Aix Marseille université
CENTRE DE PHYSIQUE DES PARTICULES DE MARSEILLE
CPPM

Thèse de doctorat

Optimal inference of cosmological parameters in preparation of the Euclid survey

Sylvain Gouyou Beauchamps

Soutenue le 15 décembre 2021 devant le jury composé de

Stéphane Plaszczynski	IJCLab, Orsay	Rapporteur
Martin Crocce	ICE-CSIC/IEEC, Barcelone	Rapporteur
Christian Marinoni	CPT, Marseille	Examinateur
Julien Lesgourgues	RWTH, Aachen	Examinateur
Sandrine Codis	LCEG/AIM-CEA, Saclay	Examinatrice
Benjamin Joachimi	UCL, Londres	Examinateur
Stéphanie Escoffier	CPPM, Marseille	Directrice de thèse
William Gillard	CPPM, Marseille	Co-directeur de thèse

Discipline
Physique et Science de la matière

Spécialité
Astrophysique et Cosmologie

École doctorale
ED 352

Outline

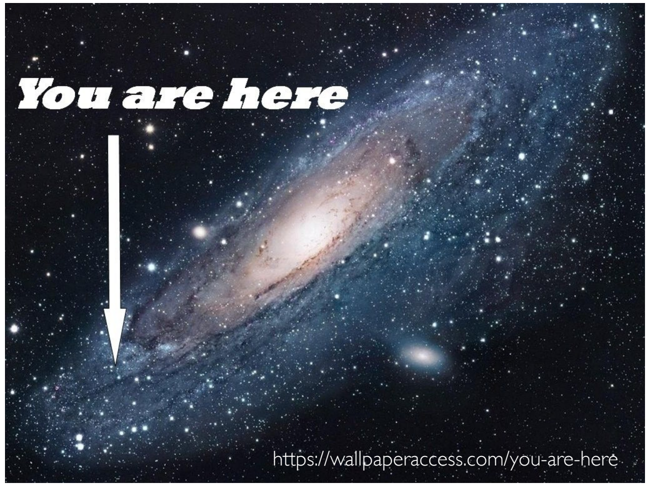
Introduction and key concepts

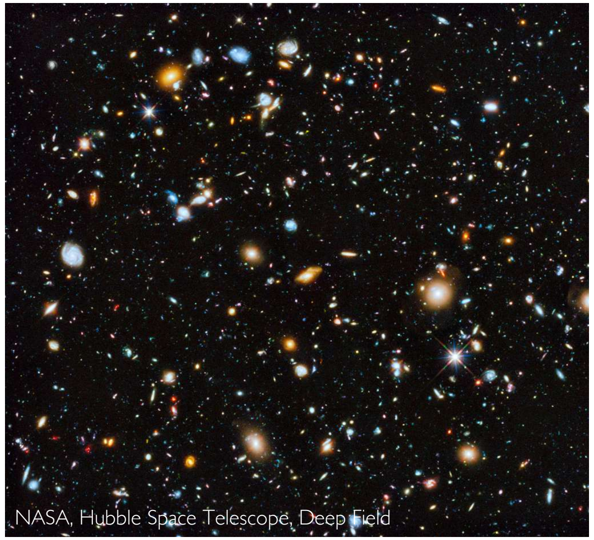
My work

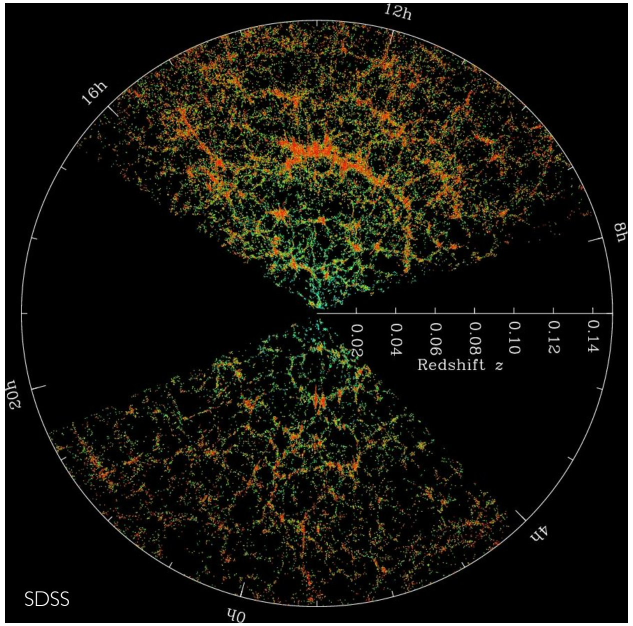
I. State of the art in observational cosmology

- The standard model of cosmology
- The Large Scale Structure and the Euclid survey

The aim of cosmology is to give a scientific description of the content and the evolution of the Universe







Cosmology is a peculiar science as it aims the description of the Universe as a whole

On the theory side

Newtonian gravity fails at predicting the behaviour of the Universe

On the observational side

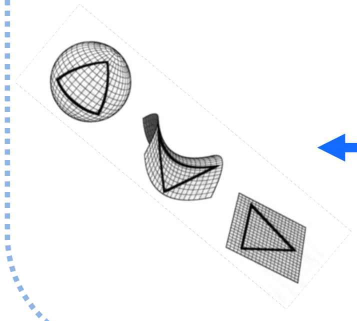
We only have one realisation of the Universe

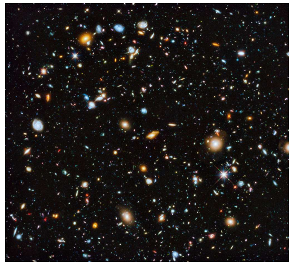
Theoretical framework

General relativity

$$R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} = 8\pi G T_{\mu\nu}$$

Geometry — Energy content

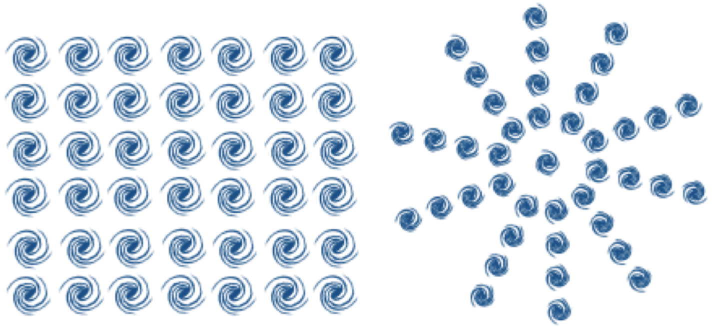




+

The cosmological principle

Homogeneity



Isotropy

Taken from the PhD thesis of Alice Pisani, 2014

$$g_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu} = \mathrm{d}t^2 - a(t)^2\left[\mathrm{d}\chi^2 + S_k^2(\chi)\mathrm{d}\Omega^2\right]$$

Scale factor — Curvature

$$H \equiv \frac{\dot{a}}{a}$$

The expansion rate of the Universe is governed by its content and curvature

$$H(a)^2 = H_0^2 \sum_i \left[\Omega_{i,0} a^{-3(1+w_i)} + \Omega_{k,0} a^{-2}\right]$$

Reduced density $\Omega_i \equiv \rho_i \frac{8\pi G}{3H^2}$ | Equation of state $w_i = p_i/\rho_i$

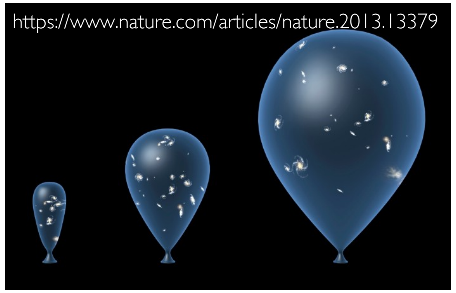
$\dot{a} = 0$ Static ?

$\dot{a} > 0$ Expanding ?

$\dot{a} < 0$ Contracting ?

For a universe filled with matter ($\Omega_m = 1$)

$\dot{a} > 0$ The universe is expanding



BUT

Einstein introduced the cosmological constant Λ to keep it static :

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} - \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}$$

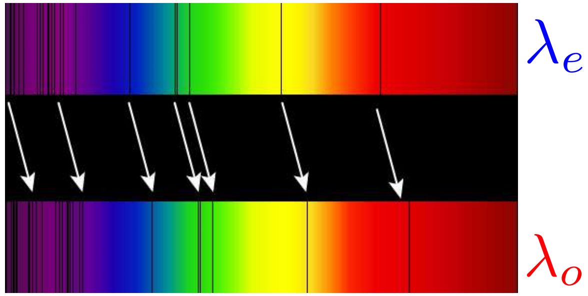
Theory

↕

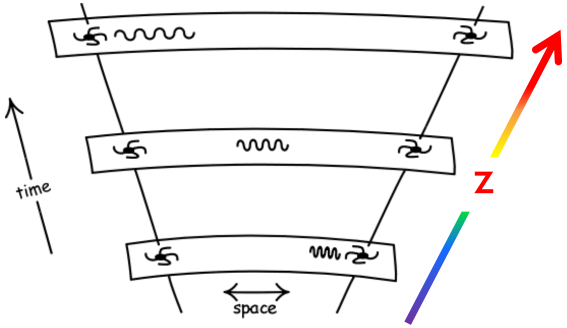
Observation

Observation of the redshift of galaxies (G. Lemaître 1927 and E. Hubble 1929)

https://en.wikipedia.org/wiki/Redshift



$$1 + z = \frac{\lambda_o}{\lambda_e}$$



The expansion stretches photon's wavelength

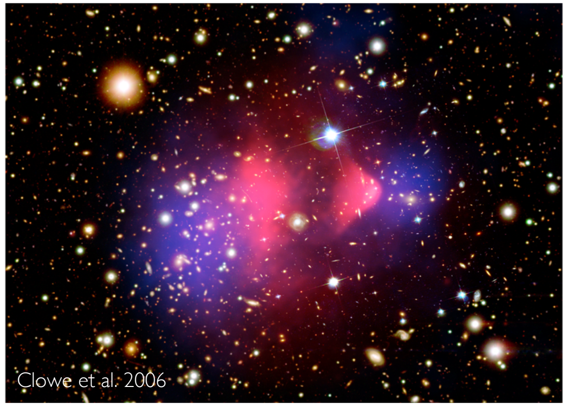
$$1 + z = \frac{a_0}{a}$$

- Observational proof of the expansion (Supression of Λ)
- Redshift is an indicator of time and distance

H_0 and $\Omega_{i,0}$ refer to their values today

The understanding of the Universe was built from a combination of observations

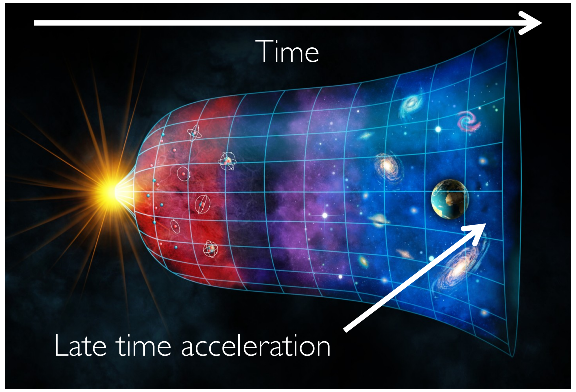
➢ The universe is filled with Cold Dark Matter (CDM)



- ➢ Velocity dispersion in Clusters (1933)
- ➢ Big Bang Nucleosynthesis (1940)
- ➢ Cosmic Microwave Background (1965)
- ➢ Galaxy rotation curves (1970)
- ➢ Lensing in clusters (2004)

→ 5% of ordinary matter and 27% of CDM

➢ The expansion of the universe is accelerating



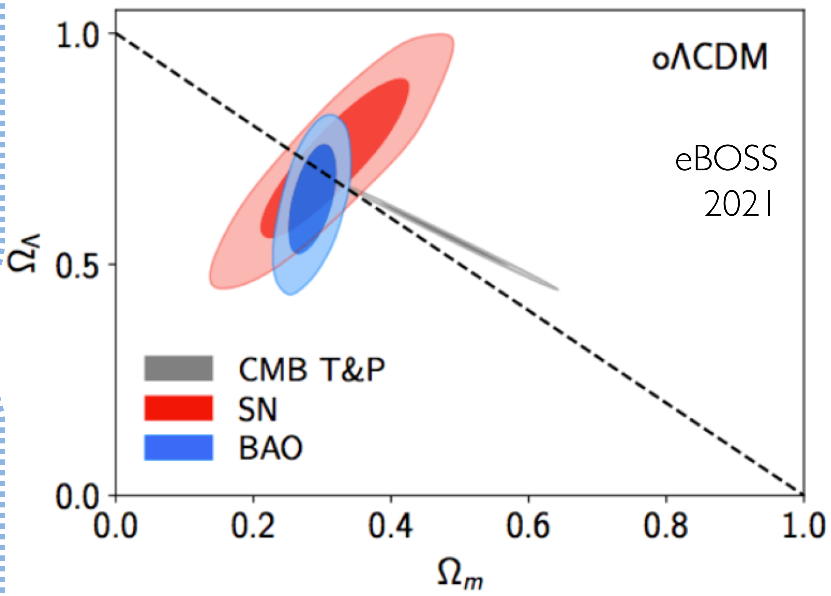
https://newatlas.com/physics/what-is-dark-energy-explained-accelerating-expansion-universe/

- ➢ Type Ia Supernovae (1998)
- ➢ Baryon acoustic peak (2005)

→ 68% of Dark Energy

➔ *Reintroduction of the cosmological constant Λ*

The ΛCDM model



95% of the universe is not understood

Some issues with ΛCDM

- Explaining the acceleration of the expansion with Λ

120 orders of magnitude difference with the theoretical prediction

- Tension on the measurement of H_0

~ 4 to 5 σ tension **within ΛCDM**

Motivations to go

Beyond ΛCDM

Alternatives to ΛCDM

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu}$$

Modified gravity models ?

Look for deviations from Einstein's Field Equations

Additional fluid : Dark energy ?

Look for deviations from

$$w = -1$$

$$w_{\mathrm{DE}}(a) = w_0 + w_a(1-a)$$

Need precise measurement of cosmological parameters to observe deviations from ΛCDM

$M_\nu \equiv \sum m_\nu$

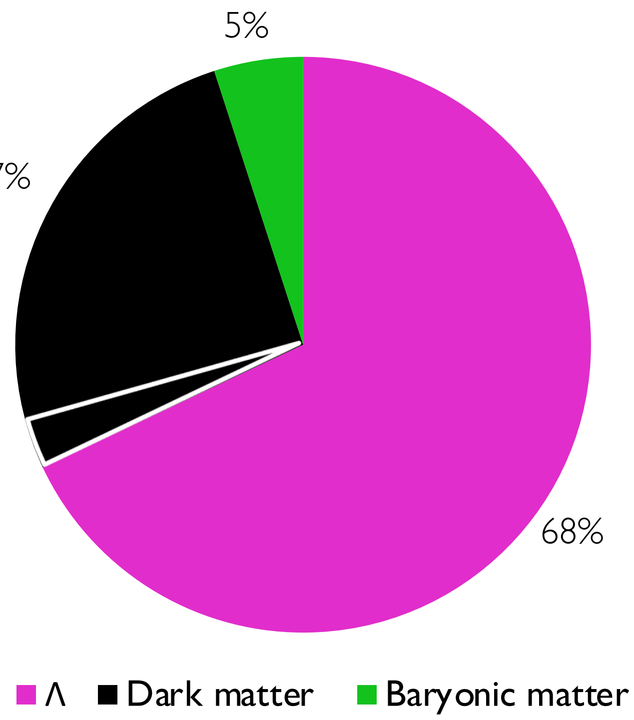
Particle Physics experiment

NH $0.056 \text{ eV} \lesssim M_\nu \lesssim 1 \text{ eV}$

IH $0.095 \text{ eV} \lesssim$

Reduced density of neutrinos in the universe $$\Omega_\nu = \frac{M_\nu}{93.14\ h^2 \text{ eV}}$$

Neutrinos : $0.1\% - 2\%$



With the expansion of the universe, massive neutrinos undergo a transition from relativistic to non-relativistic

$$1 + z_{\text{nr}} = 1890 \left(\frac{m_\nu}{1 \text{ eV}} \right)$$

Need to take this into account to measure small departures from ΛCDM

$M_\nu \equiv \sum m_\nu$

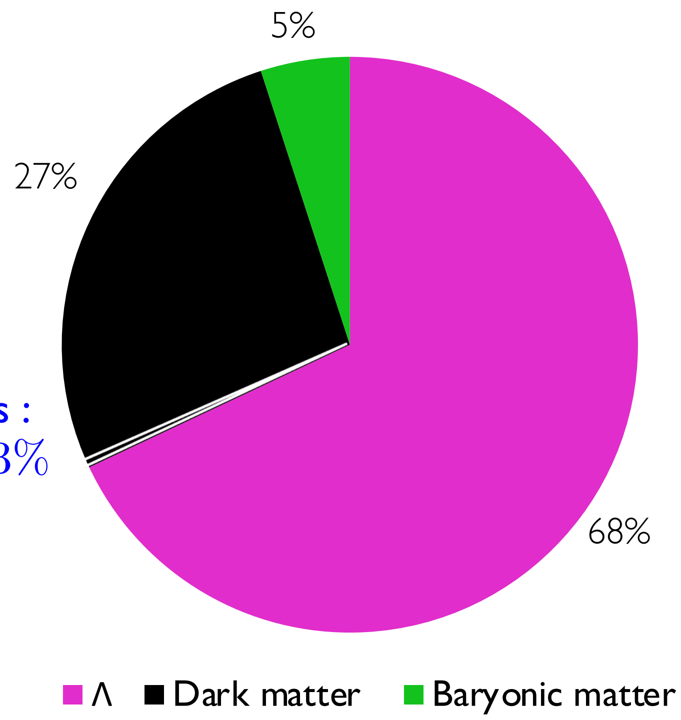
Particle Physics experiment

NH $0.056 \text{ eV} \lesssim M_\nu \lesssim 1 \text{ eV}$

IH $0.095 \text{ eV} \lesssim$

Reduced density of neutrinos in the universe $\Omega_\nu = \dfrac{M_\nu}{93.14\ h^2 \text{ eV}}$

Neutrinos : $0.1\% - 0.3\%$



With the expansion of the universe, massive neutrinos undergo a transition from relativistic to non-relativistic

$$1 + z_{\rm nr} = 1890 \left(\frac{m_\nu}{1 \text{ eV}}\right)$$

Need to take this into account to measure small departures from ΛCDM

Cosmological constraints CMB + BAO + RSD + SNIA $M_\nu < 0.1 \text{ eV}$ [eBOSS collaboration 2020]

The Large Scale Structure is the result of :

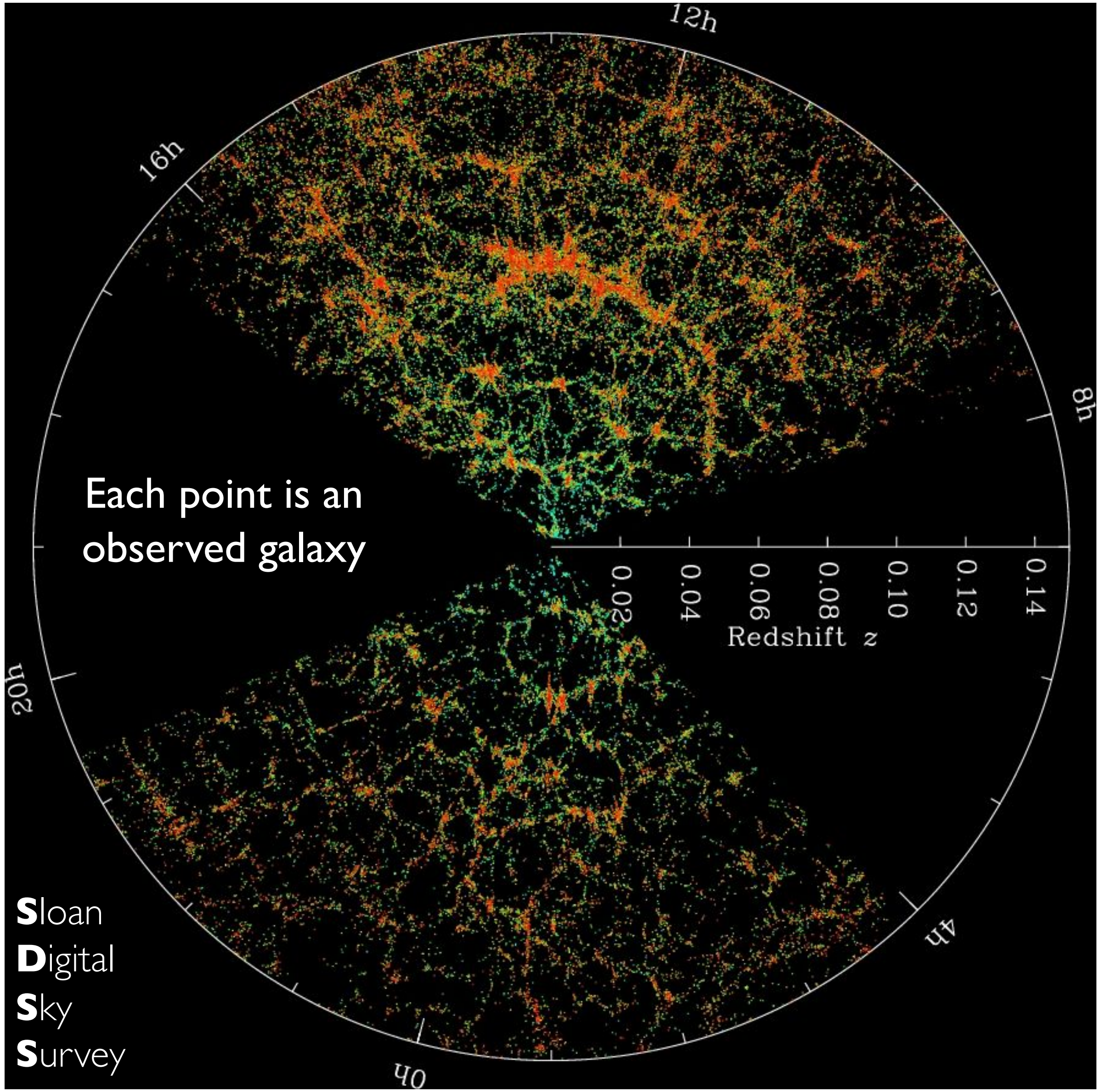
Gravity → **Modified gravity**

+ Expansion → **Dark Energy**

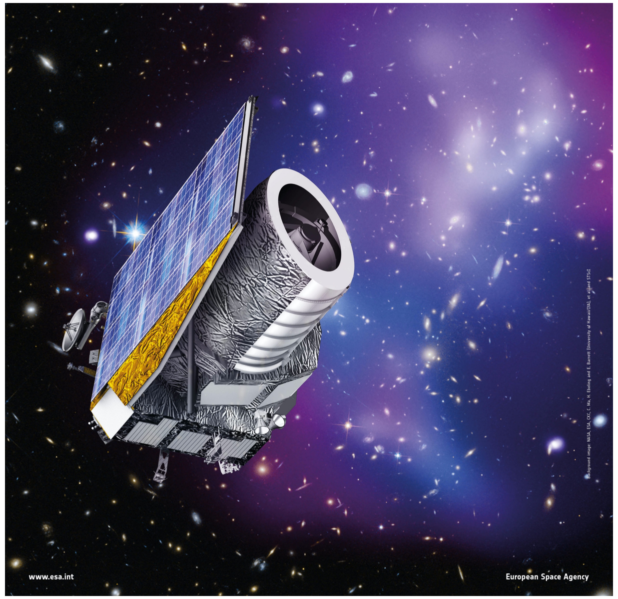
+ Interaction between baryons, photons, neutrinos, CDM → **Nature of Dark Matter, Total neutrino mass**

+ Initial conditions → **Primordial universe**

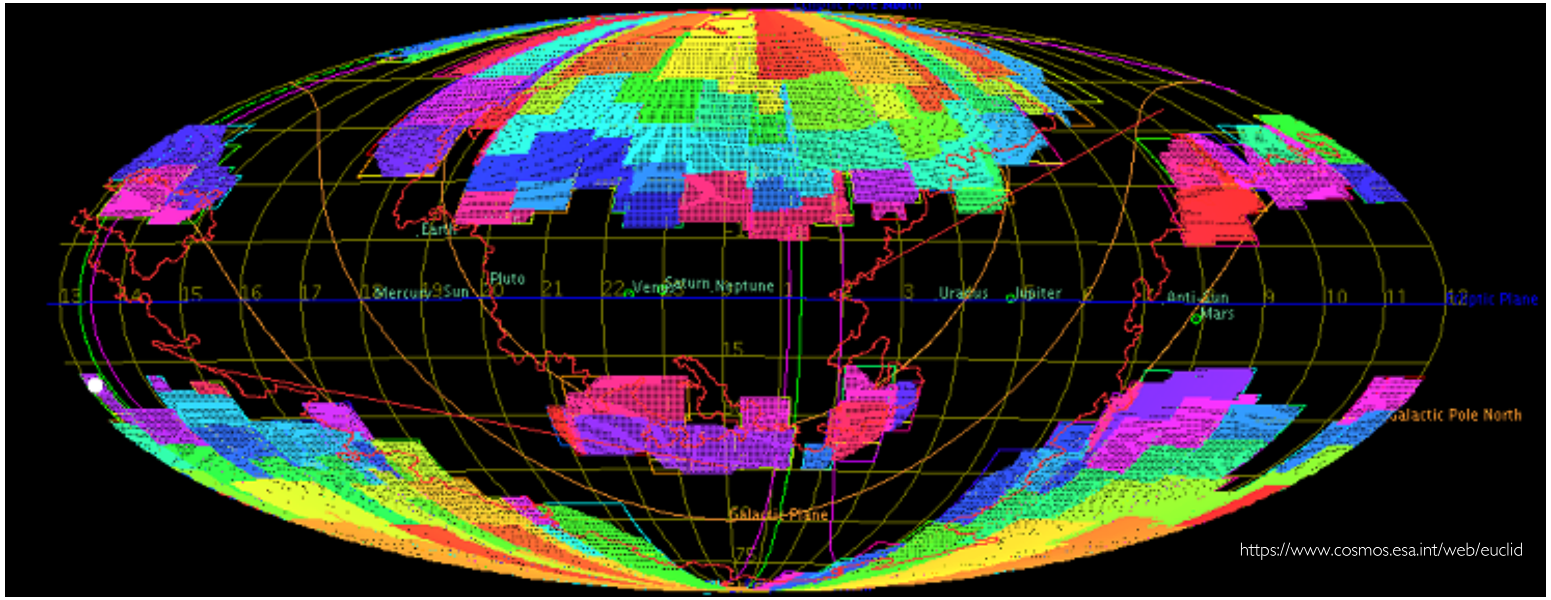
Contains a lot of cosmological information to test models



Space telescope : February 5th 2023



Sky coverage of 15 000 deg^2 for 6 years of observations



2 instruments :

- VIS (visible) : Shapes of **10^9** with **Photo-z** $\in [0.001, 2.5]$
- NISP (infrared) : Spectrum of **50×10^6** galaxies with **Spectro-z** $\in [0.9, 1.8]$

2 main probes :

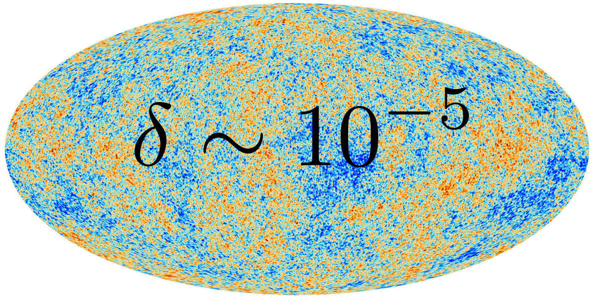
- Weak Lensing
- Galaxy Clustering

NISP = Near Infrared Spectrometer and Photometer
VIS = VISible imager

II. Cosmological parameter inference with the Large Scale Structure

- Statistics of the Large Scale Structure
- Parameter inference

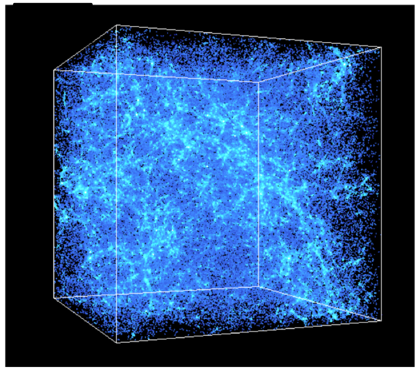
z = 1100



CMB anisotropies = seeds

Linear clustering →

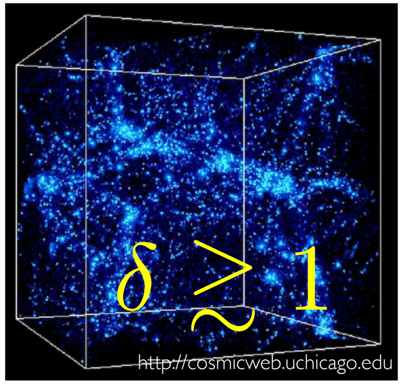
z = 5



Non-linear clustering →

on small scales

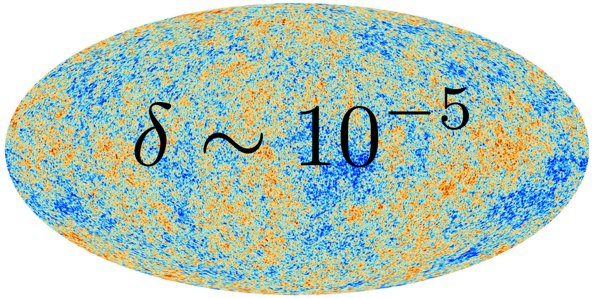
z = 0



The density contrast field

$$\delta(\boldsymbol{x},t) \equiv \frac{\rho(\boldsymbol{x},t) - \bar{\rho}(t)}{\bar{\rho}(t)}$$

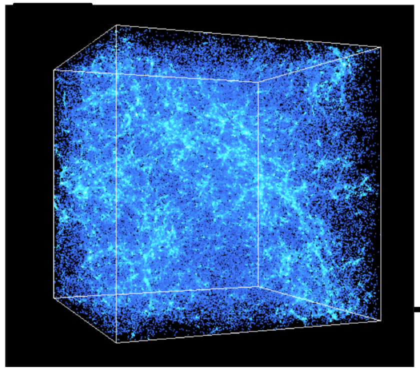
z = 1100



δ follows a Gaussian PDF

Linear clustering

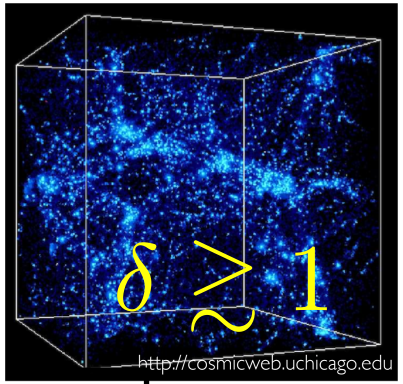
z = 5



Non-linear clustering

on small scales

z = 0



The density contrast field

$$\delta(\boldsymbol{x},t) \equiv \frac{\rho(\boldsymbol{x},t)-\bar{\rho}(t)}{\bar{\rho}(t)}$$

Statistical description

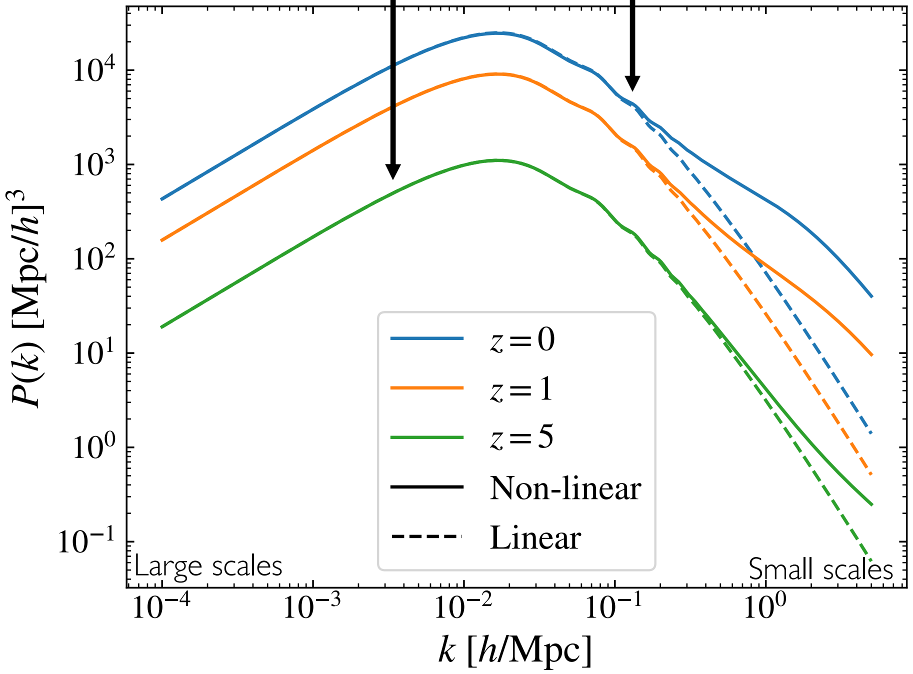
- By definition : $\langle \delta(\boldsymbol{x}) \rangle = 0$
- 2-point correlation function :

$$\langle \delta(\boldsymbol{x})\delta(\boldsymbol{x}+\boldsymbol{r}) \rangle \equiv \xi(\boldsymbol{r})$$

$$k = \frac{2\pi}{r}$$

In Fourier space : The power spectrum

$$\langle \delta(\mathbf{k})\delta(\mathbf{k}') \rangle = \delta_{\mathrm{D}}(\mathbf{k}+\mathbf{k}')P(k)$$



Inference of cosmological parameters from a data set : $\begin{cases} \boldsymbol{\theta} : \Omega_{\rm i},\ H_0,\ M_\nu ... \\ \mathbf{D} : P(k), \xi(r) ... \end{cases}$

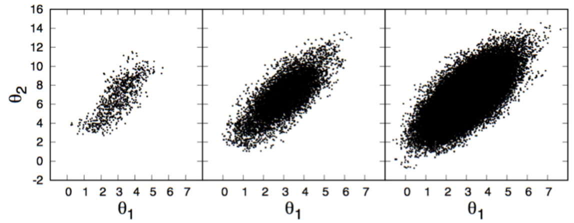
Likelihood → Prior →

$$P(\boldsymbol{\theta}|\mathbf{D}) = \frac{L(\mathbf{D}|\boldsymbol{\theta})\mathcal{P}(\boldsymbol{\theta})}{P(\mathbf{D})}$$

Normalisation

Posterior

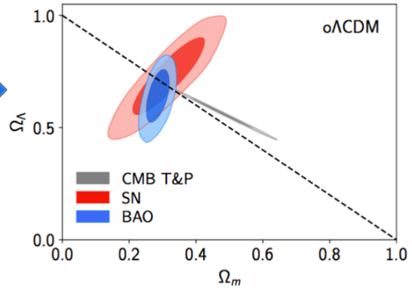
Bayes theorem



Sample the posterior through statistical processes such as Monte Carlo Markov Chains (MCMC)

Best-fit and errors

$$\theta = \hat{\theta}_{\rm bf} \pm \hat{\sigma}_\theta$$



For 2-point statistics [P(k) or ξ(r)] we generally assume a Gaussian likelihood

$$\chi^2 \equiv -2 \ln L(P(k)|\boldsymbol{\theta}) = \left[[P(k) - P(k;\boldsymbol{\theta})^{\rm theo}]^T \mathbf{C}^{-1} [P(k) - P(k;\boldsymbol{\theta})^{\rm theo}]\right] + \rm Cst.$$

Data

$$C_{ij} \equiv \langle P(k_i)P(k_j)\rangle - \langle P(k_i)\rangle\langle P(k_j)\rangle$$

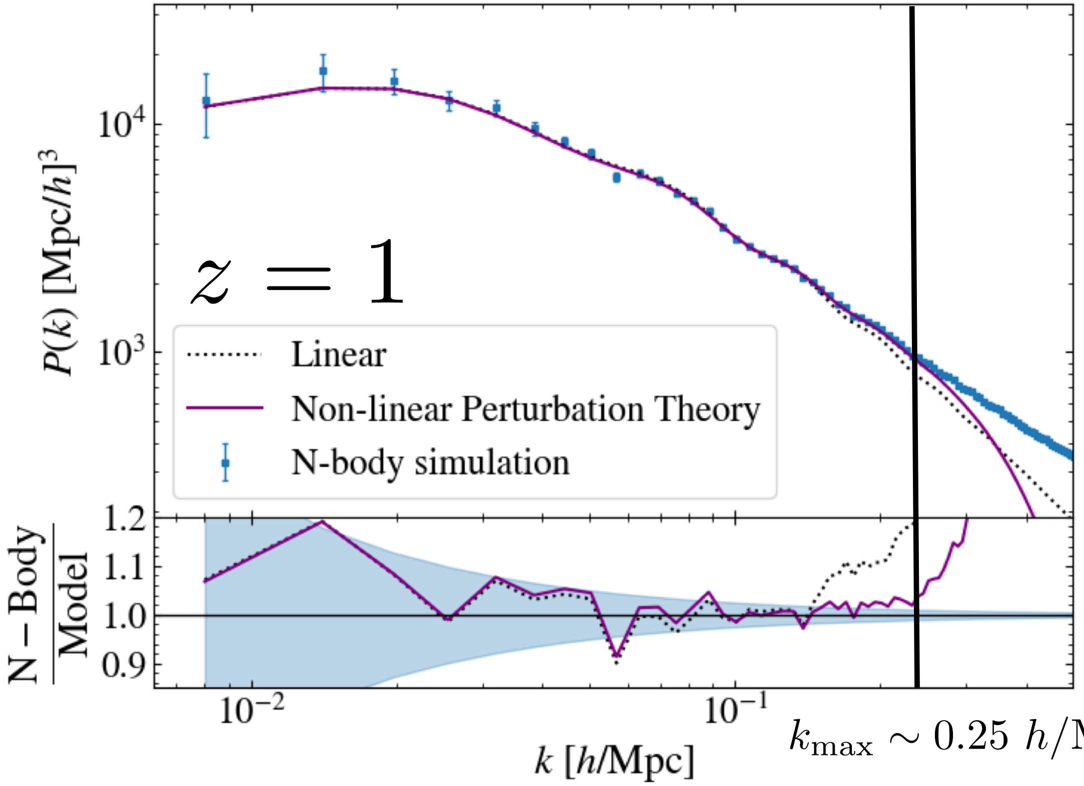
Model

Covariance : Errors and correlations

Non-linear clustering and non-Gaussian covariance

Non-linear clustering on small scales → *δ becomes non-Gaussian*

Simulate non-linear clustering with N-Body simulations → Test models on N-body simulations



$k_{\rm max} \sim 0.25\ h/{\rm Mpc}$

<u>Covariance</u> : Errors and correlations in the power spectrum [Scoccimaro et al. 1999]

$$\mathbf{C}(k_i, k_j) = \frac{P(k_i)^2}{N_{k_i}}\delta_{ij} + \bar{T}(k_i, k_j)$$

Trispectrum

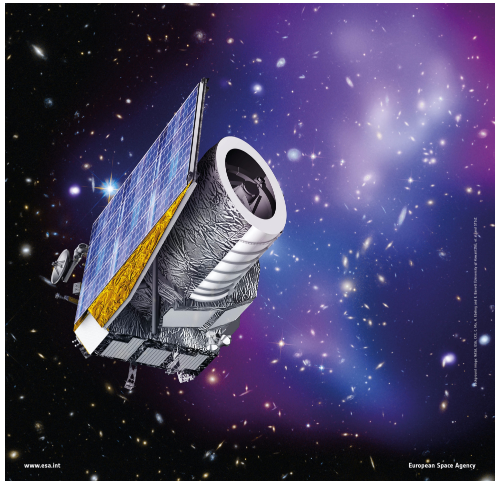
$$\langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\delta(\mathbf{k}_4)\rangle_c = \delta_{\rm D}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4)T(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$$

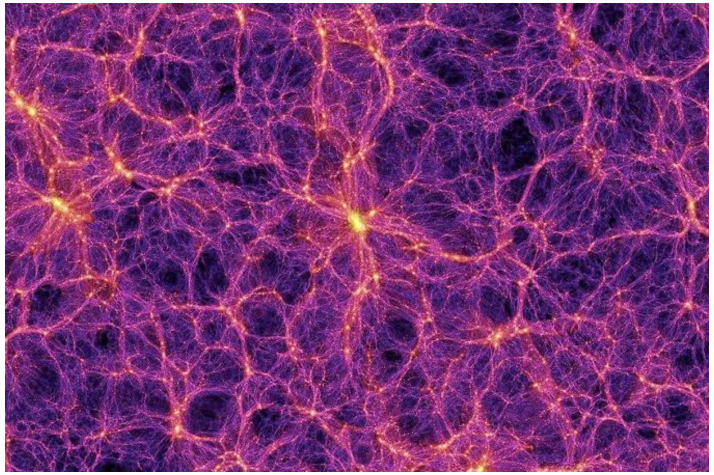
Difficult to predict analytically

Estimate the covariance from simulations

$$\hat{C}_{ij} = \frac{1}{N_m - 1}\left[\sum_n^{N_m}[P^{(n)}(k_i) - \bar{P}(k_i)][P^{(n)}(k_j) - \bar{P}(k_j)]\right]$$

With Euclid we will reach a high precision on LSS observables





To meet the challenges of modern cosmology :
Test alternative models, neutrino mass, tensions, …

→ *Important to estimate unbiased best-fit and errors on cosmological parameters !*

Challenges :

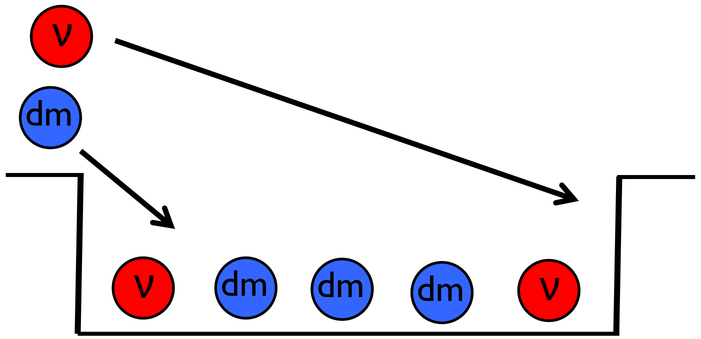
- Combination of probes ➔ **Accurate and precise covariance**
- Non-linear scales ➔ **Accurate modeling**

III. Optimal parameter inference with the power spectrum including massive neutrinos

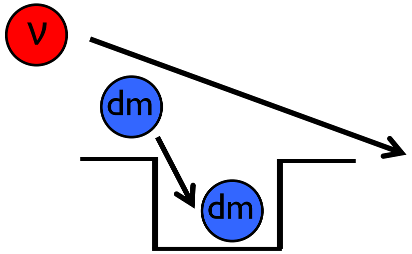
- Parameter inference and sampling noise in the covariance
- Impact of non-Gaussian covariance on cosmological constraints
- Power spectrum non-linear modeling challenge with massive neutrinos

Massive Neutrinos keep a high velocity dispersion

Free streaming : Neutrinos escape from the potential wells on small scales



<u>On large scales</u> : Neutrinos behave like CDM



<u>On small scales</u> : Neutrinos do not cluster

Smoothing of density perturbations on small scales

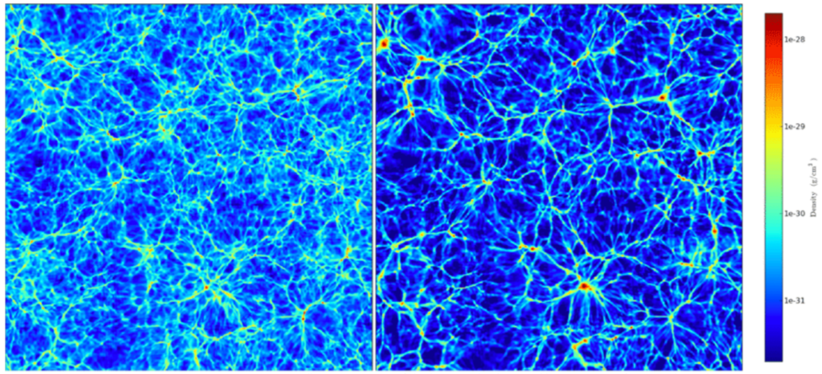
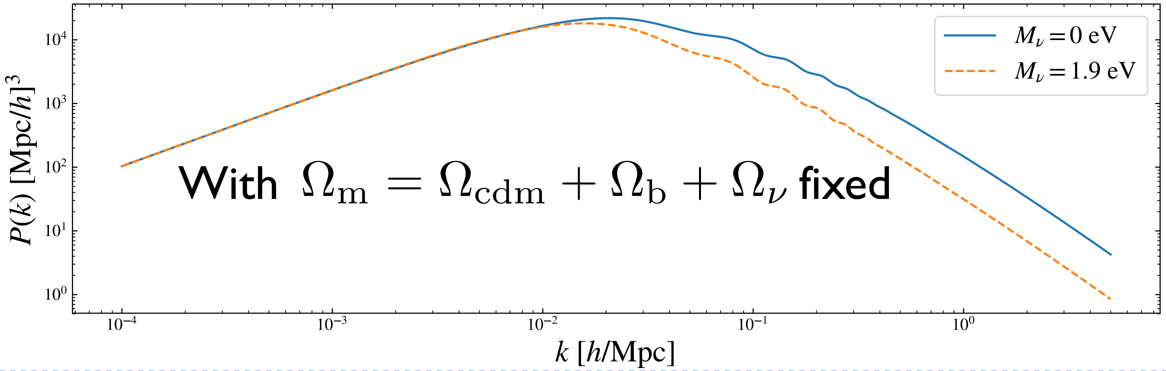


Illustration: Courtesy of Shankar Agarwal and Hume Feldman, University of Kansas; submitted to Mon. Not. R. Astron. Soc.

$M_\nu = 1.9$ eV $\qquad$ $M_\nu = 0$ eV

Imprint on the power spectrum on small scales

$$\Omega_\nu = \frac{M_\nu}{93.14\ h^2\ \mathrm{eV}}$$



DEMNUni : Dark Energy and Massive Neutrinos Universe

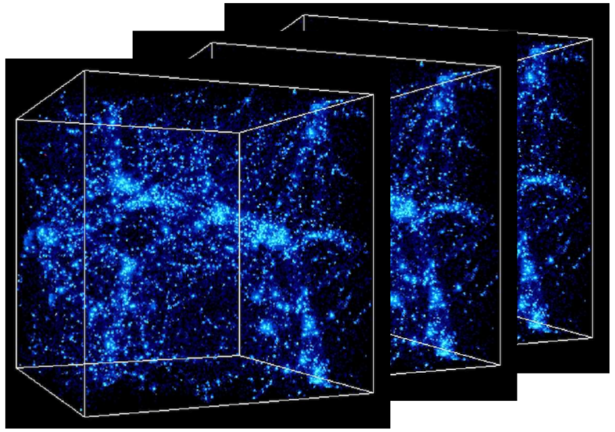
- Cosmologies : 0ν and 16ν

$$\begin{cases} \Omega_m = 0.32 \\ \Omega_b = 0.05 \\ h = 0.67 \\ n_s = 0.96 \\ A_s = 2.1265 \times 10^9 \\ \mathbf{3m_\nu = 0} \text{ or } \mathbf{0.16} \text{ eV} \\ \mathbf{\Omega_{cdm} = 0.27} \text{ or } \mathbf{0.2663} \end{cases}$$

- V = 1 [Gpc/h]^3
- N_p = 1024^3 (+ 1024^3)

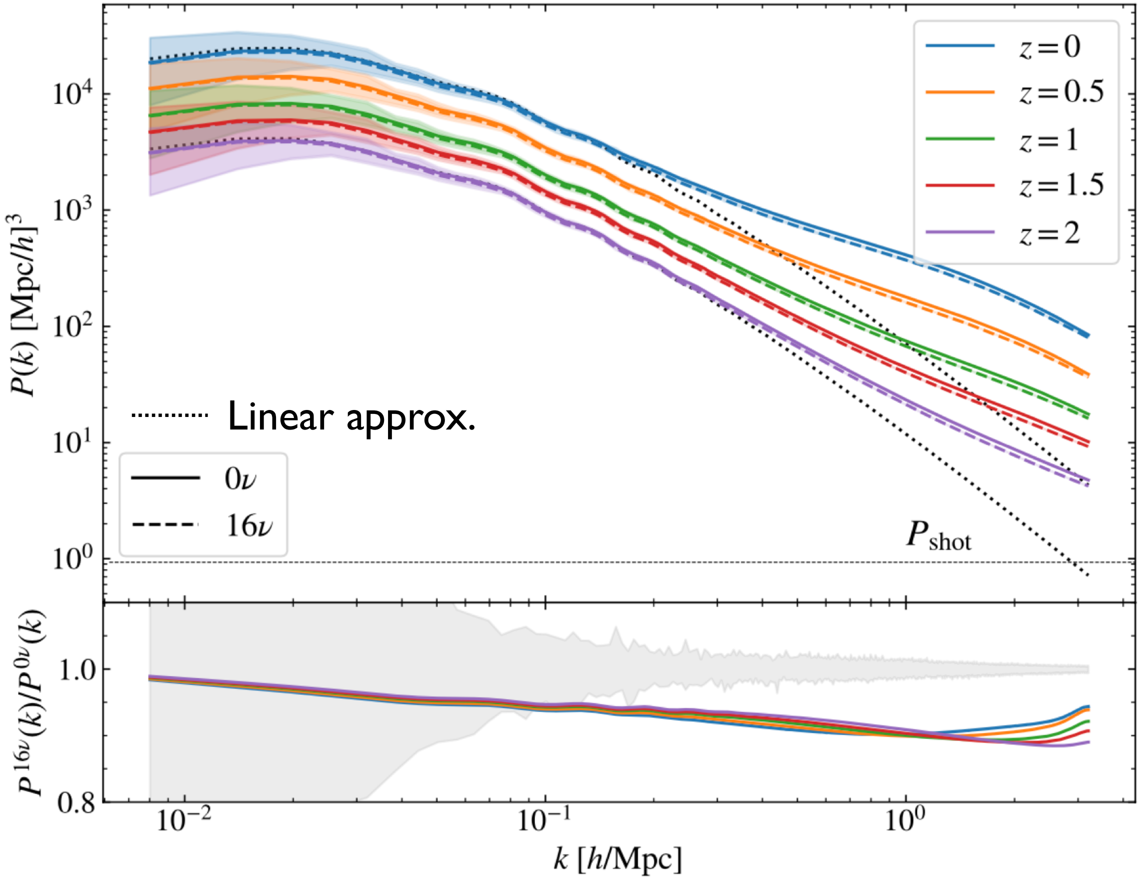
Managed by Carmelita Carbone
[Carbone et al. 2016, Castorina et al. 15]

N-body simulations of CDM + Massive neutrinos at $z \approx 0,\ 0.5,\ 1,\ 1.5,\ 2$



*DEMNUni-**Cov*** :
X 50 realisations for each cosmology

Mean CDM power spectrum on 50 DEMNUni-Cov



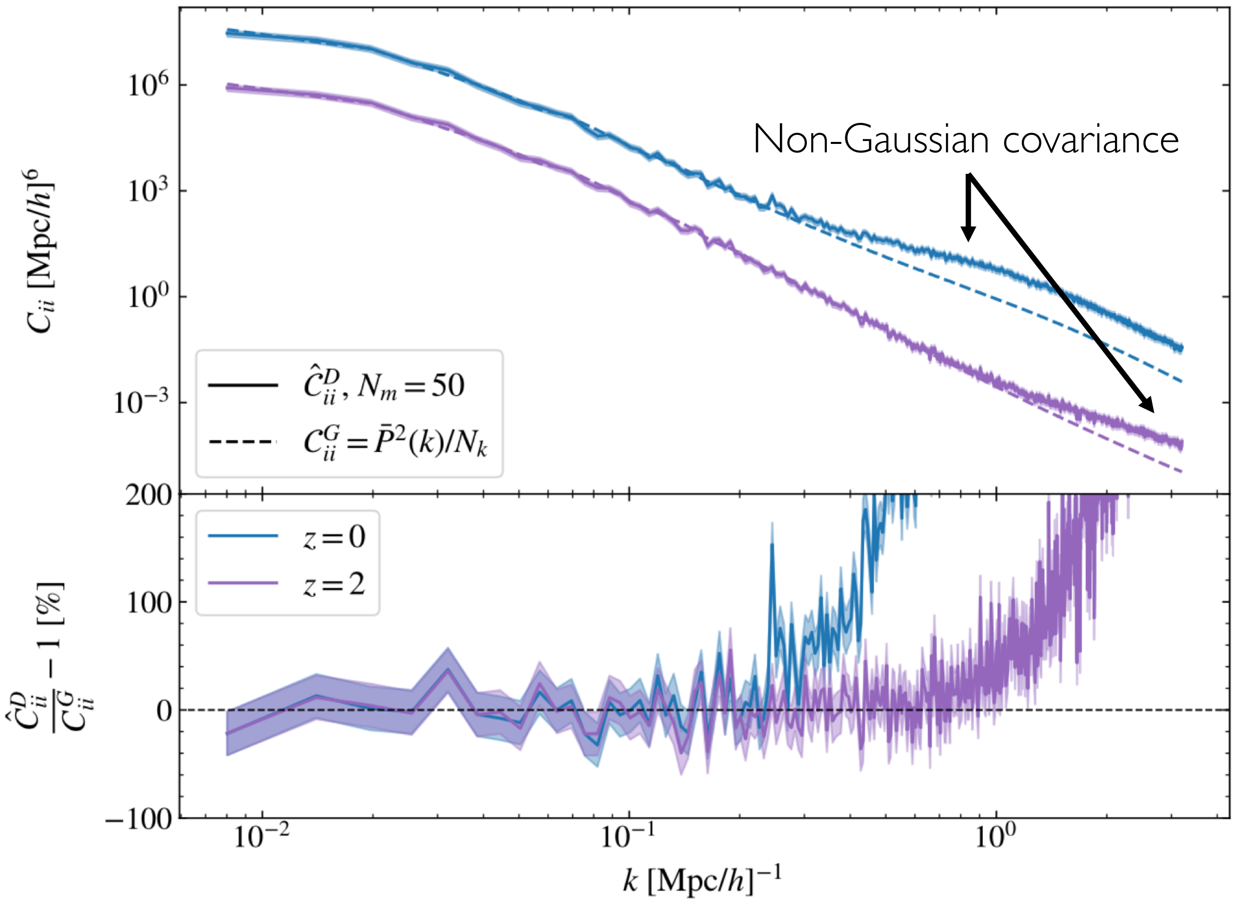
Goal →

Estimation of M_ν from the full shape of the power spectrum

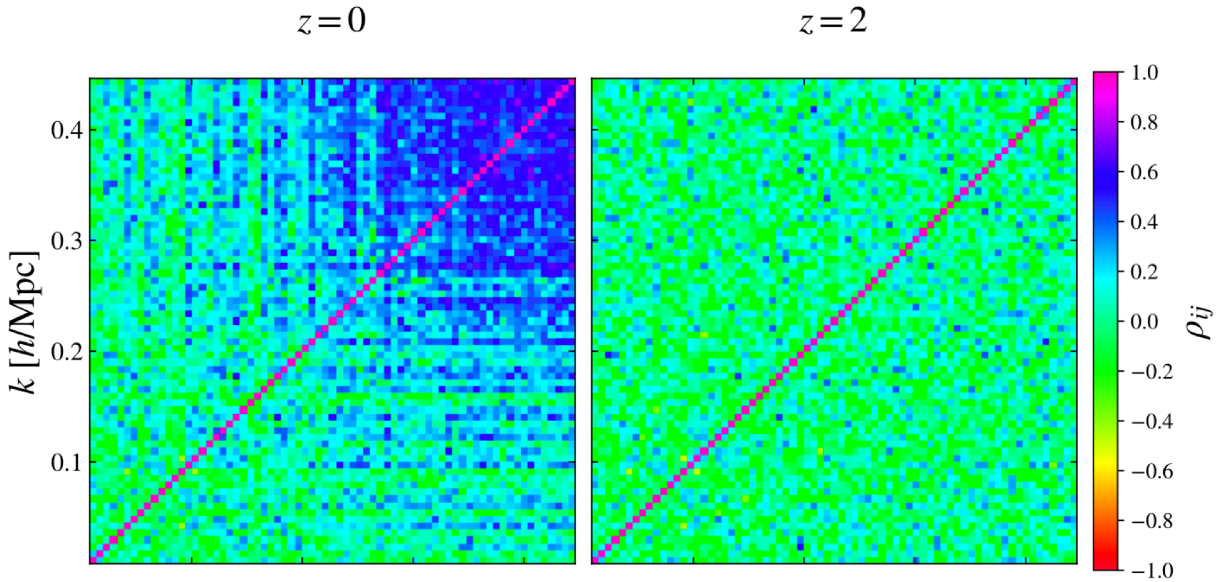
1. Test the covariance
2. Test the non-linear modeling of P(k)

Estimation of the covariance with the 50 realisations $\hat{C}_{ij} = \frac{1}{N_m - 1}\left[\sum_{n}^{N_m}[P^{(n)}(k_i) - \bar{P}(k_i)][P^{(n)}(k_j) - \bar{P}(k_j)]\right]$

Diagonal of the covariance



Correlation matrix



N-Body simulations = Accurate non-linarities and non-Gaussianities but CPU-time consuming

With only 50 realisations, the covariance is noisy

For N_m realisation of the power spectrum $\hat{C}_{ij} = \frac{1}{N_m - 1}\left[\sum_n^{N_m}[P^{(n)}(k_i) - \bar{P}(k_i)][P^{(n)}(k_j) - \bar{P}(k_j)]\right]$ ➡ **Sampling noise**

$$\chi^2 = [[P(k) - P(k;\theta)^{\rm theo}]^T \mathbf{C}^{-1}[P(k) - P(k;\theta)^{\rm theo}]] \qquad \Psi \equiv \mathbf{C}^{-1}$$

- Hartlap bias in the **precision matrix** $\mathbf{\Psi}$

[Wishart 1928, Hartlap 2007]

$$\hat{\mathbf{\Psi}} = \hat{\mathbf{C}}^{-1} \Rightarrow \langle\hat{\mathbf{\Psi}}\rangle = \frac{N_m - 1}{N_m - N_k - 2}\mathbf{\Psi}$$

- Noise in the estimated precision matrix : $\hat{\mathbf{\Psi}} = \mathbf{\Psi} + \Delta\mathbf{\Psi}$

➔ Additional variance and bias on the estimated errors and best-fit of parameters

[Taylor et al. 2013 ; Dodelson & Schneider 2013 ; Percival et al. 2014 ;Taylor & Joachimi 2014]

Solutions to reduce sampling noise effects :

- Create very large sets of mocks : **Covmos** [Baratta et al. 2019, Baratta et al. In prep.]
- Alternative estimators of the covariance : **NERCOME** [Joachimi 2016]

➡ *Test of the reliability of these methods*

Covmos : Fast mock generation

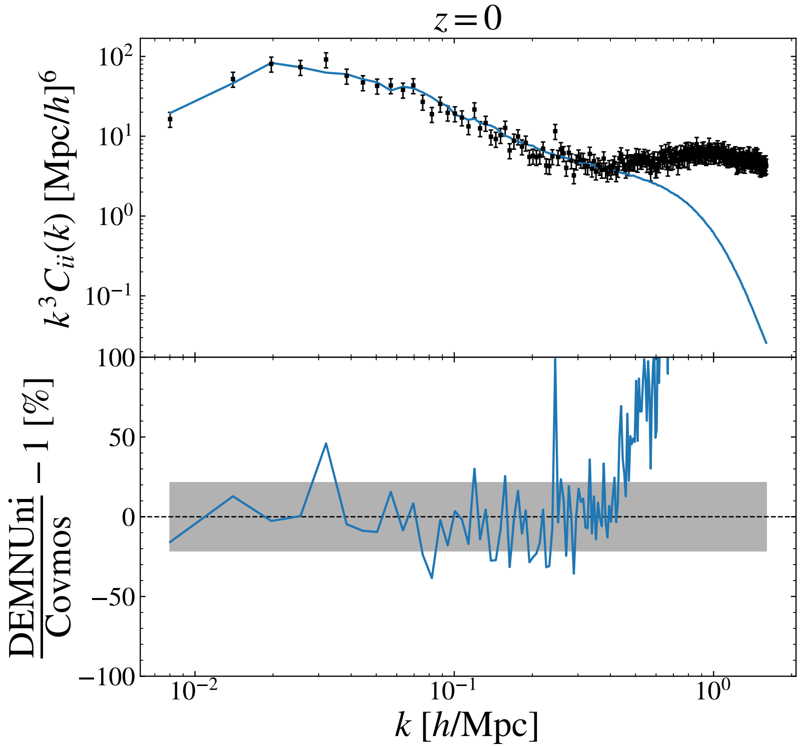
Target P(k) and PDF(δ) to clone DEMNUni-Cov

$$\mathbf{C}(k_i,k_j) = \frac{P^2(k_i)}{N_{k_i}}\delta_{ij} + \bar{T}(k_i,k_j) = \mathbf{C}^{\rm G}(k_i) + \mathbf{C}^{\rm NG}(k_i,k_j),$$

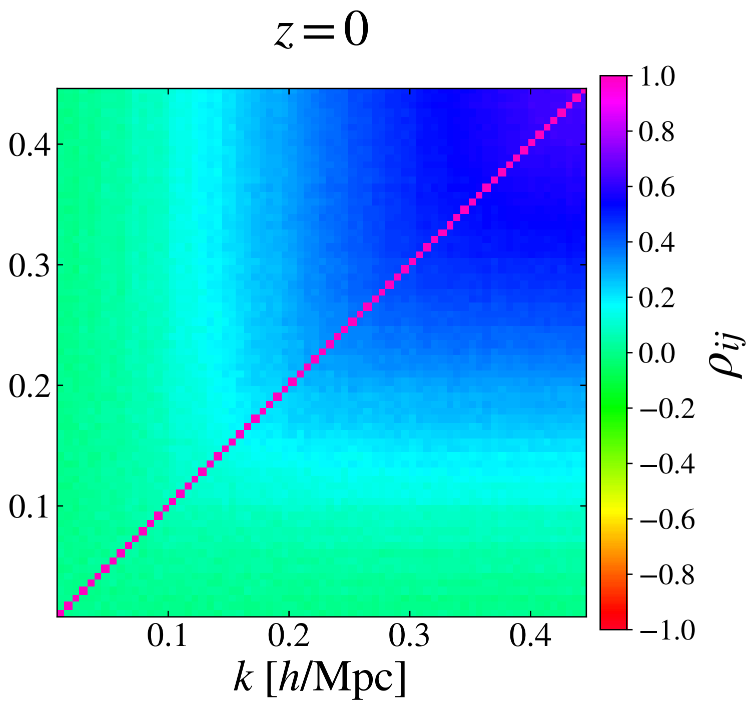
Generate **10 000** catalogs at 5 redshifts (~ 24h X 5)

Developed by Philippe Baratta

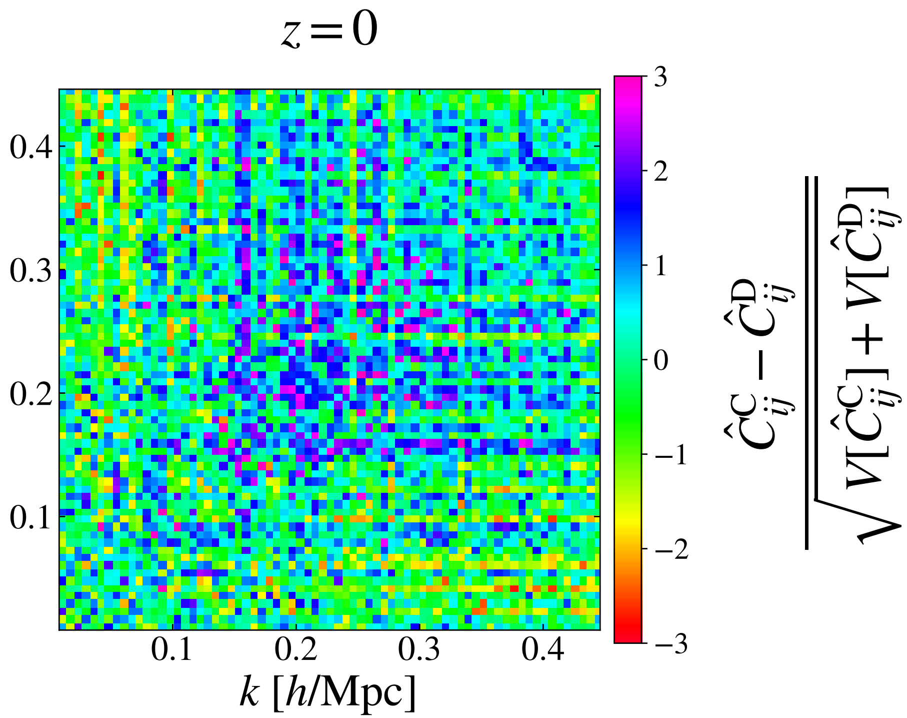
Diagonal of the covariance



Correlation matrix



Deviation with respect to DEMNUni-Cov in σ



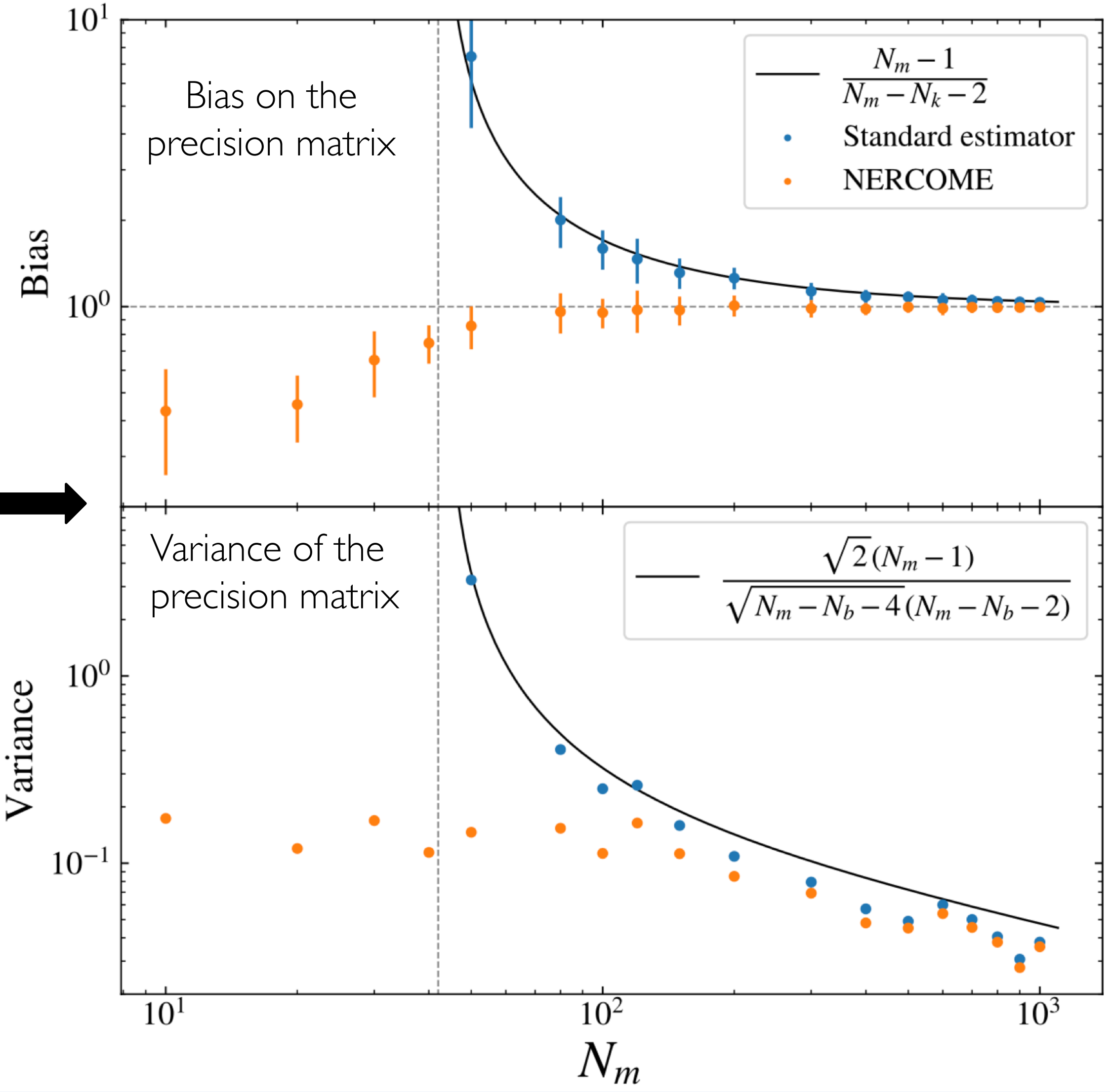
NERCOME : Non-parametric Eigenvalue-Regularised COvariance Matrix Estimator

For N_m mocks :

1. Estimate the covariance on 2 subsets of size s and N_m-s : $\hat{S}_1$ and $\hat{S}_2$
2. Eigenvalue Decomposition : $\hat{S}_i = U_i D_i U_i^T$
3. Estimate the new covariance as : $\hat{C}_N = U_1 \mathrm{diag}(U_1^T \hat{S}_2 U_1) U_1^T$

Repeat for 500 random subsets of size s and N_m-s

For N_k = 40



- Reduces the bias in the precision matrix
- Reduces the variance in the precision matrix

[Joachimi 2016]

- Tests on the covariance :
  - NERCOME : Need a true reference covariance matrix + several realisations of the covariance

- Generate 100 000 + 1 Covmos P(k) x 5 redshifts

  For the covariance (100 000) — For the data-vector (1)

- Divide into 100 subsets with $N_m = 45$

- Estimate the covariance on all subsets with NERCOME and the standard estimator

- Estimate cosmological parameters with MCMC for all the covariance matrices

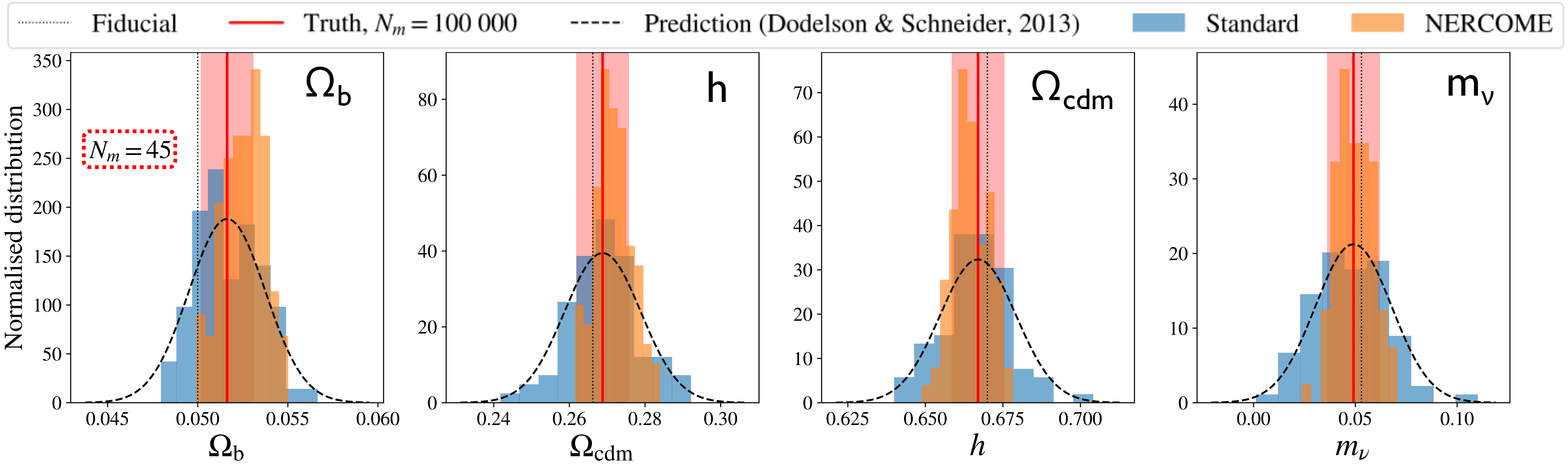
$\boldsymbol{\theta}$	Priors
Ω_b	$[0.01,\ 0.1]$
Ω_{cdm}	$[0.01,\ 0.8]$
h	$[0.3,\ 1.5]$
m_ν [eV]	$[0,\ 1]$

- $k_{max} = 0.2$ h/Mpc ➔ $N_k = 30$

- Non-linear model : Halofit

Testing NERCOME on parameter estimation

Distribution of best-fits for 100 realisations of the covariance



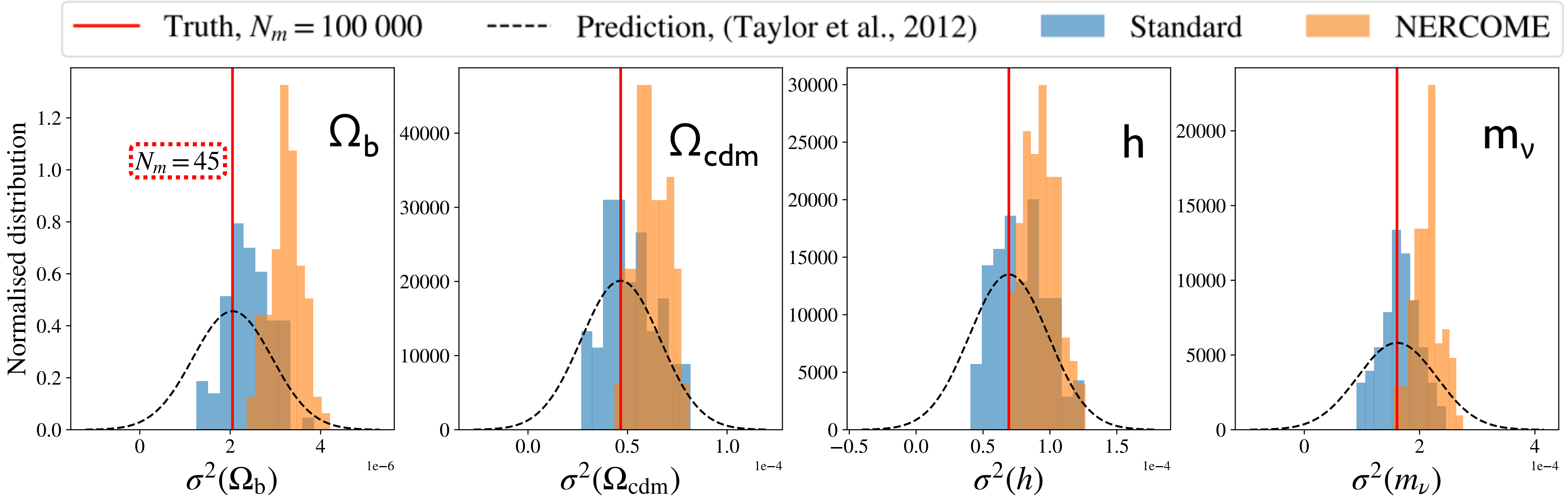
- **Standard estimator : best-fit dispersion of 1.3σ** (larger than the true error)
- **NERCOME : best-fit dispersion of ~0.7σ** (smaller than the true error)

NERCOME reduces the dispersion by 50%

- **Good agreement with the theoretical prediction** [Dodelson & Schneider 2013]

Testing NERCOME on parameter estimation

Distribution of errors for 100 realisations of the covariance



- Standard estimator : variance dispersion of 0.25σ
- NERCOME : variance dispersion of ~0.17σ

} NERCOME reduces the dispersion by 30%

- NERCOME overestimates the error-bars by ~ 30% for N_m = 45
- Poor agreement with the theoretical prediction [Taylor et al. 2012]

Testing Covmos on parameter estimation

- Tests on the covariance :
  - NERCOME ✓
  - With Covmos we can supress sampling noise. But is the covariance accurate enough ?

→ *Test against covariance from DEMNUni-Cov*

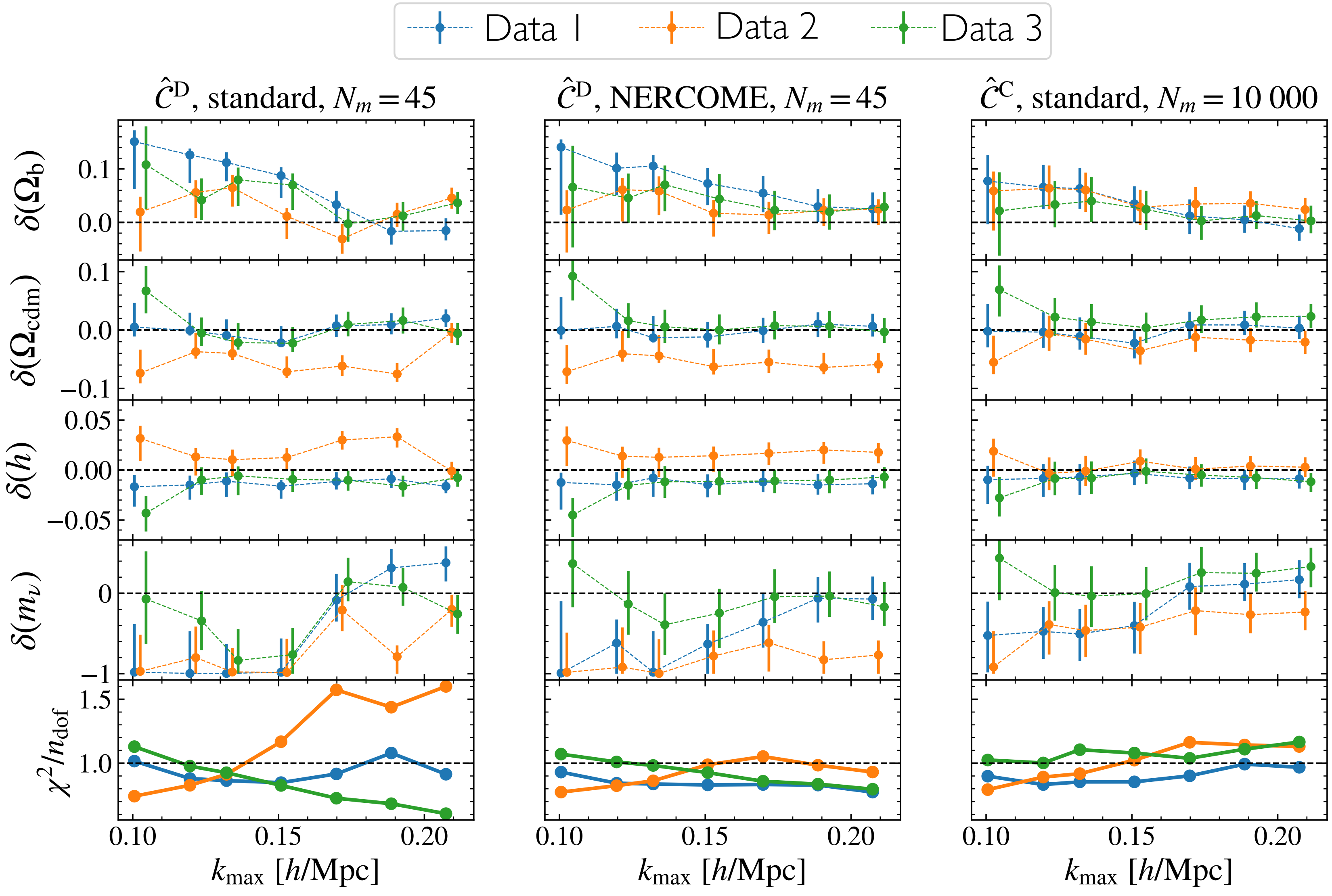
Testing Covmos on parameter estimation

- Tests on the covariance :
  - NERCOME ✅
  - With Covmos we can supress sampling noise. But is the covariance accurate enough ?

→ *Test against covariance from DEMNUni-Cov*

Parameter estimation with MCMC :

- Data-Vector : DEMNUni-Cov P(k) with 5 redshifts X *3 independent realisations of the data-vector*
- For different Covariance :
  1. DEMNUni-Cov with N_m=45 → *Accurate but noisy*
  2. DEMNUni-Cov with N_m=45, estimated with NERCOME → *Accurate and less noisy*
  3. Covmos, cloned from DEMNUni-Cov, with N_m = 10 000 → *No noise but need to test accuracy*
- k_{max} = [0.1, 0.2] h/Mpc



- **Standard DEMNUni-Cov (left):**

  Large scatter around the fiducial cosmology

- **NERCOME DEMNUni-Cov (middle):**

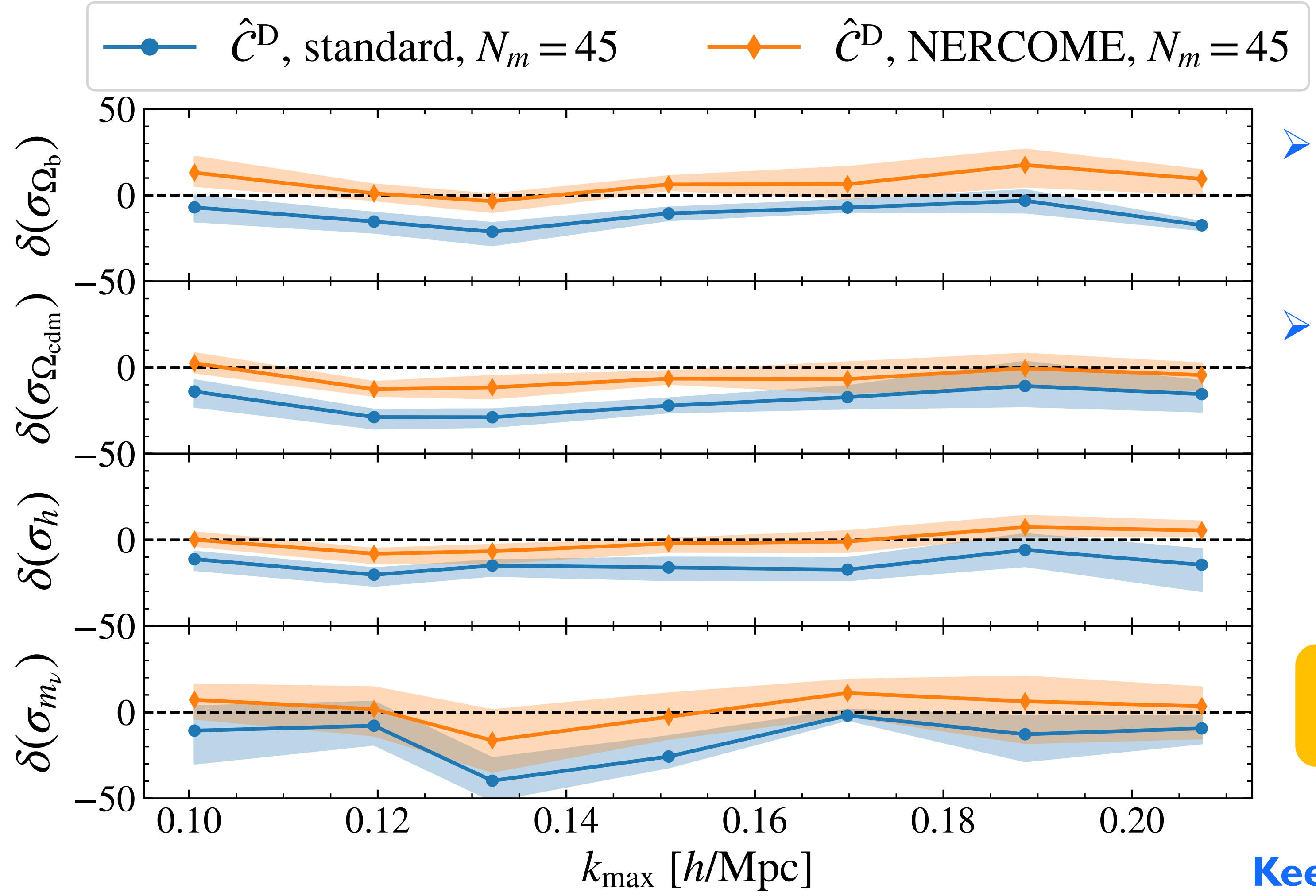
  Less scattering, but still systematic deviation for Data 2

- **Covmos (right):**

  All data-sets compatible with each other

➔ Covmos provides unbiased best-fits

Relative difference with respect to the variance obtained with Covmos N_m=10 000 : $\delta(\sigma_\theta) = \sigma_\theta/\sigma_\theta^{\text{Covmos}} - 1[\%]$



- NERCOME N_m = 45 give similar error bars than Covmos
- We found before that NERCOME leads to ~ 30% overestimation of errors for N_m = 45

Covmos slightly overestimate of error-bars

➔ overestimation the covariance

Need more N-body realisations to precisely assess bias in error with covmos

Keep covmos with Nm = 10 000 for the following

Impact of trispectrum contribution

- Tests on the covariance :
  - **Sampling noise effects** (NERCOME and Covmos) ✓
  - **Non-Gaussian covariance :** $$\mathbf{C}(k_i, k_j) = \frac{P^2(k_i)}{N_{k_i}}\delta_{ij} + \bar{T}(k_i, k_j) = \mathbf{C}^{\mathrm{G}}(k_i) + \mathbf{C}^{\mathrm{NG}}(k_i, k_j),$$

Should increase the error on cosmological parameters

By how much ?

Impact of trispectrum contribution

- Tests on the covariance :
  - **Sampling noise effects** (NERCOME and Covmos) ✓
  - Non-Gaussian covariance : $\mathbf{C}(k_i, k_j) = \frac{P^2(k_i)}{N_{k_i}}\delta_{ij} + \bar{T}(k_i, k_j) = \mathbf{C}^{\mathrm{G}}(k_i) + \mathbf{C}^{\mathrm{NG}}(k_i, k_j),$

Should increase the error on cosmological parameters

By how much ?

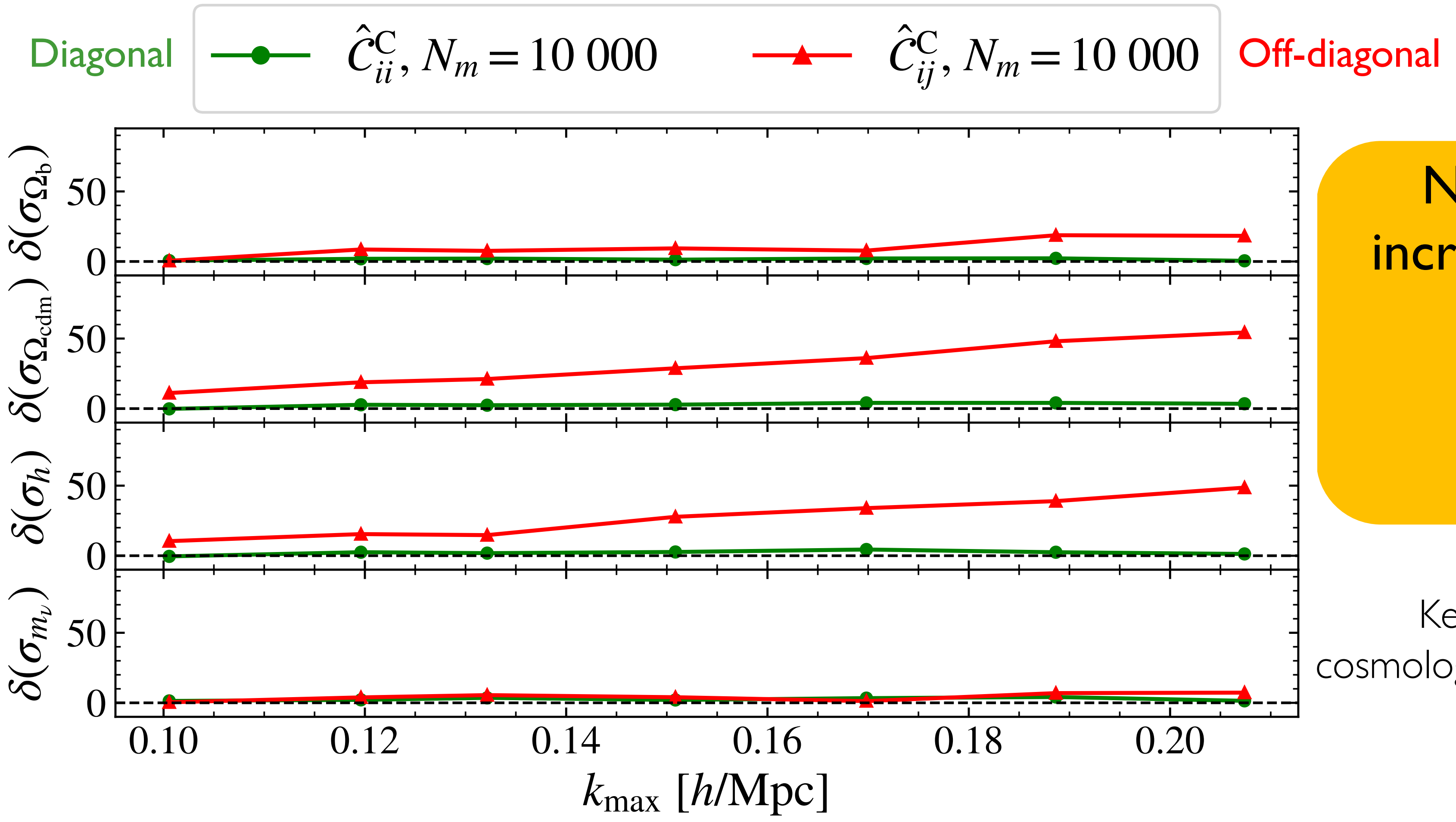
Compare constraints obtained with the Gaussian covariance only, to quantify the impact of non-Gaussian covariance

Parameter estimation with MCMC :

- Data-vector : DEMNUni-Cov at 5 redshifts : $z \approx 0,\ 0.5,\ 1,\ 1.5,\ 2$
- Non-linear model : Halofit
- **Covariance :**
  1. Gaussian only
  2. Covmos Nm = 10 000, diagonal only
  3. Covmos Nm = 10 000, full

Relative difference with respect to the error obtained with Gaussian only

$$\delta(\sigma_\theta) = \sigma_\theta / \sigma_\theta^{\mathrm{G}} - 1[\%]$$



Non-Gaussian covariance increases error bars by ~ 50% for Ω_{cdm} and h

Error on m_ν is not affected

Keep in mind that we found a slight cosmological error overestimation with Covmos

Testing non-linear models with massive neutrinos

- Tests on the covariance :
  - Sampling noise effects ✓
  - Non-Gaussian covariance ✓

→

- Test non-linear power spectrum models against DEMNUni-Cov

Testing non-linear models with massive neutrinos

- Tests on the covariance :
  - Sampling noise effects ✓
  - Non-Gaussian covariance ✓

→ Test non-linear power spectrum models against DEMNUni-Cov

The effect of massive neutrinos is more important on small scales → Non-linearities

To model the non-linear power spectrum

- Non-linear perturbation theory

Perturbative expansion of the density field

$$\langle(\delta_1 + \delta_2 + \delta_3 + ...)(\delta_1 + \delta_2 + \delta_3 + ...)\rangle = P(k)\delta_{\rm D}(\mathbf{k} + \mathbf{k}')$$

- Find appropriate fitting functions

→ Calibration on N-Body simulations

With massive neutrinos

- Non-linear perturbation theory

Difficult to account for neutrino perturbations

→ Account for massive neutrinos at the linear level only

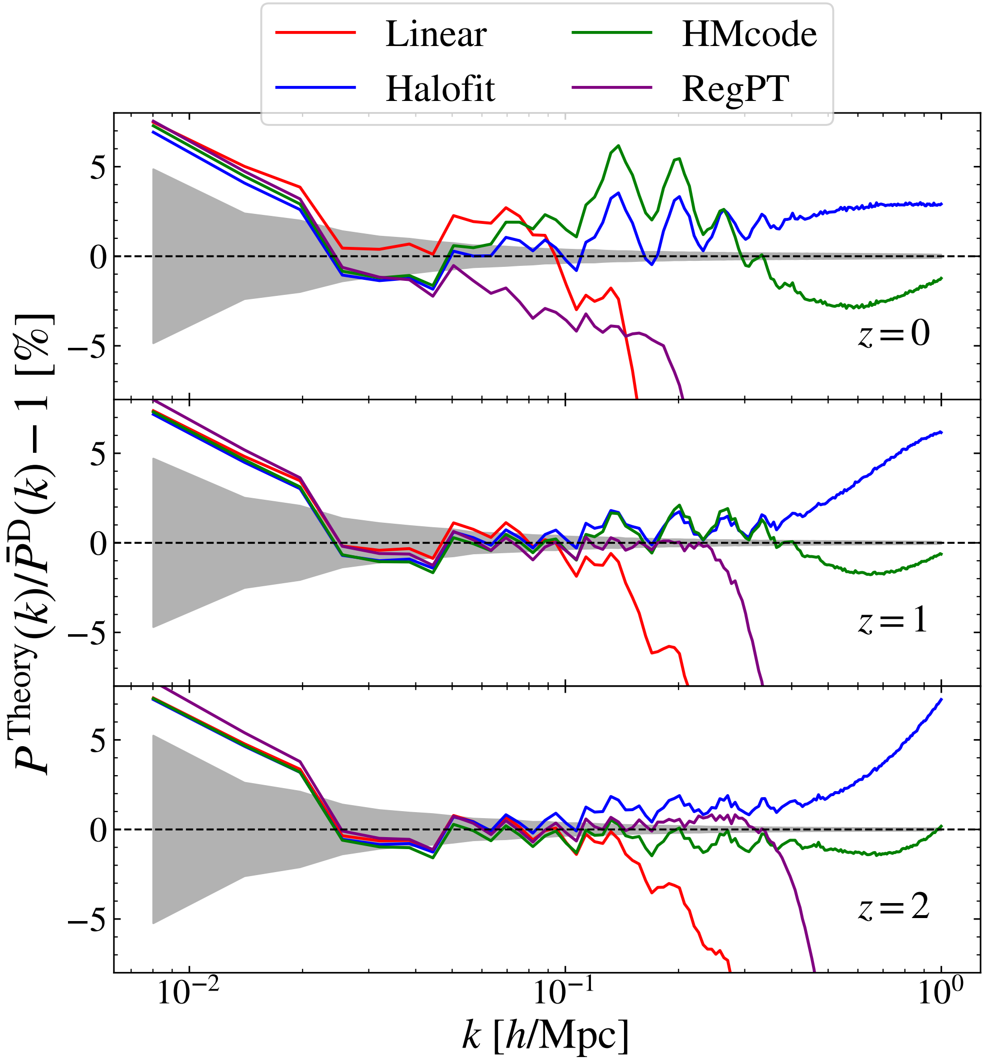
[Castorina et al. 2015]

- Fitting functions

→ Calibration on N-Body simulations with massive neutrinos

Testing non-linear models with massive neutrinos

Mean DEMNUni-Cov on 50 realisations *VS* Non-linear model



Non-linear prescriptions for the power spectrum :

- **Halofit :** Empirical fitting functions calibrated on simulations
  [Smith et al. 2003, Takahashi et al. 2012]
  Calibrated with massive neutrinos [Bird et al. 2011].

- **HMcode :** Physical fitting functions calibrated on simulations
  [Mead et al. 2015]
  Calibrated with massive neutrinos [Mead et al. 2016].

- **RegPT :** Analytic prediction with Regularised Perturbation Theory
  [Taruya et al. 2012]
  Account for massive neutrinos at the linear level [Castorina et al. 2015]

CLASS v2.9 [Blas et al. 2011]

Parameter estimation with MCMC :

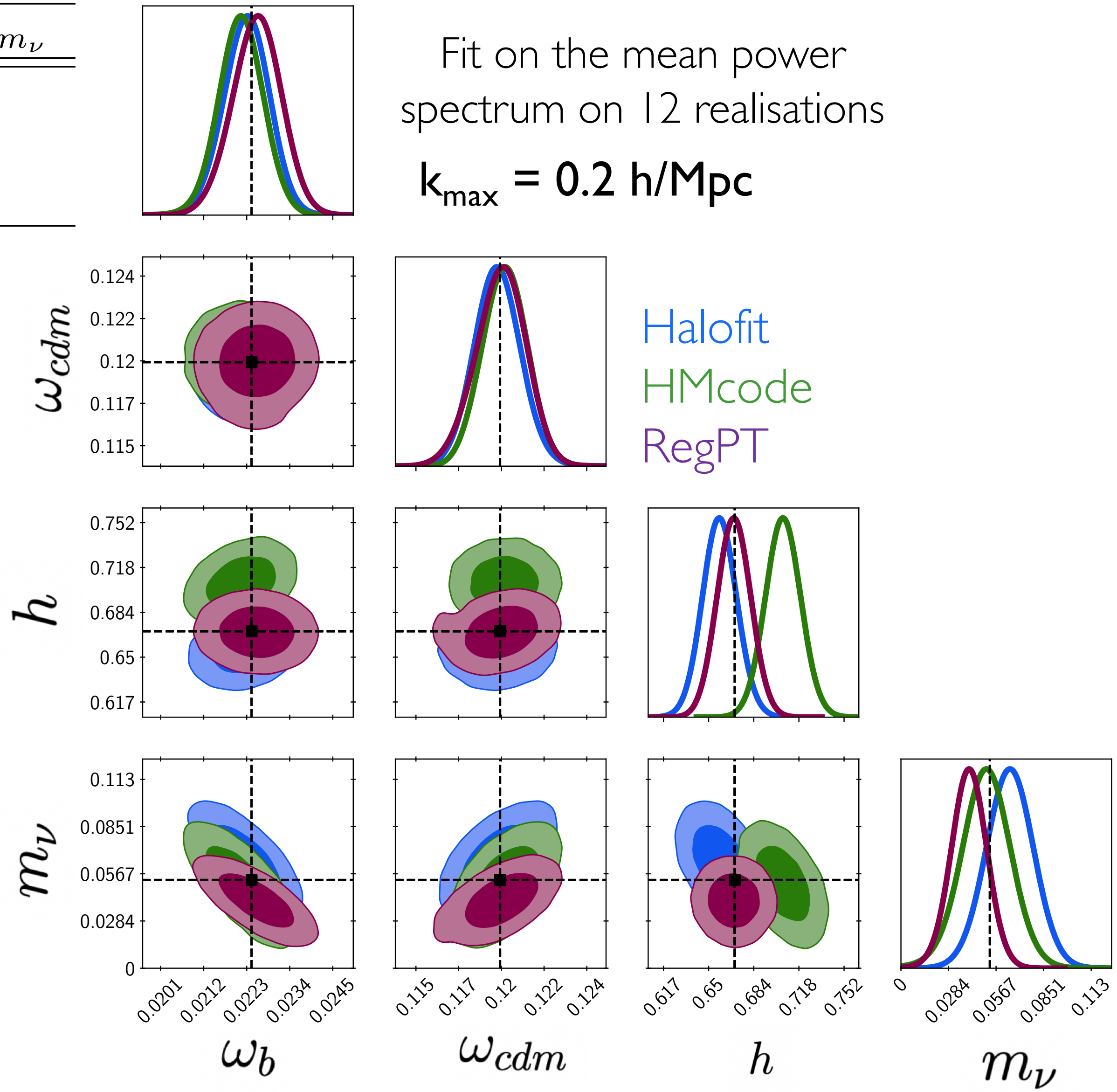
- Data-vector : **Mean DEMNUni-Cov on 12 realisations**
- 4 redshifts : $z \approx 0.5,\ 1,\ 1.5,\ 2$
- Covariance **:** Covmos Nm = 10 000
- k_{max} = [0.1, 0.3] h/Mpc
- **Non-linear models :**
  - Halofit
  - HMCODE
  - RegPT

Model	$\delta(\omega_b)/\sigma_{\omega_b}$	$\delta(\omega_{cdm})/\sigma_{\omega_{cdm}}$	$\delta(h)/\sigma_h$	$\delta(m_\nu)/\sigma_{m_\nu}$
Halofit	-0.28	-0.21	-0.90	0.86
HMcode	-0.46	0.19	2.61	-0.10
RegPT	0.24	0.05	-0.08	-1.20

- ω_b and ω_{cdm} : Unbiased for all models up to kmax = 0.22 h/Mpc
- Significant bias on h and m_ν:
  - Halofit : ~ 1σ bias on h and m_ν
  - HMcode : Unbiased estimation of m_ν, but strong bias on h
  - RegPT : Lower error-bar on m_ν ➔ ~1σ bias

Fit on the mean power spectrum on 12 realisations

k_{max} = 0.2 h/Mpc

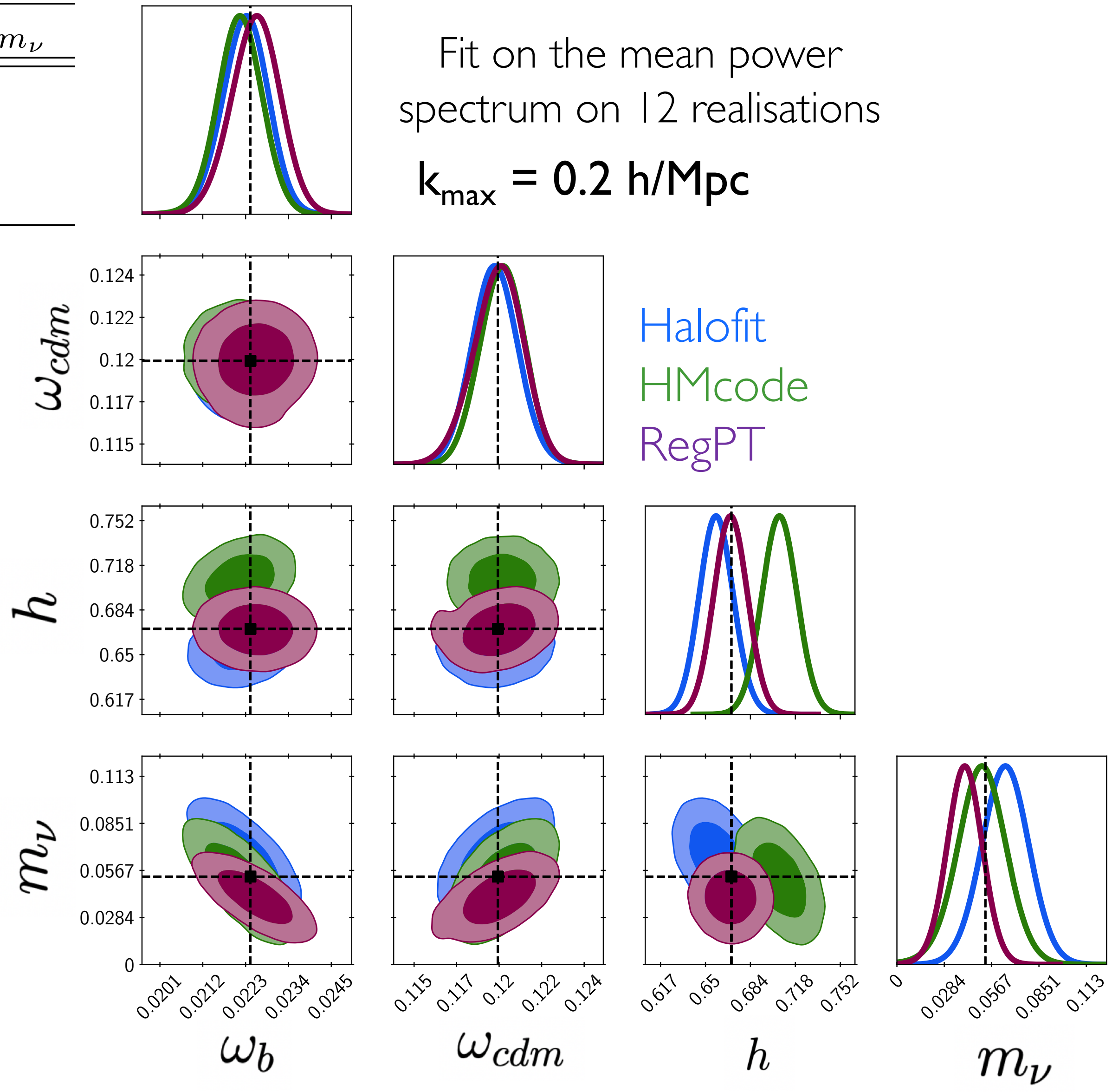


Model	$\delta(\omega_{\rm b})/\sigma_{\omega_{\rm b}}$	$\delta(\omega_{\rm cdm})/\sigma_{\omega_{\rm cdm}}$	$\delta(h)/\sigma_h$	$\delta(m_\nu)/\sigma_{m_\nu}$
Halofit	-0.28	-0.21	-0.90	0.86
HMcode	-0.46	0.19	2.61	-0.10
RegPT	0.24	0.05	-0.08	-1.20

- m_ν and h are correlated
  - Both are biased with Halofit
  - Surprising that only h is biased with HMcode
  - ➔ Test with m_ν fixed
- Can come from the treatment of m_ν in :
  - The fitting functions (New version of HMcode [Mead et al. 2020])
  - The N-Body simulations used for the calibration
- With RegPT treatement of m_ν is at the linear level
  - Explain the different correlation and error for m_ν ?
  - Explain the underestimation of m_ν ?

Fit on the mean power spectrum on 12 realisations

k_{max} = 0.2 h/Mpc



Non-linear model comparison : Only for real space matter power spectrum

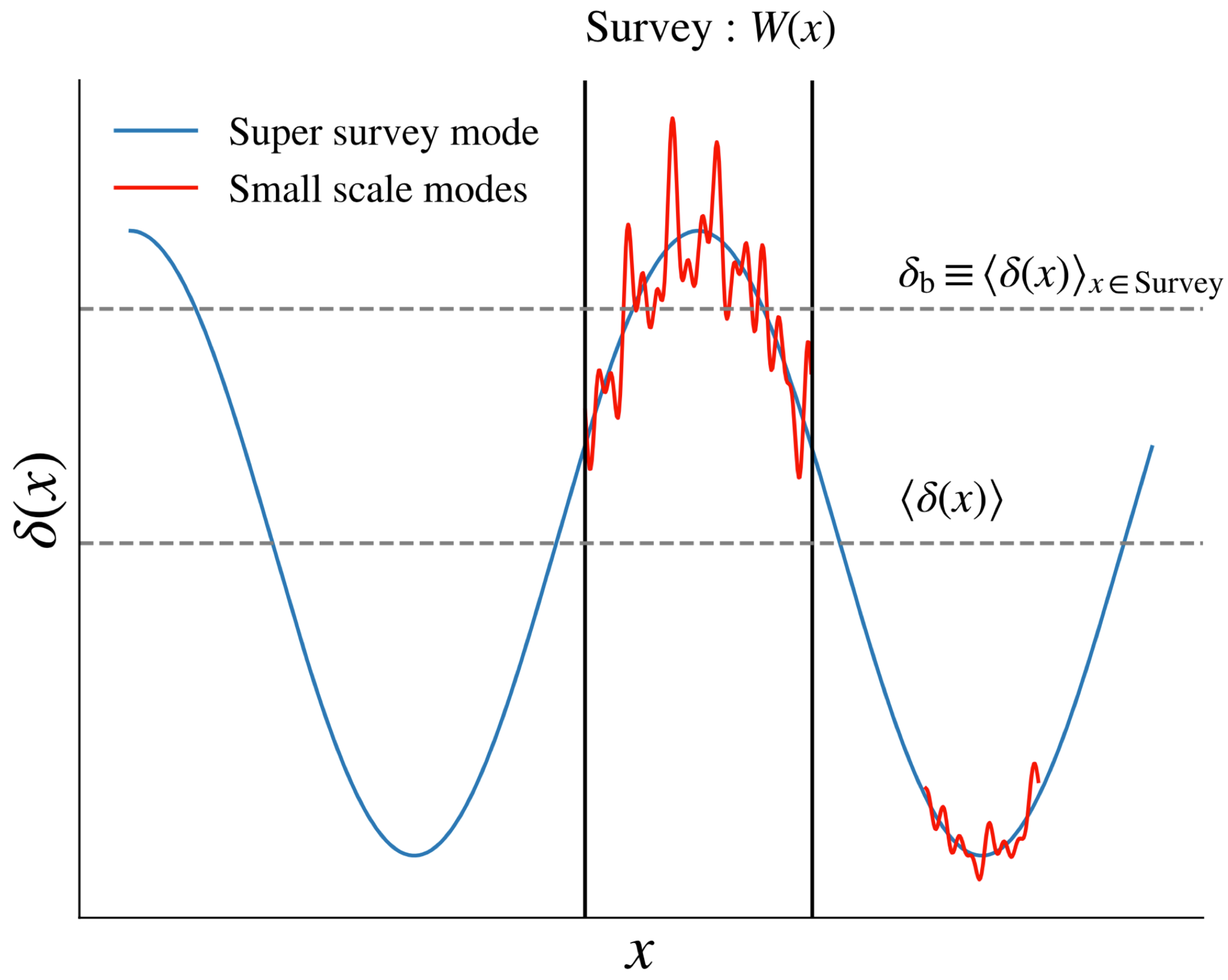
- Limited to kmax = 0.22 h/Mpc
- Revealed biases : especially for h and m_ν
  - ➔ Different hypothesis to explore to improve the treatment of massive neutrinos for non-linear model

Covariance : Sampling noise and non-Gaussian covariance

- Conclusion on NERCOME : **Handy when N_m is low but conservative in the errors.**
- Conclusion on Covmos : **Suppresion of sampling noise but conservative in the errors**
- Non-Gaussian covariance :
  - Significant impact for kmax > 0.15 h/Mpc
  - The error on m_ν is unchanged

IV. Impact of Super Sample Covariance on future photometric galaxy surveys

- Basics of SSC
- Impact of SSC on Euclid cosmological constraints
- Survey sky coverage and SSC



If we sit in an over-(under-)density, the clustering is enhanced (damped)

➔ Non-linear coupling between large scales and small scales

$$\sigma^2(z_1, z_2) = \langle \delta_b(z_1) \delta_b(z_2) \rangle = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \tilde{W}(\mathbf{k}, z_1)\, \tilde{W}^*(\mathbf{k}, z_2)\, P_m(k|z_{12})\,.$$

Where $\tilde{W}$ is the survey window function.

Important source of errors especially for photometric probes

[Barreira et al. 2018, Lacasa 2020]

Forecast of the impact of SSC on Euclid photometric probes :

Photometric Galaxy Clustering (GCph),
Weak Lensing (WL)
and Cross-Correlation (XC)

SSC Depend on the survey window function

$$\sigma^2(z_1, z_2) = \langle \delta_b(z_1)\delta_b(z_2)\rangle = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \tilde{W}(\mathbf{k}, z_1)\, \tilde{W}^*(\mathbf{k}, z_2)\, P_m(k|z_{12})\,.$$

Study the impact of the survey footprint in SSC

The S_{ij} approximation

Observables (P(k), C_l^{GCph}, C_l^{WL} ...)

Response of the observables

Covariance of δ_b

$$\mathrm{Cov}_{\mathrm{SSC}}\left(O_1, O_2\right) = \iint \mathrm{d}V_1 \mathrm{d}V_2 \frac{\partial o_1}{\partial \delta_b}\left(z_1\right) \frac{\partial o_2}{\partial \delta_b}\left(z_2\right) \sigma^2\left(z_1, z_2\right).$$

Observables (P(k), C_l^{GCph}, C_l^{WL} …)

Response of the observables

Covariance of δ_b

$$\mathrm{Cov}_{\mathrm{SSC}}\left(O_1, O_2\right) = \iint \mathrm{d}V_1 \mathrm{d}V_2 \frac{\partial o_1}{\partial \delta_b}\left(z_1\right) \frac{\partial o_2}{\partial \delta_b}\left(z_2\right) \sigma^2\left(z_1, z_2\right).$$

[F. Lacasa & J. Grain 2019] :

In the approximation where the observable's response vary slowly inside the redshift bins

$$\mathrm{Cov}_{\mathrm{SSC}}(O_1, O_2) = R_1 \; O_1 \; R_2 \; O_2 \; \boxed{S_{i_z, j_z}}$$

R is the relative response : $R = \dfrac{\partial \ln O}{\partial \delta_b}$

SSC for the angular power spectrum

The angular power spectrum : $C_\ell^{AB}(i_z, j_z) = \int \mathrm{d}V W_{i_z}^A(z) W_{j_z}^B(z) P_{AB}(k_\ell|z)$ With $k_\ell = (\ell + 1/2)/r(z)$

Projection of the 3D power spectrum to the harmonic space with a kernel W(z), characteristic of the probe considered

- A and B are the probes : Photometric Galaxy Clustering (GCph) or Weak Lensing (WL)
- i_z and j_z are the redshift bins

$$\mathrm{Cov}_{\mathrm{SSC}}\left(C_\ell^{AB}\left(i_z, j_z\right), C_{\ell'}^{CD}\left(k_z, l_z\right)\right) \approx R_\ell^{AB} C_\ell^{AB}\left(i_z, j_z\right) \times R_{\ell'}^{CD} C_{\ell'}^{CD}\left(k_z, l_z\right) \times S_{i_z, j_z; k_z, l_z}^{A,B;C,D}.$$

Ingredients :

- The C(l)
- The non-linear response R_l
- The S_{ijkl} matrix

The C(l) $C_\ell^{AB}(i_z, j_z) = \int \mathrm{d}V W_{i_z}^A(z) W_{j_z}^B(z) P_{AB}(k_\ell|z)$

- Photometric Galaxy Clustering (GCph)

$$W_i^{\rm GC}(z) = b_i \frac{n_i(z)}{r^2(z)} \frac{H(z)}{c}.$$

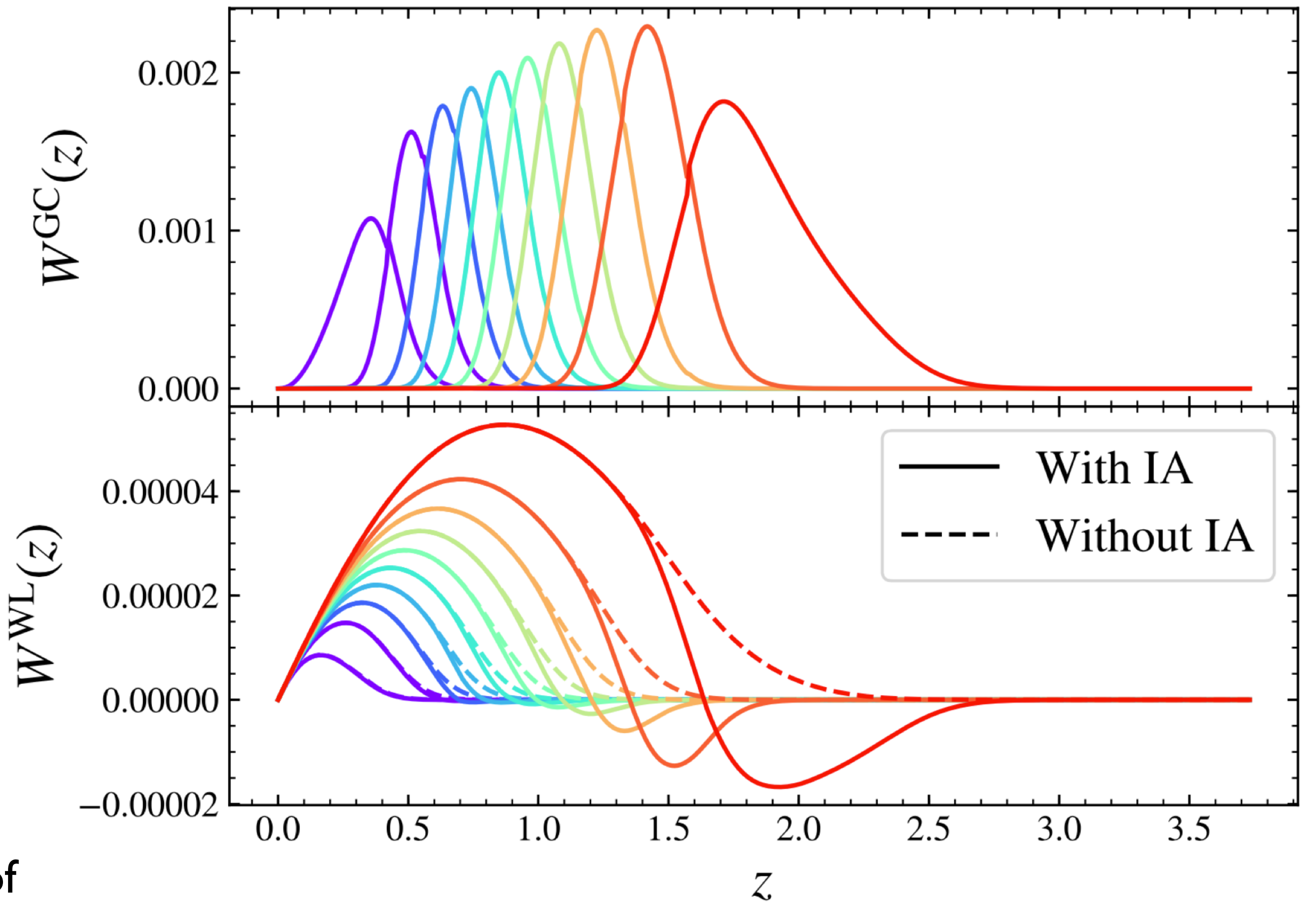
- Weak Lensing (WL)

$$W^{\rm WL}(z) = W^\gamma(z) + W^{\rm IA}(z).$$

Cosmic shear ← $W^\gamma(z)$

$W^{\rm IA}(z)$ → Intrinsic Alignment of galaxies

Kernels for 10 *Euclid-like* photometric redshift bins



The non-linear response R_l

Following [Lacasa & Grain 2019] ➔ Approximation of a constant response : $R_\ell \sim 4$

The S_{ijkl} matrix

C(l) kernels

Covariance of δ_b

$$S^{A,B;C,D}_{i_z,j_z;k_z,l_z} \equiv \int \mathrm{d}V_1 \mathrm{d}V_2 \, \frac{W^A_{i_z}(z_1)\, W^B_{j_z}(z_1)}{I^{AB}(i_z,j_z)} \frac{W^C_{k_z}(z_2)\, W^D_{l_z}(z_2)}{I^{CD}(k_z,l_z)} \sigma^2(z_1,z_2)$$

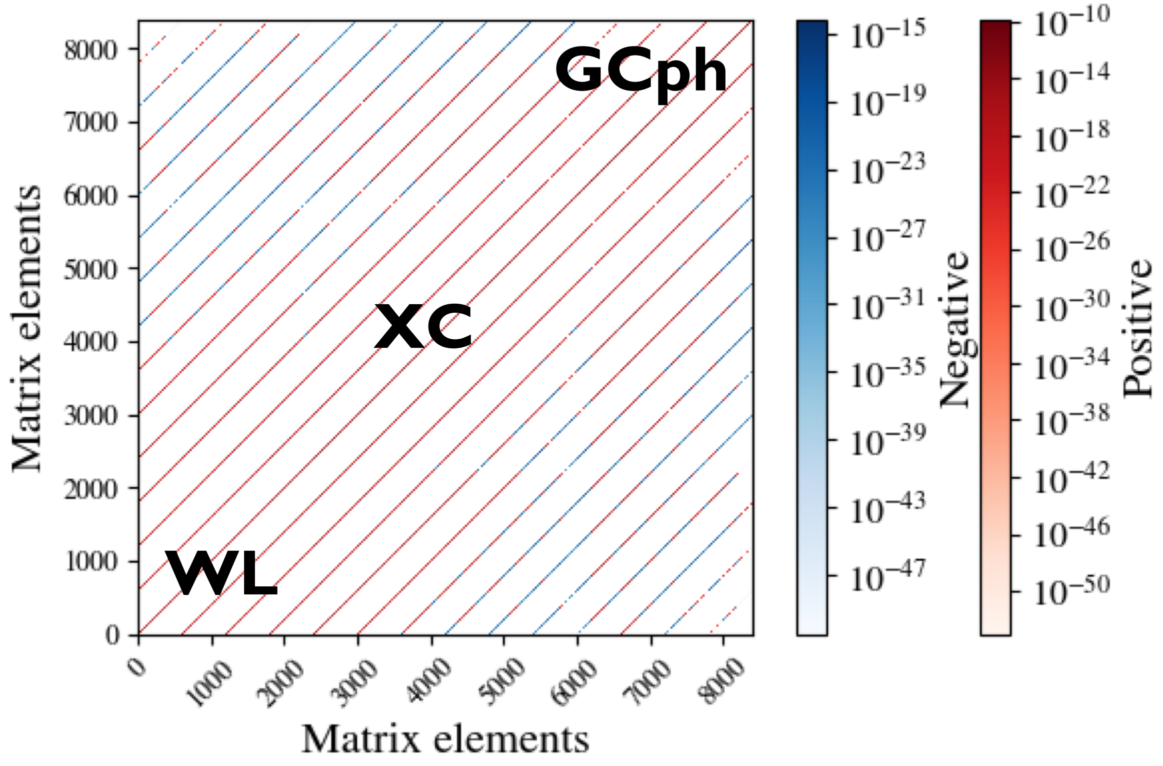
Survey area

Full-sky approximation : $\sigma^2_{\mathrm{full-sky}}(z_1,z_2) = \frac{1}{4\pi} C^{\mathrm{m}}_\ell(z_1,z_2)$ ➔ Rescale by the sky fraction $f_{\mathrm{SKY}} \equiv \Omega_S/4\pi$

Computed with PySSC
[Lacasa & Grain 2019]
https://github.com/fabienlacasa/PySSC

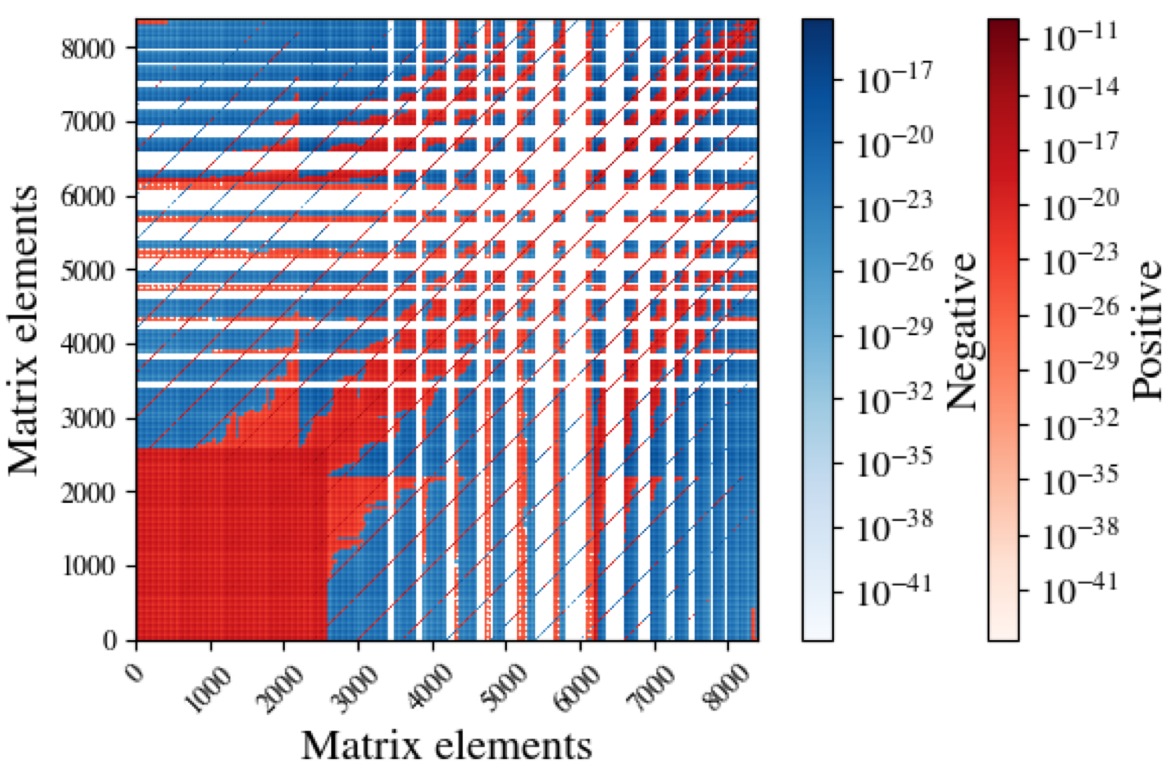
For 3x2-points : GCph+WL+XC, with 10 redshift bins

Gaussian



Diagonal in ℓ

Gaussian + SSC



Non-diagonal

Fisher forecast for Euclid

[Euclid collaboration: Blanchard at al. 2020]

$$F_{\alpha,\beta} = \sum_{i,j,k,l} \sum_{\ell,\ell'} \frac{\partial C_{ij}^{AB}(\ell)}{\partial\theta_\alpha} \mathrm{Cov}\left[C_{ij}^{AB}(\ell), C_{kl}^{CD}(\ell')\right]^{-1} \frac{C_{kl}^{CD}(\ell')}{\partial\theta_\beta},$$

- Derivatives : Theoretical C(l) for GCph + WL + XC
- 10 redshift bins :

$$z_i = \{0.0010, 0.42, 0.56, 0.68, 0.79, 0.90, 1.02, 1.15, 1.32, 1.58, 2.50\}.$$

- Multipoles : $n_\ell = 40$

$$\ell_{\min} = 10$$

$$\ell_{\max}^{GCph} = \ell_{\max}^{XC} = 3000 \qquad \ell_{\max}^{WL} = 5000$$

- Covariance :
  - Gaussian
  - Gaussian + SSC

  } *Full-sky approximation*

- Cosmological parameters : **$w_0 w_a$CDM scenario**

$$\boldsymbol{\theta}_{\rm cosmo} = \{\Omega_{\rm m}, \Omega_{\rm b}, w_0, w_a, h, n_s, \sigma_8\} = \{0.32, 0.05, -1.0, 0.0, 0.67, 0.96, 0.816\},$$

- Nuisance parameters :
  - Linear galaxy bias : $\boldsymbol{\theta}_{\rm nuisance}^{\rm GC} = \{b_i\}$, for i in $1, ..., 10$,
  - Instrinsic alignments : $\boldsymbol{\theta}_{\rm nuisance}^{\rm WL} = \{\mathcal{A}_{\rm IA}, \eta_{\rm IA}, \beta_{\rm IA}\}$,

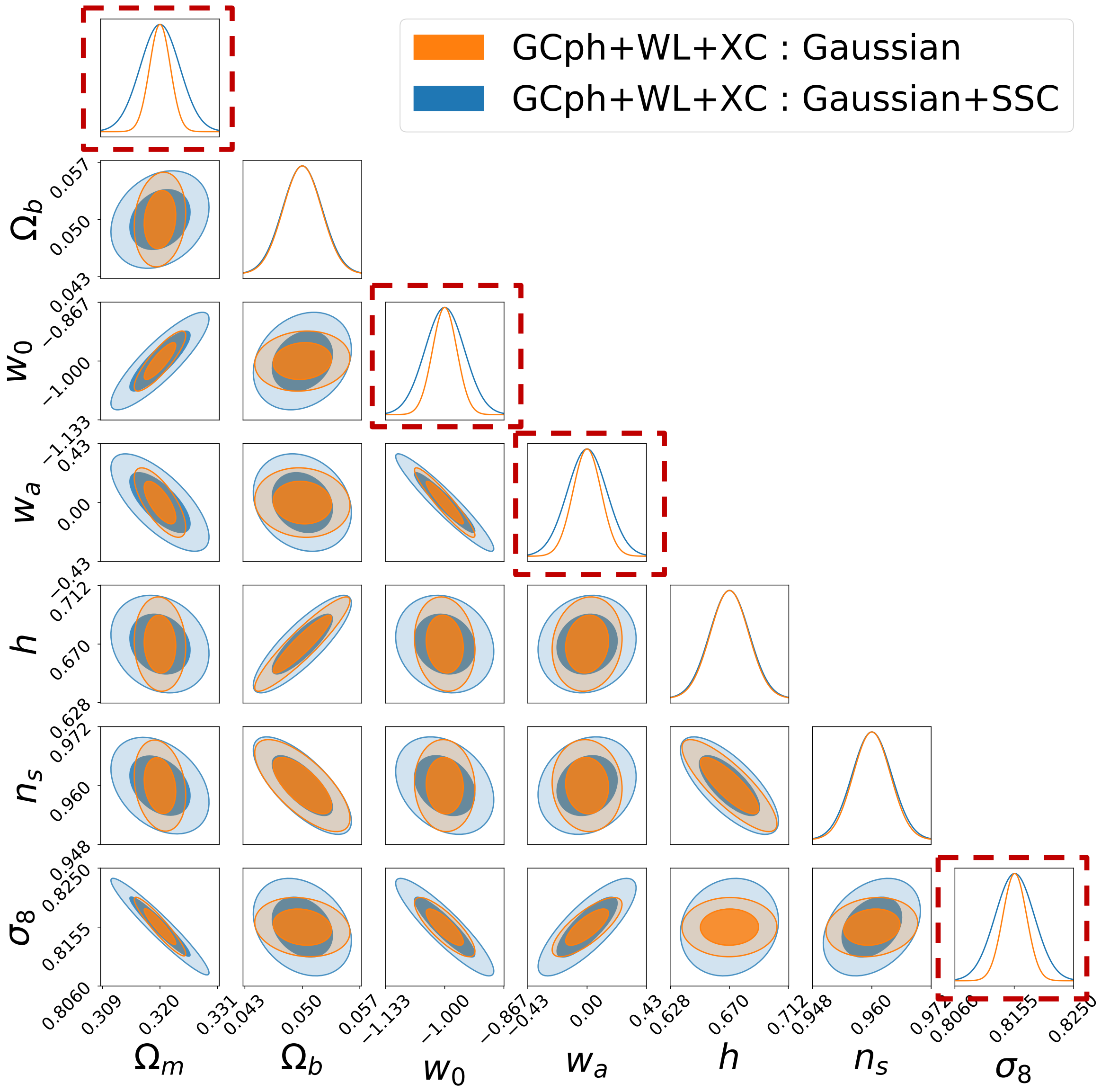
Dark Energy Figure of Merit : $\mathrm{FoM}_{w_0,w_a} = \sqrt{\det(F_{w_0,w_a})}.$

Lower = Worse constraints

Probe	Gaussian	Gaussian+SSC	Decrease due to SSC [%]
GCph+WL+XC	1038.13	454.60	−56.21

- 50% decrease of the Dark Energy FoM
- Ω_m, w_0, w_a and σ_8 are the most impacted parameters

➔ related to the amplitude of the power spectrum



SSC arises from the fact that we observe a limited portion of the universe

What is the impact of the survey footprint ?

For a survey with a mask $\mathcal{M}$ the exact derivation is :

$$\sigma^2(z_1, z_2) = \frac{1}{\Omega_S^2} \sum_\ell (2\ell+1) C_\ell(\mathcal{M}) C_\ell^m(z_1, z_2) \, ,$$

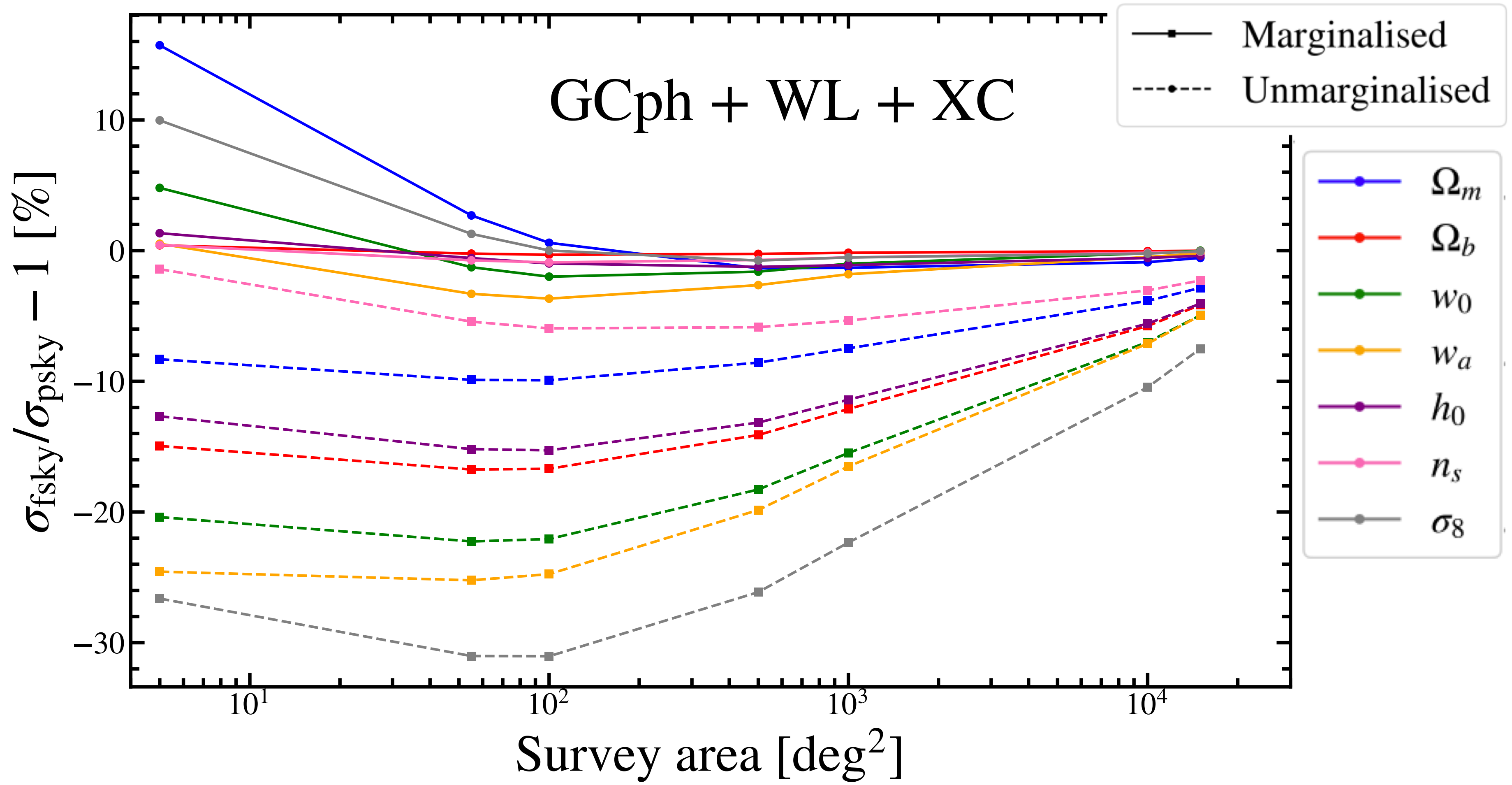
Full-sky *VS* Partial-sky

Compare the different treatments of the footprint, at the level of parameters → For a simple geometry with varying area

Same forecast setting

Circular mask with varying area

Relative difference between the error on cosmological parameters obtained with full-sky and partial-sky



Circular mask with varying area

- **Marginalised error**

$$\sigma_\alpha = \sqrt{(F^{-1})_{\alpha,\alpha}}\,.$$

➔ Small difference for wide surveys

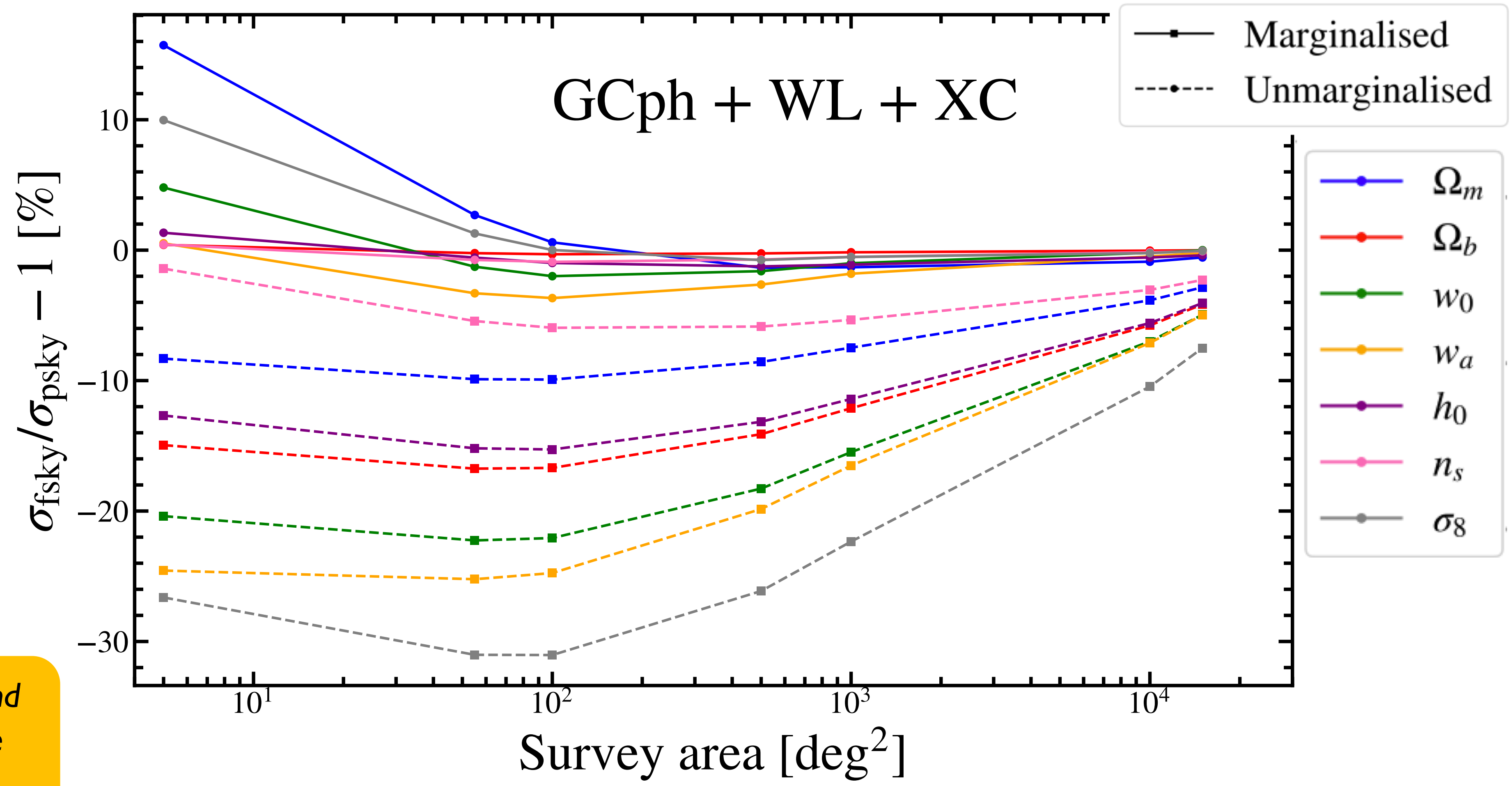
- **Unmarginalised error**

$$\sigma^U_\alpha = \sqrt{1/F_{\alpha,\alpha}}\,,$$

➔ Significant difference

Difference between full-sky and partial-sky is absorbed in the marginalisation

Relative difference between the error on cosmological parameters obtained with full-sky and partial-sky

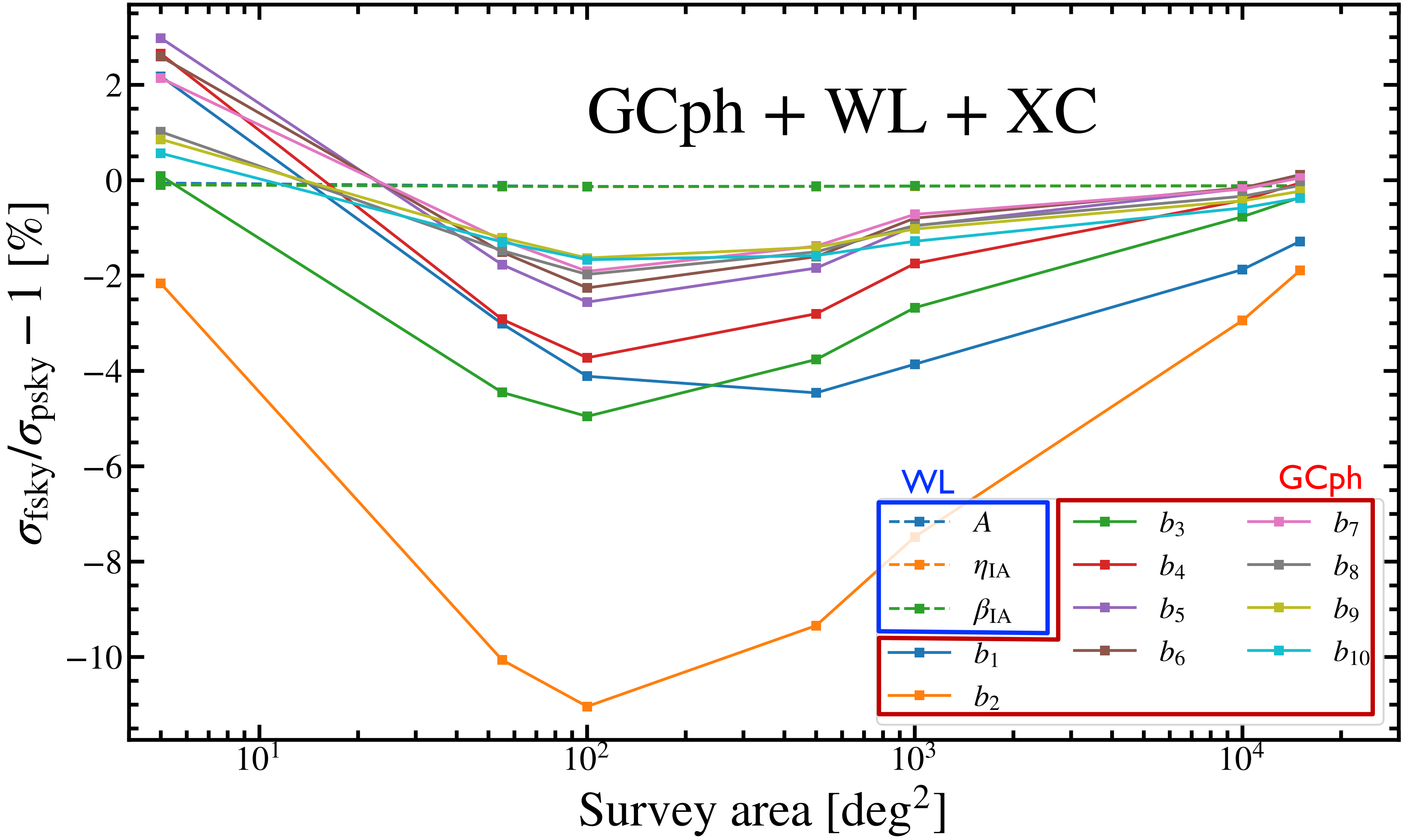


Circular mask with varying area

The linear galaxy bias is more sensitive to SSC

The difference is absorbed by the galaxy bias in the marginalisation.

Relative difference between the marginalised error on nuisance parameters obtained with full-sky and partial-sky



- Super Sample Covariance has a significant impact on Dark Energy constraints ➔ **-50% on the FoM**

  Euclid Key paper : **D.Sciotti, S. Gouyou Beauchamps et al. In prep.**

- Treatment of the survey footprint for SSC : Accepted paper **[Gouyou Beauchamps et al. 2021]** in A&A

  - Development of a new method to account for the survey footprint :

  Soon available at https://github.com/fabienlacasa/PySSC

  - For marginalised constraints : Full-sky is sufficient for area > 1000 deg^2

  - For unmarginalised constraints : More than 10% difference between full-sky and partial-sky for all areas between 5 and 15 000 deg^2

    ➔ *Might be important if tight priors are applied on nuisance parameters*

Conclusions

With Euclid we will reach a high precision on LSS observables

To exploit cosmological information on small scales and find potential deviations from LCDM and accurately measure the total neutrino mass, we need to control :

- The covariance :
  - Effects of sampling noise can be reduced with NERCOME and Covmos
  - Non-Gaussian covariance :
    - Trispectrum (Non-linear clustering) : MCMC on N-Body simulations with matter P(k)

      ➔ Significant impact of Fourier mode correlations
    - SSC (Survey limited observation) : Theoretical Fisher forecast for 3x2-points

      ➔ Significant impact on Dark Energy constraints

  Which one is the most important ?
- The modeling of non-linear scales ➔ Significant bias on the estimation of the neutrino mass

Conclusions

The inference of cosmological parameters involves

- The covariance :
  - Effects of sampling noise can be reduced with NERCOME and Covmos
  - Non-Gaussian covariance :
    - Trispectrum (Non-linear clustering) : MCMC on N-Body simulations with matter P(k)

    ➔ Significant impact of Fourier mode correlations

    - SSC (Survey limited observation) : Theoretical Fisher forecast for 3x2-points

    ➔ Significant impact on Dark Energy constraints

    } Which one is the most important ?
- The modeling of non-linear scales ➔ Significant bias on the estimation of the neutrino mass

With Euclid we will reach a high precision on LSS observables

To exploit cosmological information on small scales to find potential deviations from LCDM and accurately measure the total neutrino mass, we need to control these effects !

Perspectives

Estimation of the neutrino mass with the power spectrum

- Need more investigation for the non-linear modeling challenge
- Extend to galaxies in redshift space
- Extend to non-standard cosmologies

Super Sample Covariance

- Go beyond the constant response approximation
- Estimate the response in N-body simulations

Question the usual assumptions on the Likelihood of 2-point statistics

- Gaussian distribution of data ?
- Cosmological dependence of the covariance

➔ Additional bias in parameter inference ?

Euclid Key Papers in progress

We have to be ready for the first Euclid data : 2023 - 2024

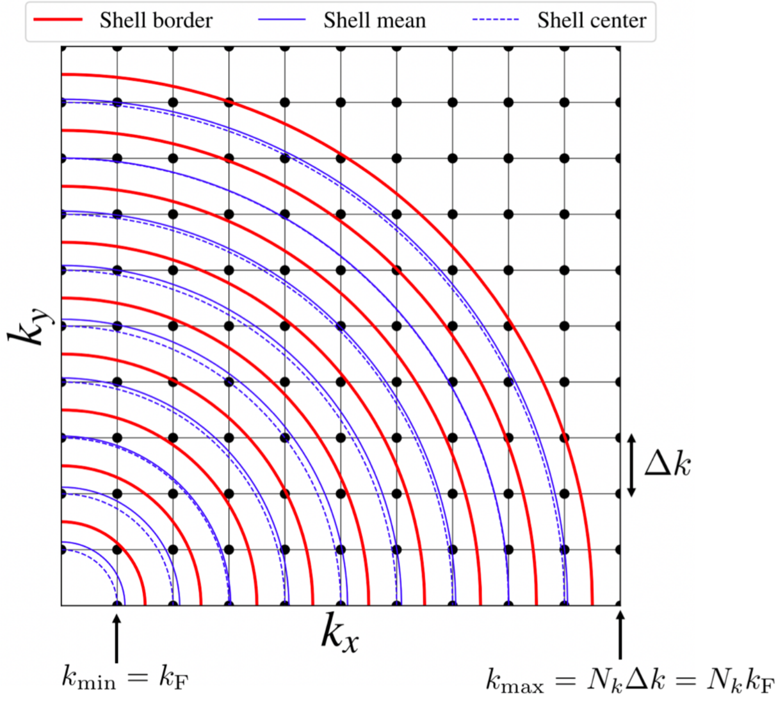
Backup

Estimate the density field on a grid and take the Fourier Transform

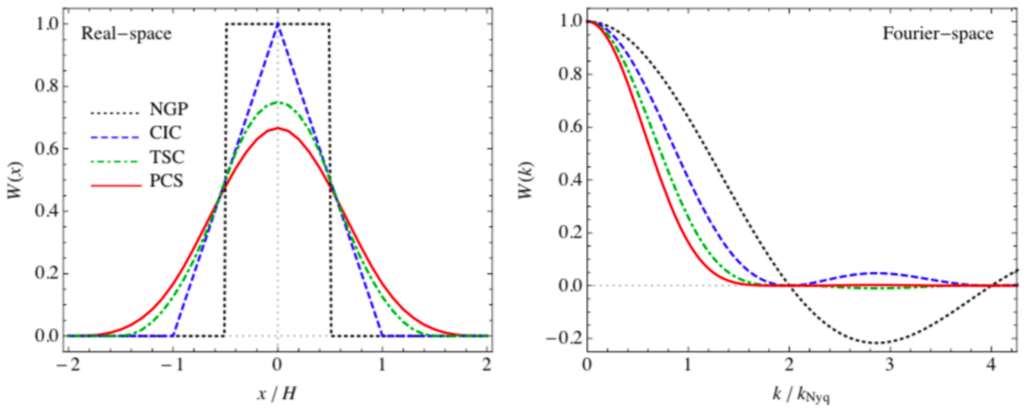
$$\hat{P}(\mathbf{k}_n) \equiv k_{\rm F}^3 \left[|\delta(\mathbf{k}_n)|^2 - \frac{1}{N_p} \right],$$

Shell average of the 3D P(k)

$$\hat{P}(k) = \frac{1}{M_k} \sum_{k<|\mathbf{k}|<k+\Delta k} \hat{P}(\mathbf{k}),$$

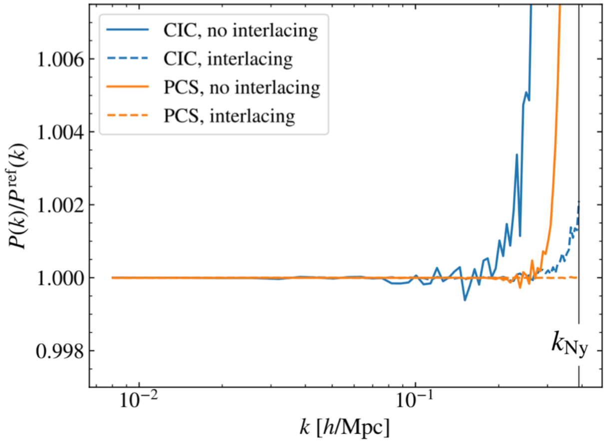


Mass Assignement Scheme

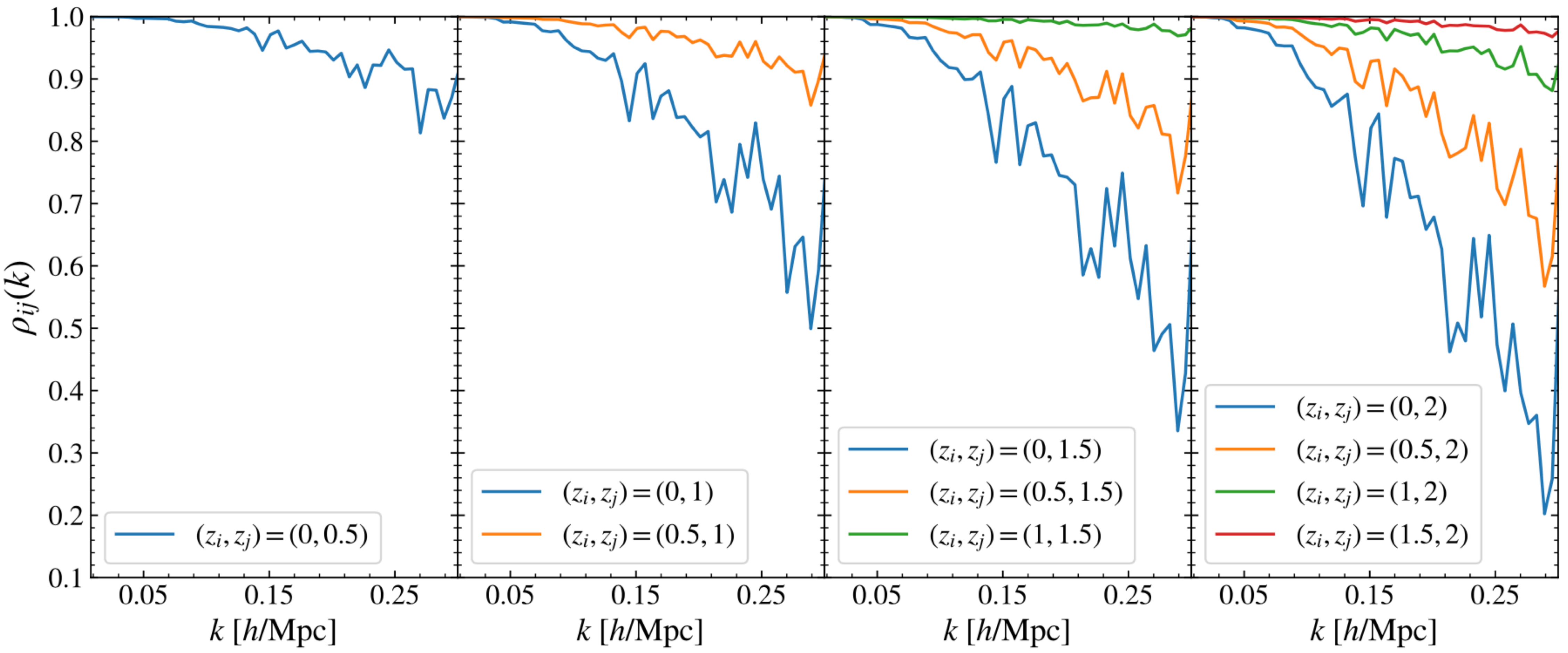


Aliasing and Interlacing

$$P^{\rm G}(\mathbf{k}) = \sum_{\mathbf{n}} |W(\mathbf{k}-\mathbf{n}k_s)|^2 \left[P(|\mathbf{k}-\mathbf{n}k_s|) + \frac{1}{(2\pi)^3 \bar{n}} \right].$$

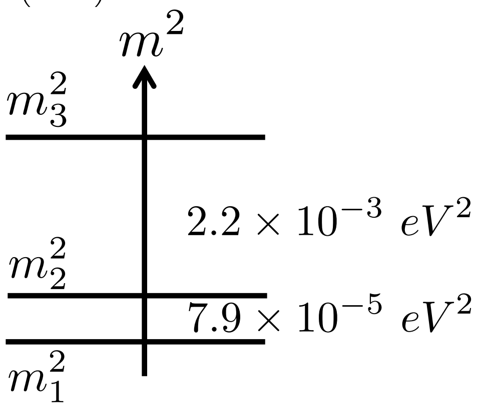


Redshift correlations



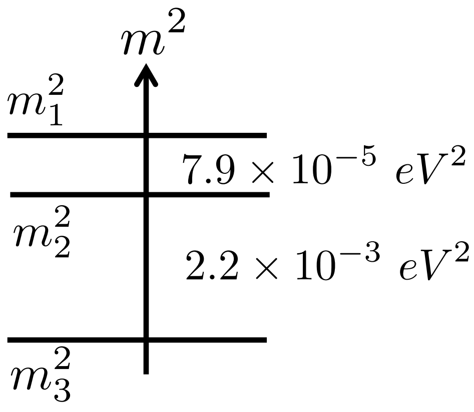
- Neutrino oscillations experiments : Super Kamiokande, T2K, KamLand…

JHEP 09 (2020) 178



Normal

$$\sum m_\nu > 0.056\ eV$$



Inverted

$$\sum m_\nu > 0.095\ eV$$

Overall particle physics experiments constraints

NH	$0.056\ \text{eV} \lesssim M_\nu \lesssim 1\ \text{eV}$
IH	$0.095\ eV$

- Direct kinematic measurments in tritium β decay : KATRIN

Phys. Rev. Lett. 123, 221802

$$\sum m_\nu < 1\ \text{eV}$$

Neutrino mass constraints : Cosmology

Constraints with the CMB [Planck Collaboration 2020]

- **CMB *T&P*** (Planck) **:** $M_\nu < 0.26 \text{ eV } (95\% \text{ C.L.})$
- **CMB *T&P*** (Planck) **+ Lensing :** $M_\nu < 0.24 \text{ eV } (95\% \text{ C.L.})$
- **CMB *T&P*** (Planck) **+ Lensing + BAO :** $M_\nu < 0.12 \text{ eV } (95\% \text{ C.L.})$ ➔ BAO break degeneracies (Ω_m, H_0)

Constraints with the LSS

- **BAO** (BOSS) **+ CMB *T&P* :** $M_\nu < 0.12 \text{ eV } (95\% \text{ C.L.})$ [Ivanov et al. 2020]
- ***P(k)*** (BOSS) **+ CMB *T&P* :** $M_\nu < 0.16 \text{ eV } (95\% \text{ C.L.})$ [Ivanov et al. 2020]
- **Ly-α *1D P(k)* :** $M_\nu < 0.71 \text{ eV } (95\% \text{ C.L.})$ [Palanque-Delabrouille et al. 2020]
- **Ly-α *1D P(k)* + CMB *T&P* :** $M_\nu < 0.1 \text{ eV } (95\% \text{ C.L.})$ [Palanque-Delabrouille et al. 2020]

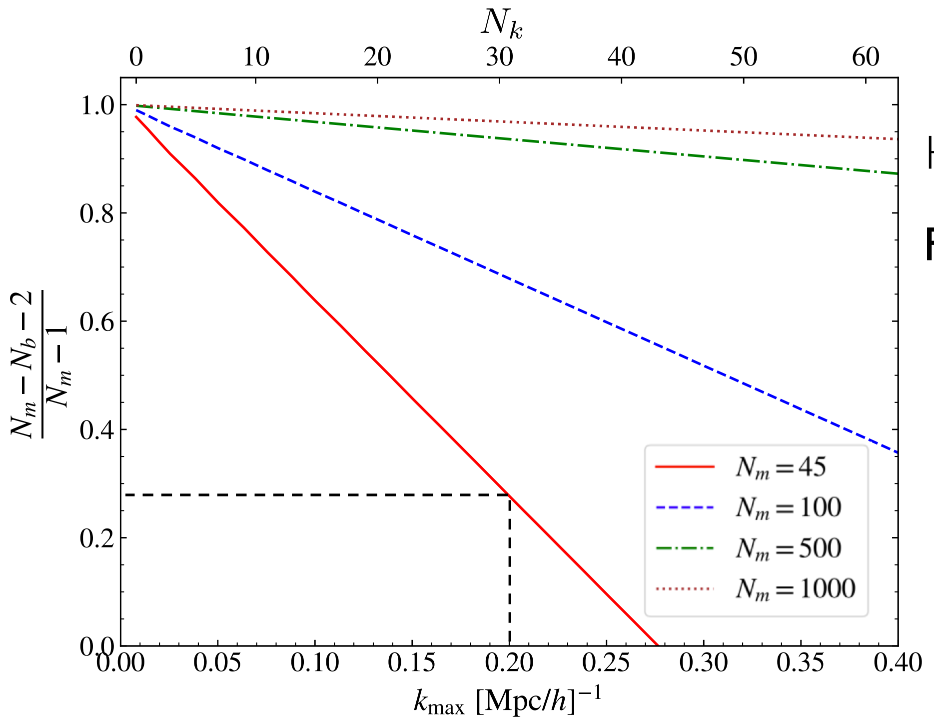
Covariance estimation and sampling noise I

For N_m realisation of the power spectrum $\hat{C}_{ij} = \frac{1}{N_m - 1}\left[\sum_{n}^{N_m}[P^{(n)}(k_i) - \bar{P}(k_i)][P^{(n)}(k_j) - \bar{P}(k_j)]\right]$ ➡ **Sampling noise**

➢ Hartlap bias in the **precision matrix** $\mathbf{\Psi}$ [Wishart 1928, Hartlap 2007]

$\hat{\mathbf{\Psi}} = \hat{\mathbf{C}}^{-1}$ Biased estimator of $\mathbf{\Psi}$: $\langle\hat{\mathbf{\Psi}}\rangle = \frac{N_m - 1}{N_m - N_k - 2}\mathbf{\Psi}$ ➡ *Sensitive to the number of simulations and bins* N_m *and* N_k

Inverse Wishart distribution



Hartlap factor to multiply to $\hat{\mathbf{\Psi}}$

For example for N_k = 30 (k_{max} ~ 0.2 h/Mpc) and N_m = 45 ➔ Divide Psi by 3

First effect of sampling noise :

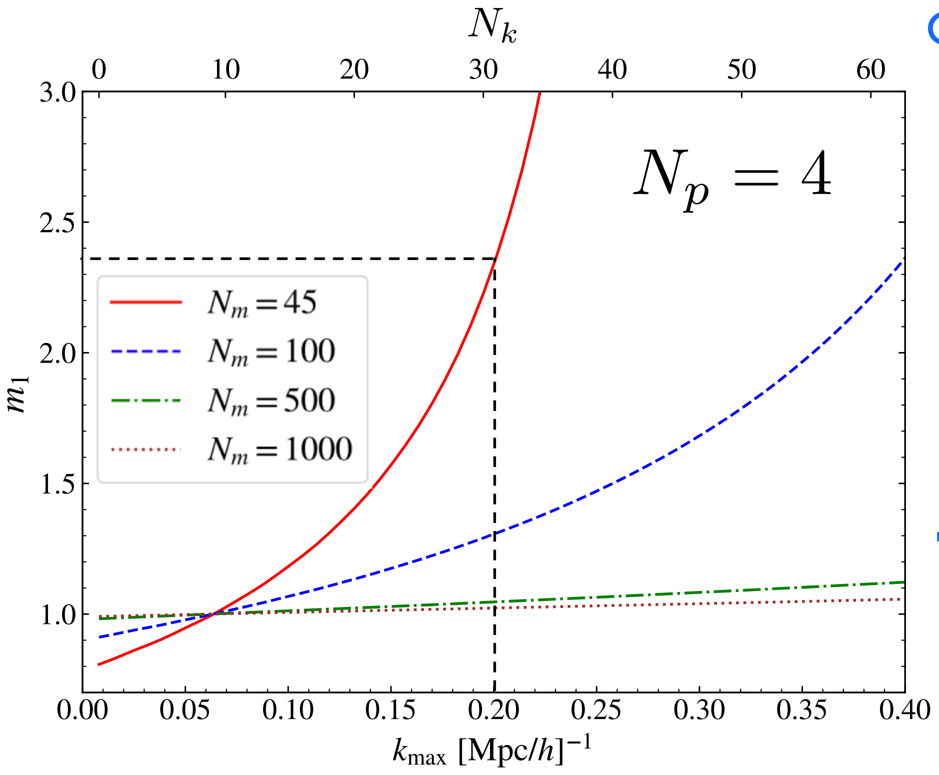
Overestimation of $\mathbf{\Psi}$ ➔ underestimation of parameter's errors

Covariance estimation and sampling noise II

Noise in the estimated precision matrix : $\hat{\boldsymbol{\Psi}} = \boldsymbol{\Psi} + \Delta\boldsymbol{\Psi}$ with $\langle \Delta\Psi_{ij}\Delta\Psi_{kl}\rangle = A(N_m, N_k)\Psi_{ij}\Psi_{kl} + B(N_m, N_k)(\Psi_{ik}\Psi_{jl} + \Psi_{il}\Psi_{jk})$

[Taylor et al. 2013]

- Sampling noise in the precision matrix propagates to the posterior distribution of parameters
[Taylor et al. 2013 ; Dodelson & Schneider 2013 ; Percival et al. 2014 ;Taylor & Joachimi 2014]



Correct by inflating the parameter's covariance matrix Φ

[Percival et al. 2014]

$$\hat{\Phi} \rightarrow m_1 \times \hat{\Phi}$$

$$m_1 = \frac{1 + B(N_b - N_p)}{1 + A + B(N_p + 1)}$$

Dispersion on the best-fit

Bias on the estimation of the error

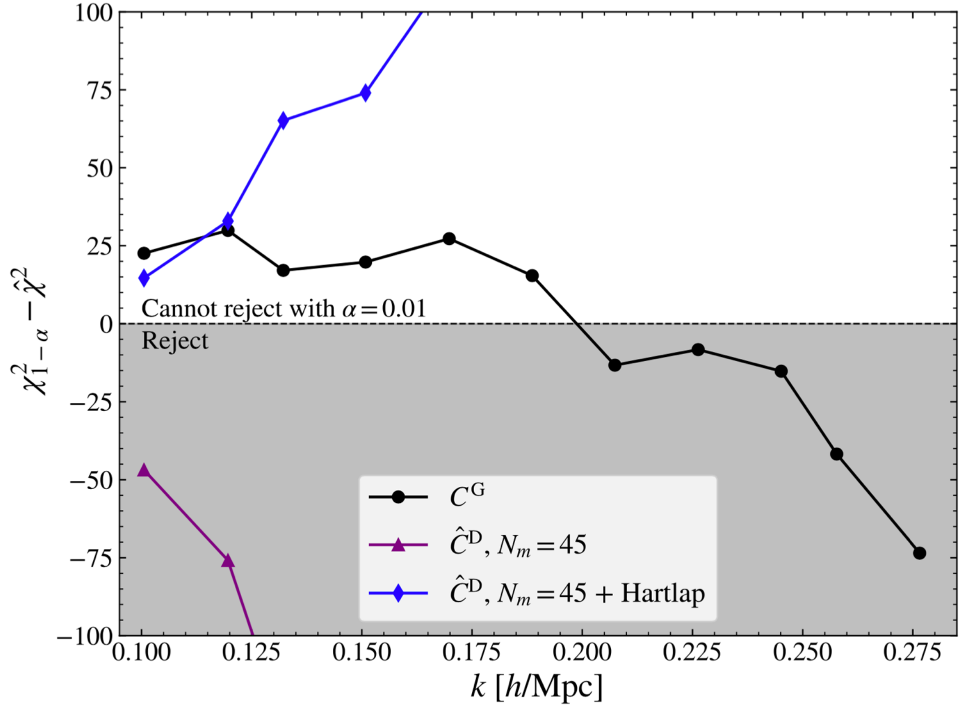
For example for
N_k = 30 (k_{max} ~ 0.2 h/Mpc)
and N_m = 45

➔ Inflate the error-bars by ~ 50%

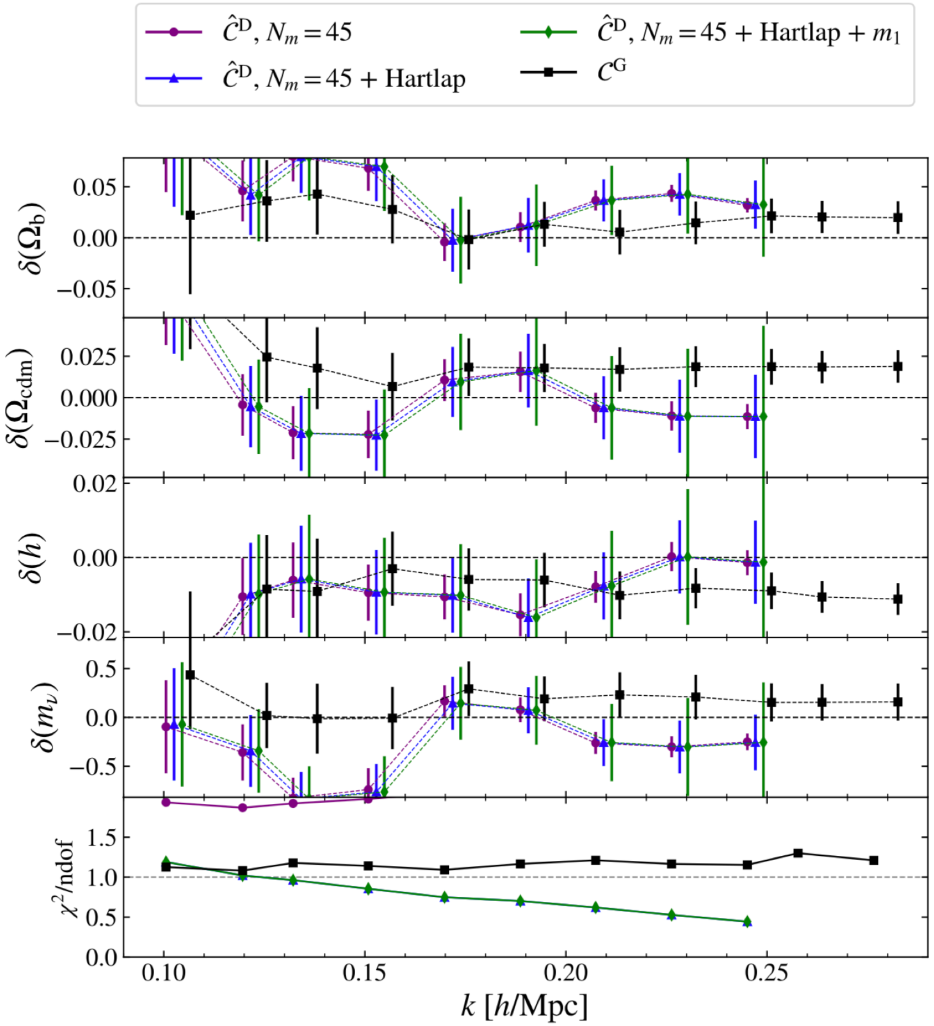
Second effect of sampling noise :

Variance and bias on the estimated variance and best-fit of parameters

Goodness of fit test



Parameter estimation



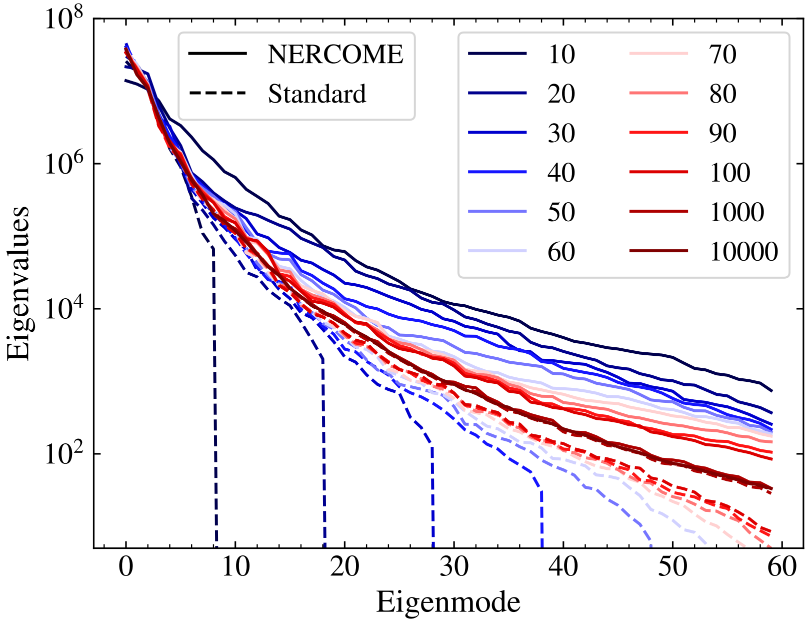
- Hartlap bias : $\langle \hat{\Psi} \rangle = \frac{N_m - 1}{N_m - N_k - 2} \Psi$

- Sampling noise : $\hat{\Phi} \rightarrow m_1 \times \hat{\Phi}$ $\quad m_1 = \frac{1 + B(N_b - N_p)}{1 + A + B(N_p + 1)}$

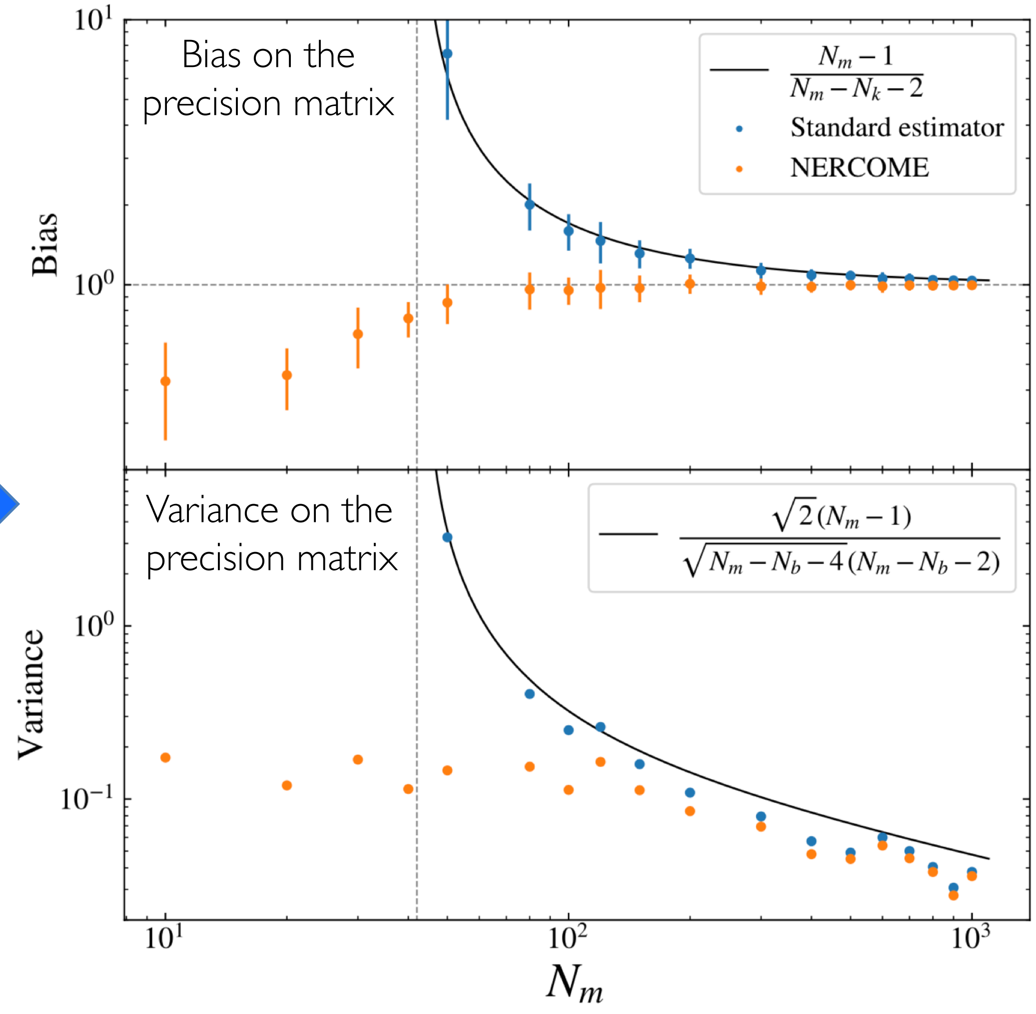
For N_m mocks :

1. Estimate the covariance on 2 subsets of size s and N_m-s : $\hat{S}_1$ and $\hat{S}_2$
2. Eigenvalue Decomposition : $\hat{S}_i = U_i D_i U_i^T$
3. Estimate the new covariance as : $\hat{C}_N = U_1 \mathrm{diag}(U_1^T \hat{S}_2 U_1) U_1^T$

Repeat for 500 random subsets of size s and N_m-s

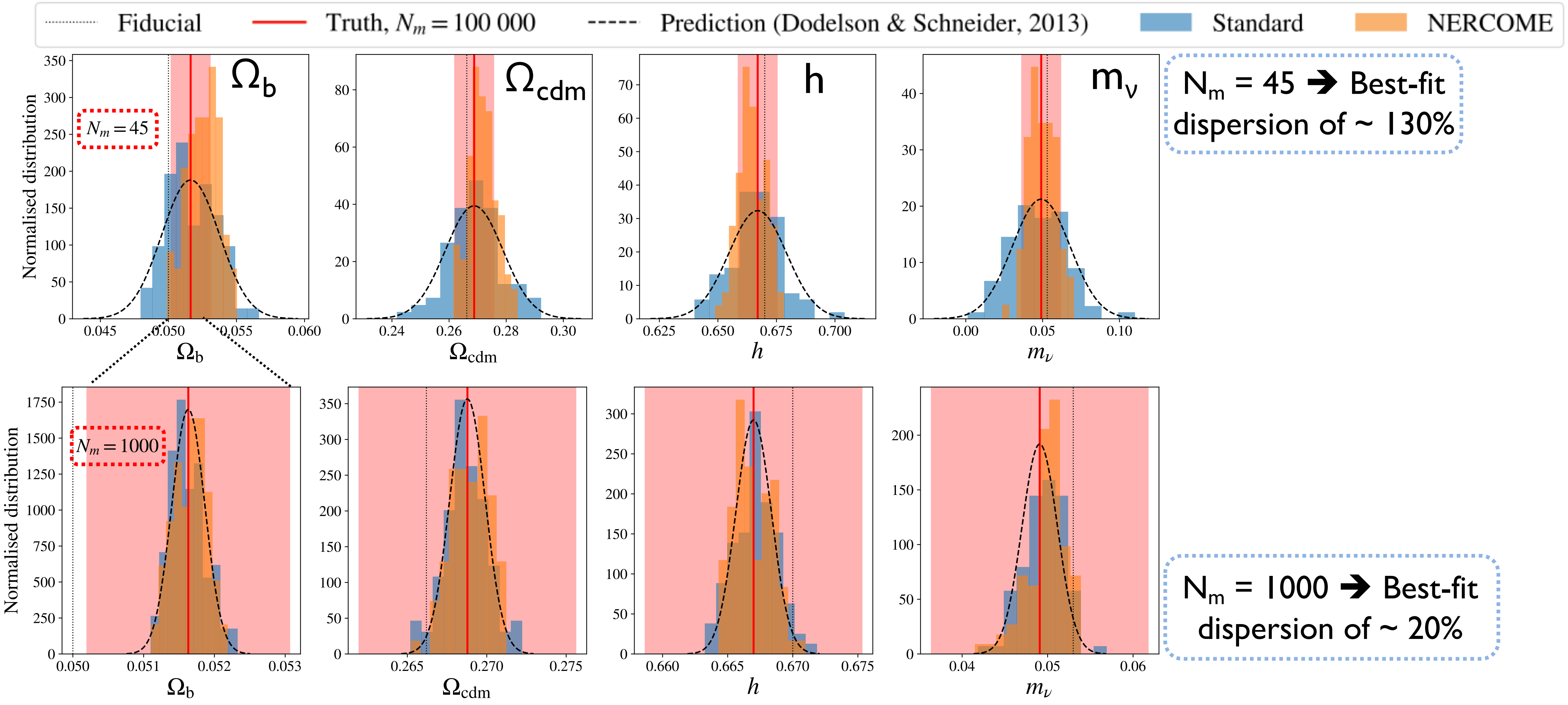


Non-linear shrinkage of low eigenvalues



Testing NERCOME on parameter estimation

Distribution of best-fits for 100 realisations of the covariance

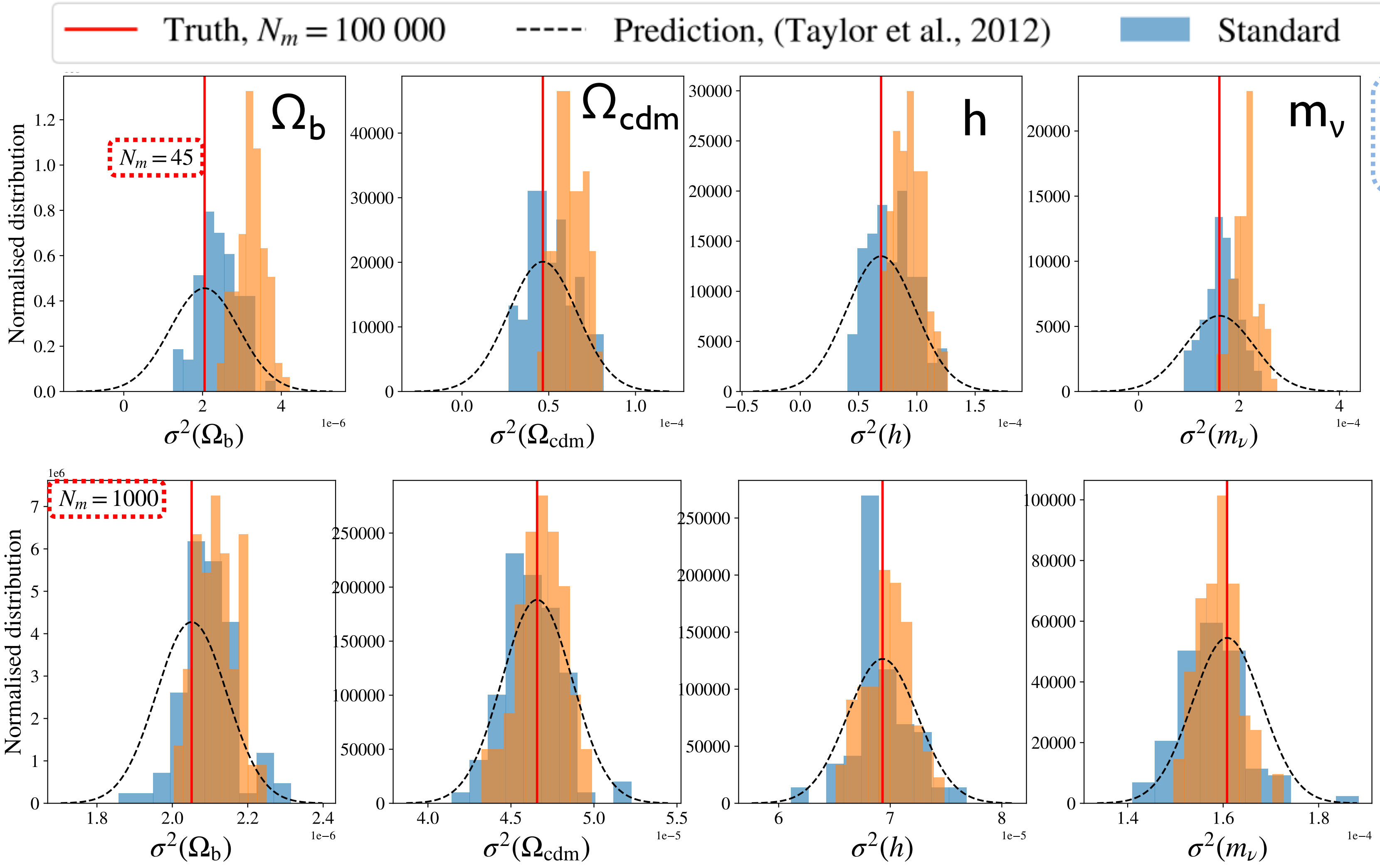


N_m = 45 ➔ Best-fit dispersion of ~ 130%

N_m = 1000 ➔ Best-fit dispersion of ~ 20%

Testing NERCOME on parameter estimation

Distribution of variances for 100 realisations of the covariance



N_m = 45 ➔ Variance dispersion of ~ 25%

N_m = 1000 ➔ Variance dispersion of ~ 3%

Testing NERCOME on parameter estimation

Distribution of best-fits for 100 realisations of the covariance

$\sqrt{\mathrm{Var}\left[\theta^{\mathrm{bf}}\right]}/\sigma_{\theta,\mathrm{Truth}}$ [%]	Ω_{b}	Ω_{cdm}	h	m_ν	$\sqrt{B(N_b-N_p)}$ [%]
Standard, $N_m = 45$	124.46	134.89	132.63	141.85	148.79
NERCOME, $N_m = 45$	78.54	69.49	70.16	68.15	
Standard, $N_m = 1000$	17.57	20.00	19.85	19.18	16.40
NERCOME, $N_m = 1000$	16.03	17.56	17.64	17.82	

[Dodelson & Schneider 2013]

$\sqrt{\mathrm{Var}\left[\theta_{\mathrm{N}}^{\mathrm{bf}}\right]/\mathrm{Var}\left[\theta_{\mathrm{S}}^{\mathrm{bf}}\right]}-1$ [%]	Ω_{b}	Ω_{cdm}	h	m_ν
$N_m = 45$	-38.89	-48.48	-47.11	-51.95
$N_m = 1000$	-8.79	-12.16	-11.14	-7.09

$\left[\langle\theta^{\mathrm{bf}}\rangle-\theta_{\mathrm{Truth}}\right]/\sigma_{\theta,\mathrm{Truth}}$	Ω_{b}	Ω_{cdm}	h	m_ν
Standard, $N_m = 45$	$<\lvert 0.1\rvert$	$<\lvert 0.1\rvert$	$<\lvert 0.1\rvert$	-0.14
NERCOME, $N_m = 45$	0.60	0.37	-0.26	$<\lvert 0.1\rvert$
Standard, $N_m = 1000$	$<\lvert 0.1\rvert$	$<\lvert 0.1\rvert$	$<\lvert 0.1\rvert$	$<\lvert 0.1\rvert$
NERCOME, $N_m = 1000$	$<\lvert 0.1\rvert$	$<\lvert 0.1\rvert$	$<\lvert 0.1\rvert$	$<\lvert 0.1\rvert$

Testing NERCOME on parameter estimation

Distribution of variances for 100 realisations of the covariance

$\sqrt{\mathrm{Var}\left[\sigma_\theta^2\right]}/\sigma^2_{\theta,\mathrm{Truth}}$ [%]	Ω_b	Ω_cdm	h	m_ν	$\sqrt{2/(N_m-N_b-4)}$ [%]
Standard, $N_m=45$	24.86	28.70	28.80	20.27	42.64
NERCOME, $N_m=45$	17.35	18.17	18.80	14.74	
Standard, $N_m=1000$	3.58	3.76	3.73	4.30	4.55
NERCOME, $N_m=1000$	2.64	3.00	3.00	2.71	

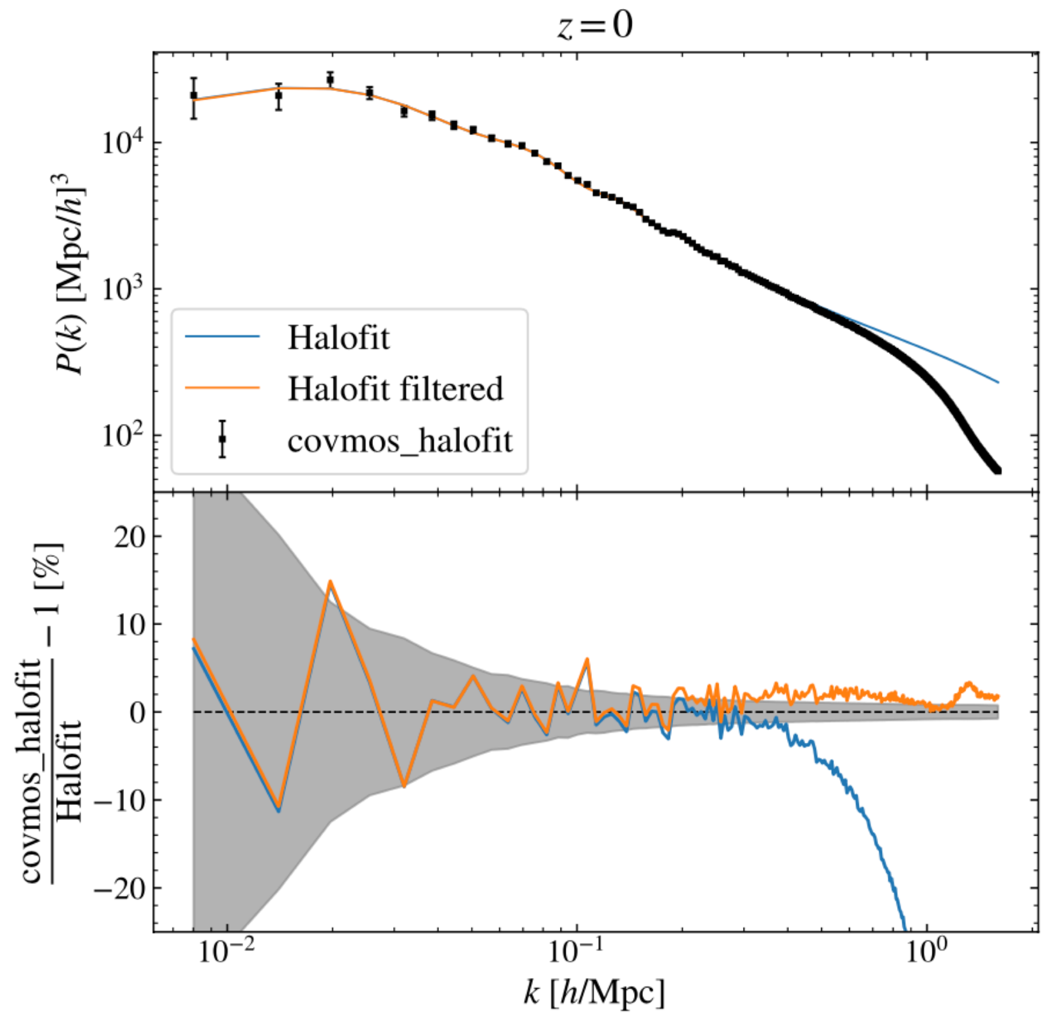
[Taylor et al. 2013]

$\sqrt{\mathrm{Var}\left[\sigma^2_{\theta,\mathrm{N}}\right]/\mathrm{Var}\left[\sigma^2_{\theta,\mathrm{S}}\right]}-1$ [%]	Ω_b	Ω_cdm	h	m_ν
$N_m=45$	-30.24	-36.68	-34.72	-27.27
$N_m=1000$	-26.34	-20.18	-20.09	-36.97

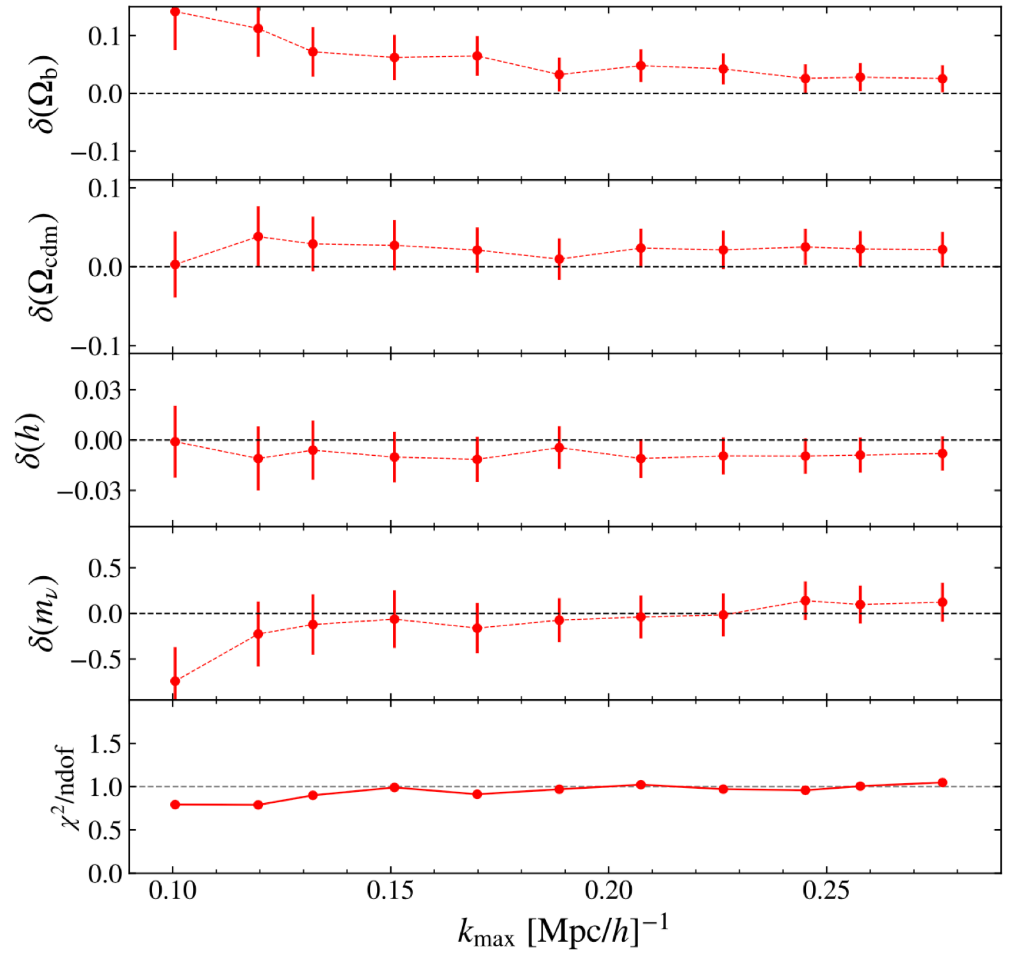
$\langle\sigma_\theta^2\rangle/\sigma^2_{\theta,\mathrm{Truth}}-1$ [%]	Ω_b	Ω_cdm	h	m_ν	$A+B(N_p+1)$ [%]
Standard, $N_m=45$	18.14	11.10	13.60	3.13	43.51
NERCOME, $N_m=45$	57.84	33.03	33.81	32.36	
Standard, $N_m=1000$	2.21	-0.74	-0.28	-1.93	0.52
NERCOME, $N_m=1000$	3.32	0.48	0.76	-0.96	

[Percival et al. 2014]

To ensure no bias from the model
➔Target Halofit P(k) with covmos



Parameter inference with covmos_halofit

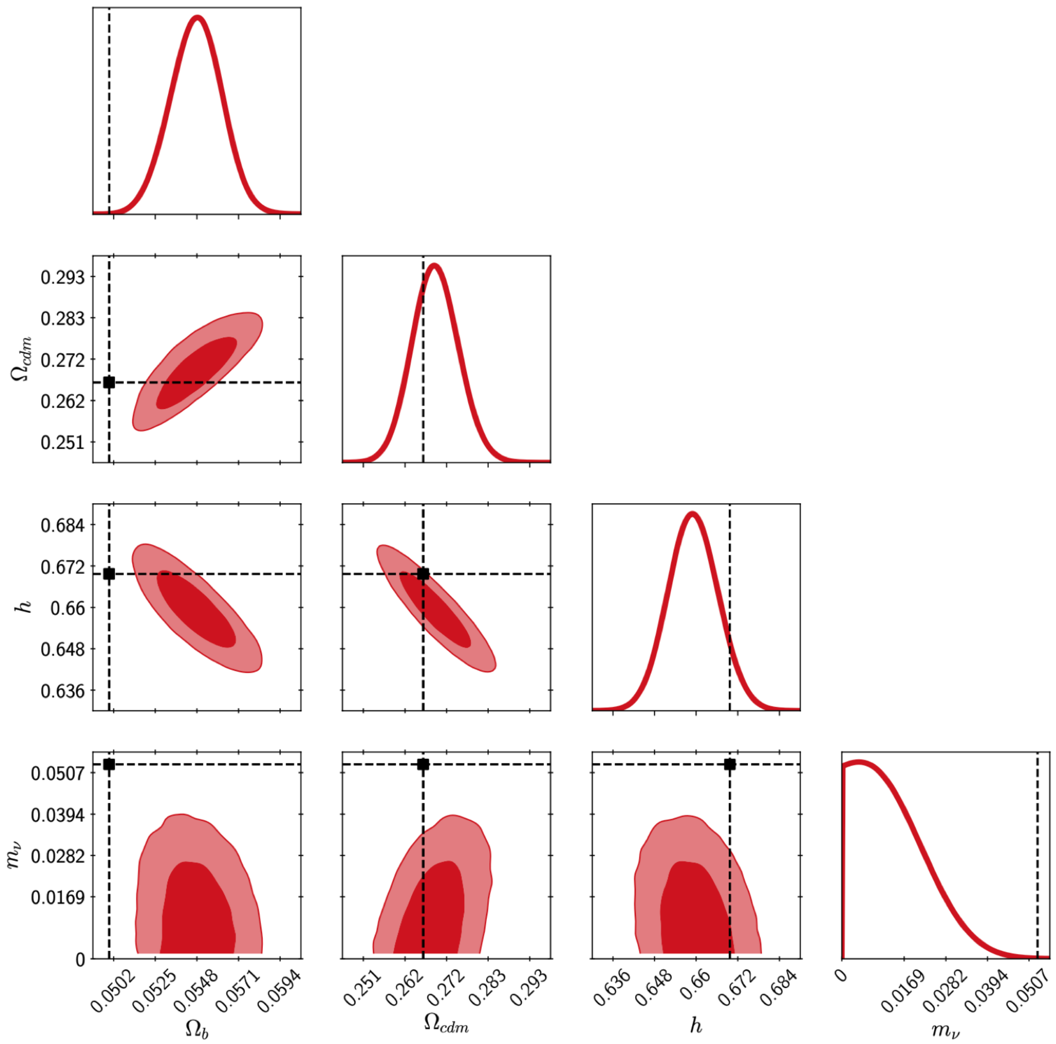


Disagreement between theoretical prediction and estimation for the variance on the error and the bias on the variance

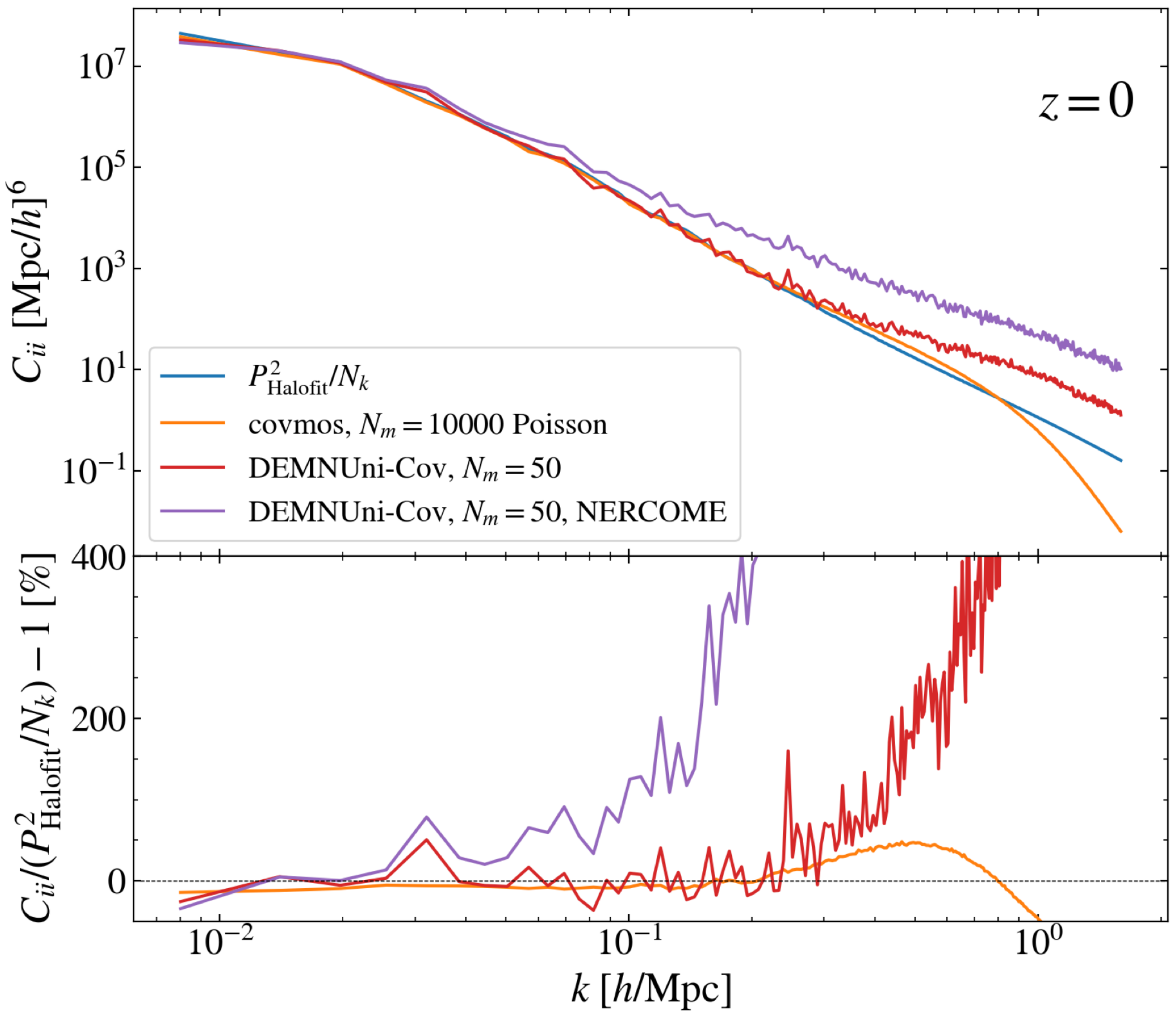
➔ Comes from non-Gaussian posteriors ?

Example of a non-Gaussian posterior

- covmos_halofit
- Standard covariance with Nm = 45



NERCOME



$z = 0$
C_{ii} [Mpc/h]^6
10^7
10^5
10^3
10^1
10^{-1}
$P^2_{\rm Halofit}/N_k$
covmos, $N_m = 10000$ Poisson
DEMNUni-Cov, $N_m = 50$
DEMNUni-Cov, $N_m = 50$, NERCOME
$C_{ii}/(P^2_{\rm Halofit}/N_k) - 1$ [%]
400
200
0
10^{-2}
10^{-1}
10^0
k [h/Mpc]

Target : P(k) and PDF

With the 1-point PDF all the cumulants should be reproduced

$$\langle\delta^4(\mathbf{x})\rangle_c = \int \mathrm{d}^3\mathbf{k}_1 \mathrm{d}^3\mathbf{k}_2 \mathrm{d}^3\mathbf{k}_3 \, T(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3).$$

We hope to reproduce the bin-averaged Trispectrum

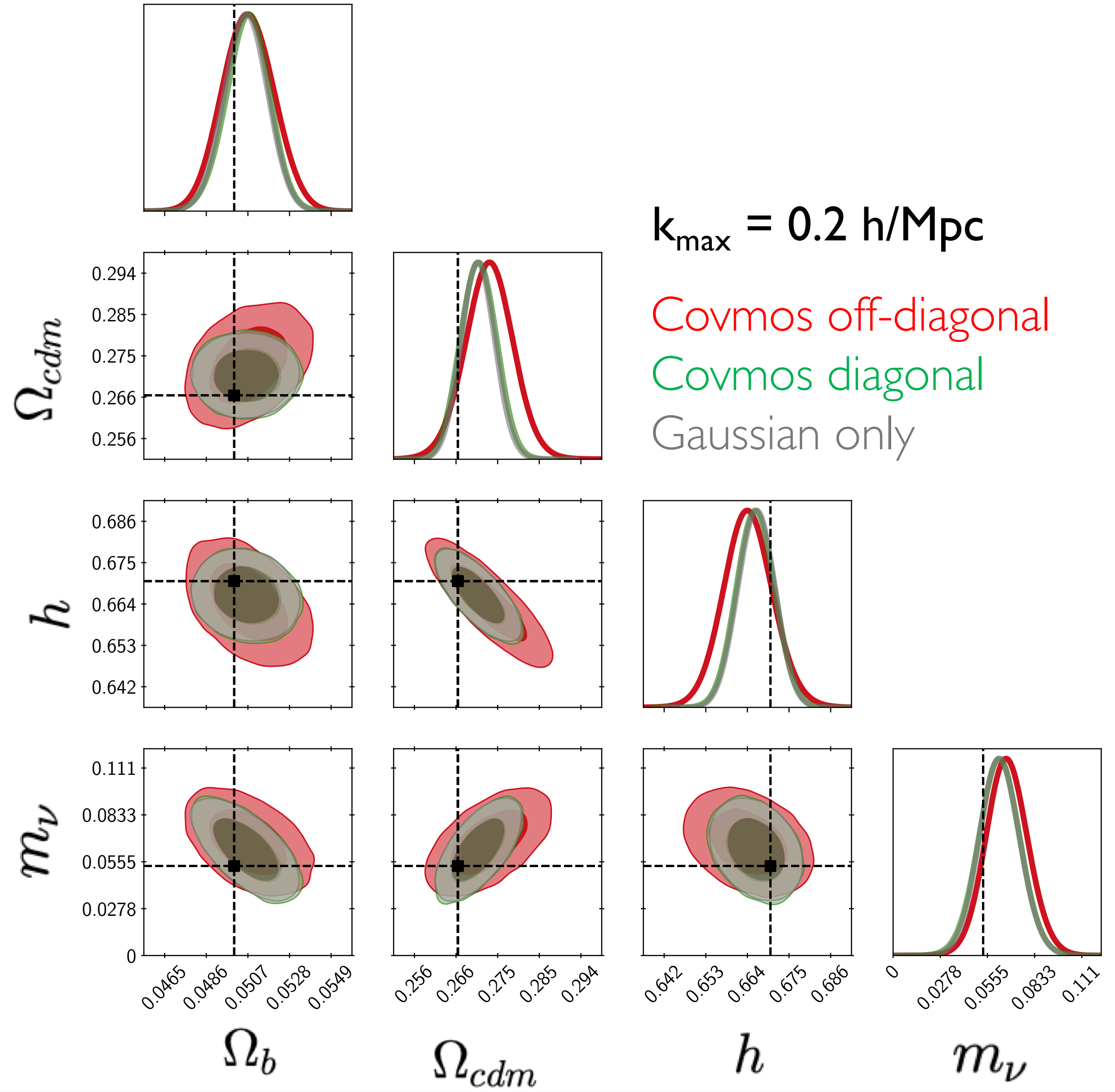
$$\bar{T}(k_i, k_j) = \int_{k_i} \frac{\mathrm{d}^3 k_1}{V_s(k_i)} \int_{k_j} \frac{\mathrm{d}^3 k_2}{V_s(k_j)} T(\mathbf{k}_1, -\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_2).$$

Compare constraints obtained with the Gaussian covariance only, to quantify the impact of non-Gaussian covariance

$$\mathbf{C}(k_i,k_j) = \frac{P^2(k_i)}{N_{k_i}}\delta_{ij} + \bar{T}(k_i,k_j) = \mathbf{C}^{\mathrm{G}}(k_i) + \mathbf{C}^{\mathrm{NG}}(k_i,k_j),$$

Parameter estimation with MCMC :

- Data : DEMNUni-Cov at 5 redshifts : $z \approx 0,\ 0.5,\ 1,\ 1.5,\ 2$
- Non-linear model : Halofit
- **Covariance :**
  1. Gaussian only
  2. Covmos Nm = 10 000, diagonal only
  3. Covmos Nm = 10 000, full



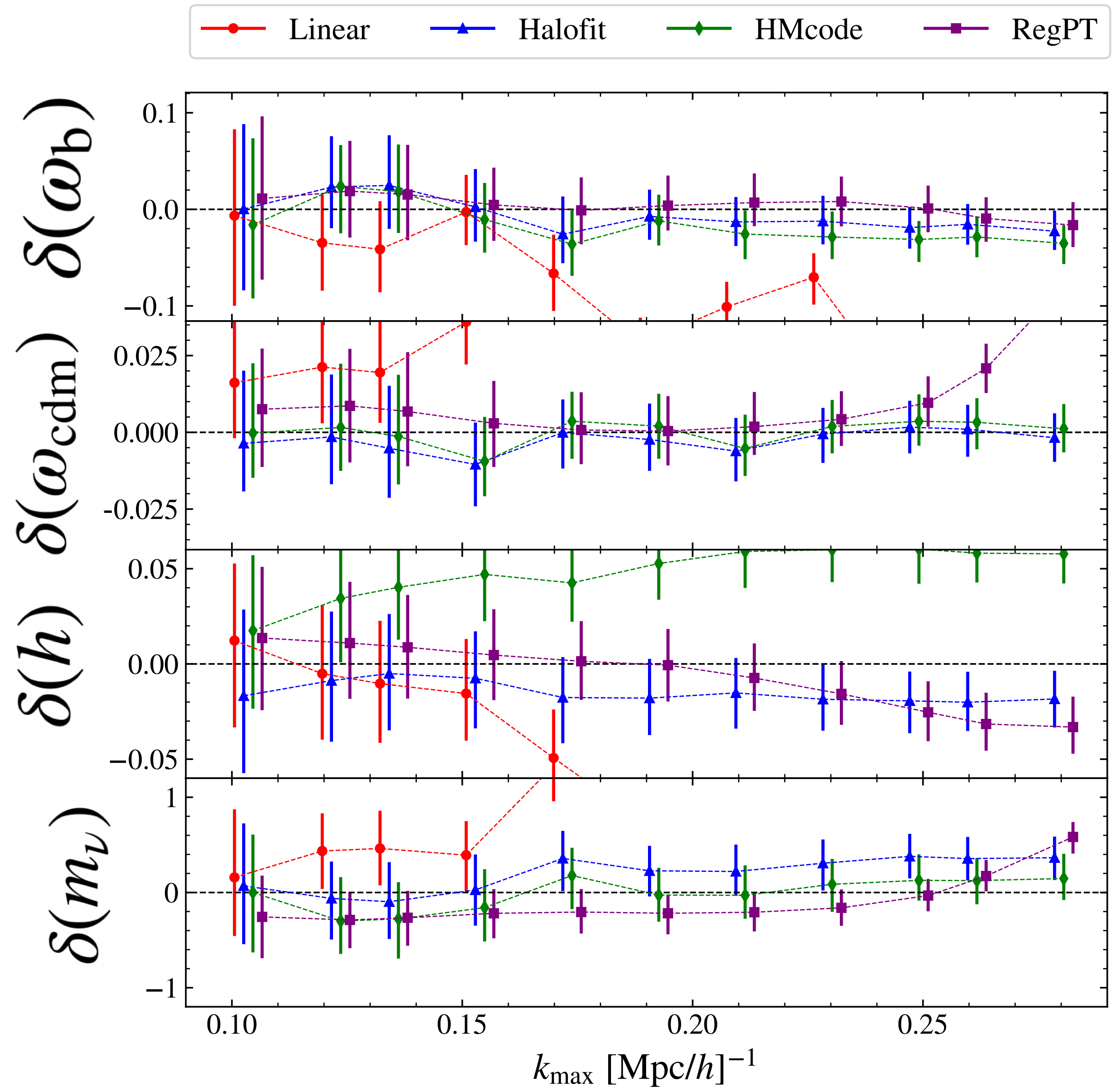
- Best-fit is unchanged
- Off-diagonal elements are the most important

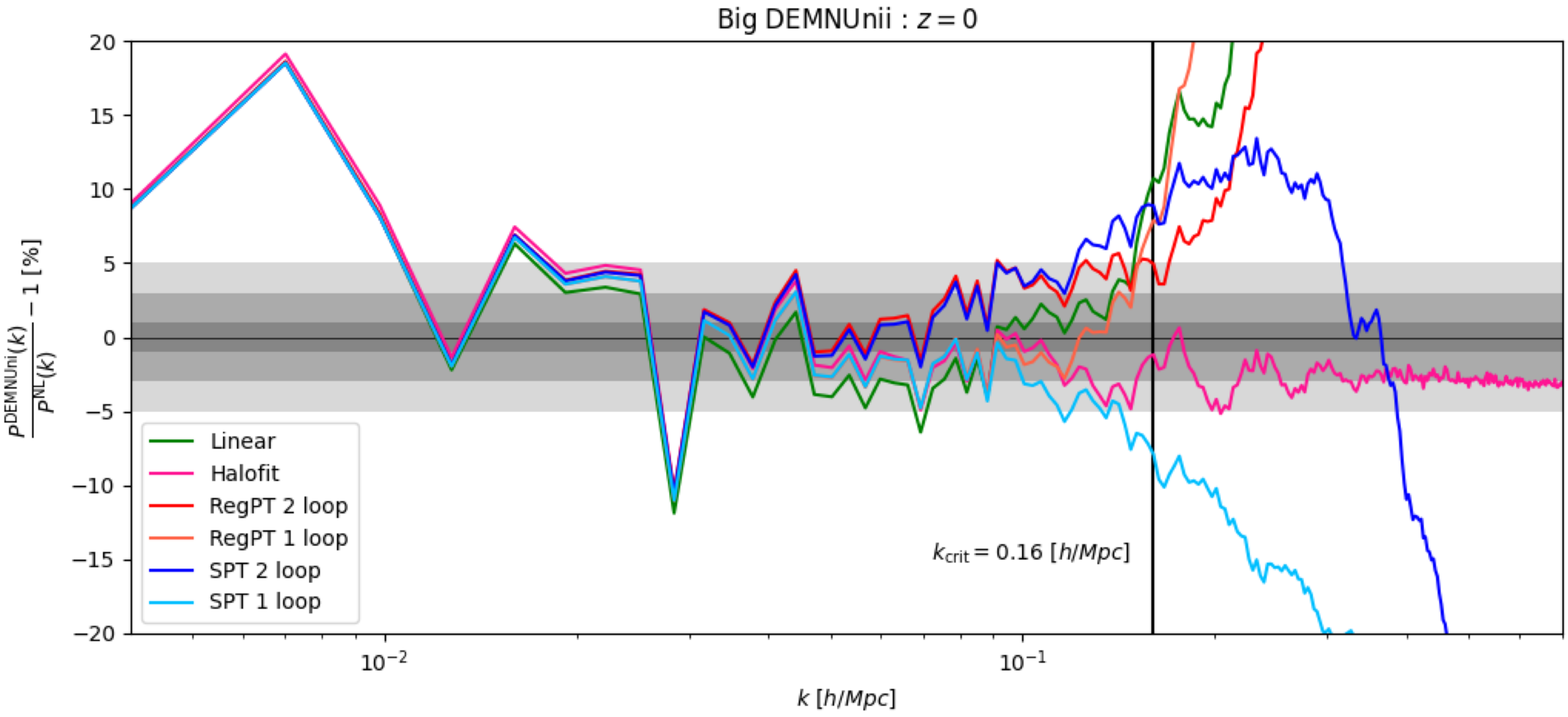
Testing non-linear models with massive neutrinos

- ω_b and ω_{cdm} : Unbiased for all models up to kmax = 0.22 h/Mpc

- <u>Significant bias on h and m_ν :</u>

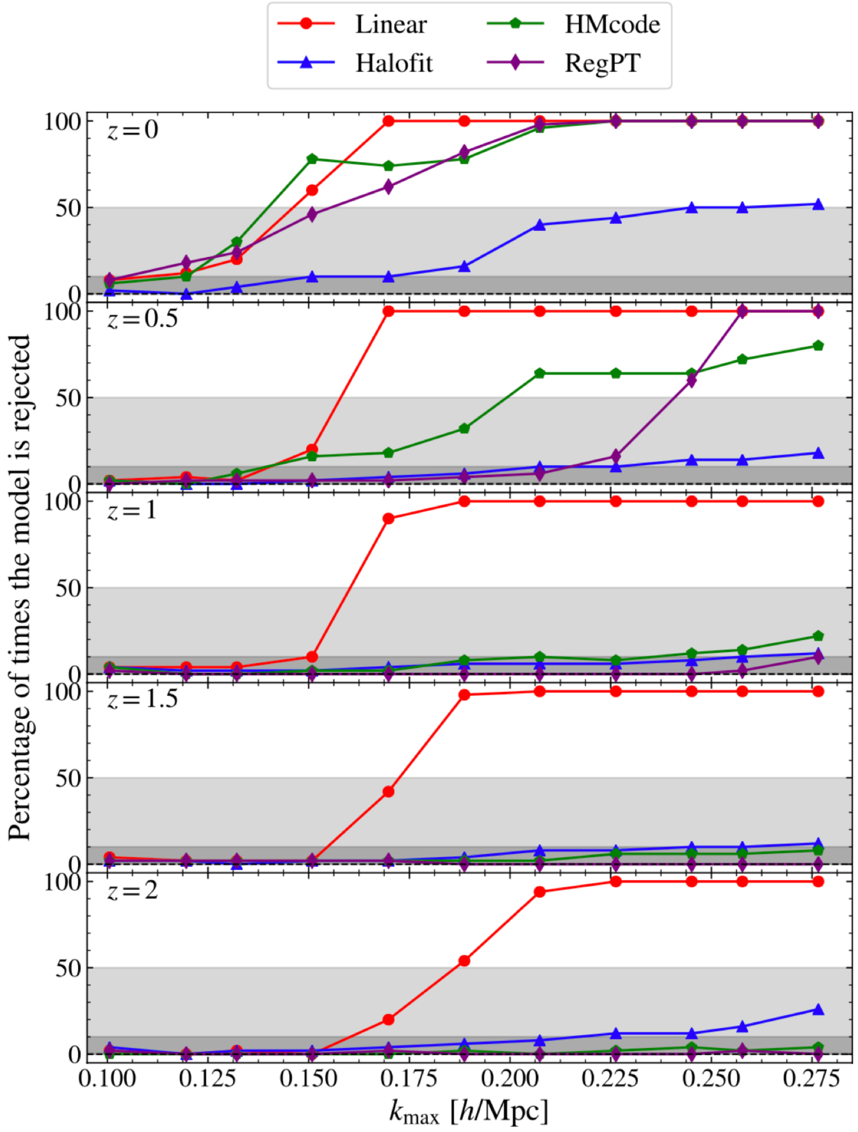
  - Halofit : ~ 1σ bias on h and m_ν
  - HMcode : Unbiased estimation of m_ν, but strong bias on h
  - RegPT : Lower error-bar on m_ν ➔ >1σ bias




Big DEMNUnii : $z = 0$
$\frac{P^{DEMNUni}(k)}{P^{NL}(k)} - 1$ [%]
20
15
10
5
0
−5
−10
−15
−20
10^{-2}
10^{-1}
k [h/Mpc]
$k_{crit} = 0.16$ [h/Mpc]
Linear
Halofit
RegPT 2 loop
RegPT 1 loop
SPT 2 loop
SPT 1 loop

Testing non-linear models with massive neutrinos

Goodness of fit test for each DEMNUni-Cov realisation



$$\text{FoM} \equiv \frac{1}{\sqrt{\det(\hat{\mathbf{\Phi}})}}.$$

$$\text{FoB} \equiv \left[\delta\boldsymbol{\theta}^T \hat{\mathbf{\Phi}}^{-1} \delta\boldsymbol{\theta}\right]^{1/2},$$

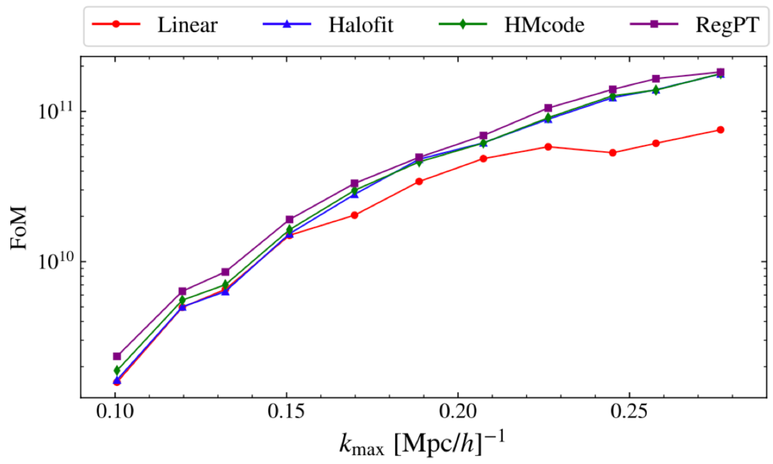


Figure 3.29.: FoM with respect to k_{max}.

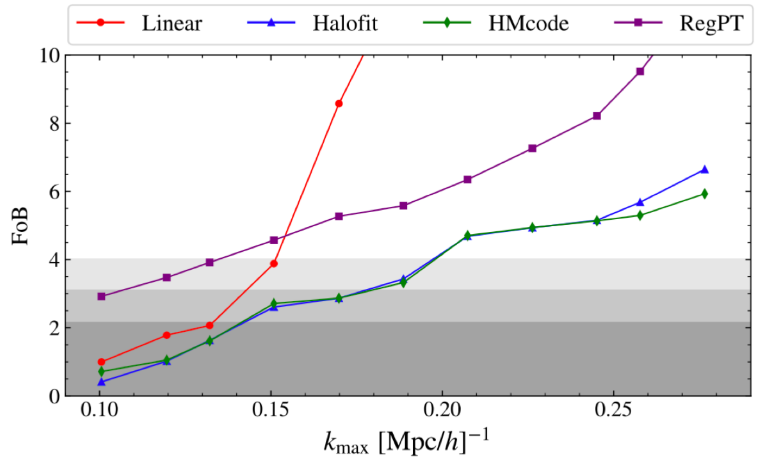
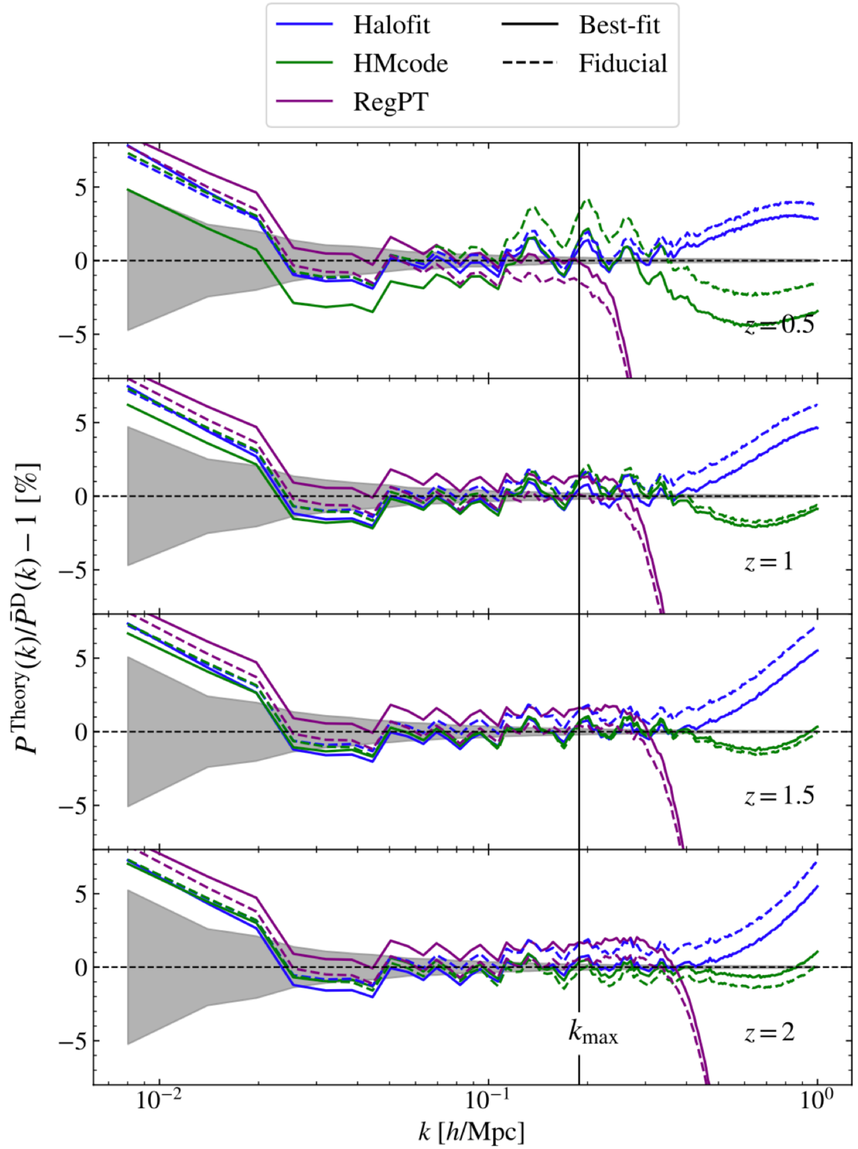
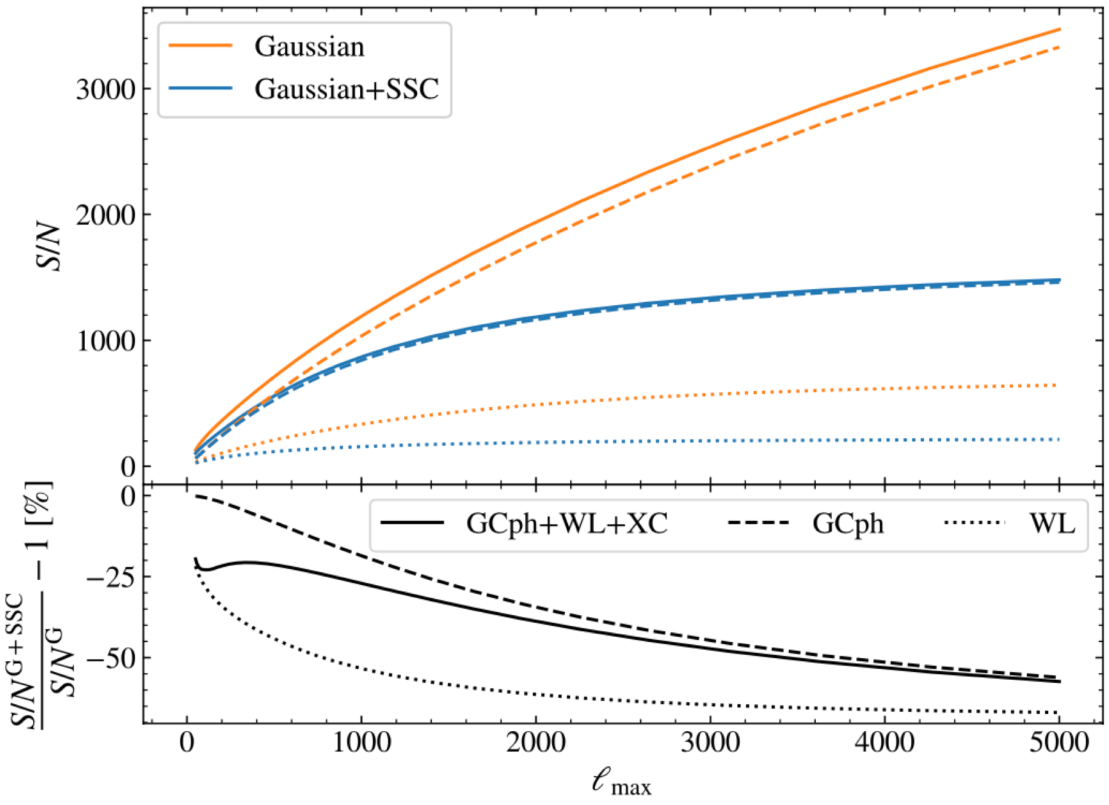


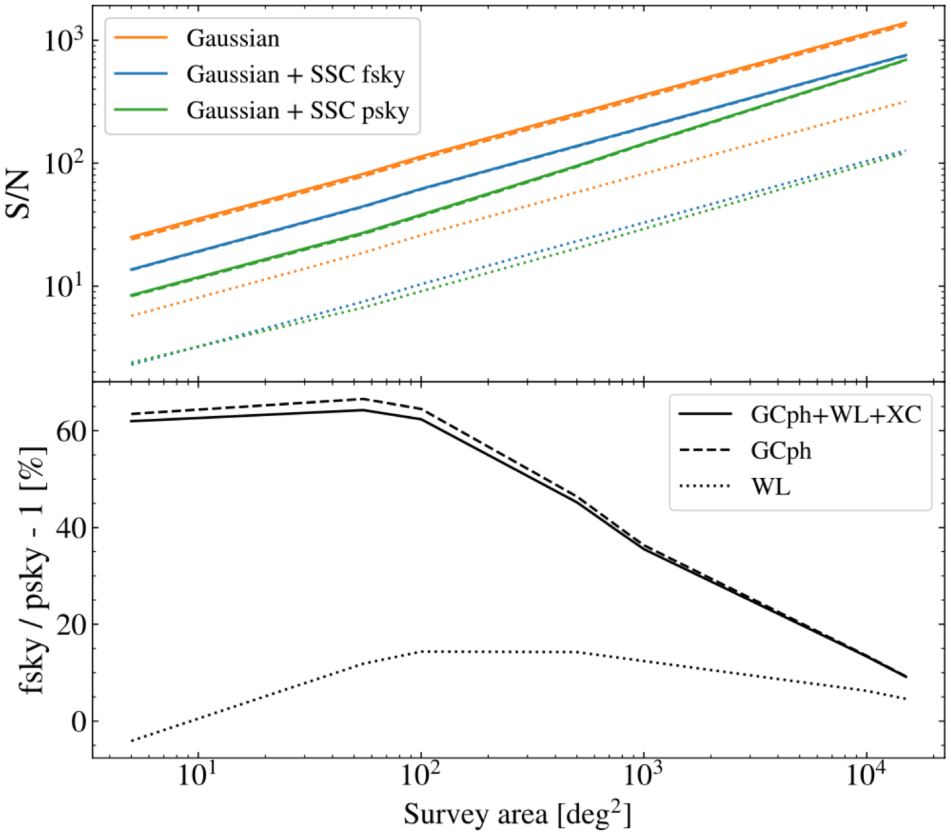
Figure 3.30.: FoB with respect to k_{max}. The grey shaded areas show the 68.3%, 95.5% and 99.7% confidence intervals, corresponding to 1σ, 2σ and 3σ

Testing non-linear models with massive neutrinos



$$(S/N)^2 = \sum_{i,j,k,l} \sum_{\ell,\ell'} C_{ij}^{AB}(\ell)\, \mathrm{Cov}\left[C_{ij}^{AB}(\ell), C_{kl}^{CD}(\ell')\right]^{-1} C_{kl}^{CD}(\ell'),$$





Weak Lensing and Intrinsic Alignments

Observed ellipticities of galaxies : $\epsilon = \gamma + \epsilon^{\mathrm{I}},$ → $W^{\mathrm{WL}}(z) = W^{\gamma}(z) + W^{\mathrm{IA}}(z).$

Cosmic shear ← γ ; ϵ^{I} → Intrinsic alignement

$$W_i^{\gamma}(z) = \frac{3}{2}\left(\frac{H_0}{c}\right)^2 \Omega_{\mathrm{m},0} \frac{1+z}{r(z)} \int_z^{z_{\max}} \mathrm{d}n_i(z') \left[1 - \frac{r(z)}{r(z')}\right],$$

$$W_i^{\mathrm{IA}}(z) = -A(z) \frac{n_i(z)}{r^2(z)} \frac{c}{H(z)}.$$

$$A(z) = \frac{\mathcal{A}_{\mathrm{IA}} \mathcal{C}_{\mathrm{IA}} \Omega_m \mathcal{F}_{\mathrm{IA}}(z)}{D(z)},$$

$$\mathcal{F}_{\mathrm{IA}} = (1+z)^{\eta_{\mathrm{IA}}} \left[\frac{\langle L|L\rangle(z)}{L_*(z)}\right]^{\beta_{\mathrm{IA}}},$$

eNLA (extended Non-Linear Alignment) model

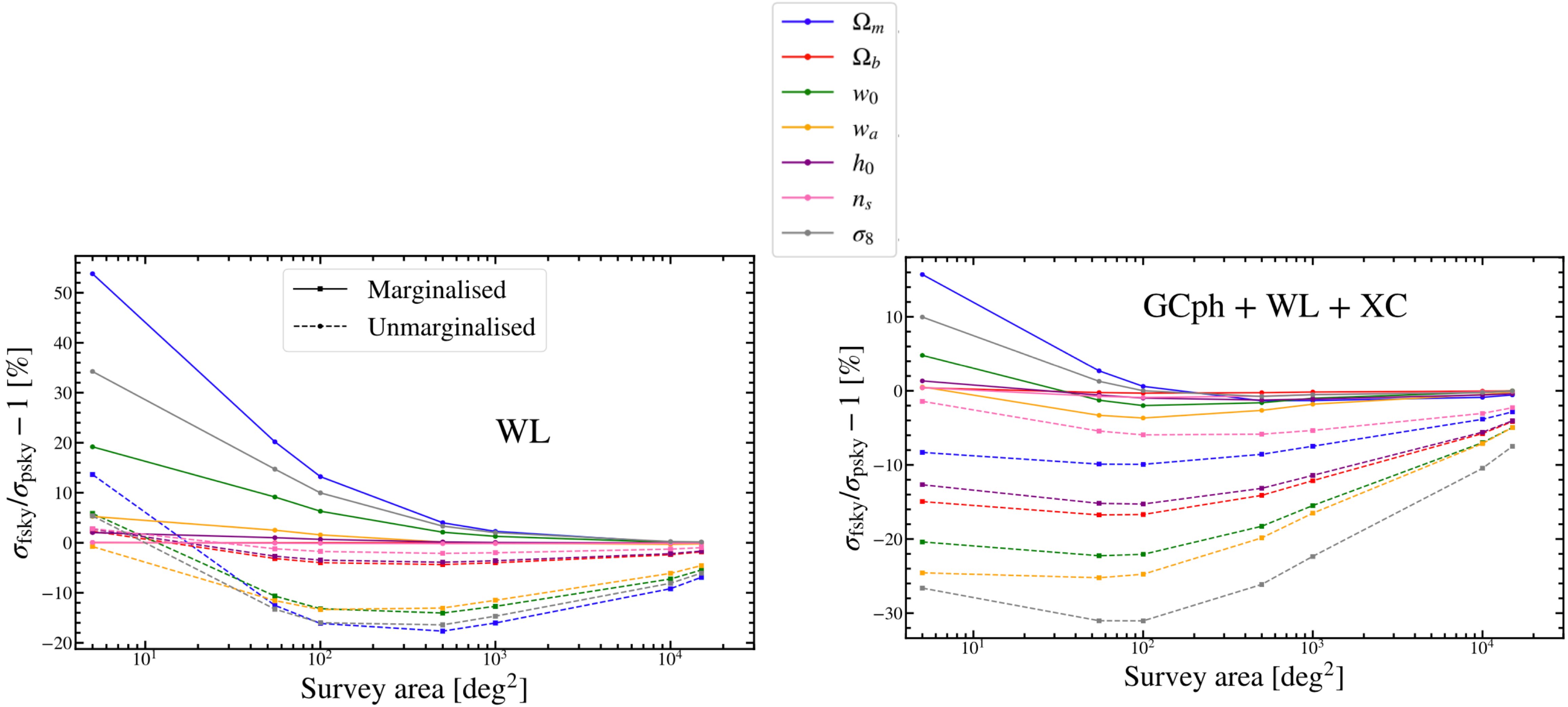
$$\mathrm{Cov_G}\left[C_{ij}^{AB}(\ell), C_{kl}^{CD}(\ell')\right] = \frac{\delta^{\mathrm{K}}_{\ell\ell'}}{(2\ell+1) f_{\mathrm{SKY}} \Delta\ell} \left\{\left[C_{ik}^{AC}(\ell) + N_{ik}^{AC}(\ell)\right]\left[C_{jl}^{BD}(\ell') + N_{jl}^{BD}(\ell')\right] + \left[C_{il}^{AD}(\ell) + N_{il}^{AD}(\ell)\right]\left[C_{jk}^{BC}(\ell') + N_{jk}^{BC}(\ell')\right]\right\},$$

Noise terms N_{ij}

$\bar{n}_i = 30\ \mathrm{gal/arcmin}^2$

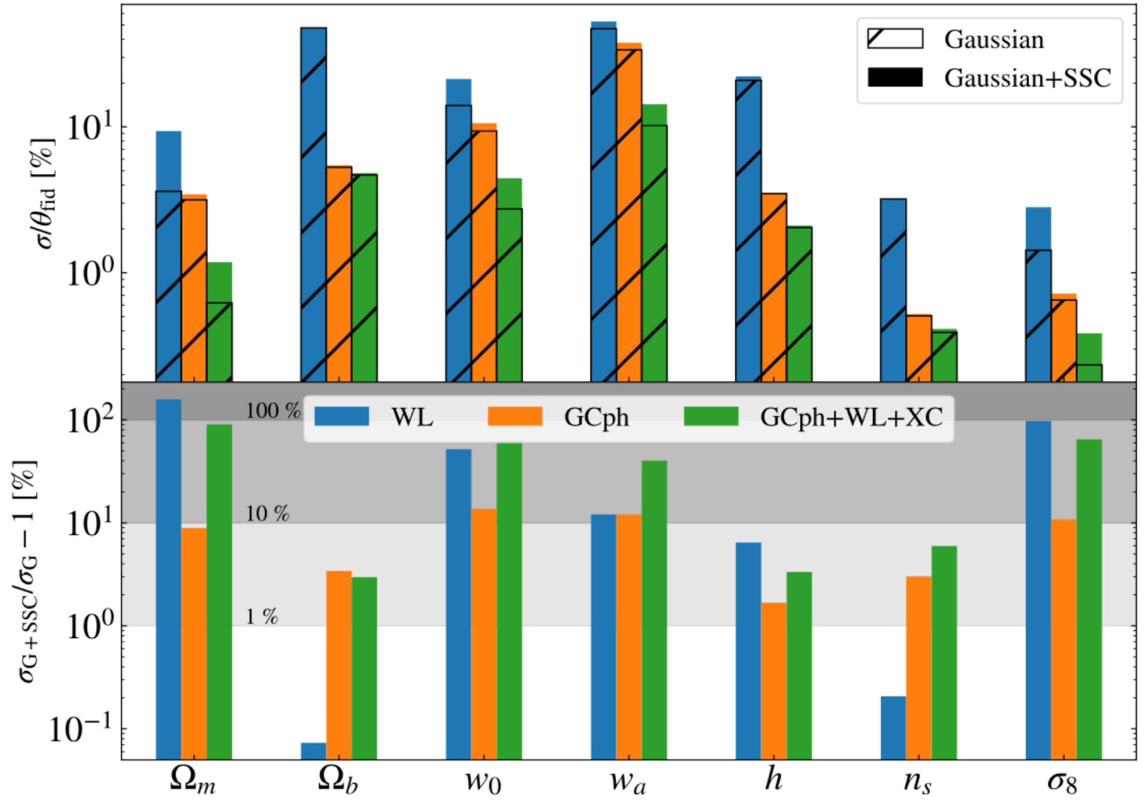
- GCph : $\delta^{\mathrm{K}}_{ij} / \bar{n}_i$ Shot-noise
- WL : $\sigma_\epsilon^2 \delta^{\mathrm{K}}_{ij} / \bar{n}_i$ Shape-noise

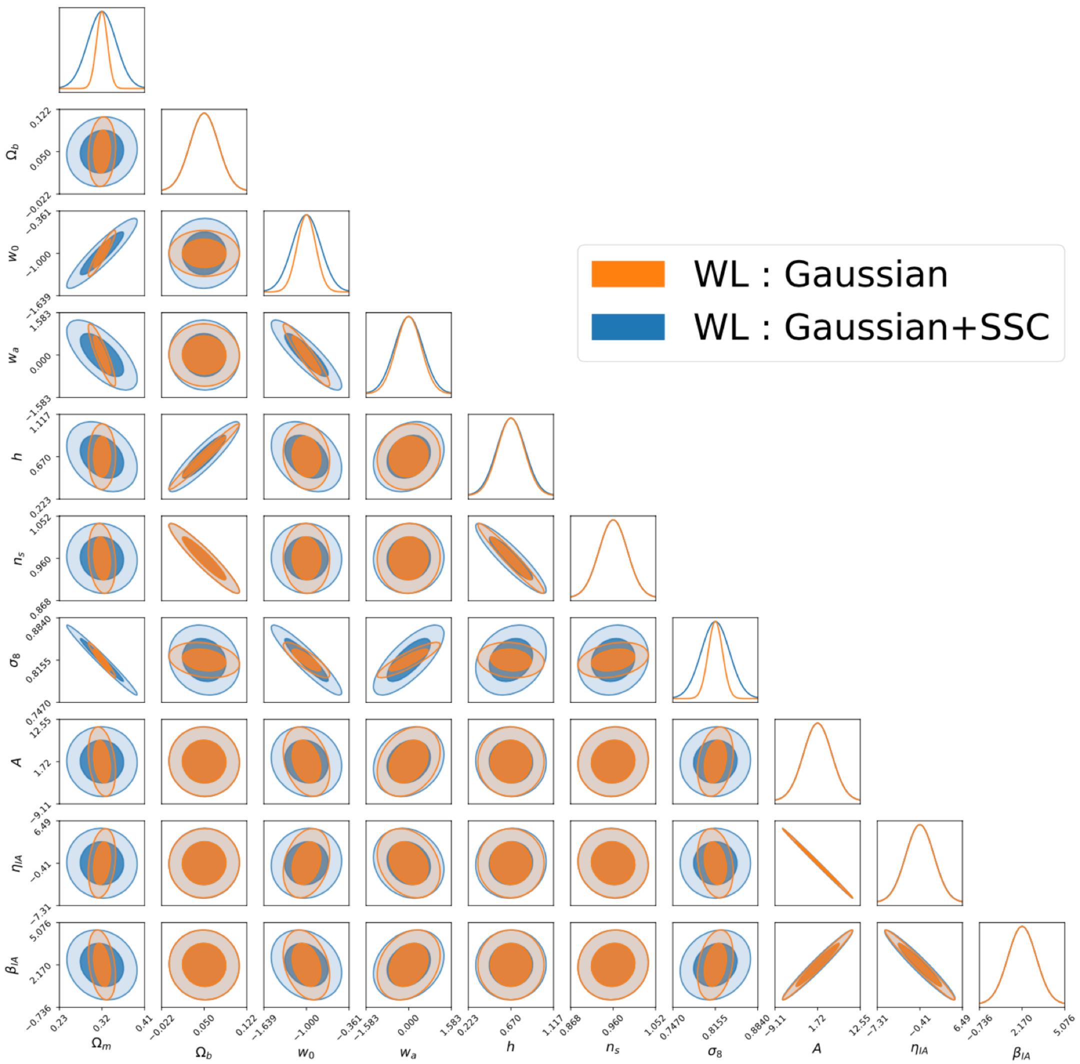
$\sigma_\epsilon^2 = 0.3$

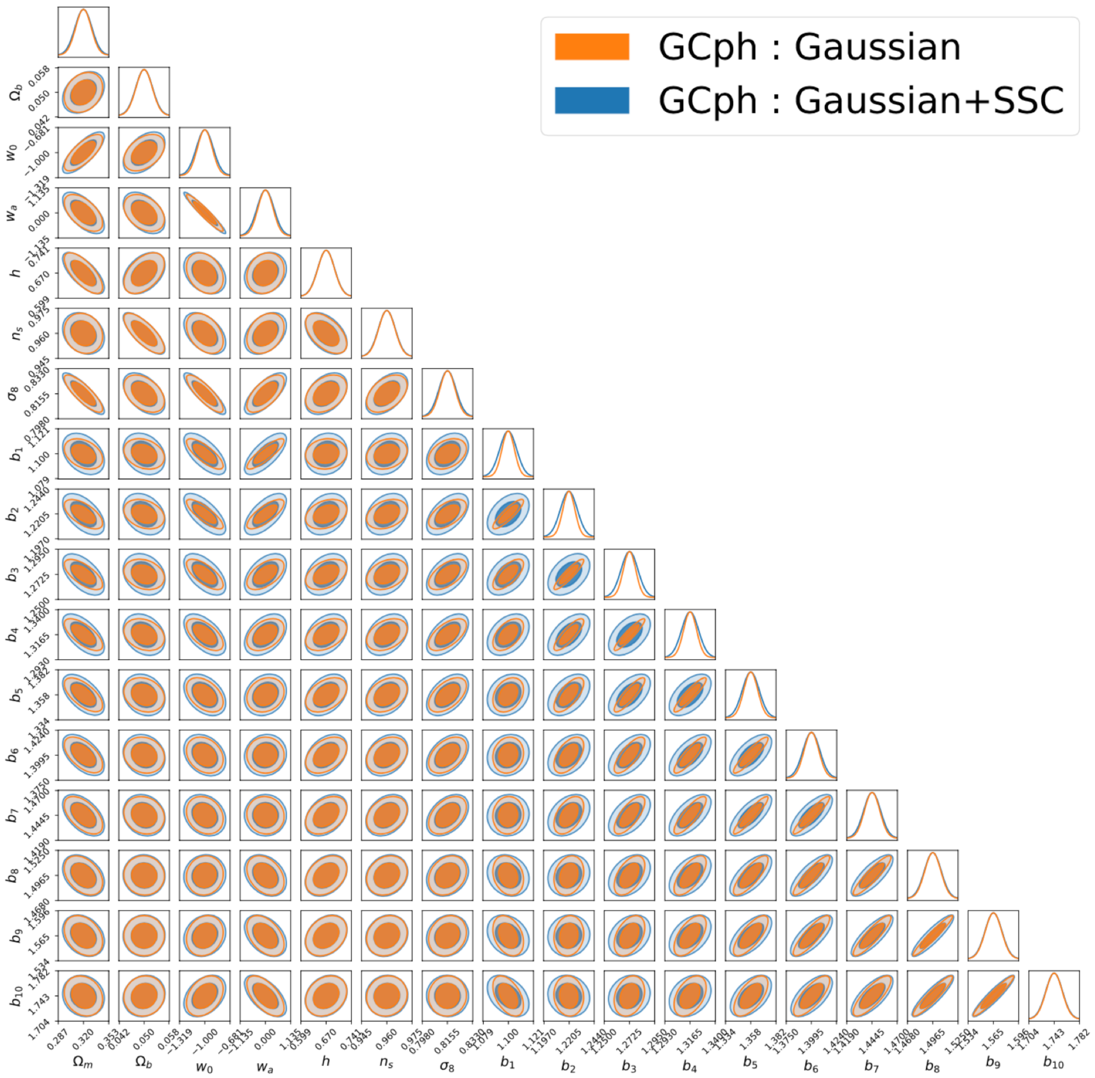

Ω_m
Ω_b
w_0
w_a
h_0
n_s
σ_8
Marginalised
Unmarginalised
WL
GCph + WL + XC
$\sigma_{\rm fsky}/\sigma_{\rm psky} - 1$ [%]
Survey area [deg^2]

Full-sky SSC : parameters

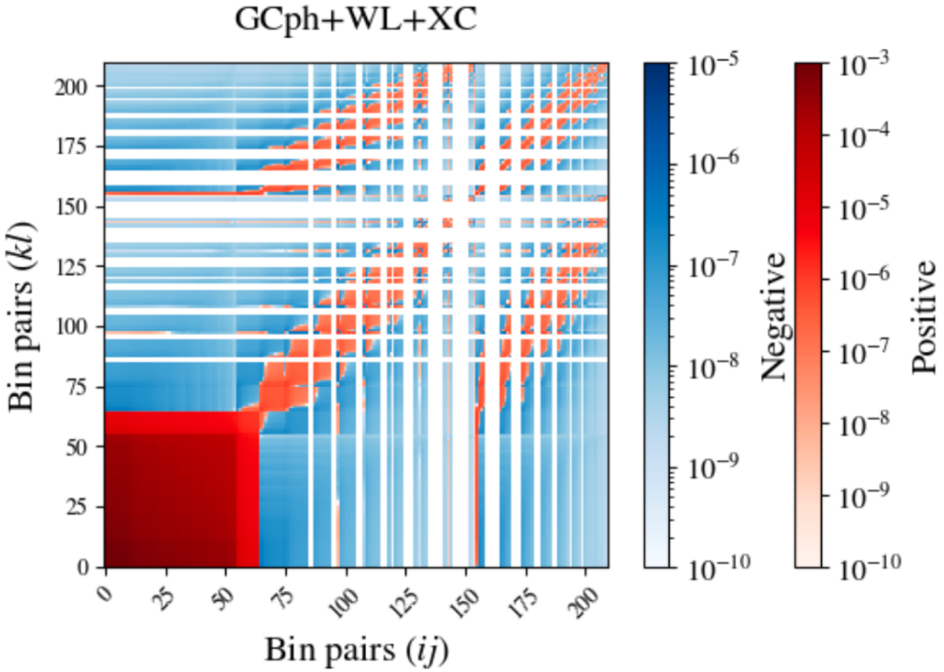
$\sigma/\theta_{\rm fid}$ [%]	Ω_m	Ω_b	w_0	w_a	h	n_s	σ_8	FoM_{w_0,w_a}
GCph, Gaussian	3.16	5.27	9.35	33.76	3.48	0.51	0.65	103.71
GCph, Gaussian+SSC	3.44	5.45	10.64	37.84	3.54	0.52	0.72	89.55
G+SSC/G - 1 [%]	8.96	3.41	13.69	12.10	1.68	3.02	10.84	-13.65
WL, Gaussian	3.61	47.74	14.03	47.11	20.87	3.20	1.42	43.12
WL, Gaussian+SSC	9.34	47.77	21.31	52.78	22.21	3.21	2.81	18.47
G+SSC/G [%]	158.71	0.07	51.90	12.05	6.44	0.21	97.40	-57.16
GCph + WL + XC, Gaussian	0.62	4.68	2.74	10.21	2.04	0.39	0.23	1038.13
GCph + WL + XC, Gaussian+SSC	1.18	4.82	4.45	14.32	2.10	0.41	0.38	490.60
G+SSC/G [%]	90.34	2.97	62.30	40.25	3.34	5.96	64.75	-52.74

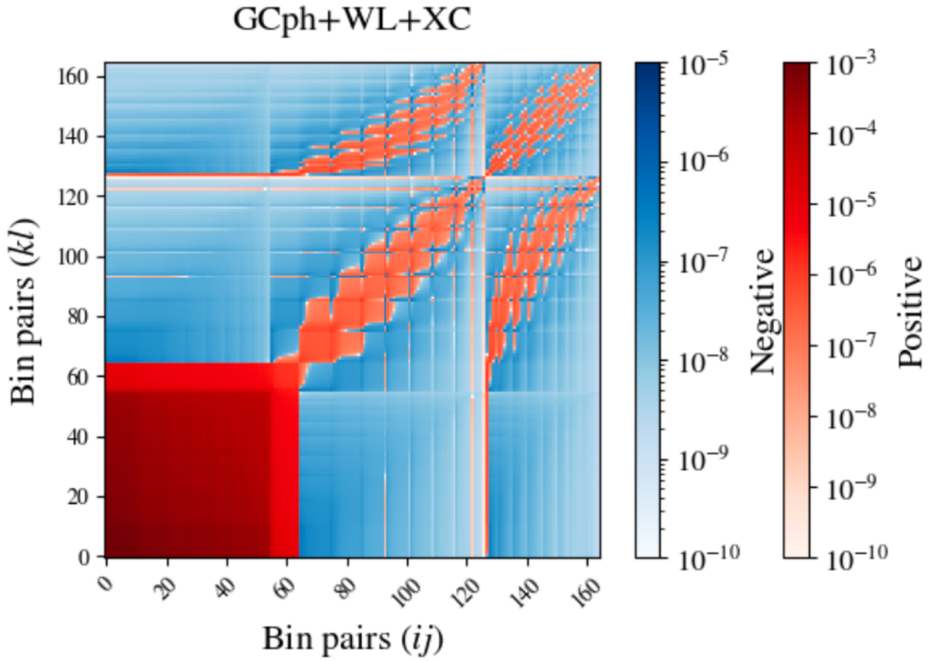



WL : Gaussian
WL : Gaussian+SSC
Ω_m
Ω_b
w_0
w_a
h
n_s
σ_8
A
η_{IA}
β_{IA}


GCph : Gaussian
GCph : Gaussian+SSC
Ω_m
Ω_b
w_0
w_a
h
n_s
σ_8
b_1
b_2
b_3
b_4
b_5
b_6
b_7
b_8
b_9
b_{10}

SSC : S_{ijkl} matrix





SSC : Full-Flat-Partial

Full-sky :

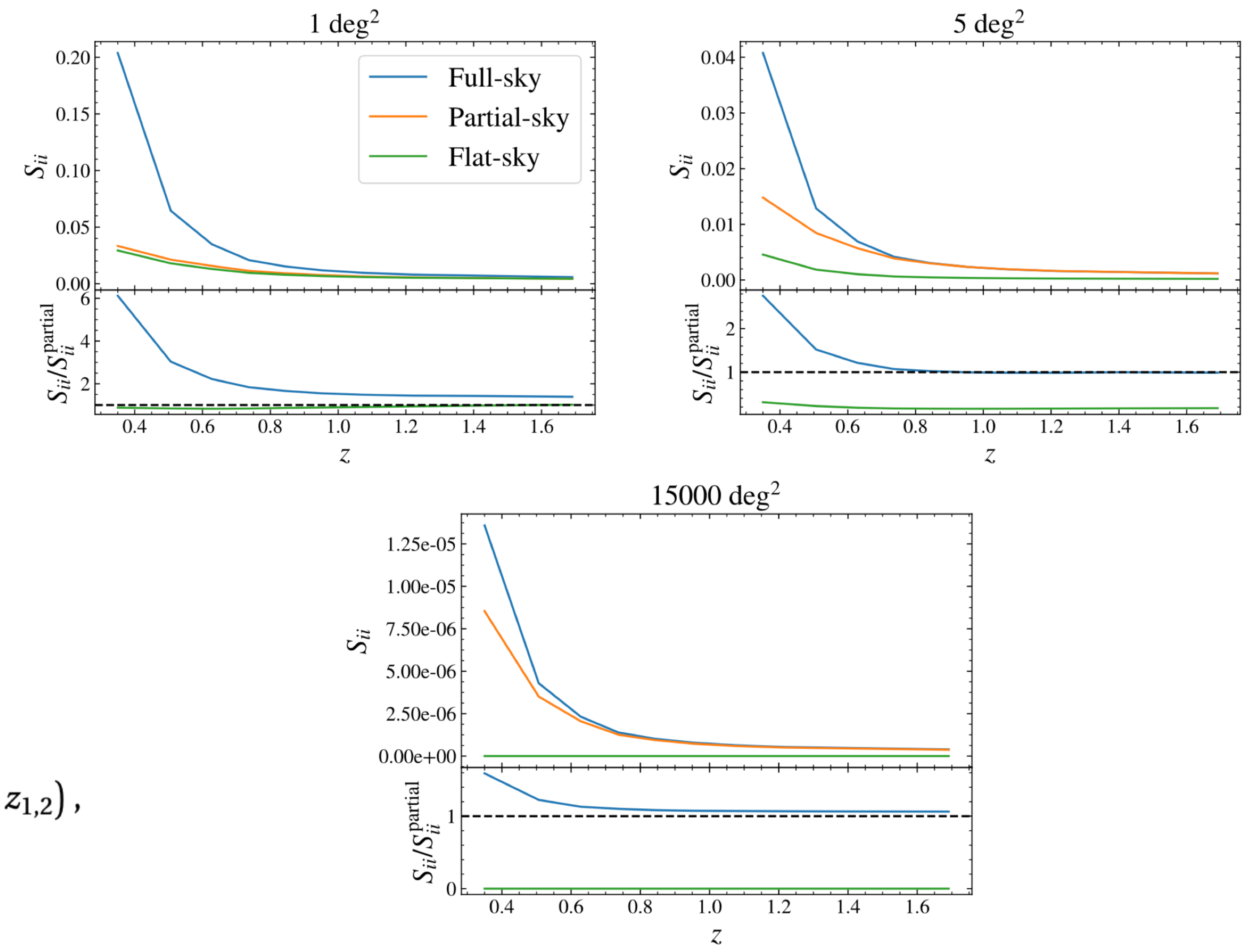
$$\sigma^2_{\text{full-sky}}(z_1, z_2) = \frac{1}{2\pi^2}\int k^2 \mathrm{d}k\; j_0(kr(z_1))\, j_0(kr(z_2)) P_m(k|z_{1,2}),$$

Partial-sky :

$$\sigma^2_{\text{part-sky}}(z_1, z_2) = \frac{1}{\Omega_S^2}\sum_\ell (2\ell+1) C^{\mathcal{W}}(\ell) C^{\mathrm{m}}_{z_{1,2}}(\ell),$$

Flat-sky :

$$S_{i,j} = \frac{1}{2\pi^2}\int k_\perp \mathrm{d}k_\perp 4 \frac{J_1(k_\perp \theta_S r_1)}{k_\perp \theta_S r_1}\frac{J_1(k_\perp \theta_S r_2)}{k_\perp \theta_S r_2} \times \int \mathrm{d}k_\parallel\, j_0\left(\frac{k_\parallel \delta r_1}{2}\right) j_0\left(\frac{k_\parallel \delta r_2}{2}\right) \cos\left[k_\parallel (r_1 - r_2)\right] P_m\left(k \mid z_{1,2}\right),$$



Full-sky :

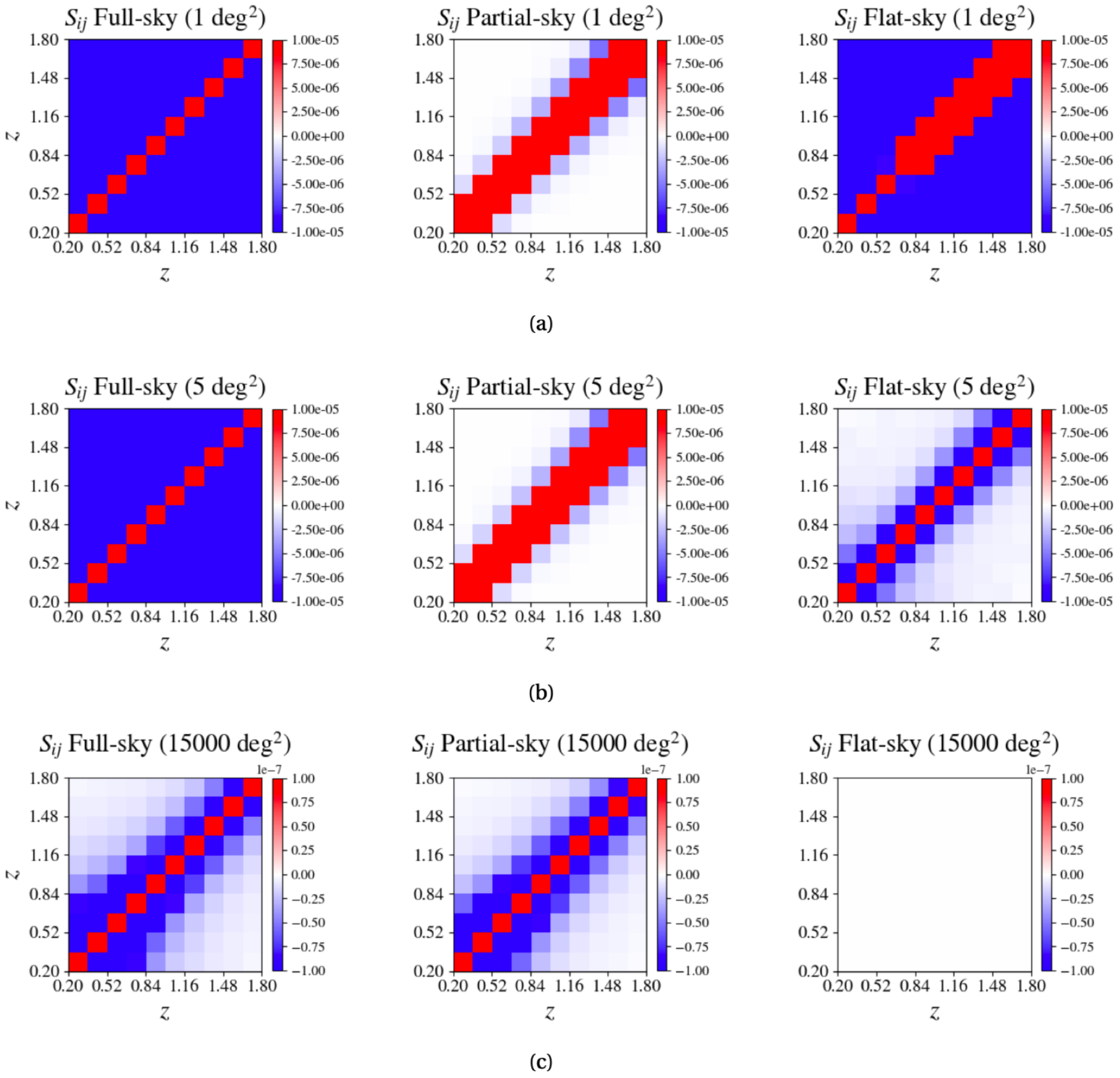
$$\sigma^2_{\mathrm{full-sky}}(z_1, z_2) = \frac{1}{2\pi^2}\int k^2 \mathrm{d}k\, j_0(kr(z_1))\, j_0(kr(z_2)) P_m(k|z_{1,2}),$$

Partial-sky :

$$\sigma^2_{\mathrm{part-sky}}(z_1, z_2) = \frac{1}{\Omega_S^2}\sum_{\ell}(2\ell+1)C^{\mathcal{W}}(\ell)C^{\mathrm{m}}_{z_{1,2}}(\ell),$$

Flat-sky :

$$S_{i,j} = \frac{1}{2\pi^2}\int k_\perp \mathrm{d}k_\perp 4\frac{J_1(k_\perp\theta_S r_1)}{k_\perp\theta_S r_1}\frac{J_1(k_\perp\theta_S r_2)}{k_\perp\theta_S r_2} \times \int \mathrm{d}k_\parallel j_0\left(\frac{k_\parallel \delta r_1}{2}\right) j_0\left(\frac{k_\parallel \delta r_2}{2}\right)\cos\left[k_\parallel (r_1 - r_2)\right] P_m\left(k \,|\, z_{1,2}\right),$$

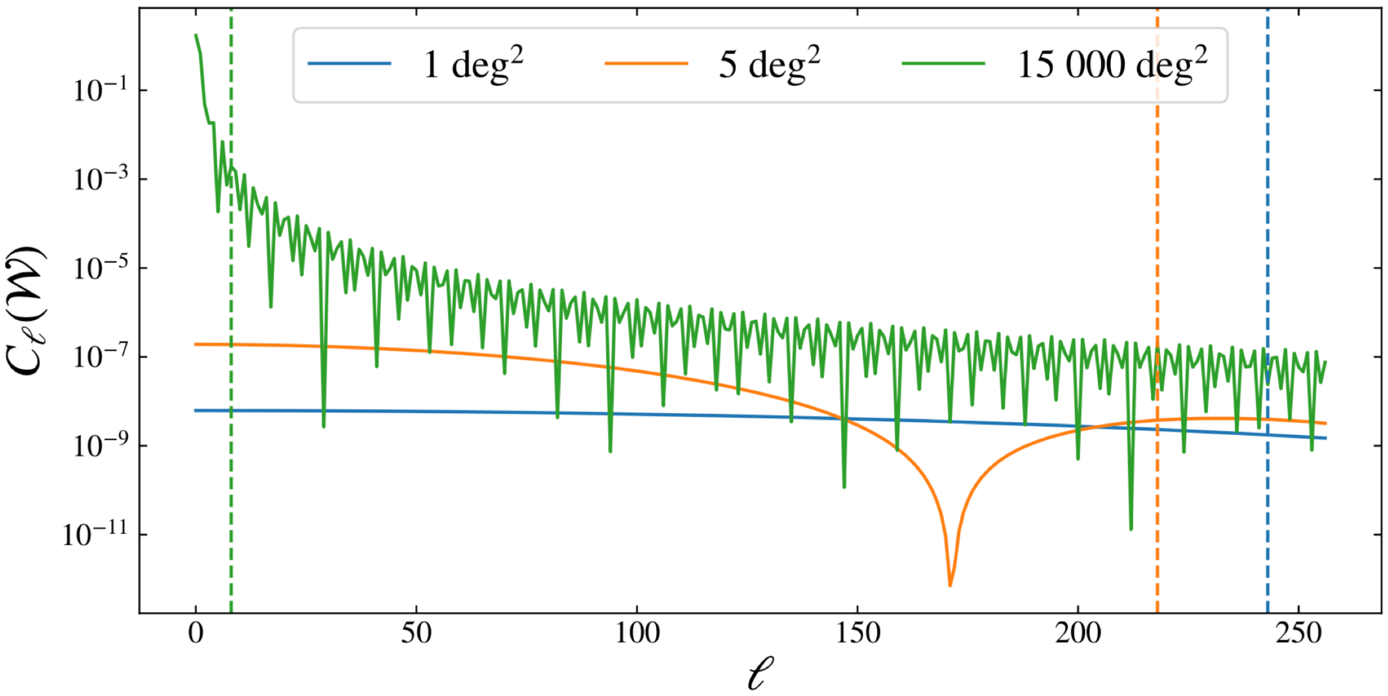


Survey area [deg^2]	Full-sky	Partial-sky	Flat-sky
1, $f_{\mathrm{SKY}} = 2.42\ 10^{-5}$	6 s	2635 s	165 s
5, $f_{\mathrm{SKY}} = 1.21\ 10^{-4}$	6 s	2371 s	167 s
15000, $f_{\mathrm{SKY}} = 0.36$	6 s	75 s	166 s

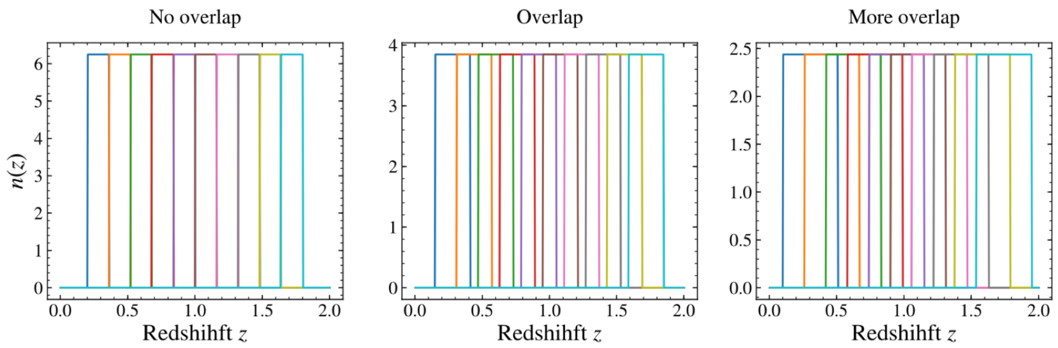
Partial-sky :

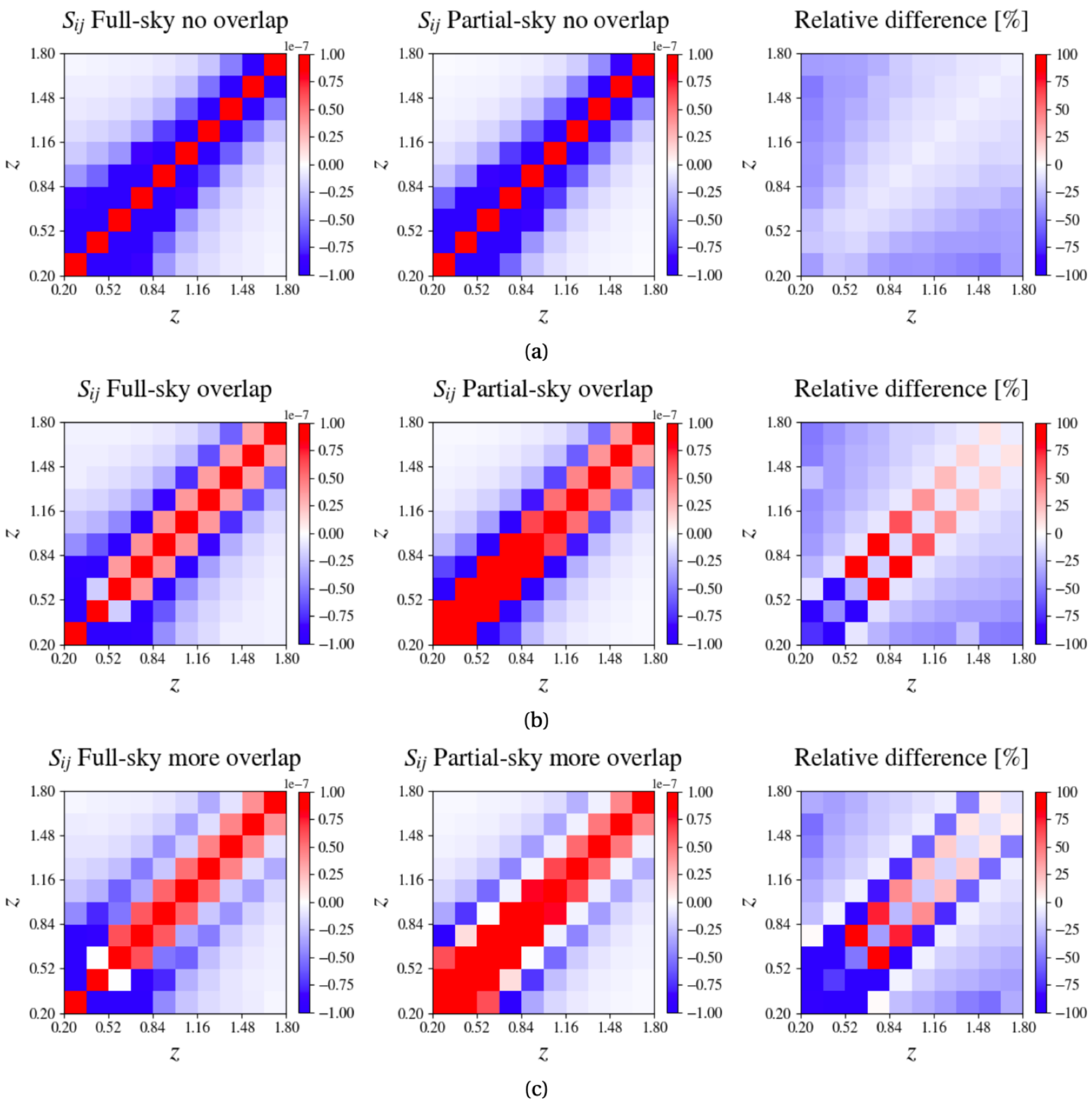
$$\sigma^2_{\mathrm{part-sky}}(z_1, z_2) = \frac{1}{\Omega_S^2} \sum_\ell (2\ell+1) C^{\mathcal{W}}(\ell) C^{\mathrm{m}}_{z_{1,2}}(\ell),$$

$$\langle \mathcal{W}^2 \rangle = \sum_\ell \frac{(2\ell+1)}{4\pi} C^{\mathcal{W}}(\ell).$$



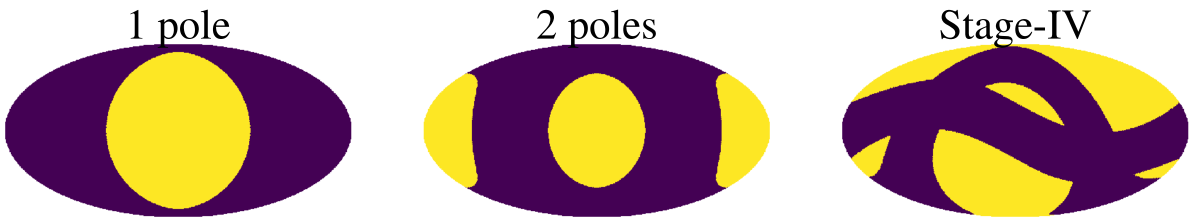
SSC : Redshift overlap





Partial-sky SSC : varying geometry

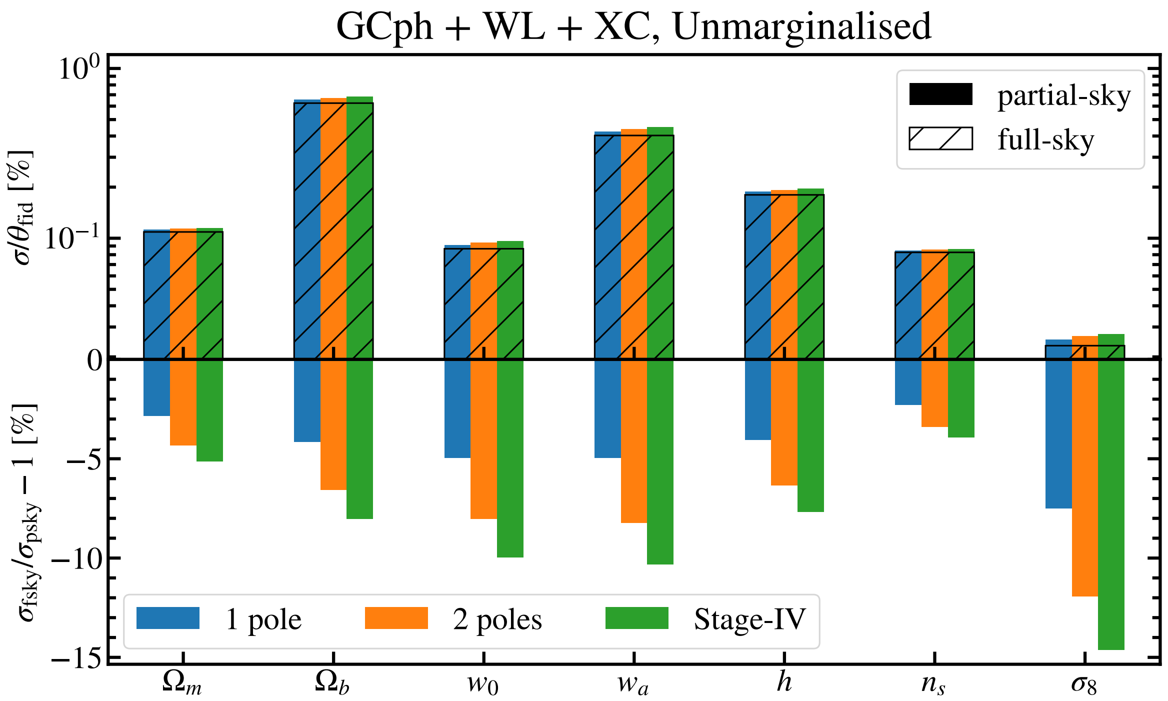
Geometry of the survey ?

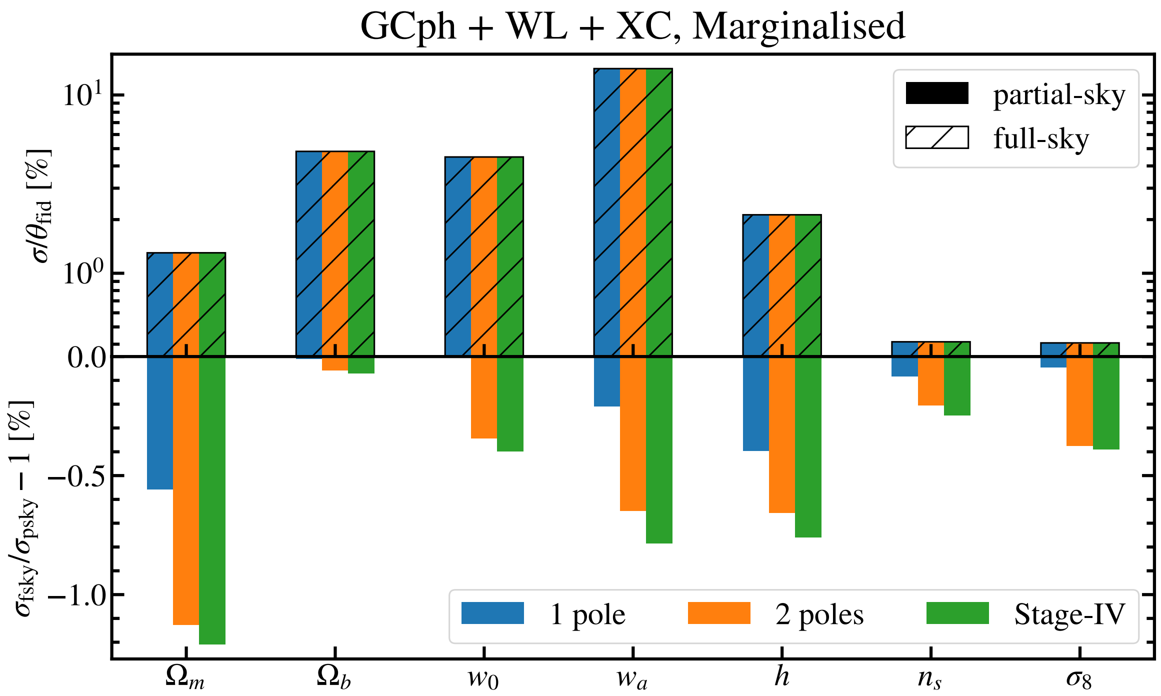


For a fixed survey area of 15 000 deg²

- **Top** : Relative error for full-sky and partial-sky
- **Bottom** : Relative difference between the two

- Difference increases for a more complex footprint
- Negligible difference for marginalised errors





$$n^{\rm true}(z) \propto \left(\frac{z}{z_0}\right)^2 \exp\left[-\left(\frac{z}{z_0}\right)^{3/2}\right],$$

$$n_i(z) = \frac{\int_{z_i^-}^{z_i^+} dz_p n^{\rm true}(z) p_{ph}(z_p|z)}{\int_{z_{min}}^{z_{max}} dz \int_{z_i^-}^{z_i^+} dz_p n^{\rm true}(z) p_{ph}(z_p|z)},$$

$$p_{ph}(z_p|z) = \frac{1-f_{out}}{\sqrt{2\pi}\sigma_b(1+z)} \exp\left\{-\frac{1}{2}\left[\frac{z-c_b z_p - z_b}{\sigma_b(1+z)}\right]^2\right\} + \frac{f_{out}}{\sqrt{2\pi}\sigma_o(1+z)} \exp\left\{-\frac{1}{2}\left[\frac{z-c_o z_p - z_o}{\sigma_o(1+z)}\right]^2\right\}.$$

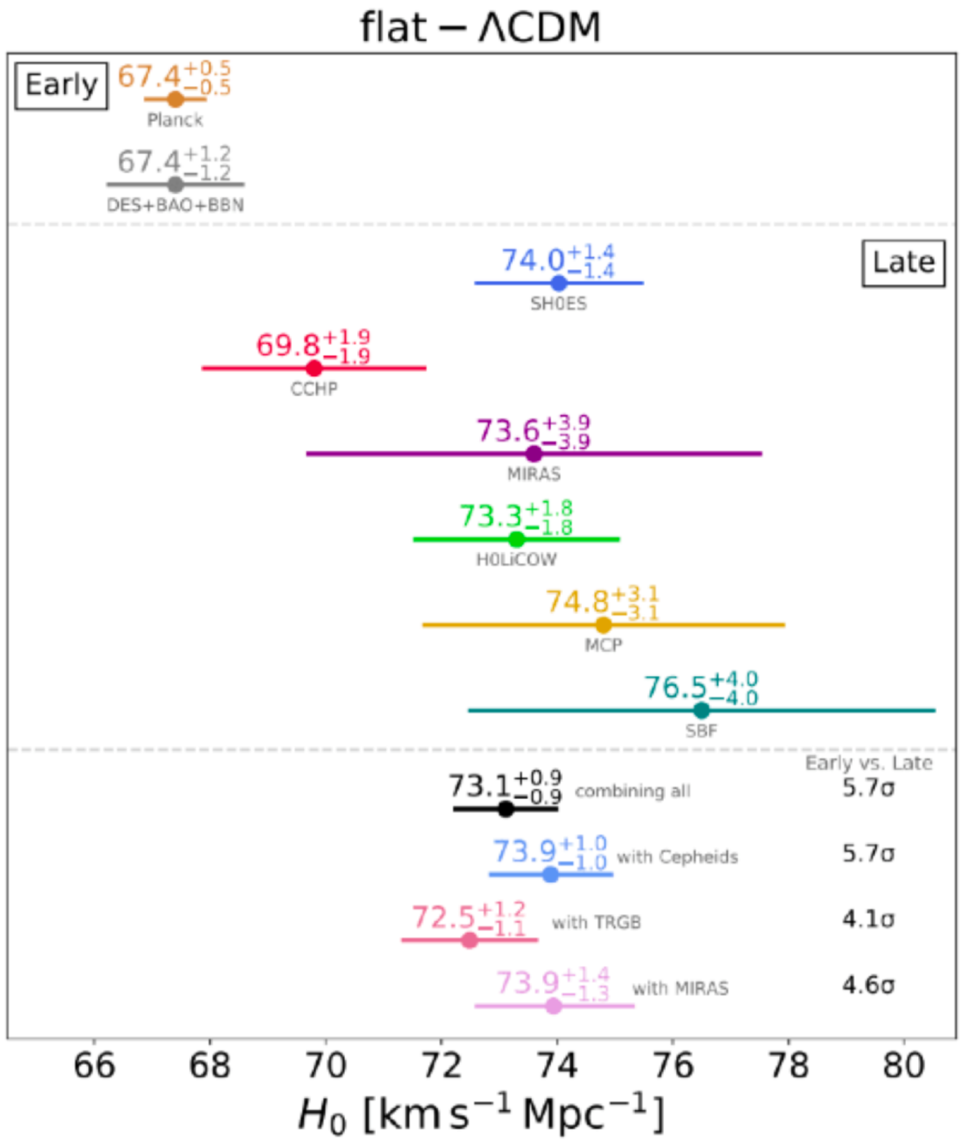
$n(z)$	c_b	z_b	σ_b	c_0	z_0	σ_0	f_{out}
Wide, $f_{out} = 0.1$	1.0	0.0	0.05	1.0	0.1	0.05	0.1
Wide, $f_{out} = 0.25$	1.0	0.0	0.05	1.0	0.1	0.05	0.25
Tight, $f_{out} = 0.1$	1.0	0.0	0.02	1.0	0.1	0.02	0.1

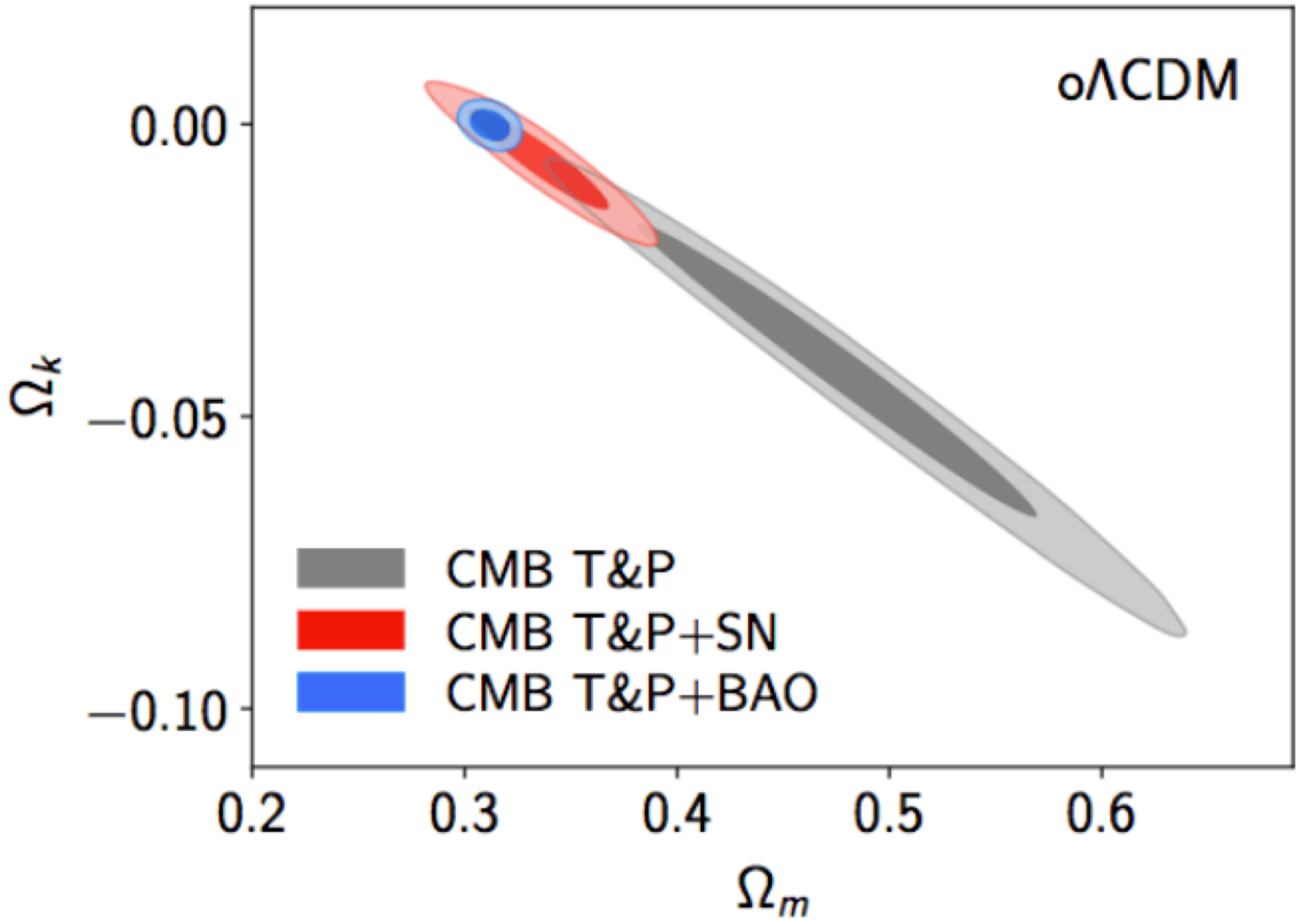
$$\bar{n}_i = 30 \text{ gal/arcmin}^2$$

H_0 ~ 67 +/- 0.5 *VS* 73 +/- 1 km/s/Mpc

Late time probes (CMB, BAO+BBN) — Early time probes (Cepheids, TGRB)

~ 4 to 5 σ tension within ΛCDM




oΛCDM
0.00
−0.05
−0.10
Ω_k
0.2
0.3
0.4
0.5
0.6
Ω_m
CMB T&P
CMB T&P+SN
CMB T&P+BAO

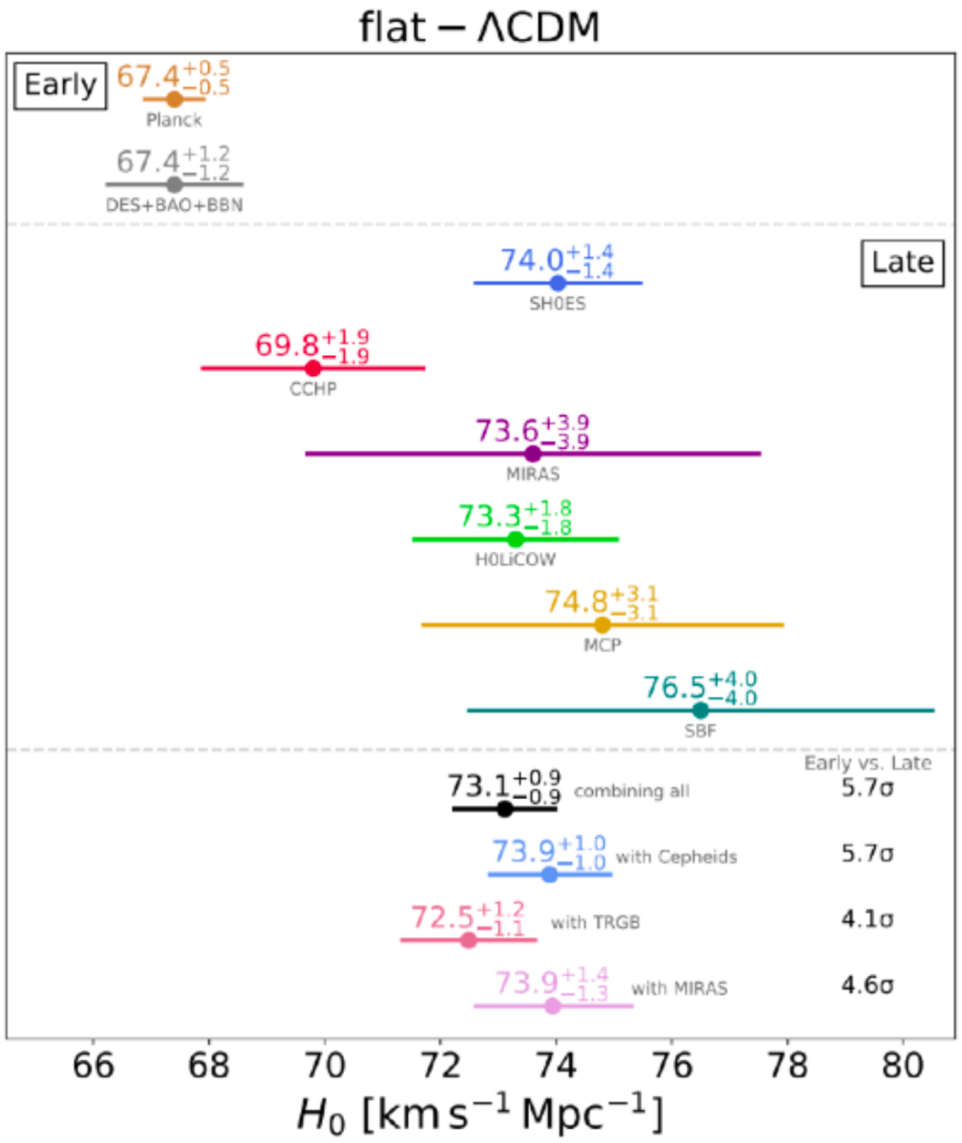
eBOSS 2021

H_0 tension

H_0 ~ 67 +/- 0.5 *VS* 73 +/- 1 km/s/Mpc

Late time probes (CMB, BAO+BBN) | Early time probes (Cepheids, TGRB)

~ 4 to 5 σ tension within ΛCDM



Observational pillars of ΛCDM

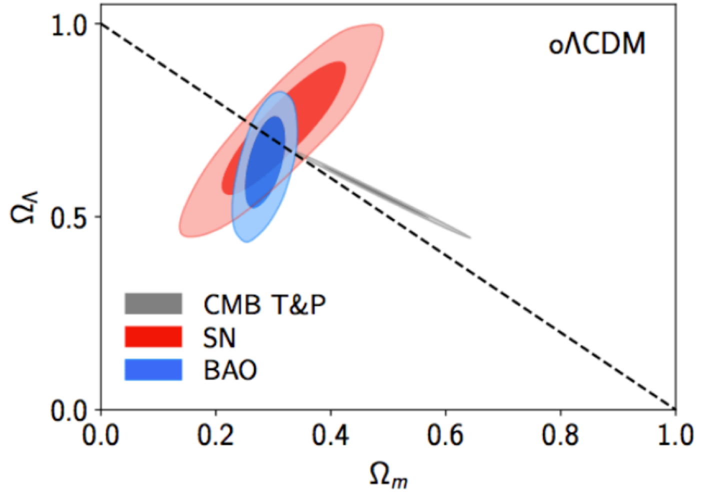
- Redshift of galaxies $1 + z \equiv \frac{\lambda_0}{\lambda} = \frac{a_0}{a}$
- Big Bang Nucleosynthesis (BBN)
- Cosmic Microwave Background (CMB)
- Galaxy rotation curves
Velocity dispersion in clusters + Lensing
- Type Ia Supernovae (SNIa)
- Baryon acoustic peak (BAO)

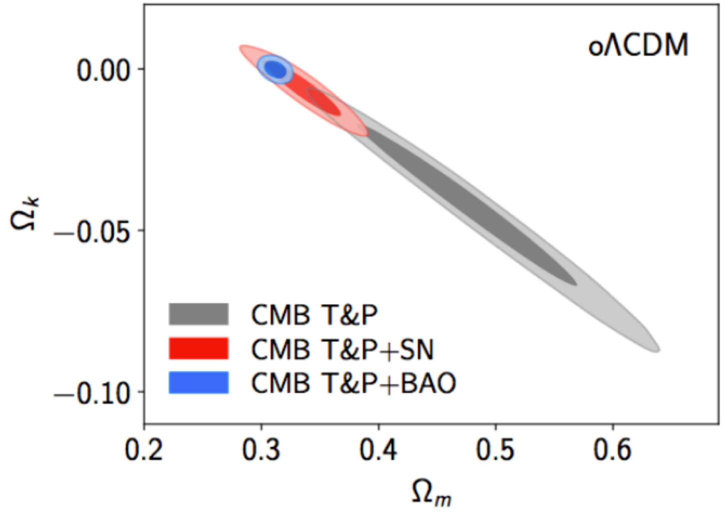
Expansion of the universe from a Big Bang
Suppression of Λ

Cold Dark Matter (CDM)

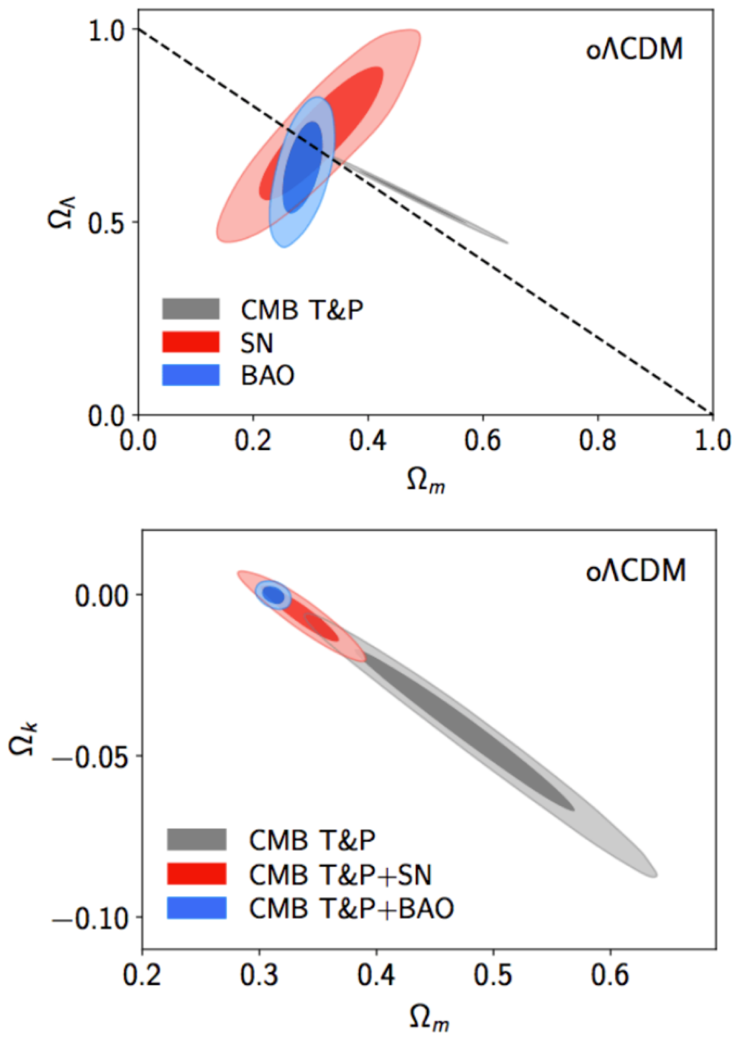
Accelerated expansion
Reintroduction of Λ

The ΛCDM model





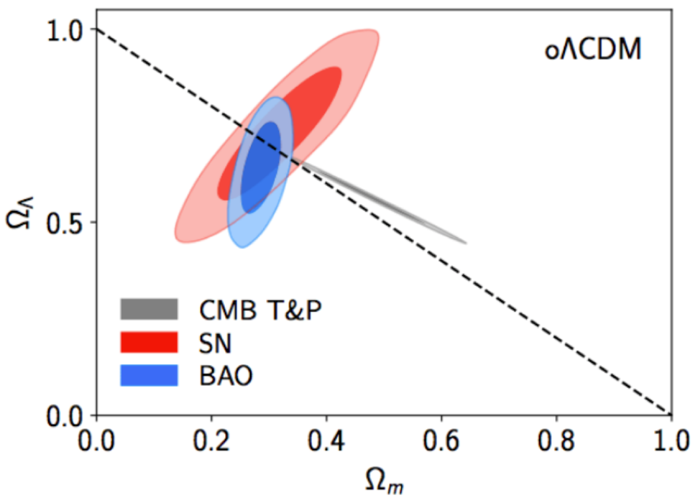
The ΛCDM model

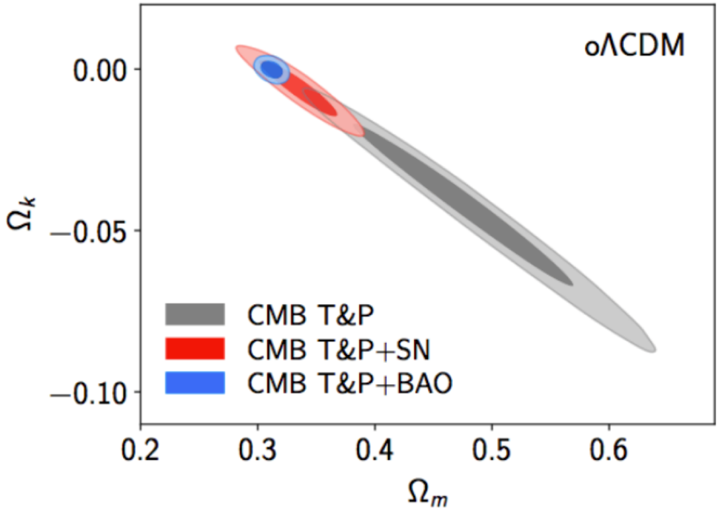


- **5 % of baryonic matter**, well described by Particle Physics
- **25 % of CDM** : Non standard matter
- **70 % of Λ** or Dark Energy or Modified Gravity ?
- **Spatially flat ?** Not with CMB alone
- **H_0 = 67 or 73 km/s/Mpc ?** Early VS Late time probes
- **What about massive neutrinos ?** Important role for cosmology
- **Inflation theory ?** No observational signature

Still interesting physics to explore and understand !

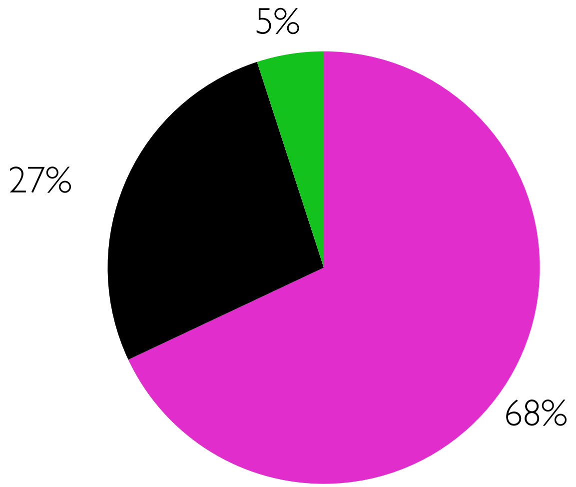
eBOSS collaboration 2021





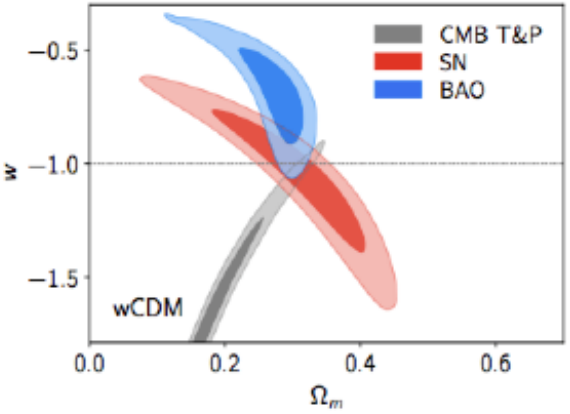
The universe is

- 5% of baryonic matter
  Well described by Particle Physics
- 27% of Cold Dark Matter
  Non standard matter
- 68% of Λ or Dark Energy
  Responsible for the accelaration
- Spatially Flat
  For CMB + SN + BAO



Dominated by the dark sector

- Λ
- Dark matter
- Baryonic matter



All probes are compatible with w=1
➔ Λ

- Explaining the acceleration of the expansion

The energy of vaccum does not fit Λ
120 orders of magnitude difference with theoretical prediction

- Tension on today's expansion rate

H_0 ~ 67 +/- 0.5 *VS* 73 +/- 1 km/s/Mpc

Late time probes (CMB, BAO+BBN) — Early time probes (Cepheids, TGRB)

~ 4 to 5 σ tension within ΛCDM

Motivation to go → Beyond ΛCDM

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu}$$

Modified gravity models ?

Look for deviations from Einstein's Field Equations

Additional fluid : Dark energy ?

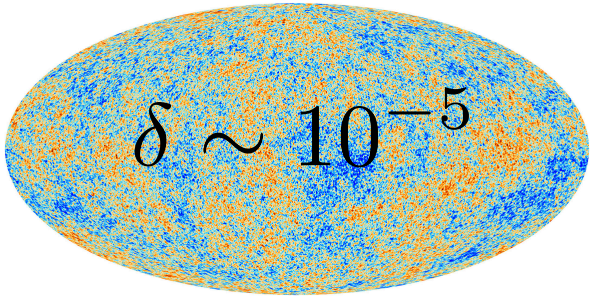
Look for deviations from

$$w = -1$$

$$w_{\mathrm{DE}}(a) = w_0 + w_a(1-a)$$

Need precise measurement of cosmological parameters to observe deviations from ΛCDM

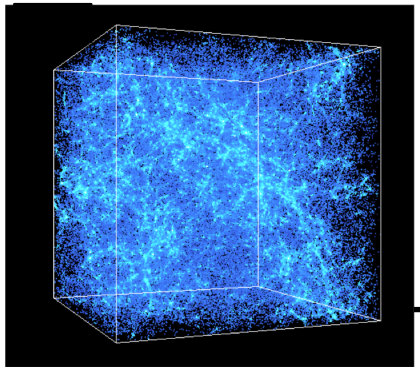
z = 1100



CMB anisotropies = seeds

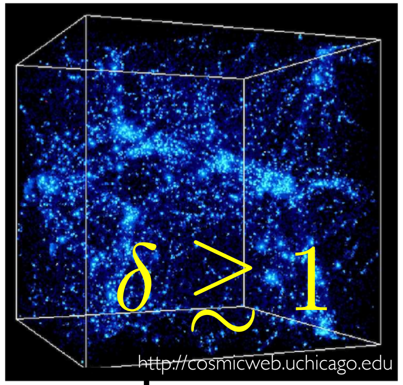
Linear clustering

z = 5



Non-linear clustering on small scales

z = 0



The density contrast field

$$\delta(\boldsymbol{x}, t) \equiv \frac{\rho(\boldsymbol{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}$$

In Fourier space : The power spectrum

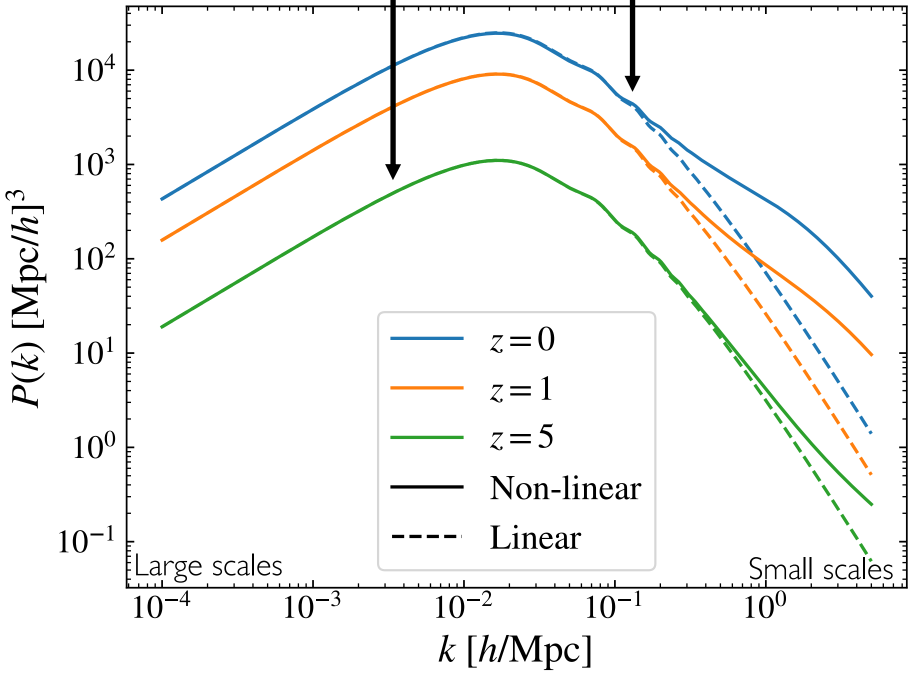
$$\langle \delta(\mathbf{k}) \delta(\mathbf{k}') \rangle = \delta_{\mathrm{D}}(\mathbf{k} + \mathbf{k}') P(k)$$

Statistical description

- By definition : $\langle \delta(\boldsymbol{x}) \rangle = 0$
- 2-point correlation function :

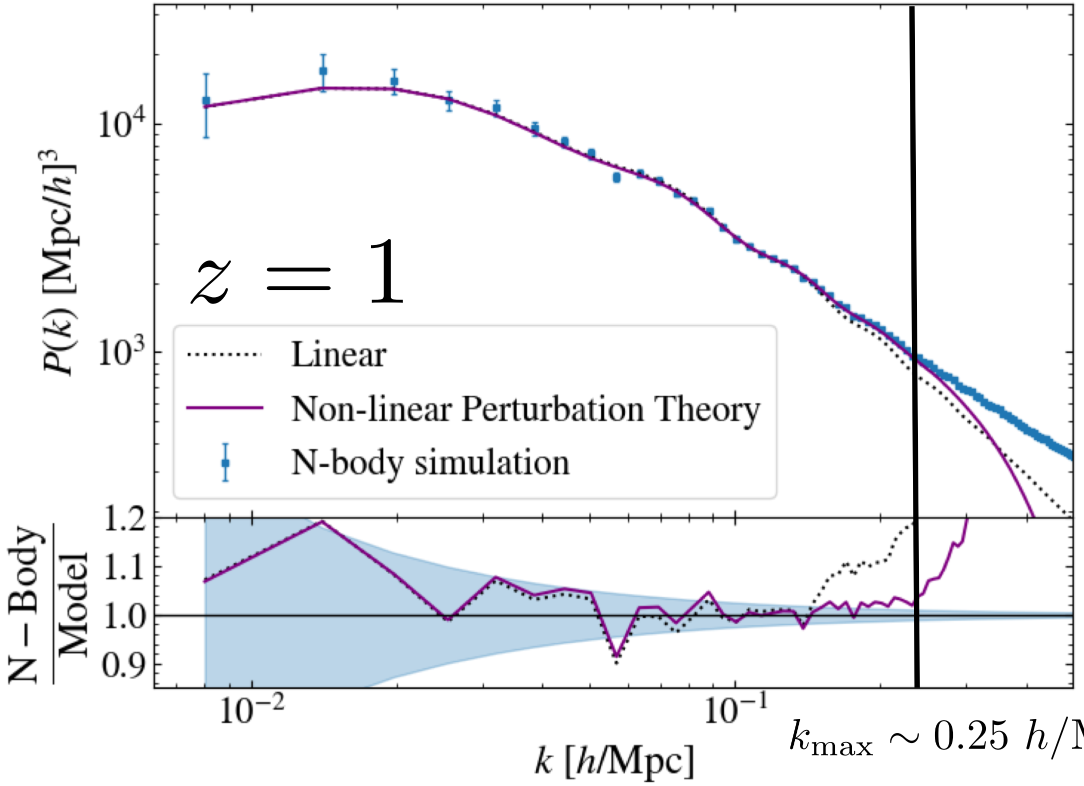
$$\langle \delta(\boldsymbol{x}) \delta(\boldsymbol{x} + \boldsymbol{r}) \rangle \equiv \xi(\boldsymbol{r})$$

$$k = \frac{2\pi}{r}$$



CMB ➔ δ follows a Gaussian PDF → Non-linear clustering on small scales → δ becomes non-Gaussian

Simulate non-linear clustering with N-Body simulations ➔ Test models on N-body simulations



$k_{\rm max} \sim 0.25\ h/{\rm Mpc}$

Covariance : Errors and correlations in the power spectrum [Scoccimaro et al. 1999]

$$\mathbf{C}(k_i, k_j) = \frac{P(k_i)^2}{N_{k_i}}\delta_{ij} + \bar{T}(k_i, k_j)$$

Trispectrum

$$\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\delta(\mathbf{k}_4)\rangle_c = \delta_{\rm D}(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3+\mathbf{k}_4)T(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$$

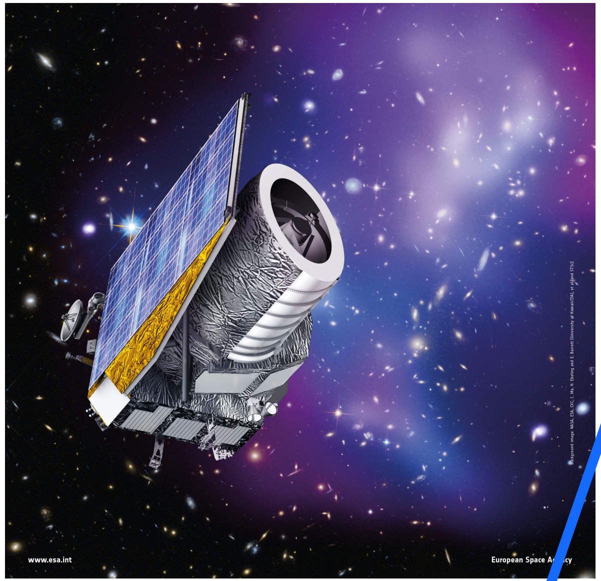
Difficult to predict analytically

Estimate the covariance from simulations

$$\hat{C}_{ij} = \frac{1}{N_m - 1}\left[\sum_n^{N_m}[P^{(n)}(k_i) - \bar{P}(k_i)][P^{(n)}(k_j) - \bar{P}(k_j)]\right]$$

The Euclid galaxy survey

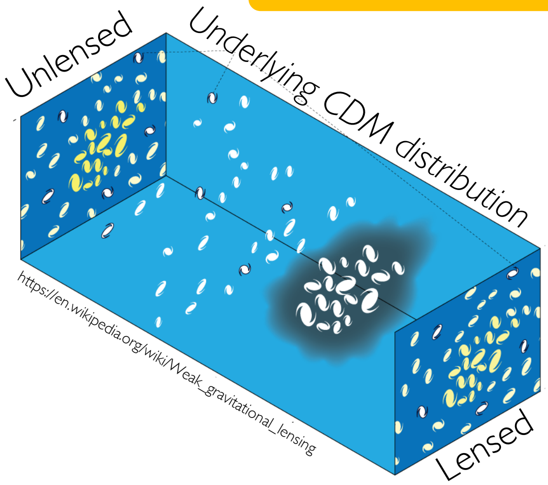
The Euclid mission



- VIS (visible) : Shapes of **10^9** with **Photo-z** $\in [0.001, 2.5]$
- NISP (infrared) : Spectrum of **$50 \text{x} 10^6$** galaxies with **Spectro-z** $\in [0.9, 1.8]$

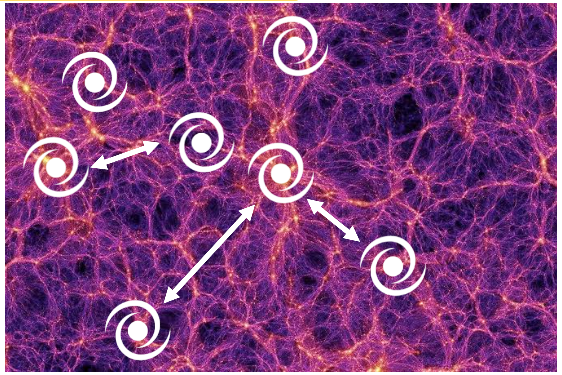
Main Probes of the LSS

Shape of galaxies : **Weak Lensing**



3D map of the position of galaxies

Galaxy Clustering



Scientific objectives

- Time dependence of Dark Energy

$$w_{\rm DE}(a) = w_0 + w_a(1-a)$$

- Modifications of gravity through the growth of structures

- Measuring the total neutrino mass

$$M_\nu \equiv \sum m_\nu$$

- **Early universe** (Inflation, primordial non-Gaussianities)

NISP = Near Infrared Spectrometer and Photometer
VIS = VISible imager

Euclid IST:Forecast : [Euclid Collaboration 2019]

Table 3: Expected number density of observed Hα emitters for the *Euclid* spectroscopic survey. This number has been updated since the Red Book (Laureijs et al. 2011) to match new observations of number densities and new instrument and survey specifications. The first two columns show the minimum, $z_{\rm min}$, and maximum, $z_{\rm max}$, redshifts of each bin. The third column is the number of galaxies per unit area and redshift intervals, $dN(z)/d\Omega dz$. The fourth column shows the number density, $n(z)$. The fifth column lists the total volume. Finally, in the sixth column we list the galaxy bias evaluated at the central redshift of the bins, $z_{\rm mean} = (1/2)(z_{\rm max} + z_{\rm min})$.

$z_{\rm min}$	$z_{\rm max}$	$dN(z_{\rm mean})/d\Omega dz$ [deg^{-2}]	$n(z_{\rm mean})$ [h^3 Mpc^{-3}]	$V_s(z_{\rm mean})$ [Gpc^3 h^{-3}]	$b(z_{\rm mean})$
0.90	1.10	1815.0	6.86×10^{-4}	7.94	1.46
1.10	1.30	1701.5	5.58×10^{-4}	9.15	1.61
1.30	1.50	1410.0	4.21×10^{-4}	10.05	1.75
1.50	1.80	940.97	2.61×10^{-4}	16.22	1.90

Table 4: Specifications for the *Euclid* photometric weak lensing survey.

	Parameter	*Euclid*
Survey area in the sky	$A_{\rm survey}$	15 000 deg^2
Galaxy number density	$n_{\rm gal}$	30 arcmin^{-2}
Total intrinsic ellipticity dispersion	σ_ϵ	0.30
Number of redshift bins	N_z	10