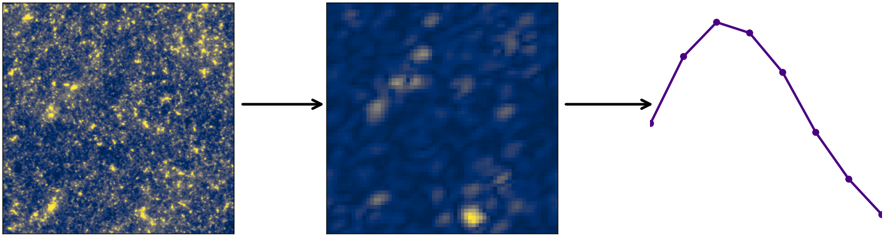
Scattering transforms for weak lensing cosmological inference



Mariia Marinichenko
Leiden University, the Netherlands

Non-Gaussian Universe ▪ June 17, 2026

Key questions

1. **Why ST?**
   - captures non-Gaussian information
   - easy to compute and interpretable

Key questions

1. **Why ST?**
   - captures non-Gaussian information
   - easy to compute and interpretable

2. **How to use it, and what can it give us?**
   - build an emulator for scattering coefficients
   - model baryonic feedback
   - forecast cosmological constraints

Key questions

1. **Why ST?**
   - captures non-Gaussian information
   - easy to compute and interpretable

2. **How to use it, and what can it give us?**
   - build an emulator for scattering coefficients
   - model baryonic feedback
   - forecast cosmological constraints

ST is a fast route from WL maps to comprehensible inference

Why the power spectrum is not enough

The power spectrum tells us **how much fluctuation power** we have at each angular scale

$$P_\kappa(\ell) = \langle |\kappa_\ell|^2 \rangle.$$

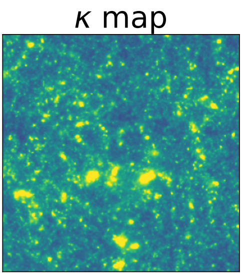
But it does not keep the Fourier phases, which carry information about **where structures are** and **how they are arranged**

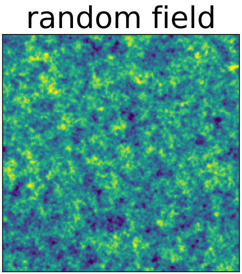
So, we lose part of the information about:

- halos, filaments and voids
- clustering of structures of different sizes
- nonlinear effects from baryons or dark sector physics

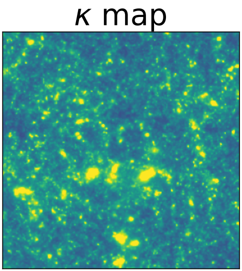
Part of this missing information can be recovered with **higher order statistics**

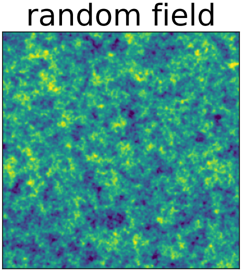
Why the scattering transform?

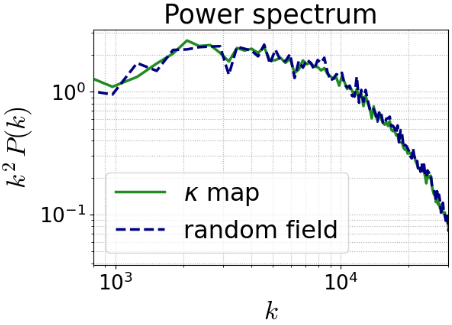




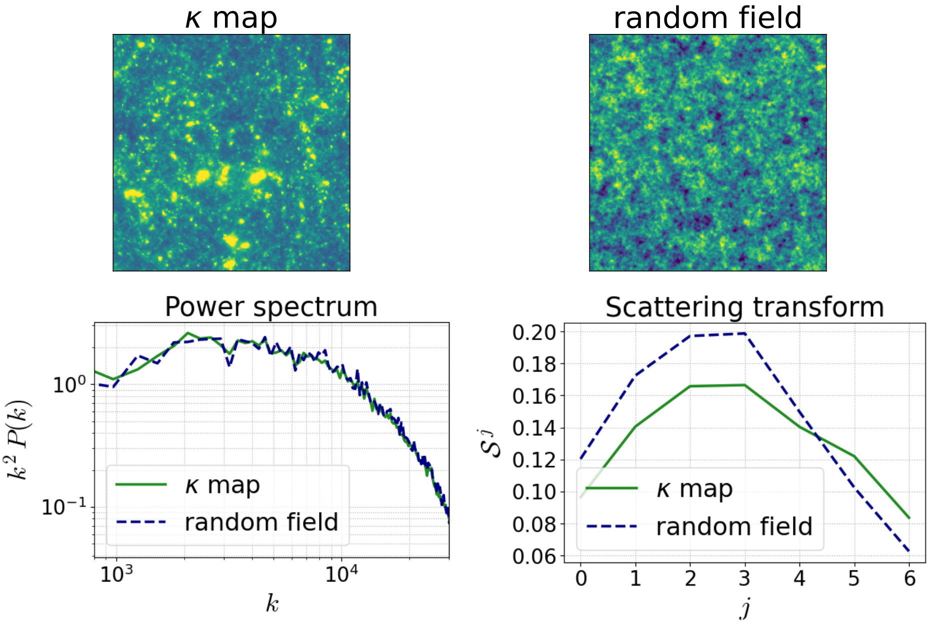
Why the scattering transform?



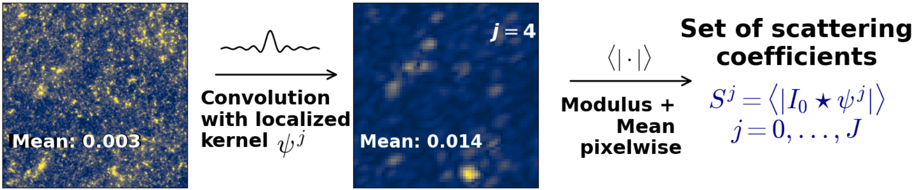




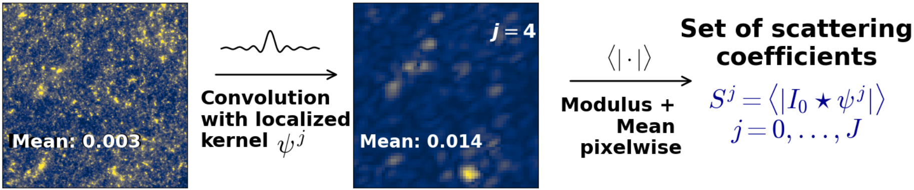
Why the scattering transform?



Scattering Transform



Scattering Transform



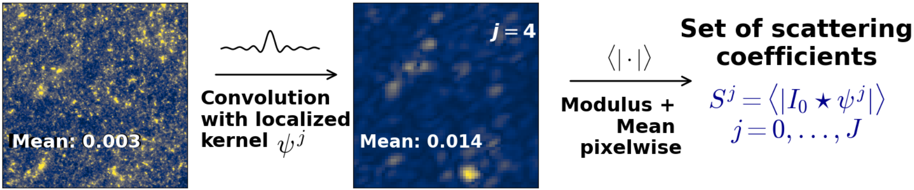
Scattering transform = wavelet convolution + modulus + mean

$$\mathcal{S}^j = \langle| I_0 \star \psi^j |\rangle$$

for a range of dyadic scales $j = 0, ..., J$

filter size = 2^j, $\quad j \sim \theta \sim \frac{1}{\ell} \sim \frac{1}{k}$

Scattering Transform



Scattering transform = wavelet convolution + modulus + mean

$$\mathcal{S}^j = \langle| \, I_0 \star \psi^j \, |\rangle$$

for a range of dyadic scales $j = 0, ..., J$

filter size $= 2^j$, $\quad j \sim \theta \sim \frac{1}{\ell} \sim \frac{1}{k}$

Advantages

- **Interpretable** (strength of fluctuations)
- **Efficient** (compact set of estimators)
- **Robust** (stable to noise)

What information does ST add?

First order: structures at each scale

$$S_1(j) = \langle |I_0 \star \psi_j| \rangle$$

- measures how much structure appears at some (angular) scale j
- similar idea to the power spectrum, but based on wavelets

What information does ST add?

First order: structures at each scale

$$S_1(j) = \langle |I_0 \star \psi_j| \rangle$$

- measures how much structure appears at some (angular) scale j
- similar idea to the power spectrum, but based on wavelets

Second order: clustering across scales

$$S_2(j_1, j_2) = \Big\langle \Big| |I_0 \star \psi_{j_1}| \star \psi_{j_2} \Big| \Big\rangle$$

- captures how small structures are grouped into larger ones
- describes the hierarchy of clustering patterns

What information does ST add?

First order: structures at each scale

$$S_1(j) = \langle |I_0 \star \psi_j| \rangle$$

- measures how much structure appears at some (angular) scale j
- similar idea to the power spectrum, but based on wavelets

Second order: clustering across scales

$$S_2(j_1, j_2) = \Big\langle \big| |I_0 \star \psi_{j_1}| \star \psi_{j_2} \big| \Big\rangle$$

- captures how small structures are grouped into larger ones
- describes the hierarchy of clustering patterns

ST describes map morphology ordered by scale

Why ST can improve cosmological constraints

Fisher forecast

$$F_{\alpha\beta} = \frac{\partial \boldsymbol{\mu}^T}{\partial \theta_\alpha} C^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \theta_\beta}$$

Why ST can improve cosmological constraints

Fisher forecast

$$F_{\alpha\beta} = \frac{\partial \boldsymbol{\mu}^T}{\partial \theta_\alpha} C^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \theta_\beta}$$

- observables that **react strongly** to cosmological parameter produce **tighter constraints**

Why ST can improve cosmological constraints

Fisher forecast

$$F_{\alpha\beta} = \frac{\partial \boldsymbol{\mu}^T}{\partial \theta_\alpha} C^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \theta_\beta}$$

- observables that **react strongly** to cosmological parameter produce **tighter constraints**
- ST is sensitive to nonlinear structure and map morphology $\Rightarrow$ additional parameter responses

Why ST can improve cosmological constraints

Fisher forecast

$$F_{\alpha\beta} = \frac{\partial \boldsymbol{\mu}^T}{\partial \theta_\alpha} C^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \theta_\beta}$$

- observables that **react strongly** to cosmological parameter produce **tighter constraints**
- ST is sensitive to nonlinear structure and map morphology $\Rightarrow$ additional parameter responses
- tomography adds redshift information, separating growth and geometry effects

Why ST can improve cosmological constraints

Fisher forecast

$$F_{\alpha\beta} = \frac{\partial \boldsymbol{\mu}^T}{\partial \theta_\alpha} C^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \theta_\beta}$$

- observables that **react strongly** to cosmological parameter produce **tighter constraints**
- ST is sensitive to nonlinear structure and map morphology $\Rightarrow$ additional parameter responses
- tomography adds redshift information, separating growth and geometry effects

ST can break power spectrum degeneracies using clustering information and redshift evolution

Why add ST to the Euclid inference pipeline?

What ST adds

- captures spatial information which κ PDF does not
- but is more compact than a densely binned bispectrum
- in addition, fewer binning choices or thresholds than with peaks or Minkowski functionals

Why add ST to the Euclid inference pipeline?

What ST adds

- captures spatial information which κ PDF does not
- but is more compact than a densely binned bispectrum
- in addition, fewer binning choices or thresholds than with peaks or Minkowski functionals

Practical for inference

- compact data vector
- natural ordering by angular scale
- easy to emulate

Why add ST to the Euclid inference pipeline?

What ST adds

- captures spatial information which κ PDF does not
- but is more compact than a densely binned bispectrum
- in addition, fewer binning choices or thresholds than with peaks or Minkowski functionals

Practical for inference

- compact data vector
- natural ordering by angular scale
- easy to emulate

ST adds map morphology in a handy way

What we want to measure

Our **goal** is to **estimate the sensitivity of the ST to cosmological parameters in a Euclid-like setup**, including evolving dark energy and decaying dark matter

What we want to measure

Our **goal** is to **estimate the sensitivity of the ST to cosmological parameters in a Euclid-like setup**, including evolving dark energy and decaying dark matter

Sensitivity

- build ST emulators from FLAMINGO
- compute Fisher forecasts
- improve with DALI

What we want to measure

Our **goal** is to **estimate the sensitivity of the ST to cosmological parameters in a Euclid-like setup**, including evolving dark energy and decaying dark matter

Sensitivity

- build ST emulators from FLAMINGO
- compute Fisher forecasts
- improve with DALI

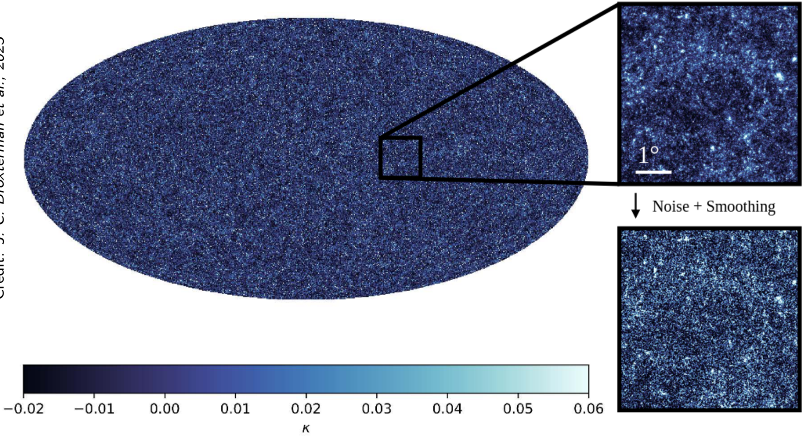
Debiasing

- model baryonic feedback bias with a transfer function
- add Euclid-like galaxy shape noise
- apply filters (Gaussian and Wiener)

FLAMINGO convergence maps

The transfer functions are derived from **Euclid-like WL mocks** from the **FLAMINGO** simulations

Credit: *J. C. Broxterman et al., 2023*



Full-sky weak lensing convergence map

Emulating scattering coefficients

Simulation set

- 100 FLAMINGO hypercube simulations
- 13 varied cosmological parameters (10 independent), including

$$\Omega_m,\ \sigma_8,\ w_0,\ w_a,\ M_\nu,\ \Gamma_{\rm dcdm},\ \dots$$

- matched initial phases for paired runs (with and without decaying DM)

Data vector

- κ maps in 5 tomographic bins
- HEALPix $N_{\rm side} = 8192$
- $3.6 \times 3.6\ {\rm deg}^2$ non-overlapping patches
- 1st ST coefficients for each patch
- Euclid-like Gaussian galaxy shape noise:
- $n_{\rm gal} = 30\,{\rm arcmin}^{-2}$, $\sigma_e = 0.26$

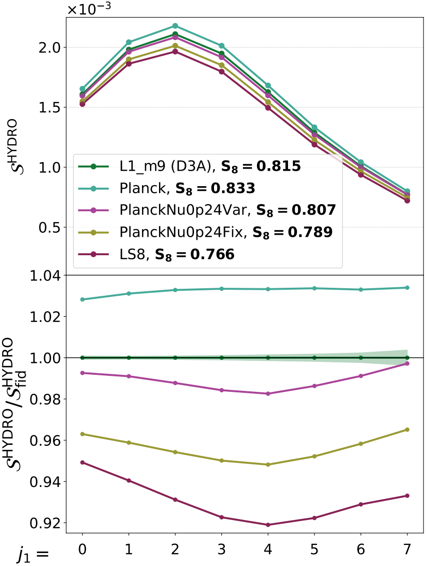
ST for cosmology variations

5 cosmology variations
in **1 cGpc** FLAMINGO runs

ST for cosmology variations

5 cosmology variations
in **1 cGpc** FLAMINGO runs

Same feedback, different cosmologies



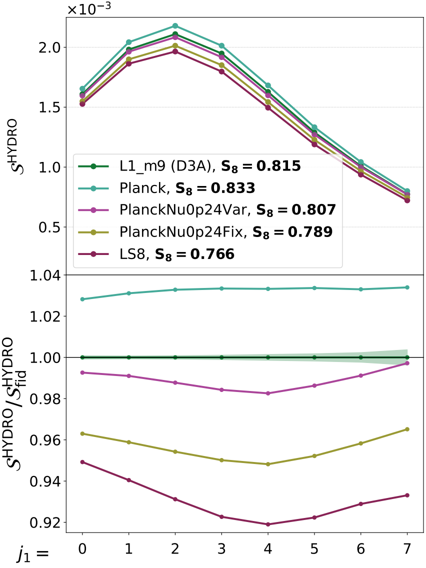
ST for cosmology variations

5 cosmology variations
in **1 cGpc** FLAMINGO runs

Result

Scattering coefficients are highly sensitive to $\mathbf{S_8}$

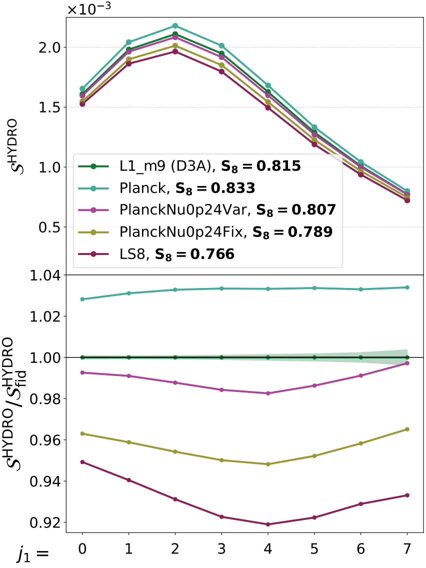
Same feedback, different cosmologies



ST for cosmology variations

5 cosmology variations
in **1 cGpc** FLAMINGO runs

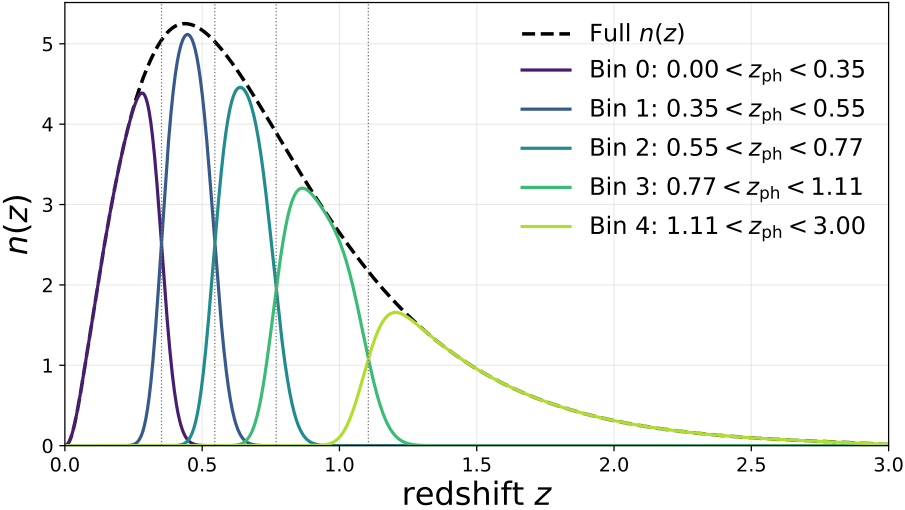
Same feedback, different cosmologies



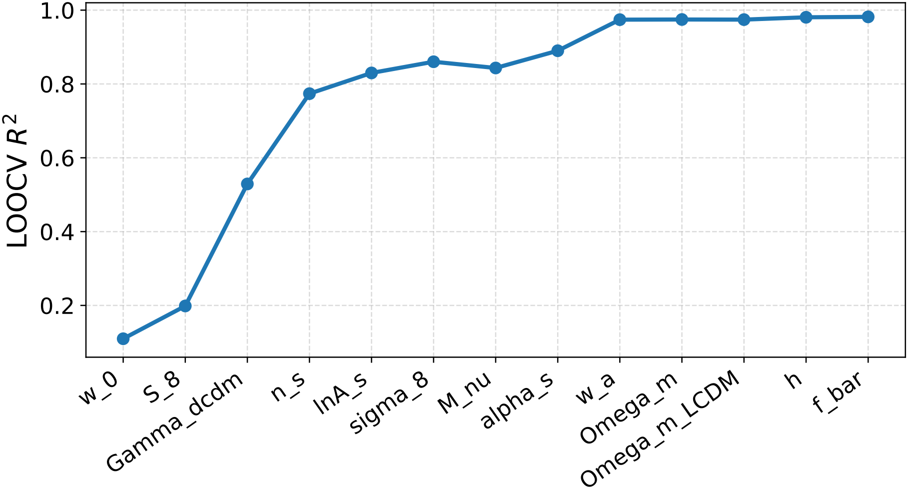
Result

Scattering coefficients are highly sensitive to $\mathbf{S_8}$
* *for the same baryonic feedback*

Tomographic bins

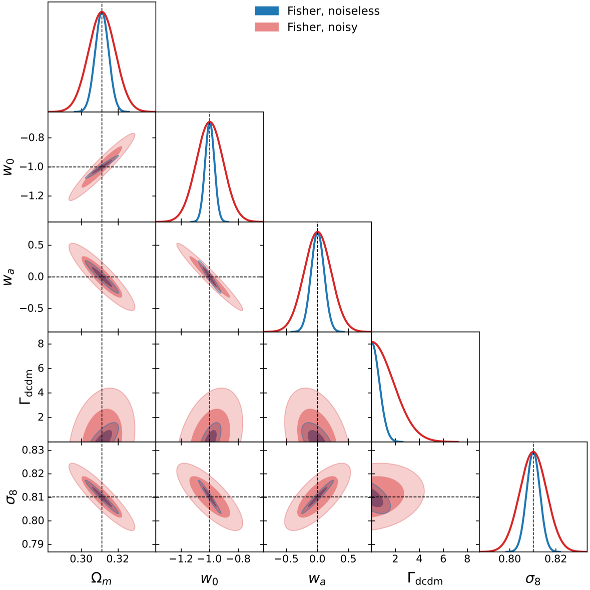


Which parameters matter for the emulator?

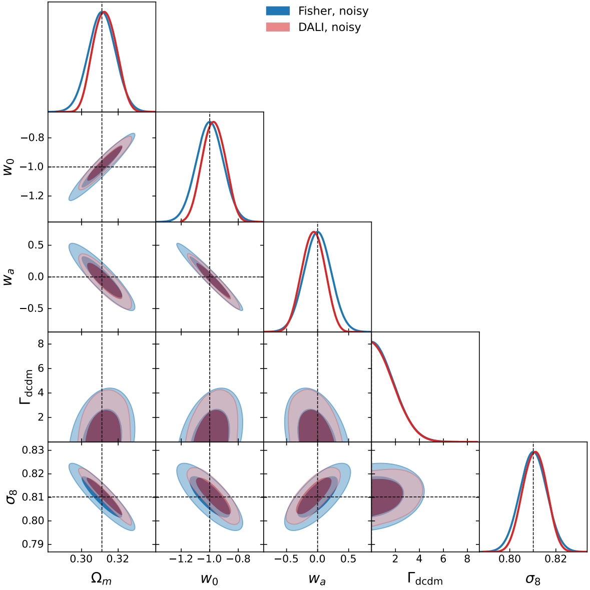


The greedy search selects the physical parameters that ST is sensitive to

ST forecast: noise



DALI improvement of the forecast



Baryonic effects on the weak lensing scattering transform

The effect of baryonic feedback on WL convergence can be quantified through a transfer function, which we extract from **FLAMINGO** simulations.

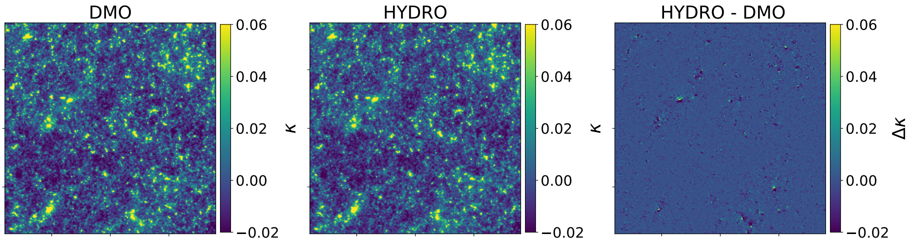


Figure: Convergence, i.e. κ, patches from FLAMINGO

Why weak lensing cares about baryons?

- **Baryonic feedback** blows gas out of halos, reshaping the total matter distribution

Why weak lensing cares about baryons?

- **Baryonic feedback** blows gas out of halos, reshaping the total matter distribution
- Harmless on large scales, but devastating at $k \gtrsim 0.1\, h$ Mpc^{-1}

Why weak lensing cares about baryons?

- **Baryonic feedback** blows gas out of halos, reshaping the total matter distribution
- Harmless on large scales, but devastating at $k \gtrsim 0.1\,h$ Mpc^{-1}
- **Stage IV surveys** (*Euclid*, *LSST*, *Roman*) will probe exactly that scales (with high S/N)

Why weak lensing cares about baryons?

- **Baryonic feedback** blows gas out of halos, reshaping the total matter distribution
- Harmless on large scales, but devastating at $k \gtrsim 0.1\,h$ Mpc^{-1}
- **Stage IV surveys** (*Euclid, LSST, Roman*) will probe exactly that scales (with high S/N)
- Ignore baryons ⇒ biased cosmology: S_8, Ω_m, w_0 off by $\gtrsim 1\sigma$

Why weak lensing cares about baryons?

- **Baryonic feedback** blows gas out of halos, reshaping the total matter distribution
- Harmless on large scales, but devastating at $k \gtrsim 0.1\, h$ Mpc^{-1}
- **Stage IV surveys** (*Euclid, LSST, Roman*) will probe exactly that scales (with high S/N)
- Ignore baryons $\Rightarrow$ biased cosmology: S_8, Ω_m, w_0 off by $\gtrsim 1\sigma$

Baryonic effects in WL **scattering transform** can be separated from cosmology through a **transfer function**

How the baryon model enters inference

Cosmological parameter inference

Observed κ maps $\Rightarrow$ ST data vector $\Rightarrow$ likelihood $\Rightarrow$ $\Omega_{\rm m}$, σ_8, w_0, $\Gamma_{\rm dcdm}$, ...

How the baryon model enters inference

Cosmological parameter inference

Observed κ maps $\Rightarrow$ ST data vector $\Rightarrow$ likelihood $\Rightarrow$ $\Omega_{\rm m}$, σ_8, w_0, $\Gamma_{\rm dcdm}$, ...

$$\mathcal{L}(x \mid \boldsymbol{\theta}_k, M_k) = P\Big(x_o \mid \boldsymbol{\mu}_k^{\rm B}(\boldsymbol{\theta}_k), C, N_r\Big)$$

How the baryon model enters inference

Cosmological parameter inference

Observed κ maps $\Rightarrow$ ST data vector $\Rightarrow$ likelihood $\Rightarrow$ Ω_m, σ_8, w_0, Γ_dcdm, ...

$$\mathcal{L}(x \mid \boldsymbol{\theta}_k, M_k) = P\Big(x_o \mid \boldsymbol{\mu}_k^\mathrm{B}(\boldsymbol{\theta}_k), C, N_r\Big)$$

$$\boxed{\mu_{k,i}^\mathrm{B}(\boldsymbol{\theta}_k) = \underbrace{\mu_i(\boldsymbol{\theta}_k)}_{\text{emulated DMO prediction}} \; \underbrace{\mathcal{T}_{k,i}}_{\text{baryonic transfer function}}}$$

How the baryon model enters inference

Cosmological parameter inference

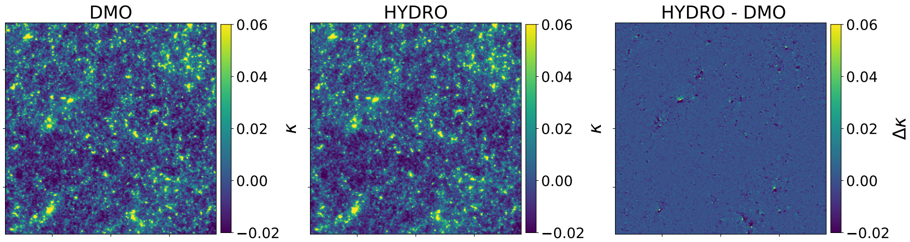
Observed κ maps $\Rightarrow$ ST data vector $\Rightarrow$ likelihood $\Rightarrow$ Ω_{m}, σ_8, w_0, Γ_{dcdm}, ...

$$\mathcal{L}(x \mid \boldsymbol{\theta}_k, M_k) = P\Big(x_o \mid \boldsymbol{\mu}_k^{\mathrm{B}}(\boldsymbol{\theta}_k), C, N_r\Big)$$

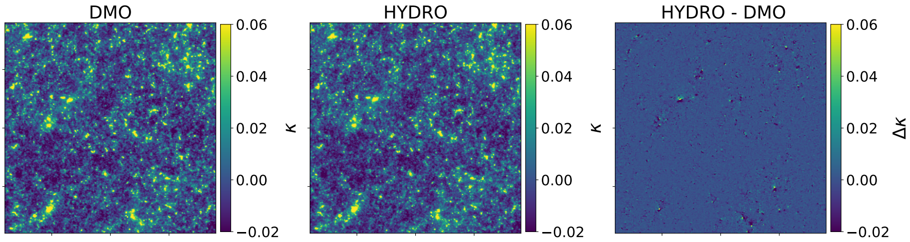
$$\mu_{k,i}^{\mathrm{B}}(\boldsymbol{\theta}_k) = \underbrace{\mu_i(\boldsymbol{\theta}_k)}_{\text{emulated DMO prediction}} \; \underbrace{\mathcal{T}_{k,i}}_{\text{baryonic transfer function}}$$

Algorithm: emulate cosmology, multiply by baryonic response, and infer parameters.

Transfer function



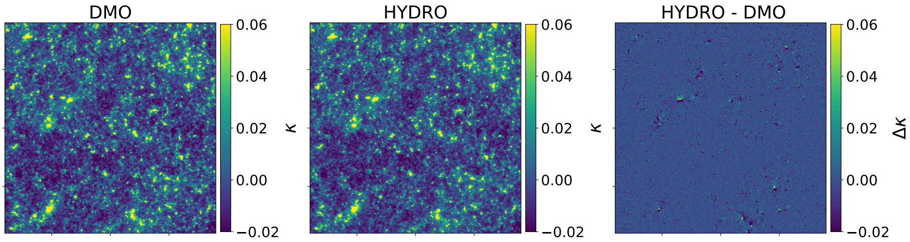
Transfer function



The baryonic suppression (or boost) factor for the total matter $P(k)$:

$$B = \frac{P^{\mathrm{HYDRO}}(k)}{P^{\mathrm{DMO}}(k)}$$

Transfer function



The baryonic suppression (or boost) factor for the total matter $P(k)$:

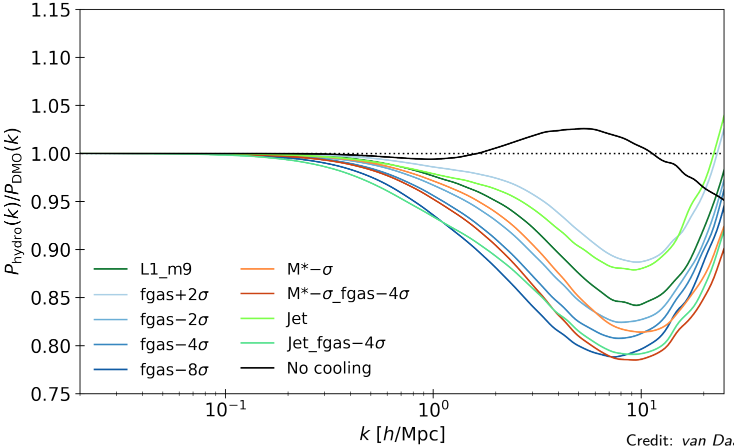
$$B = \frac{P^{\mathrm{HYDRO}}(k)}{P^{\mathrm{DMO}}(k)}$$

Similarly, we introduce **the baryonic transfer function for the scattering coefficients**:

$$\mathcal{T} = \frac{\mathcal{S}^{\mathrm{HYDRO}}}{\mathcal{S}^{\mathrm{DMO}}}$$

Baryonic feedback in FLAMINGO

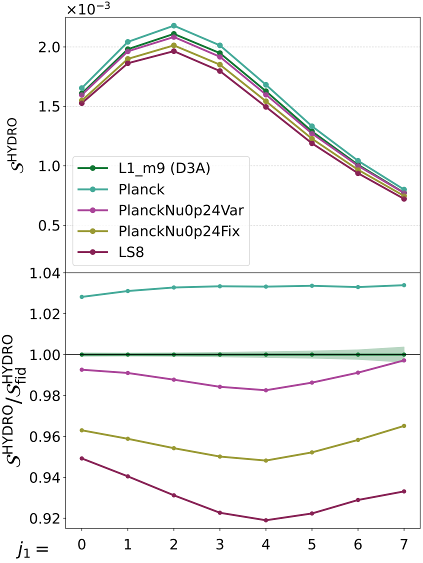
Calibrated to cluster gas fraction, fgas, and stellar mass function, $\mathrm{M}^\star$ at $z \approx 0$



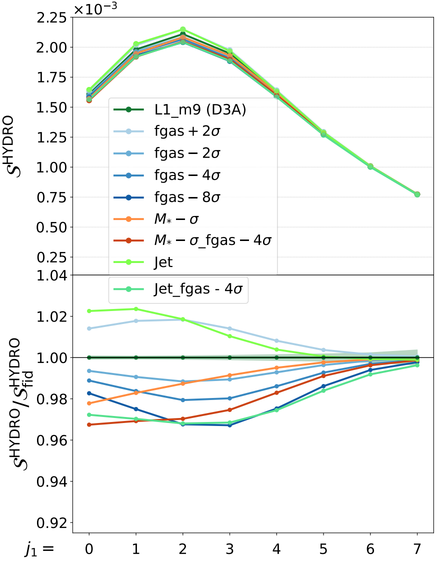
Credit: *van Daalen et al. (2025)*

Impact of baryons in ST

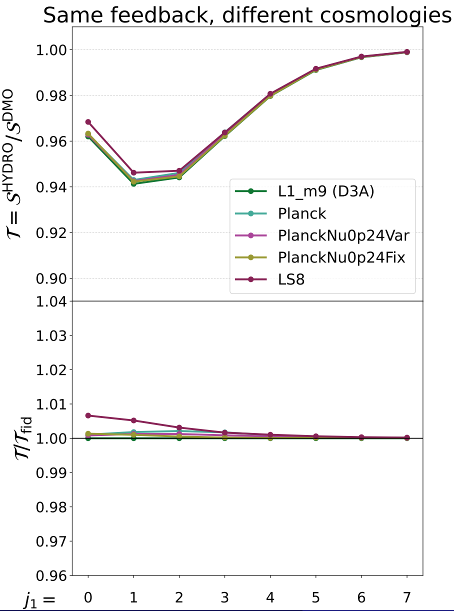
Same feedback, different cosmologies



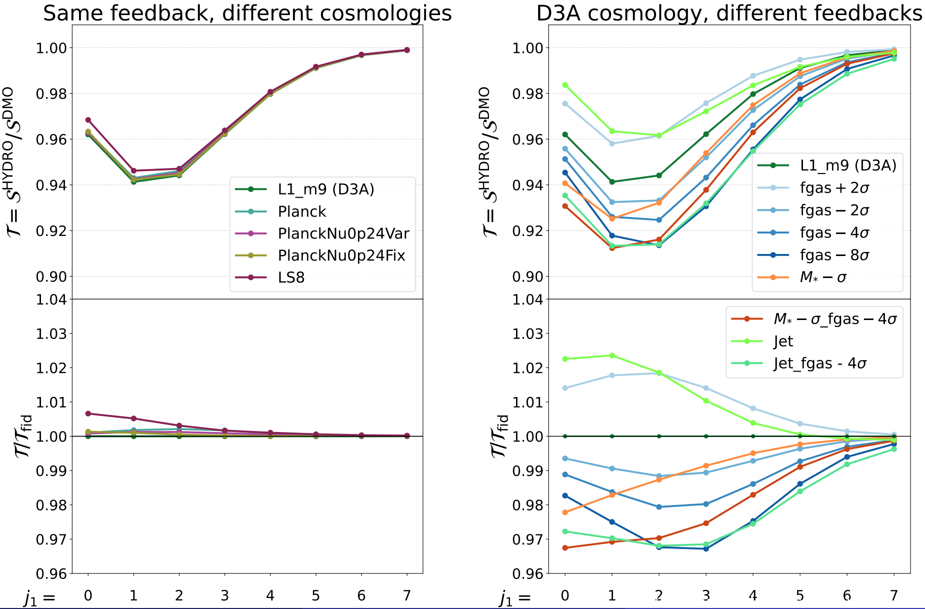
D3A cosmology, different feedbacks



ST transfer function $\mathcal{T} = \mathcal{S}^{HYDRO}/\mathcal{S}^{DMO}$

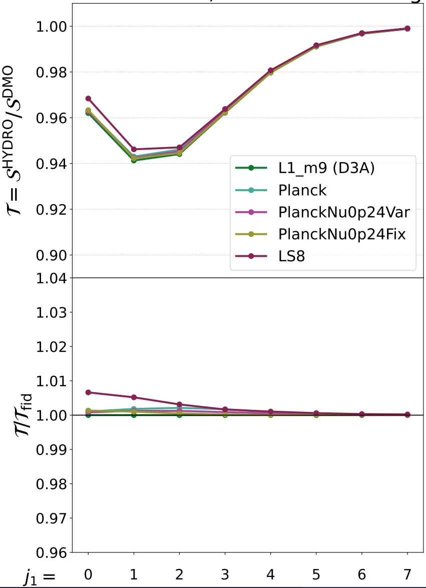


ST transfer function $\mathcal{T} = \mathcal{S}^{HYDRO}/\mathcal{S}^{DMO}$

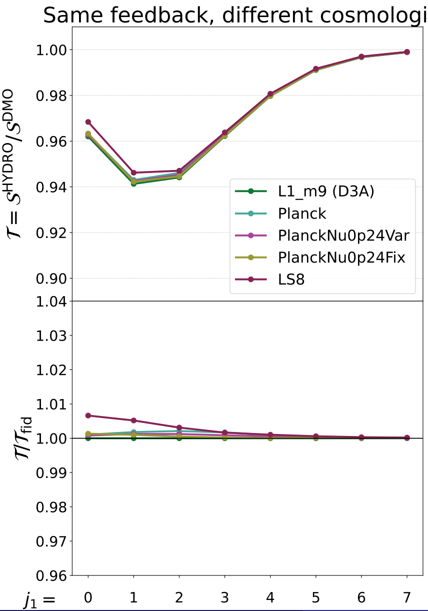


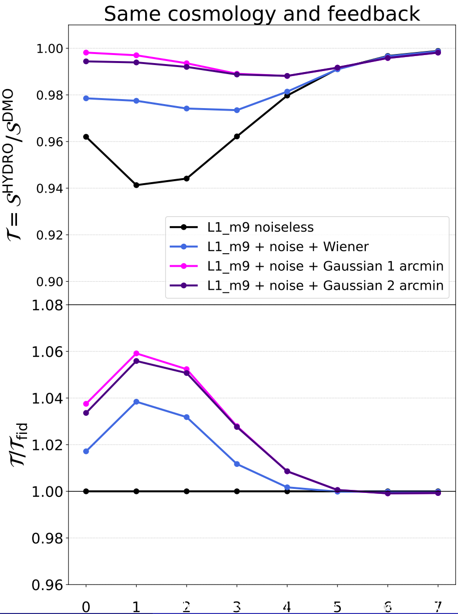
Noise and smoothing

Same feedback, different cosmologies

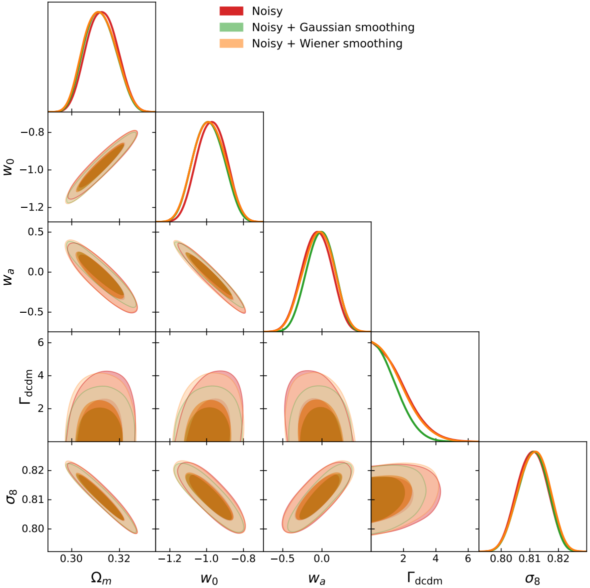


Noise and smoothing

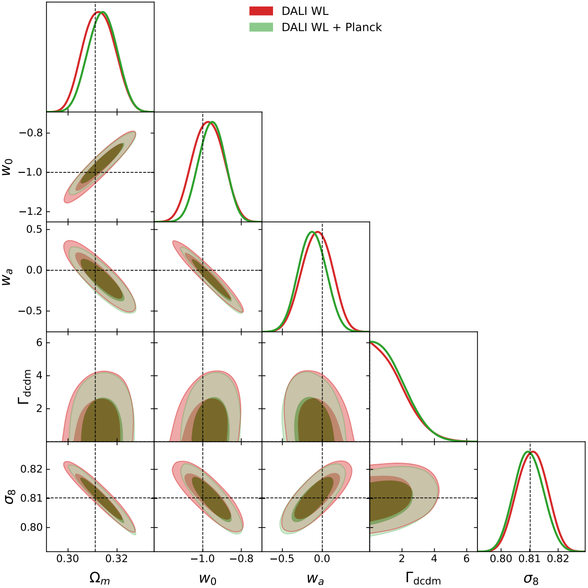




DCDM forecast: smoothing choices



DCDM constraints from ST + Planck

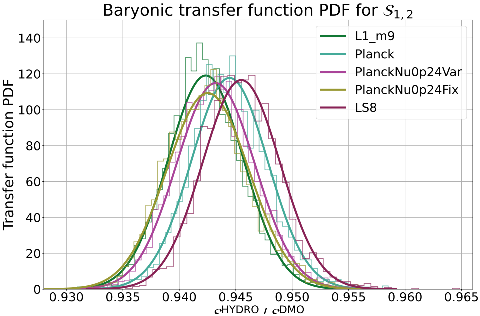


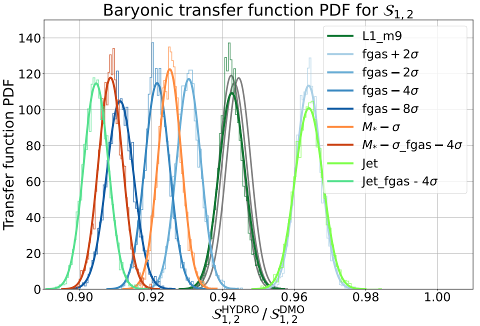
Take-home messages

- Scattering coefficients add **non-Gaussian information** beyond the power spectrum
- ST is **compact, stable and interpretable**
- ST alone provides tight constraints on cosmological parameters such as Γ_{dcdm}, Ω_{m}, σ_8, w_0, w_a
- Baryonic feedback can be modelled in a handy way by transfer function $\mathcal{T}$, but requires smart denoising techniques (e.g., Wiener filtering)

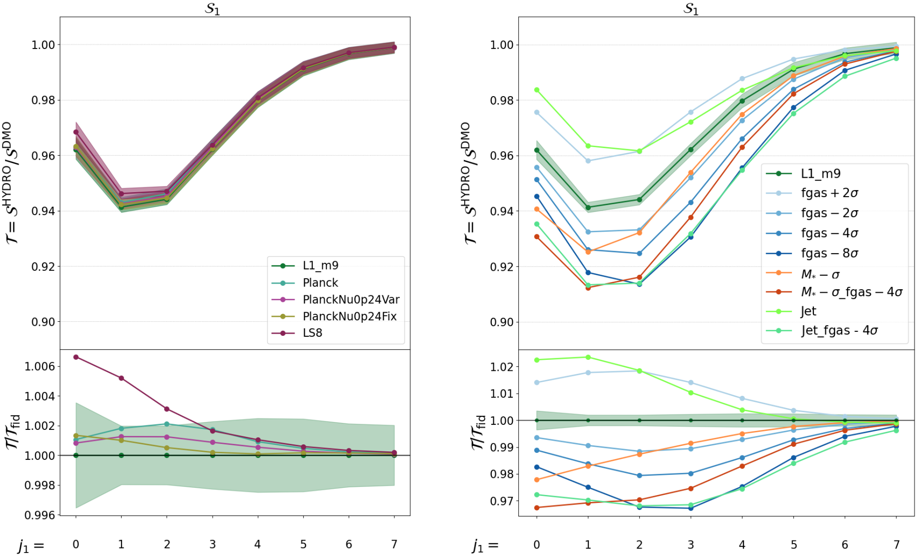
Thank you!

Backup slides


Baryonic transfer function PDF for $\mathcal{S}_{1,2}$
Transfer function PDF
$\mathcal{S}^{\mathrm{HYDRO}}_{1,2} / \mathcal{S}^{\mathrm{DMO}}_{1,2}$
0.930 0.935 0.940 0.945 0.950 0.955 0.960 0.965
0 20 40 60 80 100 120 140
L1_m9
Planck
PlanckNu0p24Var
PlanckNu0p24Fix
LS8

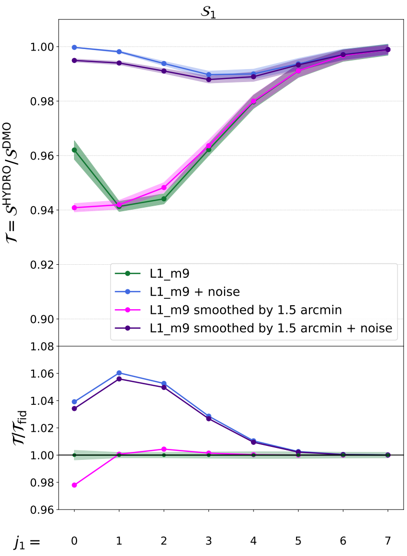

Baryonic transfer function PDF for $\mathcal{S}_{1,2}$
Transfer function PDF
$\mathcal{S}^{\mathrm{HYDRO}}_{1,2} / \mathcal{S}^{\mathrm{DMO}}_{1,2}$
0.90 0.92 0.94 0.96 0.98 1.00
0 20 40 60 80 100 120 140
L1_m9
fgas $+ 2\sigma$
fgas $- 2\sigma$
fgas $- 4\sigma$
fgas $- 8\sigma$
$M_* - \sigma$
$M_* - \sigma$_fgas $- 4\sigma$
Jet
Jet_fgas - 4σ

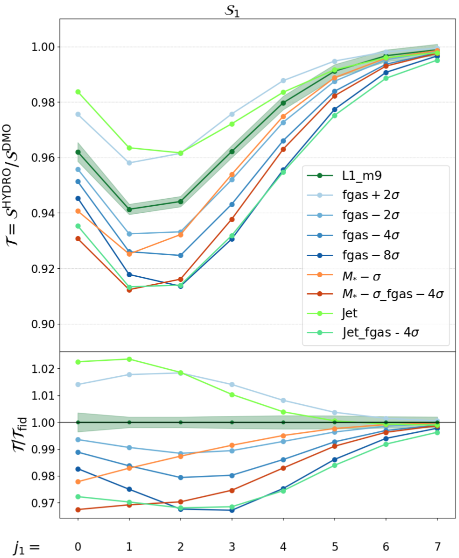
With std

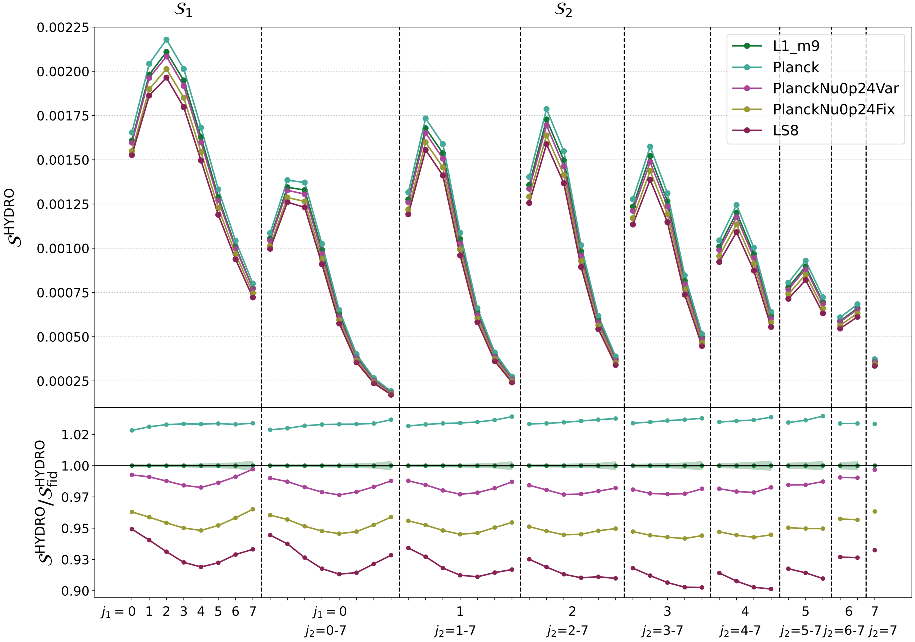


$\mathcal{S}_1$
$\mathcal{T} = S^{\rm HYDRO}/S^{\rm DMO}$
$\mathcal{T}/\mathcal{T}_{\rm fid}$
L1_m9
Planck
PlanckNu0p24Var
PlanckNu0p24Fix
LS8
$\mathcal{S}_1$
L1_m9
fgas + 2σ
fgas − 2σ
fgas − 4σ
fgas − 8σ
$M_* - \sigma$
$M_* - \sigma$_fgas − 4σ
Jet
Jet_fgas - 4σ
$j_1 =$
0 1 2 3 4 5 6 7

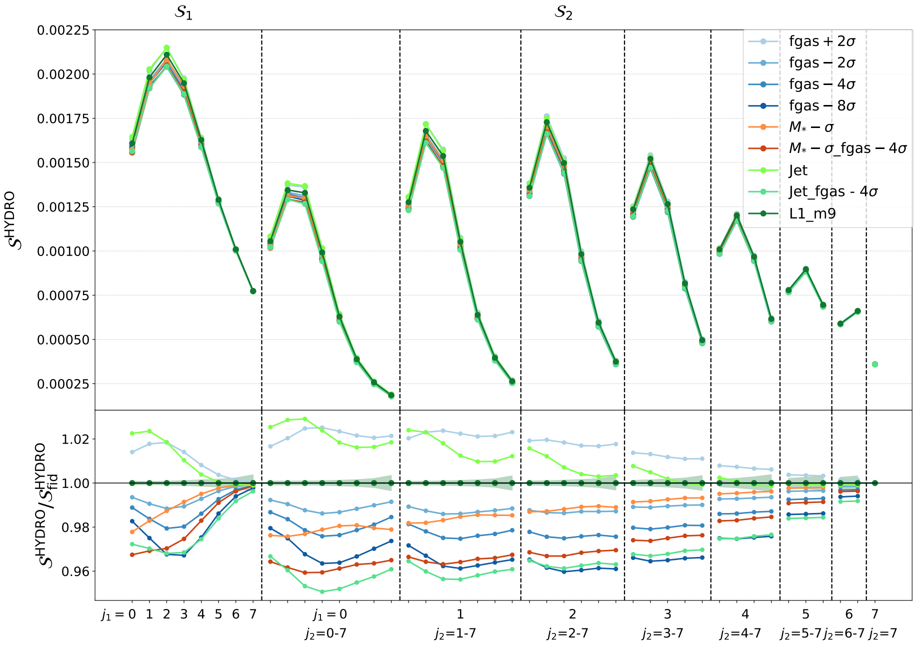
With std

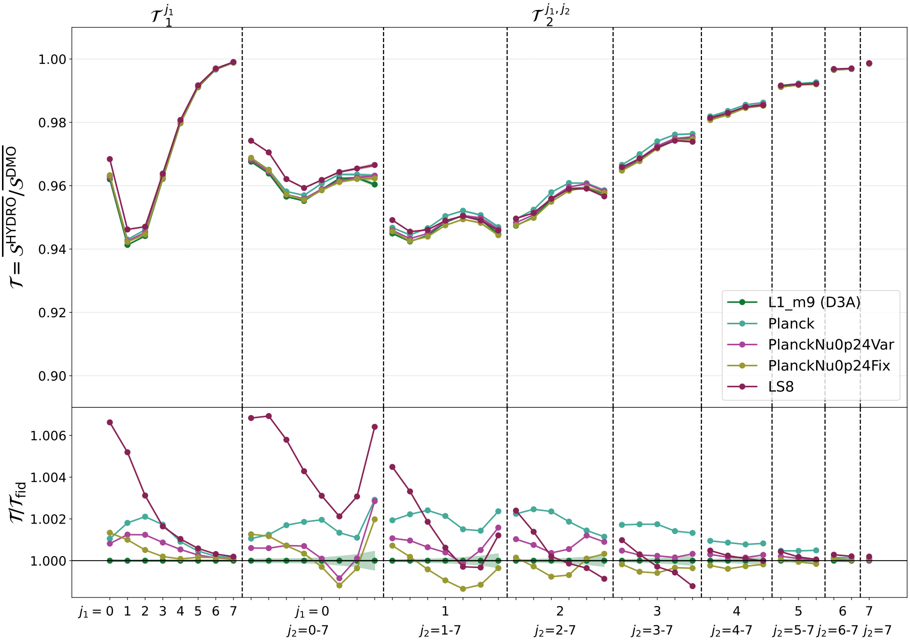

$\mathcal{S}_1$
$\mathcal{T} = \mathcal{S}^{\mathrm{HYDRO}}/\mathcal{S}^{\mathrm{DMO}}$
L1_m9
L1_m9 + noise
L1_m9 smoothed by 1.5 arcmin
L1_m9 smoothed by 1.5 arcmin + noise
$\mathcal{T}/\mathcal{T}_{\mathrm{fid}}$
$j_1 =$ 0 1 2 3 4 5 6 7

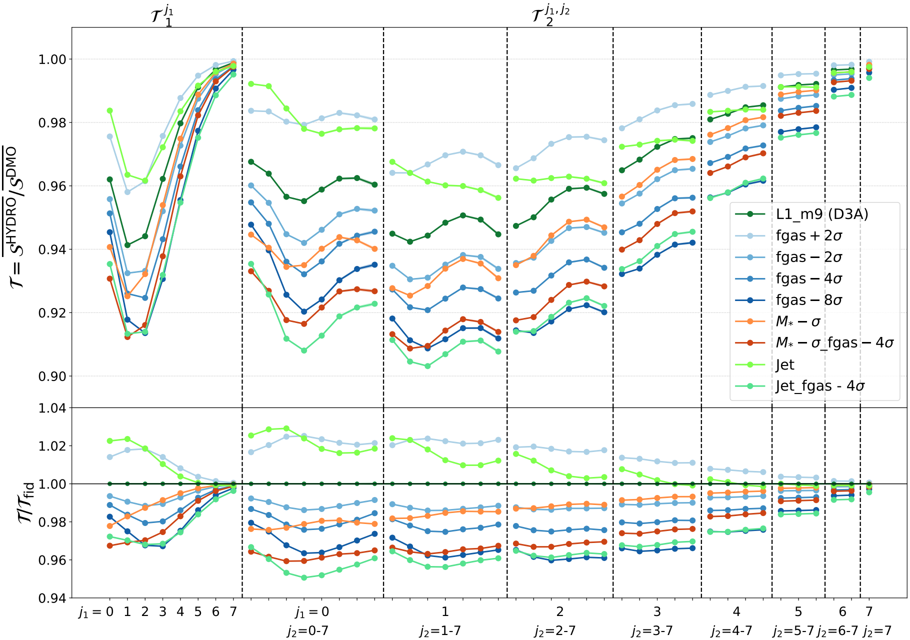

$\mathcal{S}_1$
$\mathcal{T} = \mathcal{S}^{\mathrm{HYDRO}}/\mathcal{S}^{\mathrm{DMO}}$
L1_m9
fgas + 2σ
fgas − 2σ
fgas − 4σ
fgas − 8σ
$M_* - \sigma$
$M_* - \sigma$_fgas − 4σ
Jet
Jet_fgas - 4σ
$\mathcal{T}/\mathcal{T}_{\mathrm{fid}}$
$j_1 =$ 0 1 2 3 4 5 6 7



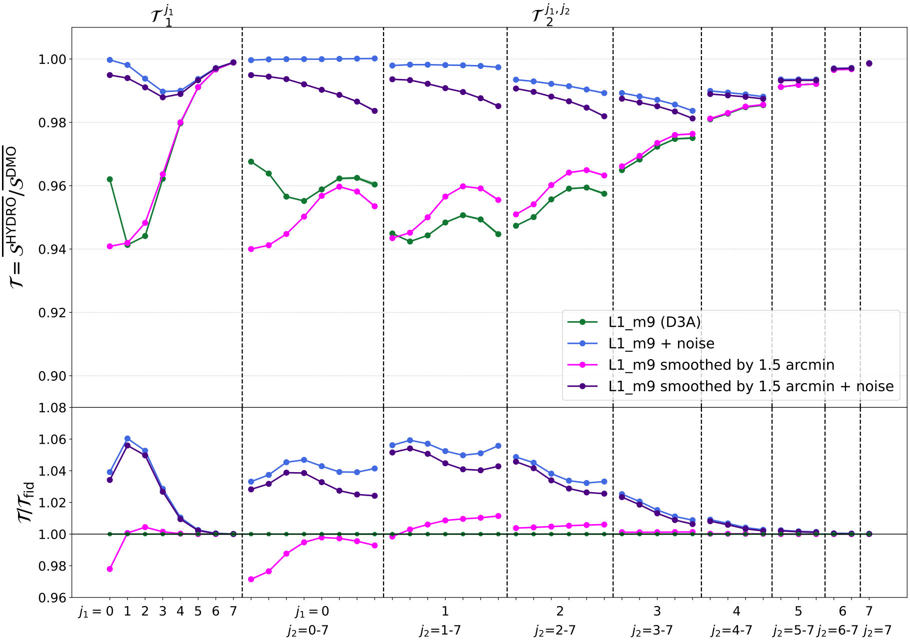
$\mathcal{S}_1$
$\mathcal{S}_2$
L1_m9
Planck
PlanckNu0p24Var
PlanckNu0p24Fix
LS8
$\mathcal{S}^{\rm HYDRO}$
$\mathcal{S}^{\rm HYDRO}/\mathcal{S}^{\rm HYDRO}_{\rm fid}$
0.00225
0.00200
0.00175
0.00150
0.00125
0.00100
0.00075
0.00050
0.00025
1.02
1.00
0.97
0.95
0.93
0.90
$j_1 = 0$ 1 2 3 4 5 6 7
$j_1 = 0$
j_2=0-7
1
j_2=1-7
2
j_2=2-7
3
j_2=3-7
4
j_2=4-7
5
j_2=5-7
6
j_2=6-7
7
j_2=7


$\mathcal{S}_1$
$\mathcal{S}_2$
$\mathcal{S}^{\mathrm{HYDRO}}$
$\mathcal{S}^{\mathrm{HYDRO}}/\mathcal{S}^{\mathrm{HYDRO}}_{\mathrm{fid}}$
fgas + 2σ
fgas − 2σ
fgas − 4σ
fgas − 8σ
$M_* - \sigma$
$M_* - \sigma$_fgas − 4σ
Jet
Jet_fgas - 4σ
L1_m9
$j_1 = 0$ 1 2 3 4 5 6 7
$j_1 = 0$ j_2=0-7
1 j_2=1-7
2 j_2=2-7
3 j_2=3-7
4 j_2=4-7
5 j_2=5-7
6 j_2=6-7
7 j_2=7

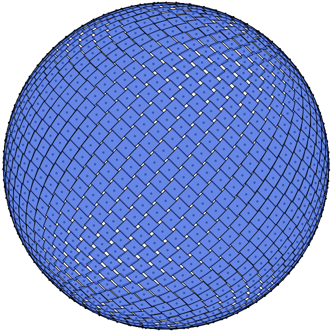

$\mathcal{T}_1^{j_1}$
$\mathcal{T}_2^{j_1,j_2}$
$\mathcal{T} = \overline{S^{\mathrm{HYDRO}}} / \overline{S^{\mathrm{DMO}}}$
$\mathcal{T}/\mathcal{T}_{\mathrm{fid}}$
L1_m9 (D3A)
Planck
PlanckNu0p24Var
PlanckNu0p24Fix
LS8
1.00
0.98
0.96
0.94
0.92
0.90
1.006
1.004
1.002
1.000
$j_1 = 0$ 1 2 3 4 5 6 7
$j_1 = 0$
j_2=0-7
1
j_2=1-7
2
j_2=2-7
3
j_2=3-7
4
j_2=4-7
5
j_2=5-7
6
j_2=6-7
7
j_2=7

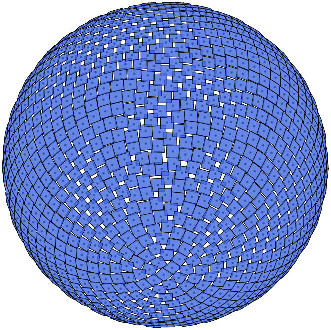


$\mathcal{T}_1^{j_1}$
$\mathcal{T}_2^{j_1,j_2}$
$\mathcal{T} = \overline{S^{\mathrm{HYDRO}}}/\overline{S^{\mathrm{DMO}}}$
$\mathcal{T}/\mathcal{T}_{\mathrm{fid}}$
L1_m9 (D3A)
fgas + 2σ
fgas − 2σ
fgas − 4σ
fgas − 8σ
$M_* - \sigma$
$M_* - \sigma$_fgas − 4σ
Jet
Jet_fgas - 4σ
$j_1 = 0$ 1 2 3 4 5 6 7
$j_1 = 0$ 1 2 3 4 5 6 7
j_2=0-7 j_2=1-7 j_2=2-7 j_2=3-7 j_2=4-7 j_2=5-7 j_2=6-7 j_2=7



$\mathcal{T}_1^{j_1}$
$\mathcal{T}_2^{j_1,j_2}$
$\mathcal{T} = \overline{S^{\mathrm{HYDRO}}} / \overline{S^{\mathrm{DMO}}}$
$\mathcal{T}/\mathcal{T}_{\mathrm{fid}}$
L1_m9 (D3A)
L1_m9 + noise
L1_m9 smoothed by 1.5 arcmin
L1_m9 smoothed by 1.5 arcmin + noise
$j_1 = 0$ 1 2 3 4 5 6 7
$j_1 = 0$ j_2=0-7
1 j_2=1-7
2 j_2=2-7
3 j_2=3-7
4 j_2=4-7
5 j_2=5-7
6 j_2=6-7
7 j_2=7





Scales

Table 2. Scale definitions used throughout the paper, with corresponding diadic scales j, Fourier wavenumbers k, and real-space scales D at $z \approx 1$.

Name	j	k ($h\,\mathrm{Mpc}^{-1}$)	D ($\mathrm{Mpc}\,h^{-1}$)
Smallest	0	3.6	0.3
Small	1 – 2	1.8 - 0.9	0.6 – 1
Intermediate	3 – 5	0.4 – 0.1	2 – 9
Large	6 – 7	0.06 – 0.03	18 – 36