Weak lensing cosmology beyond two-point functions

Zoltán Haiman
(Columbia University)

Weak lensing cosmology beyond two-point functions

Zoltán Haiman
(Columbia University)

Unsolved problems

- How much cosmological information is contained, in principle, in a (perfect) weak lensing map?

- How well can we constrain background cosmology, in practice, from observed lensing data?

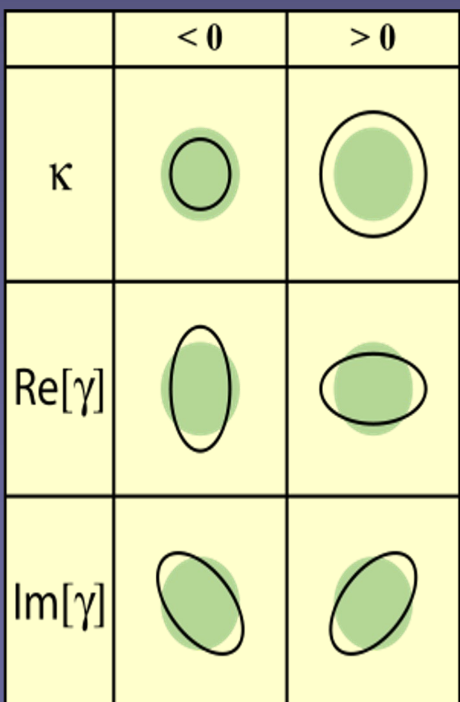
Gravitational Lensing

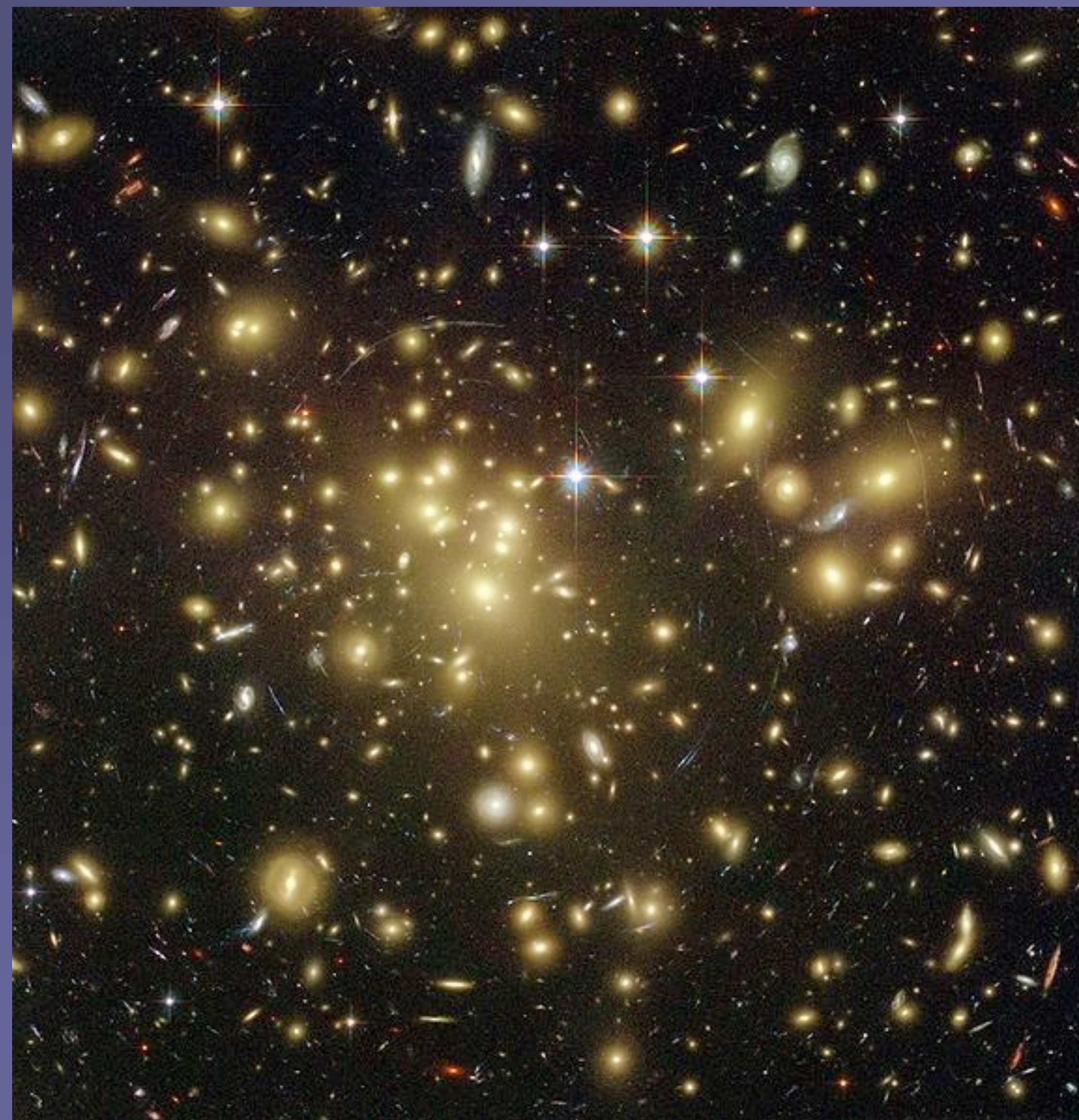
Abell 1689; Benitez et al. (2003)

Unlensed position(θ_j)
Observed position (θ_i)

$$f_{obs}(\theta_i) = f_s(A_{ij}\theta_j)$$

$$A_{ij} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}$$



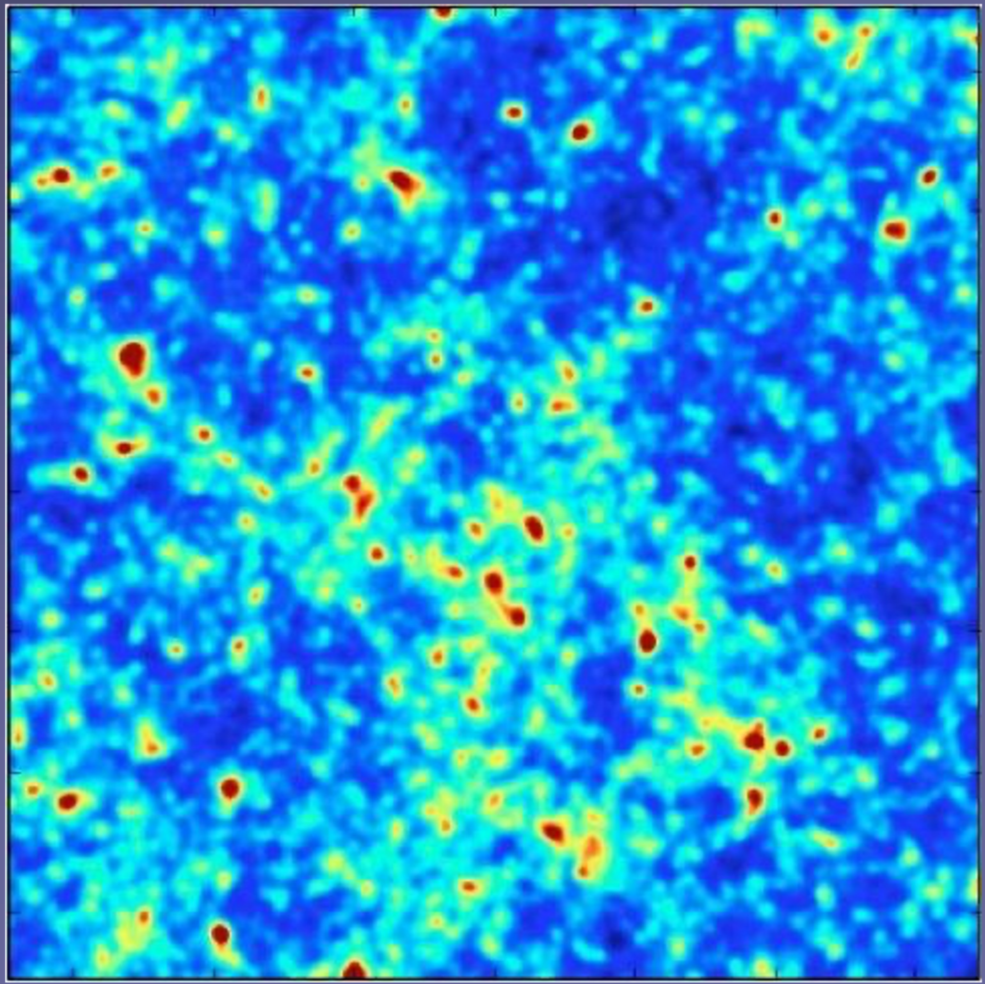


Weak lensing: convergence map

- Measure ellipticities of galaxies
- Convert to convergence (=magnification)
- Smooth over ~arcmin^2 patches

$$\hat{\kappa}(\mathbf{s}) = \frac{1}{2}\left(\frac{k_1^2 - k_2^2}{k_1^2 + k_2^2}\right)\hat{\gamma}_1(\mathbf{s}) + \frac{k_1 k_2}{k_1^2 + k_2^2}\hat{\gamma}_2(\mathbf{s})$$

Kaiser & Squires 1993



Workhorse: 2-point functions

- **Real-space: two-point correlation functions**

  $\xi(\Theta) = \langle \kappa(\vec{\theta})\, \kappa(\vec{\theta}+\vec{\Theta}) \rangle$

  $\langle \vec{\gamma}\,\vec{\gamma} \rangle$ **in principle, same information**

- **Fourier space: convergence power spectrum**

  $\langle \kappa(\vec{l})\, \kappa^*(\vec{l}-\vec{l}') \rangle = 2\pi\delta(\vec{l}-\vec{l}')\, P(l)$

$$P_\kappa(l) = \frac{9}{4}\Omega_m^2 \frac{H_0^4}{c^4}\int_0^\infty dz \left[\frac{d\chi(z)}{dz}\right] \frac{\xi^2\left[\chi(z)\right]}{a^2(z)} P_{3D}\left(\frac{l}{\chi(z)};z\right),$$

$$\xi(\chi) = \int_z^\infty dz'\, n_{gal}(z')\, \frac{\chi(z')-\chi(z)}{\chi(z')}.$$

Workhorse: 2-point functions

- **Real-space: two-point correlation functions**

  $\xi(\Theta) = \langle \kappa(\vec{\theta})\, \kappa(\vec{\theta}+\vec{\Theta}) \rangle$

  $\langle \vec{\gamma}\,\vec{\gamma} \rangle$ **in principle, same information**

- **Fourier space: convergence power spectrum**

  $\langle \kappa(\vec{l})\, \kappa^*(\vec{l}-\vec{l}') \rangle = 2\pi\delta(\vec{l}-\vec{l}')\, P(l)$

$$P_\kappa(l) = \frac{9}{4}\Omega_m^2 \frac{H_0^4}{c^4} \int_0^\infty dz \left[\frac{d\chi(z)}{dz}\right] \frac{\xi^2[\chi(z)]}{a^2(z)} P_{3D}\left(\frac{l}{\chi(z)}; z\right),$$

$$\xi(\chi) = \int_z^\infty dz'\, n_{\mathrm{gal}}(z') \frac{\chi(z') - \chi(z)}{\chi(z')}.$$

geometry

Workhorse: 2-point functions

- **Real-space: two-point correlation functions**

  $\xi(\vec{\Theta}) = \langle \boldsymbol{\kappa}(\vec{\boldsymbol{\theta}})\, \boldsymbol{\kappa}(\vec{\boldsymbol{\theta}}+\vec{\Theta})\rangle$

  $\langle \vec{\gamma}\,\vec{\gamma} \rangle$ **in principle, same information**

- **Fourier space: convergence power spectrum**

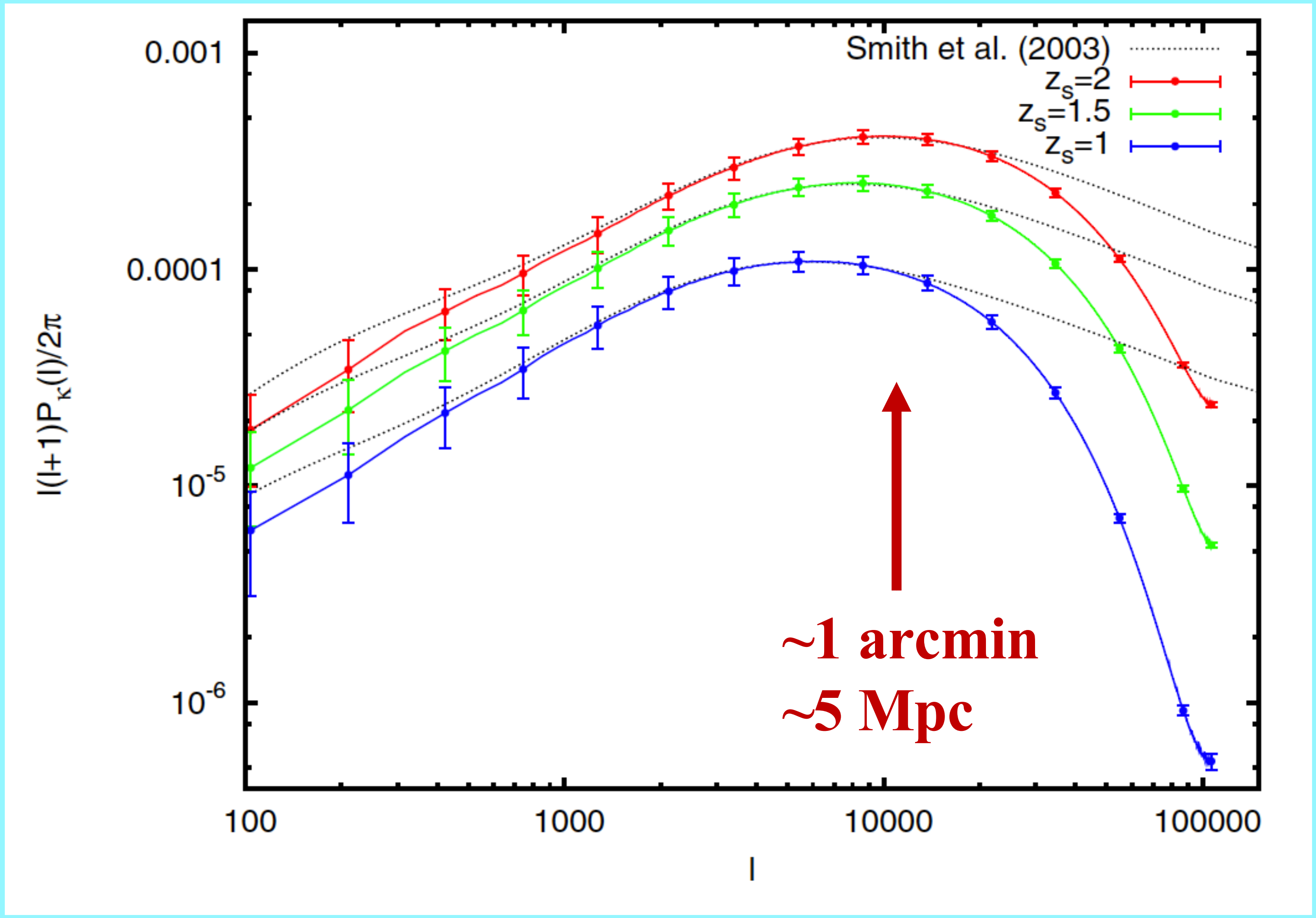
  $\langle \boldsymbol{\kappa}(\vec{l})\, \boldsymbol{\kappa}^*(\vec{l}-\vec{l}')\rangle = 2\pi\delta(\vec{l}-\vec{l}')\, P(l)$

$$P_\kappa(l) = \frac{9}{4}\Omega_m^2 \frac{H_0^4}{c^4}\int_0^\infty dz \left[\frac{d\chi(z)}{dz}\right] \frac{\xi^2\left[\chi(z)\right]}{a^2(z)} P_{3D}\left(\frac{l}{\chi(z)};z\right),$$

$$\xi(\chi) = \int_z^\infty dz'\, n_{\rm gal}(z')\, \frac{\chi(z')-\chi(z)}{\chi(z')}.$$

geometry
growth

Convergence power spectrum



Kratochvil et al. 2012

Cosmology results

signal is weak (~1%), must average over many galaxies:

(0.3/√900) → 900 galaxies for S/N=1 detection of a systematic γ ~0.01

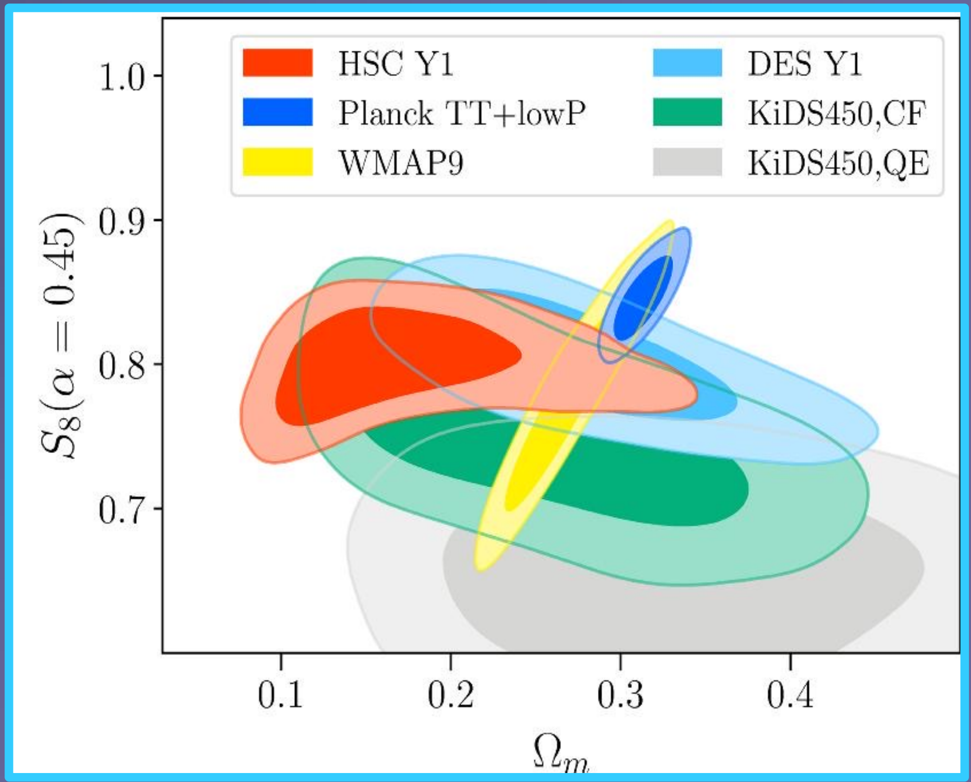
→ 900×10^4 ~ 10^7 galaxies for ~1% error on γ → need ~100 deg^2

Canada-France-Hawaii Telescope (CFHTLenS)
154 deg^2 imaging (6×10^6 gals)
Kilbinger et al. (2013)

Kilo Degree Survey (KiDS-1000)
1006 deg^2 imaging in 4 bands (25×10^6 gals)
Heymans et al. (2020)

Dark Energy Survey (DES; Year 3)
4143 deg^2 imaging in 5 bands (100×10^6 gals)
Amon et al. (2021)

Subaru Hyper Suprime-Cam (HSC; Year 1)
137 deg^2 imaging in 5 bands (9×10^6 gals)
Hikage et al. (2019)

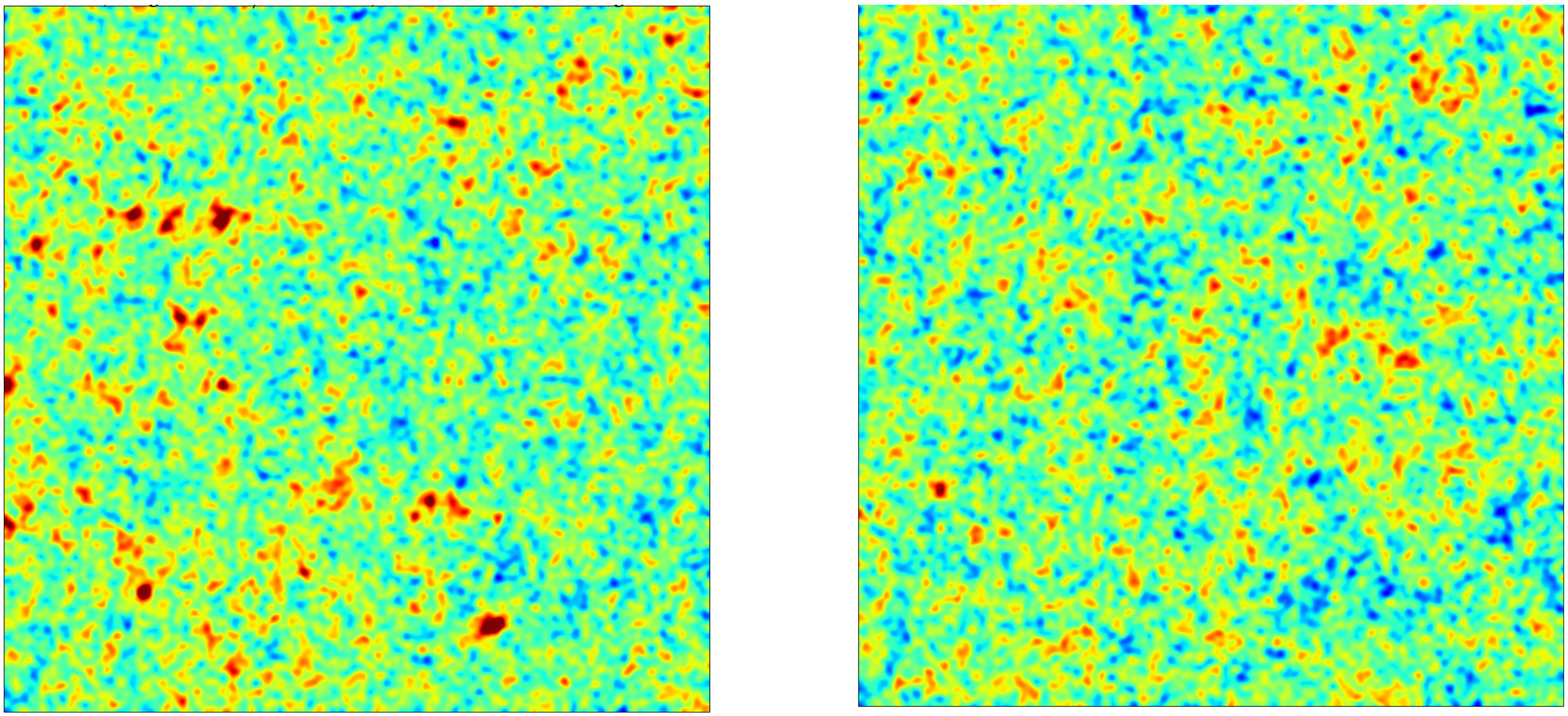


The Future: Full HSC, Euclid, LSST, Roman 10^7 → 10^8 → 10^9+ gals

Cosmic Shear is Not Gaussian

Millennium simulation – Volker Springel, MPA

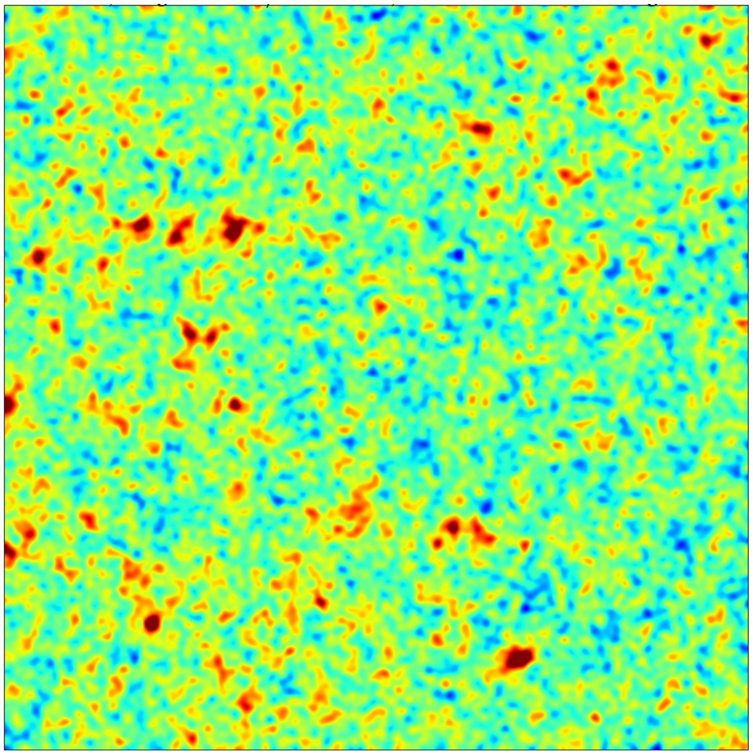
Cosmic Shear is Not Gaussian



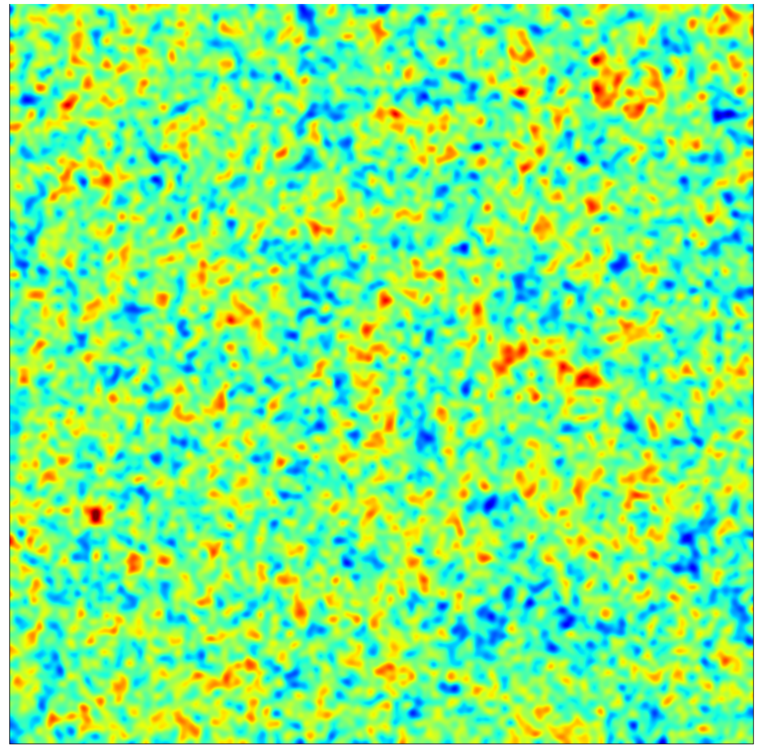
Millennium simulation – Volker Springel, MPA

Cosmic Shear is Not Gaussian

lensing by cosmic structures



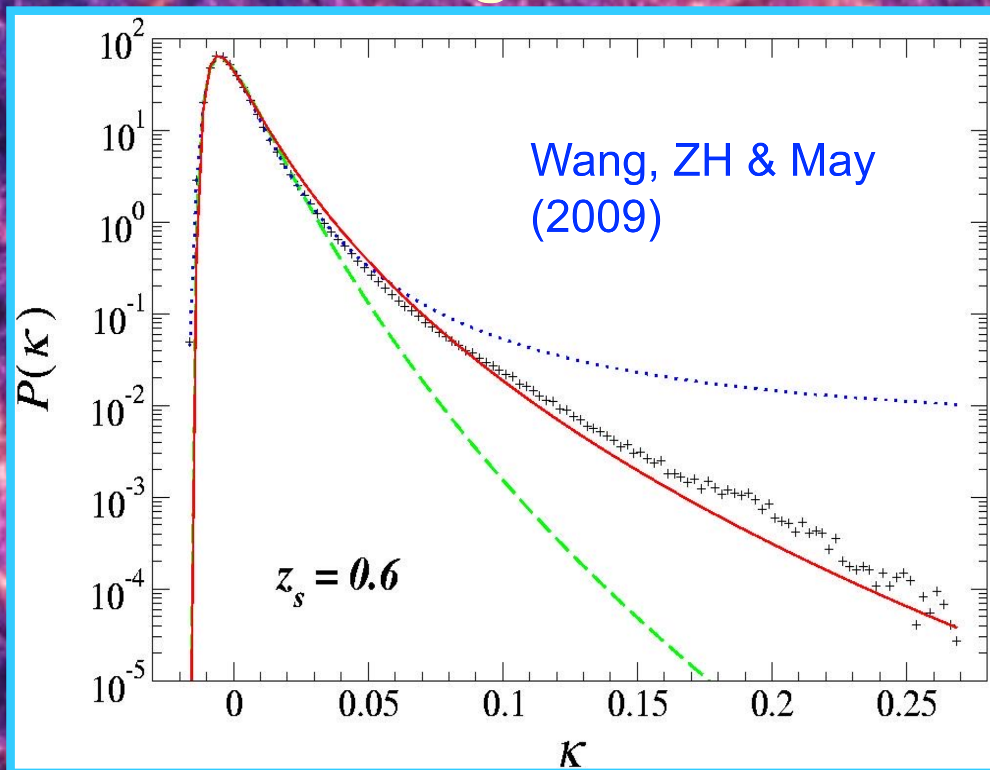
mock Gaussian equivalent



Millennium simulation – Volker Springel, MPA

Cosmic Shear is Not Gaussian

PDF of convergence:



Millennium simulation – Volker Springel, MPA

Looking for beyond-Gaussian info

Approaches:

1 perturbative expansions:

higher-order moments (skewness, kurtosis …)

higher-order correlation functions (3pt, 4pt ….)

Fourier counterparts (bispectra, trispectra ….)

2 Other morphological ”features”:

peaks, Minkowski functionals, shapelets …

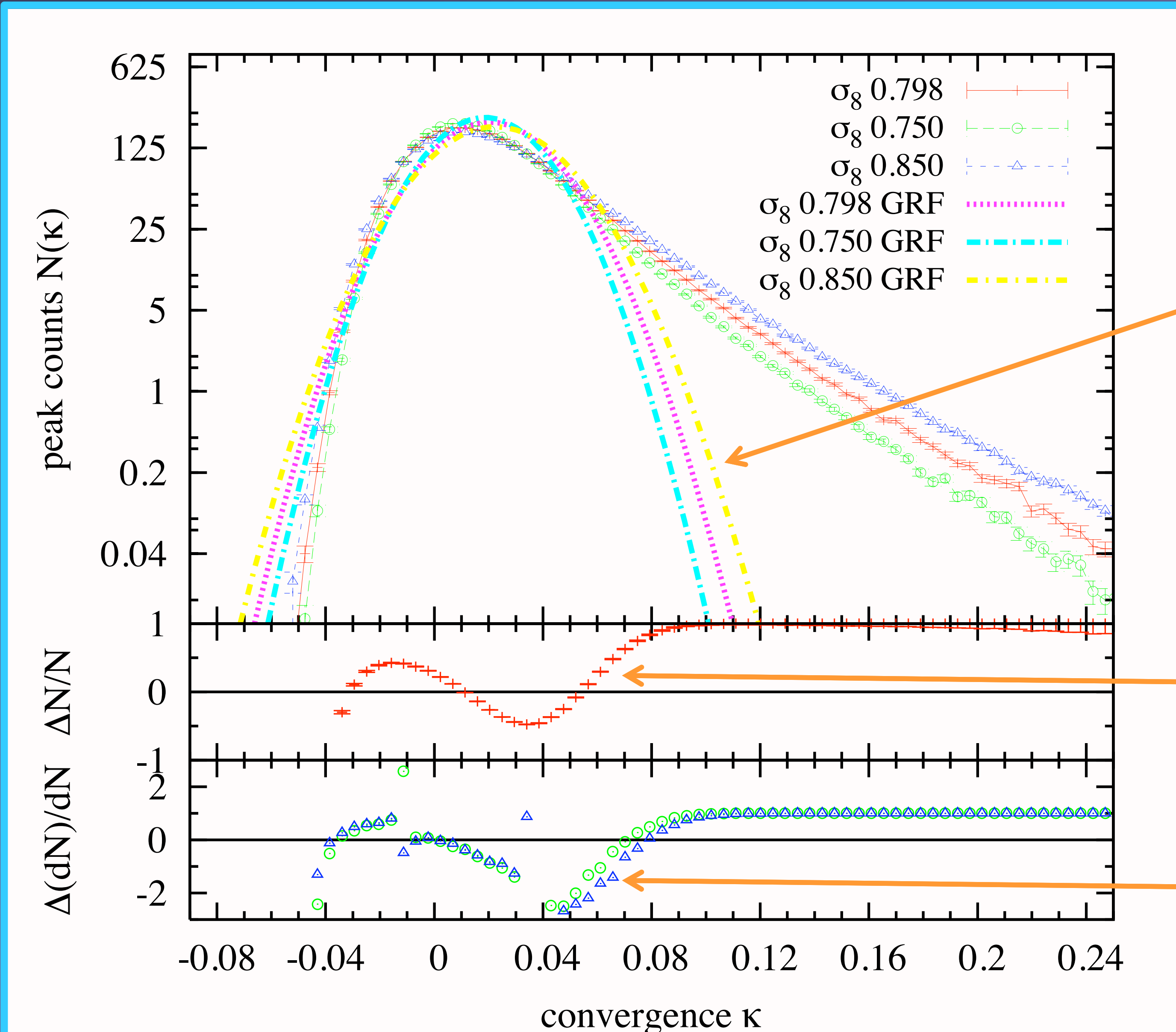
3 ”Gaussianization”: transform lensing field locally

4 machine learning: *can be cast as 2D image classification*

Questions:

- how do these respond to cosmology *vs* systematics
- extra info is from small, nonlinear scales - modeling
- how do you tell whether most info has been found?

Peak Counts

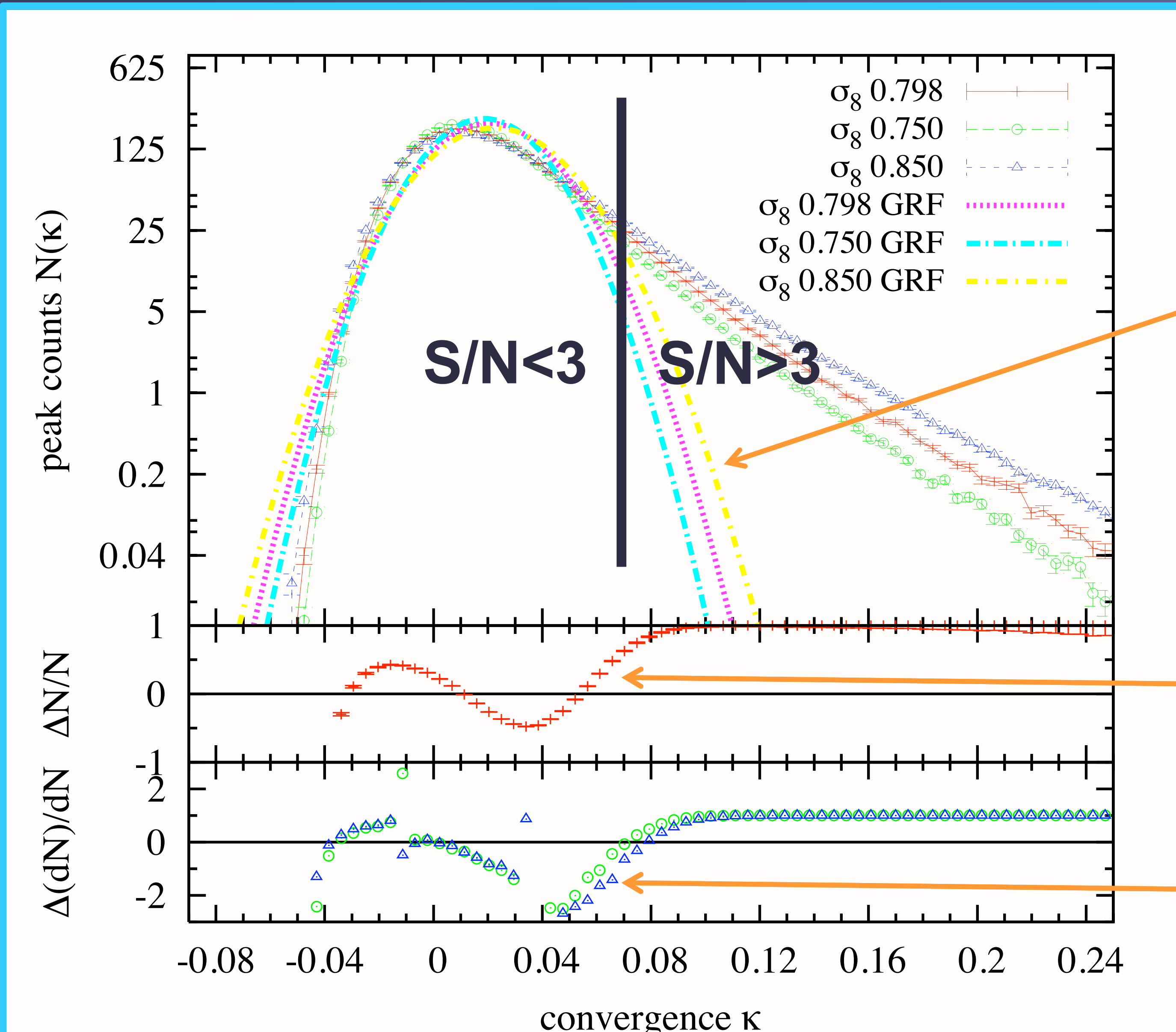


analytic predictions for GRF with same power spec.

Peak counts Non-Gaussian

Cosmology dependence Non-Gaussian

Peak Counts



analytic predictions for GRF with same power spec.

Peak counts Non-Gaussian

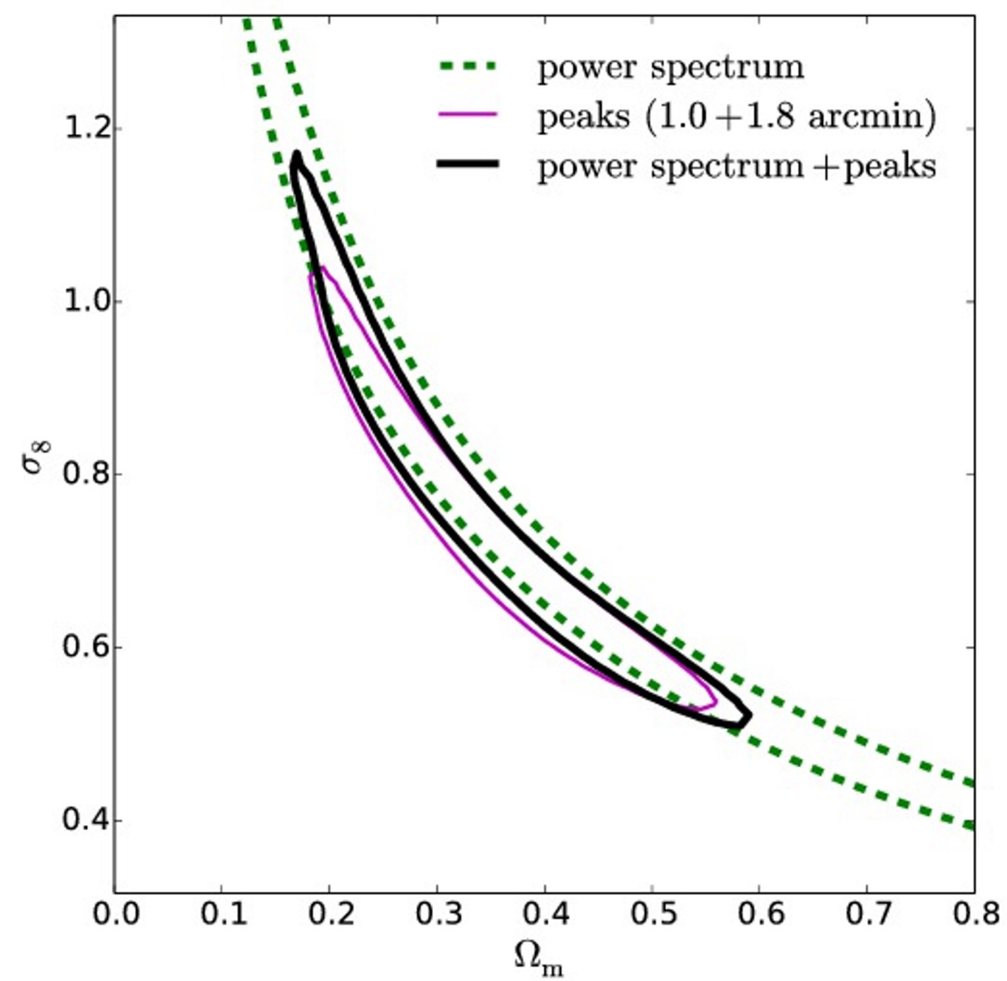
Cosmology dependence Non-Gaussian

Forward-modeling CFHTLenS

Liu et al. (2015)

- w unconstrained
- Adding peaks improves constraint by factor ~2
- Majority of constraint is coming from low peaks
- No tension with Planck

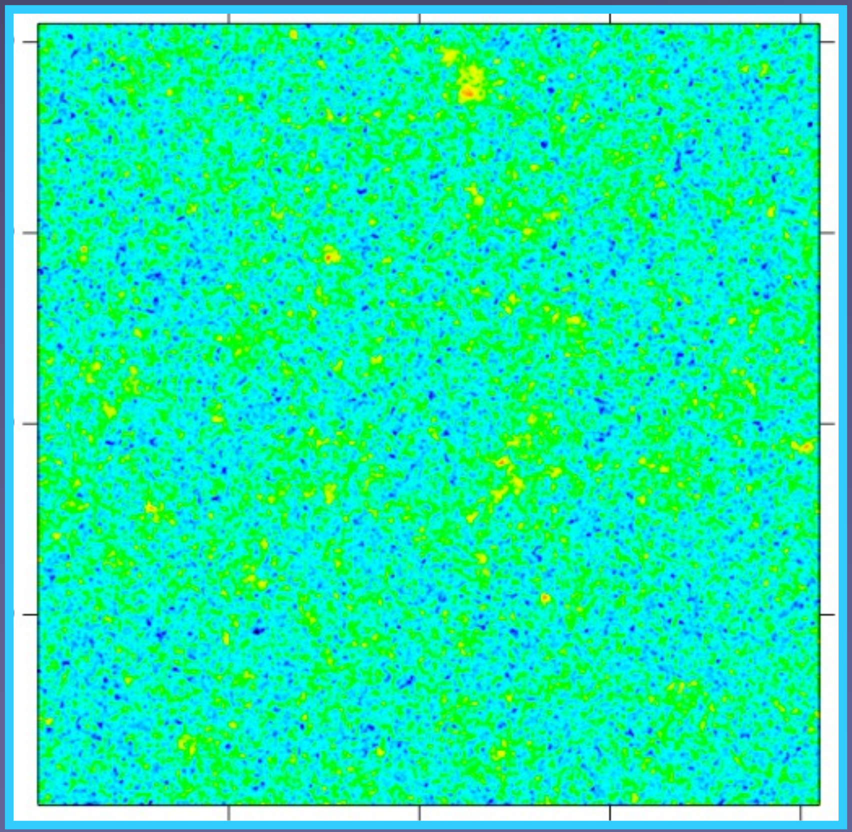
	w–Ω_m		Ω_m–σ_8	
	68%	95%	68%	95%
power spectrum	1.00	1.74	1.00	1.99
peak counts	0.41	1.01	0.59	1.51
combined	0.42	1.05	0.61	1.46



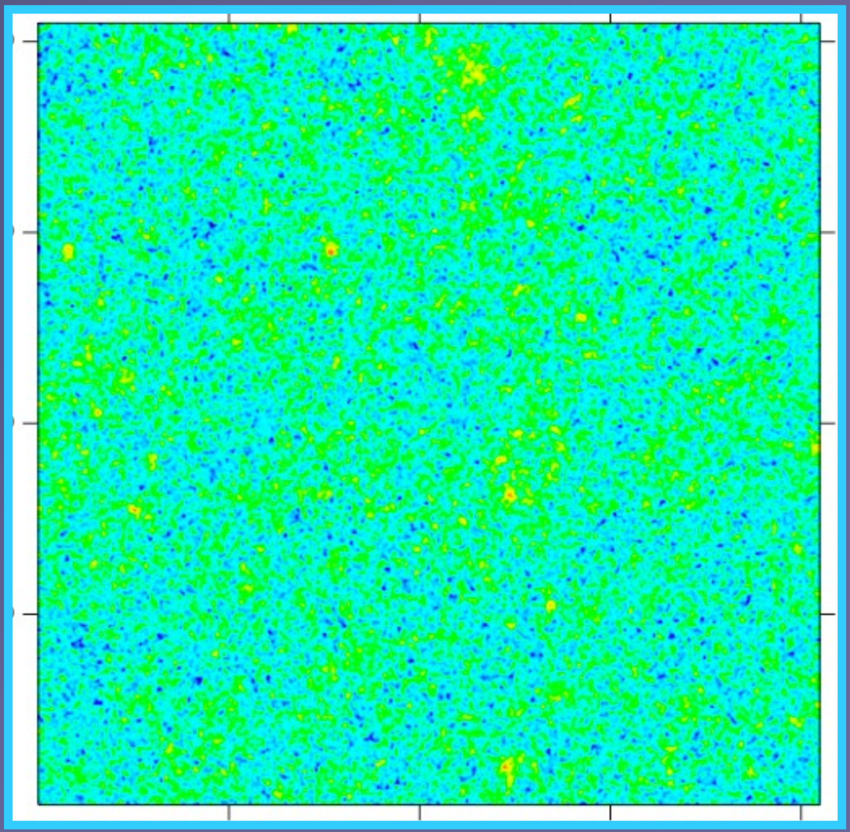
How to go beyond this?

Not quite “cats vs dogs” but these 2D images do look different..

w=-1



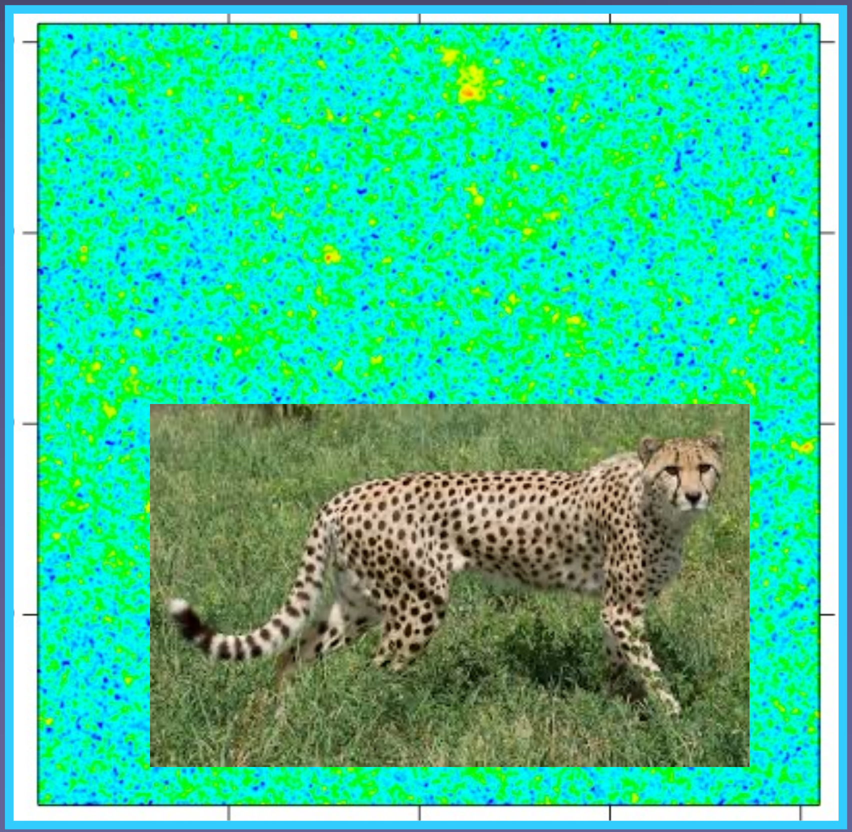
w=-0.8



How to go beyond this?

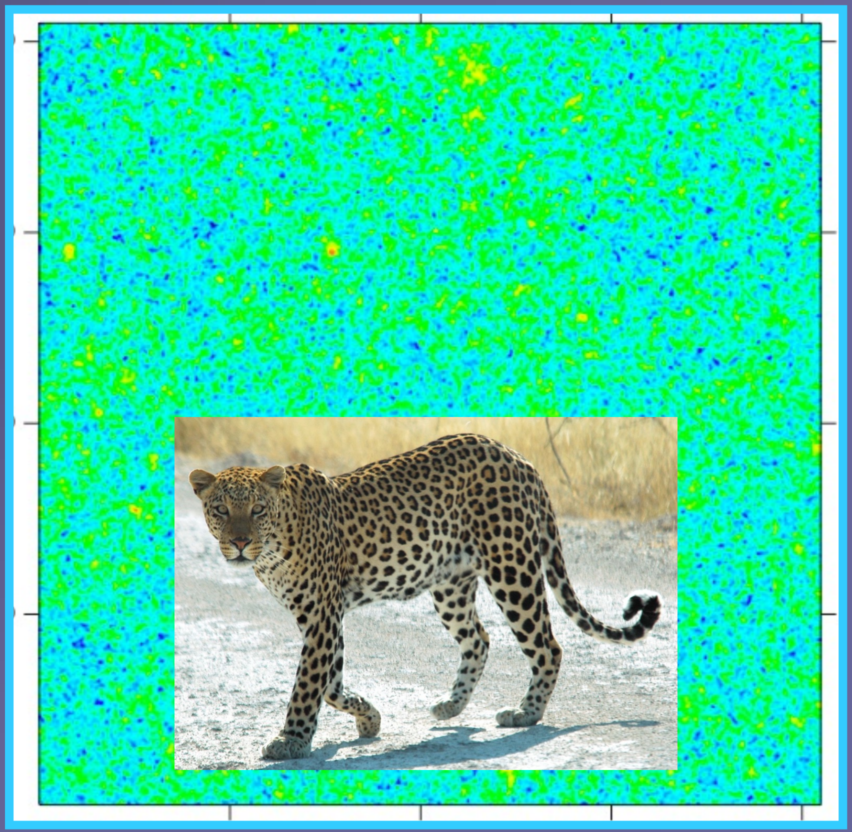
Not quite "cats vs dogs" but these 2D images do look different..

w=-1



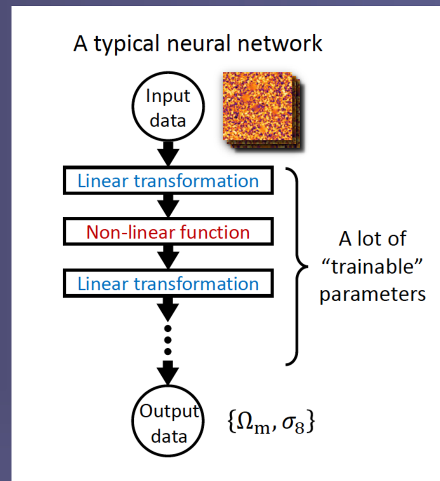
Cheetah? Leopard?

w=-0.8



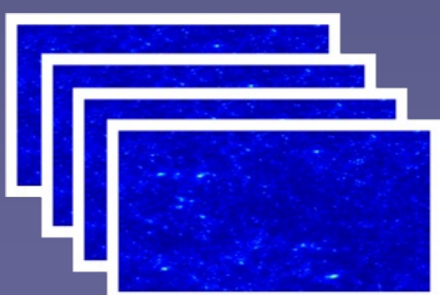
Leopard? Cheetah?

Deep convolutional neural network



Training

Training set of maps (50-70%)



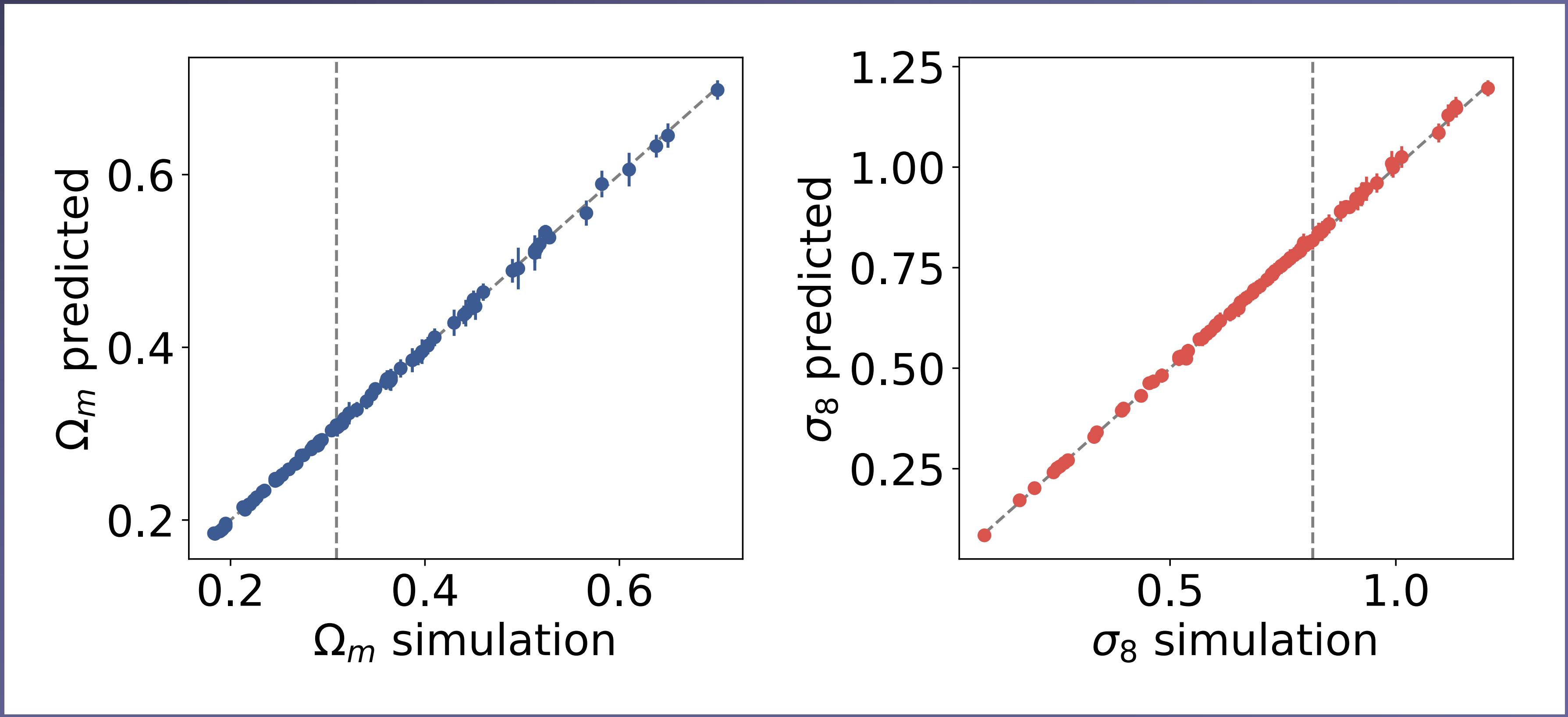
CNN (black box)

Loss function (to be minimized)

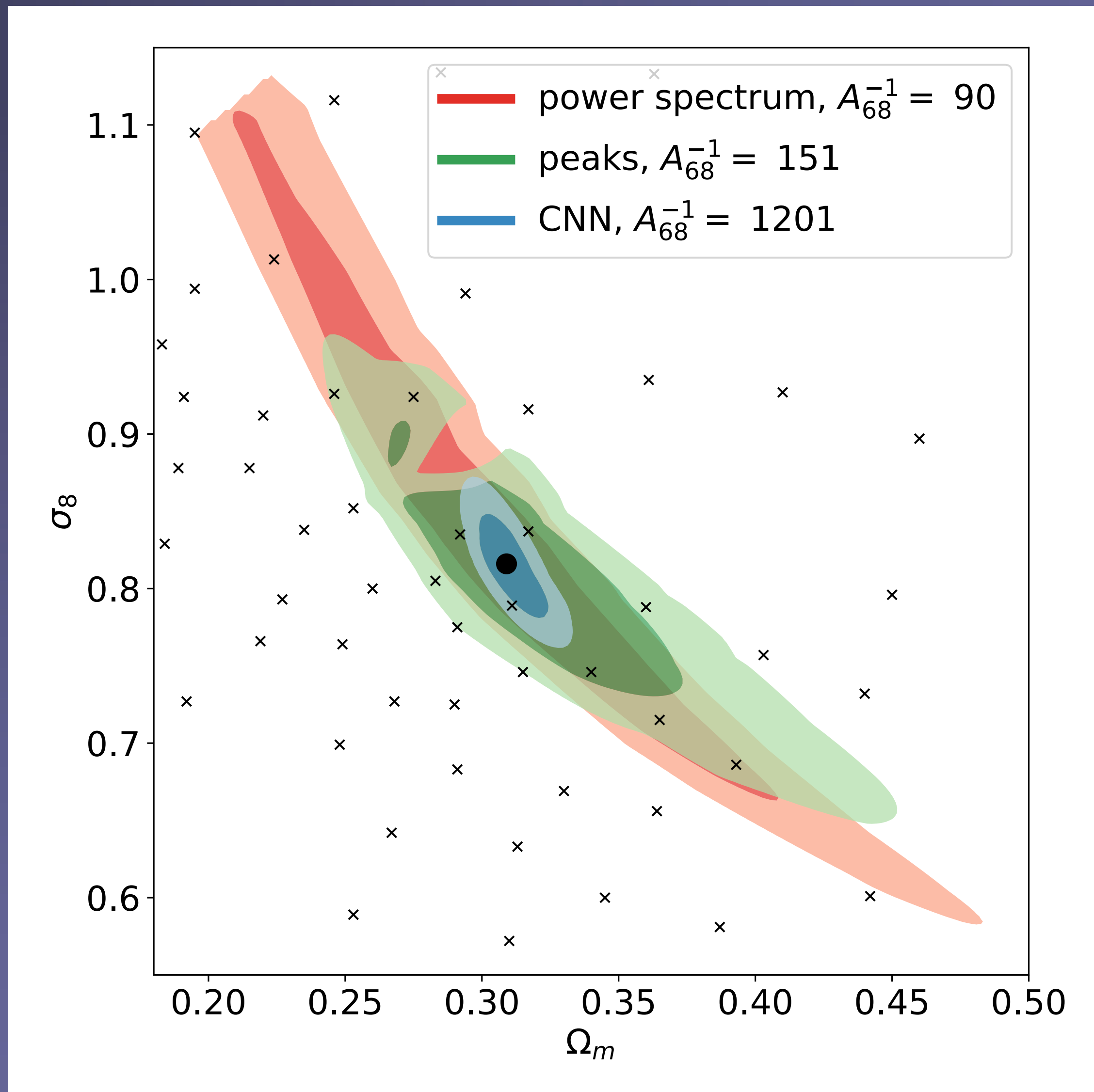
$$\sum_{\text{map}\in\text{batch}} W_{\text{cosmo}}\left(|\sigma_8^{\text{pred}} - \sigma_8^{\text{true}}| + |\Omega_m^{\text{pred}} - \Omega_m^{\text{true}}|\right)$$

Update weights (e.g. Adams optimizer, stochastic gradient descent)

Constraints from CNN



Noiseless maps



Constraints improve by factor of 13(!). Passes Gaussian test

Unsolved problems

- How much cosmological information is contained, in principle, in a (perfect) weak lensing map?

- How well can we constrain background cosmology, in practice, from observed lensing data?

Unsolved problems

- How much cosmological information is contained, in principle, in a (perfect) lensing map?

→ **At least an order of magnitude more than in power spectrum**

- How well can we constrain background cosmology, in practice, from observed lensing data?

CNN on noisy maps

Confidence range ratios around two input cosmologies
$(\Omega_m, \sigma_8) = (0.26, 0.8) - (0.309, 0.816)$

Table 2. The table lists the relative sizes of the 68 percent credible contour areas of the power spectrum and peak counts compared to the CNN. The CNN achieves smaller 68 percent credible contour areas than the power spectrum for any noise level, and also outperforms the peak counts when the galaxy density is at least 30 arcmin^{-2}.

A_{68} ratio	Noiseless	100 gal arcmin^{-2}	75 gal arcmin^{-2}	50 gal arcmin^{-2}	30 gal arcmin^{-2}	10 gal arcmin^{-2}
Power spectrum / CNN	13	3.7–4.6	3.5–4.1	3–3.6	2.4–2.8	1.4–1.5
Peak counts / CNN	8	1.5–2.1	1.4–1.9	1.2–1.7	1.05–1.42	0.9–1.1

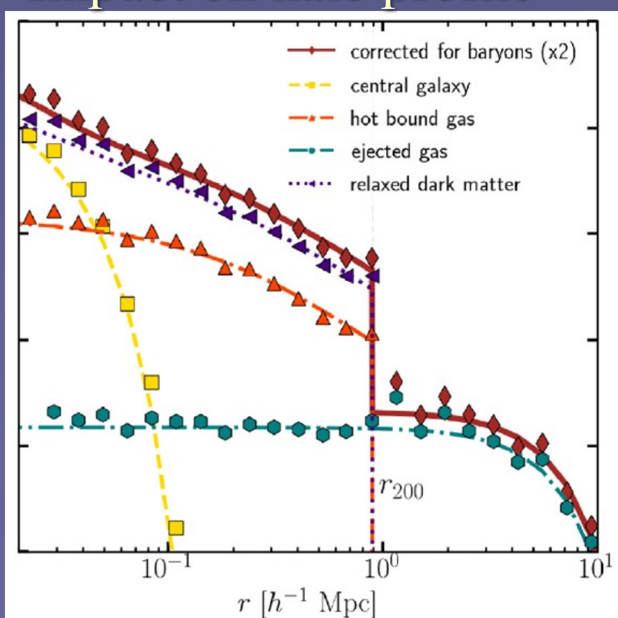
Baryons

- Hydro simulations *vs*
- Baryon correction models (BCM)

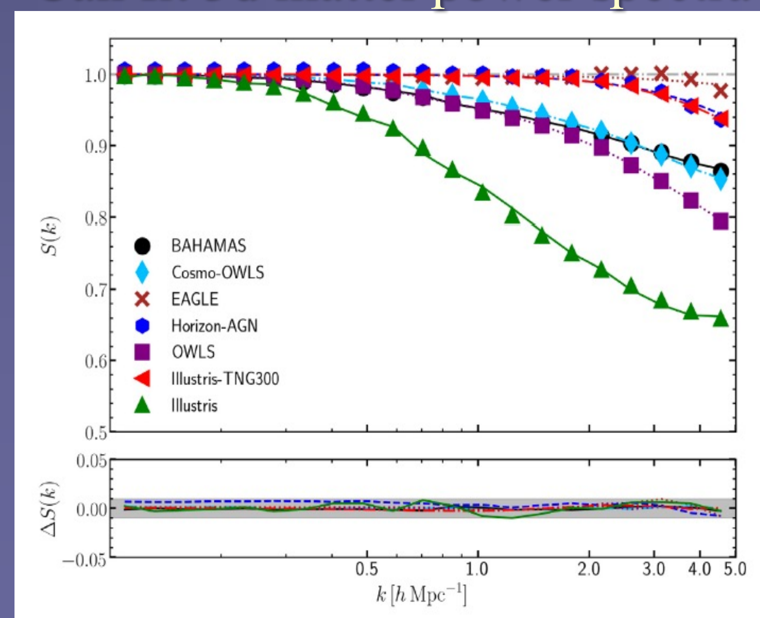
Arico+ 2020
Schneider & Teyssier 2015

Parameter	Description	Fiducial Value ($z = 0$)
M_c	Halo mass scale for retaining half of the total gas	$3.3 \times 10^{13}\, h^{-1}\, M_\odot$
M_1	Characteristic halo mass for a galaxy mass fraction $\epsilon = 0.023$	$8.63 \times 10^{11}\, h^{-1}\, M_\odot$
η	Maximum distance of gas ejection in terms of the halo escape radius	0.54
β	Slope of the gas fraction as a function of halo mass	0.12

Impact on halo profile



Can fit 3d matter power spectra

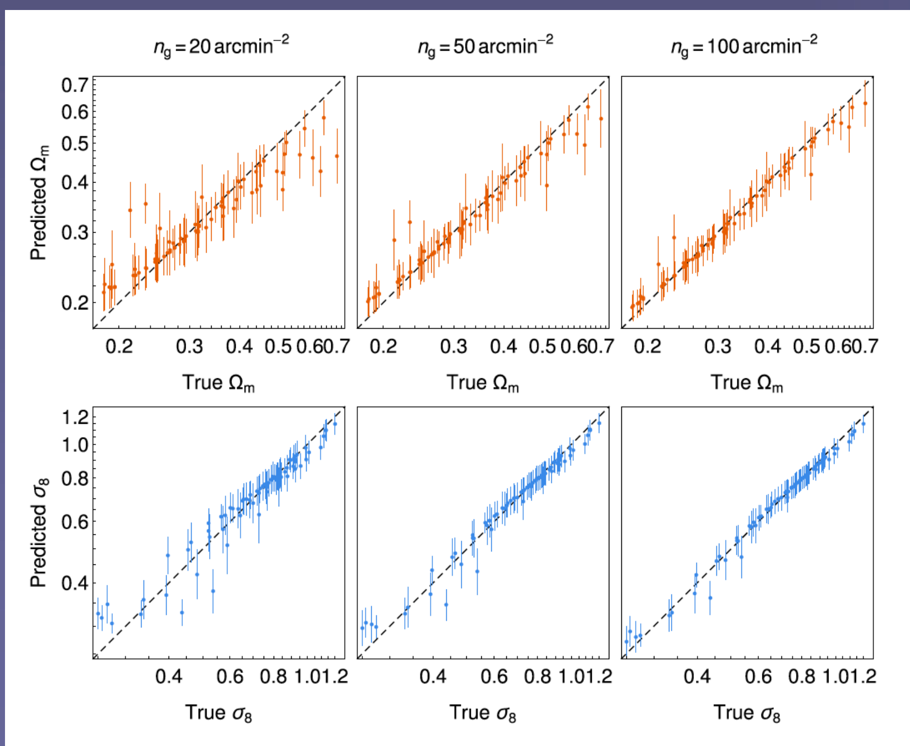


Jointly fit cosmology & baryons

Lu, ZH & Zorrilla 2022

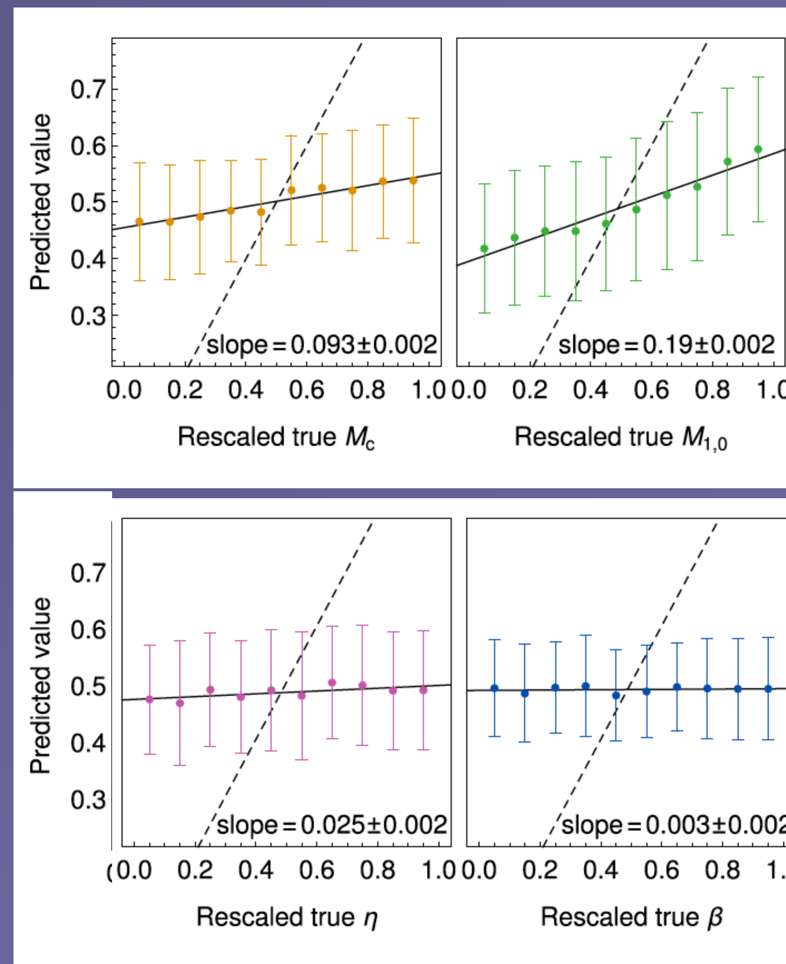
Cosmology

- Can predict parameters,
- tilt/bias (corrected in likelihood)



Baryons

- Network can learn $M_c + M_{1,0}$
- but not β or η



Baryons with machine learning

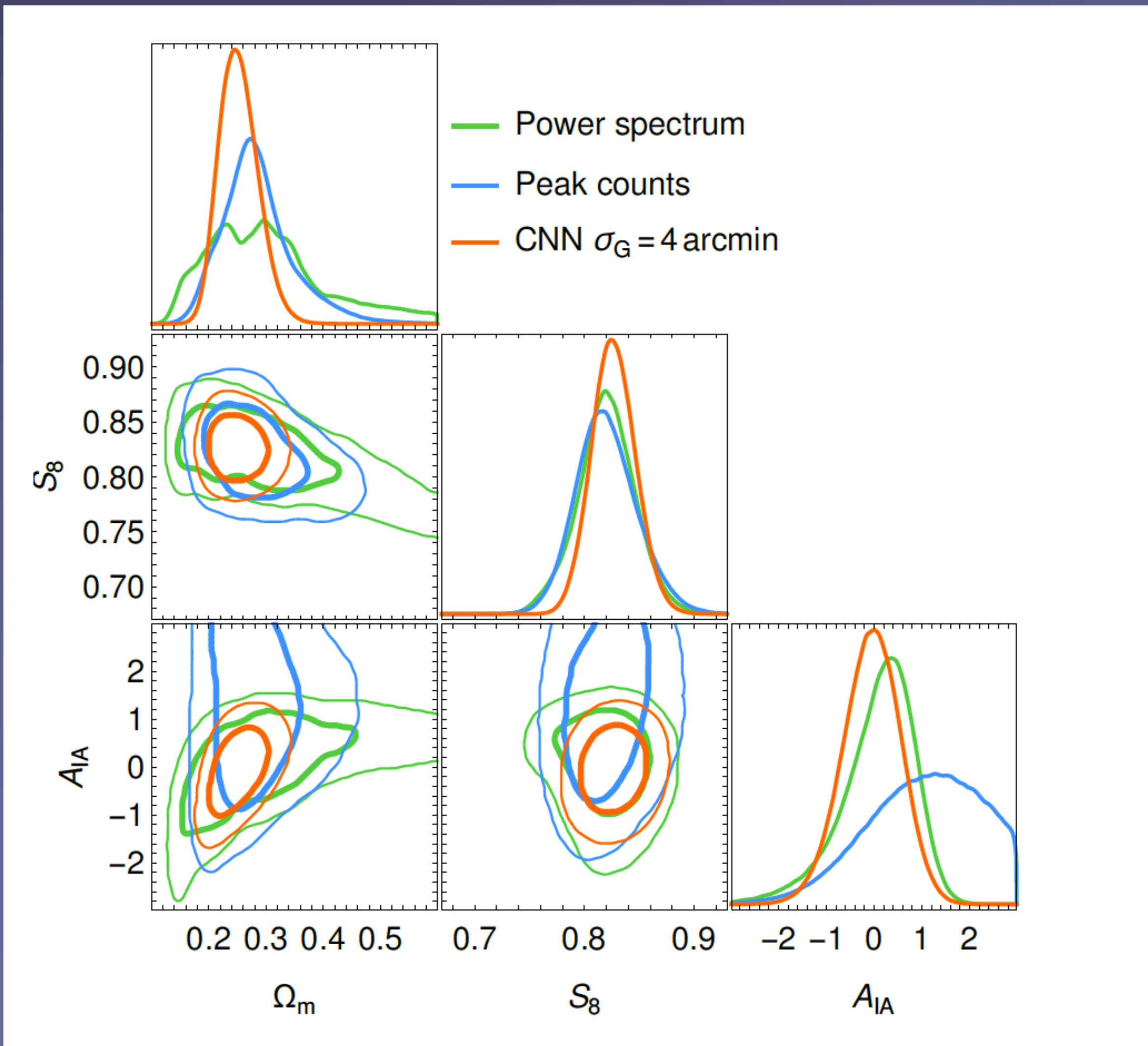
Lu, ZH & Zorrilla 2022

Methods	$\Omega_m - \sigma_8$			$M_{1,0} - \eta$		
	S_{full} ($\times 10^{-4}$)	S_{fid} ($\times 10^{-4}$)	S_{full}/S_{fid}	S_{full} ($\times 10^{-2}$)	S_{fid} ($\times 10^{-2}$)	S_{full}/S_{fid}
Power spectrum	3.45	0.93	3.71	10.4	3.6	2.88
Peak counts	5.89	0.94	6.28	30.6	7.3	4.16
CNN	2.08	0.44	4.70	13.0	3.7	3.48
CNN + power spectrum (L)	1.27	0.44	2.91	7.1	2.6	2.69
CNN + power spectrum (M)	1.11	0.42	2.61	6.9	2.8	2.41
CNN + power spectrum (S)	1.74	0.41	4.23	9.7	3.0	3.26
CNN + power spectrum (L, M)	1.01	0.42	2.39	5.2	2.3	2.24
CNN + power spectrum (full)	0.96	0.40	2.41	4.6	2.1	2.24

- CNN improves over peaks/power spectrum by factor of ~1.8.
- With baryons, peaks degrade the most
- CNN was unable to learn the medium and large-scale power spectrum – so their combination mitigates degradation
- For baryon parameters, CNN comparable to power spectrum but independent

Fitting HSC data with CNN

Lu, ZH & Li, in prep



Include:

1. Photo-z errors
2. Baryon effects
3. Intrinsic alignments

Unsolved problems

- How much cosmological information is contained, in principle, in a (perfect) lensing map?

→ **At least an order of magnitude more than in power spectrum**

- How well can we constrain background cosmology, in practice, from observed lensing data?

Unsolved problems

- How much cosmological information is contained, in principle, in a (perfect) lensing map?

→ **At least an order of magnitude more than in power spectrum**

- How well can we constrain background cosmology, in practice, from observed lensing data?

→ **Factor of ~two better than power spectrum**

Conclusions

- Beyond-Gaussian info: Peaks constrain Ω_m, σ_8 tighter than the power spectrum – errors improve by up to a factor of ~2

- Baryons: can be modeled with a flexible parameterized model, generally degrade constraints by a factor of a few.

- Neural networks: can improve constraints by a factor of >10 in perfect simulations, and by factor of ~2 in presence of noise and/or baryons

- Just the beginning: $10^{7\text{-}8}$ → few x 10^9 gals with LSST, Euclid, Roman

The End