Projection effects in likelihood analyses of galaxy clustering

Andrea Pezzotta (MPE)

with Benjamin Camacho, Martin Crocce, Alexander Eggemeier, Giosue Gambardella, Ariel Sanchez, Roman Scoccimarro, Agne Semenaite

Sesto Workshop, 01/07/2024

Outline

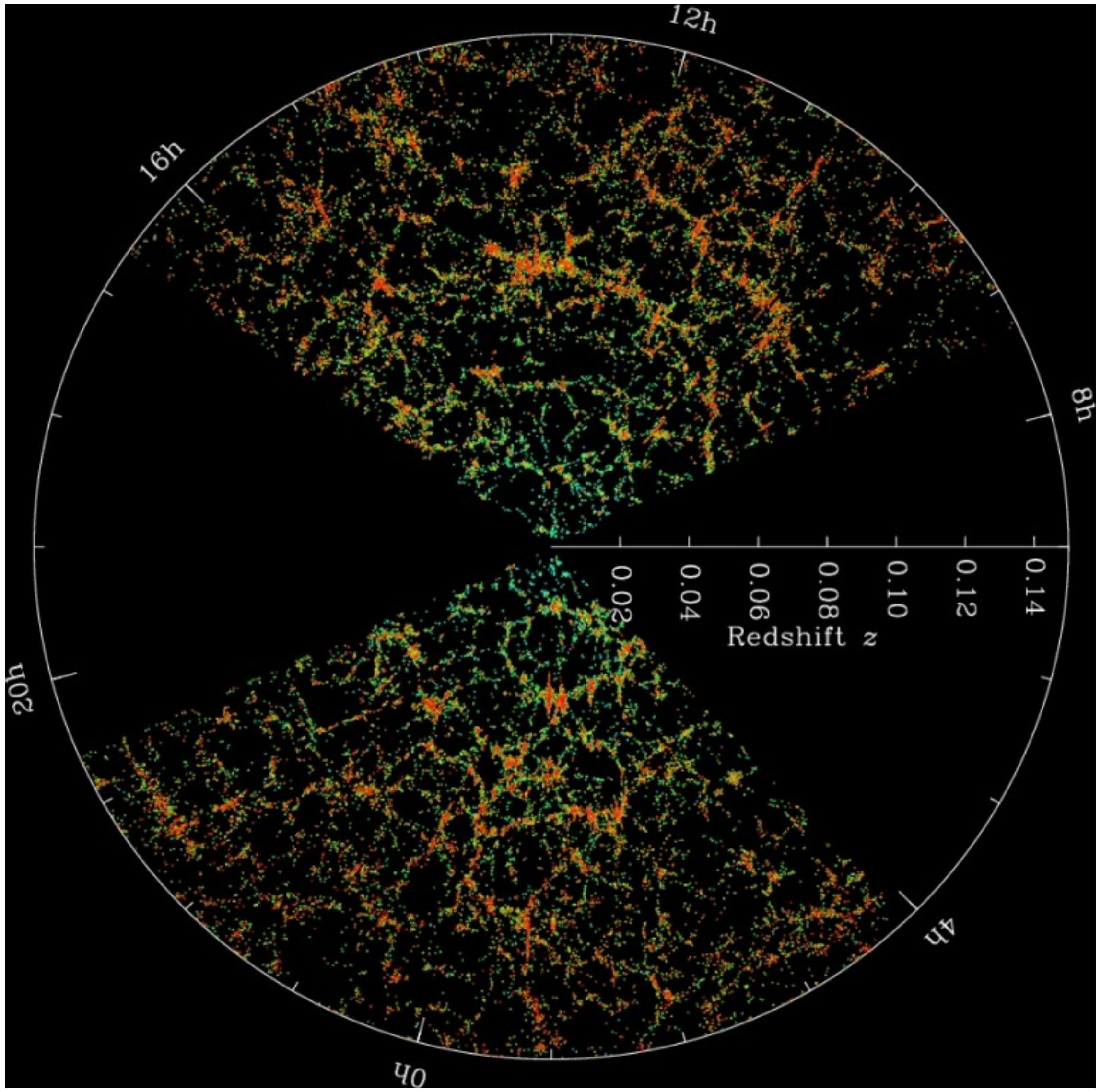
- Cosmology from galaxy clustering
- State-of-art models for galaxy clustering
- Faster model evaluation with evolution mapping and extension to massive neutrinos
- Prior volume effects in cosmological parameter inference
- Conclusions

Galaxy clustering

Why galaxy clustering?

One of the two main cosmological probes to characterise the Large Scale Structure

Galaxy clustering



(SDSS)

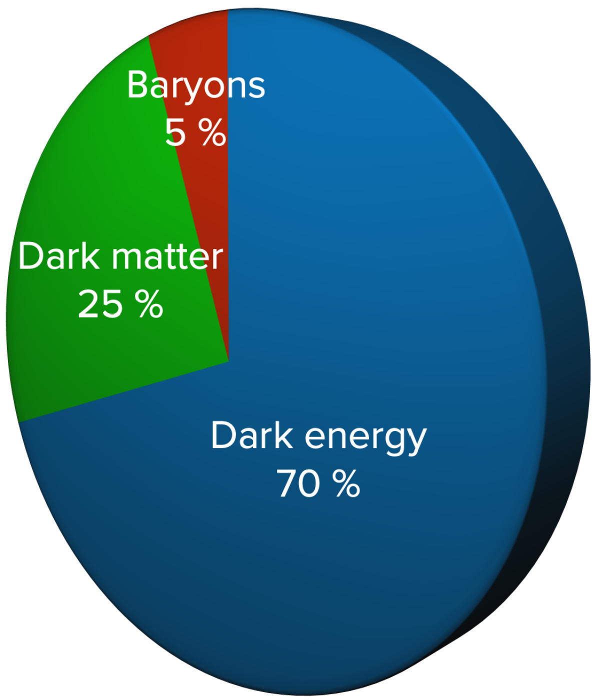
Weak gravitational lensing

(HST)

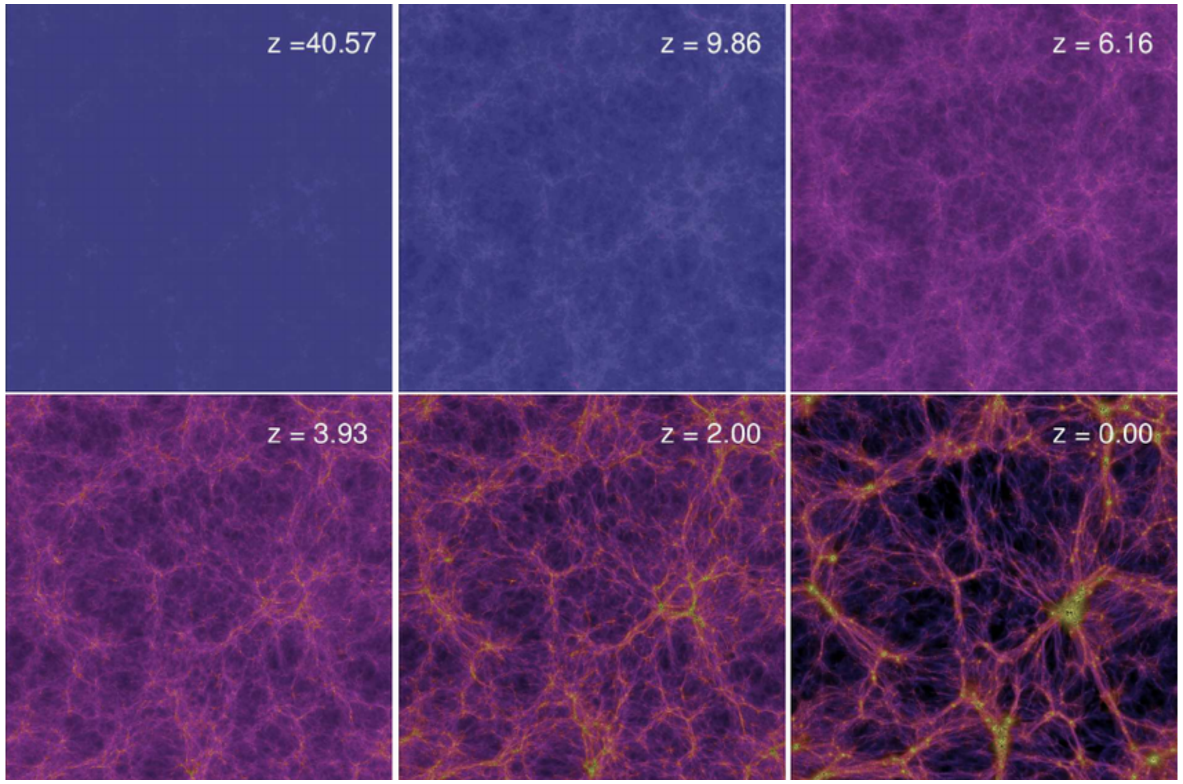
Why galaxy clustering?

One of the two main cosmological probes to characterise the Large Scale Structure

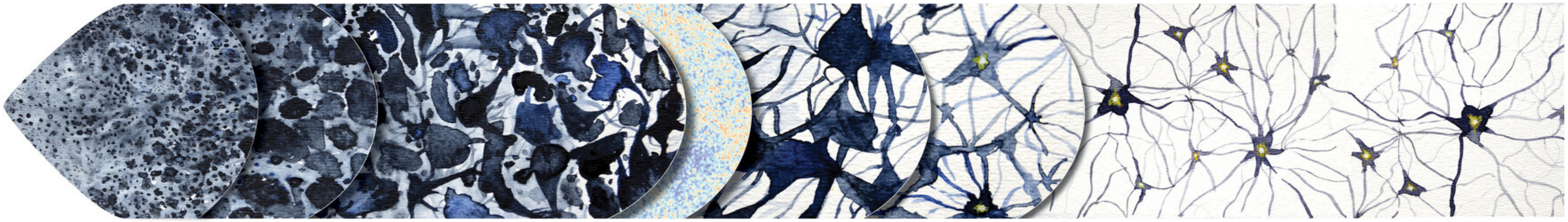
Energy/matter components



Growth of structures



Expansion history



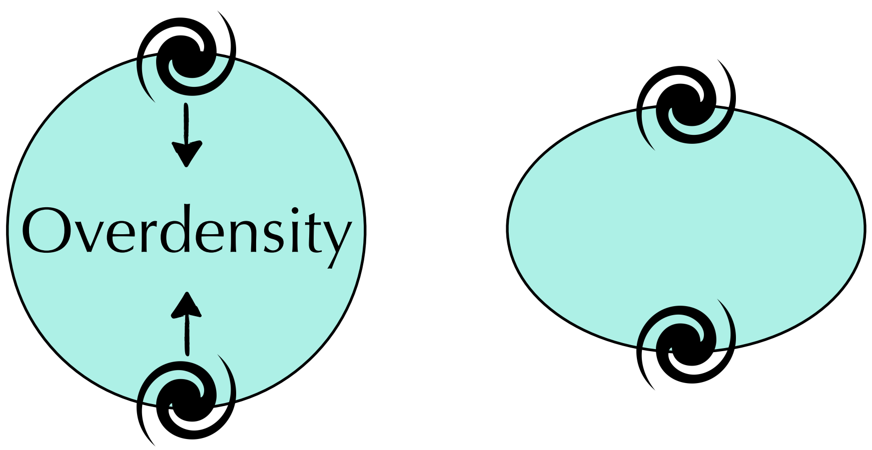
Redshift-space distortions

Peculiar velocities distort the observed distance inferred from the redshift of luminous sources

$$\vec{s} = \vec{r} + \frac{v_{\parallel}}{a\,H(a)}\hat{e}_r$$

Traditionally split into 2 different regimes (but cannot be treated as independent)

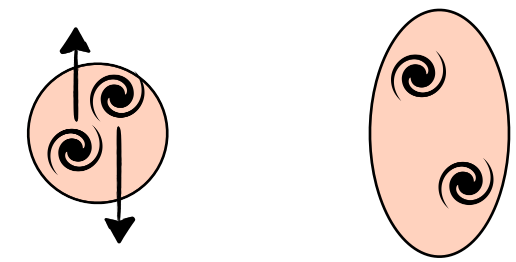
Linear regime ($\gg 1\,h^{-1}\,\mathrm{Mpc}$)



Kaiser effect

Flattening of large overdense regions due to infall towards overdose regions

Nonlinear regime ($\sim 1\,h^{-1}\,\mathrm{Mpc}$)



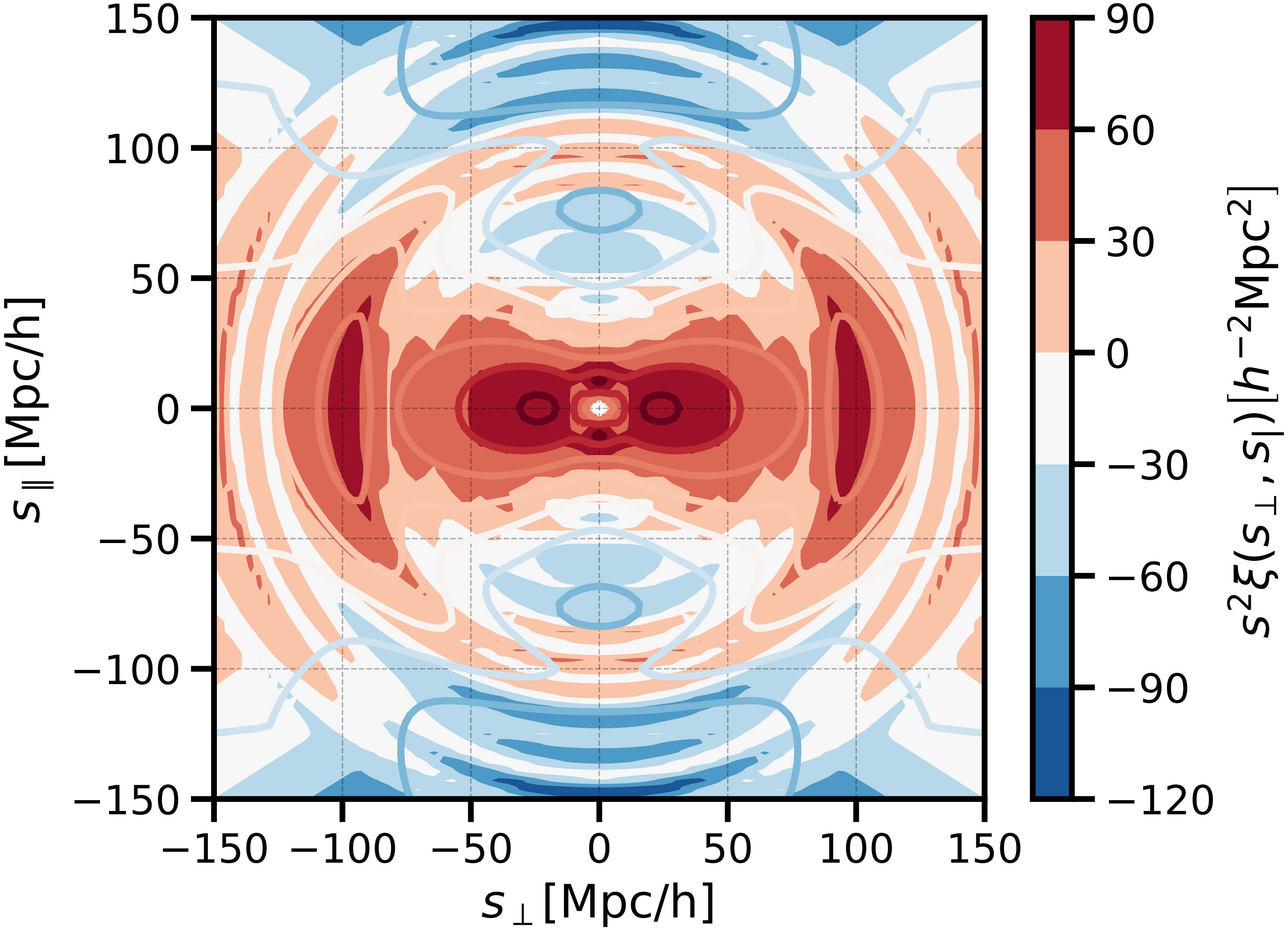
Fingers of God

Stretching of virialised structures along the line-of-sight due to viral velocities

Clustering observables

Anisotropy observable from clustering measurements (e.g. 2PCF, galaxy power spectrum)

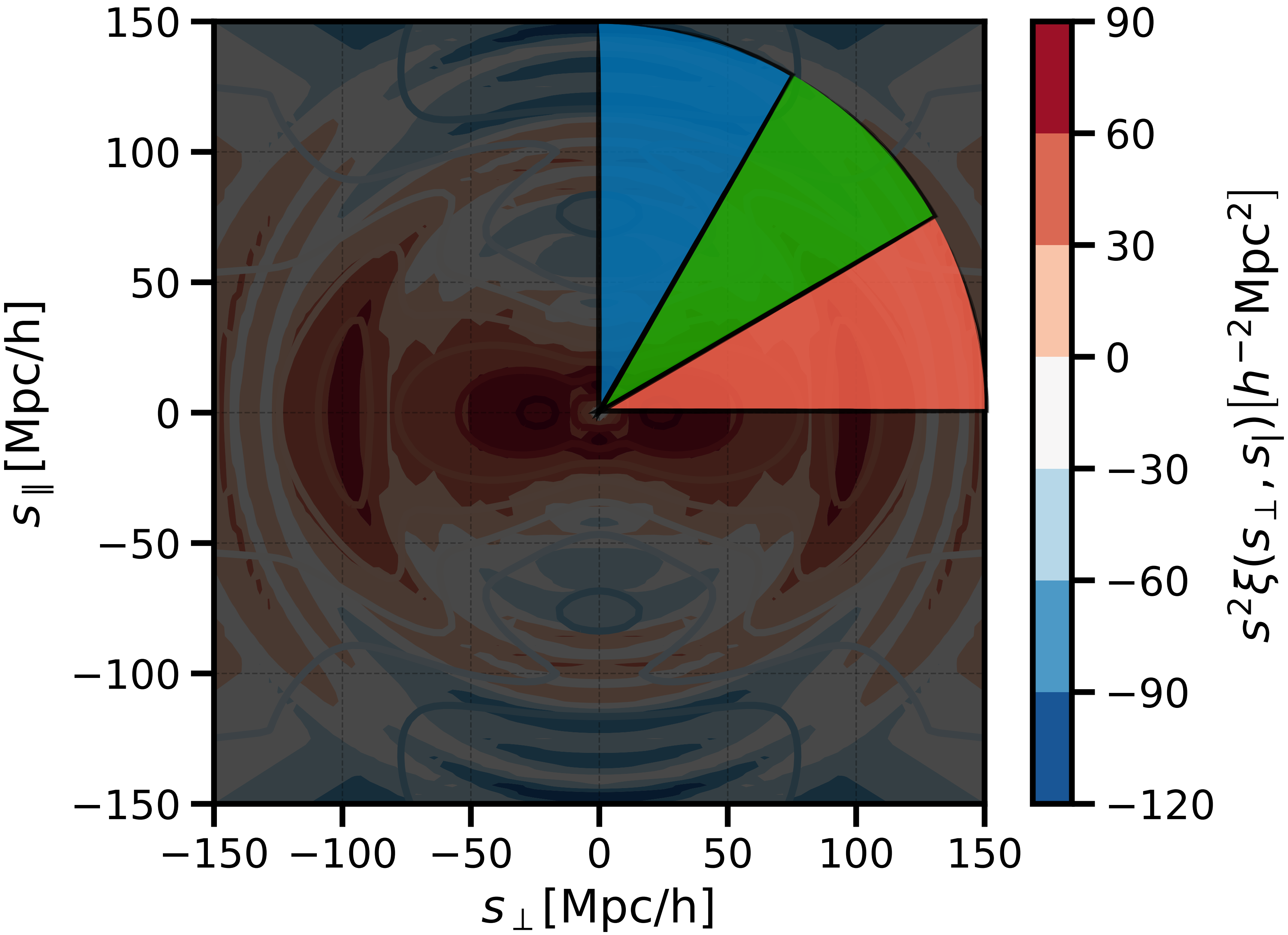
eBOSS LRGs (Hou et al. 2020)



Clustering observables

Anisotropy observable from clustering measurements (e.g. 2PCF, galaxy power spectrum)

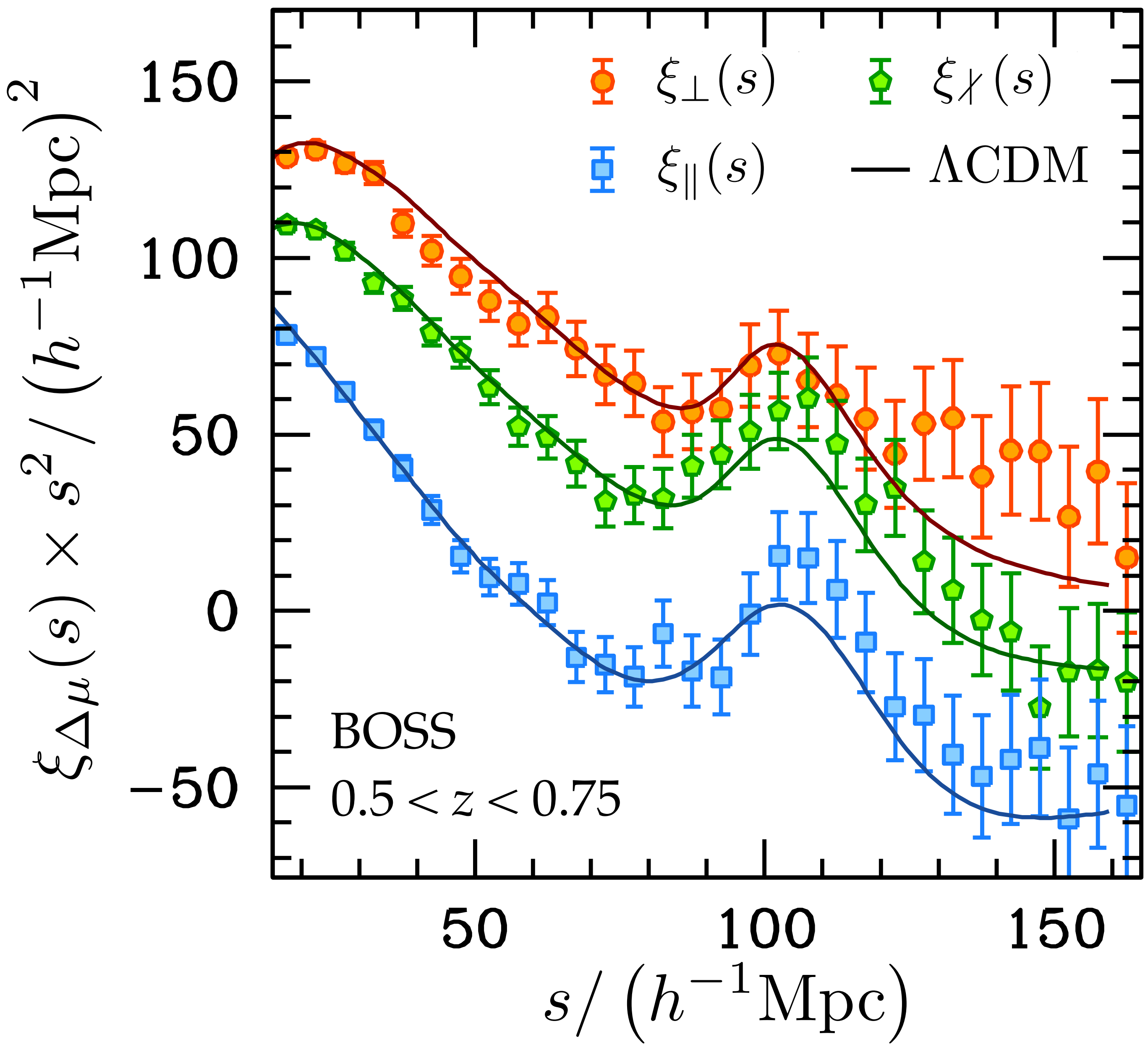
eBOSS LRGs (Hou et al. 2020)



Clustering observables

Standard analyses focus on the projection of anisotropic clustering measurements

(Sánchez et al. 2017a)



- Clustering wedges

$$\xi_{\Delta\mu}(s) = \frac{1}{\Delta\mu}\int_{\Delta\mu} \xi(s,\mu)\,\mathrm{d}\mu$$

- Legendre multipoles

$$\xi_\ell(s) = \frac{2\ell+1}{2}\int_{-1}^{+1} \xi(s,\mu)\,\mathscr{L}_\ell(\mu)\,\mathrm{d}\mu$$

Perturbation theory

Perturbative approaches can be employed to describe clustering measurements

$$\delta_{\rm g} = b_1\,\delta + b_{\nabla^2\delta}\nabla^2\delta + \varepsilon + \frac{b_2}{2}\delta^2 + b_{\mathcal{G}_2}\left[\left(\partial_i\,\partial_j\,\Phi\right)^2 - \left(\nabla^2\Phi\right)^2\right] + \ldots$$

- **Nonlinear evolution** of the matter density field
- **Galaxy bias** (relationship between galaxy and matter density field)
- **Redshift-space distortions**

Employed for the analysis of Stage-III data (e.g. BOSS)

Perturbation theory

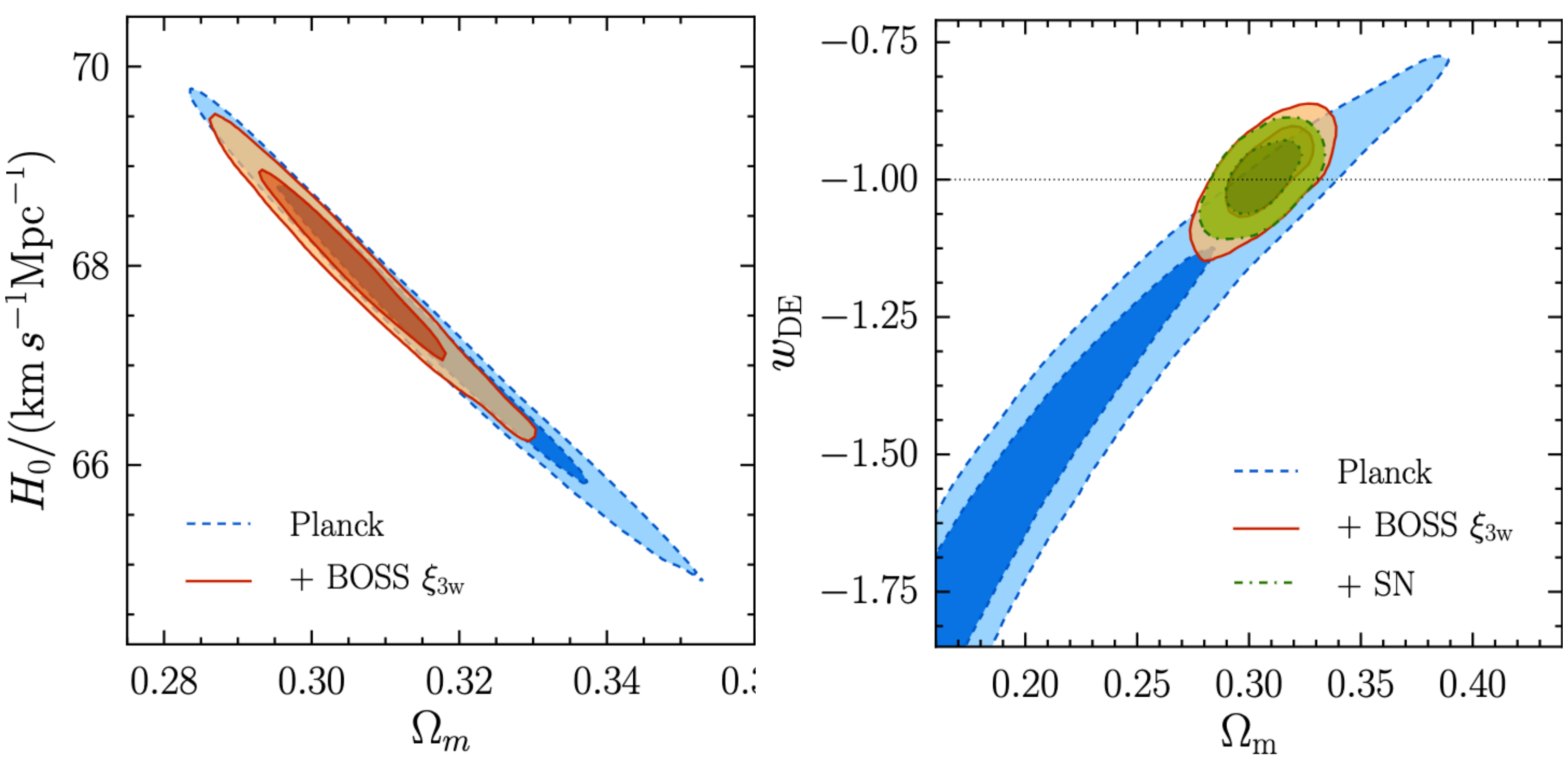
Perturbative approaches can be employed to describe clustering measurements

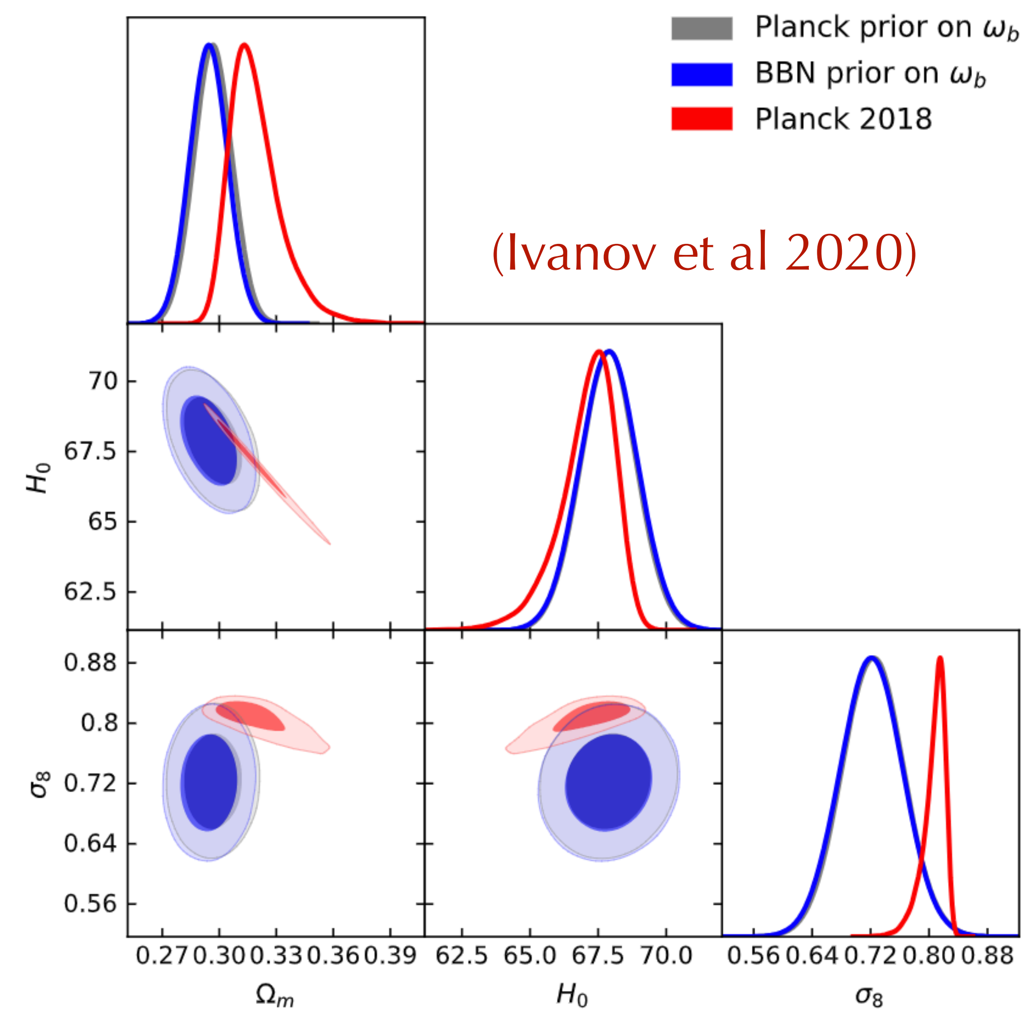
$$\delta_{\rm g} = b_1\,\delta + b_{\nabla^2\delta}\nabla^2\delta + \varepsilon + \frac{b_2}{2}\delta^2 + b_{\mathcal{G}_2}\left[\left(\partial_i\partial_j\Phi\right)^2 - \left(\nabla^2\Phi\right)^2\right] + \ldots$$

- **Nonlinear evolution** of the matter density field
- **Galaxy bias** (relationship between galaxy and matter density field)
- **Redshift-space distortions**

Employed for the analysis of Stage-III data (e.g. BOSS)

(Sánchez et al. 2017a)





(Ivanov et al 2020)

Perturbation theory

State-of-art modelling of redshift-space galaxy power spectrum: **EFTofLSS**

$$P^{s}_{\rm gg}(k,\mu) = P^{s,\,\rm tree}_{\rm gg}(k,\mu) + P^{s,\,\rm 1\text{-}loop}_{\rm gg}(k,\mu) + P^{s,\,\rm noise}_{\rm gg}(k,\mu) + P^{s,\,\rm ctr}_{\rm gg}(k,\mu)$$

Perturbation theory

State-of-art modelling of redshift-space galaxy power spectrum: **EFTofLSS**

$$P^s_{\rm gg}(k,\mu) = P^{s,\,\rm tree}_{\rm gg}(k,\mu) + P^{s,\,\rm 1\text{-}loop}_{\rm gg}(k,\mu) + P^{s,\,\rm noise}_{\rm gg}(k,\mu) + P^{s,\,\rm ctr}_{\rm gg}(k,\mu)$$

Tree level
(leading-order)
$\{\, b_1 \,\}$

Loops
(nonlinear corrections)
$\{\, b_2\,,\, b_{\mathcal{G}_2}\,,\, b_{\Gamma_3} \,\}$

Shot-noise
(high-frequency modes)
$\{\, \alpha_P\,,\, \epsilon_0\,,\, \epsilon_2 \,\}$

Counterterms
(FOGs, failure of SPT)
$\{\, c_0\,,\, c_2\,,\, c_4\,,\, c_{\rm nlo} \,\}$

Fully perturbative model including:

- **Local** and **non-local bias**
- **Shot-noise** corrections (constant and scale-dependent)
- **Counterterms** (from expanding pairwise velocity difference, i.e. FOGs; integration of UV modes, higher-derivatives, velocity bias)
- **Resummation of infrared modes** (need recipe to obtain the no-wiggle $P(k)$)
- 11 nuisance parameters, some of them highly degenerate in terms of $P^s_{\rm gg}(k,\mu)$

Speeding up model evaluation

Using PT: is it worth?

IT IS ACCURATE!

- Optimal performances over mildly nonlinear scales
- Flexible enough to accommodate different models
- Extensively employed to derive cosmological constraints from GC data

IT CAN ALSO BE SLOW…

- Standard bayesian inference can request up to $\mathcal{O}(10^6)$ likelihood evaluations
- Evaluation time for a single model can be relevant $\mathcal{O}(\mathrm{s})$
- Even more expensive when considering higher-order statistics
- Additional time lost to account for observational systematics

One solution (out of fews): we can make use of **EMULATORS**

Evolution Mapping (Sánchez et al. 2021)

Based on the split of cosmological parameters into **shape** and **evolution**

$\mathbf{\Theta}_\mathrm{s} = \{\omega_\mathrm{c}, \omega_\mathrm{b}, \omega_\gamma, n_\mathrm{s}, \ldots\}$

Shape of matter **transfer function** and **primordial power spectrum**

$\mathbf{\Theta}_\mathrm{e} = \{H_0, \omega_\mathrm{K}, \omega_\mathrm{DE}, w(a), A_\mathrm{s}, \ldots\}$

Overall **amplitude of power spectrum** at any redshift z

New parameter describing the **degeneracy among evolution parameters** (and redshift):

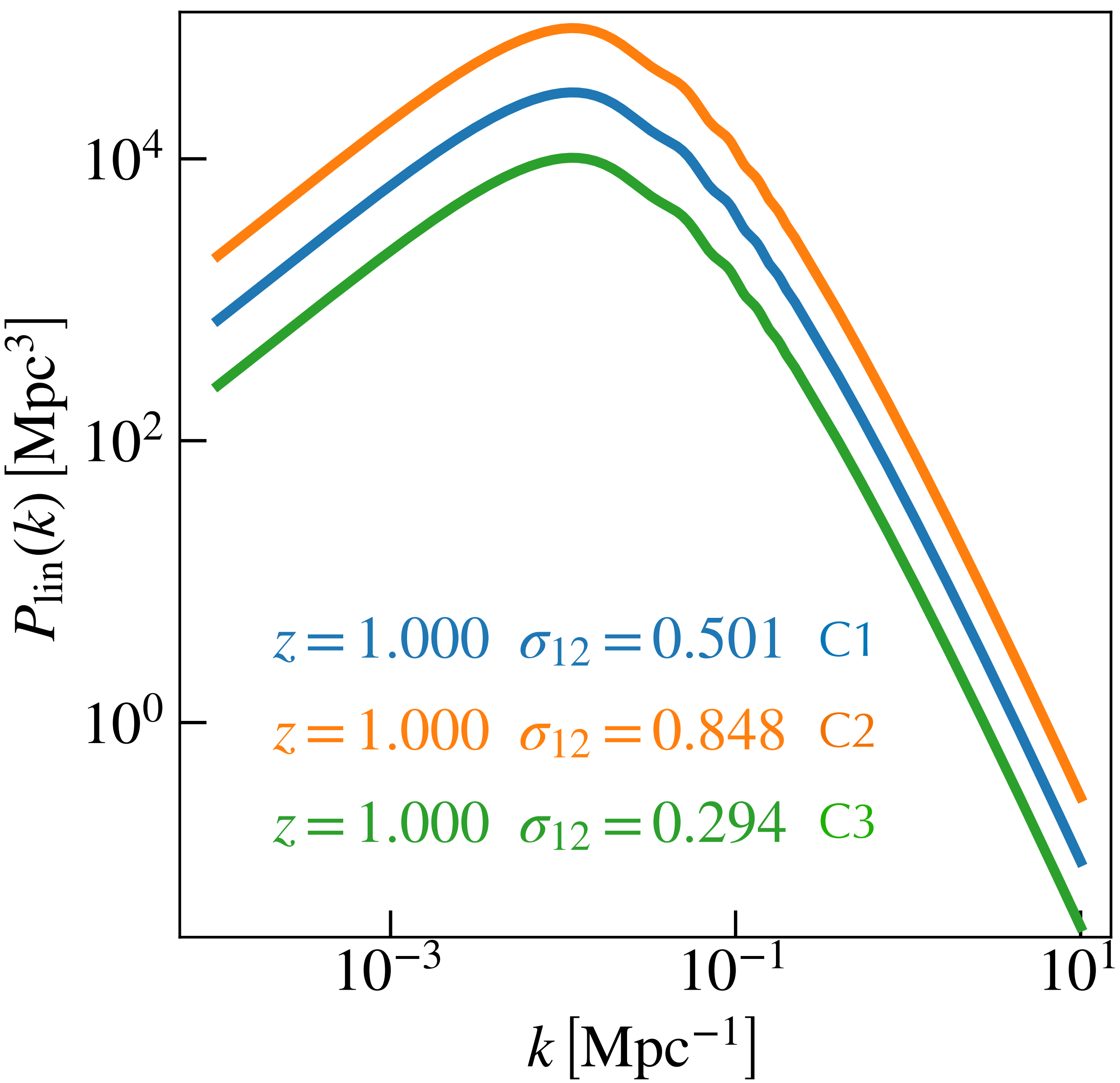
$$\sigma_{12} = \int \mathrm{d}k\, k^2\, P(k)\, W^2_\mathrm{TH}(kR) \Big|_{R=12\mathrm{Mpc}}$$

Same power spectrum when considering models with same shape and σ_{12}

$$P_\mathrm{lin}(k\,|\,\mathbf{\Theta}_\mathrm{s}, \mathbf{\Theta}_\mathrm{e}, z) = P_\mathrm{lin}(k\,|\,\mathbf{\Theta}_\mathrm{s}, \sigma_{12}(\mathbf{\Theta}_\mathrm{s}, \mathbf{\Theta}_\mathrm{e}, z))$$

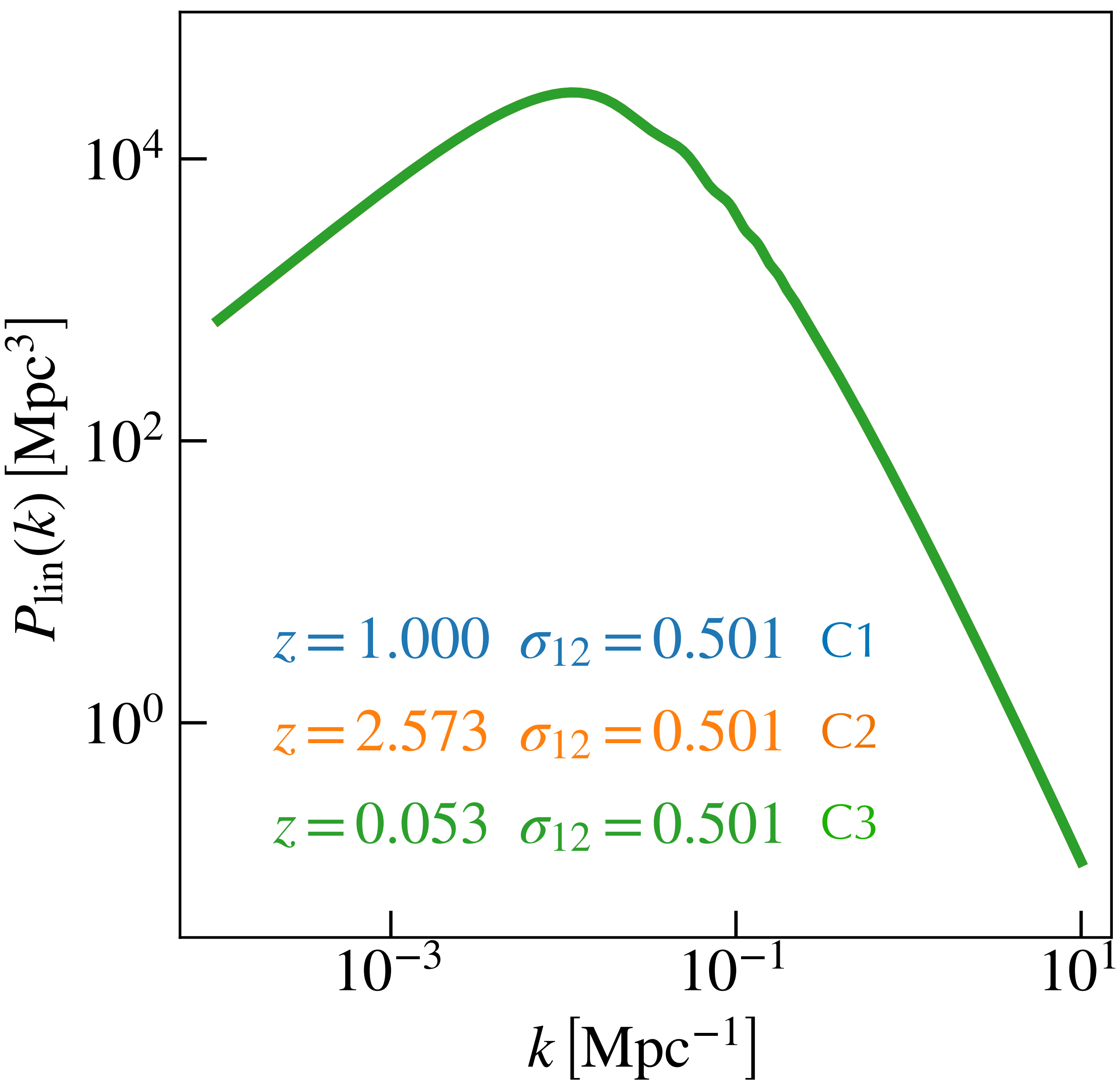
Evolution Mapping (Sánchez et al. 2021)

Let's see it directly with an example: three cosmologies with same $\boldsymbol{\Theta}_s$ but different $\boldsymbol{\Theta}_e$



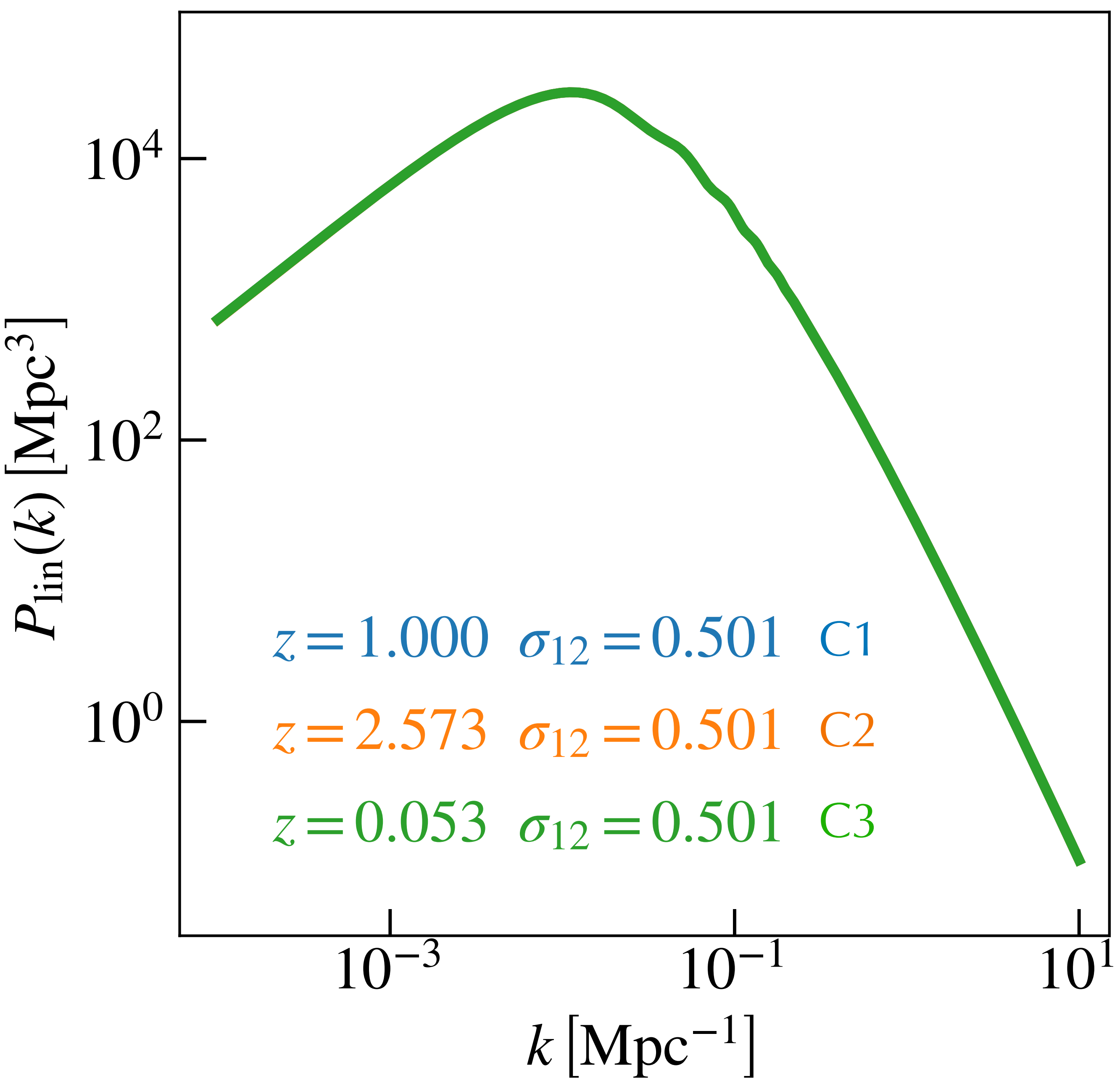
Evolution Mapping (Sánchez et al. 2021)

Let's see it directly with an example: three cosmologies with same $\mathbf{\Theta}_s$ but different $\mathbf{\Theta}_e$



Evolution Mapping (Sánchez et al. 2021)

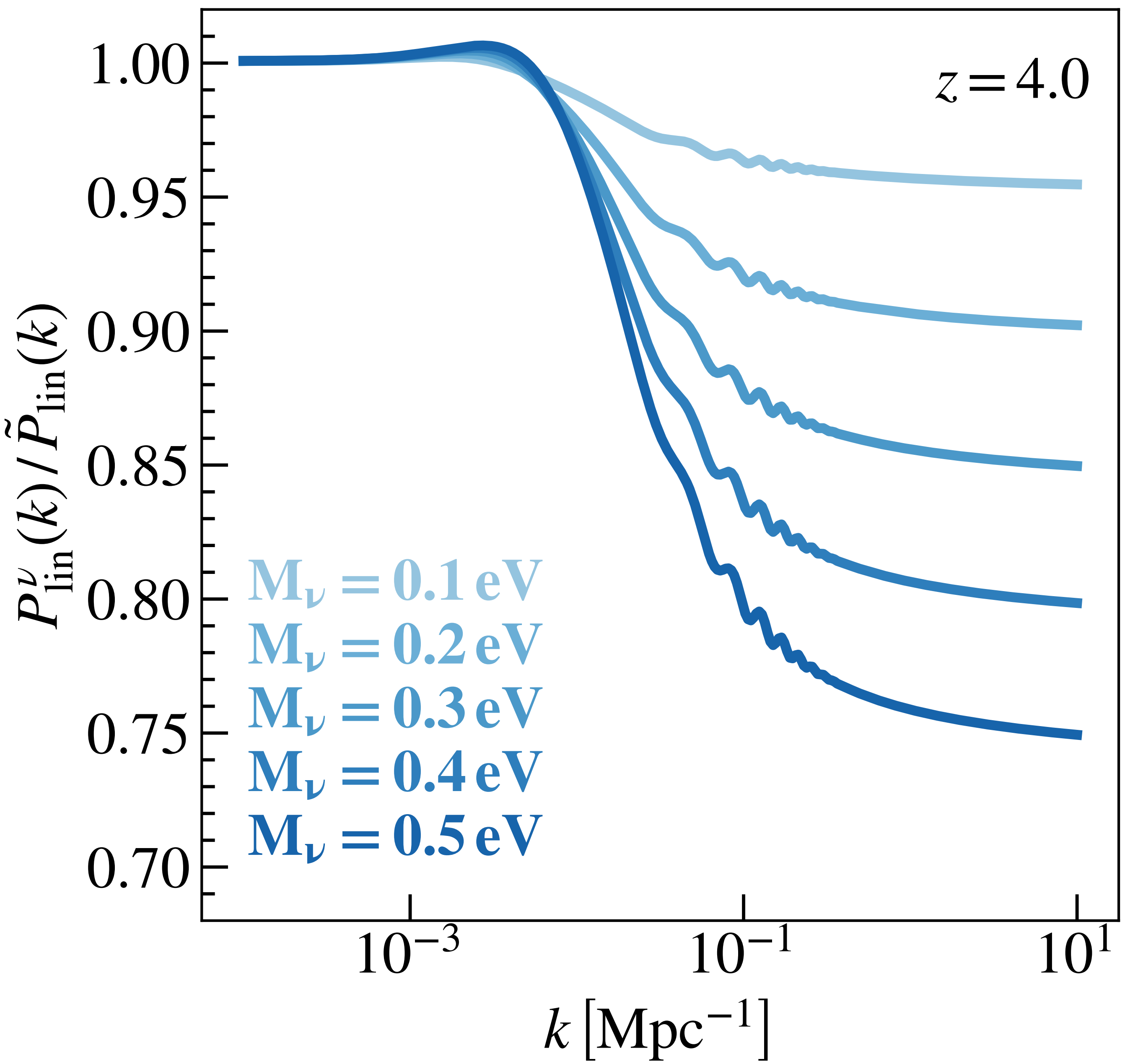
Let's see it directly with an example: three cosmologies with same $\mathbf{\Theta}_\mathrm{s}$ but different $\mathbf{\Theta}_\mathrm{e}$



- Mapping is propagated to the non-linear regime (need to correct for different history of structure formation, though)
- Applies to other statistics, like velocity power spectra (Esposito et al. 2024), higher-order statistics, halo mass functions, etc.
- Use of Mpc/h units to express distances breaks the degeneracy

Problems with massive neutrinos

Massive neutrinos break the original formulation of evolution mapping



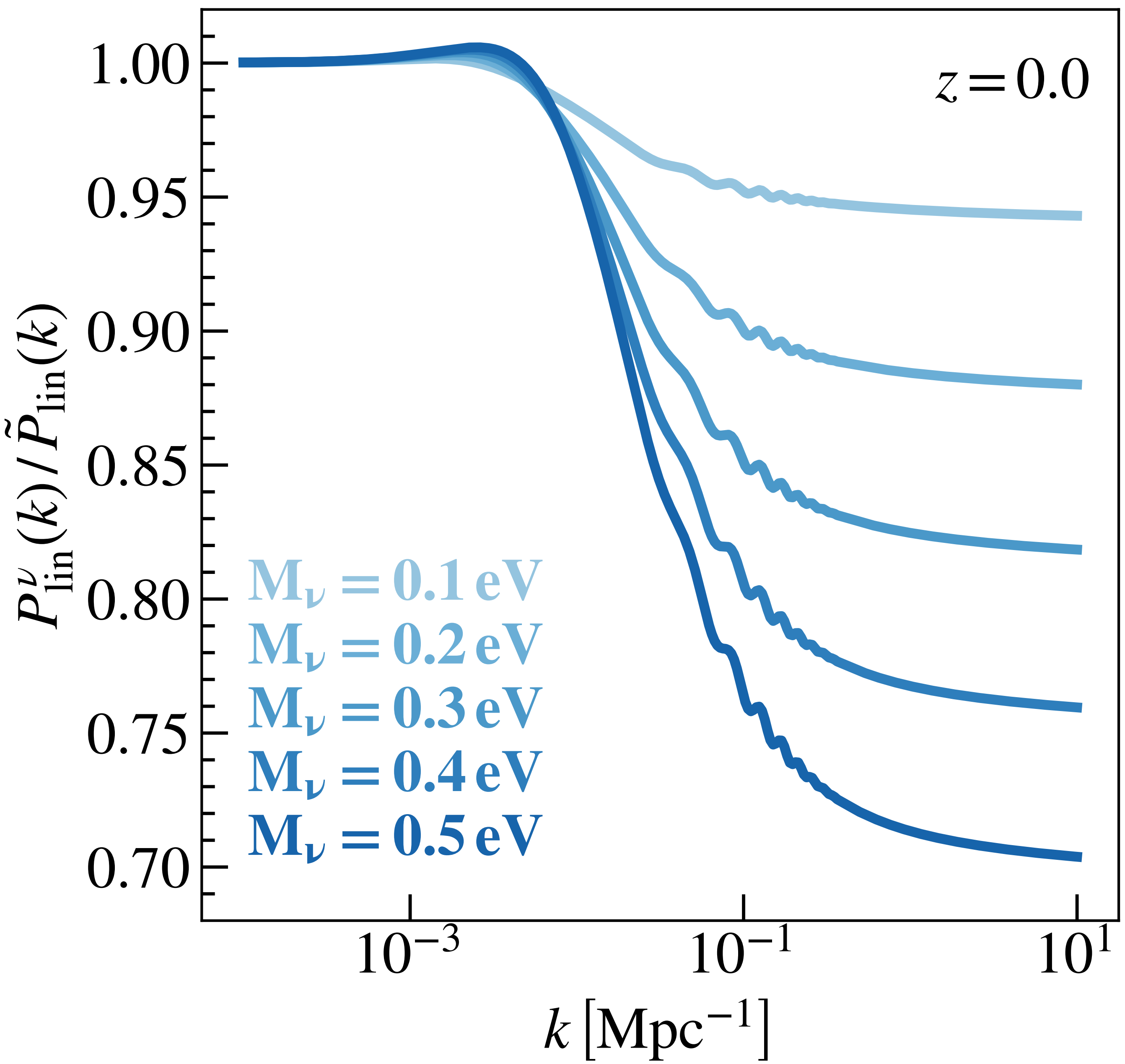
Redshift evolution induces a modification in the shape of the matter power spectrum

↓

Time variables (z, a) can no longer be described in terms of σ_{12}

Problems with massive neutrinos

Massive neutrinos break the original formulation of evolution mapping



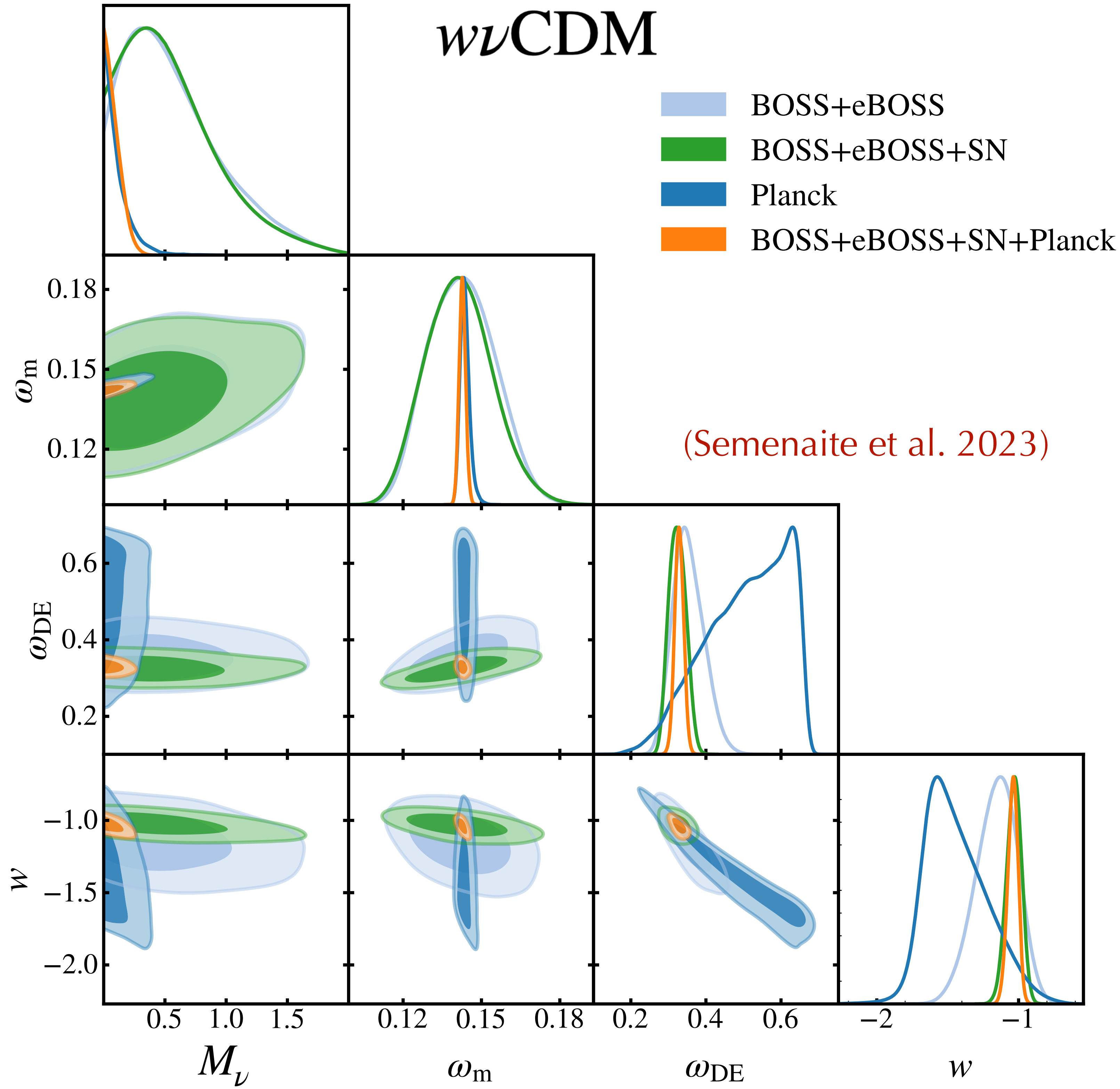
Redshift evolution induces a modification in the shape of the matter power spectrum

↓

Time variables (z, a) can no longer be described in terms of σ_{12}

Problems with massive neutrinos

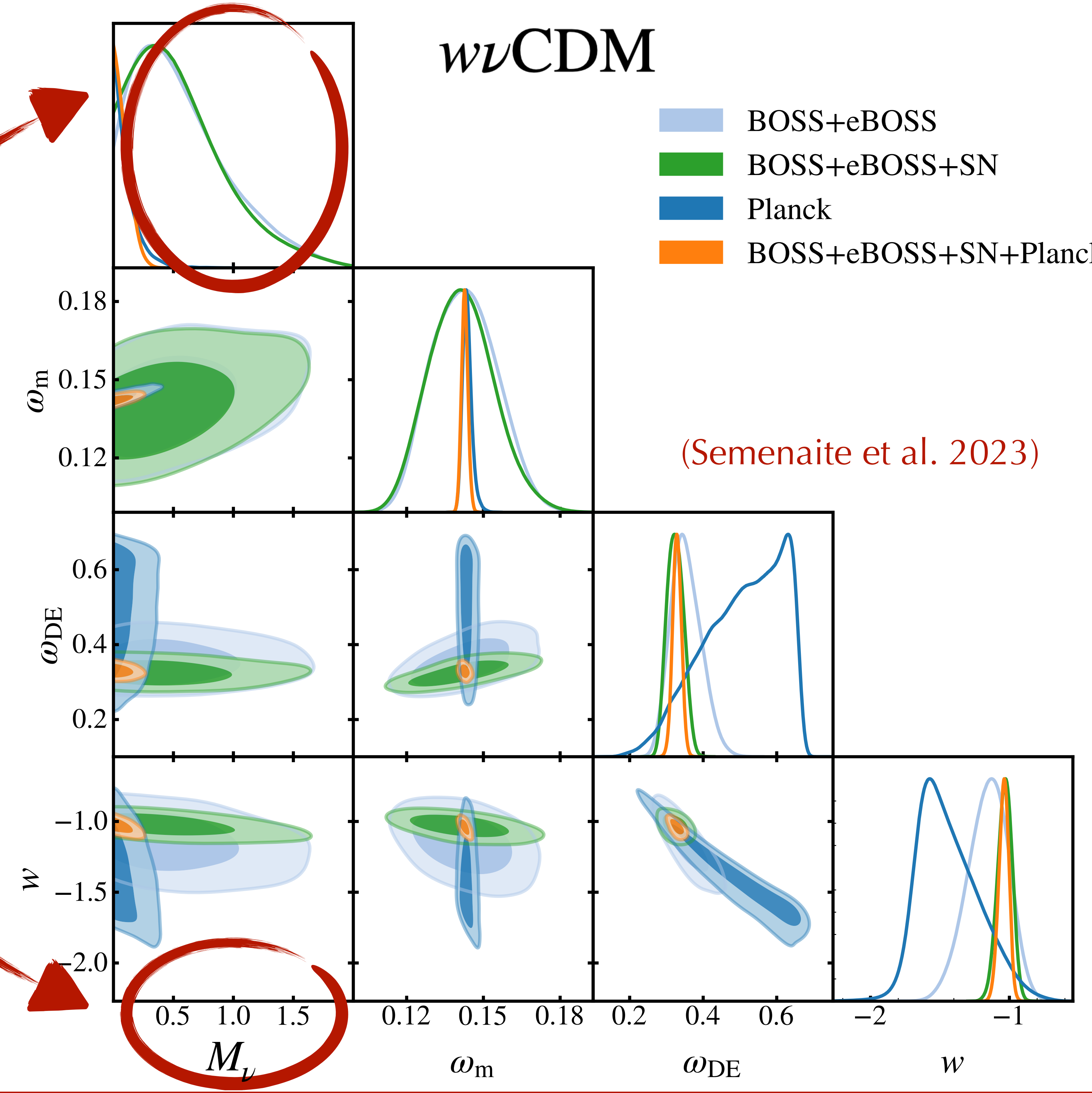
Relevant for Stage-III experiments, since constraining power on M_ν from GC-only data is loose



(Semenaite et al. 2023)

Problems with massive neutrinos

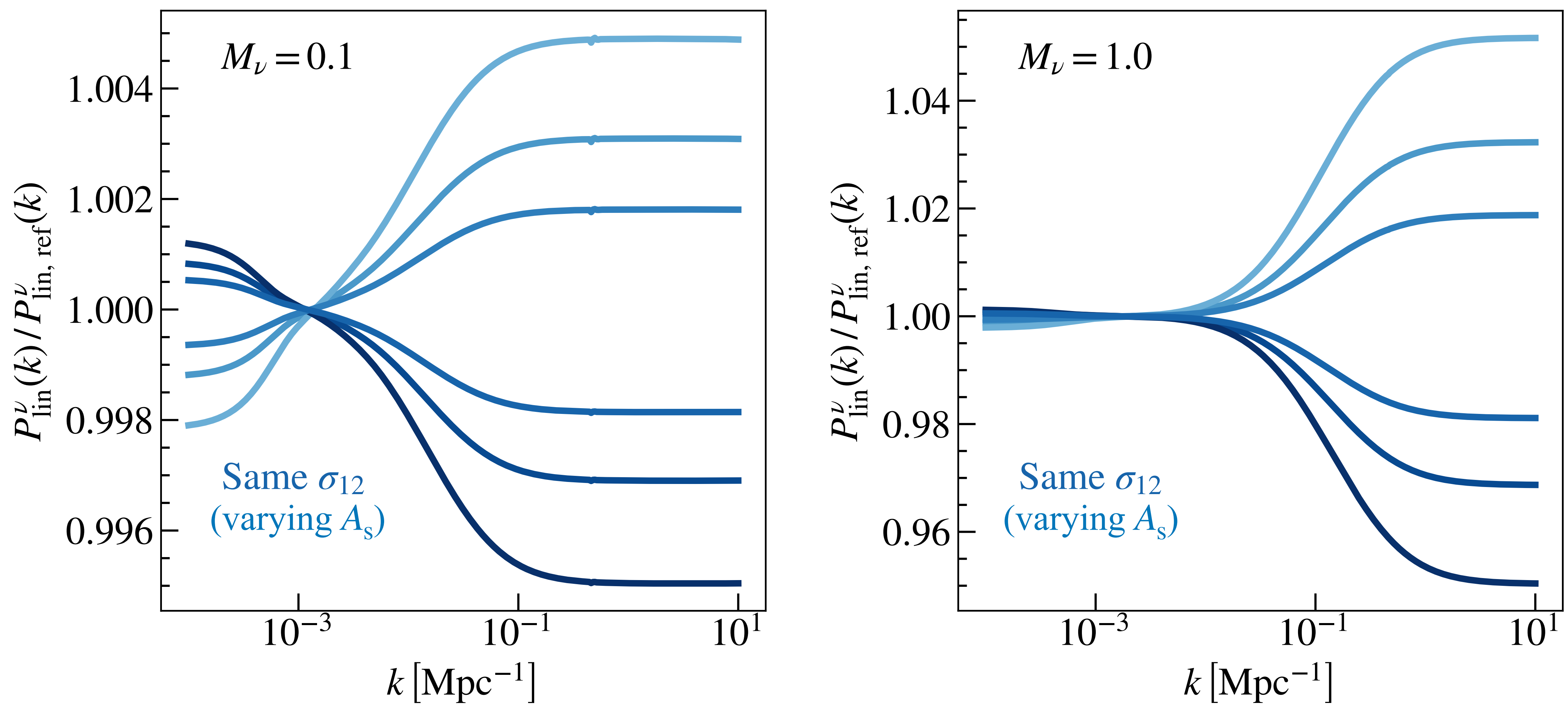
Relevant for Stage-III experiments, since constraining power on M_ν from GC-only data is loose



(Semenaite et al. 2023)

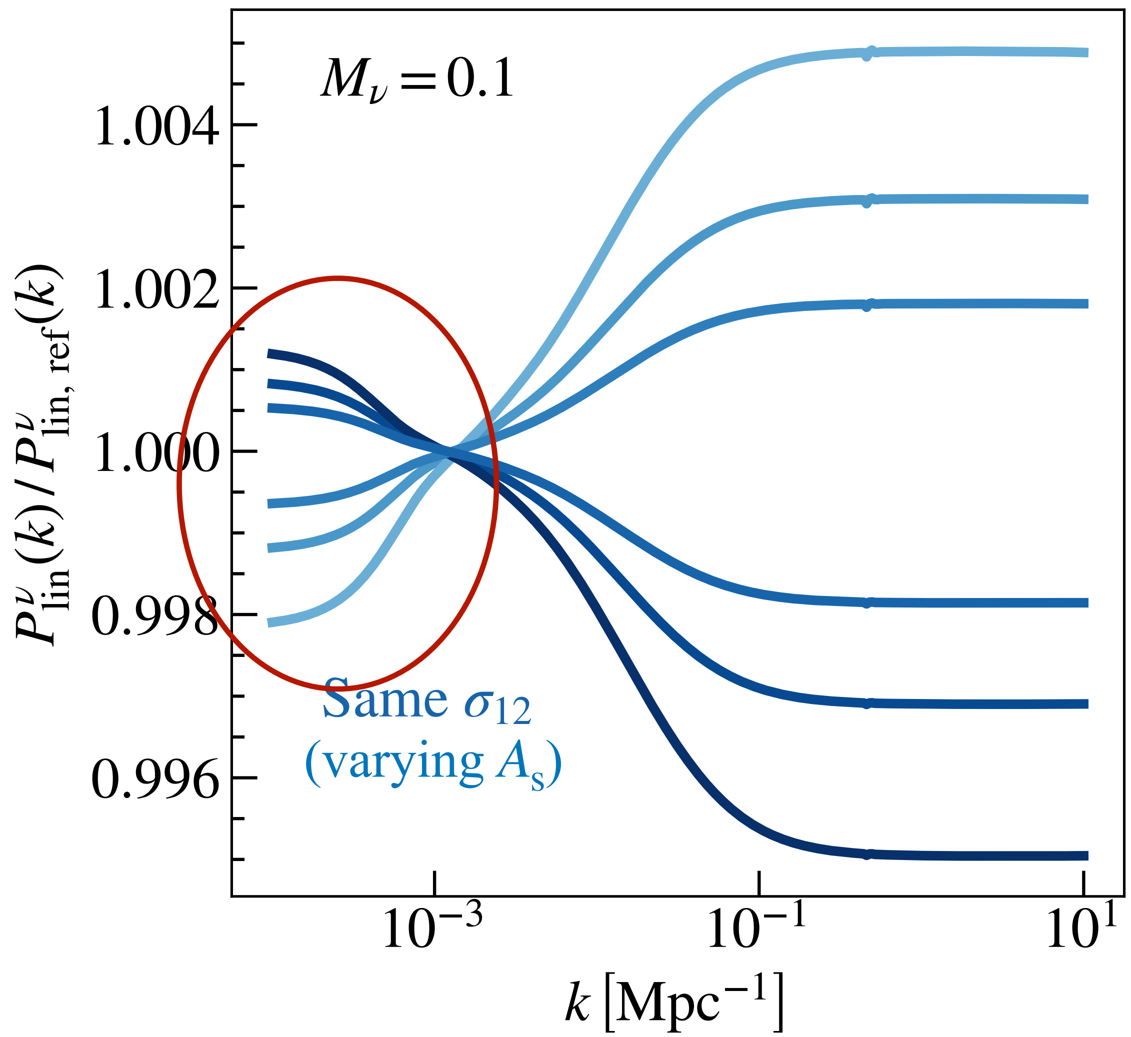
Problems with massive neutrinos

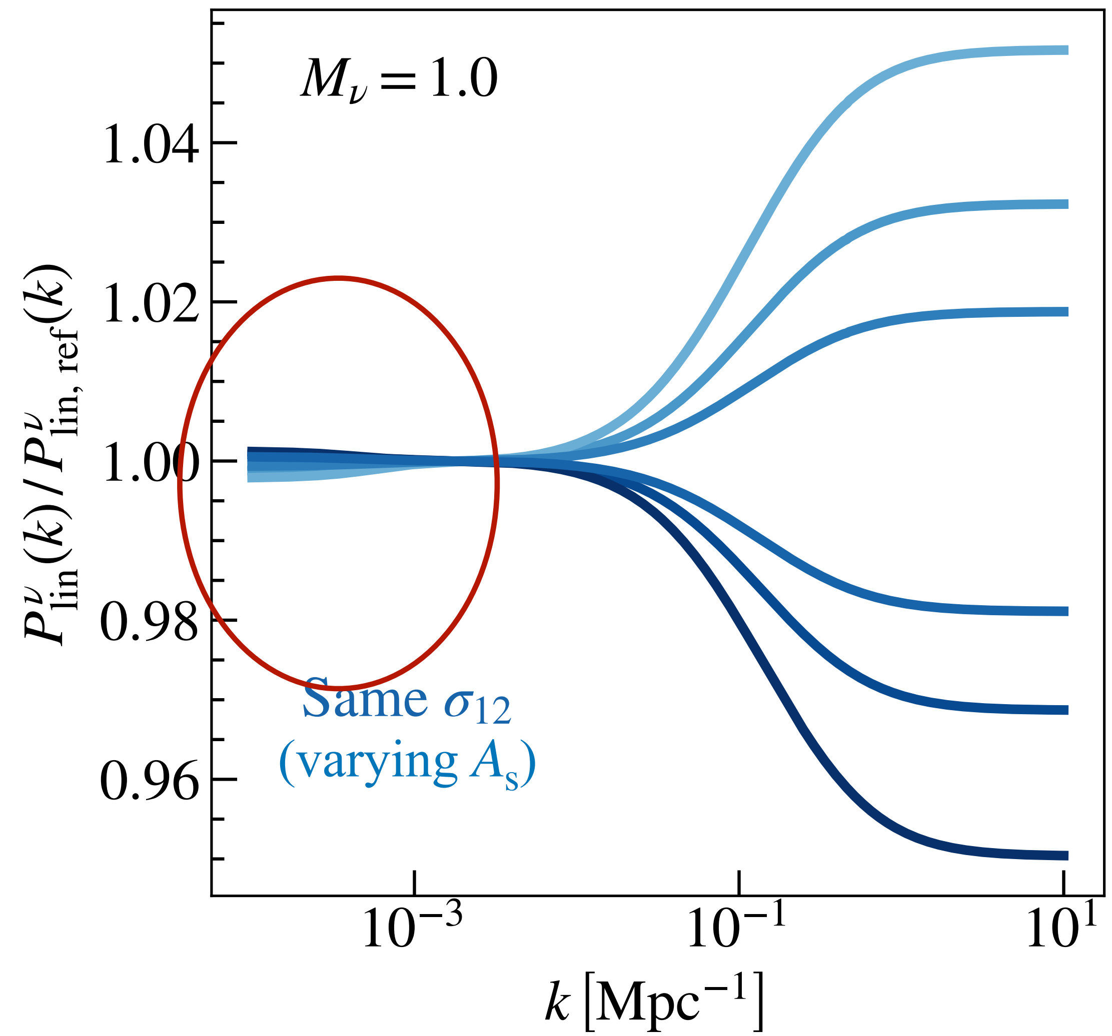
Large-scale amplitude is (partially) recovered only matching σ_{12} of the neutrino-free cosmology ($\omega_c \to \omega_c + \omega_\nu$, $\omega_\nu = 0$), but small-scale differences are important!



Problems with massive neutrinos

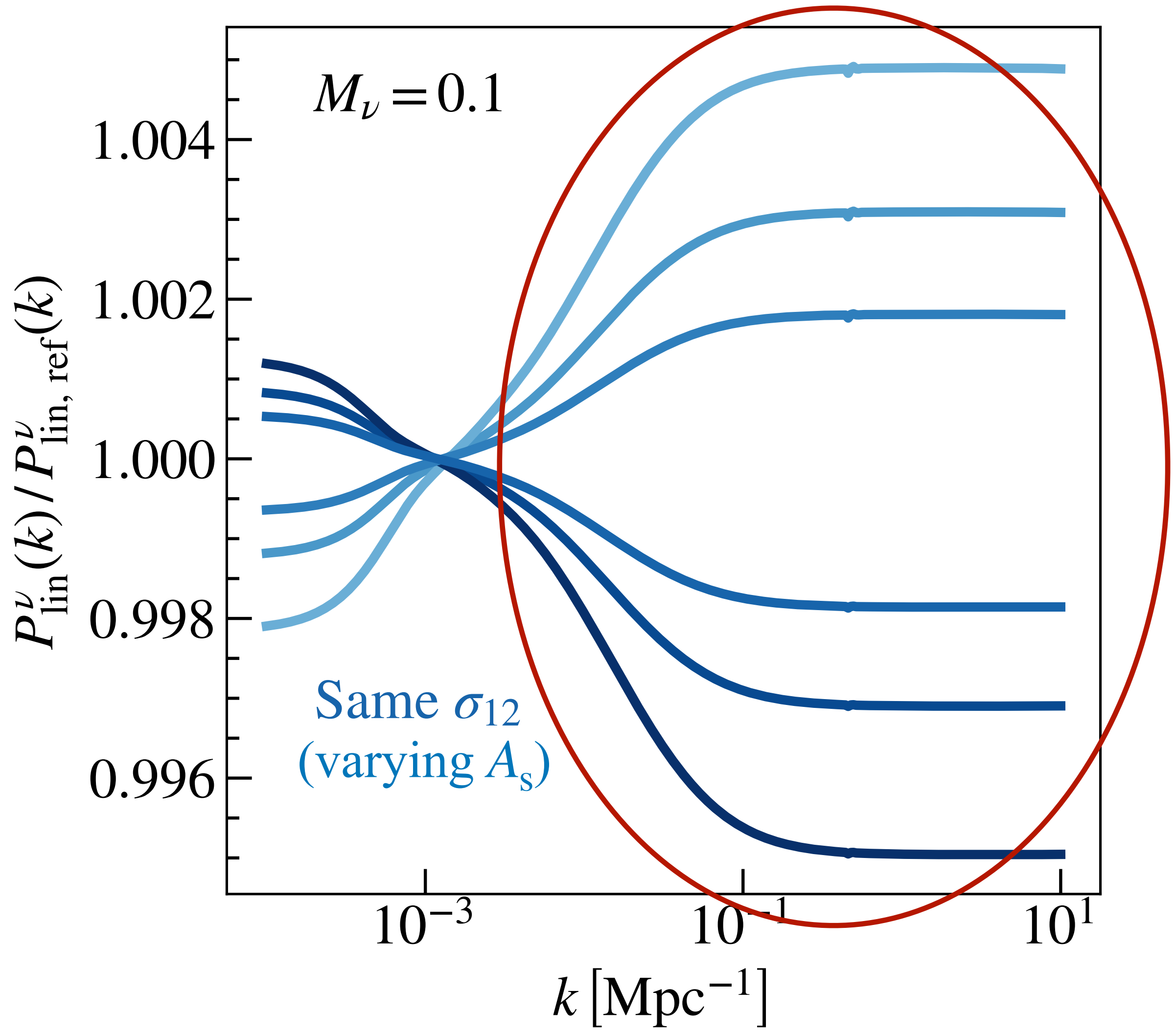
Large-scale amplitude is (partially) recovered only matching σ_{12} of the neutrino-free cosmology ($\omega_c \to \omega_c + \omega_\nu$, $\omega_\nu = 0$), but small-scale differences are important!

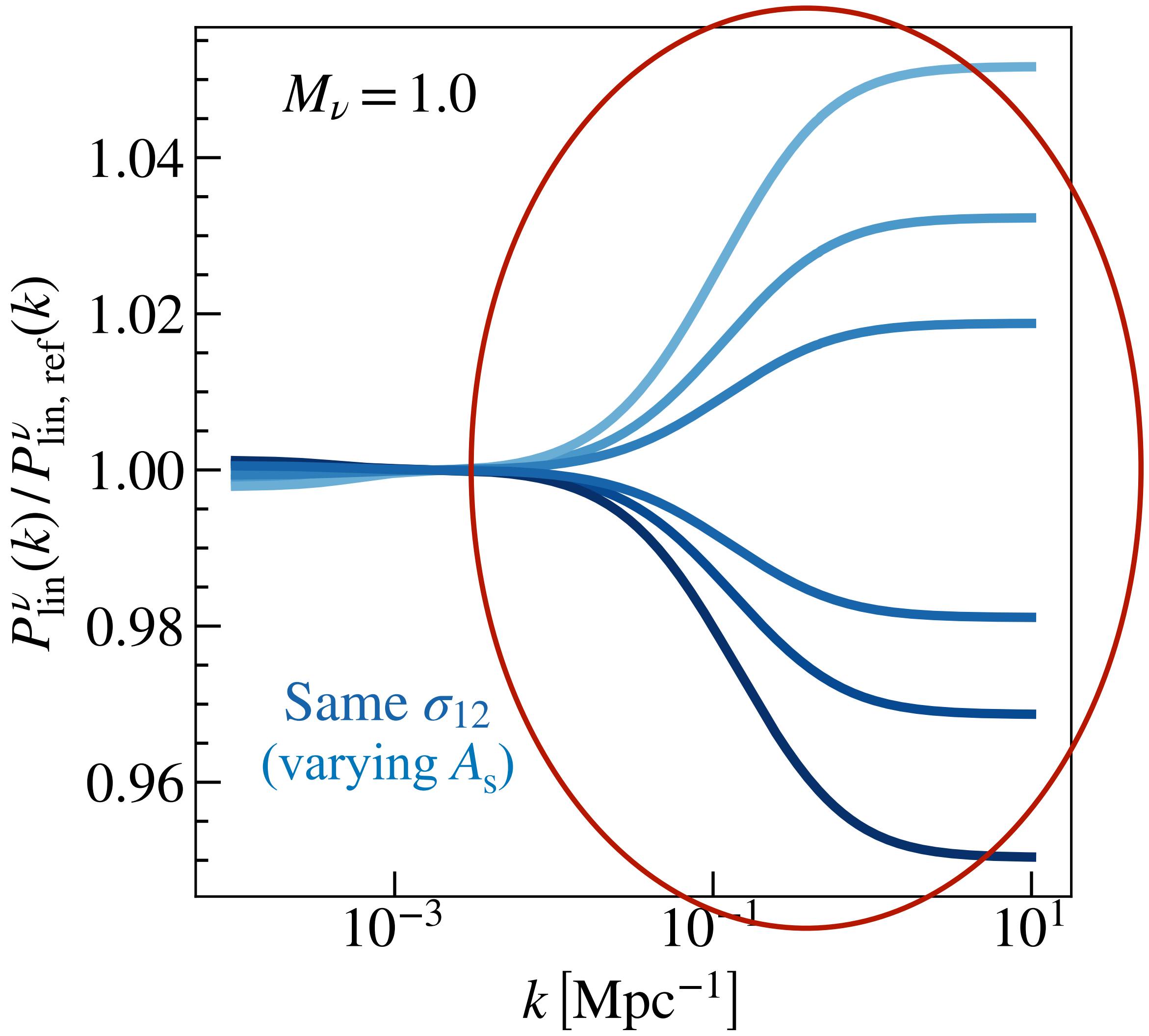




Problems with massive neutrinos

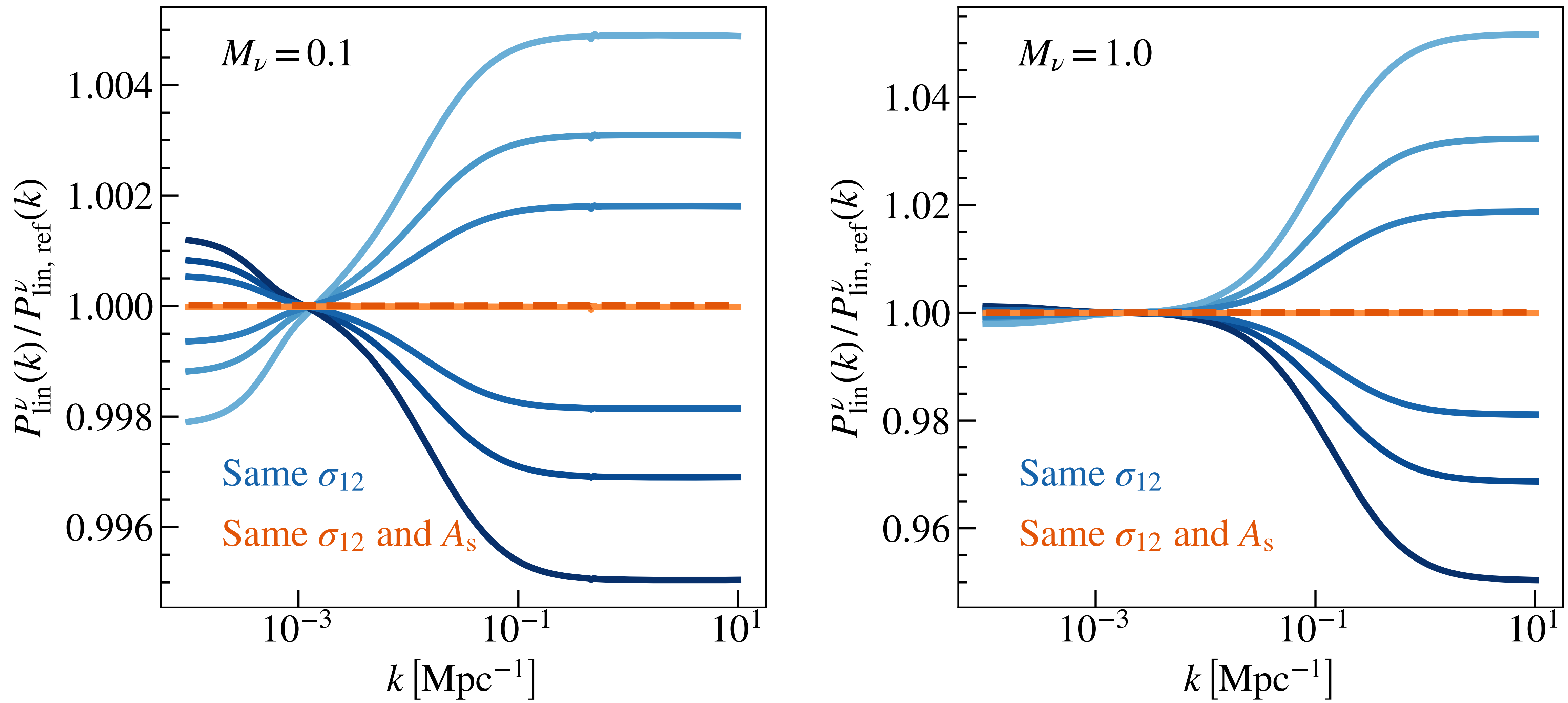
Large-scale amplitude is (partially) recovered only matching σ_{12} of the neutrino-free cosmology ($\omega_c \to \omega_c + \omega_\nu$, $\omega_\nu = 0$), but small-scale differences are important!





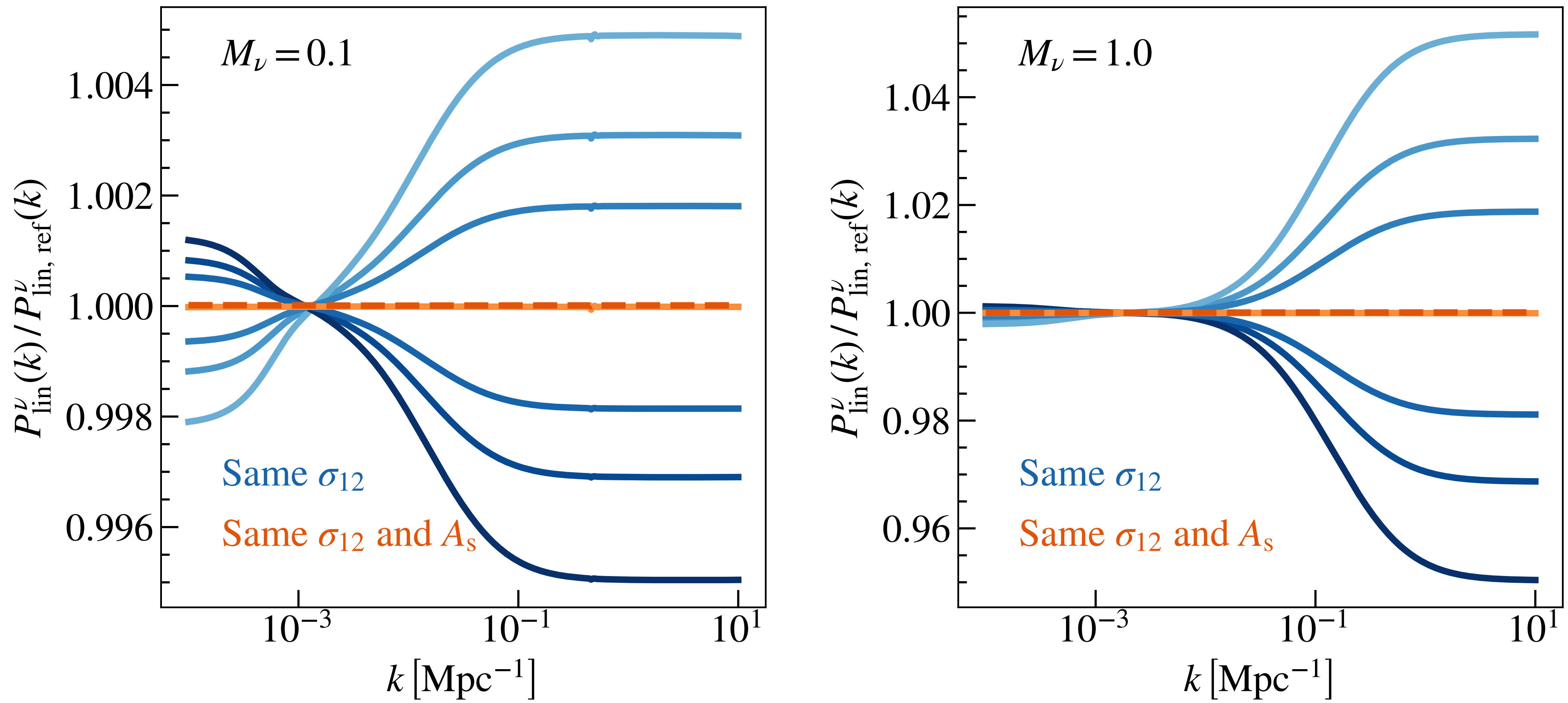
Problems with massive neutrinos

Large-scale amplitude is (partially) recovered only matching σ_{12} of the neutrino-free cosmology ($\omega_c \rightarrow \omega_c + \omega_\nu$, $\omega_\nu = 0$), but small-scale differences are important!



Problems with massive neutrinos

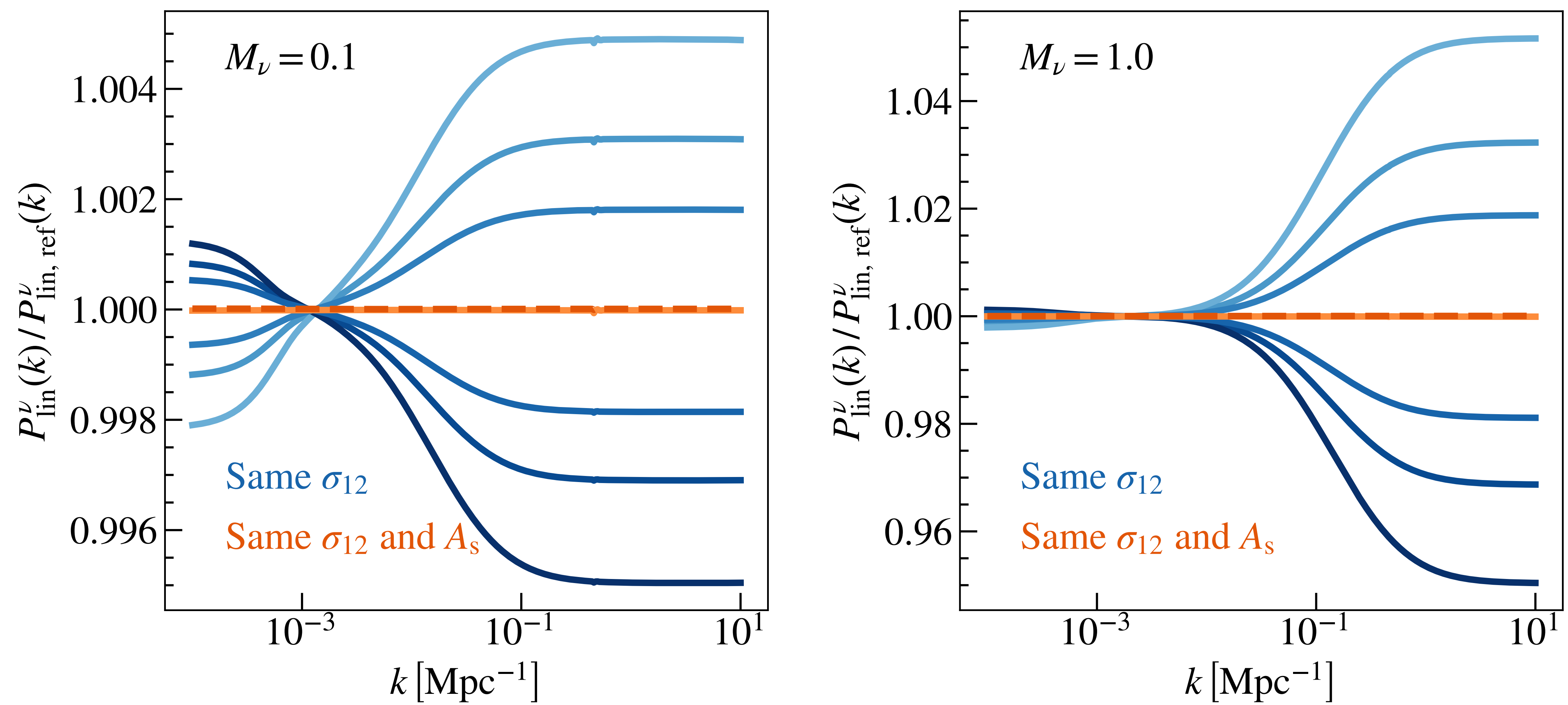
Large-scale amplitude is (partially) recovered only matching σ_{12} of the neutrino-free cosmology ($\omega_c \rightarrow \omega_c + \omega_\nu$, $\omega_\nu = 0$), but small-scale differences are important!



$$P_{\rm lin}(k\,|\,\mathbf{\Theta}_s\,,\mathbf{\Theta}_e\,,z) = P_{\rm lin}(k\,|\,\mathbf{\Theta}_s\,,\sigma_{12}(\mathbf{\Theta}_s\,,\mathbf{\Theta}_e\,,z))$$

Problems with massive neutrinos

Large-scale amplitude is (partially) recovered only matching σ_{12} of the neutrino-free cosmology ($\omega_c \rightarrow \omega_c + \omega_\nu$, $\omega_\nu = 0$), but small-scale differences are important!



$$P_{\rm lin}(k\,|\,\boldsymbol{\Theta}_s^\star, \boldsymbol{\Theta}_e^\star, z) = P_{\rm lin}(k\,|\,\boldsymbol{\Theta}_s^\star, \tilde{\sigma}_{12}(\boldsymbol{\Theta}_s^\star, \boldsymbol{\Theta}_e^\star, z))$$

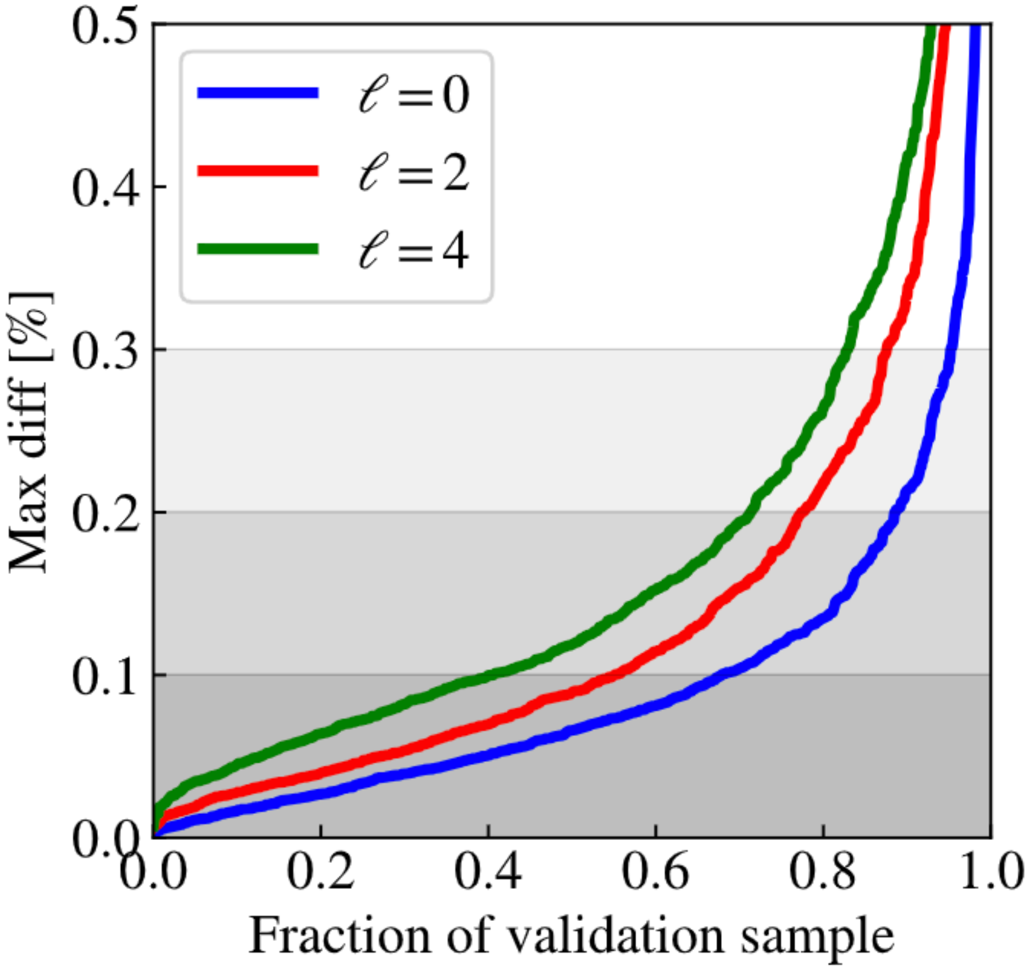
New implementation in COMET

- Emulation of perturbation theory to speed up model evaluation
- Based on evolution mapping to reduce dimensionality of parameter space
- Flexibility to incorporate different PT models (EFT, VDG, …)
- V1 (without massive neutrinos) public (Eggemeier et al. 2023)
  - `git clone git@gitlab.com:aegge/comet-emu.git`
  - `pip install comet-emu`
- Finalising V2 (massive neutrinos + extra features)

$$\{\omega_c, \omega_b, n_s, \sigma_{12}, f\} \rightarrow \{\omega_c, \omega_b, M_\nu, n_s, A_s, \sigma_{12}, f\}$$

Accuracy better than 0.2% for ~90% of the validation sample (**monopole**)

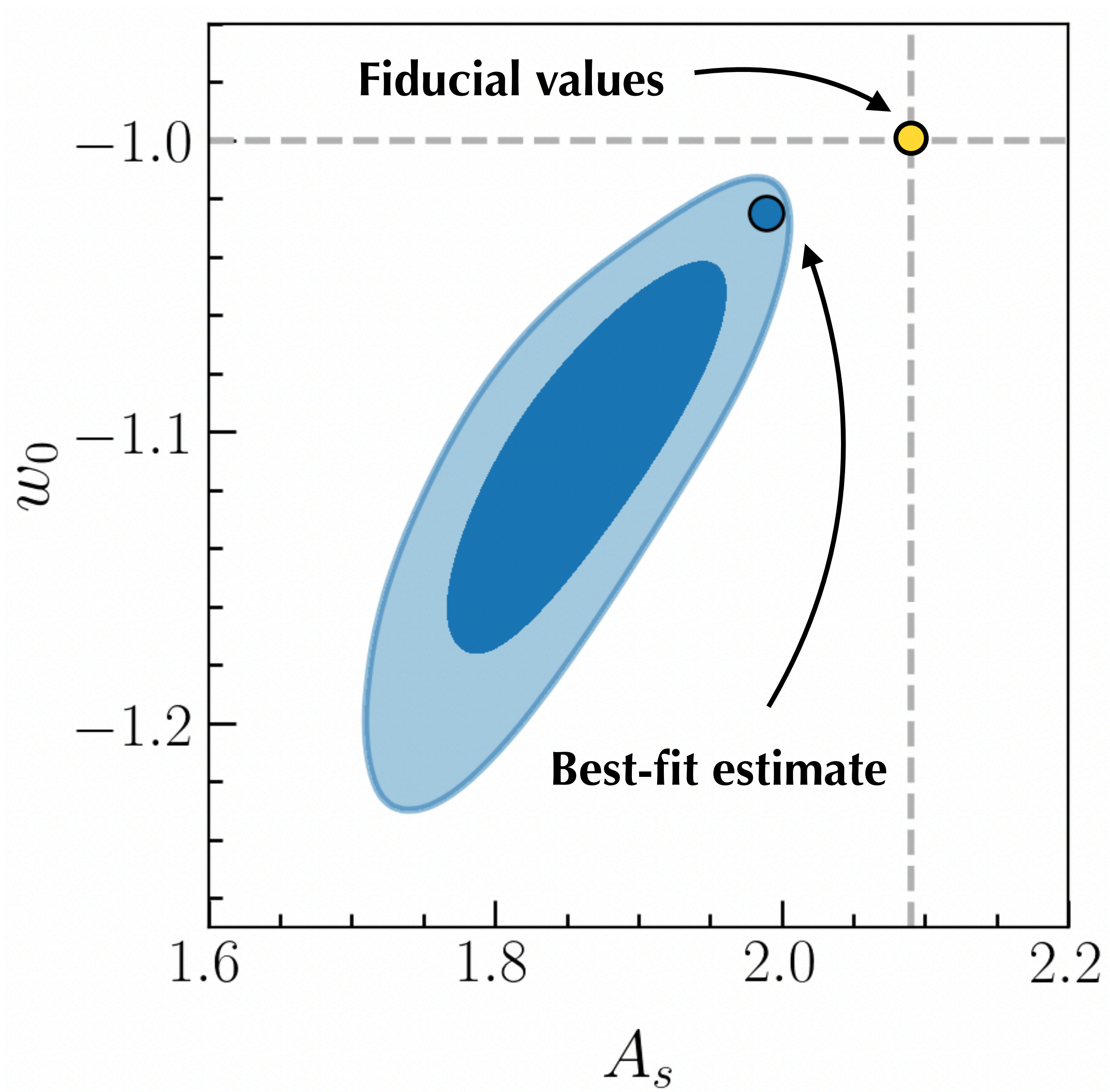
Evaluation time decreases to $\mathcal{O}(10\,\mathrm{ms})$



Prior volume effects

Prior volume effects

Bayesian analyses can be affected by **projection effects** during the marginalisation over nuisance parameters



Test with **synthetic data vectors**

- 4 redshift bins $\bar{z} = \{1, 1.2, 1.4, 1.65\}$
- Final mission sky fraction ($\sim 15000\,\mathrm{deg}^2$)
- wCDM cosmology
- Marginalisation over 6 nuisance parameters per bin $(b_1\,, b_2\,, c_0\,, c_2\,, c_4\,, \alpha_P)$

The importance of projection effects is more relevant when

- exploring high dimensional parameter space
- exploring non-Gaussian likelihoods (which is typically the case for spectroscopic galaxy clustering)

Prior volume effects

Let's assume to have a model (with parameters θ) that describes some data $\mathbf{d}$

$$\vec{\theta} = \{\vec{\Omega}, \vec{n}\}$$

Parameters of interest
(**cosmology**)

Parameters to marginalise over
(**nuisance, bias, RSD, systematics**)

The marginalised posterior distribution of $\vec{\Omega}$ is $\mathbf{P}(\vec{\Omega}\,|\,\mathbf{d}) = \int \mathrm{d}\vec{n}\;\mathbf{P}(\vec{\Omega}, \vec{n}\,|\,\mathbf{d})$

We can expand $\chi^2 \equiv -2\log\mathbf{P}$ around the nuisance parameters $\vec{n}_\star(\vec{\Omega})$ that maximise the posterior

$$\chi^2(\vec{\Omega}, \vec{n}) \simeq \chi^2_\star(\vec{\Omega}) + \Delta\vec{n}^T\mathscr{F}_\star\Delta\vec{n} + \ldots$$

After marginalisation of $\vec{n}$, this reads

$$\chi^2_{\mathrm{marg}}(\vec{\Omega}) \simeq \chi^2_\star(\vec{\Omega}) + \log\{\det[\mathscr{F}_\star(\vec{\Omega})]\}$$

Profile likelihood
(only depends on data)

Laplace term
(prior volume of nuisance parameters)

Prior volume effects

Let's assume to have a model (with parameters θ) that describes some data $\mathbf{d}$

$$\overrightarrow{\theta} = \{\overrightarrow{\Omega}, \overrightarrow{n}\}$$

Parameters of interest
(**cosmology**)

Parameters to marginalise over
(**nuisance, bias, RSD, systematics**)

The marginalised posterior distribution of $\overrightarrow{\Omega}$ is $\mathbf{P}(\overrightarrow{\Omega}\,|\,\mathbf{d}) = \int \mathrm{d}\overrightarrow{n}\, \mathbf{P}(\overrightarrow{\Omega}, \overrightarrow{n}\,|\,\mathbf{d})$

We can expand $\chi^2 \equiv -2\log \mathbf{P}$ around the nuisance parameters $\overrightarrow{n}_\star(\overrightarrow{\Omega})$ that maximise the posterior

$$\chi^2(\overrightarrow{\Omega}, \overrightarrow{n}) \simeq \chi^2_\star(\overrightarrow{\Omega}) + \Delta \overrightarrow{n}^T \mathscr{F}_\star \Delta \overrightarrow{n} + \ldots$$

After marginalisation of $\overrightarrow{n}$, this reads

$$\chi^2_{\mathrm{marg}}(\overrightarrow{\Omega}) \simeq \chi^2_\star(\overrightarrow{\Omega}) + \log\{\det[\mathscr{F}_\star(\overrightarrow{\Omega})]\}$$

Profile likelihood
(only depends on data)

Laplace term
(prior volume of nuisance parameters)

Prior volume effects

Let's assume to have a model (with parameters θ) that describes some data $\mathbf{d}$

$$\vec{\theta} = \{\vec{\Omega}, \vec{n}\}$$

Parameters of interest
(**cosmology**)

Parameters to marginalise over
(**nuisance, bias, RSD, systematics**)

The marginalised posterior distribution of $\vec{\Omega}$ is $\mathbf{P}(\vec{\Omega}\,|\,\mathbf{d}) = \int \mathrm{d}\vec{n}\; \mathbf{P}(\vec{\Omega}, \vec{n}\,|\,\mathbf{d})$

We can expand $\chi^2 \equiv -2\log \mathbf{P}$ around the nuisance parameters $\vec{n}_\star(\vec{\Omega})$ that maximise the posterior

$$\chi^2(\vec{\Omega}, \vec{n}) \simeq \chi^2_\star(\vec{\Omega}) + \Delta\vec{n}^T \mathscr{F}_\star \Delta\vec{n} + \ldots$$

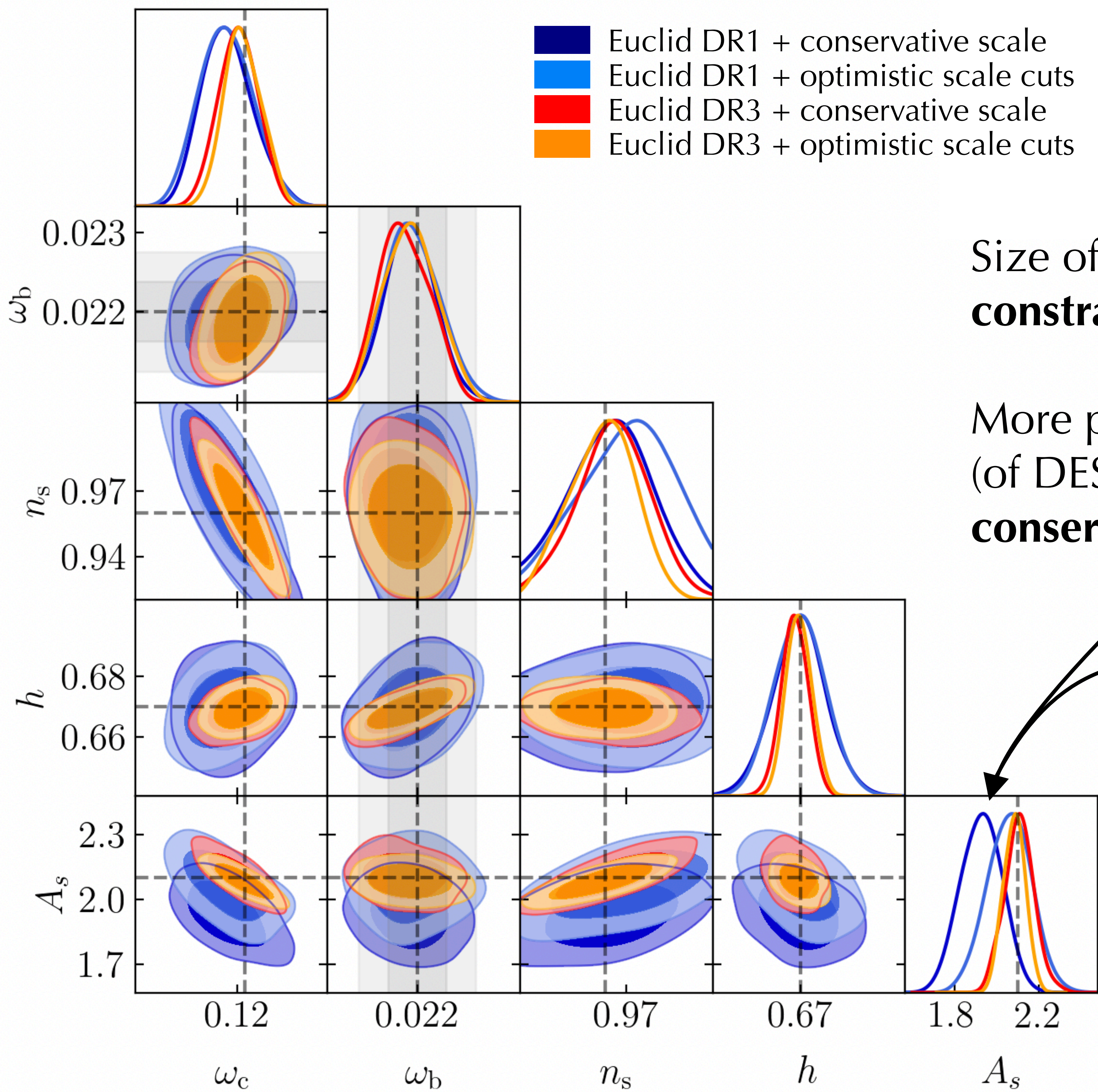
After marginalisation of $\vec{n}$, this reads

$$\chi^2_{\rm marg}(\vec{\Omega}) \simeq \chi^2_\star(\vec{\Omega}) + \log\{\det[\mathscr{F}_\star(\vec{\Omega})]\}$$

Profile likelihood
(only depends on data)

Laplace term
(prior volume of nuisance parameters)

Prior volume effects

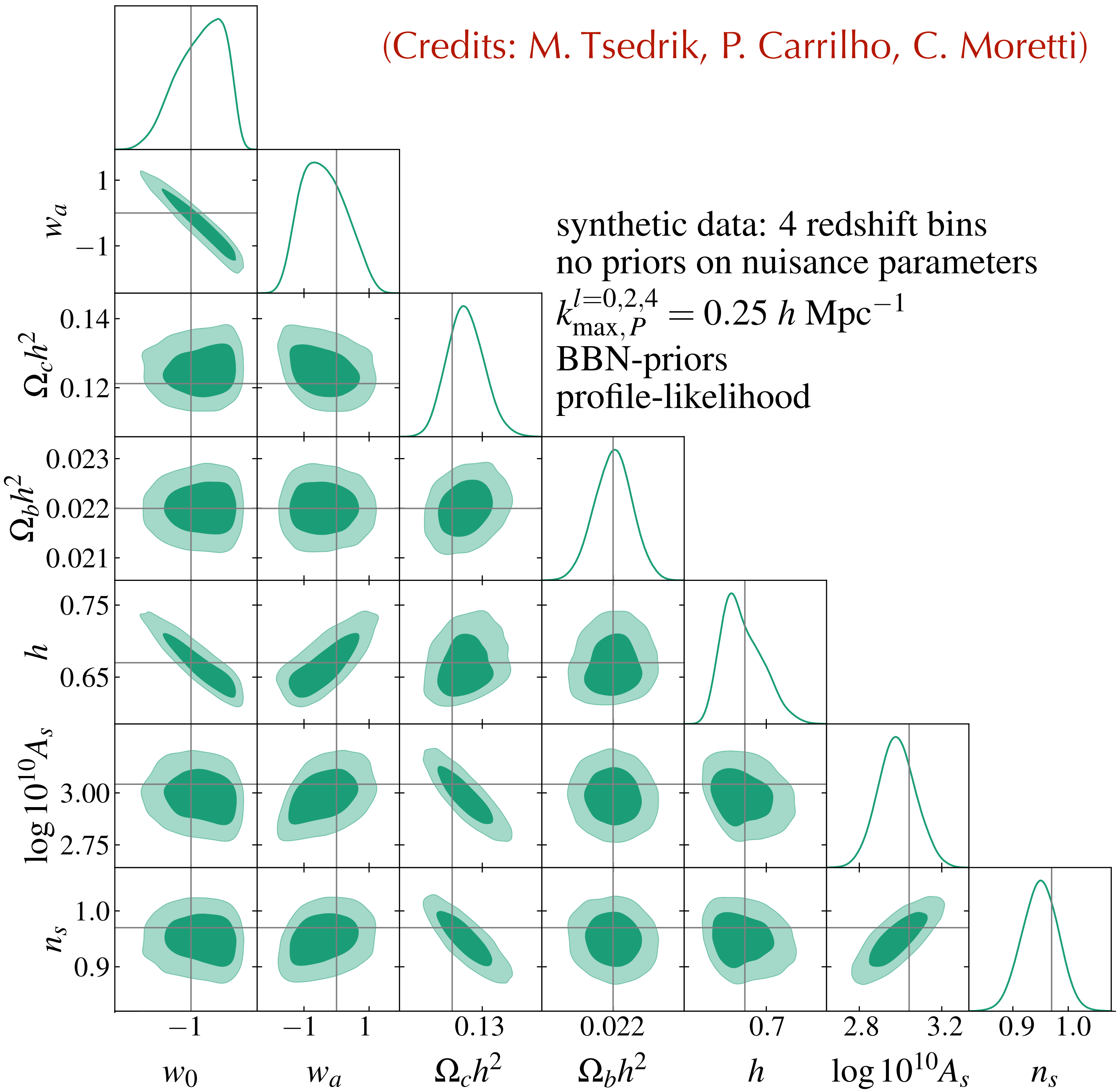


Size of projection effects related to **constraining power** of data vectors

More problematic for **first Data Releases** (of DESI, Euclid, …) and for more **conservative scale cuts**

Some parameters are more affected, i.e. A_s , due to stronger dependence of Laplace term on the nuisance parameters

Potential solutions



1) Using profile likelihood

Removal of Laplace term (the one sourcing the prior volume effects)

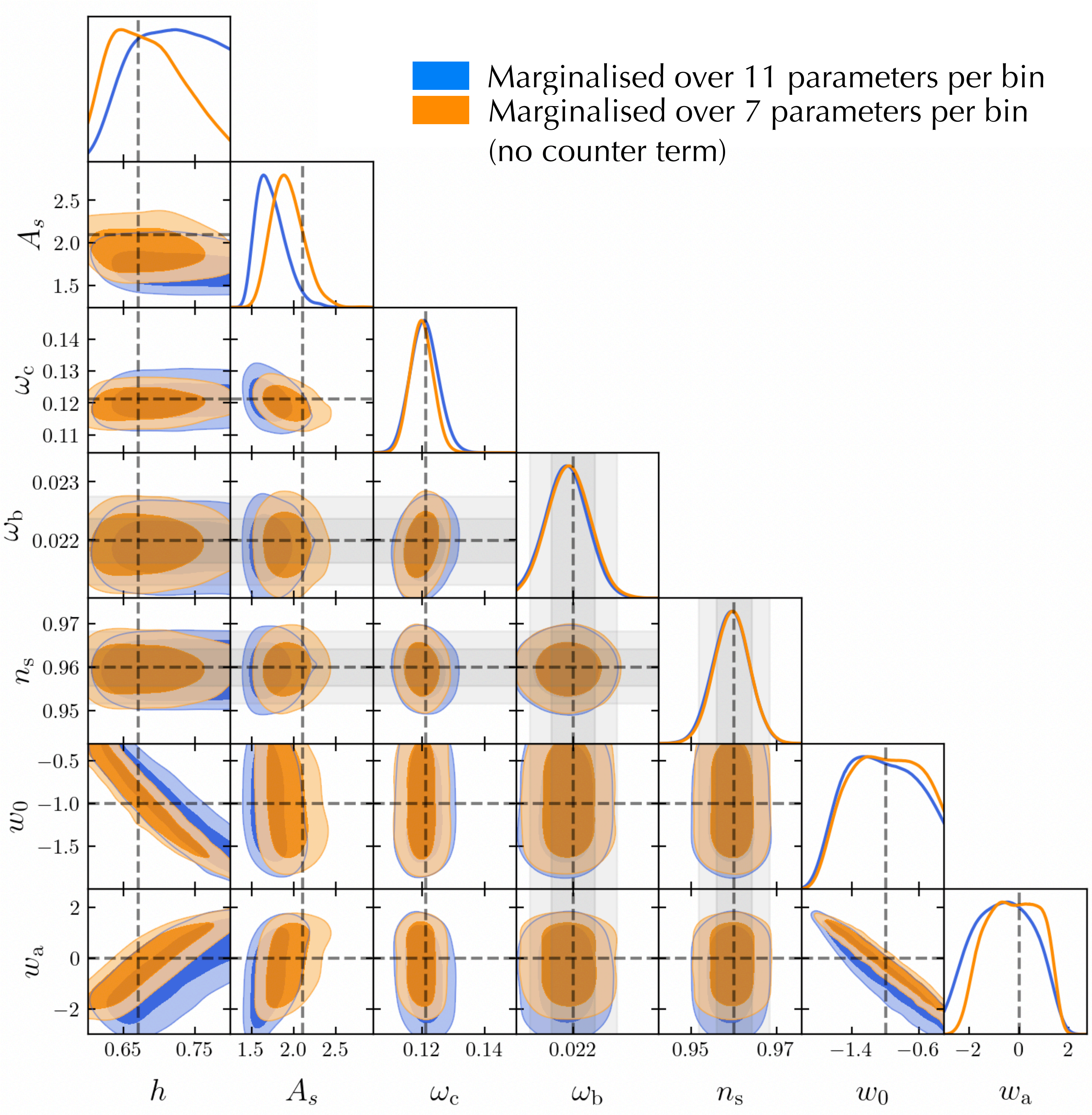
Pros

- Can be done efficiently with optimised tools (e.g. χ^2 minimisers)
- Removes the problem even for extended models (e.g. $w_0 w_a$CDM)

Cons

- Not easy to interpret error bars
- Frequentist approach instead of Bayesian

Potential solutions



2) Using priors on nuisance parameters

Tight priors on parameters that are mostly responsible for appearance of projection effects (e.g. counterterms)

Pros

- Optimal choices such as **Jeffreys priors**
- Encouraged when using RSD models where counterterms are not essential (e.g. VDG)

Cons

- Not in agreement with Bayesian formalism

Conclusions

- Perturbative models will still play a pivotal role in the analysis of Stage-IV data
  - Already reached a sufficient level of maturity
  - Different models to test scale cuts, theory systematics, constraining power, etc
  - Posterior validation thanks to many public codes
- Speeding up the model evaluation is essential
  - Use of theory emulators to maintain accuracy of PT-based models
  - Evolution mapping is an easy way to reduce the number of parameters of the emulators (works with massive neutrinos)
- Prior volume effects as one of the major limitations in the use of these models
  - Due to non-Gaussianity of likelihood
  - Potential ways to reduce their impact requires a questioning of standard Bayesian methods:

    - profile likelihood
    - Jeffrey's priors
    - tight priors on nuisances
    - reparametrization
    - use of alternative models