Breaking degeneracies in modified gravity cosmologies with WL

Austin Peel

Outline

1. Modified gravity simulations
2. Aperture mass statistics
3. Distinguishing cosmological models
4. A machine learning approach
5. Summary

Outline

1. Modified gravity simulations
2. Aperture mass statistics
3. Distinguishing cosmological models
4. A machine learning approach
5. Summary

Motivation

Dark energy is still a (big) problem.

If a **non-standard gravity** universe is masquerading as ΛCDM, can we find out using **weak lensing** ?

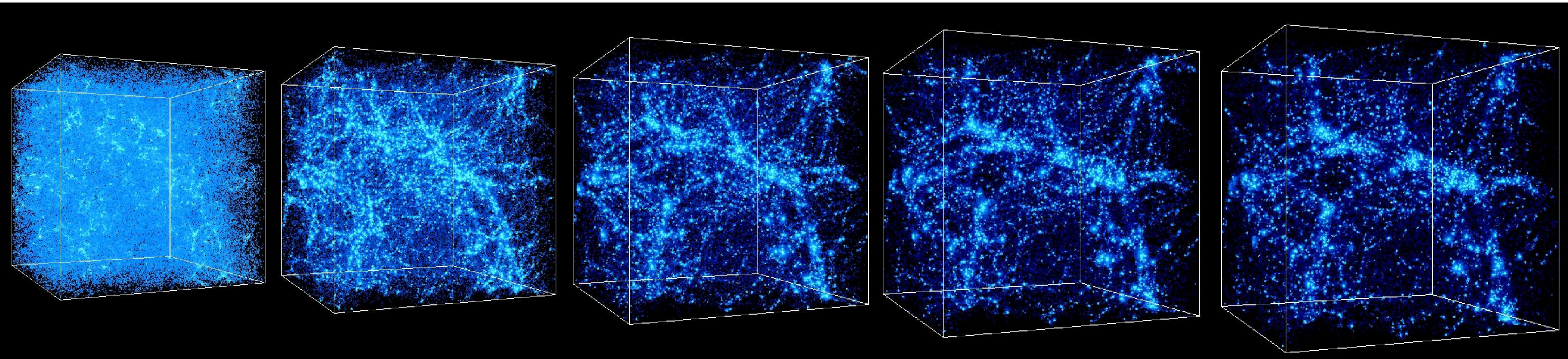
Reference

A. Peel et al., in press A&A (2018) [arXiv:1805.05146]

DUSTGRAIN-*pathfinder* simulations C. Giocoli et al. 2018 [arXiv:1806.04681]

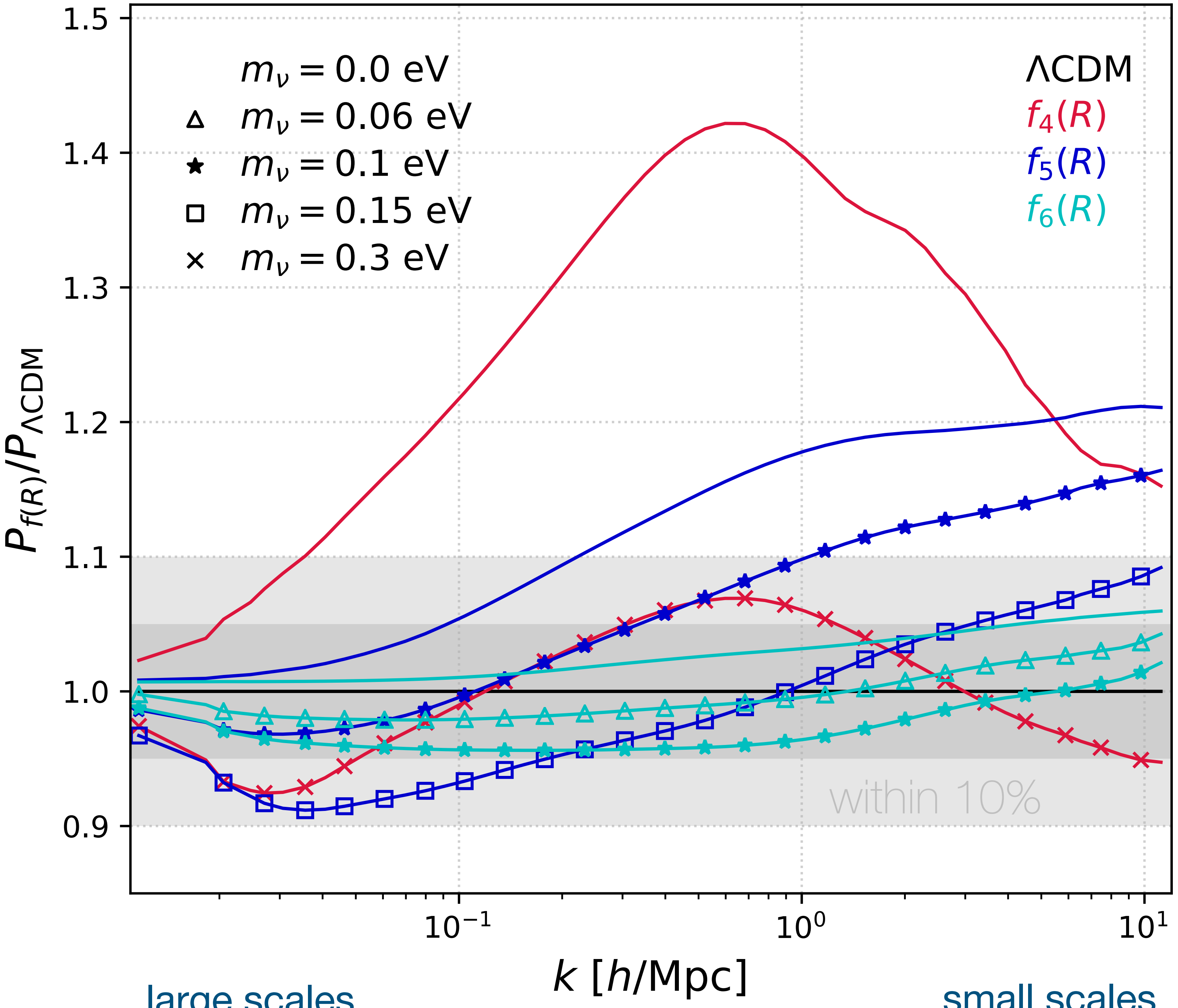
Sample joint parameter space of $f(R)$ gravity and massive neutrino cosmologies

Performed with MG-Gadget code that implements the extra fifth-force and screening



NOT DUSTGRAIN snapshots—for illustration only (http://cosmicweb.uchicago.edu/filaments.html)

Simulation Name	Gravity type	f_{R0}	m_ν [eV]	$\Omega_{\rm CDM}$	Ω_ν	$m^p_{\rm CDM}$ [M$_\odot$/h]	m^p_ν [M$_\odot$/h]
ΛCDM	GR	–	0	0.31345	0	8.1×10^{10}	0
fR4	$f(R)$	-1×10^{-4}	0	0.31345	0	8.1×10^{10}	0
fR5	$f(R)$	-1×10^{-5}	0	0.31345	0	8.1×10^{10}	0
fR6	$f(R)$	-1×10^{-6}	0	0.31345	0	8.1×10^{10}	0
fR4-0.3eV	$f(R)$	-1×10^{-4}	0.3	0.30630	0.00715	7.92×10^{10}	1.85×10^{9}
fR5-0.15eV	$f(R)$	-1×10^{-5}	0.15	0.30987	0.00358	8.01×10^{10}	9.25×10^{8}
fR5-0.1eV	$f(R)$	-1×10^{-5}	0.1	0.31107	0.00238	8.04×10^{10}	6.16×10^{8}
fR6-0.1eV	$f(R)$	-1×10^{-6}	0.1	0.31107	0.00238	8.04×10^{10}	6.16×10^{8}
fR6-0.06eV	$f(R)$	-1×10^{-6}	0.06	0.31202	0.00143	8.07×10^{10}	3.7×10^{8}



large scales (linear)

small scales (nonlinear)

matter power spectra (relative to LCDM)

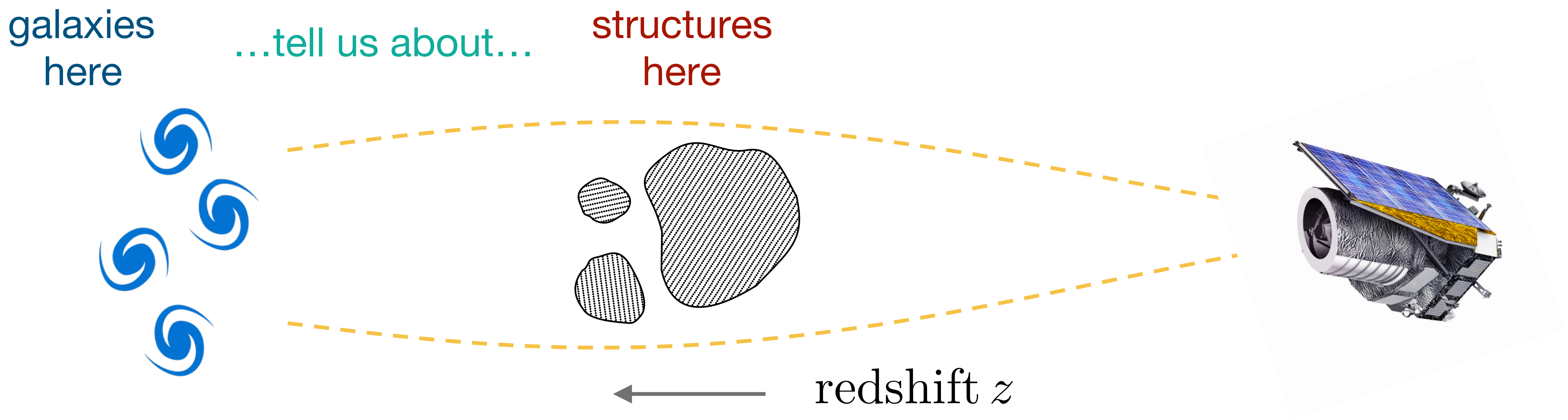
$f_4(R)$ farther from LCDM

$f_5(R)$ intermediate

$f_6(R)$ closer to LCDM

neutrinos suppress the growth of structure

Weak lensing maps from ray tracing

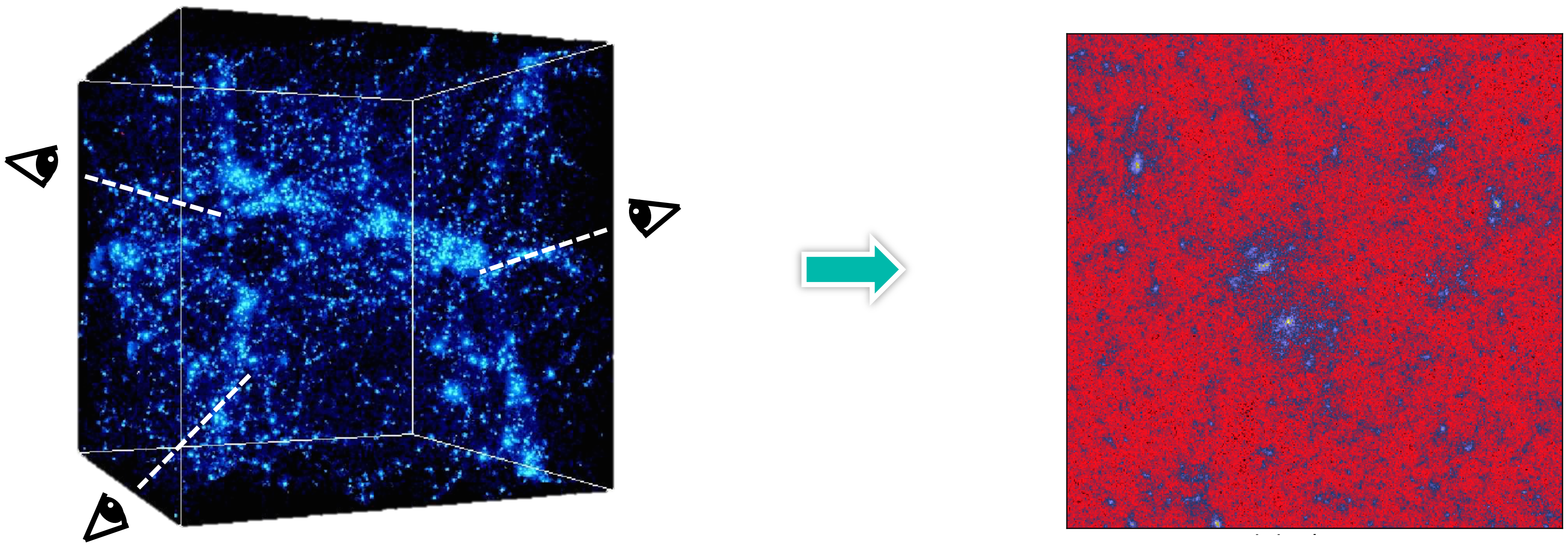


galaxies here

…tell us about…

structures here

redshift z



$\times\ 256$ per redshift, per model

WL convergence map $\kappa(\boldsymbol{\theta})$

Outline

1. Modified gravity simulations
2. **Aperture mass statistics**
3. Distinguishing cosmological models
4. A machine learning approach
5. Summary

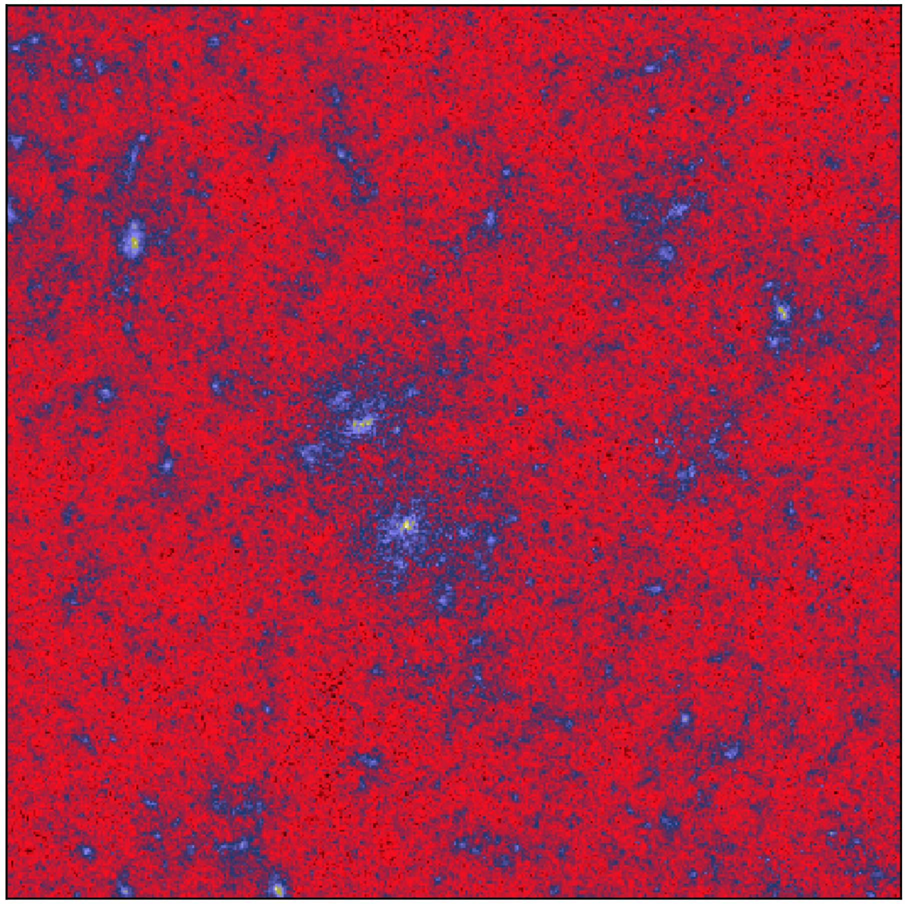
isotropic filter function

aperture mass :

$$M_{\rm ap}(\boldsymbol{\theta};\vartheta) = \int \mathrm{d}^2\theta' \, U_\vartheta(|\boldsymbol{\theta}' - \boldsymbol{\theta}|) \, \kappa(\boldsymbol{\theta}')$$

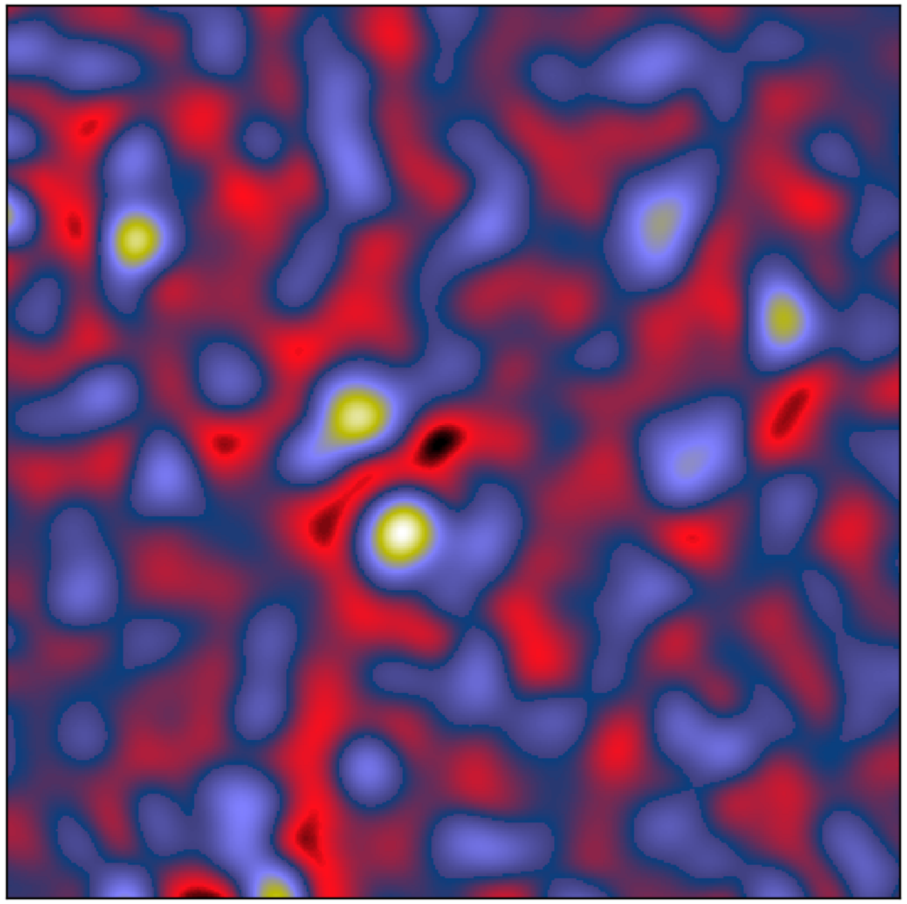
mass map

implemented as a **wavelet** transform (starlet)



original κ map

$$= \sum$$

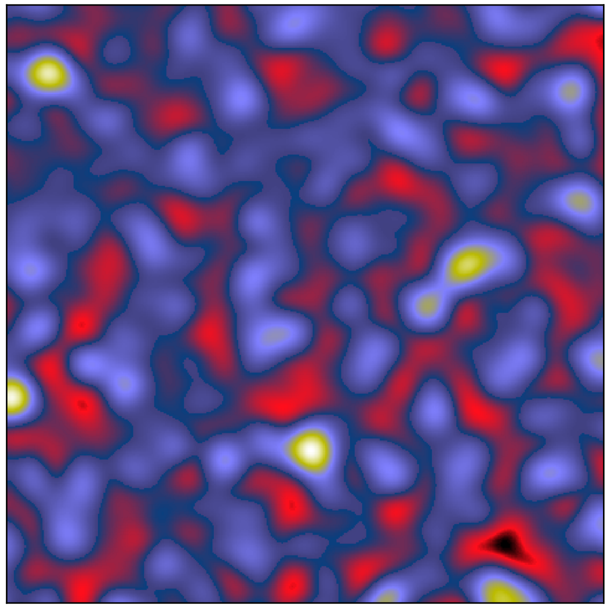


aperture size $\vartheta_5 = 4.69$ arcmin

Breaking Degeneracies in MG with WL

aperture mass map

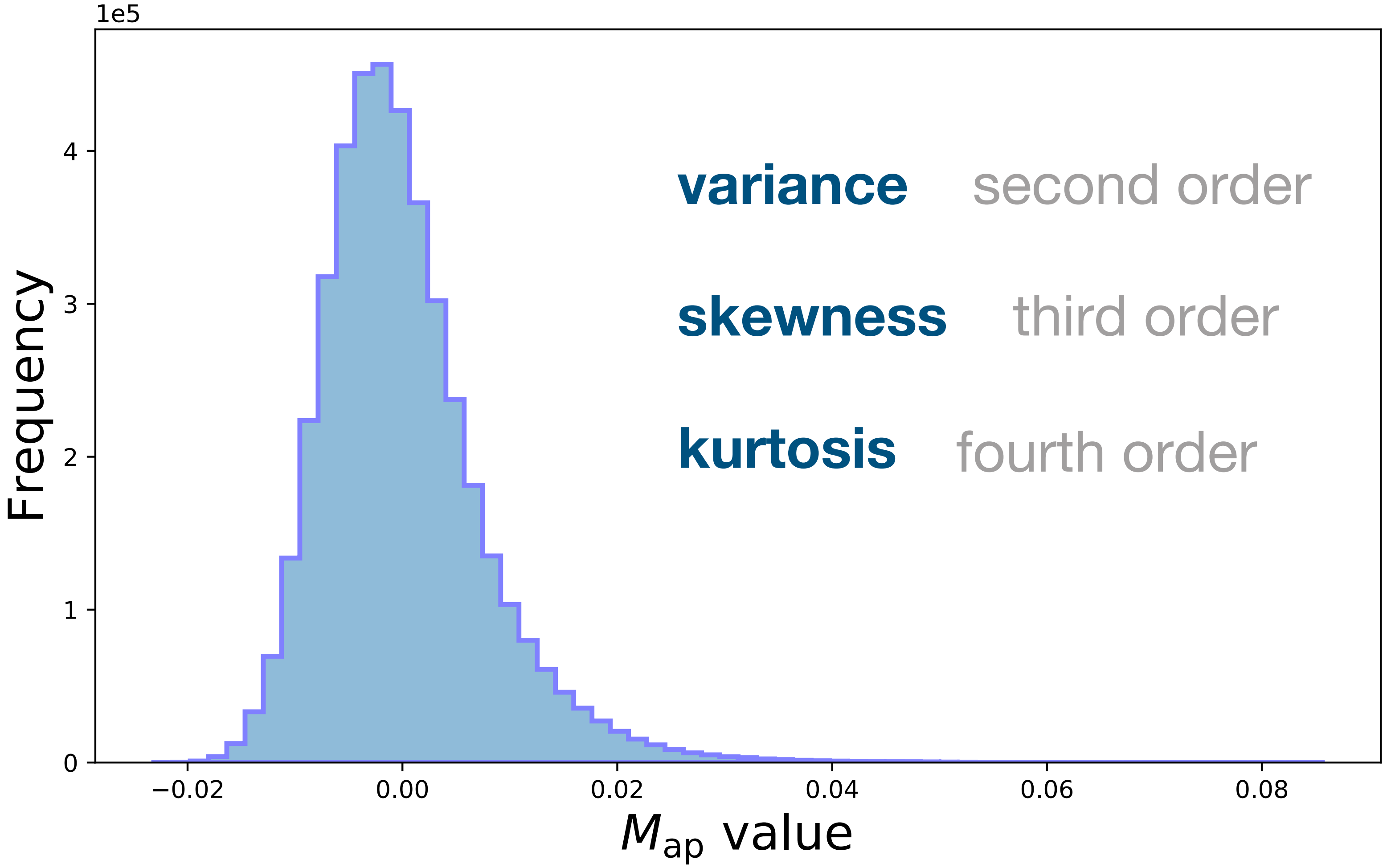
$$M_{\mathrm{ap}}(\mathrm{model}, \vartheta_j, z_s) =$$



$5 \times 5\ \mathrm{deg}^2$

400×400 shown, but

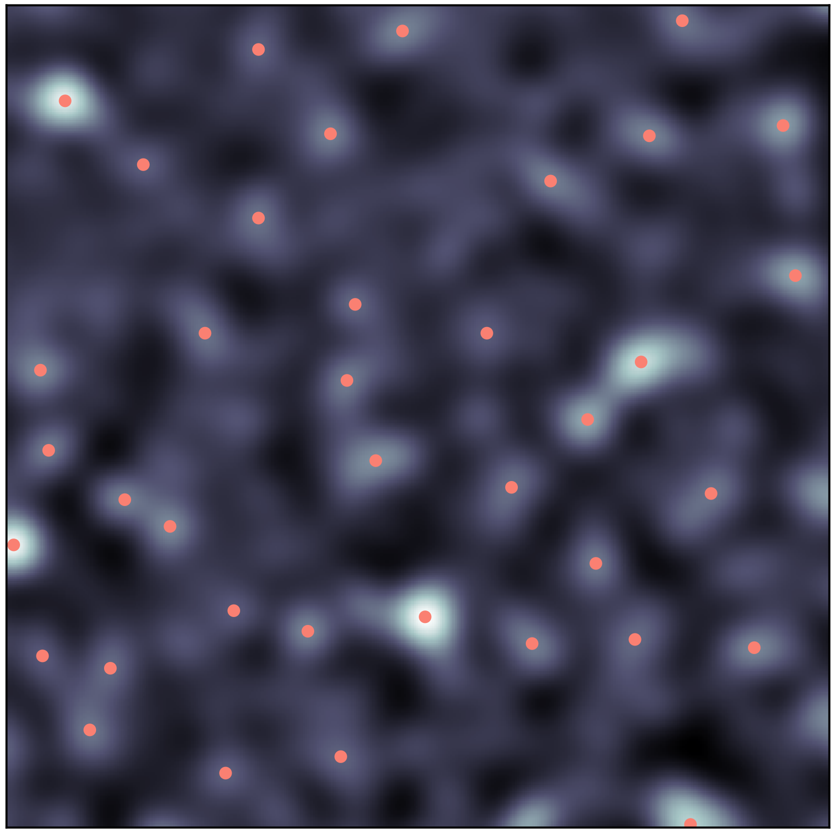
2048×2048 in practice



variance second order

skewness third order

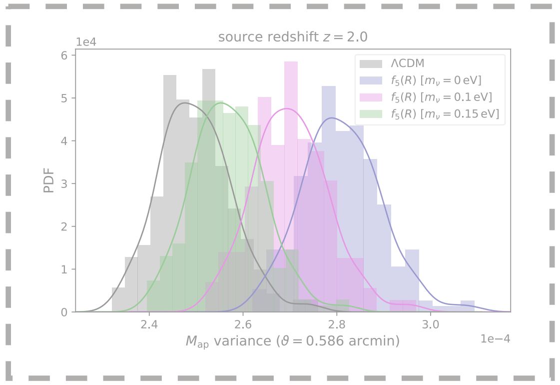
kurtosis fourth order

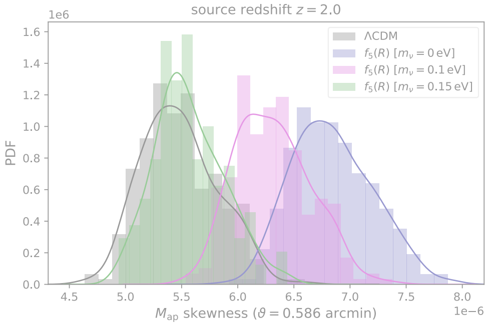


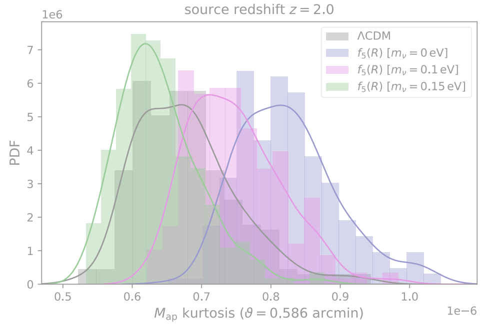
peak count

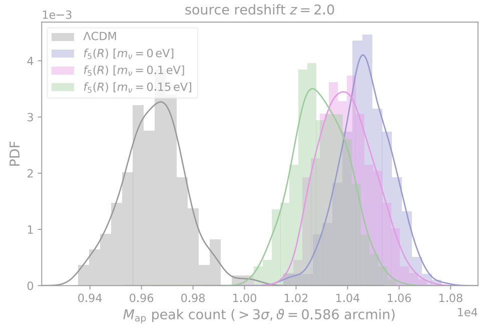
Outline

Breaking Degeneracies in MG with WL

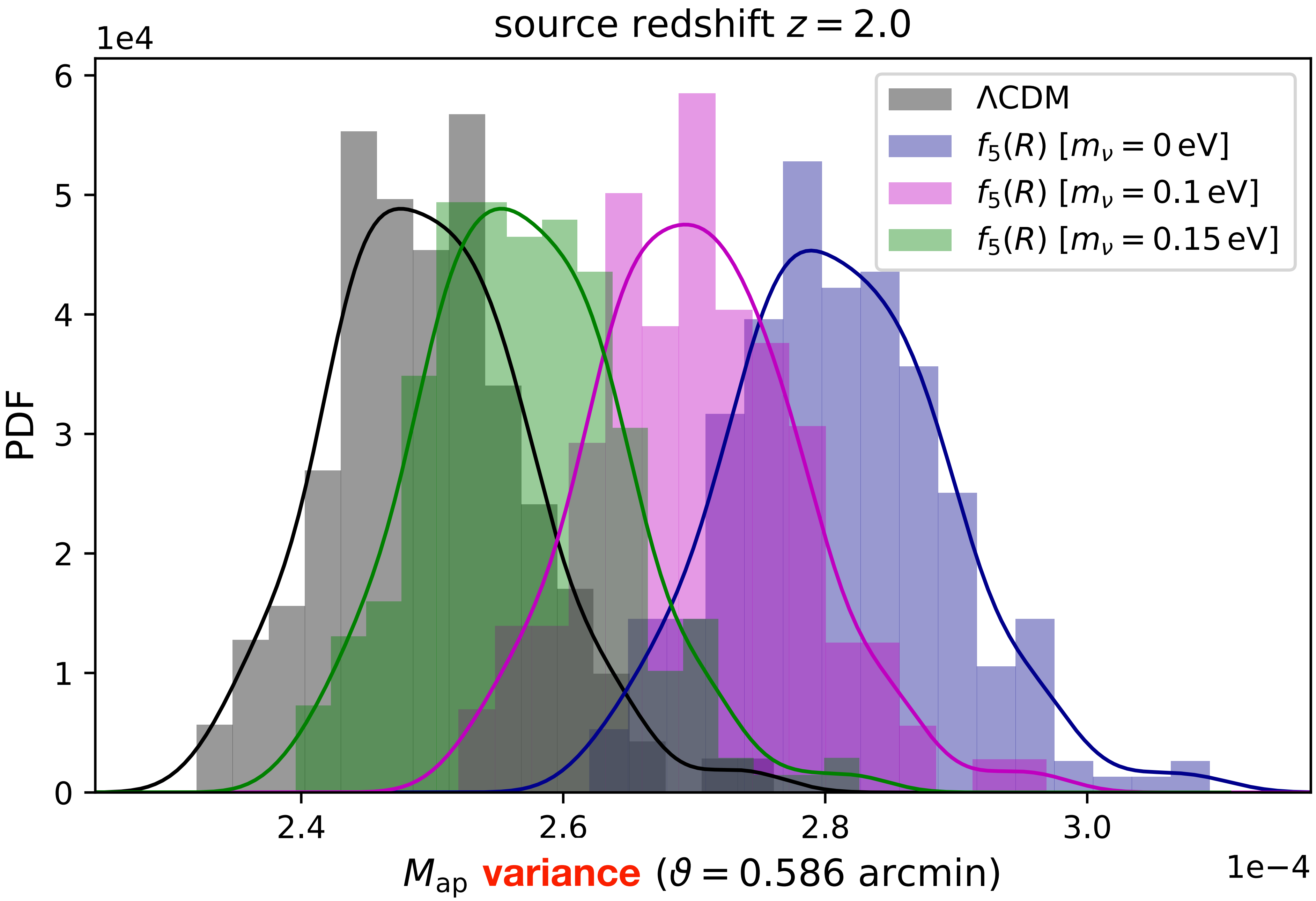




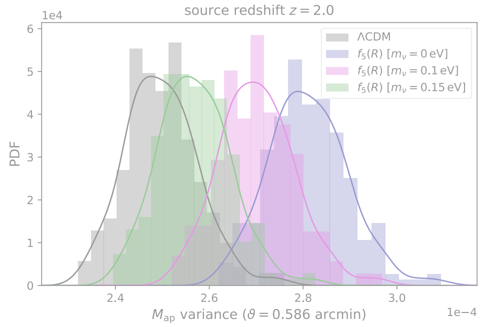


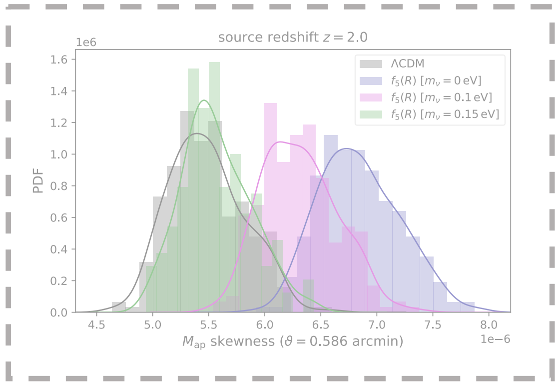


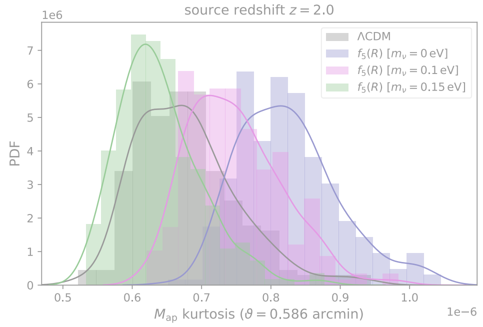
distributions of observables

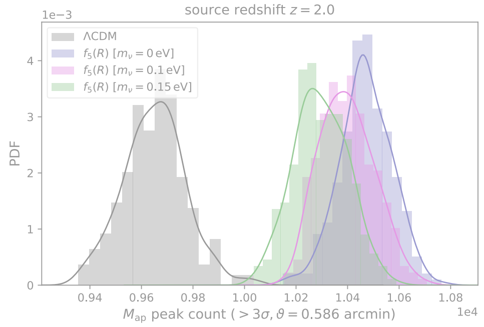


Breaking Degeneracies in MG with WL

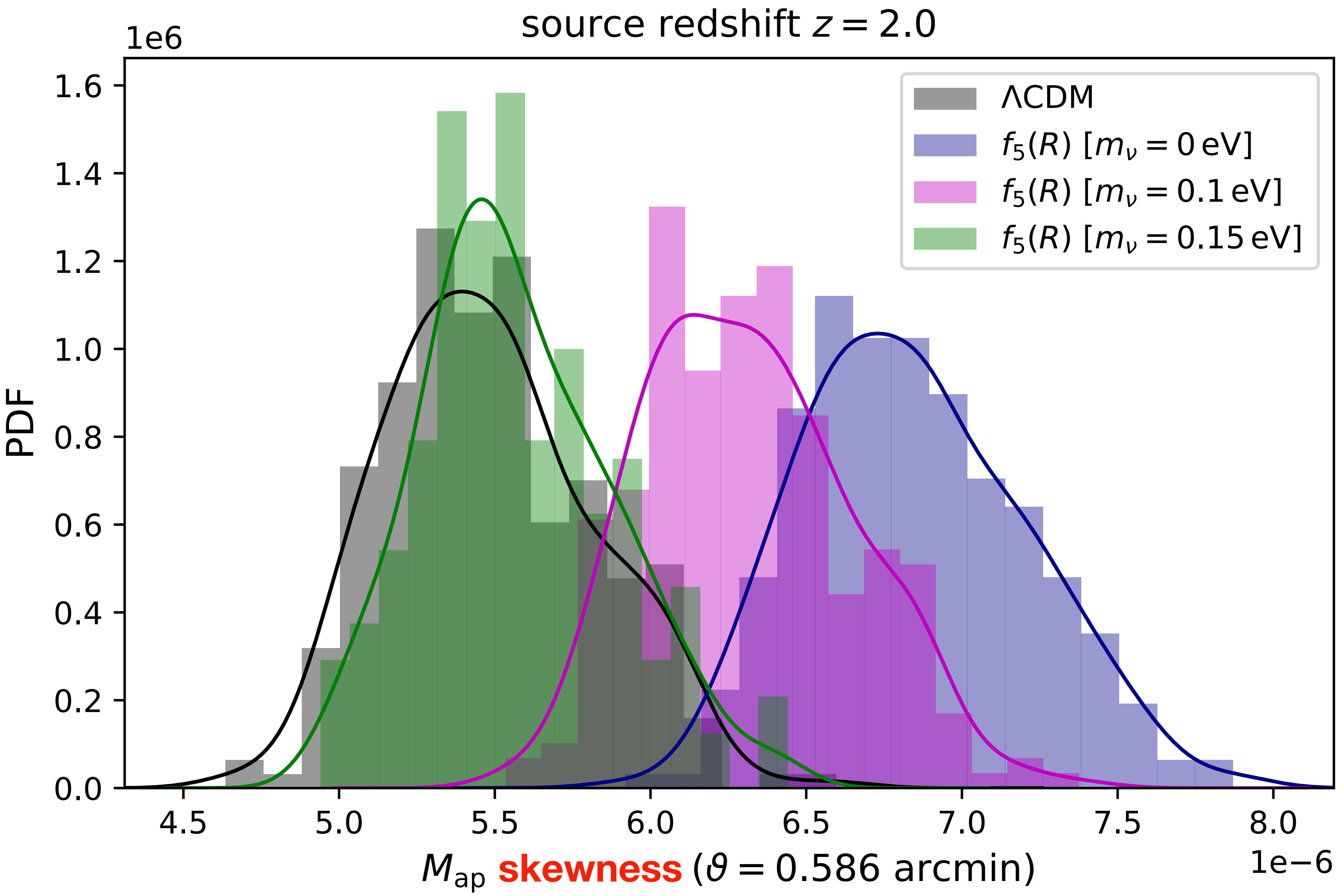




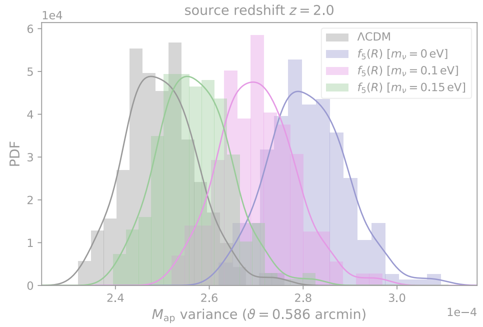


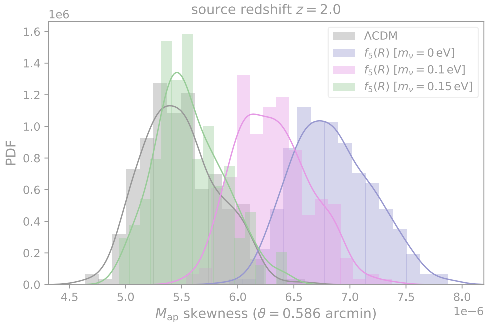


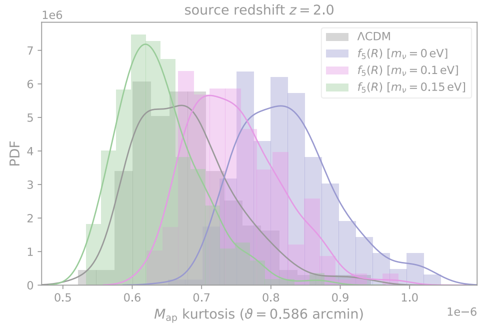
distributions of observables

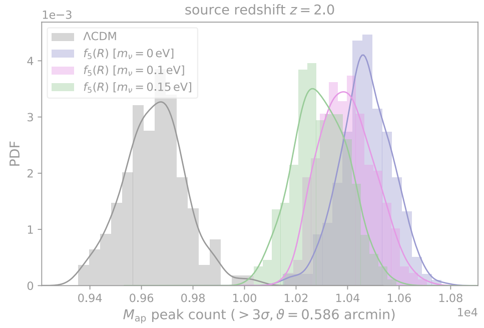


Breaking Degeneracies in MG with WL

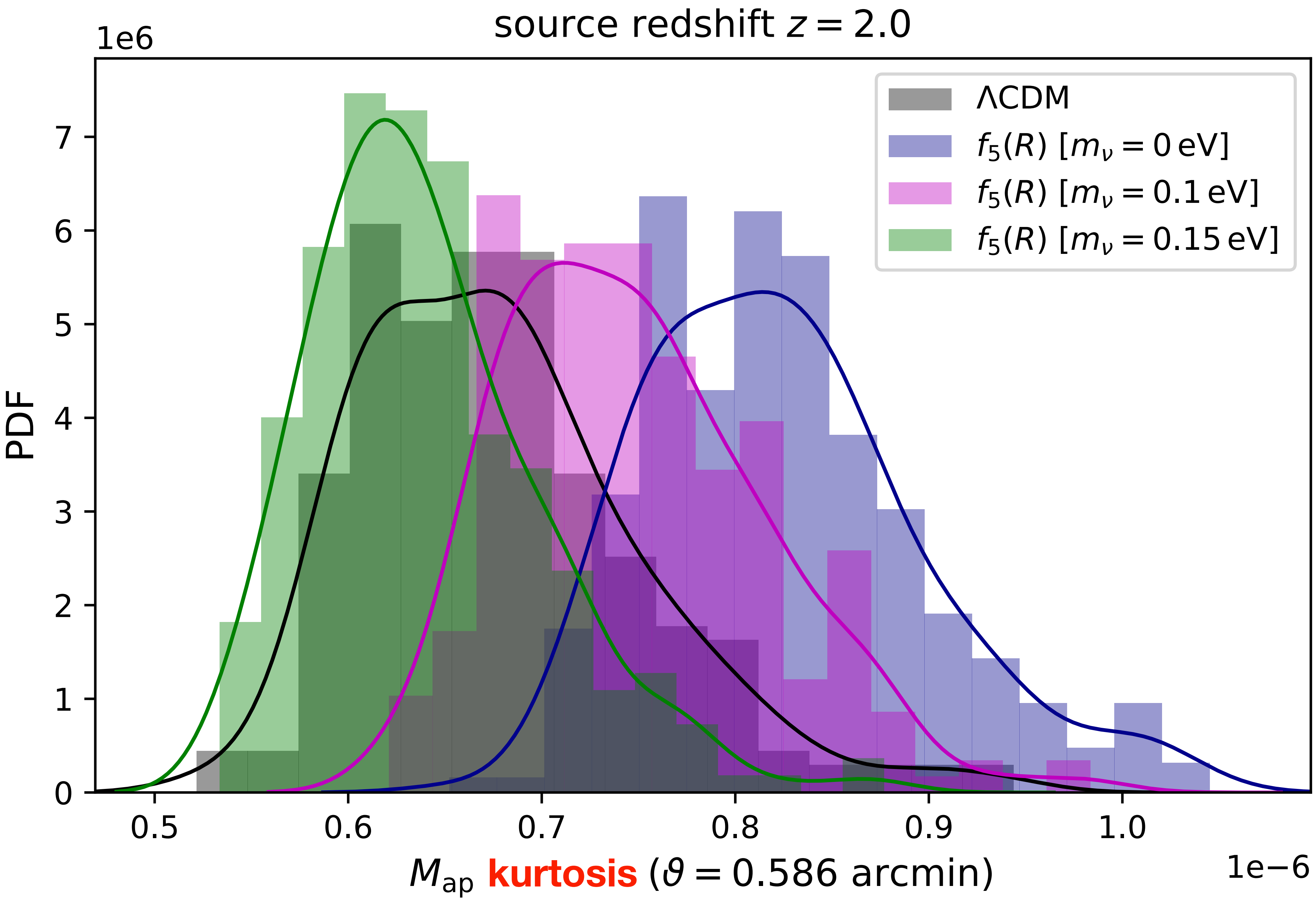




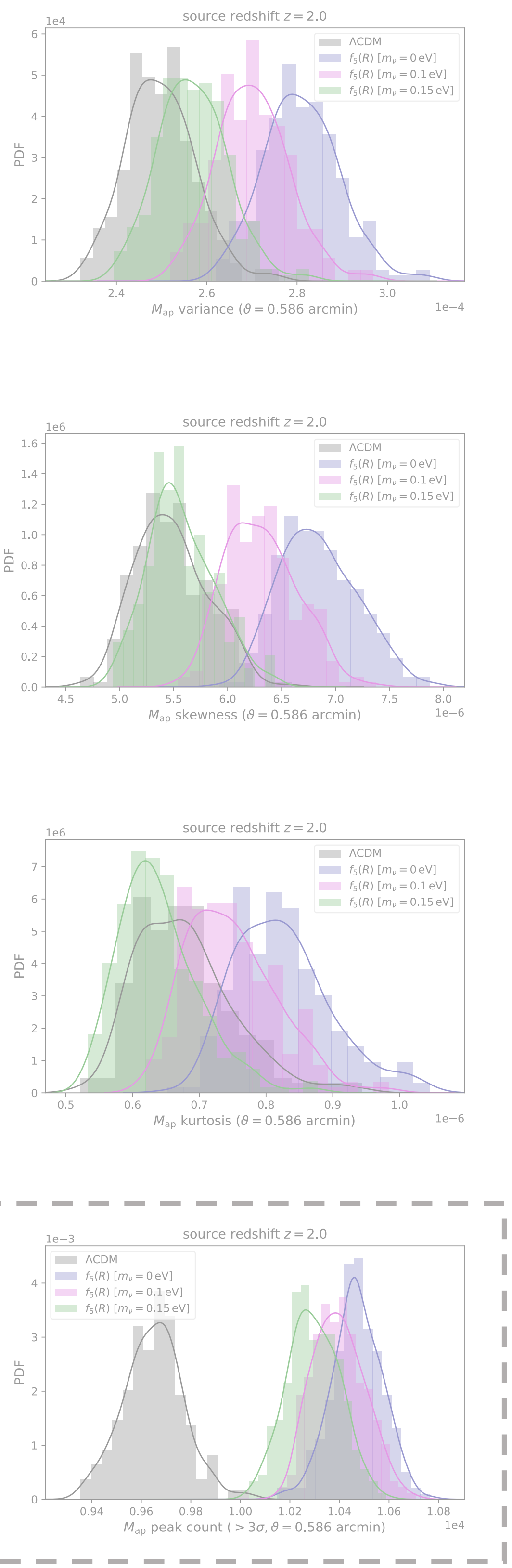




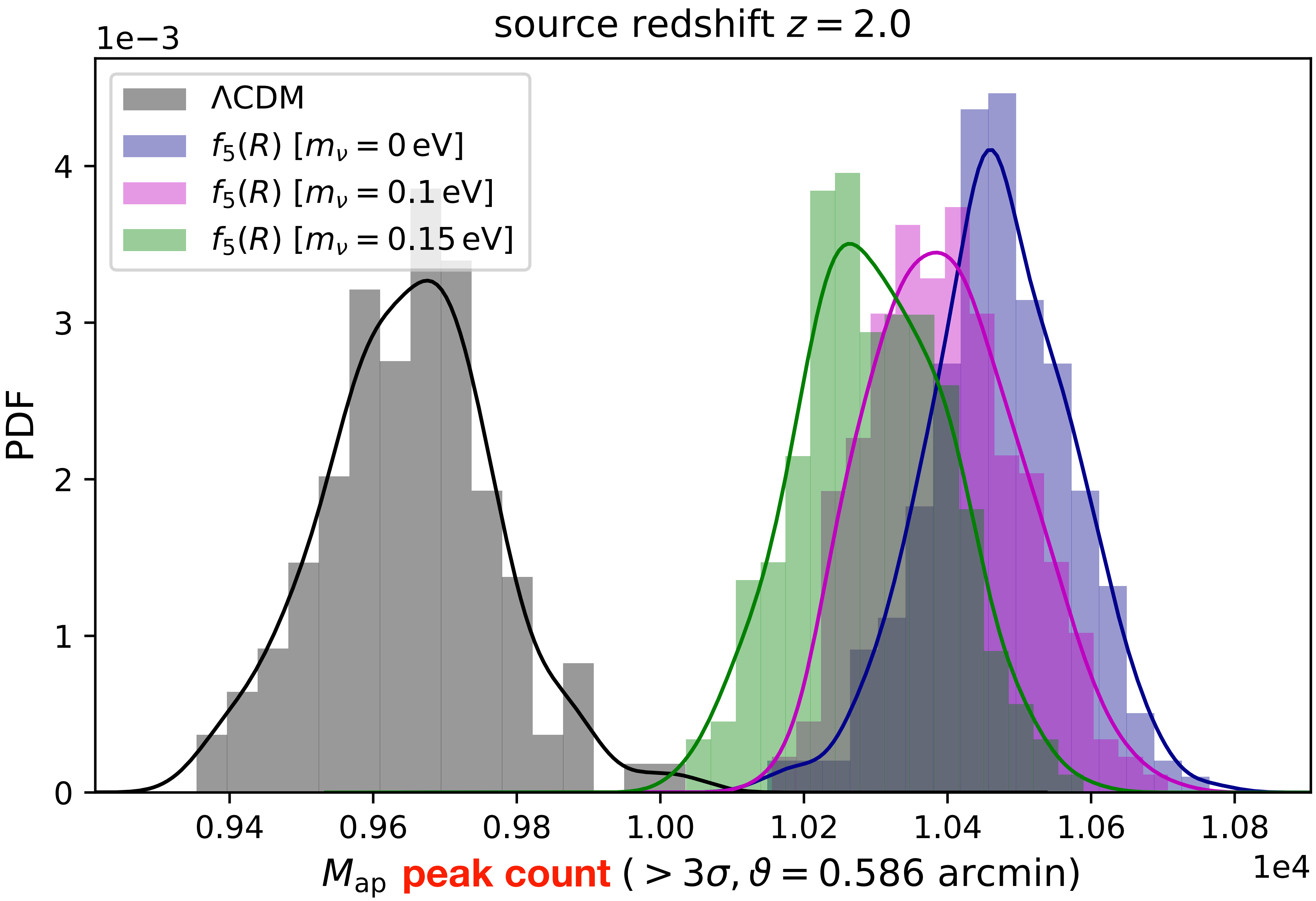
distributions of observables



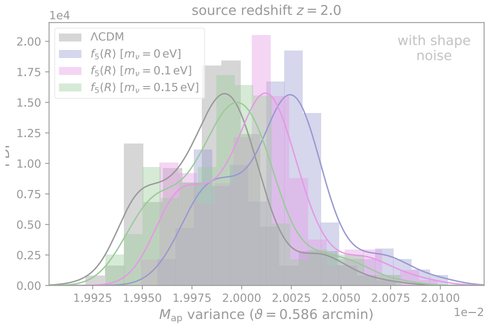
Breaking Degeneracies in MG with WL

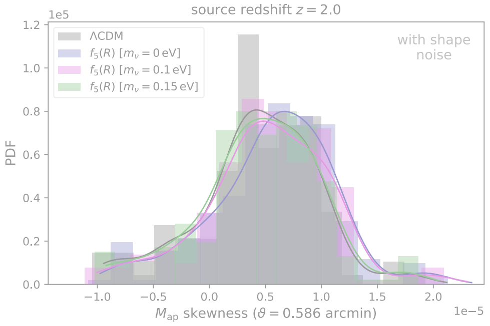


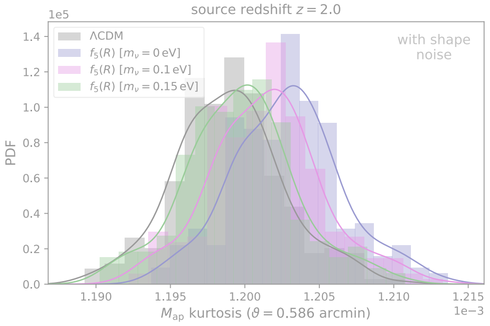
distributions of observables

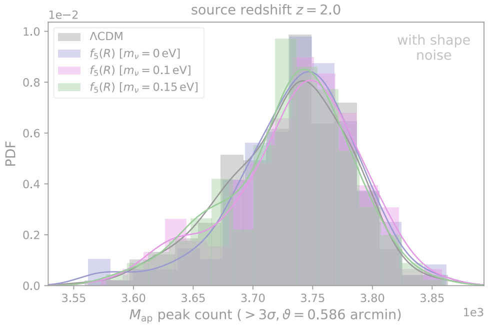


Breaking Degeneracies in MG with WL

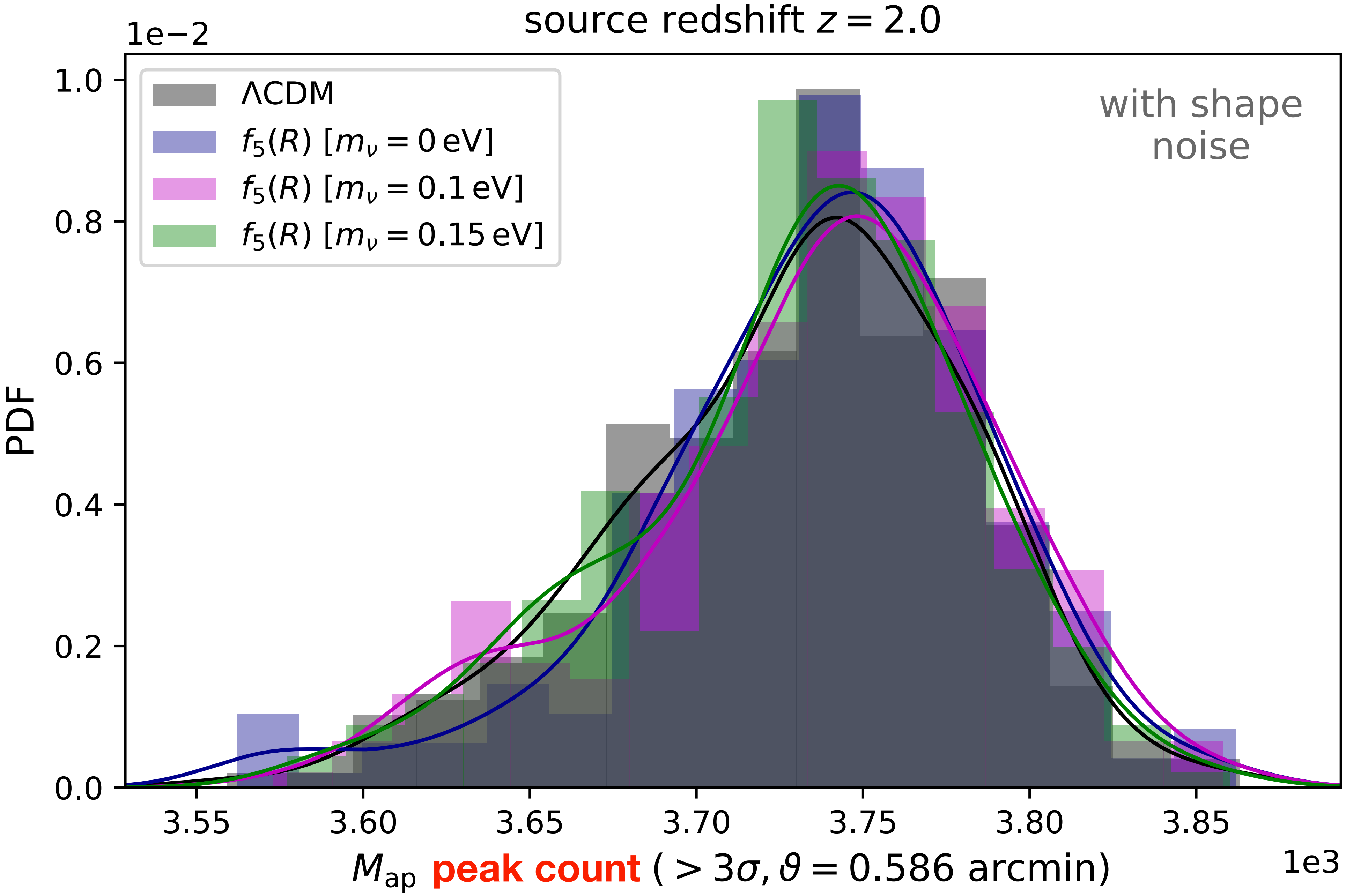








noisy distributions of observables

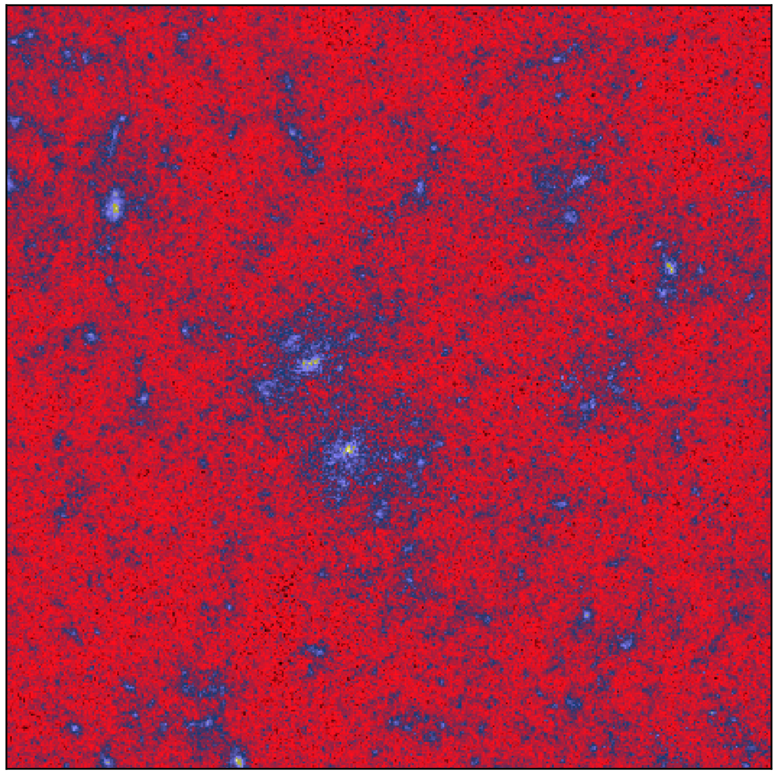


Outline

1. Modified gravity simulations
2. Aperture mass statistics
3. Distinguishing cosmological models
4. **A machine learning approach**
5. Summary

Can a neural network do better ?

Let’s recall the data



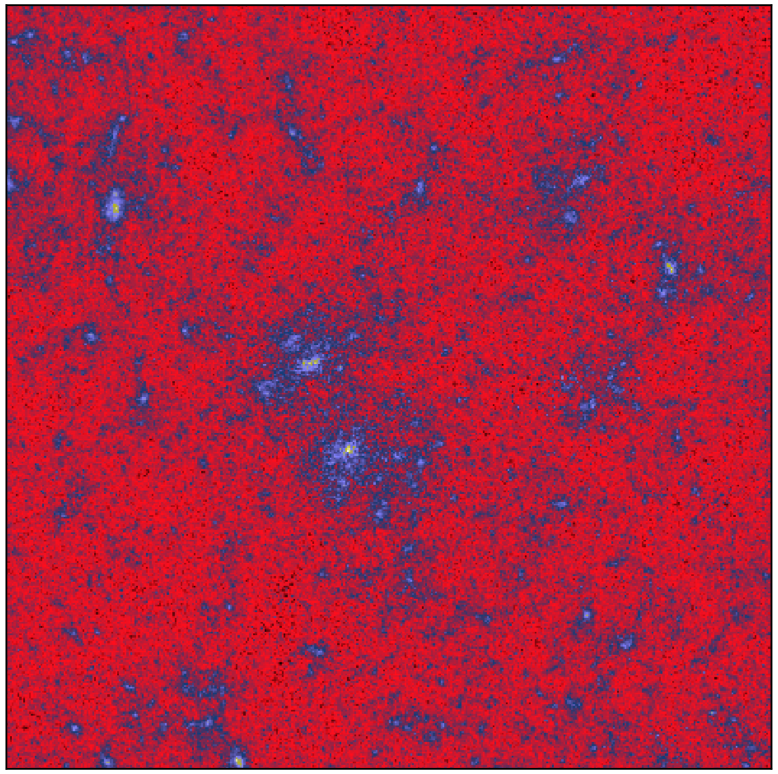
4,194,304 pixels
per map

x (256 maps per model)
x (4 source redshifts)
x (4 cosmological models)

=

Can a neural network do better ?

Let's recall the data

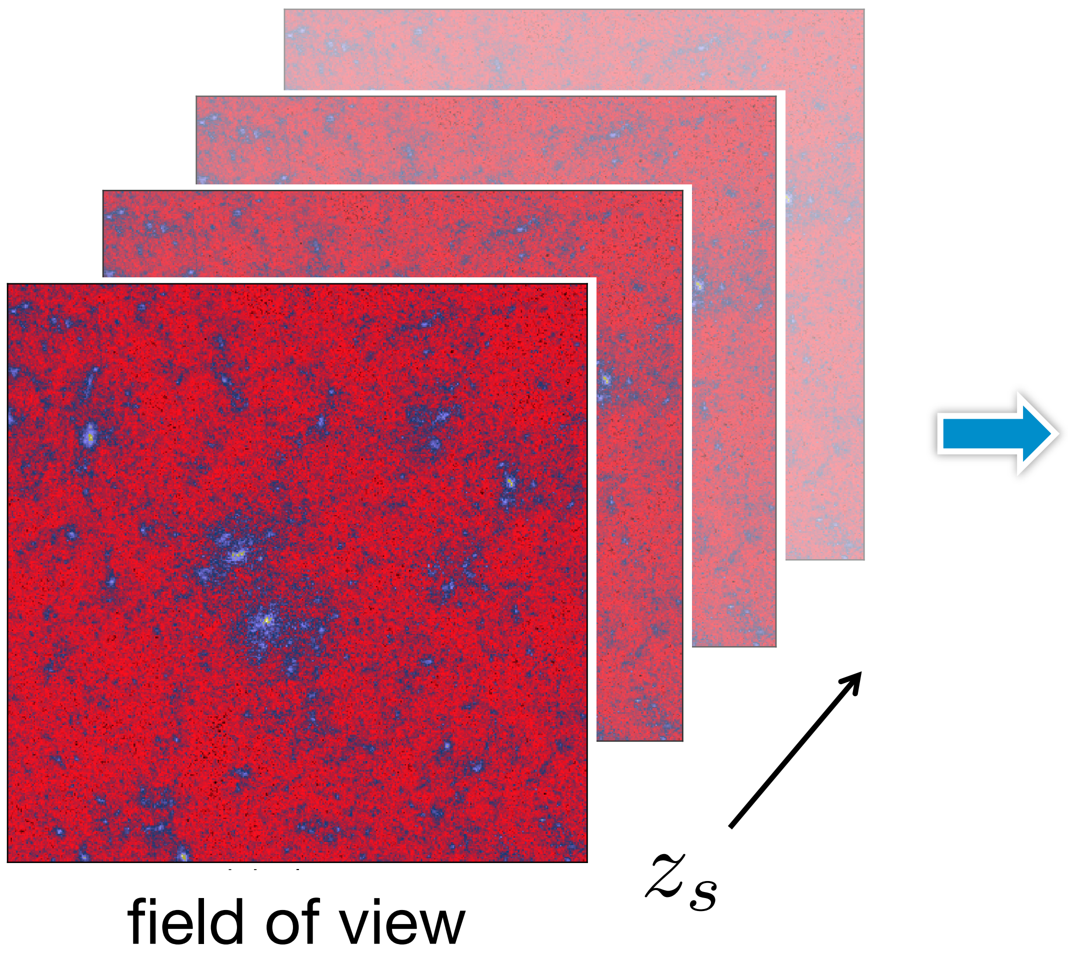


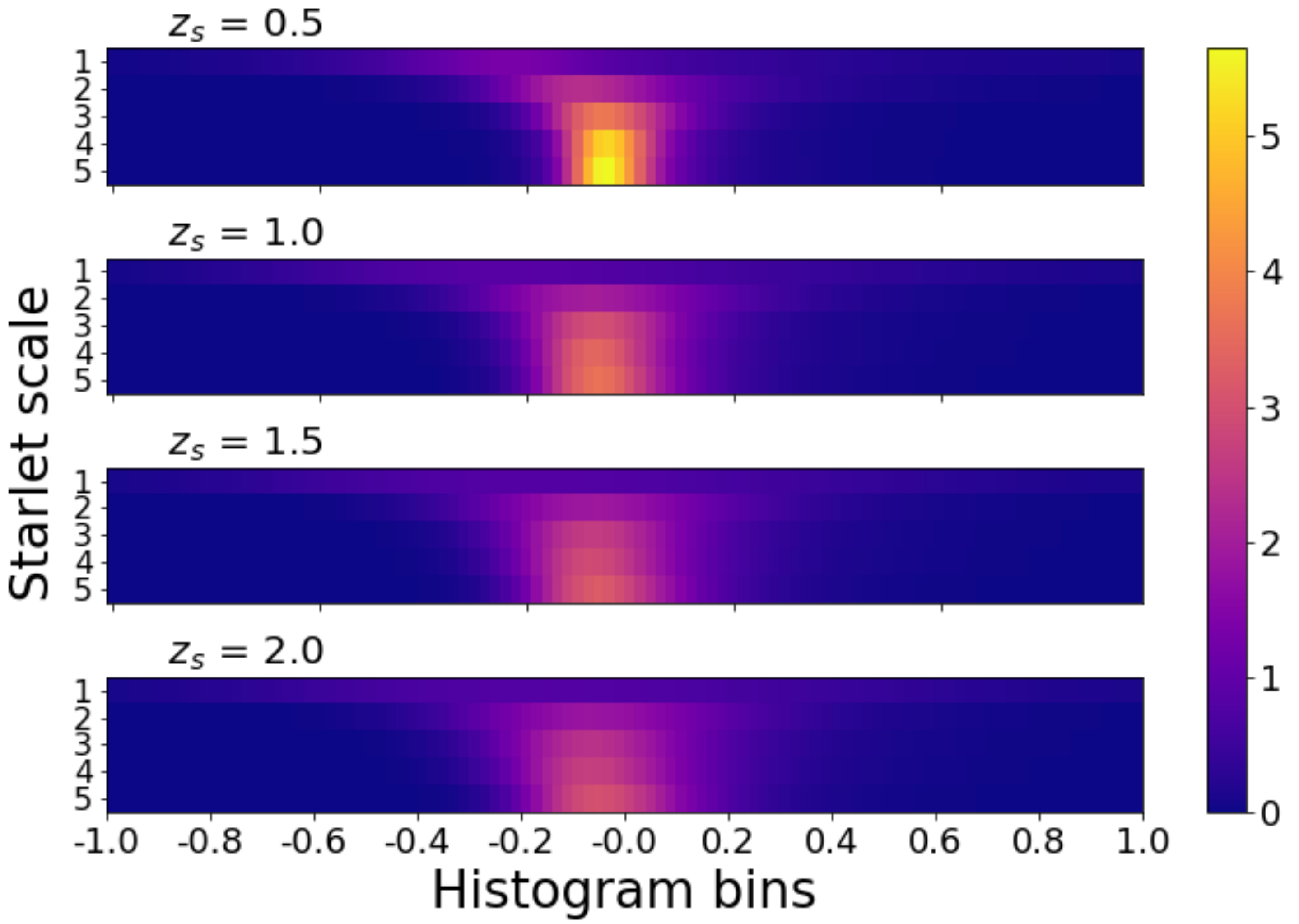
4,194,304 pixels
per map

x (256 maps per model)
x (4 source redshifts)
x (4 cosmological models)

= a computational challenge

A new data representation

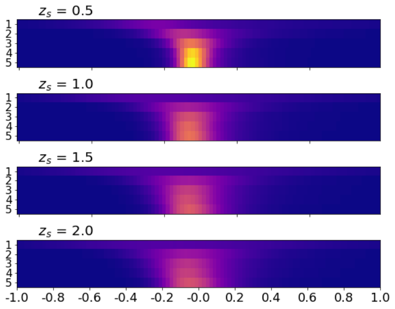




[A. Peel, F. Lalande, et al., submitted PRL (2018)]

Network architecture

Convolutional neural network (CNN) classification problem



input data

75% train
25% test

Layer type	Output shape	# params
Input layer	$1 \times 4 \times 5 \times 100$	0
Conv 3D $[2 \times 3 \times 10]$	$8 \times 4 \times 5 \times 100$	448
Conv 3D $[2 \times 3 \times 10]$	$8 \times 4 \times 5 \times 100$	3848
Max pooling $[1 \times 1 \times 5]$	$8 \times 4 \times 5 \times 20$	0
Conv 3D $[2 \times 3 \times 10]$	$8 \times 4 \times 5 \times 20$	3848
Max pooling $[1 \times 1 \times 2]$	$8 \times 4 \times 5 \times 10$	0
Dropout [0.3]	$8 \times 4 \times 5 \times 10$	0
Flatten	1600	0
Fully connected	32	51232
Fully connected	16	528
Fully connected	4	68

$\longrightarrow$ ΛCDM

? $f_5(R)$

$\longrightarrow$ $M_\nu = 0$ eV

$\longrightarrow$ $M_\nu = 0.1$ eV

$\longrightarrow$ $M_\nu = 0.15$ eV

[A. Peel, F. Lalande, et al., submitted PRL (2018)]

Breaking Degeneracies in MG with WL

noise level ▨ ☐ ☐ zero

Convolutional neural network

Prediction

Truth

$\sigma_{\text{noise}} = 0$	ΛCDM	$f_5(R)$ $M_\nu = 0$ eV	$f_5(R)$ $M_\nu = 0.1$ eV	$f_5(R)$ $M_\nu = 0.15$ eV
ΛCDM	1.00	0.00	0.00	0.00
$f_5(R)$ $M_\nu = 0$ eV	0.00	0.88	0.12	0.00
$f_5(R)$ $M_\nu = 0.1$ eV	0.00	0.13	0.83	0.04
$f_5(R)$ $M_\nu = 0.15$ eV	0.00	0.00	0.04	0.96

Peak statistics (best case)

Prediction

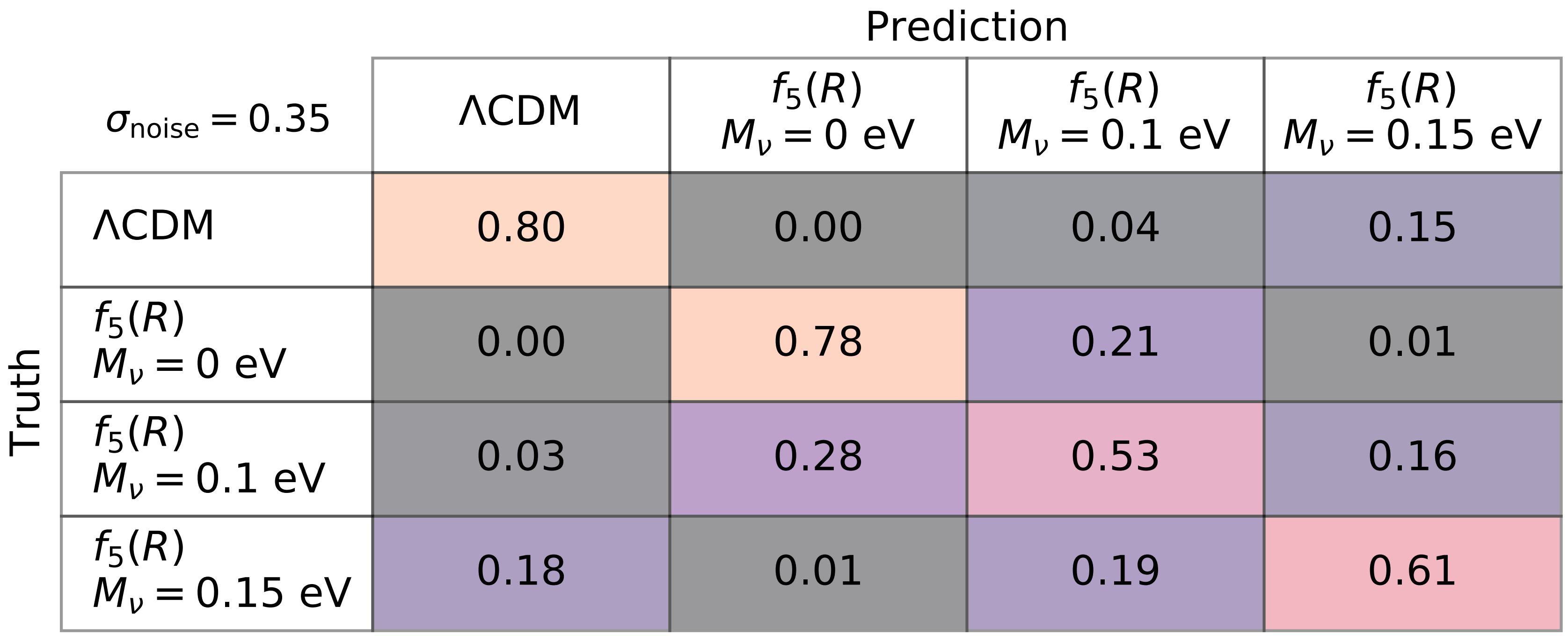
Truth

$\sigma_{\text{noise}} = 0$	ΛCDM	$f_5(R)$ $M_\nu = 0$ eV	$f_5(R)$ $M_\nu = 0.1$ eV	$f_5(R)$ $M_\nu = 0.15$ eV
ΛCDM	1.00	0.00	0.00	0.00
$f_5(R)$ $M_\nu = 0$ eV	0.00	0.49	0.42	0.09
$f_5(R)$ $M_\nu = 0.1$ eV	0.00	0.33	0.45	0.22
$f_5(R)$ $M_\nu = 0.15$ eV	0.00	0.09	0.25	0.66

Breaking Degeneracies in MG with WL

noise level ☐▨☐ optimistic

Convolutional neural network

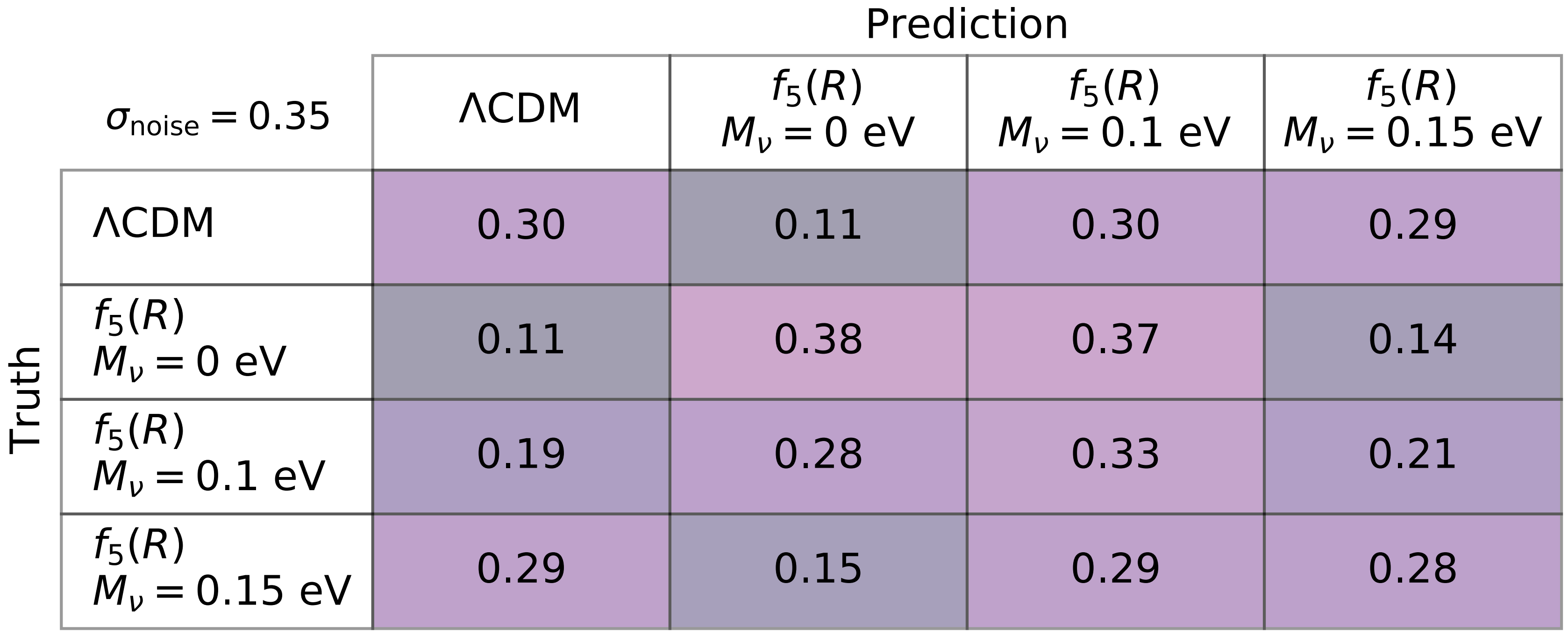


Prediction

Truth

$\sigma_{\text{noise}} = 0.35$	ΛCDM	$f_5(R)$ $M_\nu = 0$ eV	$f_5(R)$ $M_\nu = 0.1$ eV	$f_5(R)$ $M_\nu = 0.15$ eV
ΛCDM	0.80	0.00	0.04	0.15
$f_5(R)$ $M_\nu = 0$ eV	0.00	0.78	0.21	0.01
$f_5(R)$ $M_\nu = 0.1$ eV	0.03	0.28	0.53	0.16
$f_5(R)$ $M_\nu = 0.15$ eV	0.18	0.01	0.19	0.61

Peak statistics (best case)



Prediction

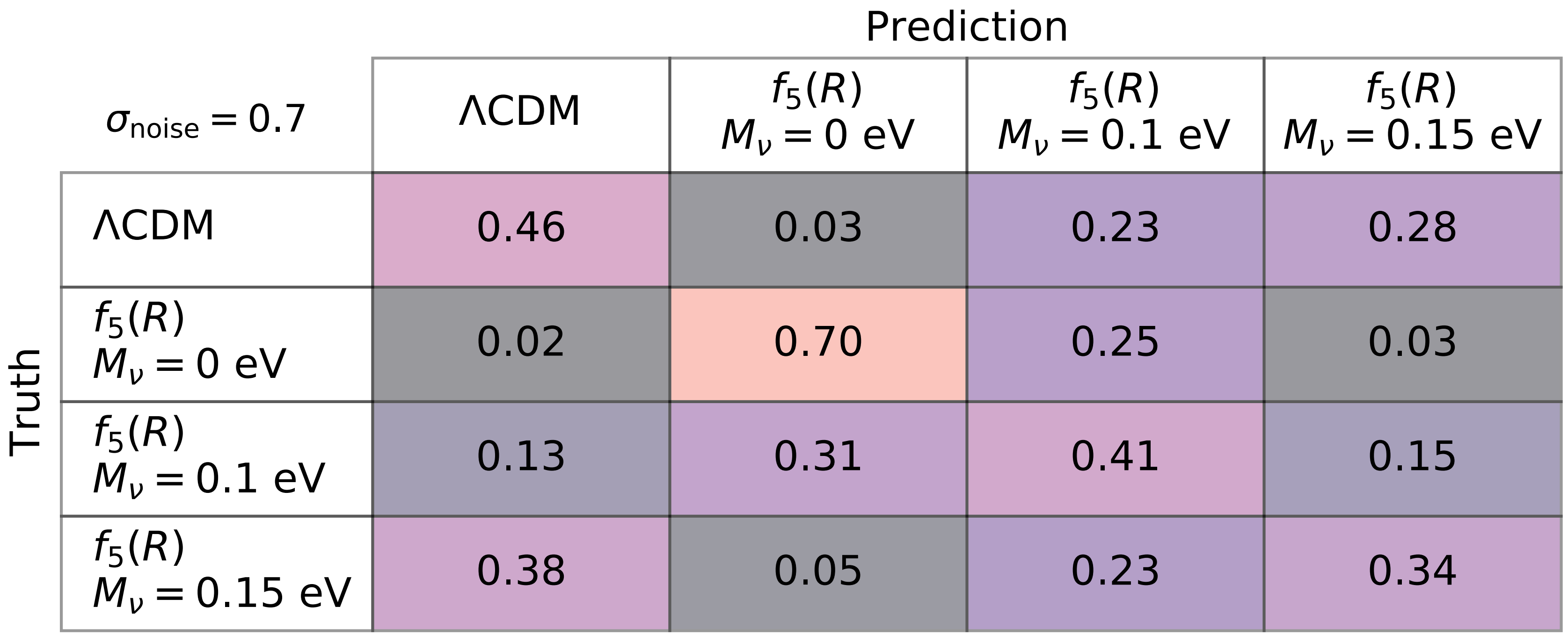
Truth

$\sigma_{\text{noise}} = 0.35$	ΛCDM	$f_5(R)$ $M_\nu = 0$ eV	$f_5(R)$ $M_\nu = 0.1$ eV	$f_5(R)$ $M_\nu = 0.15$ eV
ΛCDM	0.30	0.11	0.30	0.29
$f_5(R)$ $M_\nu = 0$ eV	0.11	0.38	0.37	0.14
$f_5(R)$ $M_\nu = 0.1$ eV	0.19	0.28	0.33	0.21
$f_5(R)$ $M_\nu = 0.15$ eV	0.29	0.15	0.29	0.28

Breaking Degeneracies in MG with WL

noise level ☐☐▨ pessimistic

Convolutional neural network

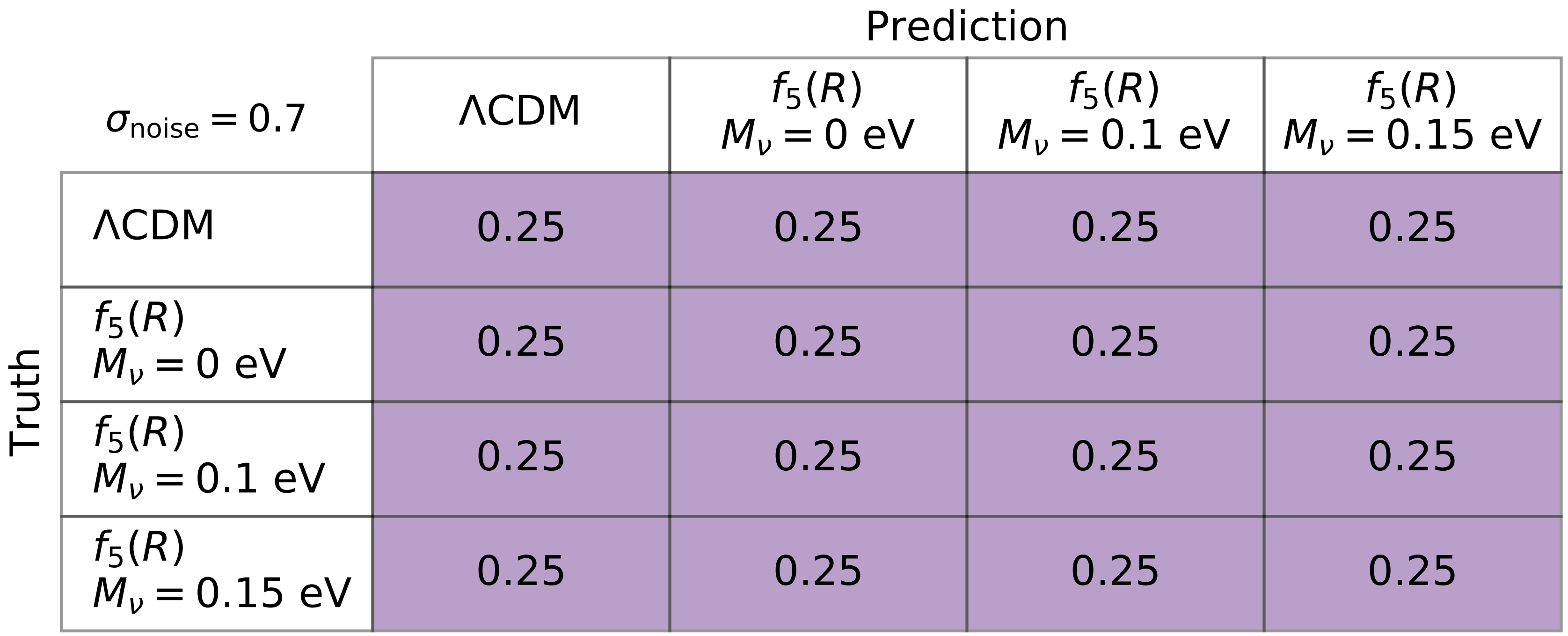


Prediction

Truth

$\sigma_{\text{noise}} = 0.7$	ΛCDM	$f_5(R)$ $M_\nu = 0$ eV	$f_5(R)$ $M_\nu = 0.1$ eV	$f_5(R)$ $M_\nu = 0.15$ eV
ΛCDM	0.46	0.03	0.23	0.28
$f_5(R)$ $M_\nu = 0$ eV	0.02	0.70	0.25	0.03
$f_5(R)$ $M_\nu = 0.1$ eV	0.13	0.31	0.41	0.15
$f_5(R)$ $M_\nu = 0.15$ eV	0.38	0.05	0.23	0.34

Peak statistics (best case)



Prediction

Truth

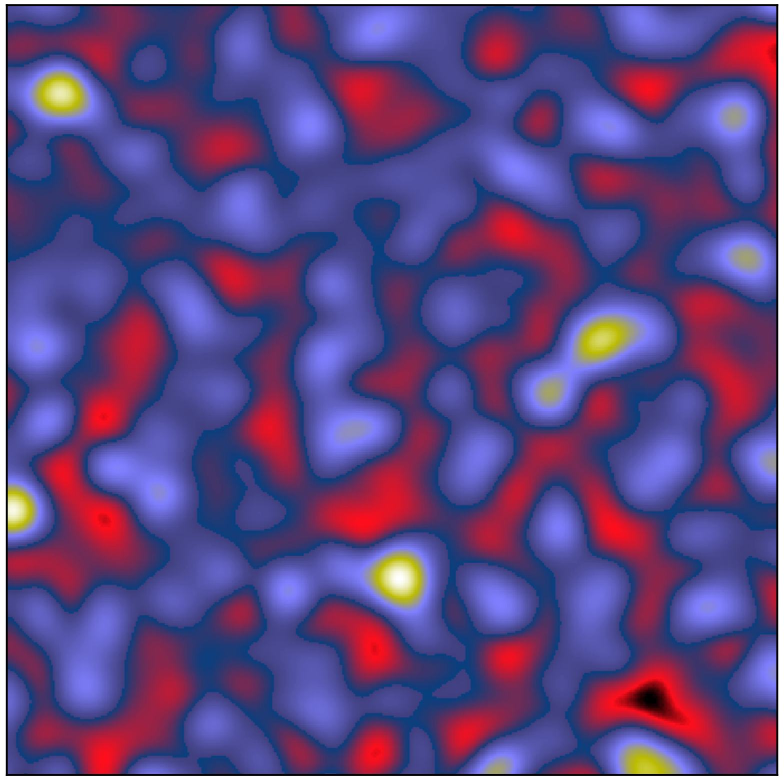
$\sigma_{\text{noise}} = 0.7$	ΛCDM	$f_5(R)$ $M_\nu = 0$ eV	$f_5(R)$ $M_\nu = 0.1$ eV	$f_5(R)$ $M_\nu = 0.15$ eV
ΛCDM	0.25	0.25	0.25	0.25
$f_5(R)$ $M_\nu = 0$ eV	0.25	0.25	0.25	0.25
$f_5(R)$ $M_\nu = 0.1$ eV	0.25	0.25	0.25	0.25
$f_5(R)$ $M_\nu = 0.15$ eV	0.25	0.25	0.25	0.25

Outline

1. Modified gravity simulations
2. Aperture mass statistics
3. Distinguishing cosmological models
4. A machine learning approach
5. **Summary**

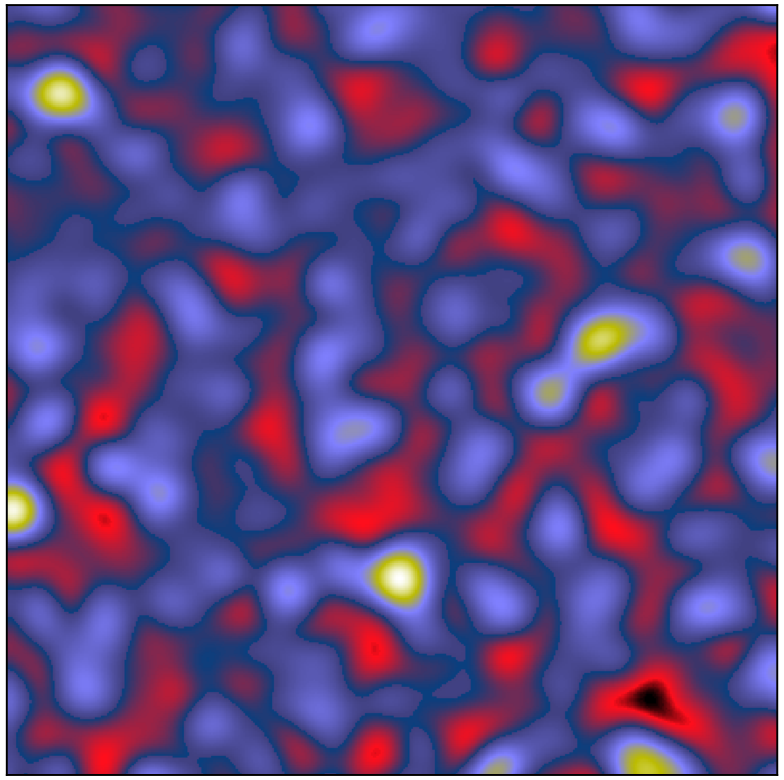
To sum up

- MG + neutrinos can **mimic ΛCDM** at the background and linear level
- Weak-lensing observations accessing **non-Gaussian information** can be used to break degeneracies
- In particular, **peak counts** generally outperform higher (than second) order moments of the aperture mass
- **Machine learning** can do even better, especially in the presence of noise



To sum up

- MG + neutrinos can **mimic ΛCDM** at the background and linear level
- Weak-lensing observations accessing **non-Gaussian information** can be used to break degeneracies
- In particular, **peak counts** generally outperform higher (than second) order moments of the aperture mass
- **Machine learning** can do even better, especially in the presence of noise



Thank you !