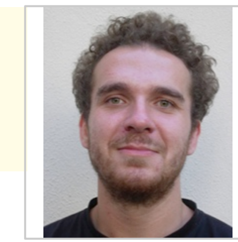
*modified gravity

Distinguishing standard and MG* cosmologies with machine learning

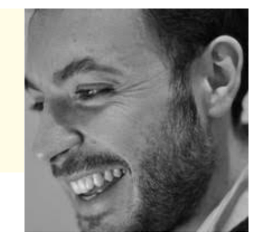
Austin Peel

Collaborators

Florian Lalande
ENSAI / CosmoStat / ESA



Julian Merten
INAF-OAS Bologna

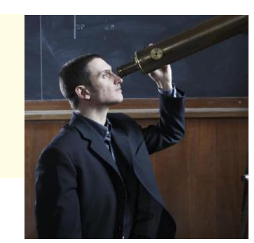


Carlo Giocoli
DIFA Università di Bologna

Jean-Luc Starck
CosmoStat / CEA Saclay

Marco Baldi
DIFA Università di Bologna

Valeria Pettorino
CosmoStat / CEA Saclay



Massimo Meneghetti
DIFA Università di Bologna

Peel et al. 2018 [arXiv:1810.11030]

Outline

1. Modified gravity simulations
2. Data representations
3. ML network architecture
4. Classifying cosmological models
5. Summary

Outline

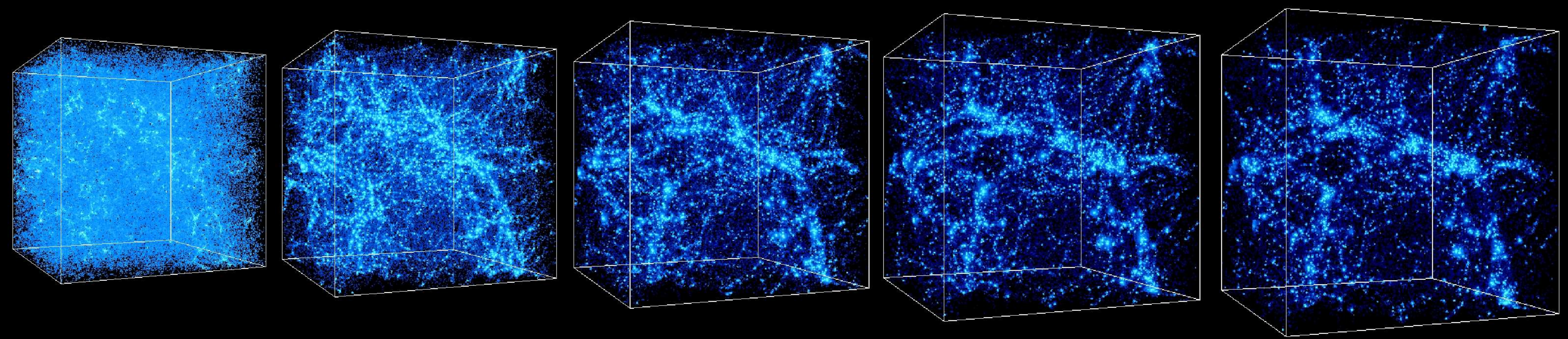
1. **Modified gravity simulations**
2. Data representations
3. ML network architecture
4. Classifying cosmological models
5. Summary

DUSTGRAIN-*pathfinder* simulations

C. Giocoli et al. 2018 [arXiv:1806.04681]

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

Performed with MG-Gadget code, which 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}

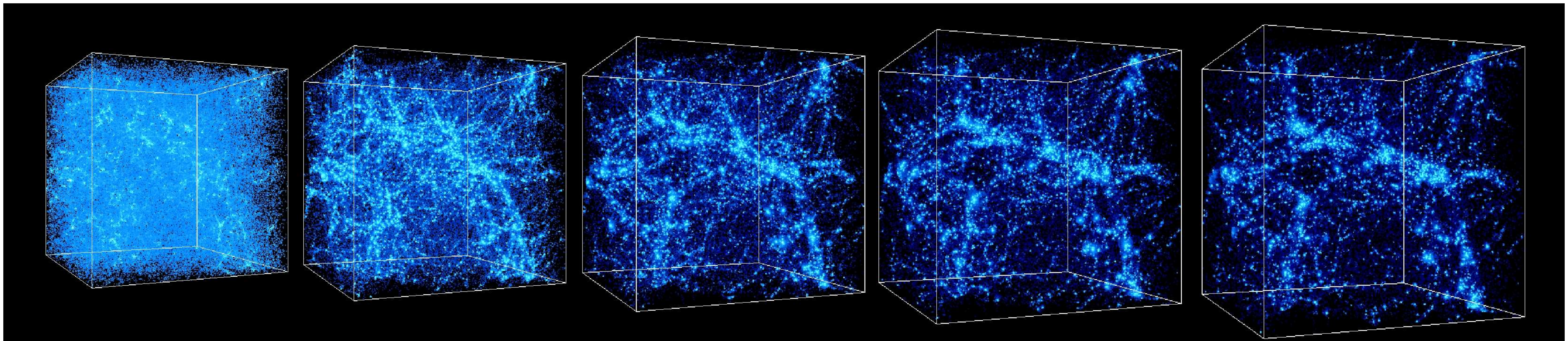
ML to distinguish standard and MG cosmologies

DUSTGRAIN-*pathfinder* simulations

C. Giocoli et al. 2018 [arXiv:1806.04681]

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

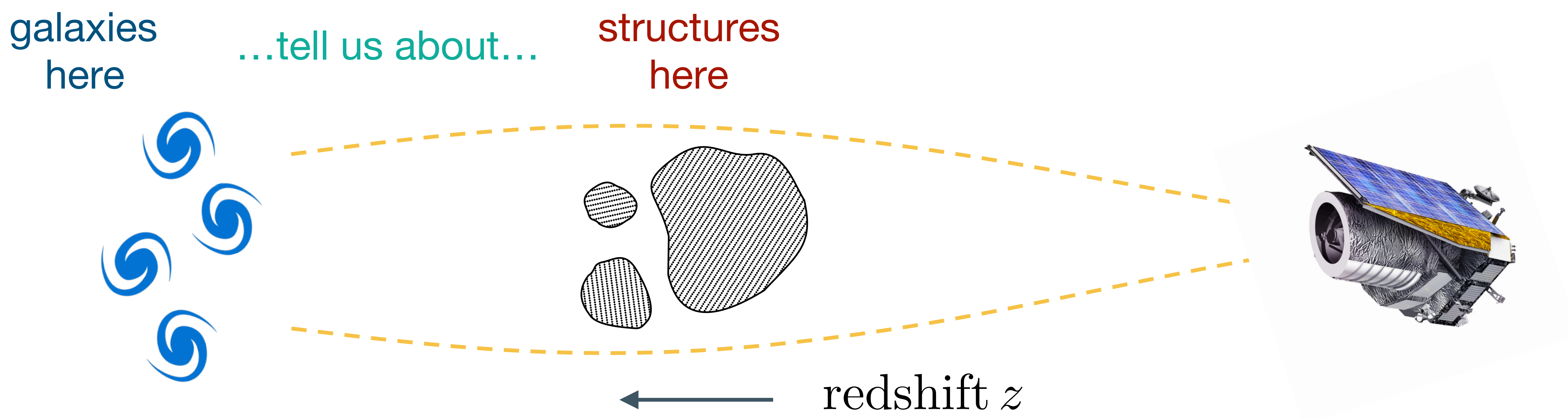
Performed with MG-Gadget code, which implements the extra fifth-force and screening

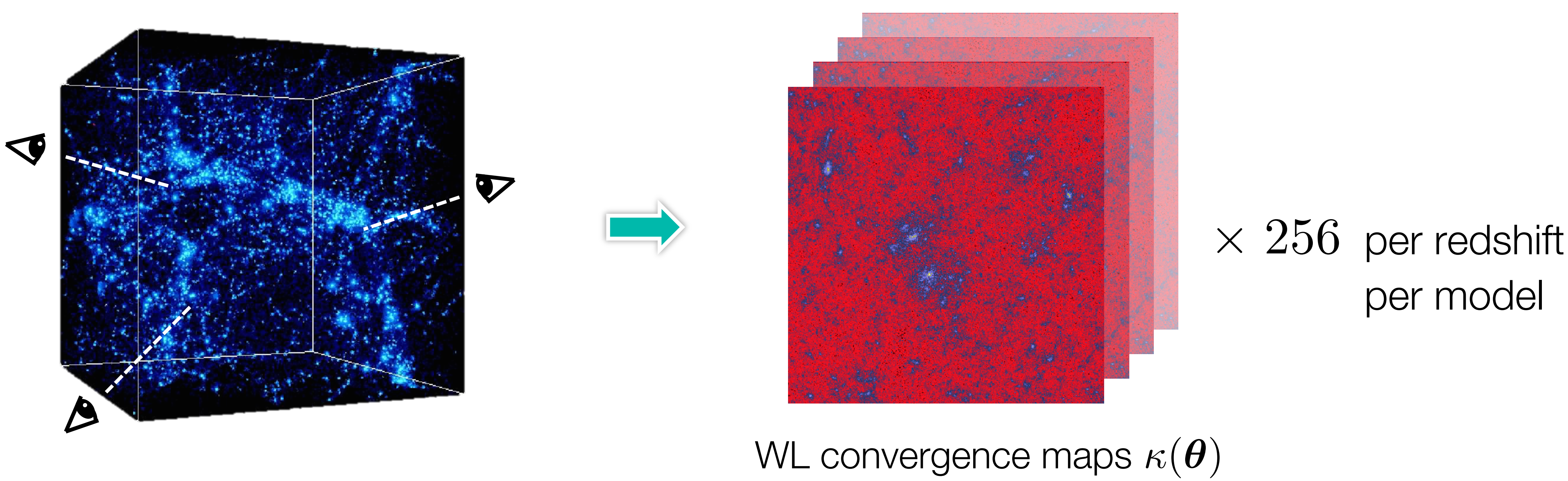


NOT DUSTGRAIN snapshots—for illustration only (http://cosmicweb.uchicago.edu/filaments.html)

	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]
1	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
2	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}
3	fR5-0.15eV	$f(R)$	-1×10^{-5}	0.15	0.30987	0.00358	8.01×10^{10}	9.25×10^{8}
4	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}

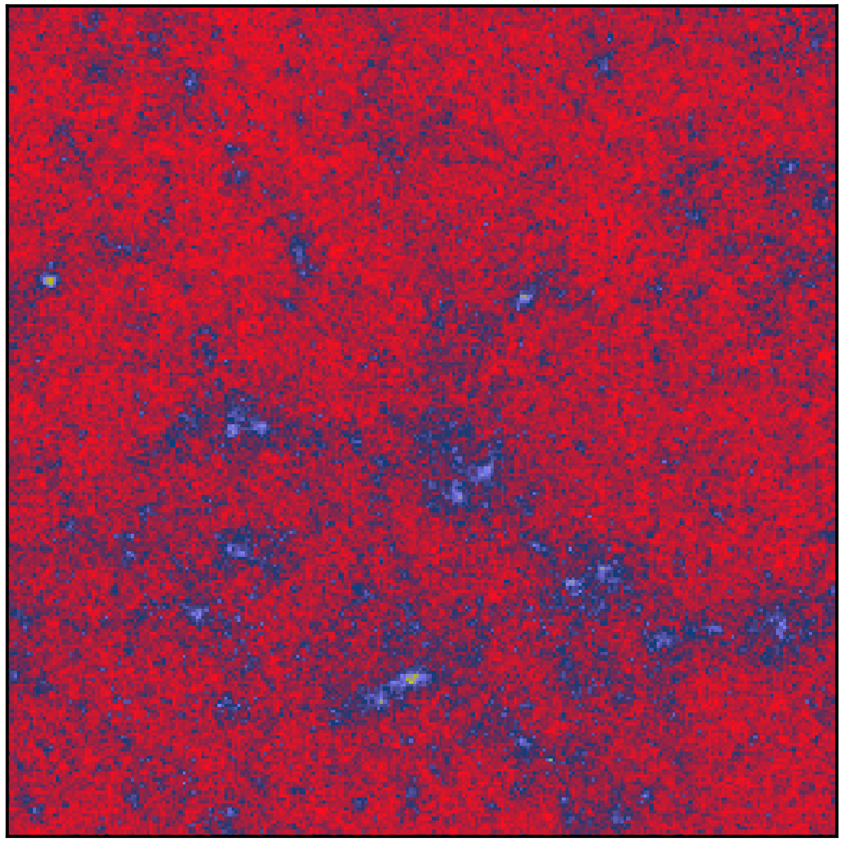
Weak lensing maps from ray tracing

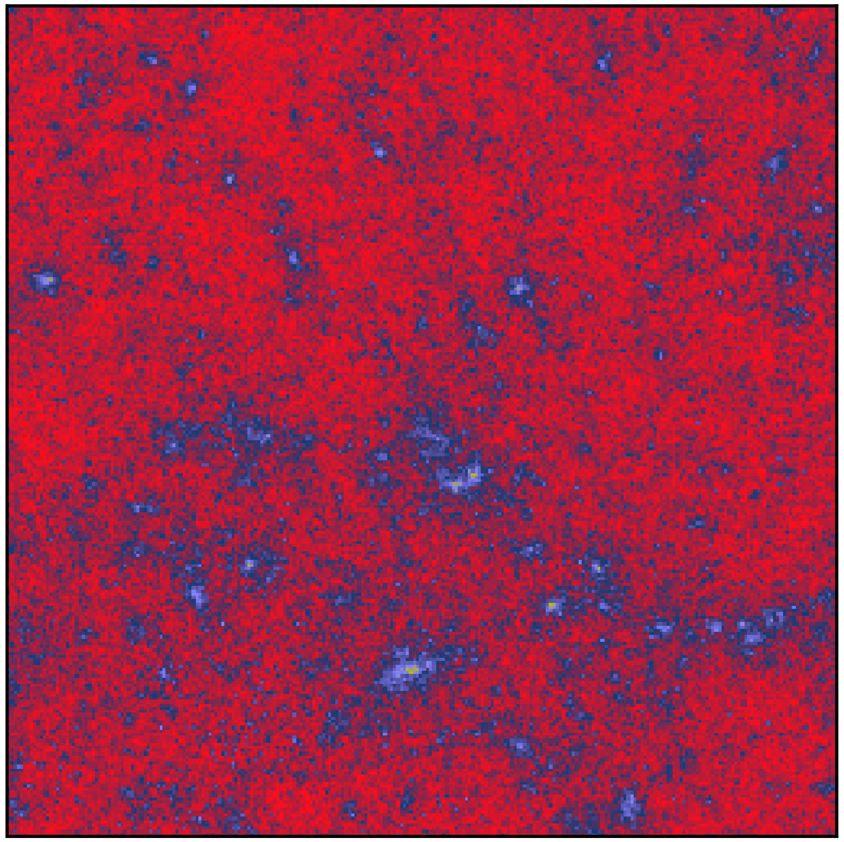




ML to distinguish standard and MG cosmologies

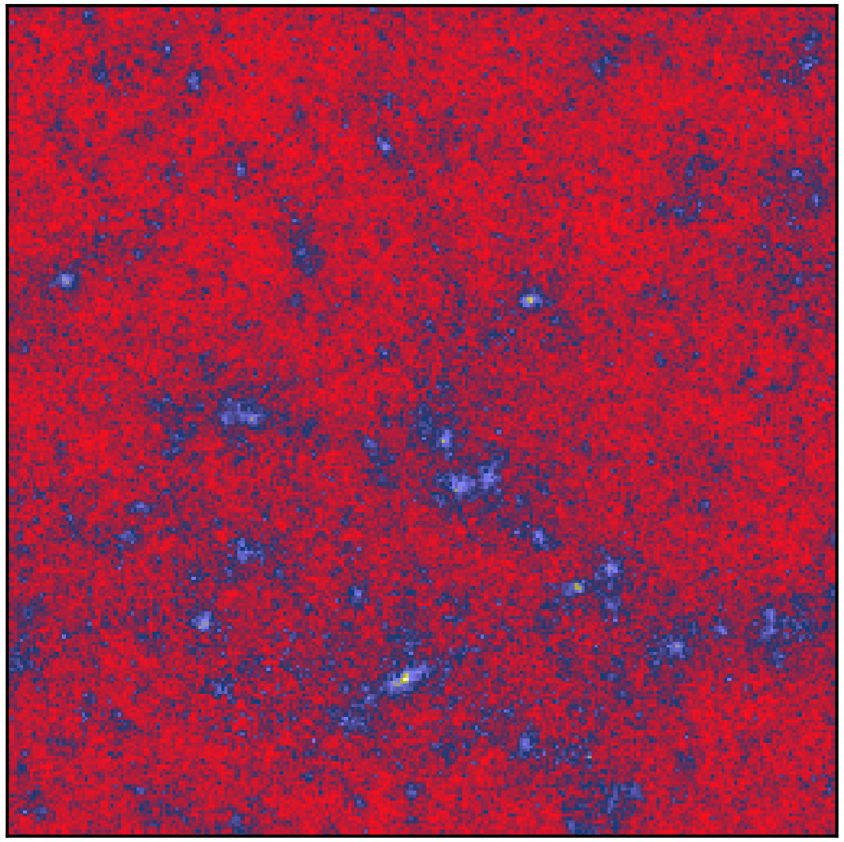
$\Lambda\mathrm{CDM}$ $[M_\nu = 0 \text{ eV}]$

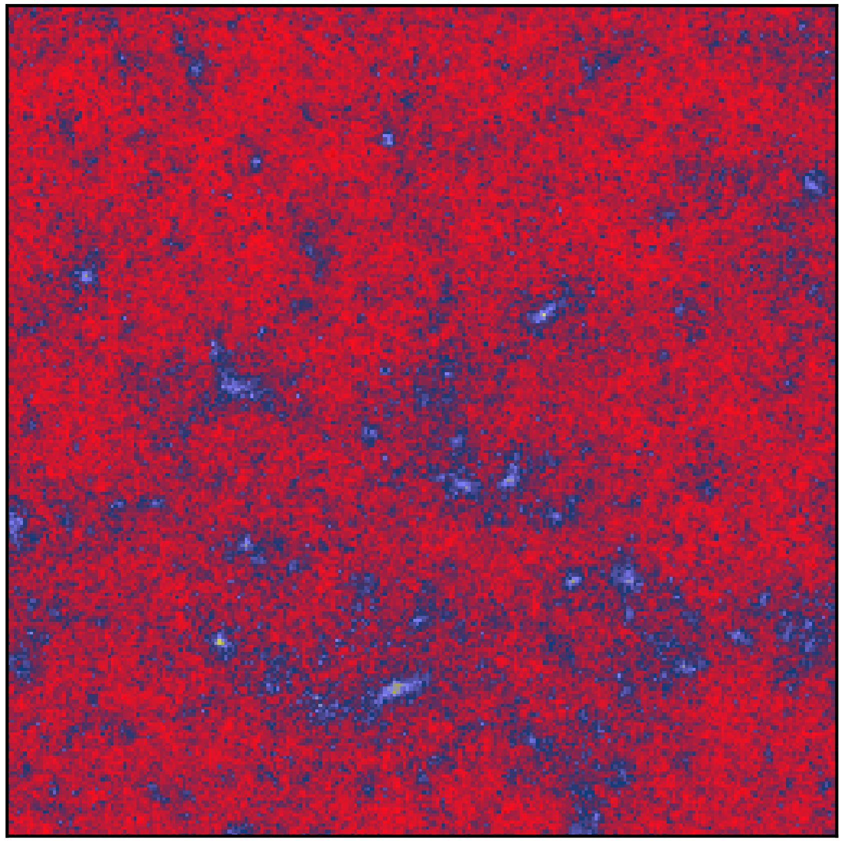




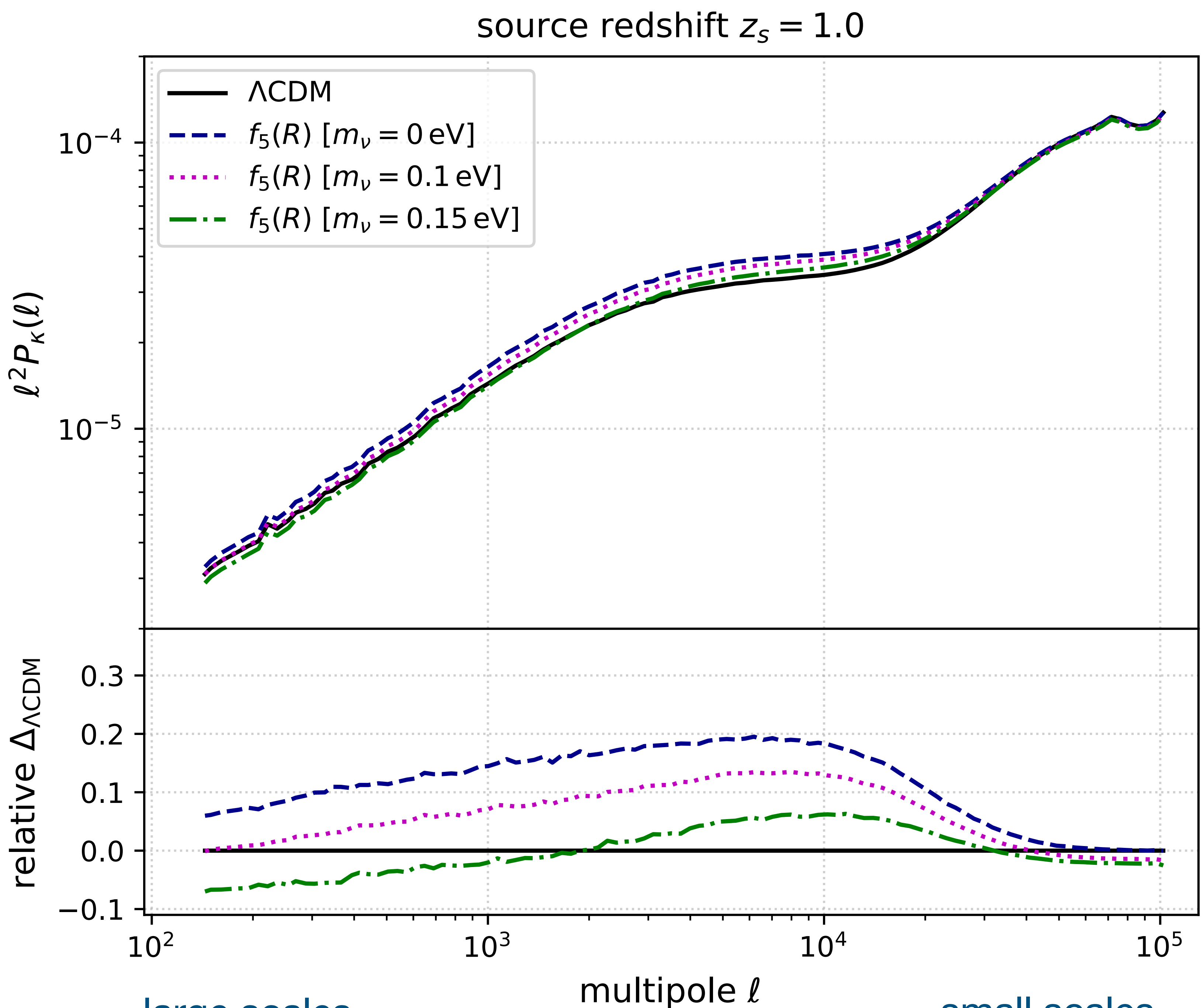
$f_5(R)$ $[M_\nu = 0 \text{ eV}]$

$f_5(R)$ $[M_\nu = 0.1 \text{ eV}]$





$f_5(R)$ $[M_\nu = 0.15 \text{ eV}]$



large scales (linear)

small scales (nonlinear)

Convergence power spectra

M_ν [eV]	
0	farther from ΛCDM
0.1	intermediate
0.15	closer to ΛCDM

neutrinos suppress the growth of structure

Outline

1. Modified gravity simulations
2. **Data representations**
3. ML network architecture
4. Classifying cosmological models
5. Summary

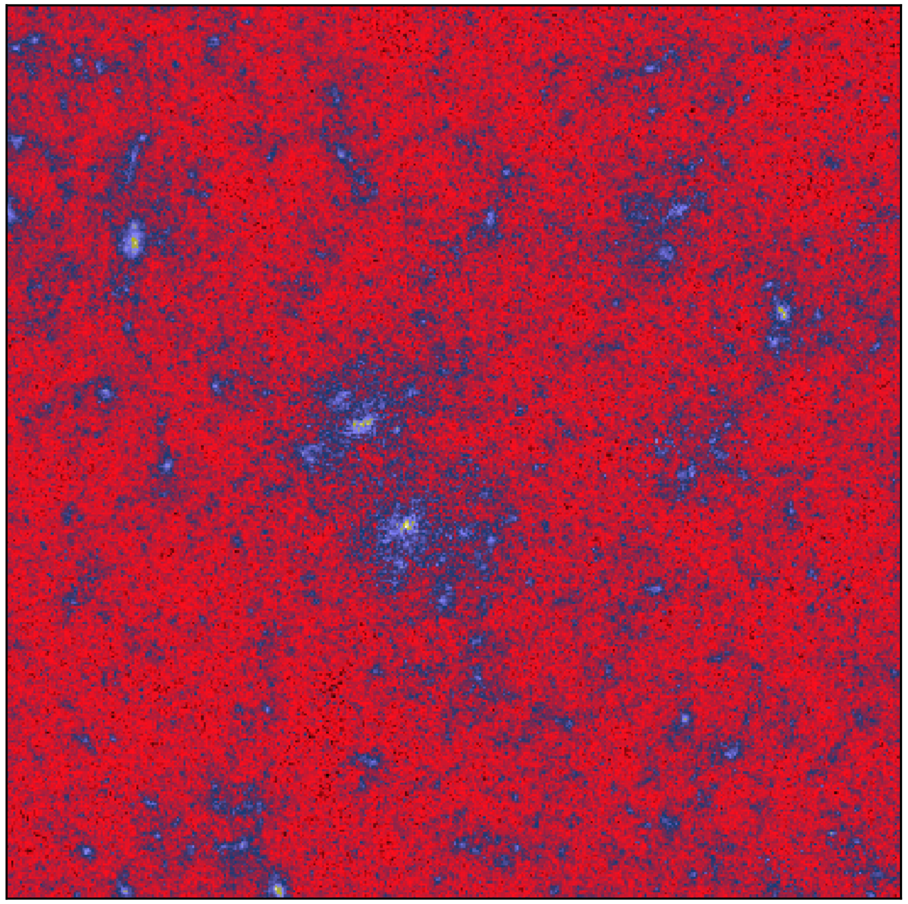
Aperture mass

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

isotropic filter function

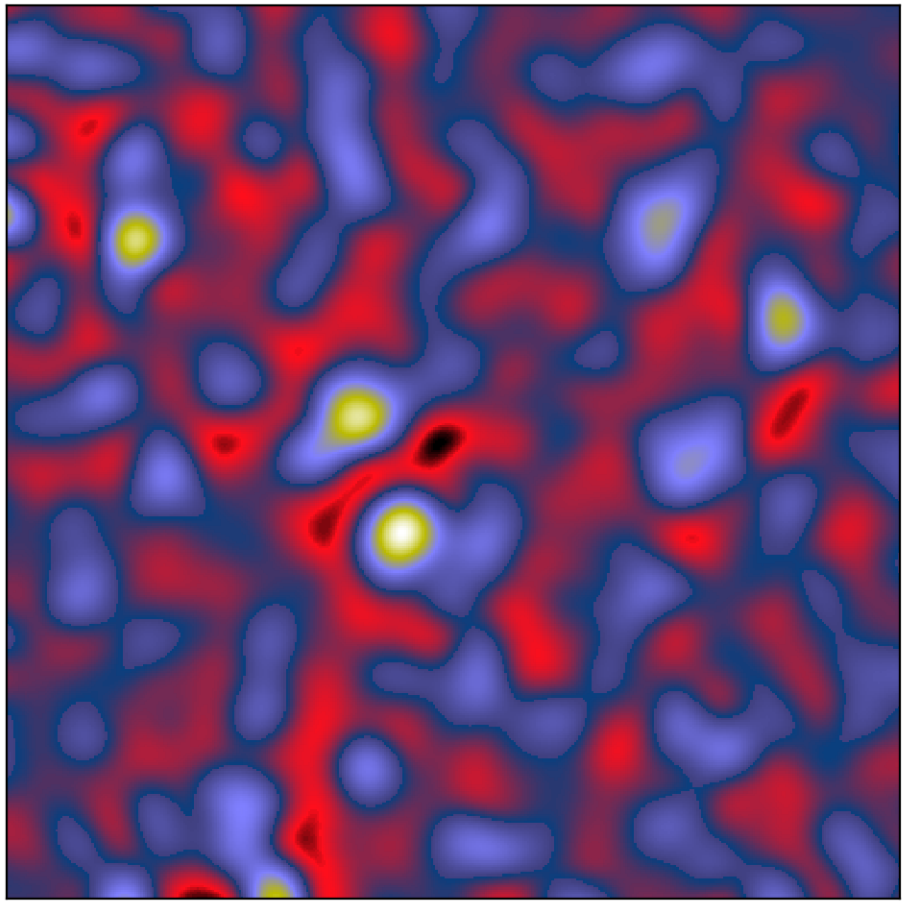
convergence map

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



original κ map

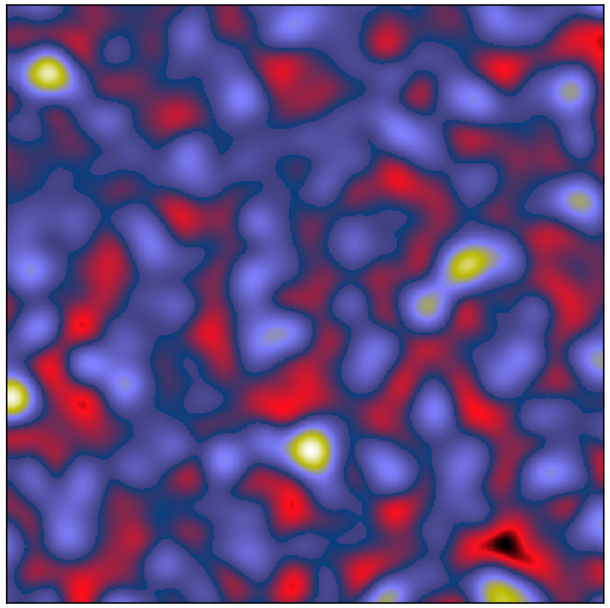
$$= \sum$$



aperture size $\vartheta_5 = 4.69$ arcmin

Aperture mass map at scale j

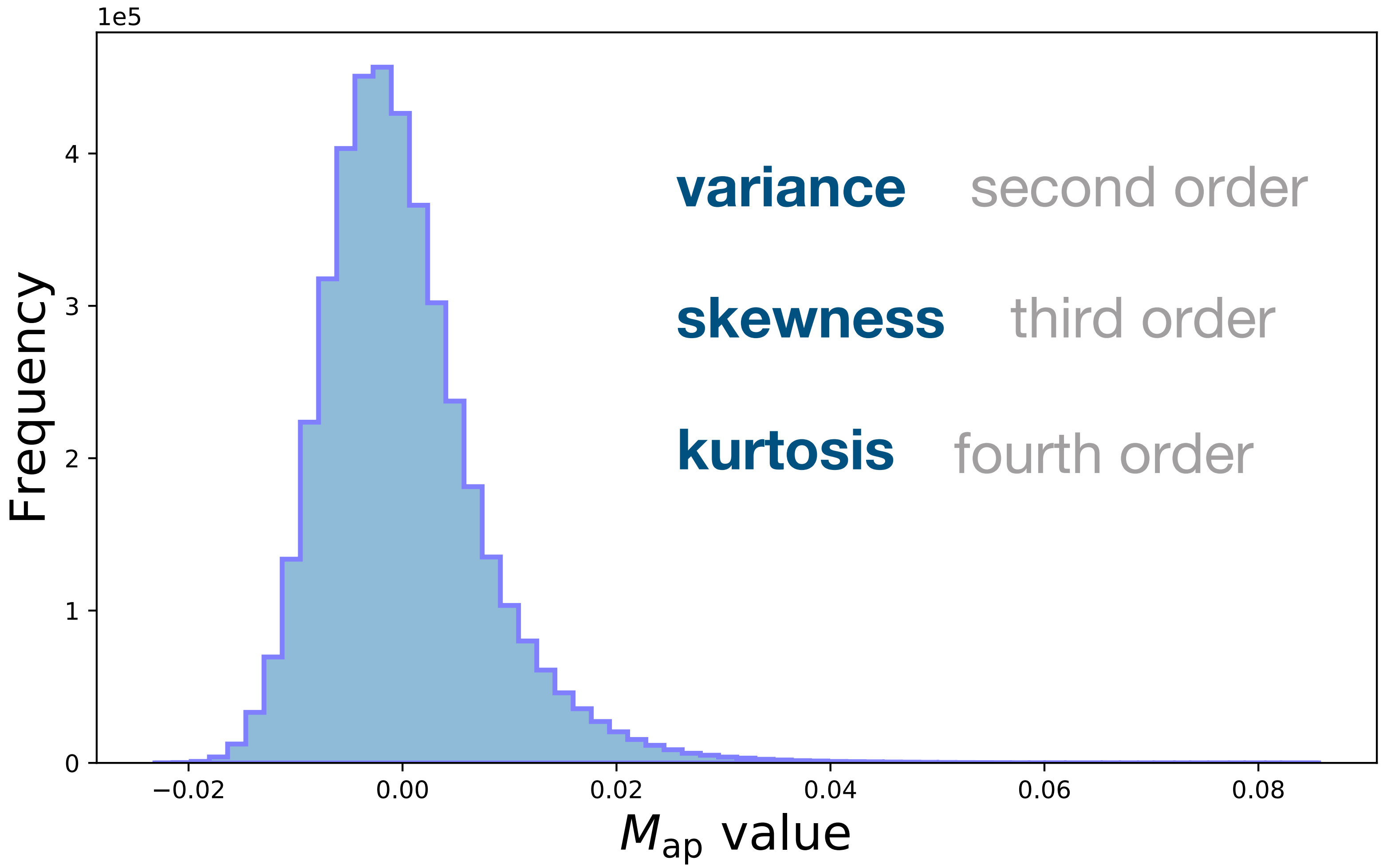
$$M_{\rm ap}({\rm model}, \vartheta_j, z_s) =$$

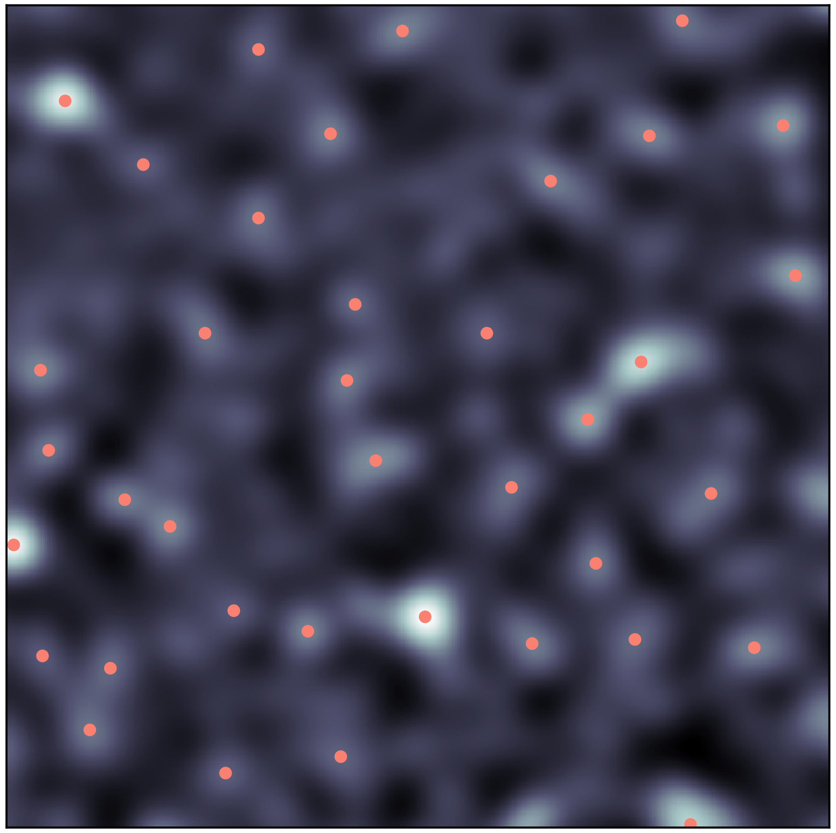


$5 \times 5 \ {\rm deg}^2$

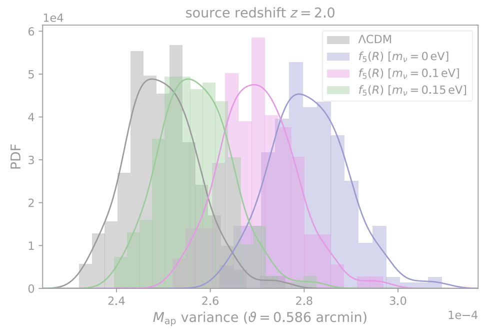
400×400 shown, but

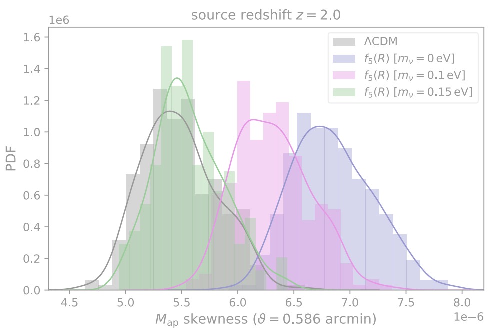
2048×2048 in practice

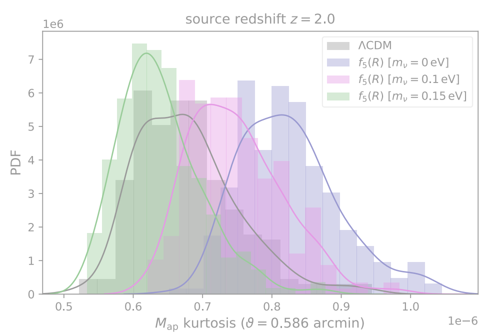


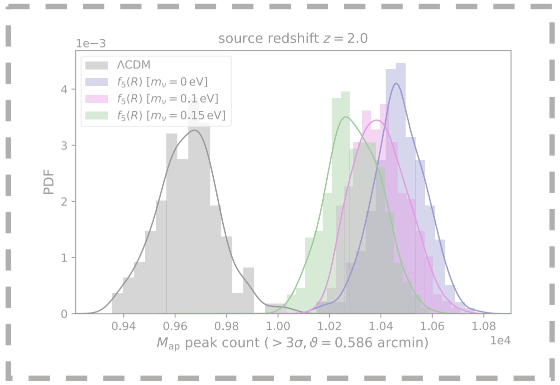


peak count

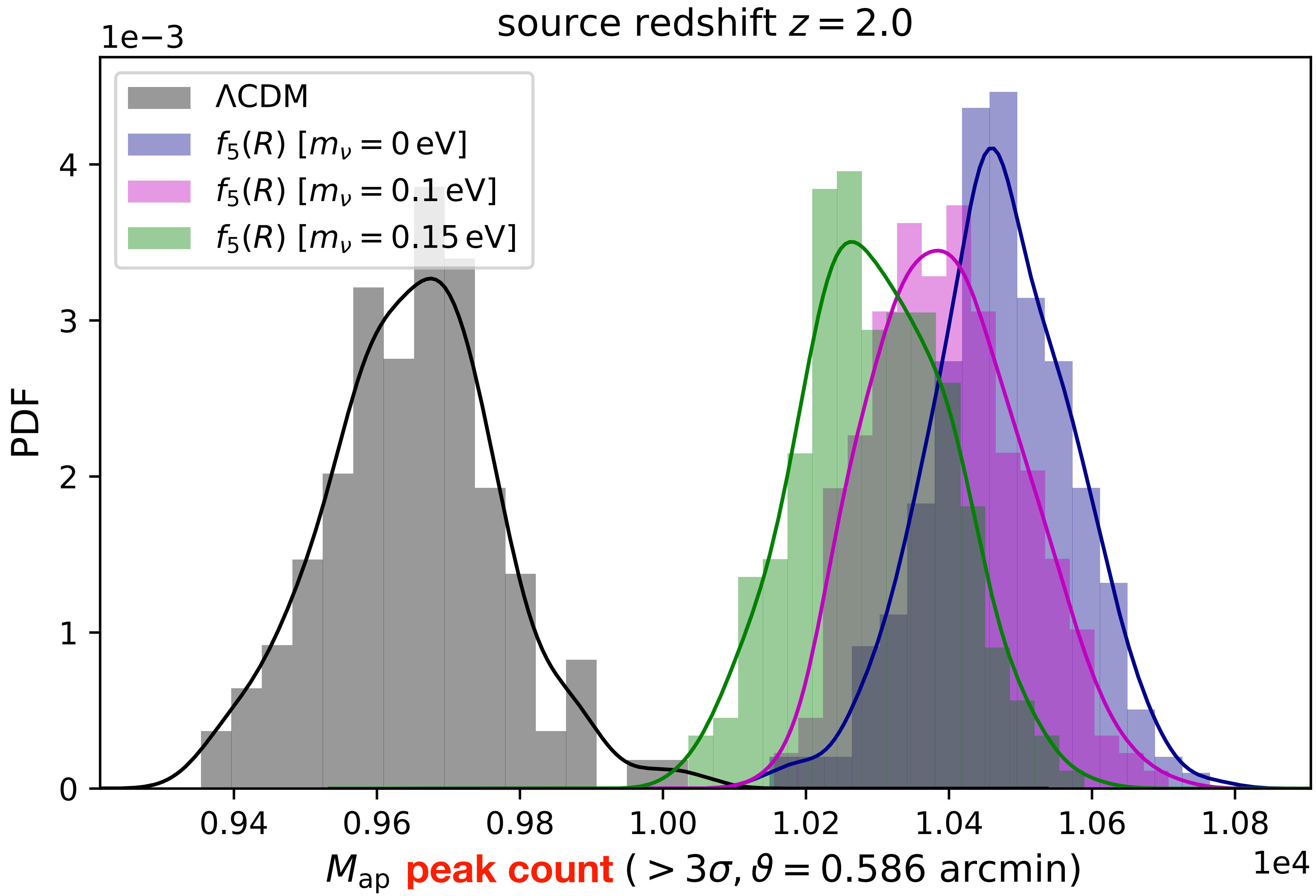






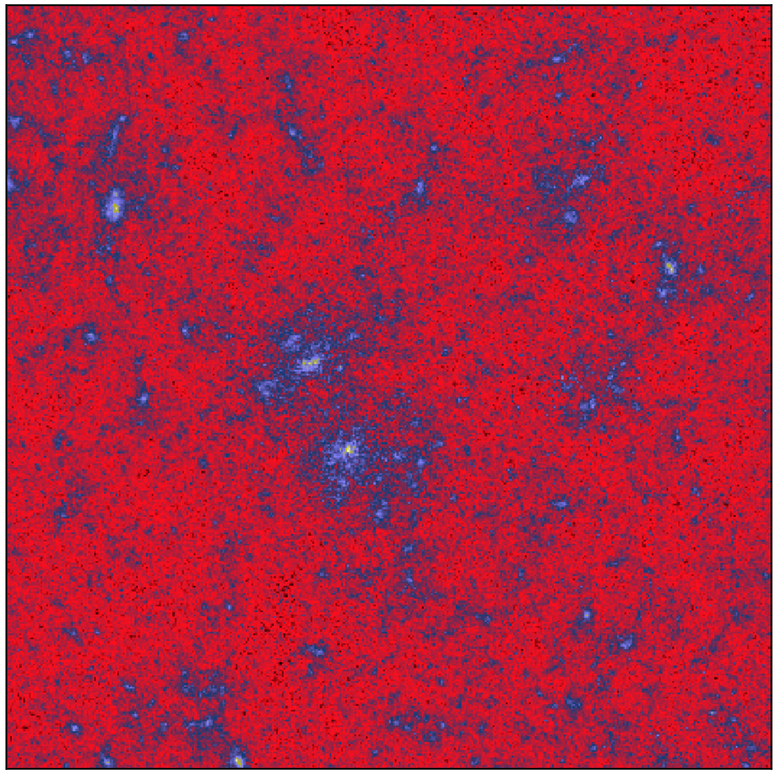


Distributions of observables



Can a neural network do better ?

Let's recall the data



4,194,304 pixels
per map

x (256 maps per model)

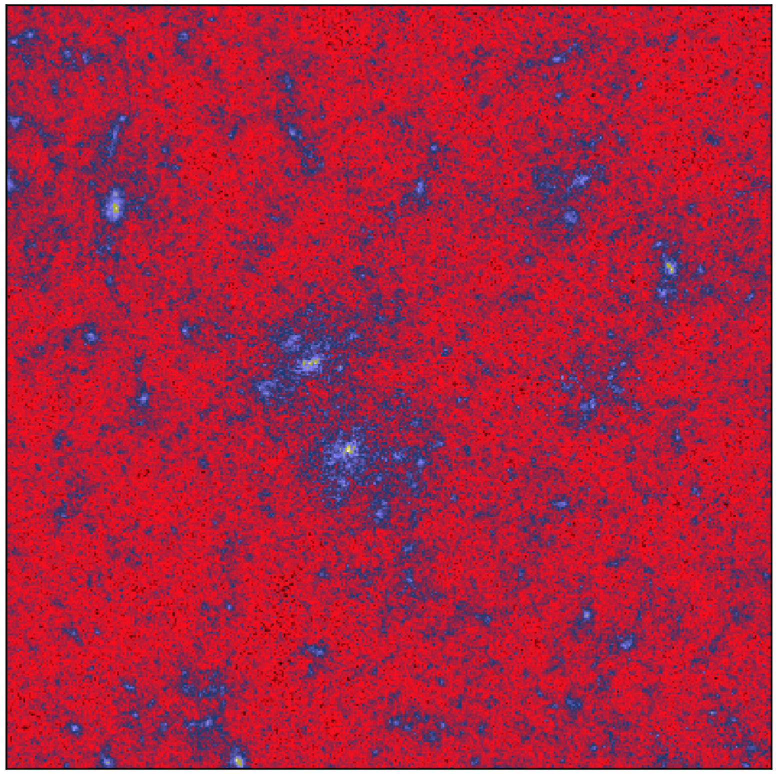
x (4 source redshifts)

x (4 cosmological models)

=

Can a neural network do better ?

Let's recall the data



4,194,304 pixels
per map

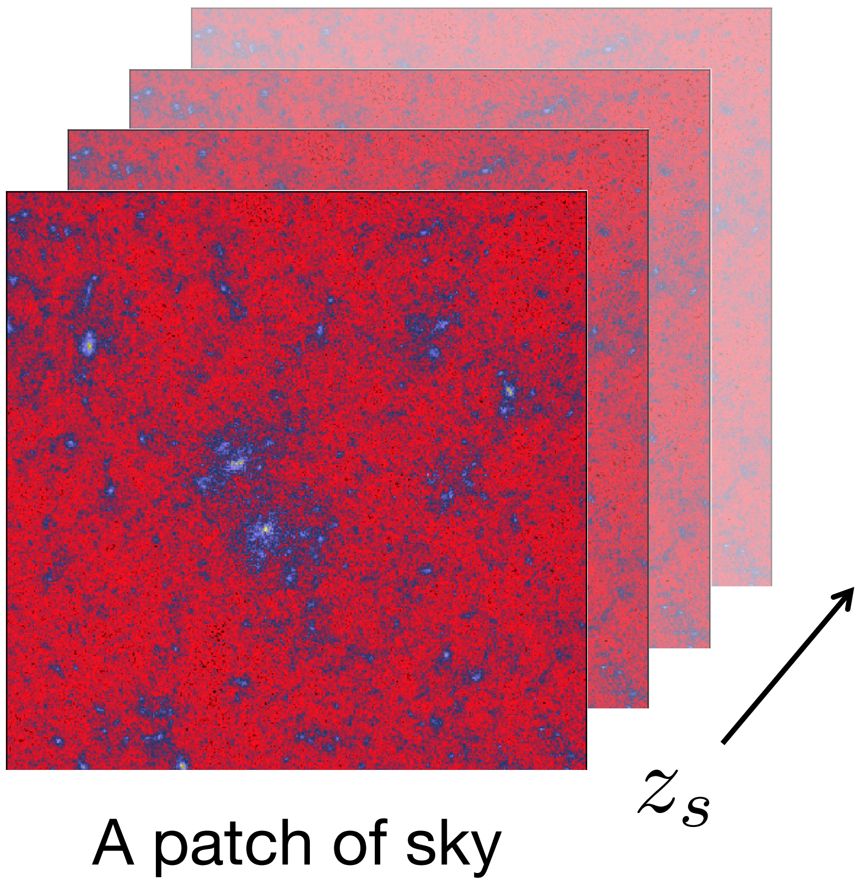
x (256 maps per model)

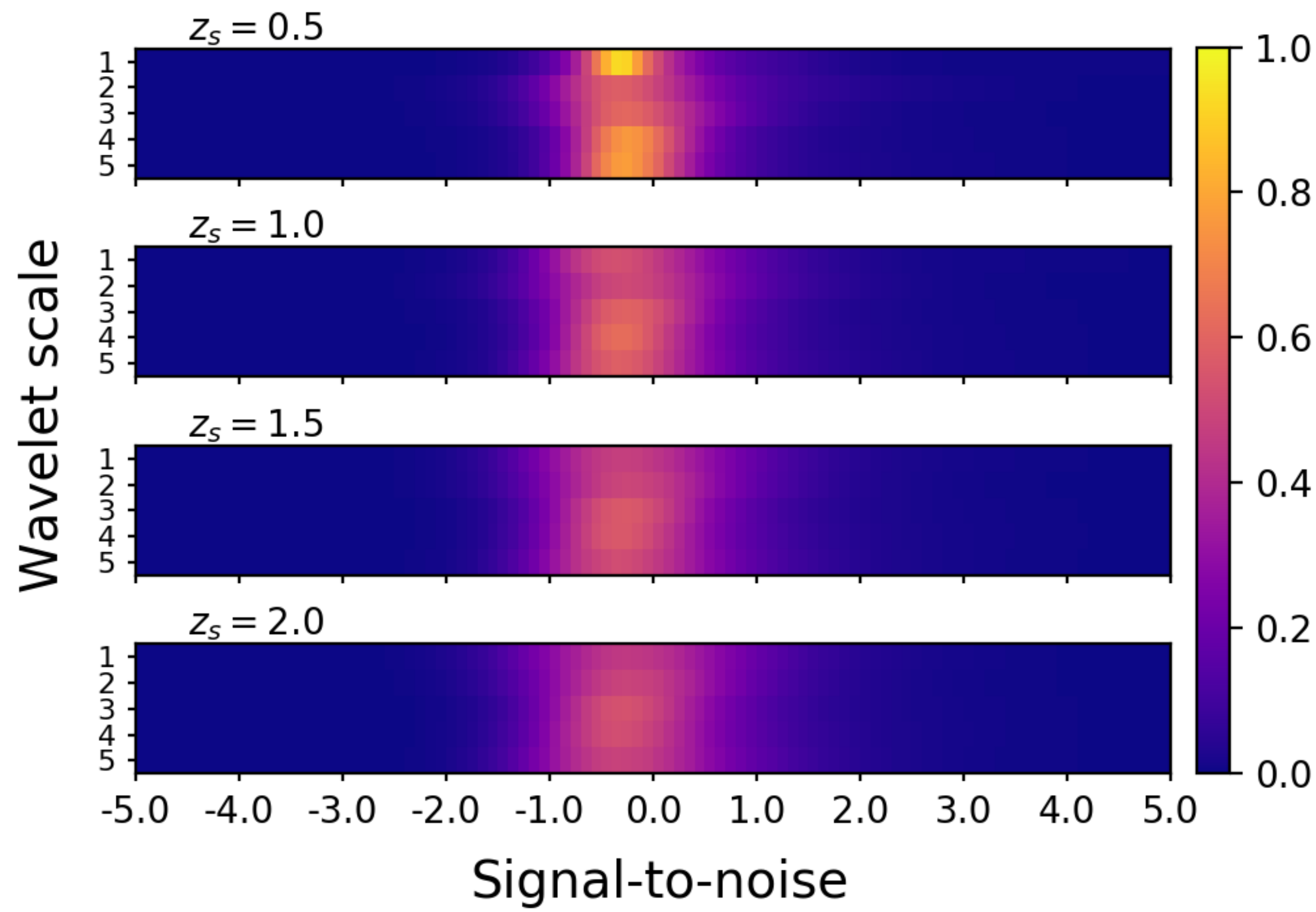
x (4 source redshifts)

x (4 cosmological models)

= a computational challenge

A dimensionally reduced data representation





reduction factor of ~8400

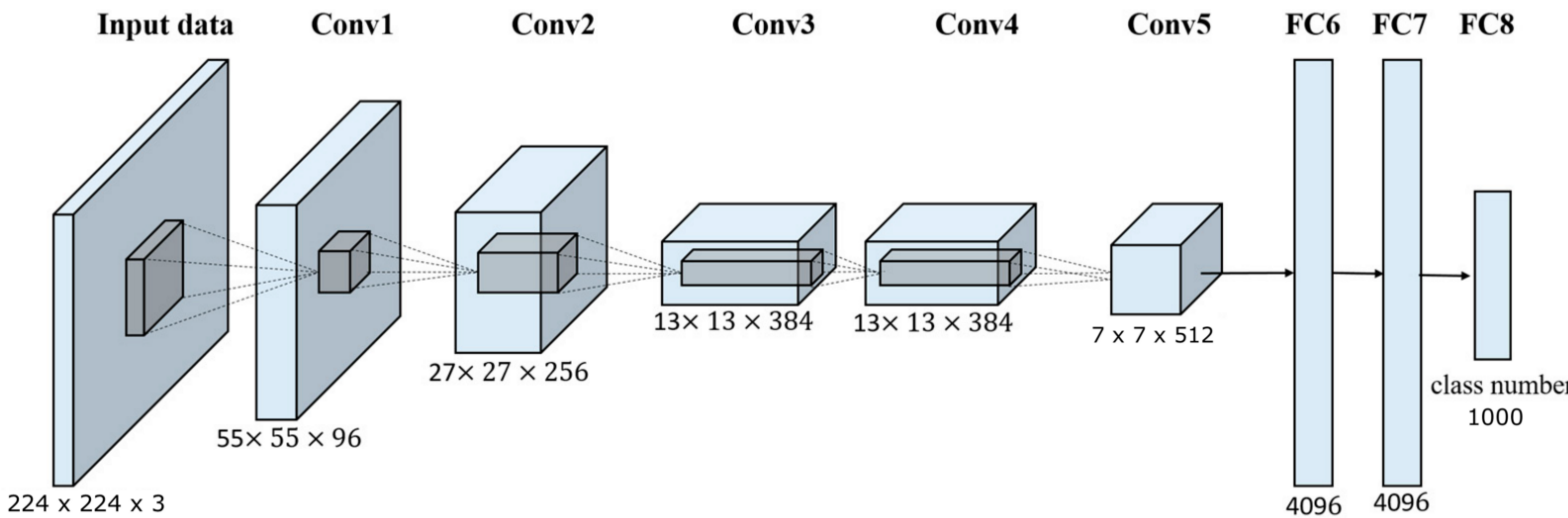
Outline

1. Modified gravity simulations
2. Data representations
3. **ML network architecture**
4. Classifying cosmological models
5. Summary

Network architecture

Network architecture figure

Network architecture

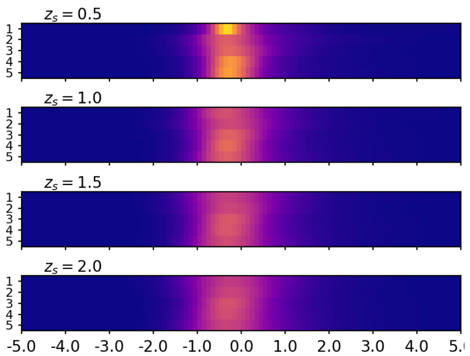


(the dream)

Network architecture

Convolutional neural network (CNN) classification problem

condensed input datacube



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

?

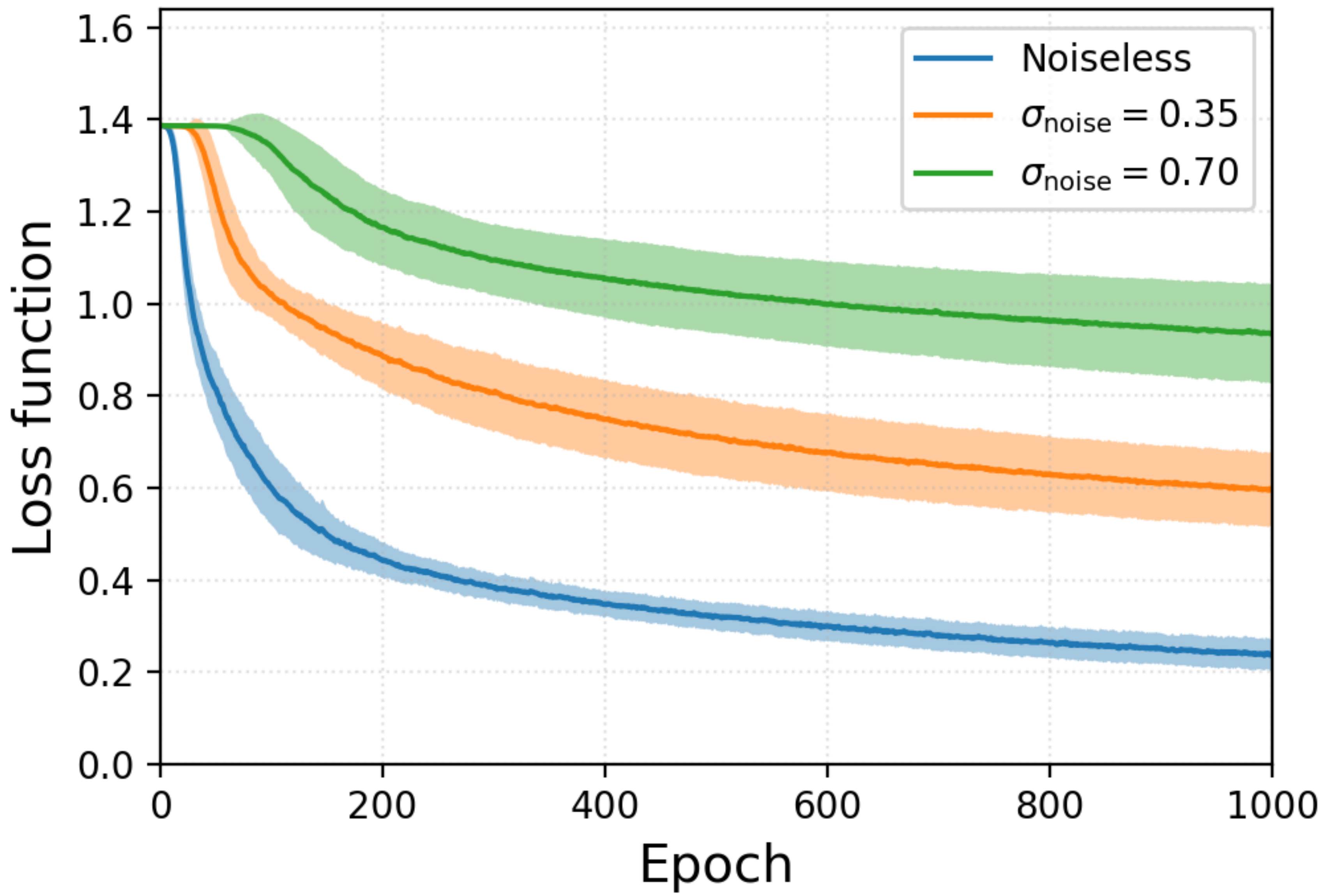
$f_5(R)$

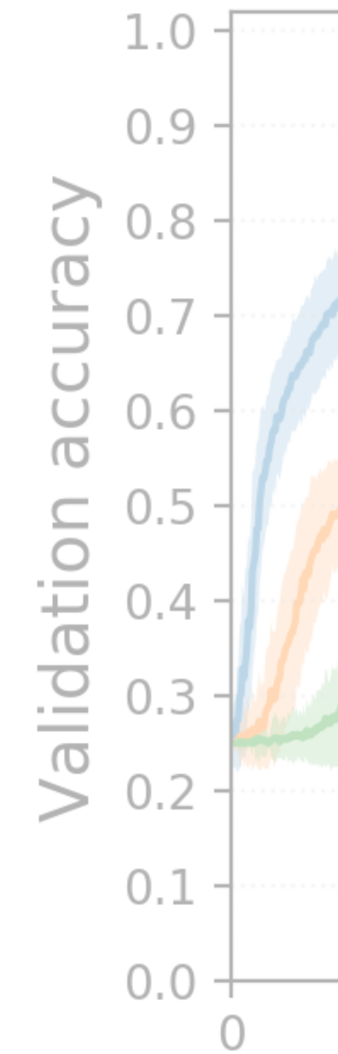
$\longrightarrow$ $M_\nu = 0$ eV

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

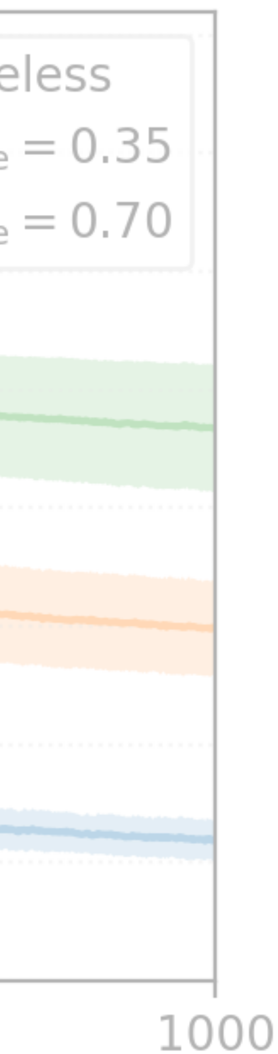
$\longrightarrow$ $M_\nu = 0.15$ eV

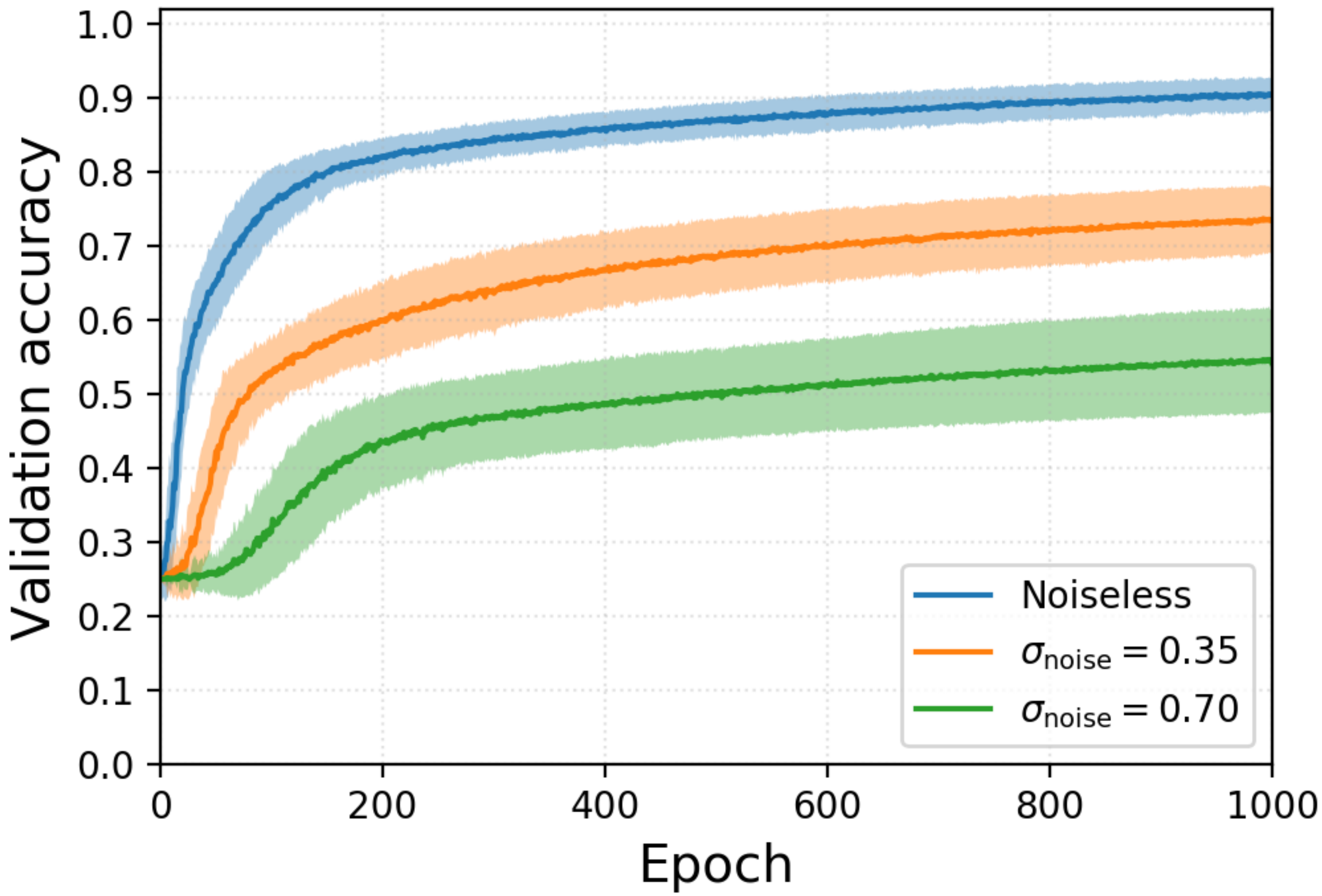
Performance measures





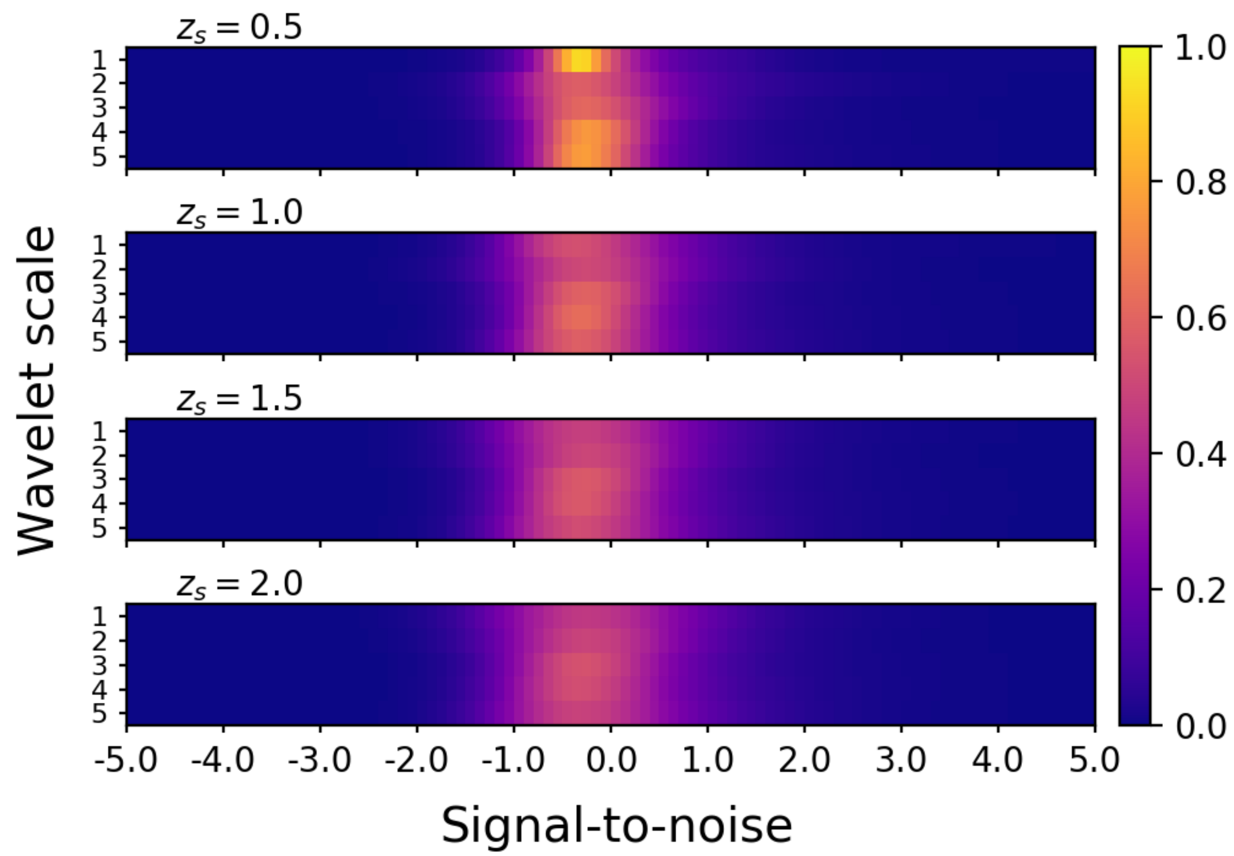
Performance measures



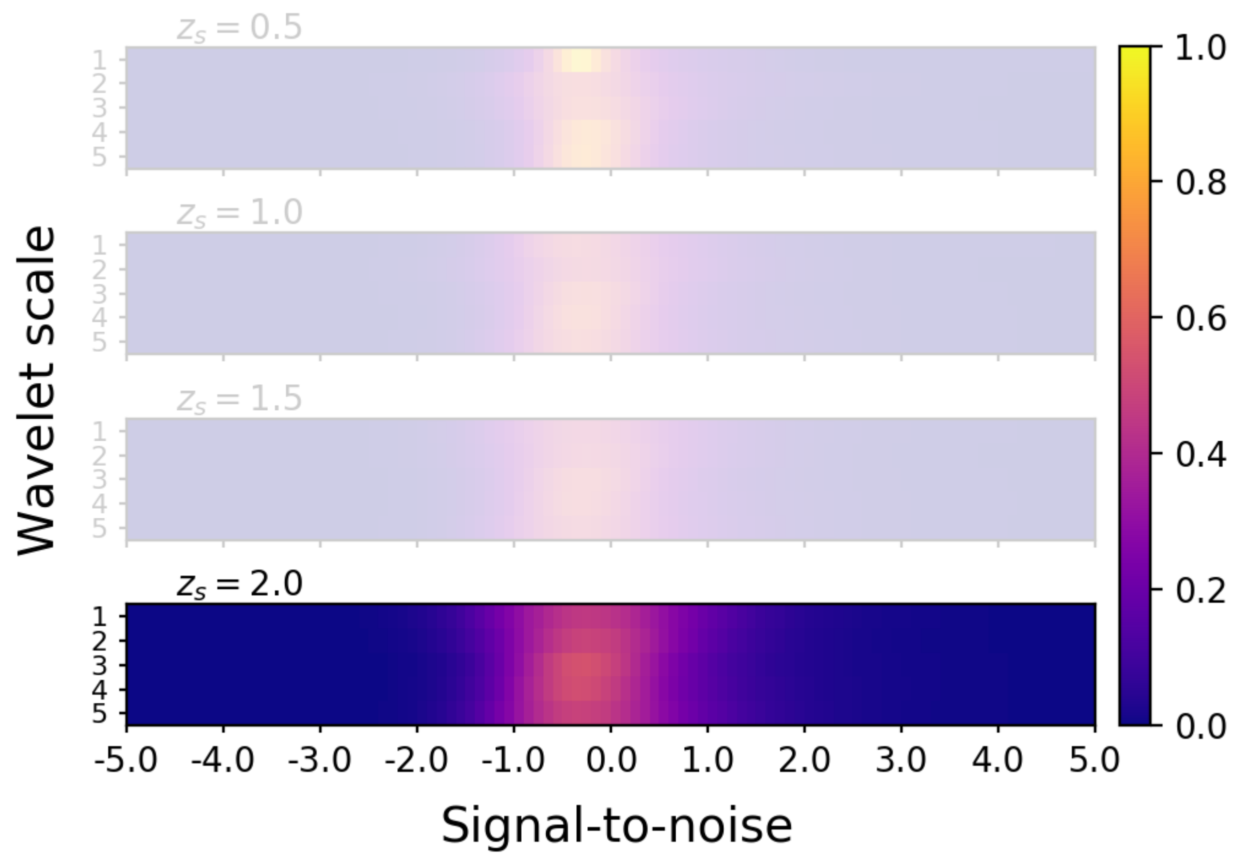


Outline

A warm up with only one redshift



A warm up with only one redshift



↓

Simpler 2D problem

noise level

zero

Single source redshift (z_s = 2.0)

Convolutional neural network (2D)

Prediction (columns) / Truth (rows)

$\sigma_{noise} = 0$	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	0.98	0.00	0.00	0.02
$f_5(R)$ $M_\nu = 0$ eV	0.00	0.83	0.17	0.00
$f_5(R)$ $M_\nu = 0.1$ eV	0.00	0.19	0.72	0.09
$f_5(R)$ $M_\nu = 0.15$ eV	0.01	0.00	0.08	0.90

Peak statistics (best case)

Prediction (columns) / Truth (rows)

$\sigma_{noise} = 0$	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

noise level

zero

Single source redshift ($z_s = 2.0$)

Convolutional neural network (2D)

Prediction (columns) / Truth (rows)

$\sigma_{noise} = 0$	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	0.98	0.00	0.00	0.02
$f_5(R)$ $M_\nu = 0$ eV	0.00	0.83	0.17	0.00
$f_5(R)$ $M_\nu = 0.1$ eV	0.00	0.19	0.72	0.09
$f_5(R)$ $M_\nu = 0.15$ eV	0.01	0.00	0.08	0.90

Peak statistics (best case)

Prediction (columns) / Truth (rows)

$\sigma_{noise} = 0$	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

noise level

optimistic

Single source redshift ($z_s = 2.0$)

Convolutional neural network (2D)

Prediction (columns) / Truth (rows)

$\sigma_{\mathrm{noise}} = 0.35$	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	0.79	0.00	0.03	0.18
$f_5(R)$ $M_\nu = 0$ eV	0.00	0.76	0.22	0.01
$f_5(R)$ $M_\nu = 0.1$ eV	0.02	0.26	0.54	0.17
$f_5(R)$ $M_\nu = 0.15$ eV	0.22	0.01	0.18	0.59

Peak statistics (best case)

Prediction (columns) / Truth (rows)

$\sigma_{\mathrm{noise}} = 0.35$	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	0.14	0.29	0.28

noise level

pessimistic

Single source redshift (z_s = 2.0)

Convolutional neural network (2D)

$\sigma_{\text{noise}} = 0.7$ (Truth \ Prediction)	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	0.44	0.02	0.18	0.36
$f_5(R)$ $M_\nu = 0$ eV	0.02	0.69	0.24	0.04
$f_5(R)$ $M_\nu = 0.1$ eV	0.12	0.31	0.39	0.18
$f_5(R)$ $M_\nu = 0.15$ eV	0.32	0.02	0.19	0.47

Peak statistics (best case)

$\sigma_{\text{noise}} = 0.7$ (Truth \ Prediction)	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

Back to the 3D problem

noise level

zero

Convolutional neural network (2D)

Prediction

$\sigma_{\text{noise}} = 0$ (Truth ↓)	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	0.98	0.00	0.00	0.02
$f_5(R)$ $M_\nu = 0$ eV	0.00	0.83	0.17	0.00
$f_5(R)$ $M_\nu = 0.1$ eV	0.00	0.19	0.72	0.09
$f_5(R)$ $M_\nu = 0.15$ eV	0.01	0.00	0.08	0.90

Convolutional neural network (3D)

Prediction

$\sigma_{\text{noise}} = 0$ (Truth ↓)	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	0.86	0.14	0.00
$f_5(R)$ $M_\nu = 0.1$ eV	0.00	0.15	0.80	0.05
$f_5(R)$ $M_\nu = 0.15$ eV	0.00	0.00	0.04	0.96

noise level

optimistic

Convolutional neural network (2D)

Prediction

$\sigma_{\text{noise}} = 0.35$ (Truth)	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	0.79	0.00	0.03	0.18
$f_5(R)$ $M_\nu = 0$ eV	0.00	0.76	0.22	0.01
$f_5(R)$ $M_\nu = 0.1$ eV	0.02	0.26	0.54	0.17
$f_5(R)$ $M_\nu = 0.15$ eV	0.22	0.01	0.18	0.59

Convolutional neural network (3D)

Prediction

$\sigma_{\text{noise}} = 0.35$ (Truth)	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	0.87	0.00	0.02	0.11
$f_5(R)$ $M_\nu = 0$ eV	0.00	0.76	0.23	0.01
$f_5(R)$ $M_\nu = 0.1$ eV	0.03	0.23	0.58	0.16
$f_5(R)$ $M_\nu = 0.15$ eV	0.11	0.00	0.15	0.73

noise level

pessimistic

Convolutional neural network (2D)

Prediction (columns); Truth (rows)

$\sigma_{\mathrm{noise}} = 0.7$	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	0.44	0.02	0.18	0.36
$f_5(R)$ $M_\nu = 0$ eV	0.02	0.69	0.24	0.04
$f_5(R)$ $M_\nu = 0.1$ eV	0.12	0.31	0.39	0.18
$f_5(R)$ $M_\nu = 0.15$ eV	0.32	0.02	0.19	0.47

Convolutional neural network (3D)

Prediction (columns); Truth (rows)

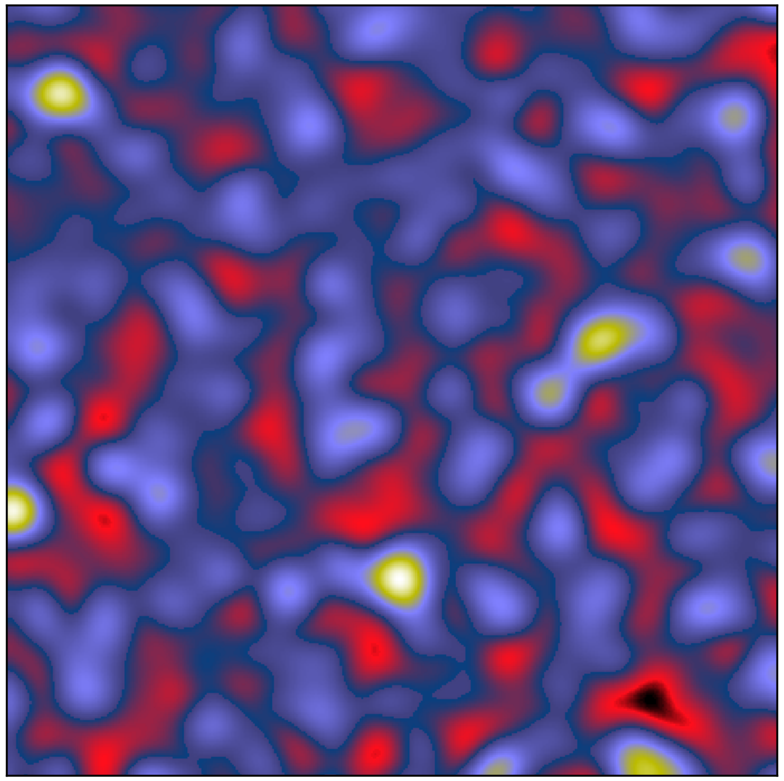
$\sigma_{\mathrm{noise}} = 0.7$	ΛCDM	$f_5(R)$ $M_\nu = 0$ eV	$f_5(R)$ $M_\nu = 0.1$ eV	$f_5(R)$ $M_\nu = 0.15$ eV
ΛCDM	0.50	0.02	0.18	0.30
$f_5(R)$ $M_\nu = 0$ eV	0.02	0.70	0.25	0.03
$f_5(R)$ $M_\nu = 0.1$ eV	0.15	0.28	0.42	0.15
$f_5(R)$ $M_\nu = 0.15$ eV	0.28	0.02	0.15	0.55

Outline

1. Modified gravity simulations
2. Data representations
3. ML network architecture
4. Classifying cosmological models
5. **Summary**

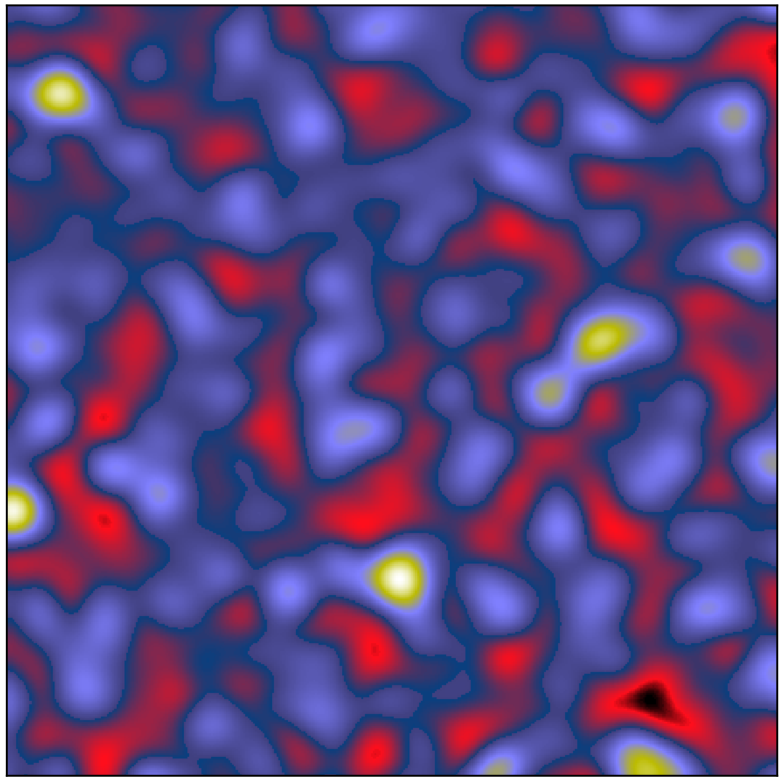
To sum up

- 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



Room for improvement

- The noise-free case still isn't perfect
- May be worth including a separate denoising step
- Regression vs. classification
- Ultimately test on real data

Thank you