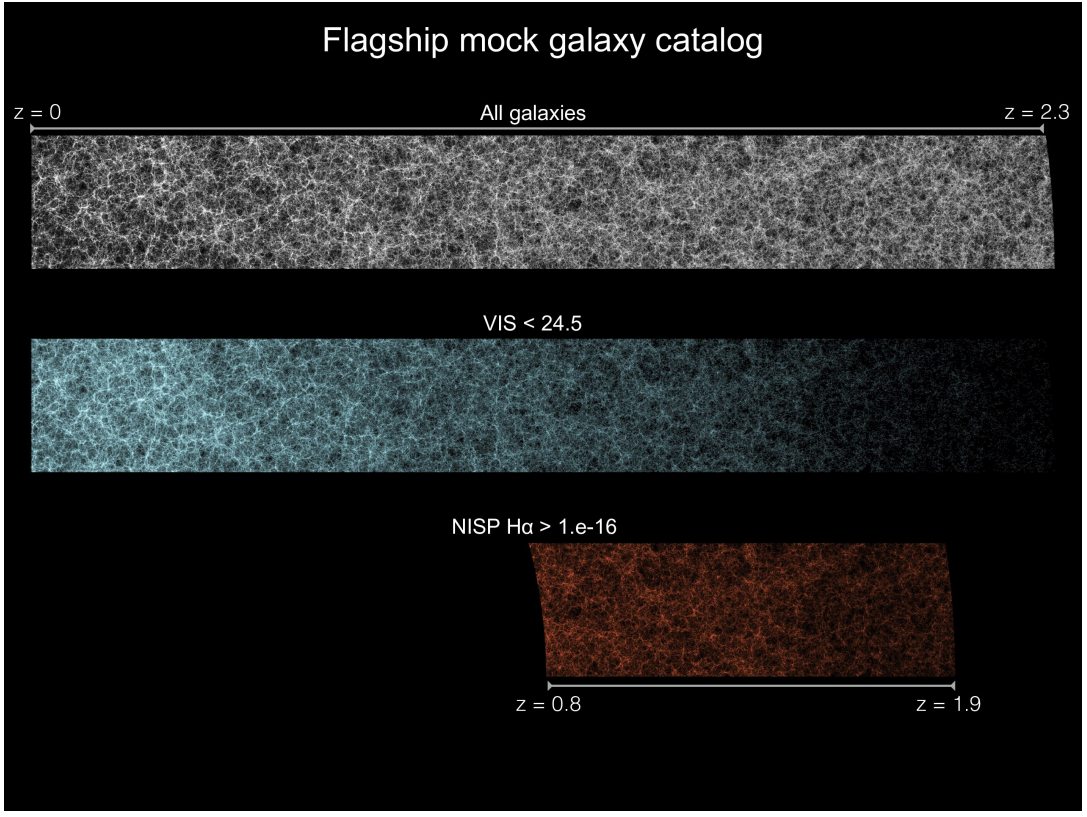
Outline

1) Photometric redshifts calibration

2) Data : Flagship and 2pcf measurement

3) Full-shape analysis for Euclid GCph KP 3

4) BAO analysis for Euclid GCph KP 10



Photometric redshifts calibration

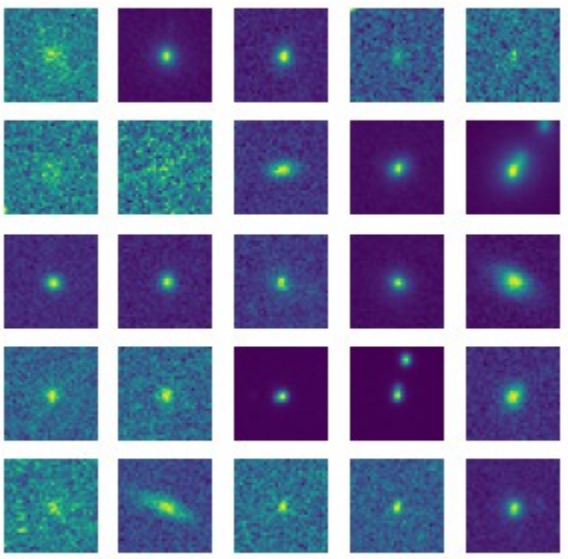
Goal : find the relationship between the color space (magnitudes) and redshift.

Template fitting : a set of SED templates is made using observations or modelizations. They can be shifted to any redshift and convolved with the transmission curves of a telescope before minimizing a χ^2 between templates and observations to infer a redshift

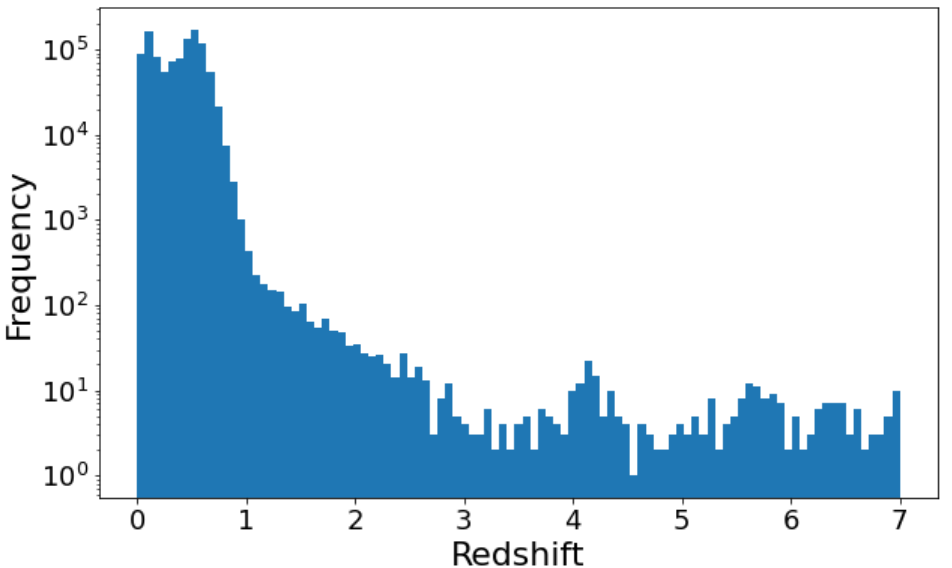
ML/DL : the relationship between colors and redshift is learnt by the algorithm thanks to a training on a galaxy dataset for which we have a spectroscopic redshift.

Photometric redshifts calibration

Idea : exploit images rather than extracted photometry → more information
SDSS data release 12
Input : 1059678 galaxy images in u,g,r,i,z bands of size 32 × 32 px
Labels : spectroscopic redshifts

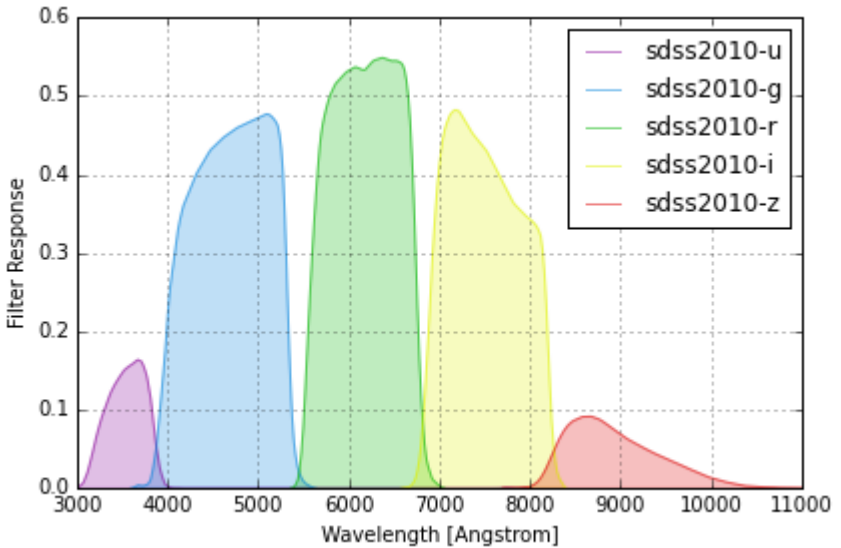


r band



Photometric redshifts calibration

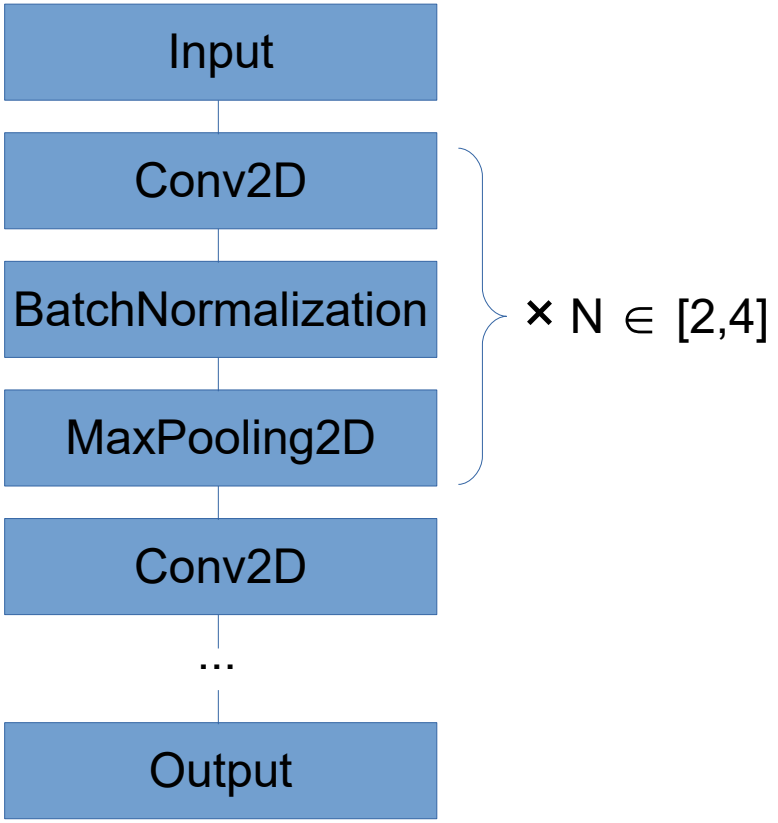
Idea : exploit images rather than extracted photometry → more information
SDSS data release 12
Input : 1059678 galaxy images in u,g,r,i,z bands of size 32 × 32 px
Labels : spectroscopic redshifts



SDSS u,g,r,i,z bands filter response

Photometric redshifts calibration

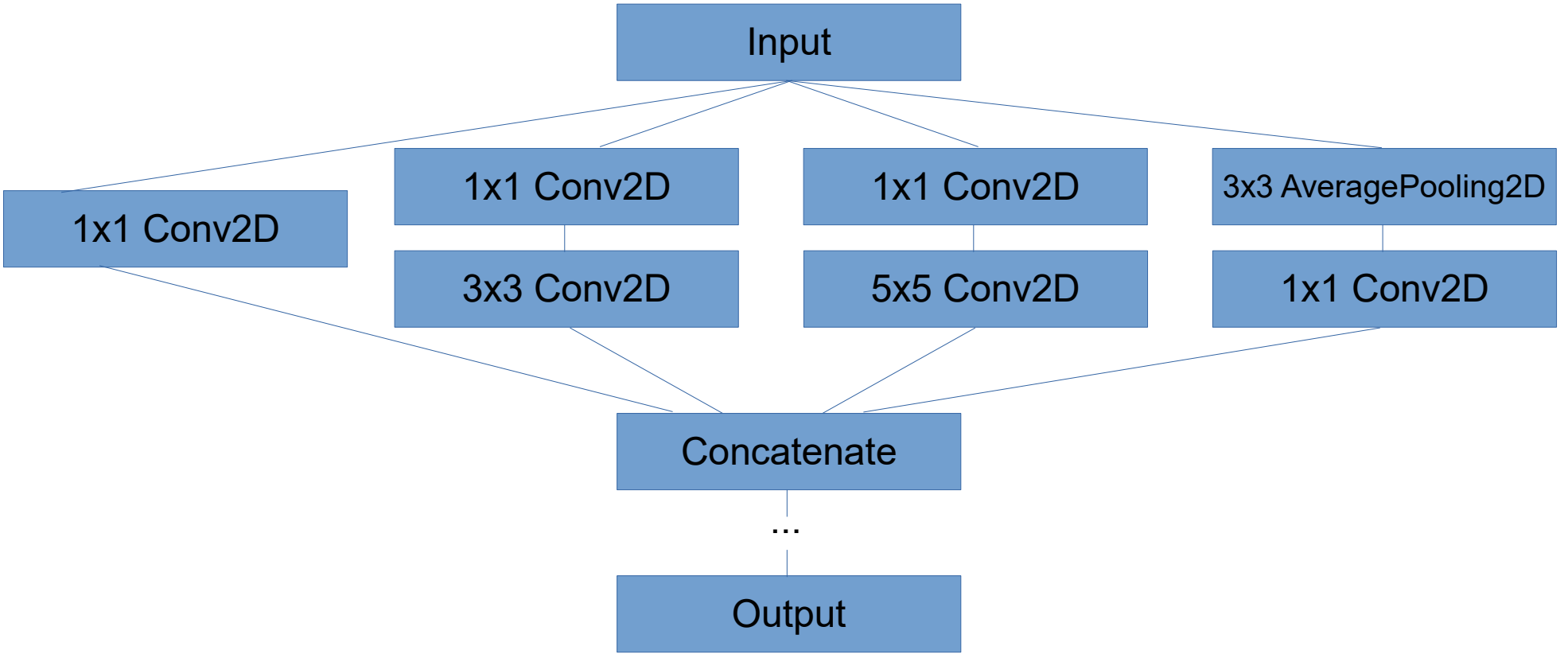
Tested neural networks : sequential CNN



Sequential CNN architecture

Photometric redshifts calibration

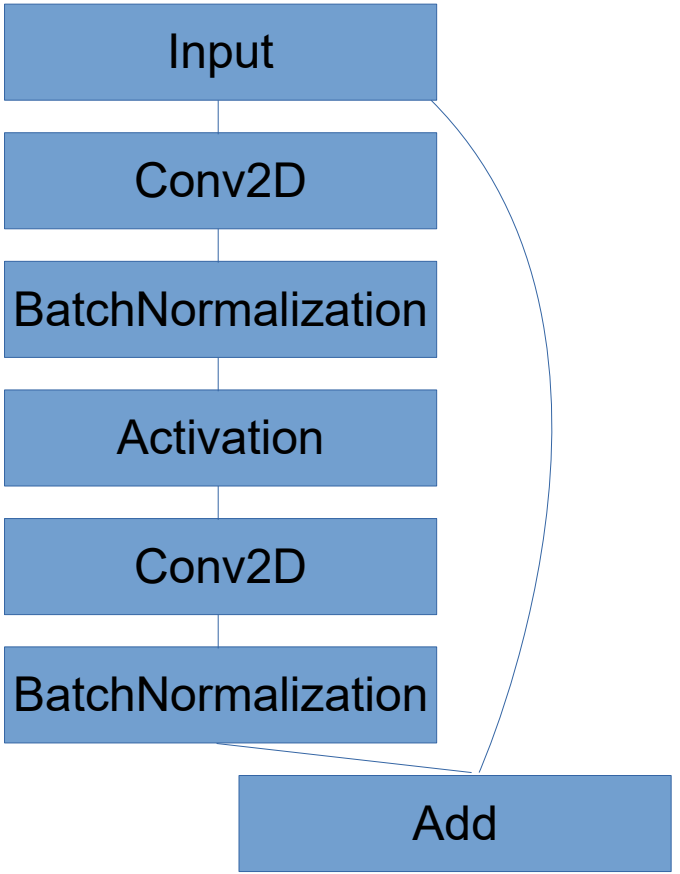
Tested neural networks : sequential CNN, inception CNN



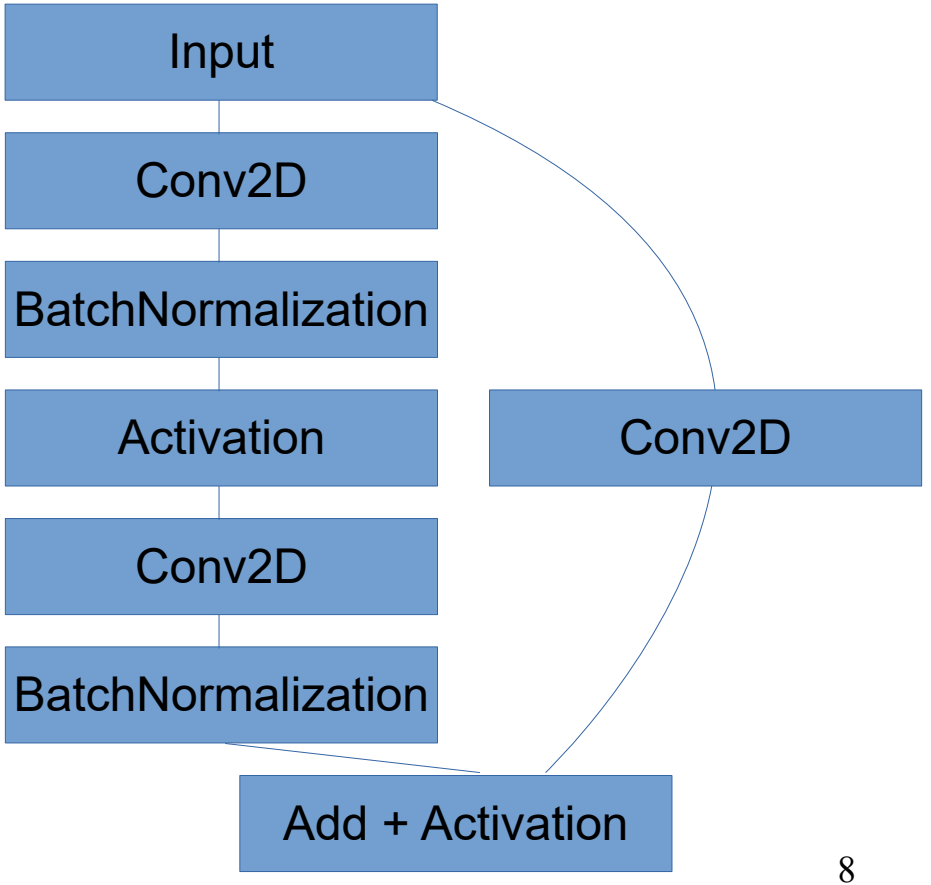
Inception block architecture (arXiv:1512.00567)

Photometric redshifts calibration

Tested neural networks : sequential CNN, inception CNN, ResNet34



Identity block



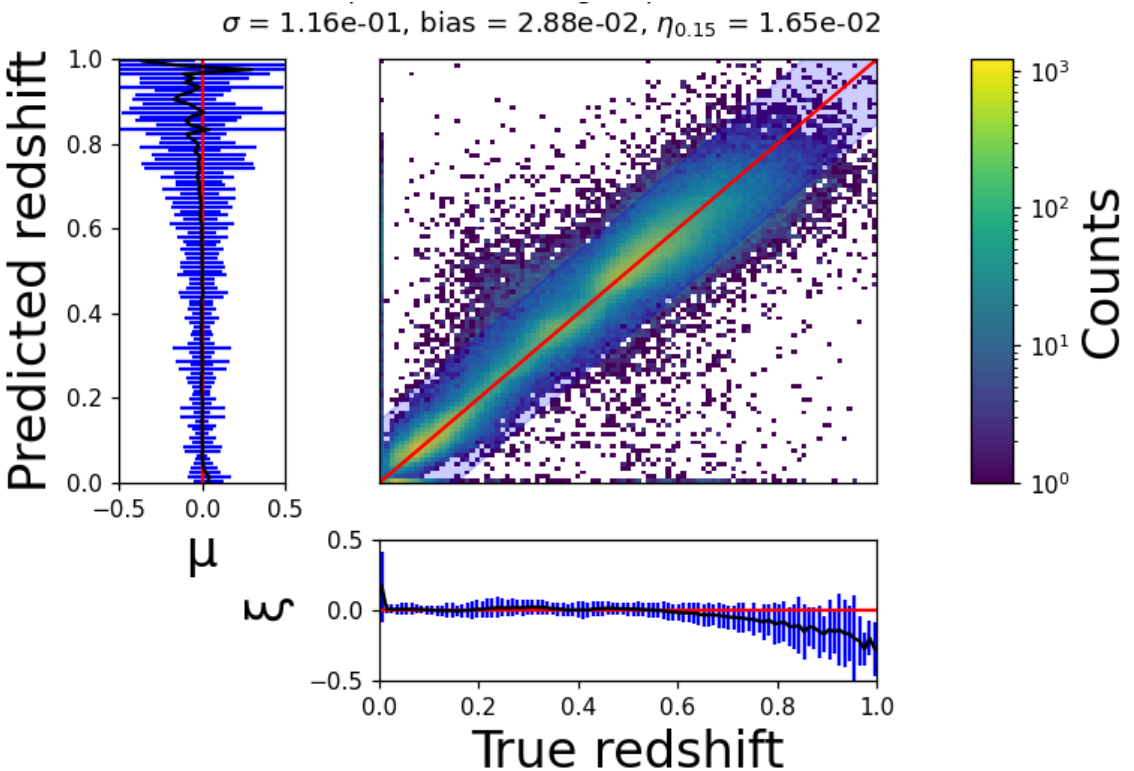
Conv block

Photometric redshifts calibration

Results :

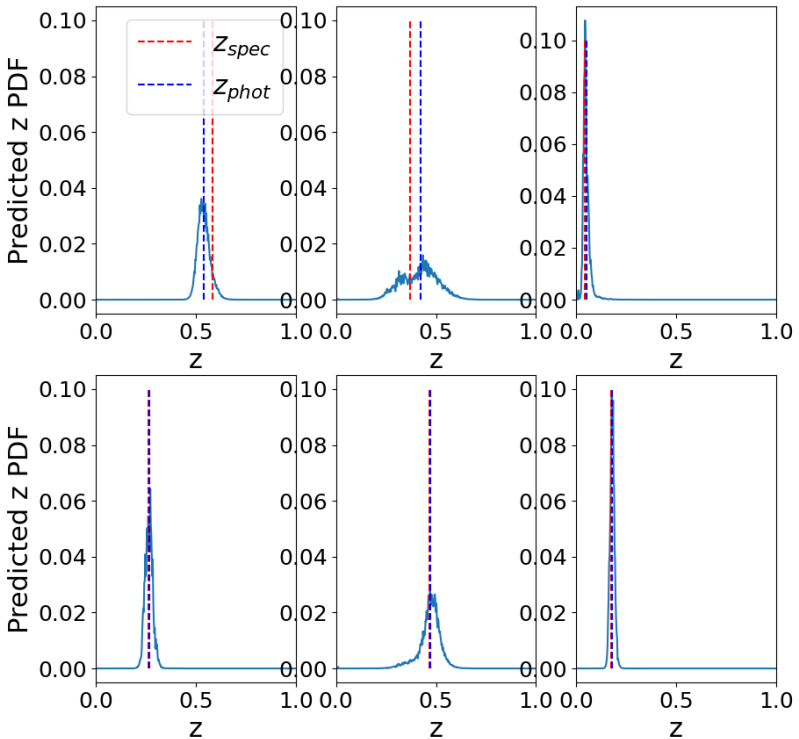
Obtained with ResNet34
$\sigma = 1.16 \times 10^{-1}$
Bias = 2.88×10^{-2}
$\eta_{0.15} = 1.65 \times 10^{-2}$

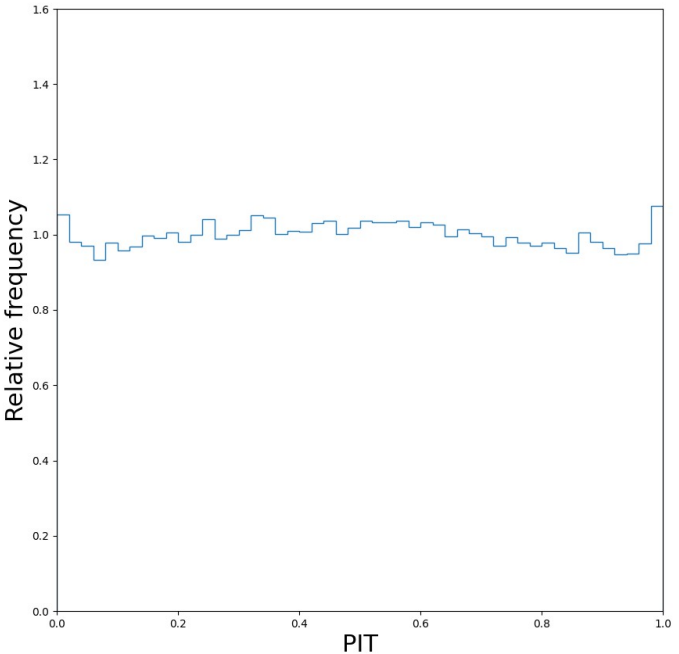
Comparison with a traditional approach of random forest
σ : -13.7 %
Bias : - 27.4 %
$\eta_{0.15}$: - 61.2 %



Photometric redshifts calibration

Example of PDFs produced after adaptation of the networks :





PIT distribution of the PDFs

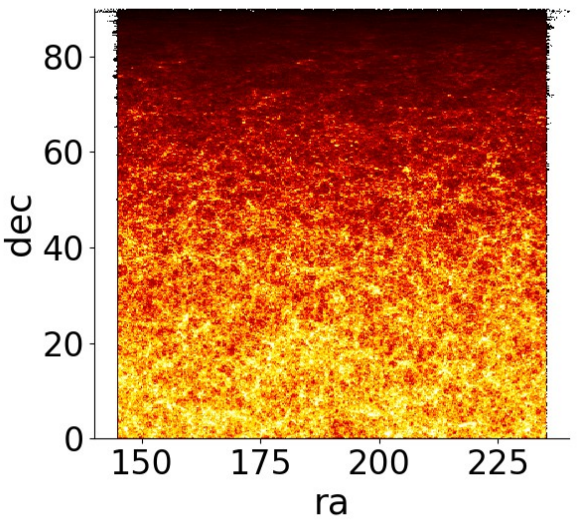
Flagship 2.1

- one octant of the sky, 145 < ra < 235 deg, 0 < dec < 90 deg

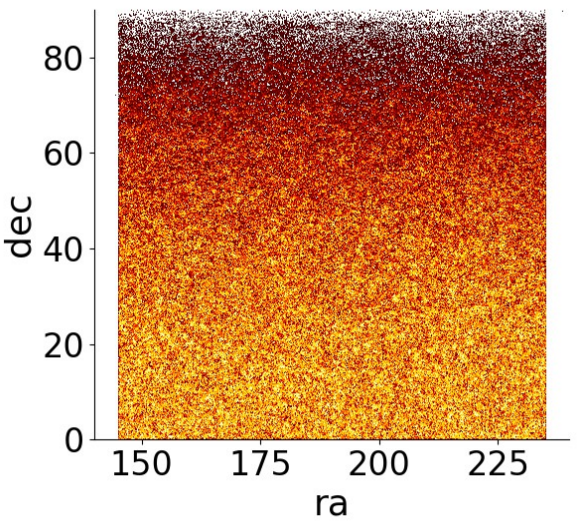
- 500 × 10^6 galaxies with VIS < 24.5 and photo-zs.

- fiducial cosmology : $\Omega_b = 0.049$
  $\Omega_c = 0.27$
  $h = 0.67$
  $A_s = 2.1 \times 10^9$
  $n_s = 0.96$

- 13 bins between 0.2 < z < 2.54



0.2 < z < 0.38

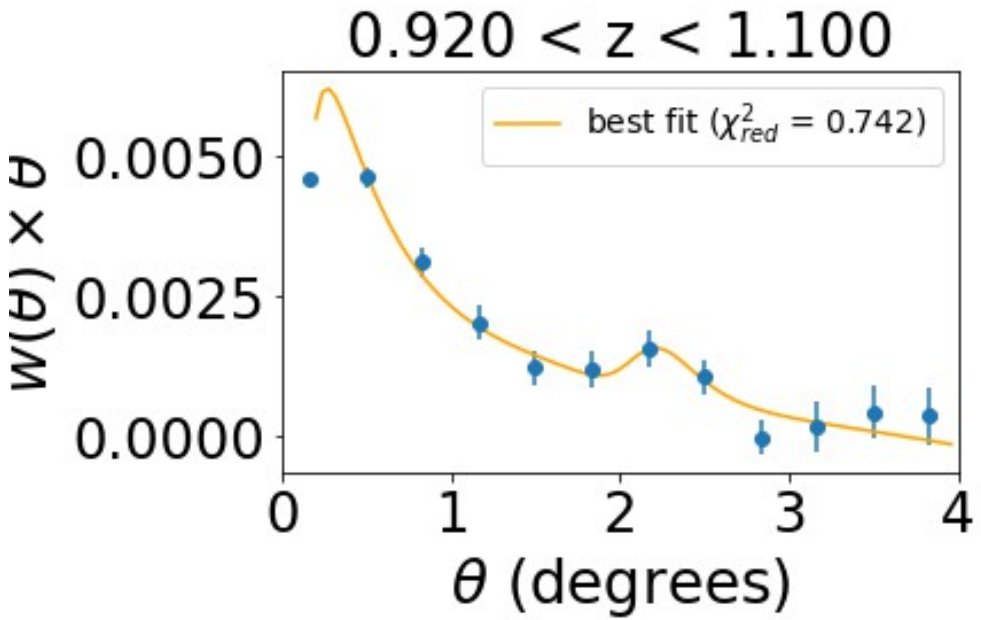


2.36 < z < 2.54

2pcf measurement

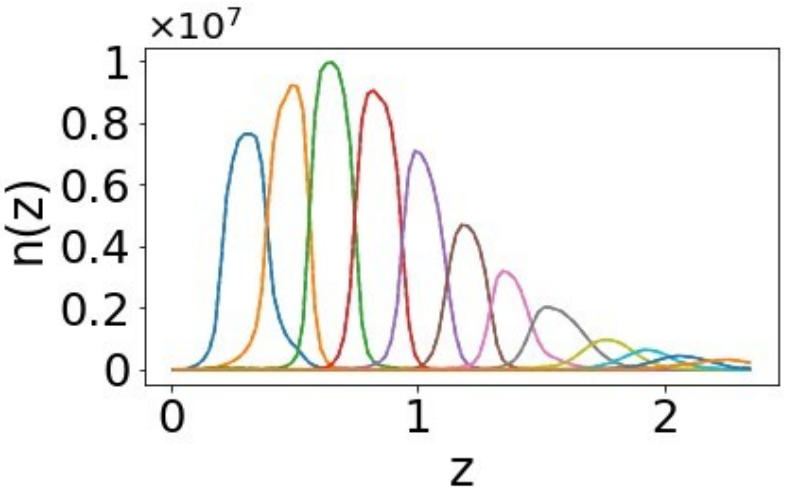
- Landy-Szalay $w(\theta) = \frac{DD-2DR+RR}{RR}$

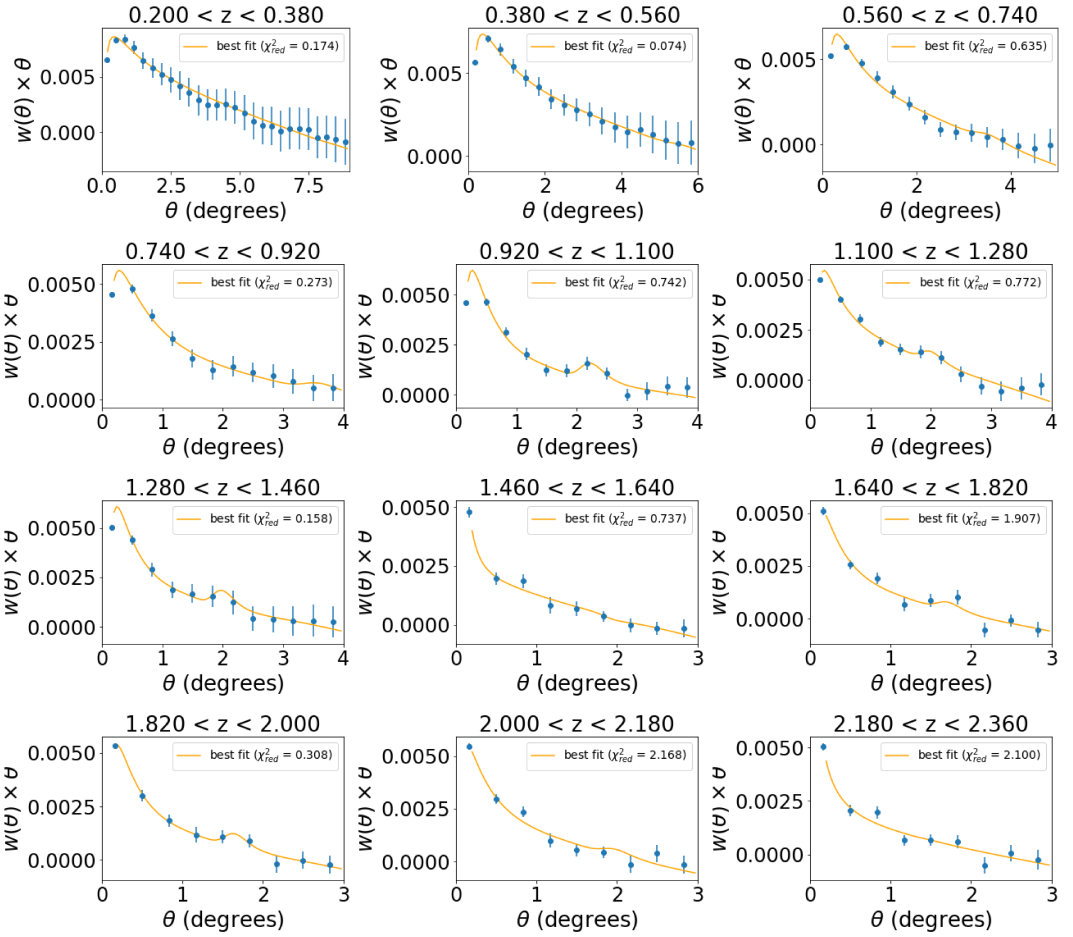
- Errors : jackknife



2pcf measurement

- n(z) from Euclid preparation XII Optimizing the photometric sample of the Euclid survey for galaxy clustering and galaxy-galaxy lensing analyses

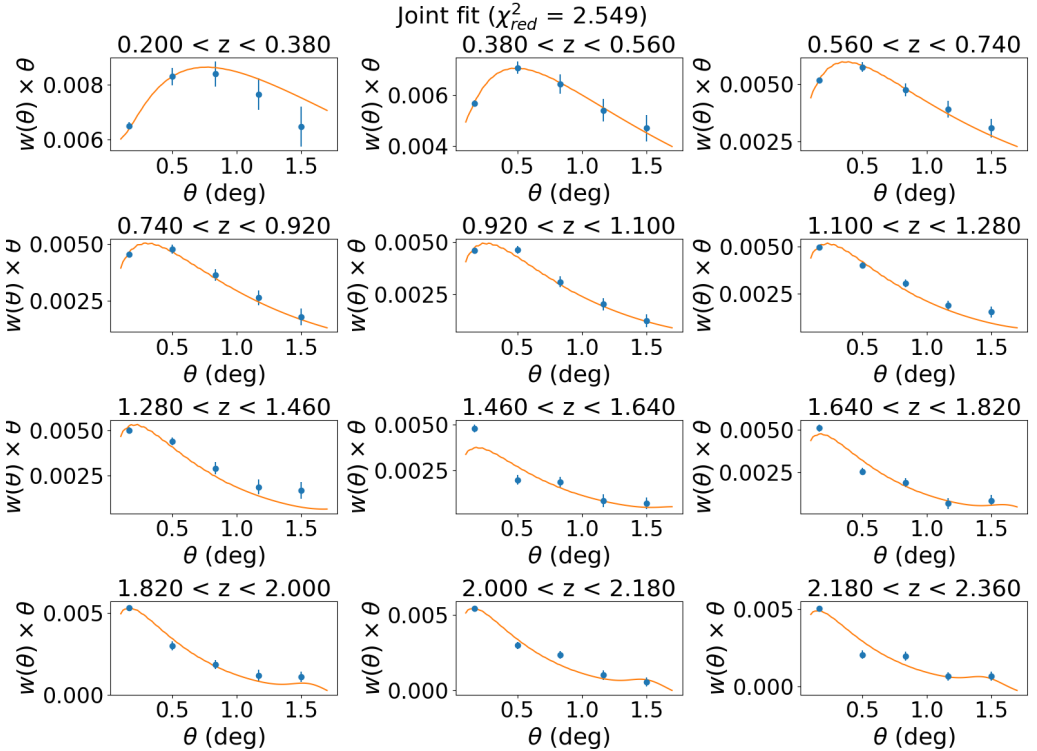




Full-shape analysis

- Full-shape : restriction to 0.48° < θ < 1.7 °
Euclid forecasts defined an optimistic and pessimistic scenarios for GCph with
$l_{max} = 3000 \rightarrow \theta_{min} = 0.12°$ or $l_{max} = 750 \rightarrow \theta_{min} = 0.48°$

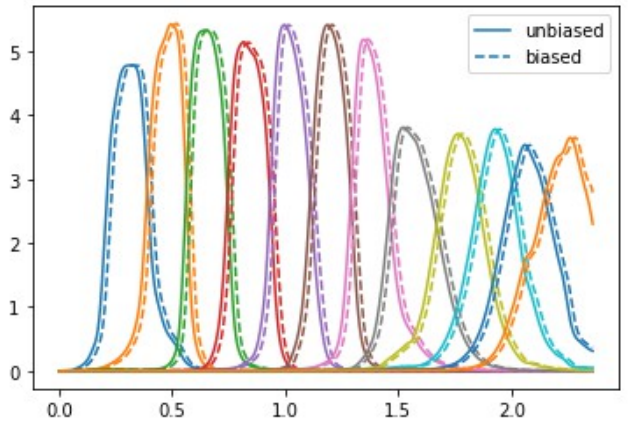
- Joint fit :



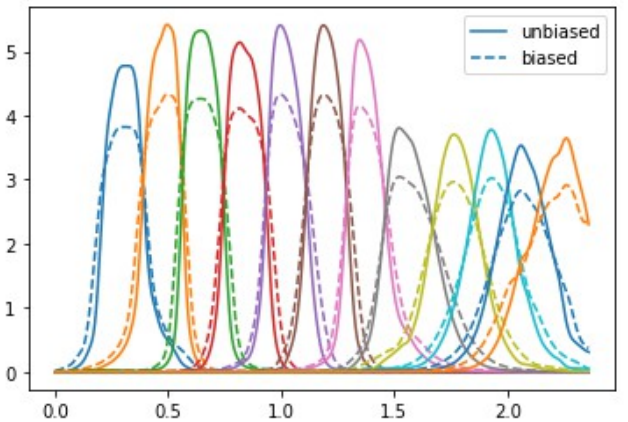
Full-shape analysis with modified n(z)

Goal of GCPHz WP paper 3 : study systematic uncertainties like n(z) model misspecifications

Modifications of n(z) :



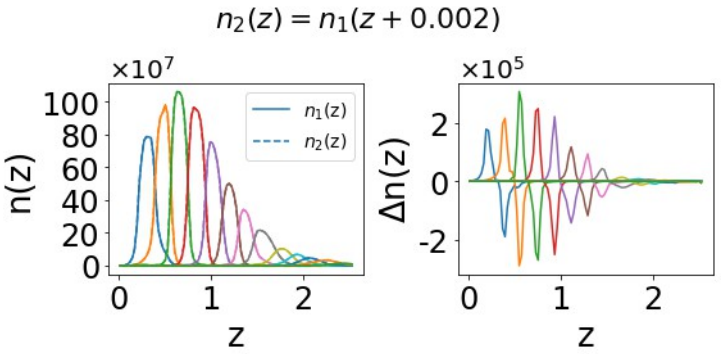
Additive bias

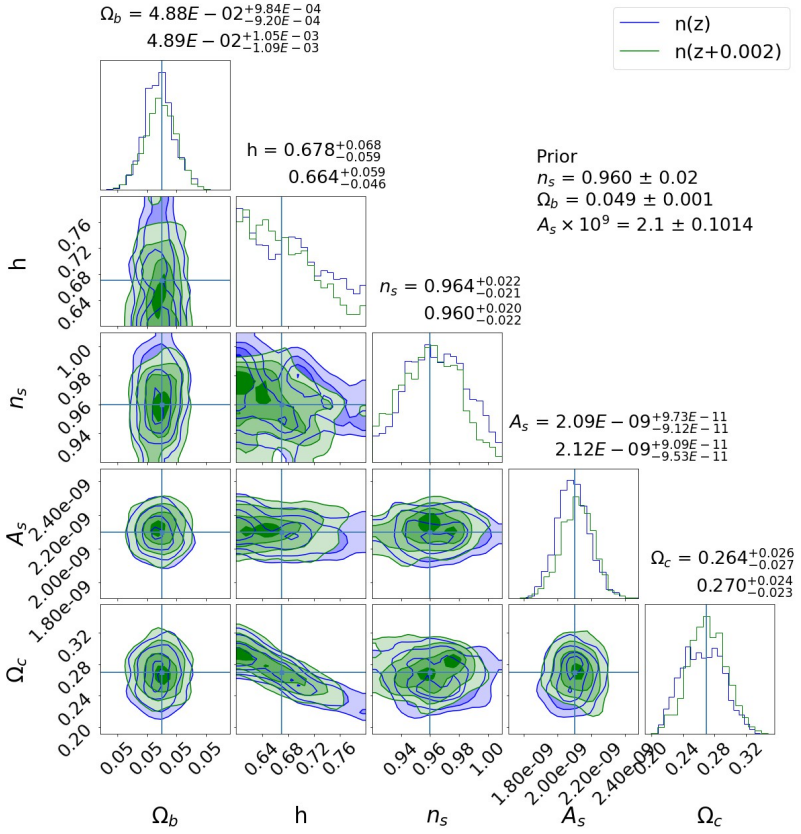


Broadening

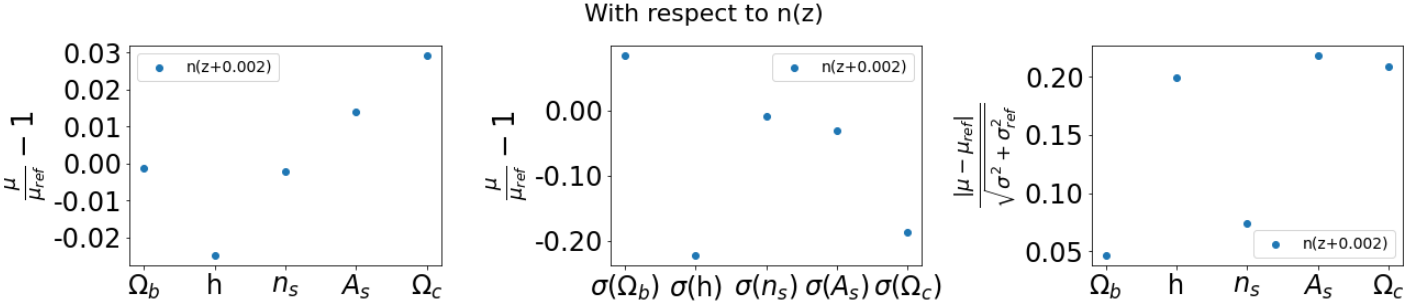
Full-shape analysis with modified n(z)

Bias of n(z) :





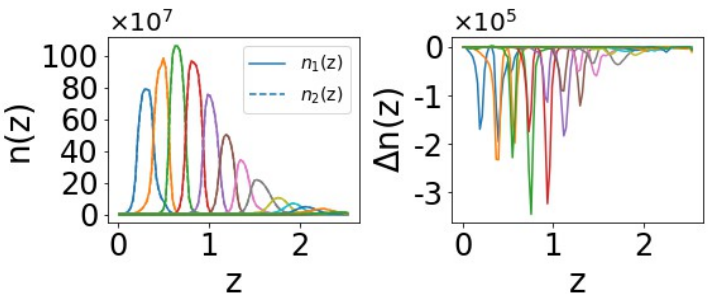
Shift of **0.2 σ** on h, A_s et Ω_c

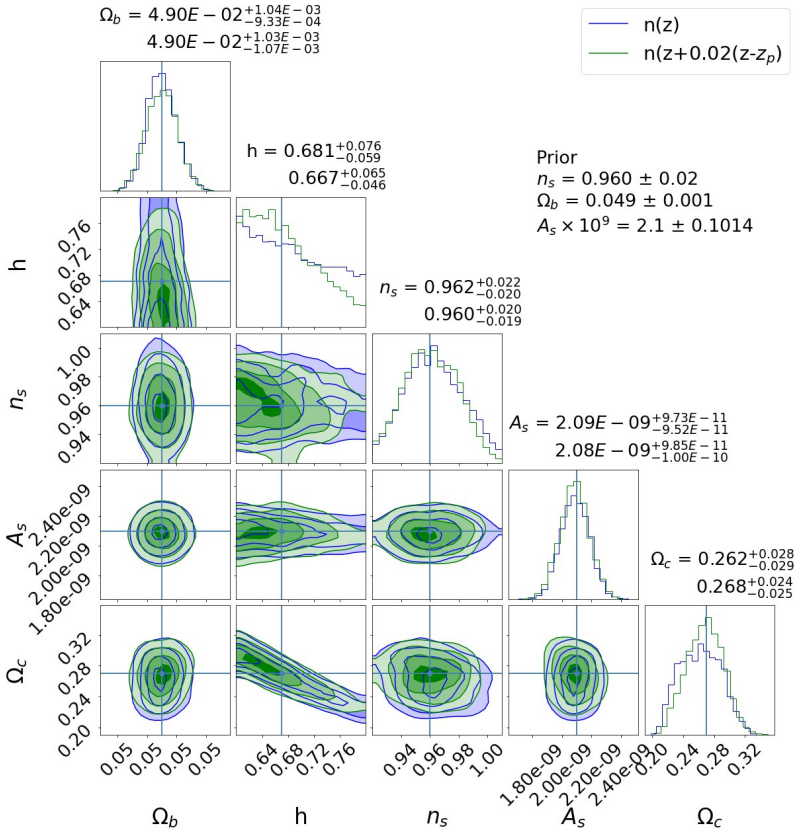


Full-shape analysis with modified n(z)

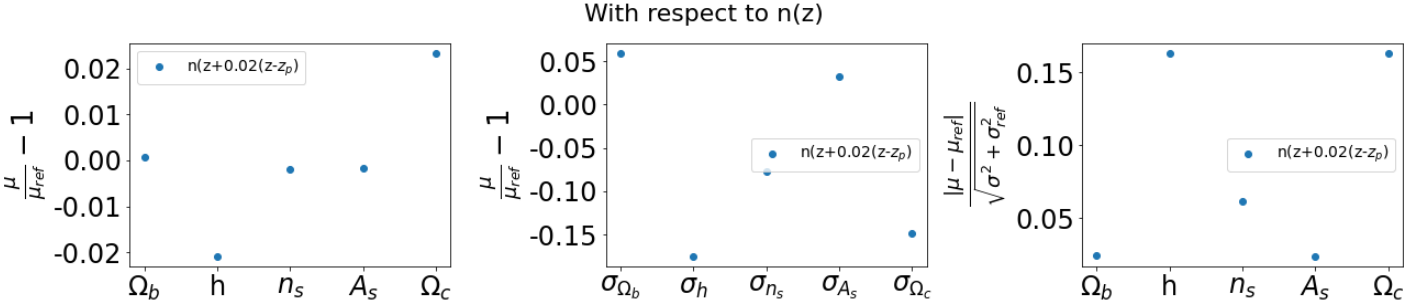
Broadening of n(z) :

$$n_2(z) = n_1(z + 0.02(z - z_p))$$





Shift of **0.15 σ** on h and Ω_c



Planned work for the full-shape analysis

- study of the effect of priors over the constraints.

- comparison of the optimistic and pessimistic scale cuts.

- analysis using the CPL dark energy parametrization with w_0, w_a.

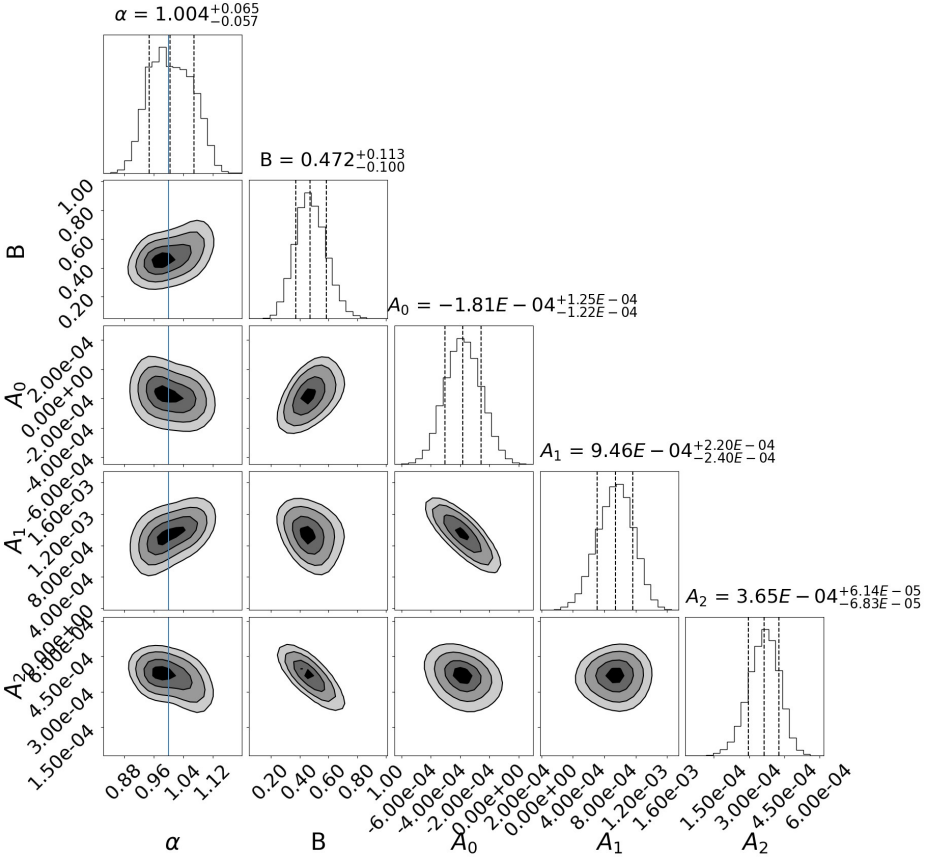
BAO analysis

- No restriction to small scales since we're interested in the BAO peak (≠ full-shape).

- Template : $B \times w(\alpha\theta) + A_0 + \frac{A_1}{\theta} + \frac{A_2}{\theta^2}$

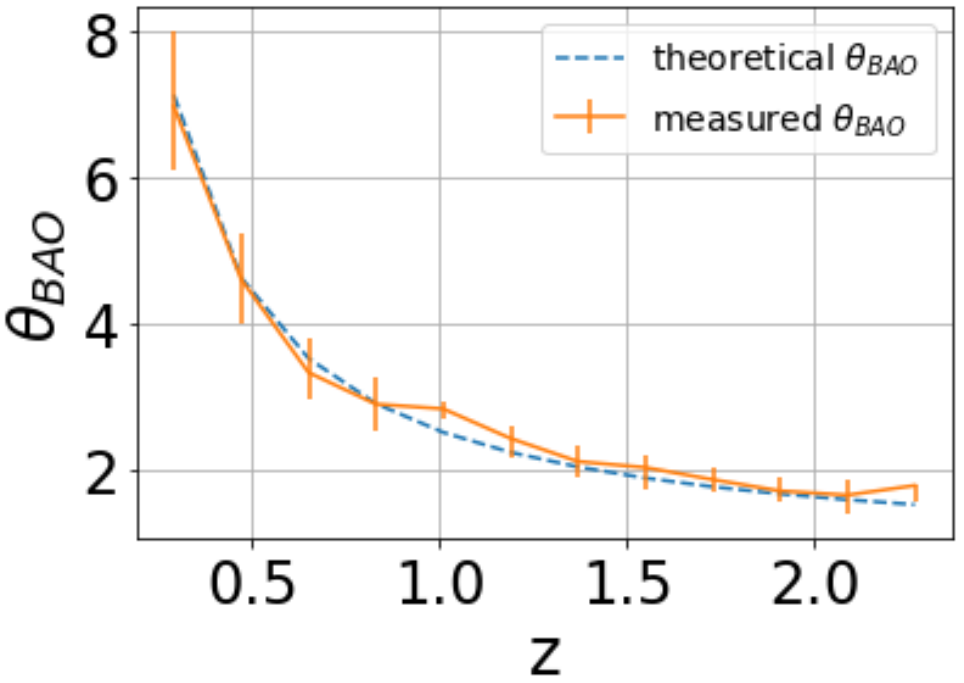
α quantifies an eventual shift of the BAO peak in the data with respect to the fiducial cosmology. Since the 2pcf is measured on Flagship, we expect α = 1.

B is a nuisance parameters accounting for corrections of the amplitude.



BAO analysis

BAO extracted from the 2pcf measured on Flagship, in each bin of redshift



θ_{BAO} and its error are obtained by MCMC with the previous template.

BAO analysis

- Exploration of different templates :

Templates 1-4 :

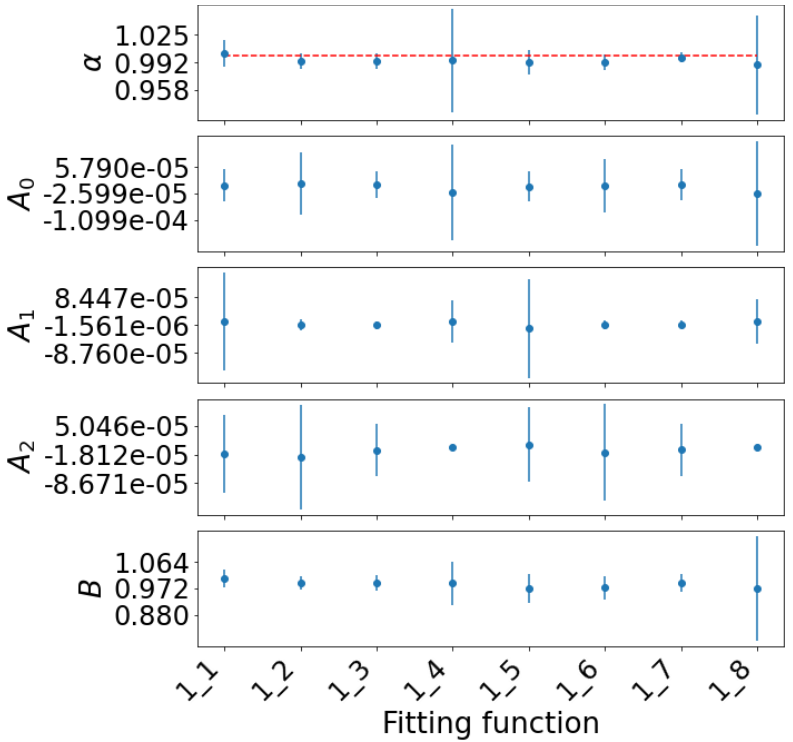
$$B \times w(\alpha\theta) + A_0 + \frac{A_1}{\theta} + \frac{A_2}{\theta^2}$$
$$B \times w(\alpha\theta) + A_0 + A_1\theta + \frac{A_2}{\theta}$$
$$B \times w(\alpha\theta) + A_0 + A_1\theta + \frac{A_2}{\theta^2}$$
$$B \times w(\alpha\theta) + A_0 + A_1\theta + A_2\theta^2$$

Templates 5-8 : $B \to \frac{B}{\alpha^2}$

The introduction of α^{-2} improved the constraints in previous analyses but this trend was not observed in the following results.

BAO analysis

- Validation test made by replacing the measured 2pcf by a theoretical one with gaussian noise of $\sigma = \sigma_{measured,Flagship}$. The mean over 100 realizations of this noise is in agreement with $\alpha = 1$.

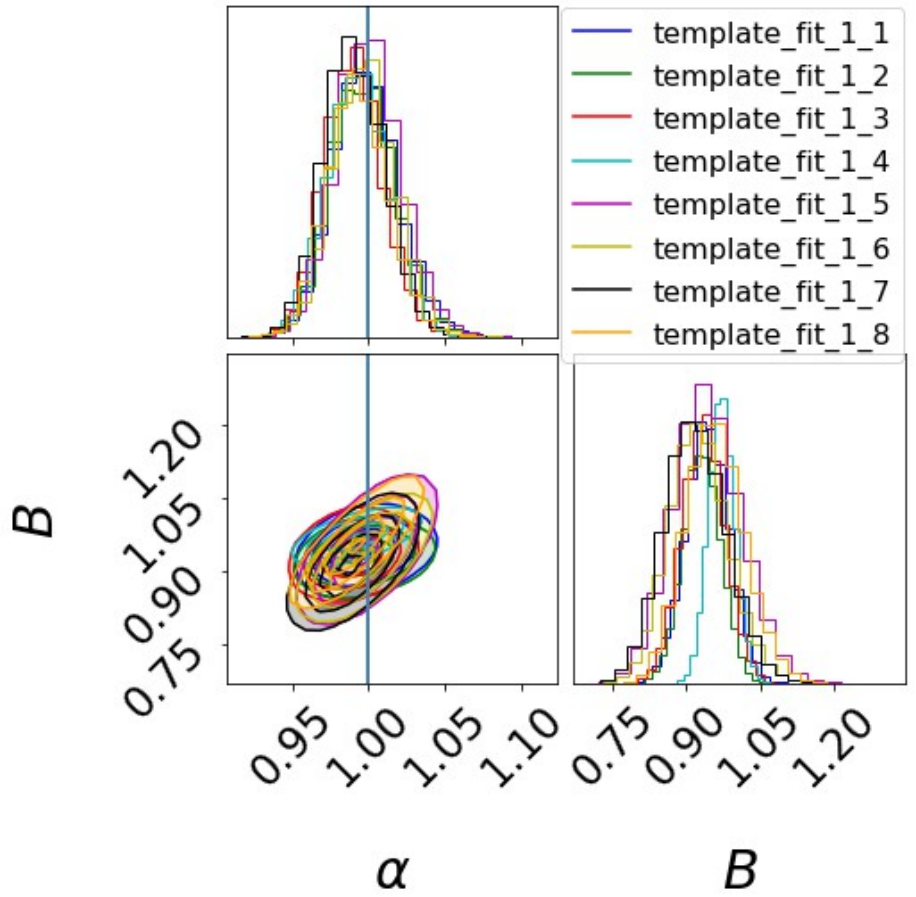


	1	2	3	4	5	6	7	8
α	1.003	0.993	0.993	0.994	0.992	0.992	0.998	0.988
±	0.016	0.009	0.009	0.063	0.015	0.010	0.006	0.061

Mean best fit α and its associated scatter

BAO analysis

- Exploration of different templates with the measured 2pcf, joint MCMC :

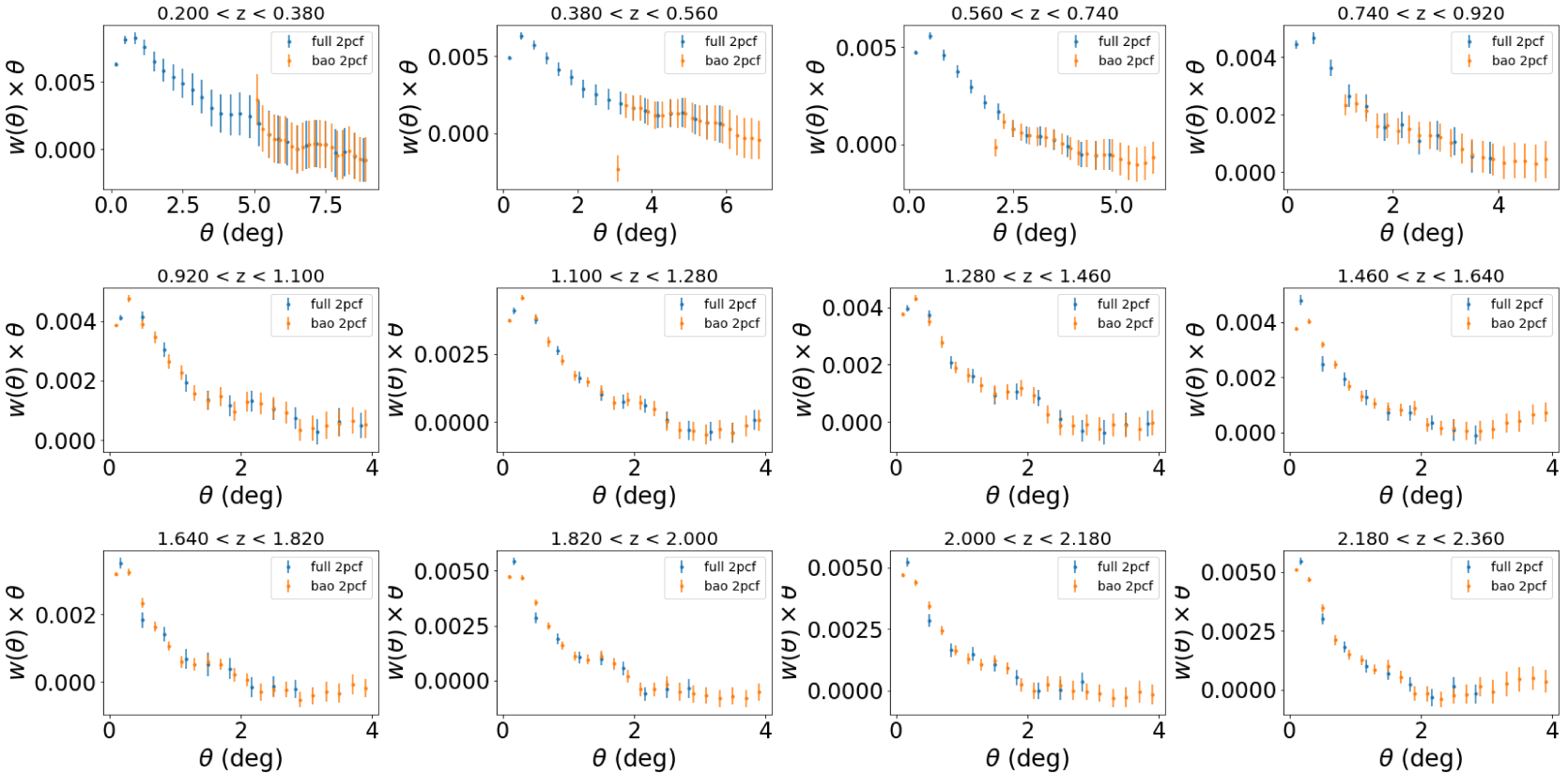


	α	B
template_1_1	$0.997^{+0.022}_{-0.020}$	$0.941^{+0.044}_{-0.044}$
template_1_2	$0.996^{+0.021}_{-0.020}$	$0.928^{+0.038}_{-0.038}$
template_1_3	$0.989^{+0.019}_{-0.018}$	$0.933^{+0.043}_{-0.043}$
template_1_4	$0.992^{+0.020}_{-0.019}$	$0.970^{+0.027}_{-0.027}$
template_1_5	$0.999^{+0.021}_{-0.020}$	$0.941^{+0.073}_{-0.066}$
template_1_6	$0.997^{+0.021}_{-0.019}$	$0.923^{+0.062}_{-0.058}$
template_1_7	$0.989^{+0.020}_{-0.019}$	$0.916^{+0.066}_{-0.061}$
template_1_8	$0.991^{+0.020}_{-0.019}$	$0.955^{+0.062}_{-0.058}$

α and B are in agreement for all templates

BAO analysis

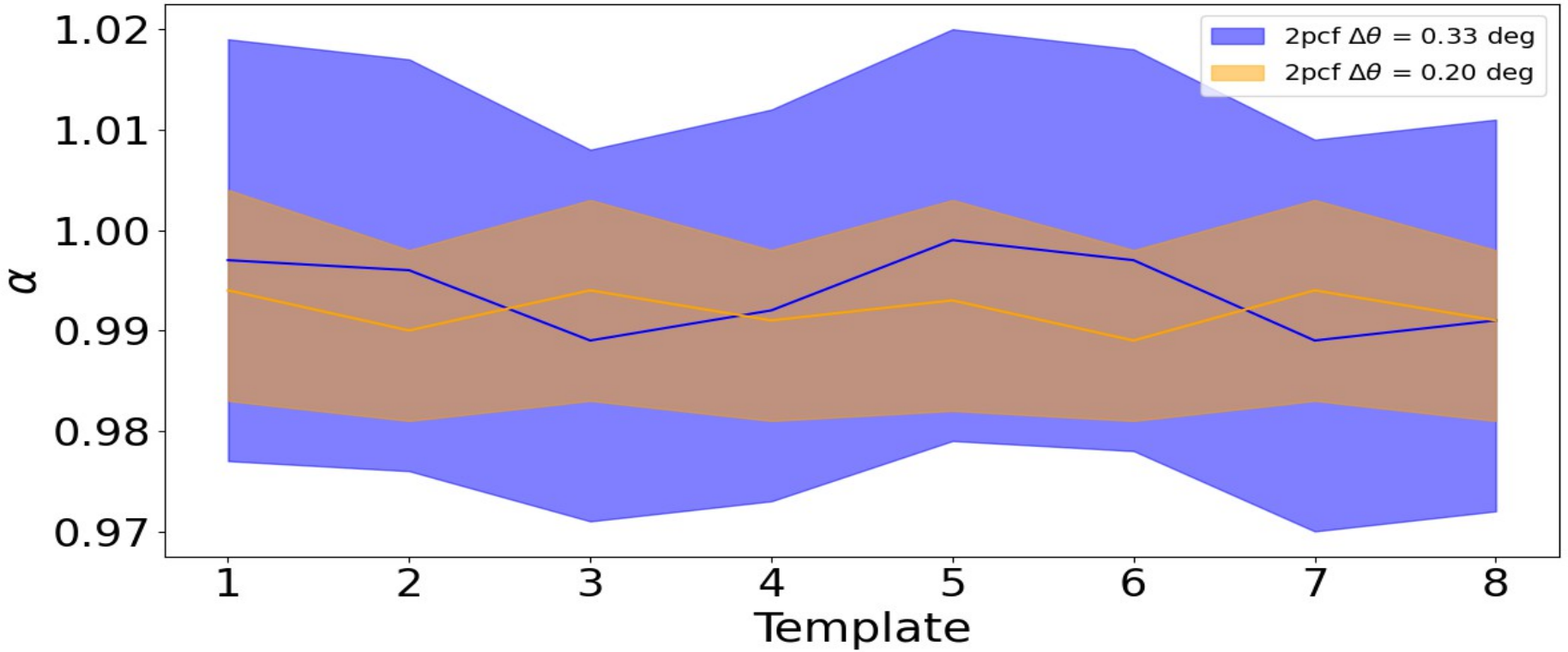
New measurement of the 2pcf with an improved resolution of 0.2 ° :



Range : $\theta_{BAO,th}$ ± 2° if $\theta_{BAO,th}$ > 2°, else [0°,4°]

BAO analysis

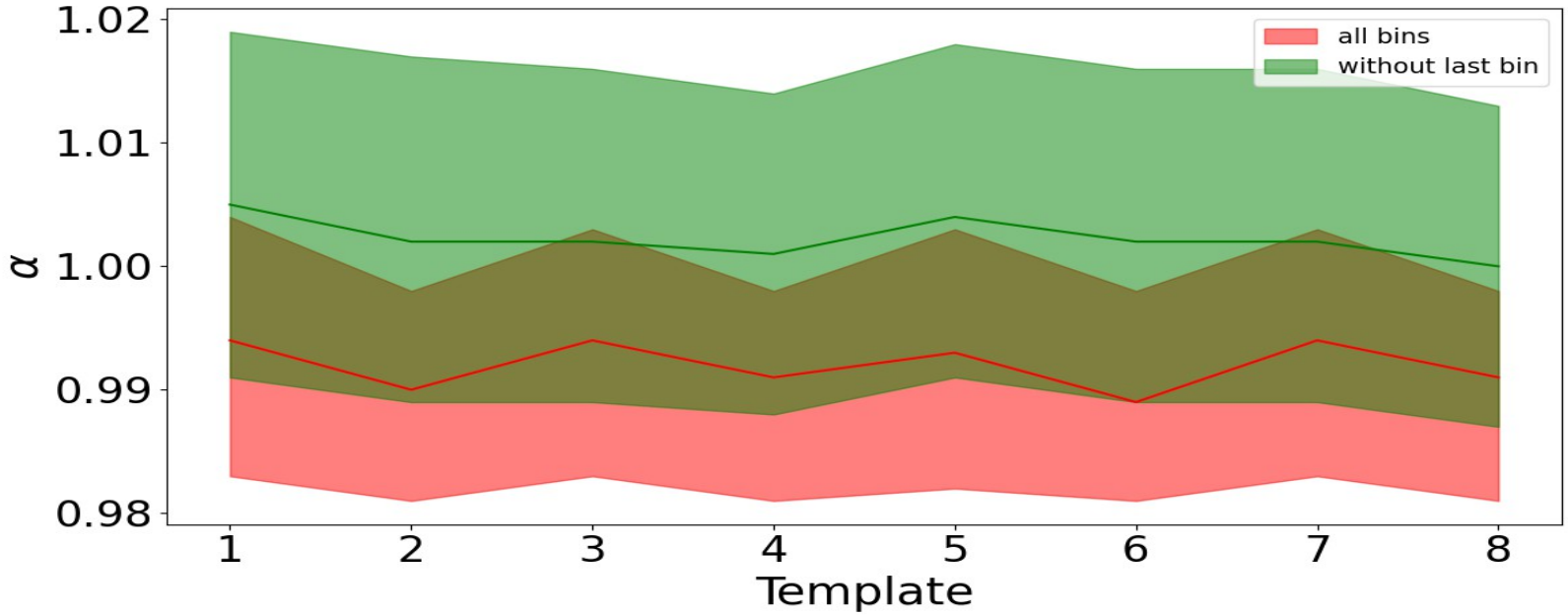
MCMC with the previous measurement (left) and the new one (right) :



The error on α is divided by 2 with the new measurement.

BAO analysis

Comparison including or excluding the last redshift bin :



In agreement at 1σ but there is an obvious systematic shift towards larger α and errors.
The robustness of the results with respect to the redshift bins used should be checked.

Planned work for the BAO analysis

- robustness validation with respect to the redshift bins

- study of the scale cuts influence

- study of the impact of RSD

- study of the influence of the Limber approximation

Thank you for your attention !

Questions ?

Back-up

Photometric redshifts

Loss function used to train : mean squared error

Metrics :

- Standard deviation of residuals σ = std(Δz) wtih $\Delta z = z_{phot} - z_{spec}$

- Bias : mean($|\Delta z| / (1 + z_{spec})$)

- Outlier fraction at 15 % : #(bias > 0.15) / #(test set)
+ fractions at 10 % and 5 %

- σ_{NMAD} = 1.4826 $\times$ median($|\Delta z|$ - median(Δz))

- σ_{MAD} = 1.48 $\times$ median($|\Delta z|$)

Photometric redshifts

Side plots :

Learning error $\xi = p(z_{phot} - z_{spec} \mid z_{spec})$
→ in each bin of the histogram, I compute the mean and standard deviation of the $z_{predicted,i} - z_{bin}$ for all $z_{spec,i}$ falling into that bin

Prediction uncertainty $\mu = p(z_{phot} - z_{spec} \mid z_{phot})$
→ in each bin of the histogram, I compute the mean and standard deviation of the $z_{spec,i} - z_{bin}$ for all $z_{predicted,i}$ falling into that bin

Additional statistics on ξ and μ :

Avg % error	Min % error	Max % error	Avg % error without high z bins	Min % error without	Max % error without
16.05	5.26	90.56	19.23	10.97	90.56

Photometric redshifts

Characterization of the PDFs :

Probability Integral Transform (PIT), for a galaxy i of redshift $z_{spec} = z_i$

$$CDF_i(z_i) = \int_0^{z_i} PDF_i(z)\,dz$$

If PDFs are often too narrow then the z_{spec} will more often be under/overestimated and the PIT value will be close to 0 or 1.
If they are too wide then z_{spec} will often be in the PDF, which favors intermediate PIT values

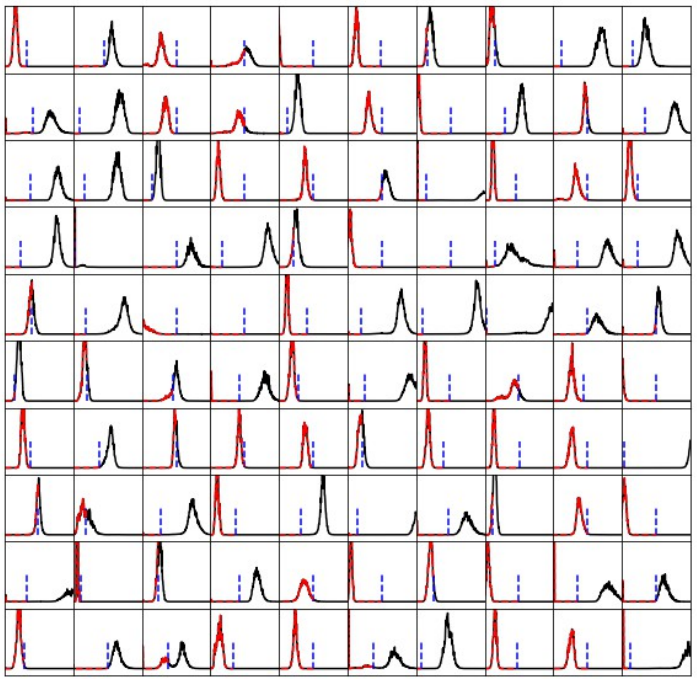
→ study of the PITs distribution :
- if PDFs have inadequate shapes then the distribution will either be concave or convex.
- if there is a bias between the predicted redshifts and z_{spec} then it creates a slope

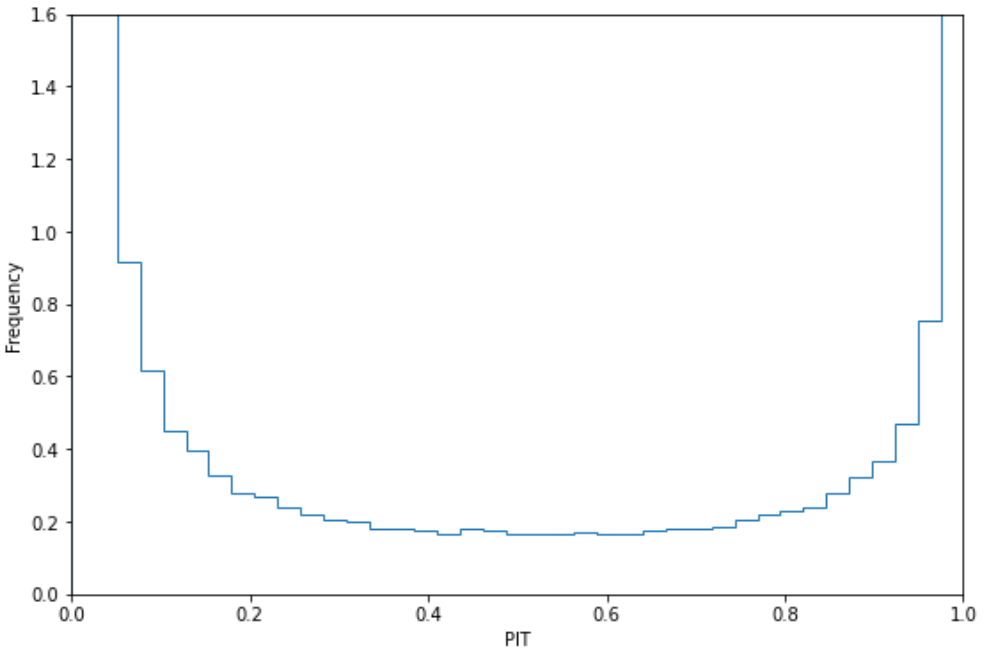
→ an ideal PIT distribution is horizontal and has no curvature.

Photometric redshifts

Example of a bad PIT distribution :

Many PDFs miss z_{spec} → The PIT distribution is convex

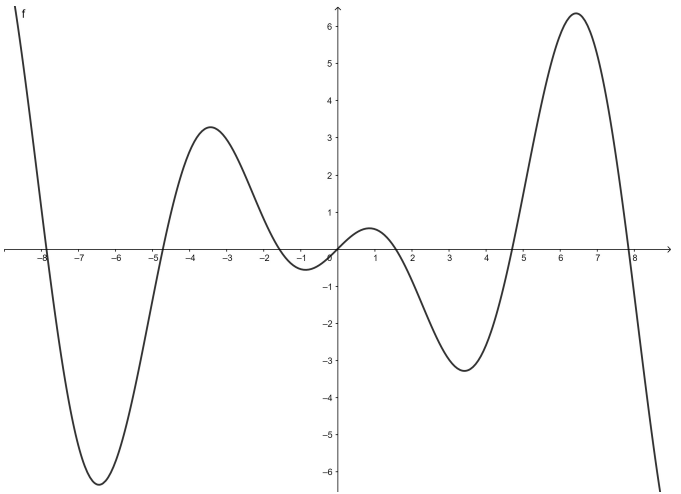




Vanishing gradients

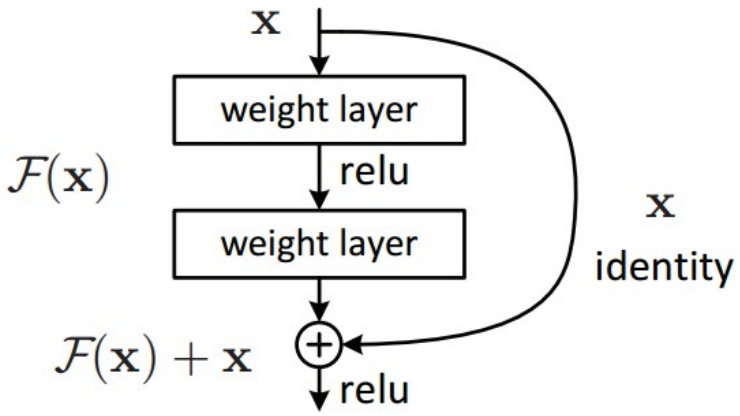
The update of weights is proportional to the gradient of the loss function with respect to current weights. In the backpropagation, the chain rule for partial derivatives is used, which implies that we can end up multiplying very small gradients in chain. This entails the death of some neurons because their weights no longer change.

As for exploding gradients, Rectified activation functions like ReLu limit this issue because they can only saturate by negative values but the issue can still appear. Some oscillating functions can be used to counter this problem like the Growing Cosine Unit



https://commons.wikimedia.org/wiki/File:Growing_Cosine_Unit_(GCU)_activation_function.png

Residual blocks



https://arxiv.org/abs/1512.03385

The layer n give its output to layer n+1 and layer n+5 (in ResNet34) or n+3,… depending on the architecture

Benefit : when the number of layers is increased in a neural network, results improve before reaching a maximum and then degrade (vanishing gradients).

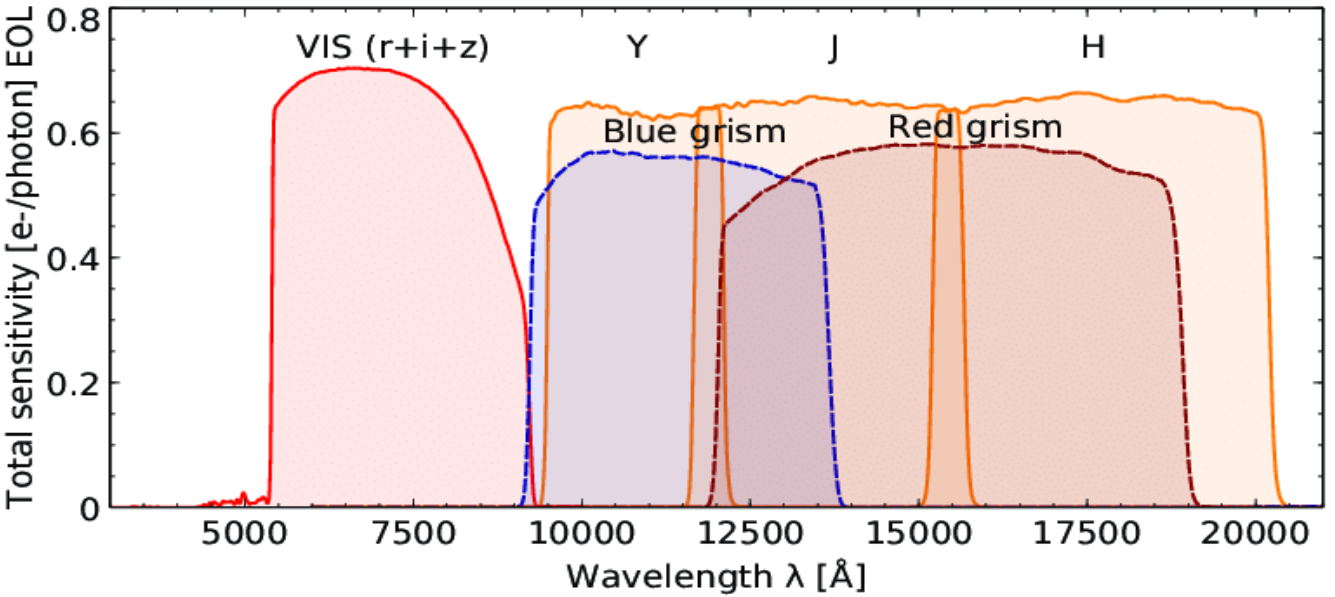
Idea :
residual = output – input ↔ output = residual + input

This enables the identity operation when the residual is fixed to 0. This is useful since the identity can't be the output of a neural network if there is no skip connection (non linear activation functions) → the least useful layers have weights close to 0 but won't make gradients vanish because the skip connection will have larger weights.

Photometric redshifts

Euclid bands :

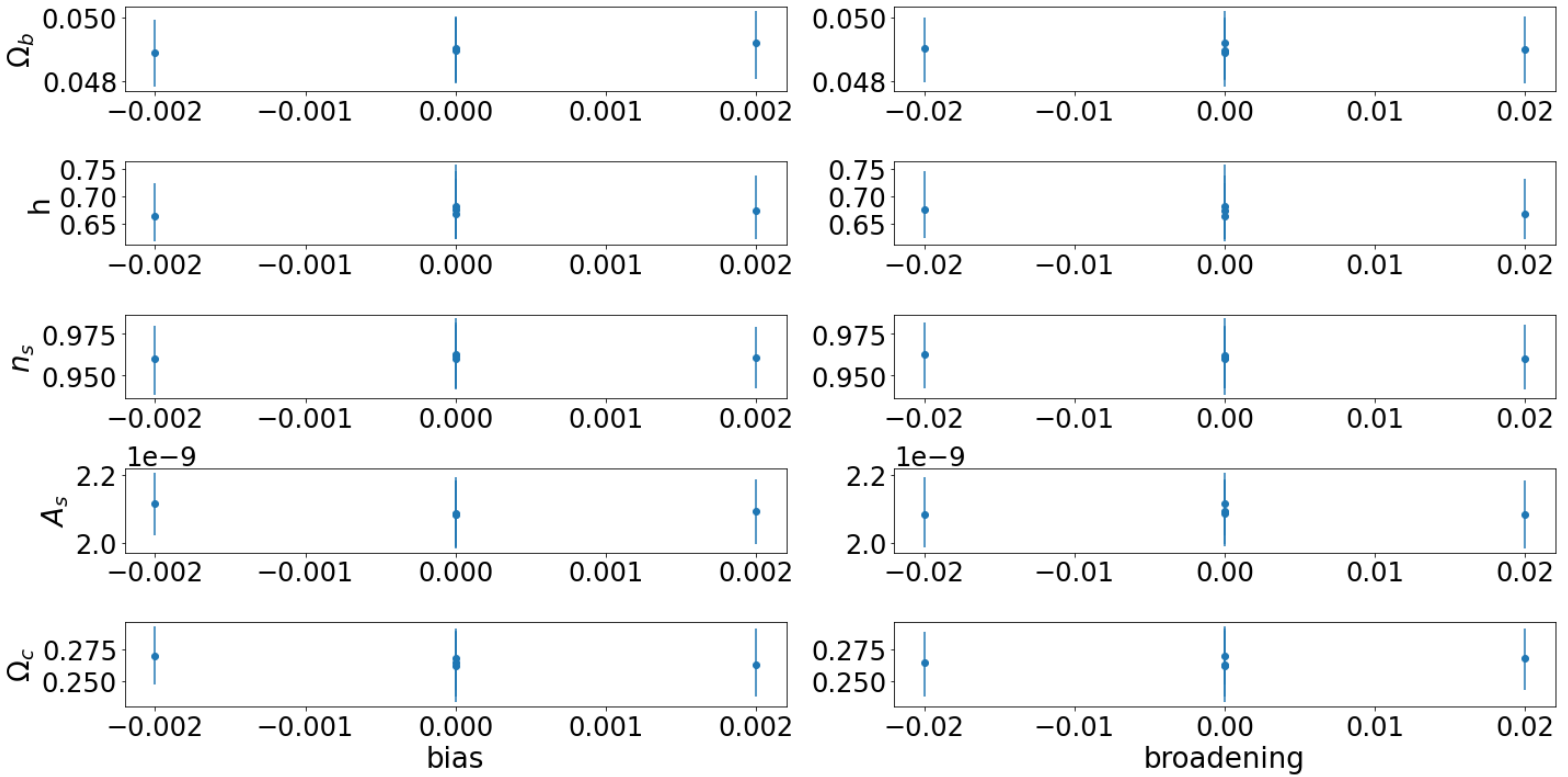
VIS 550-900 nm
Y 920-1146 nm
J 1146-1372 nm
H 1372-2000 nm



Euclid preparation: I. The Euclid Wide Survey (arXiv:2108.01201)

Full-shape analysis with modified n(z)

Influence of n(z) model misspecifications



Angular power spectra

Core Cosmology Library : precision cosmological predictions for LSST (arXiv:1812.05995)

$$\langle a_{\ell m} b^*_{\ell m} \rangle \equiv C^{ab} \delta_{\ell \ell'} \delta_{m m'}$$

$$C_\ell^{ab} = 4\pi \int_0^\infty \frac{dk}{k} \mathcal{P}_\Phi(k) \Delta_\ell^a(k) \Delta_\ell^b(k)$$

$$\Delta_\ell^{\mathrm{D}}(k) = \int dz \, p_z(z) \, b(z) \, T_\delta(k, z) \, j_\ell(k\chi(z))$$

p(z) : normalized distribution of sources in redshift

Limber approximation : small angles (large l)

$$j_\ell(x) \simeq \sqrt{\frac{\pi}{2\ell + 1}} \, \delta\left(\ell + \frac{1}{2} - x\right)$$

Takahashi Halofit model :

Fitting formula for the dimensionless non-linear power spectrum (Takahashi et al. 2018) : using the notation $\Delta^2(k) = k^3P(k)/(2\pi^2)$, Q denoting the two-halo term, H the one-halo term and L the linear power spectrum :

$\Delta^2(k) = \Delta_Q^2(k) + \Delta_H^2(k)$ (two-halo and one-halo terms)

$$\Delta^2_{\rm Q}(k) = \Delta^2_{\rm L}(k)\left[\frac{\{1+\Delta^2_{\rm L}(k)\}^{\beta_{\rm n}}}{1+\alpha_{\rm n}\Delta^2_{\rm L}(k)}\right]e^{-f(y)}$$ with $f(y) = y/4 + y^2/8$

$$\Delta^2_{\rm H}(k) = \frac{\Delta'^2_{\rm H}(k)}{1+\mu_{\rm n}y^{-1}+\nu_{\rm n}y^{-2}} \quad \text{with} \quad \Delta'^2_{\rm H}(k) = \frac{a_{\rm n}y^{3f_1(\Omega_{\rm m})}}{1+b_{\rm n}y^{f_2(\Omega_{\rm m})}+[c_{\rm n}f_3(\Omega_{\rm m})y]^{3-\gamma_{\rm n}}}$$

with $y = k / k_\sigma$

k_σ is defined so that $\sigma^2(k_\sigma^{-1}) = 1$ with $\sigma^2(R) = \int d\ln k\ \Delta^2_{\rm L}(k)e^{-k^2R^2}$

Takahashi Halofit model :

Defining $n_{\rm eff} + 3 = -\left.\frac{d\ln\sigma^2(R)}{d\ln R}\right|_{\sigma=1}$, $C = -\left.\frac{d^2\ln\sigma^2(R)}{d\ln R^2}\right|_{\sigma=1}$

$$f_1(\Omega_{\rm m}) = \Omega_{\rm m}^{-0.0307},\ f_2(\Omega_{\rm m}) = \Omega_{\rm m}^{-0.0585},\ f_3(\Omega_{\rm m}) = \Omega_{\rm m}^{0.0743}$$

the best fit parameters of the Takahashi Halofit model are then :

$$\log_{10} a_{\rm n} = 1.5222 + 2.8553 n_{\rm eff} + 2.3706 n_{\rm eff}^2 + 0.9903 n_{\rm eff}^3 + 0.2250 n_{\rm eff}^4 - 0.6038 C + 0.1749 \Omega_{\rm w}(z)\,(1+w)\,,$$

$$\log_{10} b_{\rm n} = -0.5642 + 0.5864 n_{\rm eff} + 0.5716 n_{\rm eff}^2 - 1.5474 C + 0.2279 \Omega_{\rm w}(z)\,(1+w)\,,$$

$$\log_{10} c_{\rm n} = 0.3698 + 2.0404 n_{\rm eff} + 0.8161 n_{\rm eff}^2 + 0.5869 C,$$

$$\gamma_{\rm n} = 0.1971 - 0.0843 n_{\rm eff} + 0.8460 C,$$

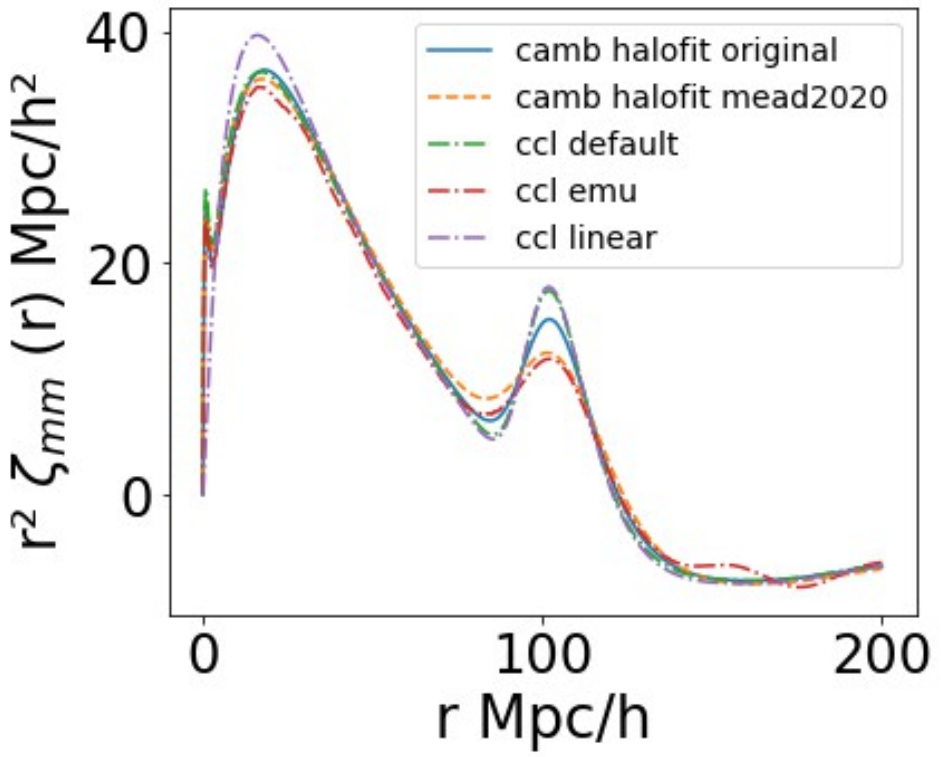
$$\alpha_{\rm n} = \left|6.0835 + 1.3373 n_{\rm eff} - 0.1959 n_{\rm eff}^2 - 5.5274 C\right|,$$

$$\beta_{\rm n} = 2.0379 - 0.7354 n_{\rm eff} + 0.3157 n_{\rm eff}^2 + 1.2490 n_{\rm eff}^3 + 0.3980 n_{\rm eff}^4 - 0.1682 C,$$

$$\mu_{\rm n} = 0,$$

$$\log_{10} \nu_{\rm n} = 5.2105 + 3.6902 n_{\rm eff},$$

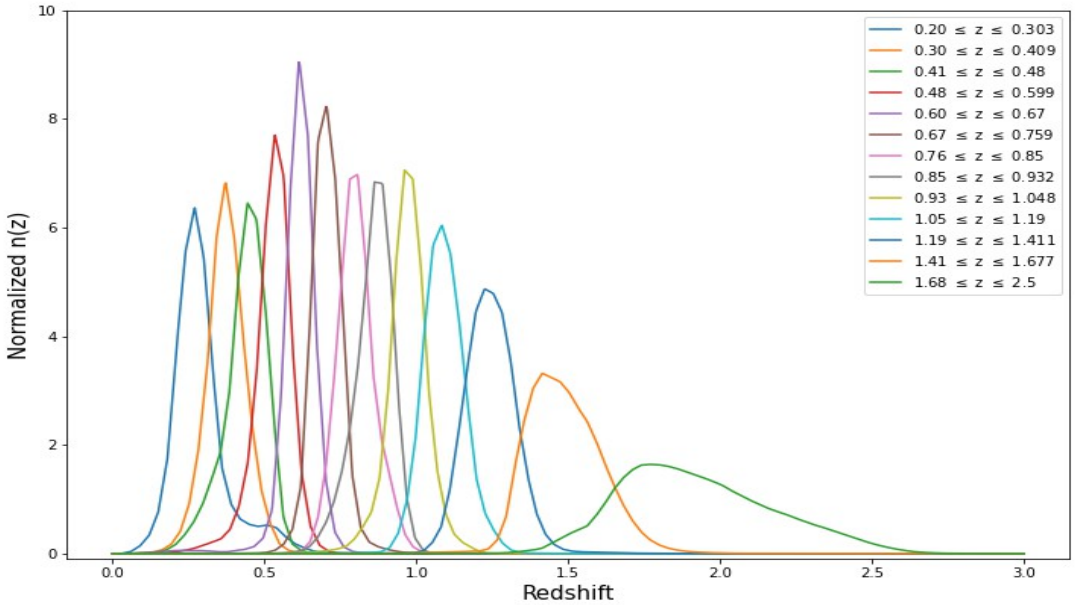
Model under-damping :



→ Use of Mead2020 (arXiv:2009.01858)

Flagship 2.1

Equipopulated bins n(z) :



Measured galaxy bias: