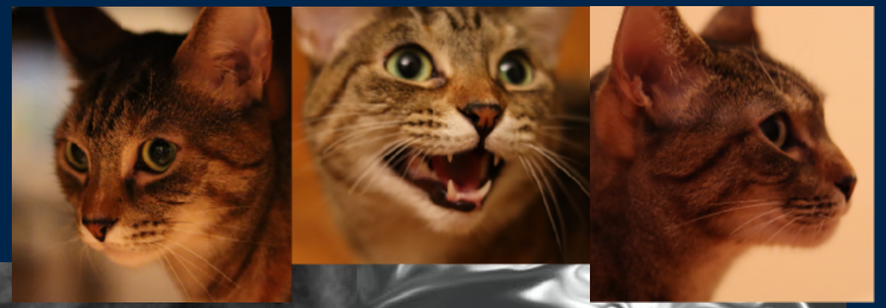
The scattering transform: a new vocabulary for textures

Sihao Cheng (程思浩)
Johns Hopkins University

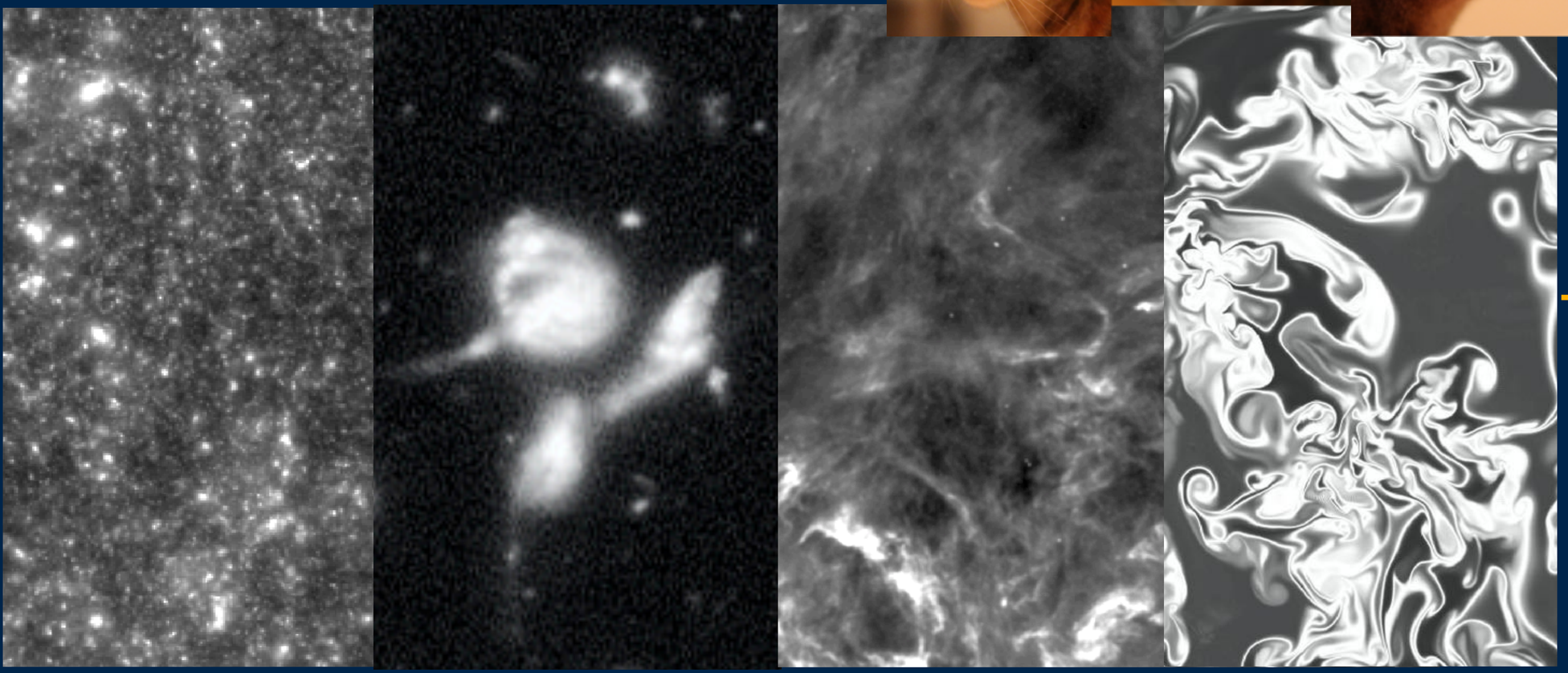
TsingHua University
Mar 16th, 2021

with Brice Ménard, Yuan-Sen Ting, & Joan Bruna

How do we extract information from a field?



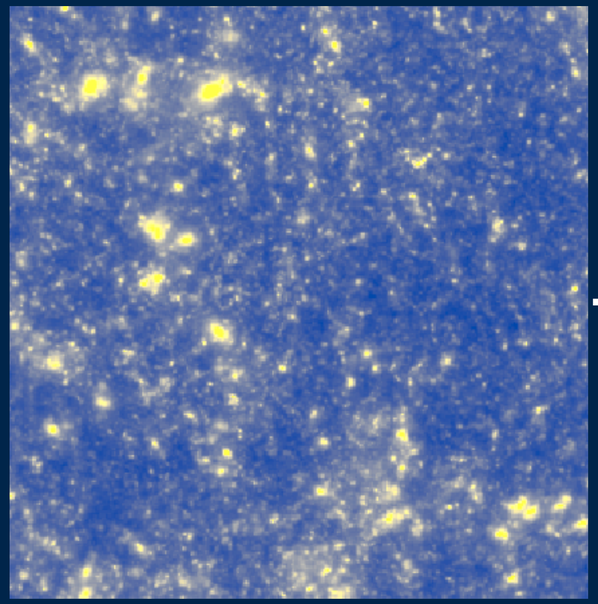
daily-life information



physical parameters

1. compression
2. better metric

How do we characterize a field?



a description
using a vocabulary

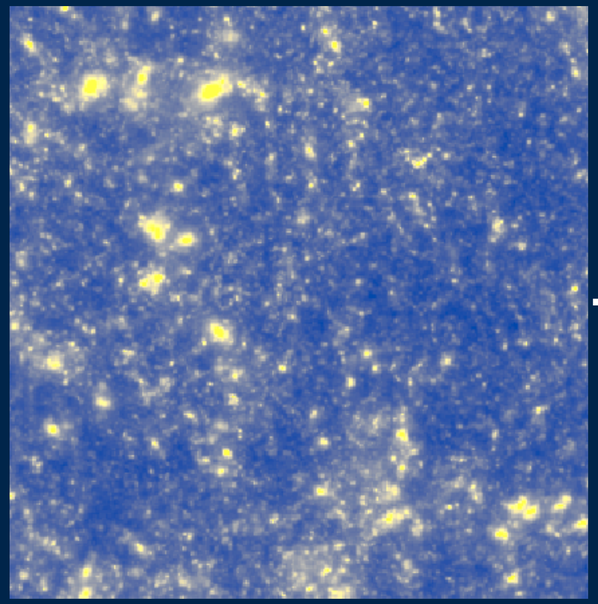
physical
parameters

data

extract patterns

information

How do we characterize a field?



power spectrum

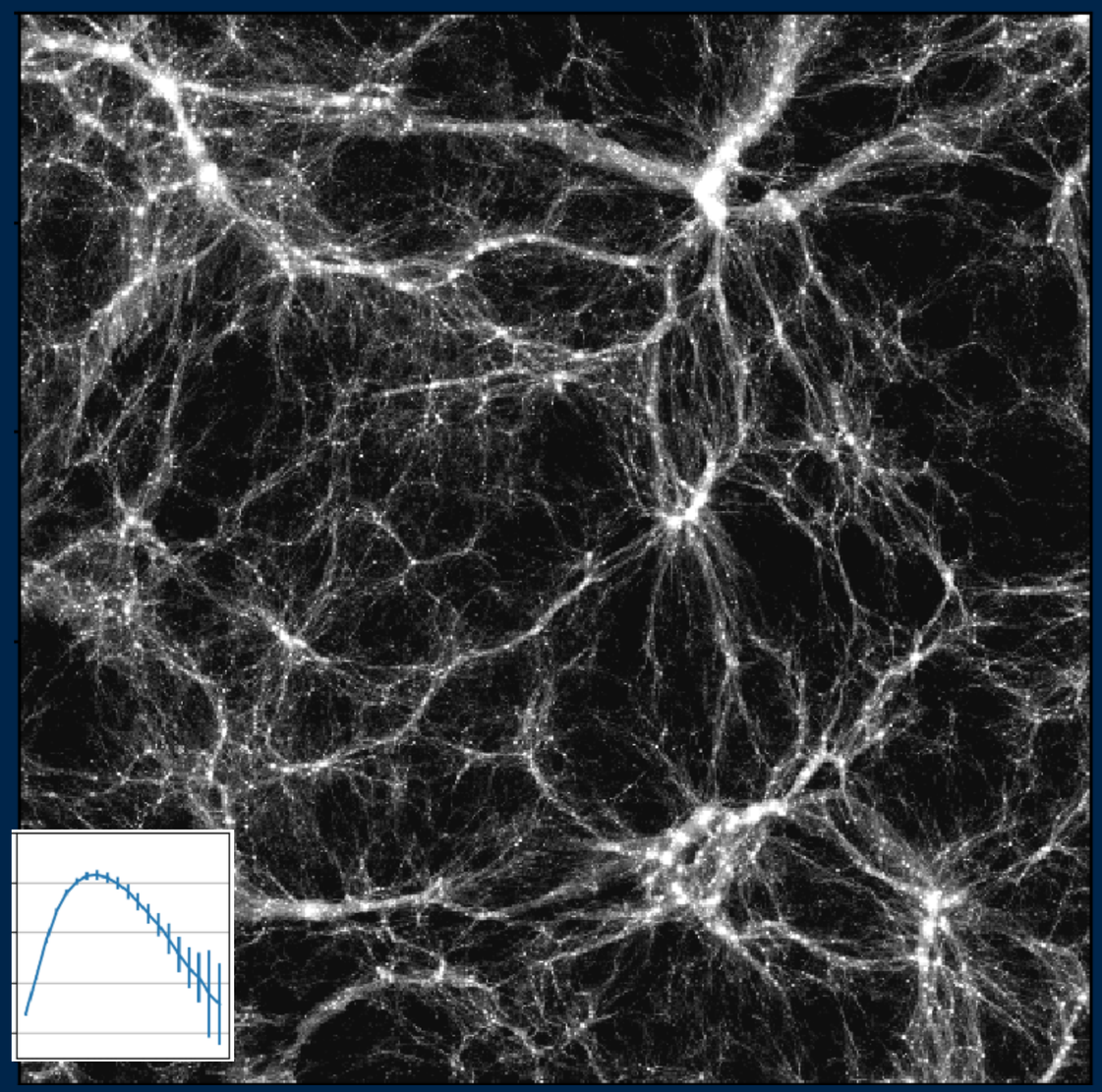
$$P(k) = \langle \delta_k \delta_{-k} \rangle$$

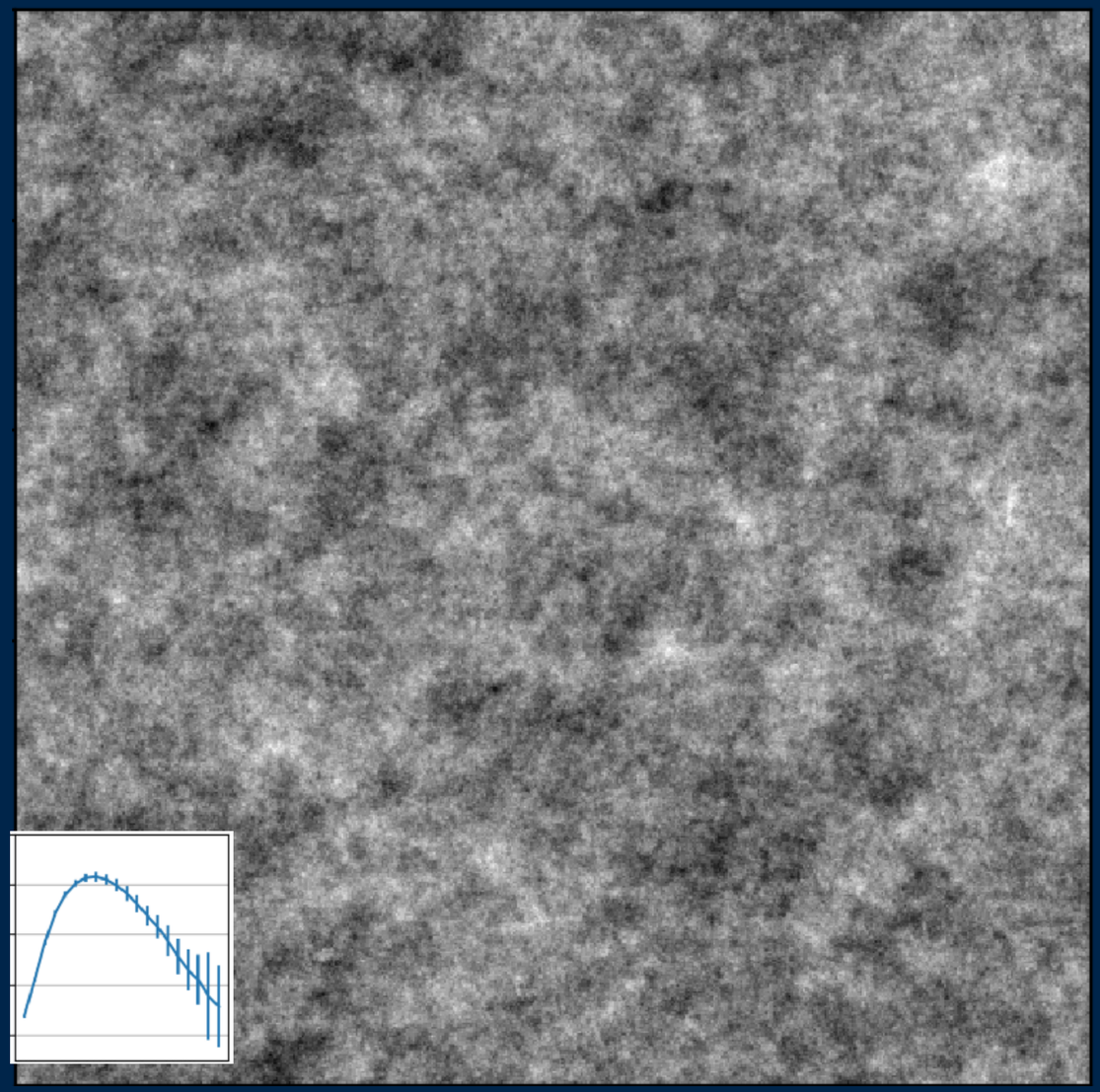
physical parameters

data

extract patterns

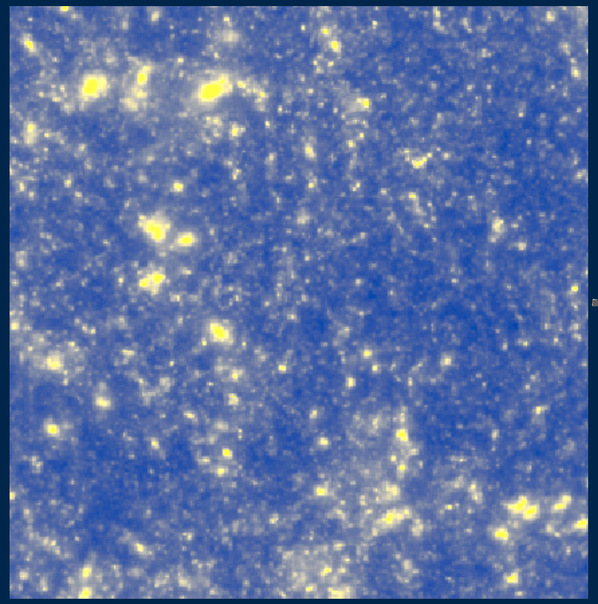
information





We need to go beyond $P(k)$ to capture non-Gaussianity.

How do we characterize a field?



$P(k)$ plus
classic statistics

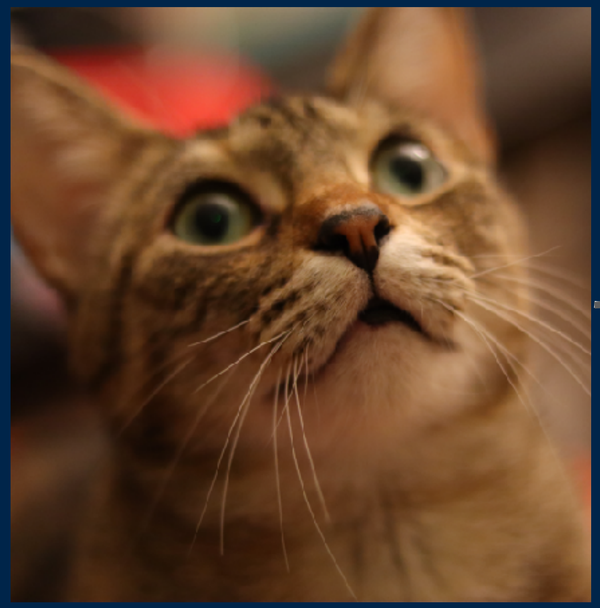
$\langle \delta_1 \delta_2 \ldots \delta_n \rangle$

CNN

physical
parameters

How do we characterize a field?

1. lose information (Carron 2011)
2. too many configurations



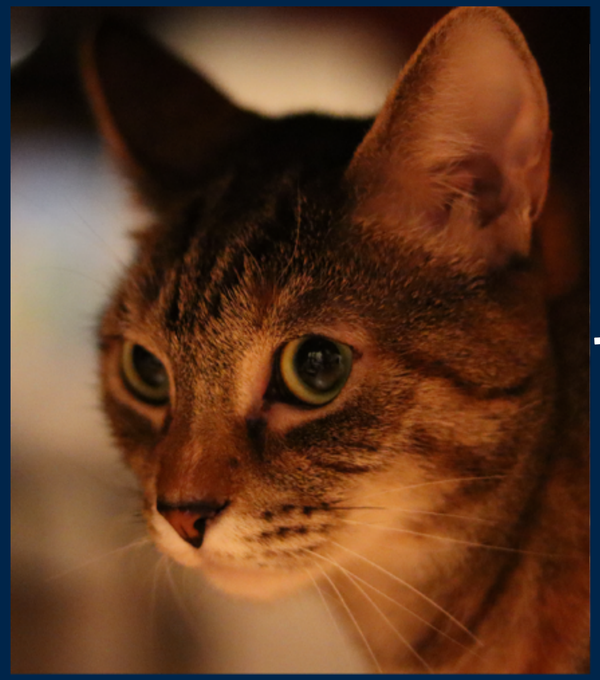
$P(k)$ plus
classic statistics

$\langle \delta_1 \delta_2 \ldots \delta_n \rangle$

CNN

physical
parameters

How do we characterize a field?



$P(k)$ plus
classic statistics
$\langle \delta_1 \delta_2 \ldots \delta_n \rangle$

CNN

physical
parameters

convolutional neural network

CNN learns & extracts patterns.

Conv. layer → Conv. layer → … → fully connected →

my cat : 99%
your cat : 1%

data — extract patterns — information

convolutional neural network

CNN learns & extracts patterns.

1. convolution
2. nonlinearity
3. average
4. iteration

kernels learned by AlexNet

Conv. layer

…

fully connected

my cat : 99%
your cat : 1%

data

extract patterns

information

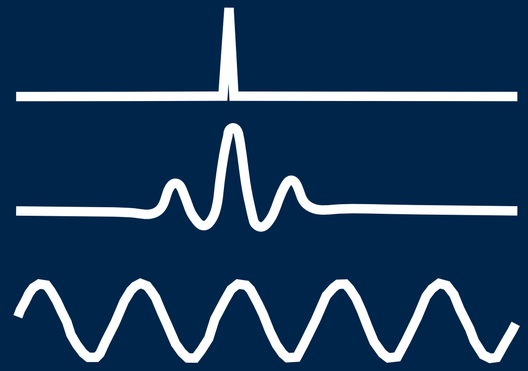
What are wavelets?

band-pass filters

delta function

wavelet

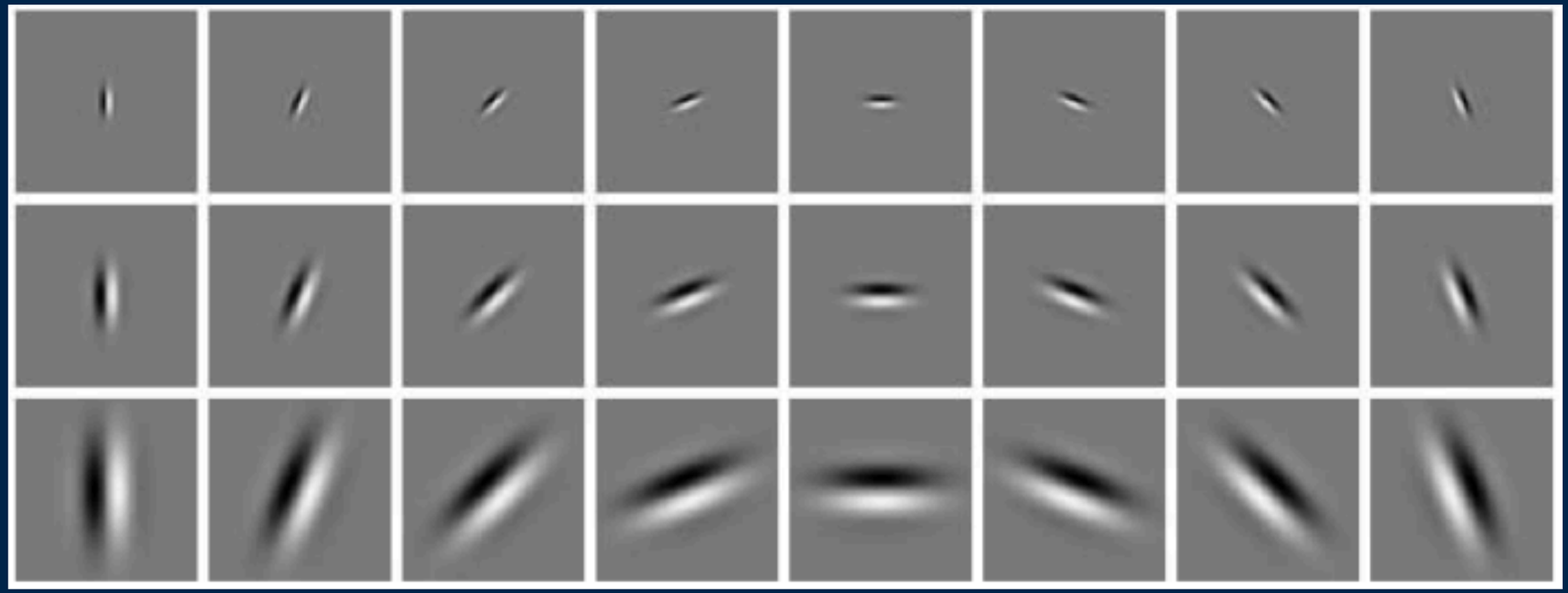
sine function



$\psi^{j,l}$

j: scales

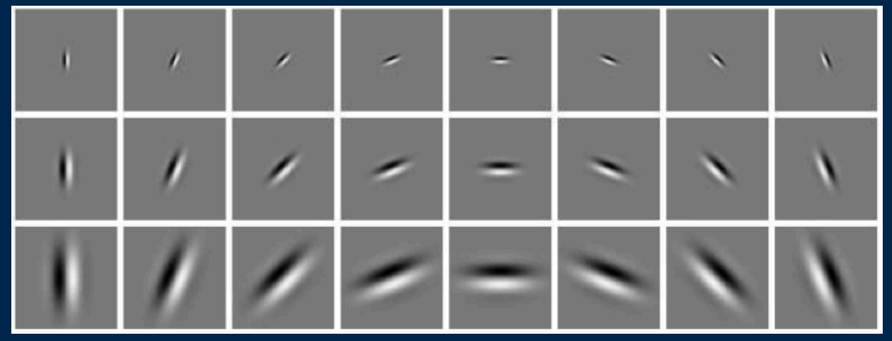
l: directions



receptive fields of mammal vision

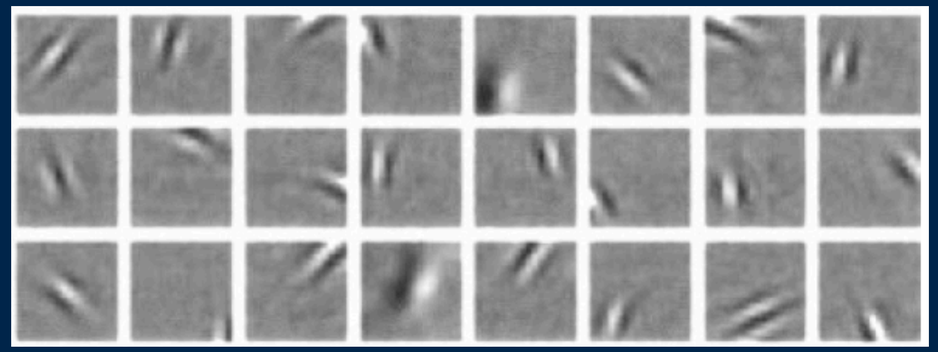
(Hubel & Wiesel 1968)

close to Gabor wavelets



sparse representation of natural images

(Olshausen & Field 1996)



kernels learned in AlexNet

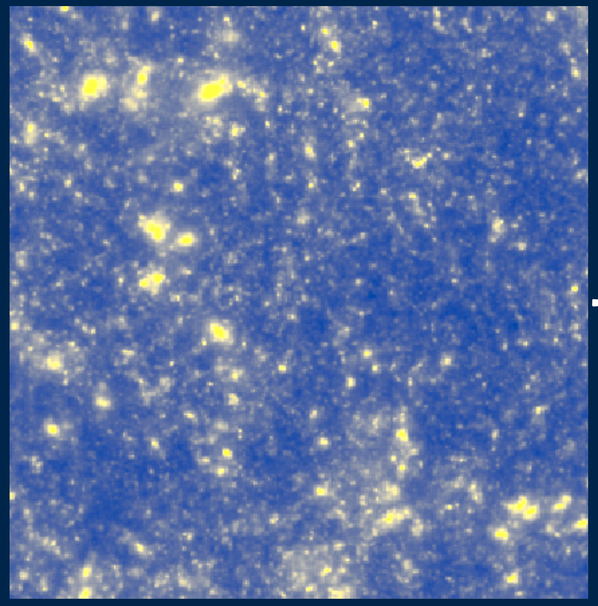
(Krizhevsky, Sutskever, & Hinton 2012)

How do we characterize a field?

limited information

but deterministic, controllable

$P(k)$ plus
classic statistics



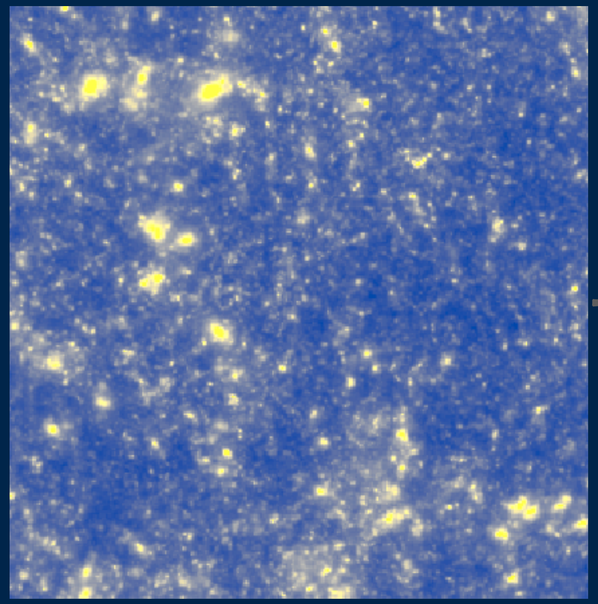
physical
parameters

CNN

wavelet, nonlinearity, iteration

but need learning

How do we characterize a field?

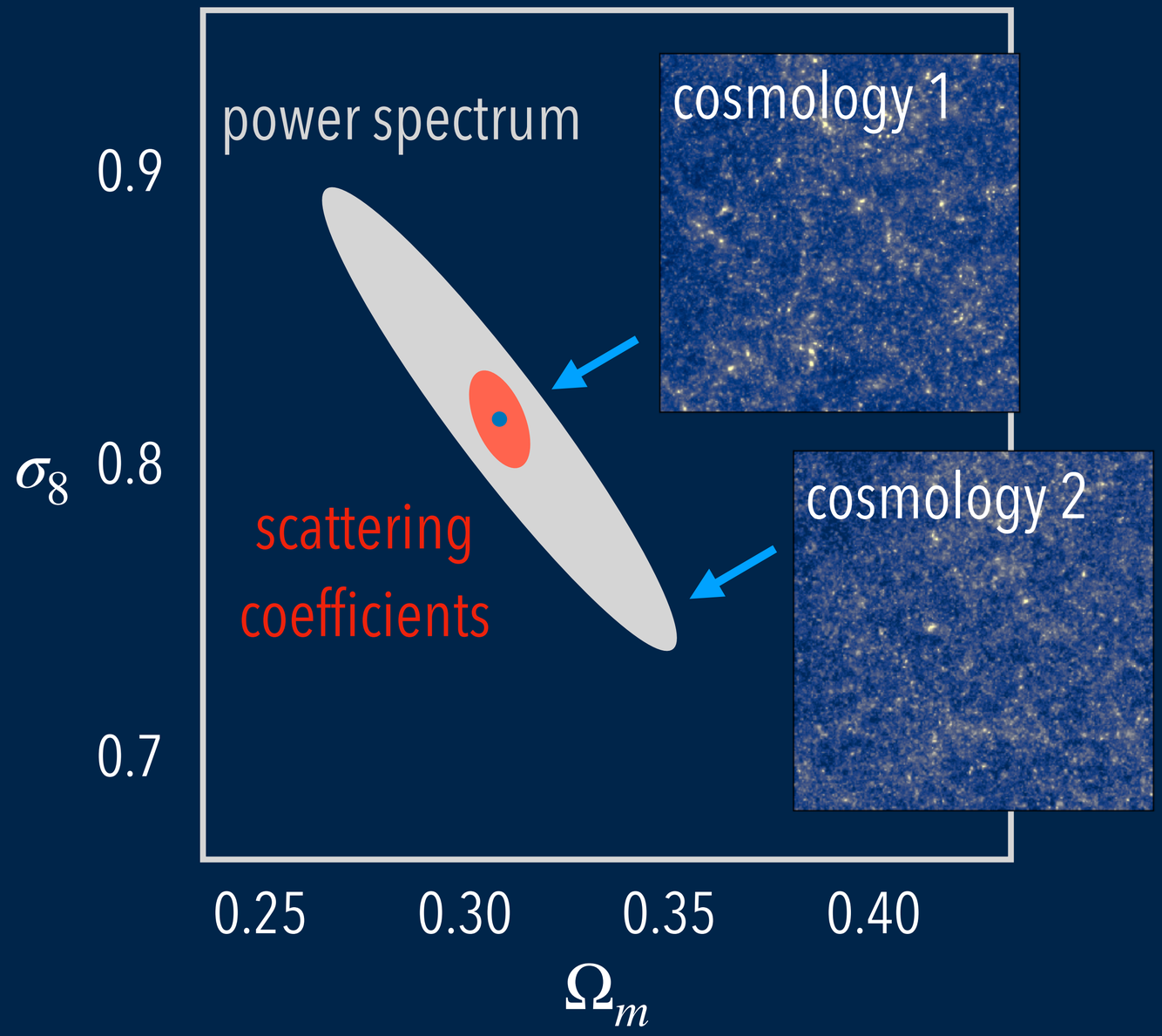


$P(k)$ plus
classic statistics

scattering transform
wavelets, nonlinearity, iteration, average
(Mallat 2012)

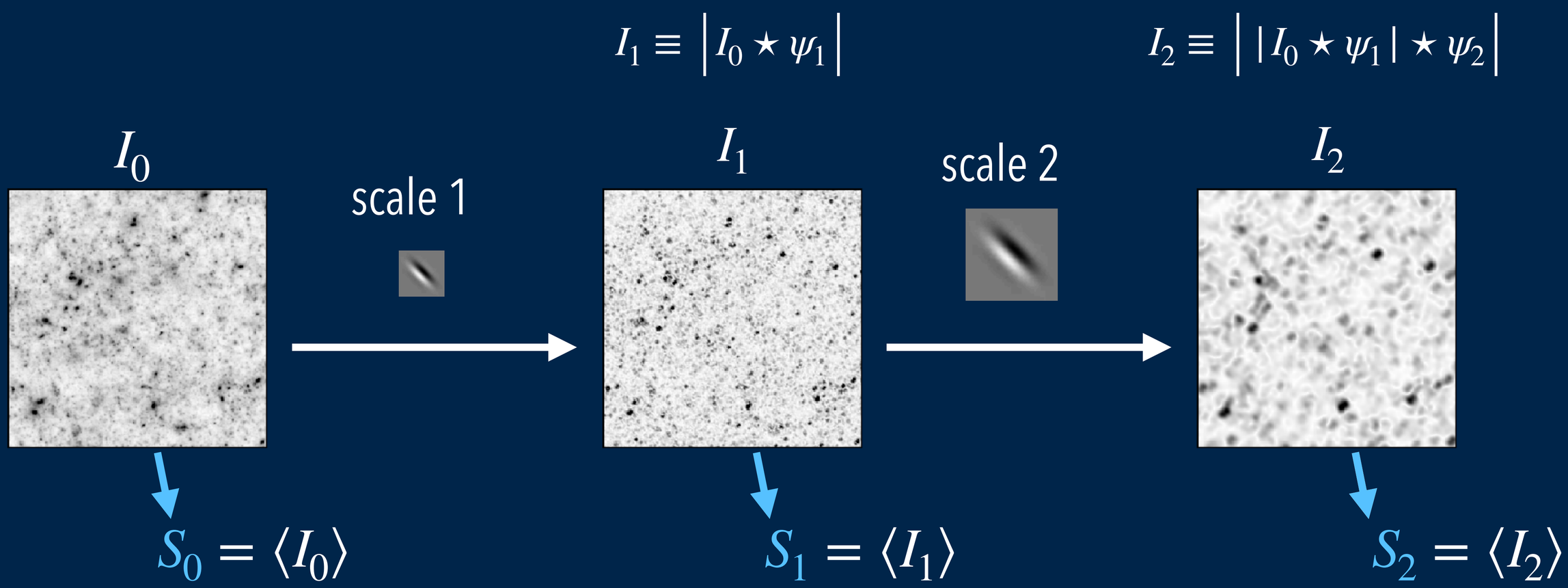
CNN

physical
parameters


power spectrum
cosmology 1
cosmology 2
scattering
coefficients
0.9
0.8
0.7
σ_8
0.25
0.30
0.35
0.40
Ω_m

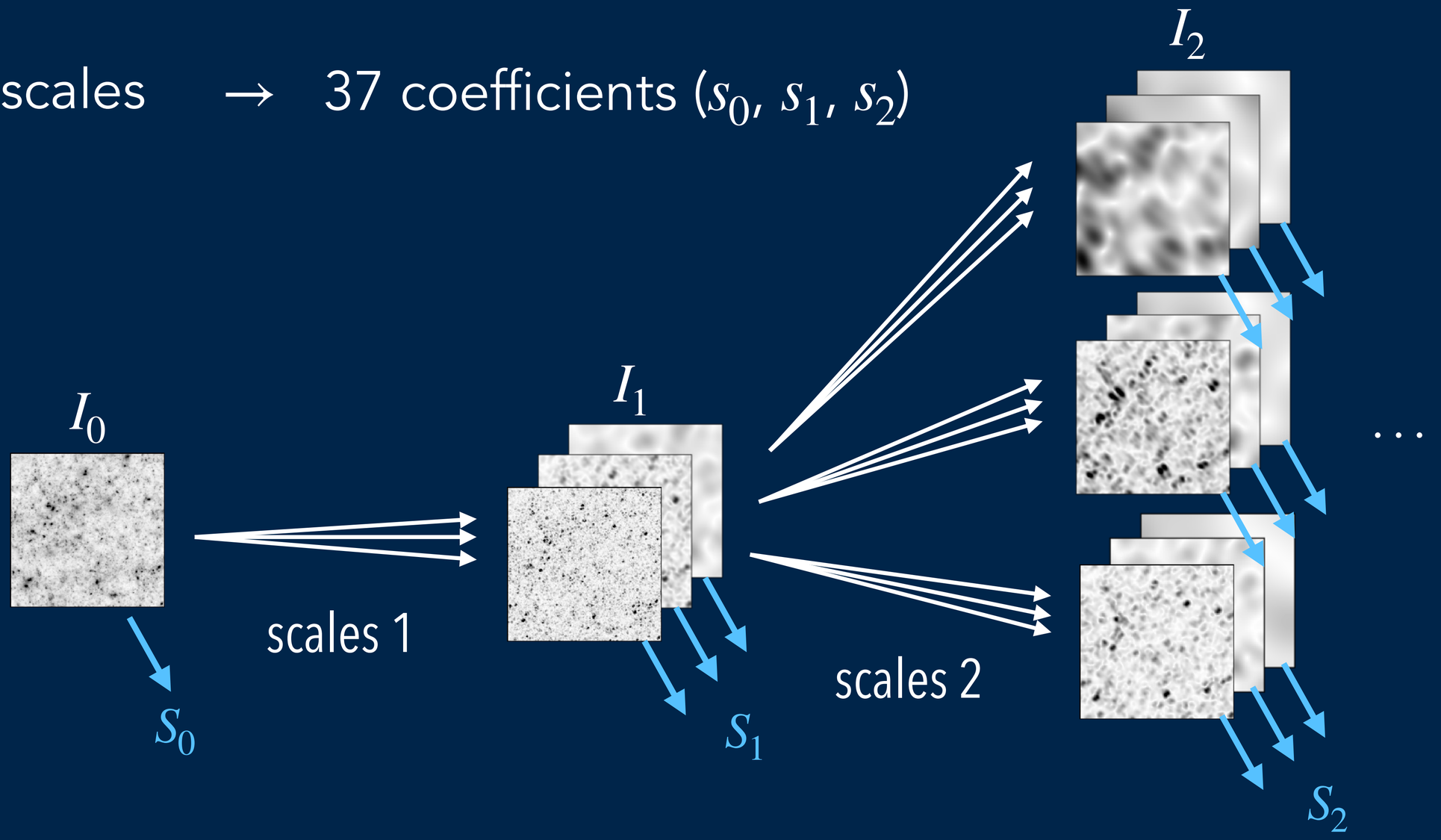
scattering transform

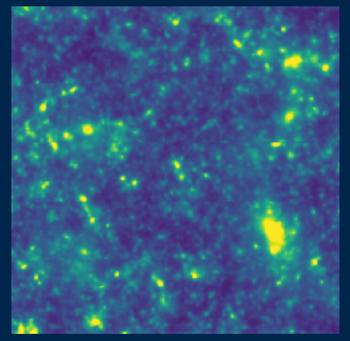
wavelets, modulus, iteration, average (Mallat 2012)



number of coefficients

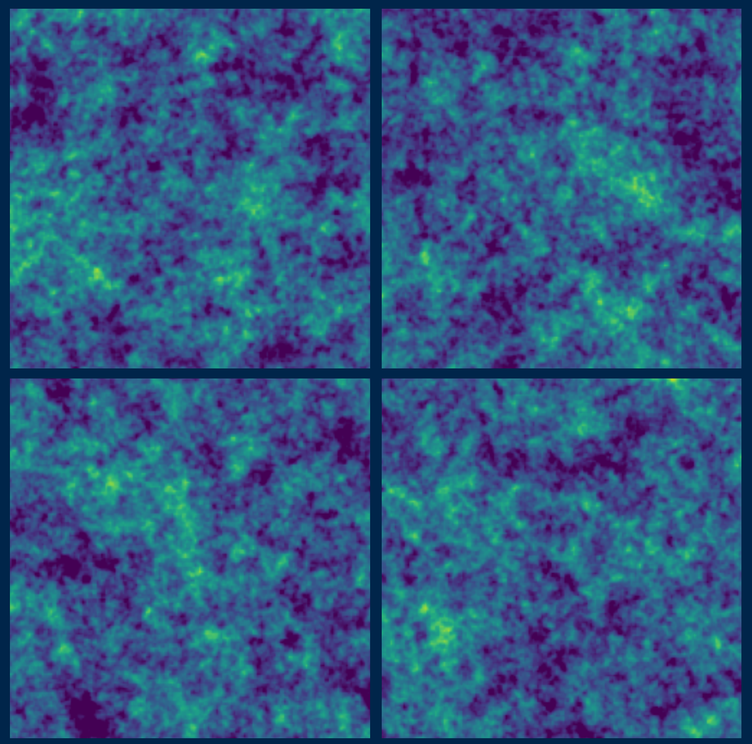
e.g., 8 scales → 37 coefficients (s_0, s_1, s_2)



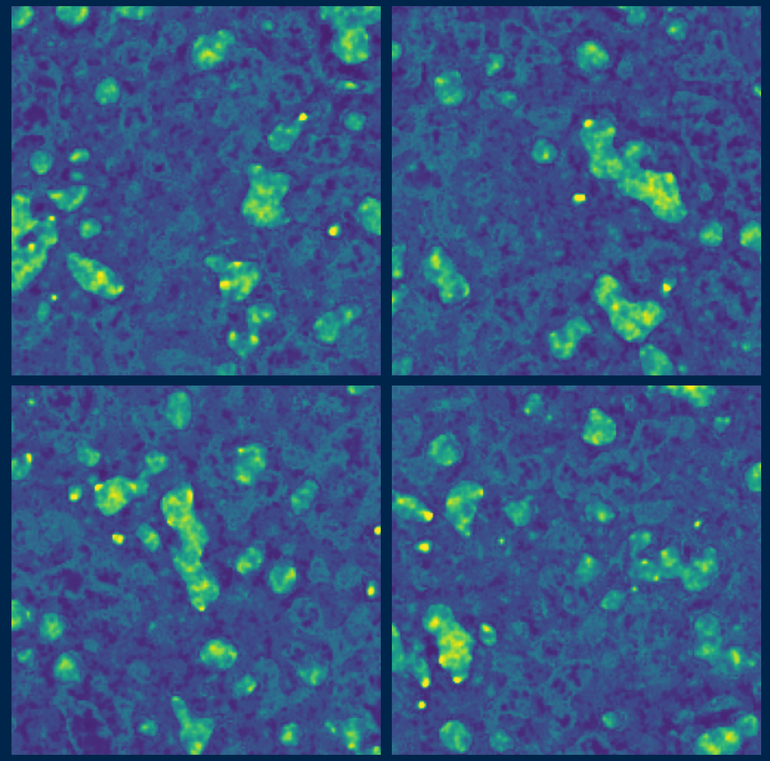


lensing map

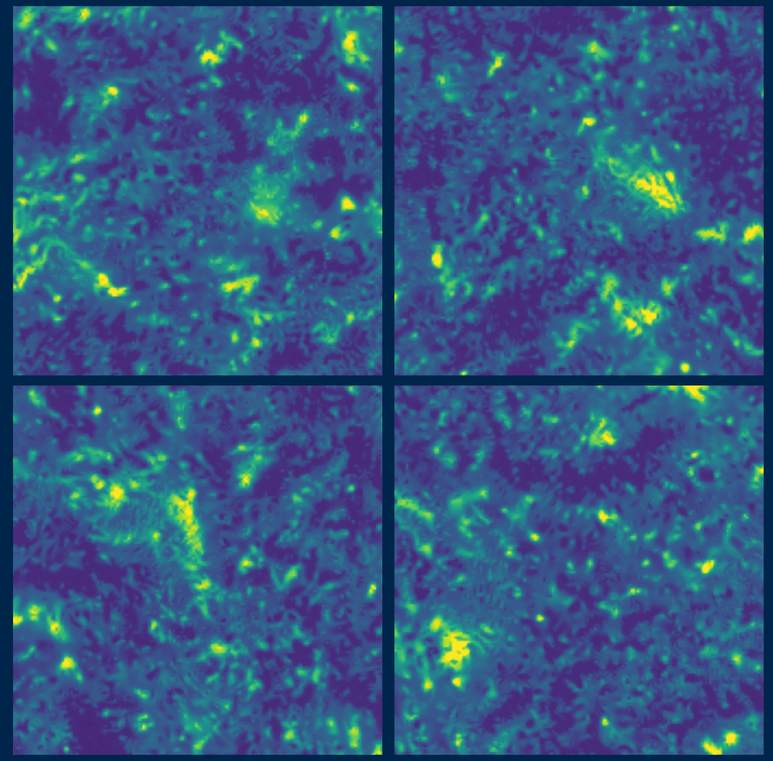
with P(l)



with B(l_1, l_2, l_3) and P(l)

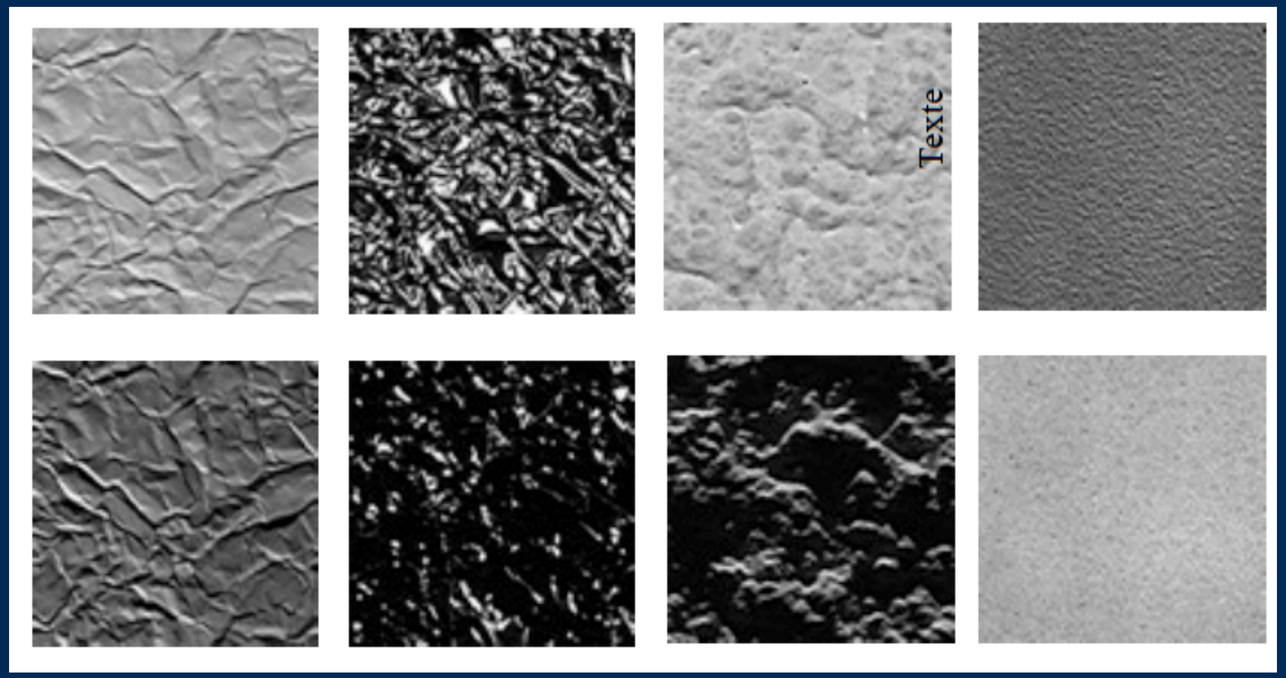


with S_1, S_2



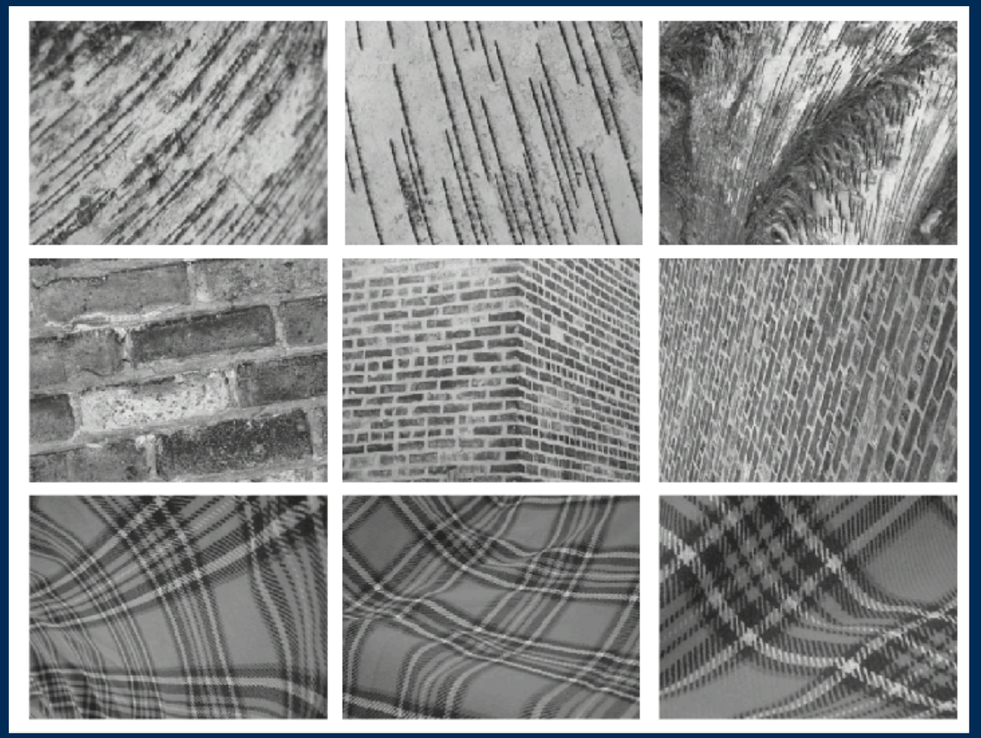
scattering transform in computer vision

CUREt database



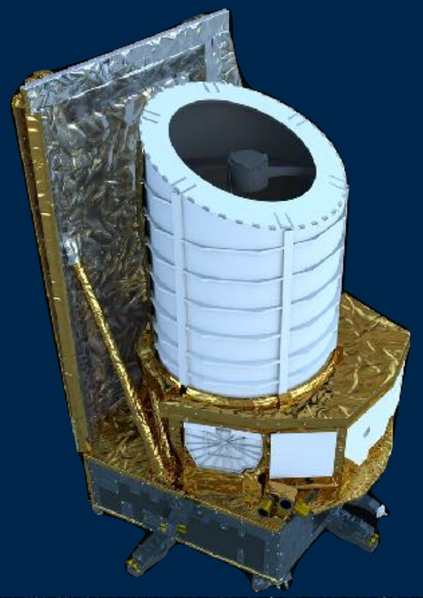
Bruna & Mallat 2013

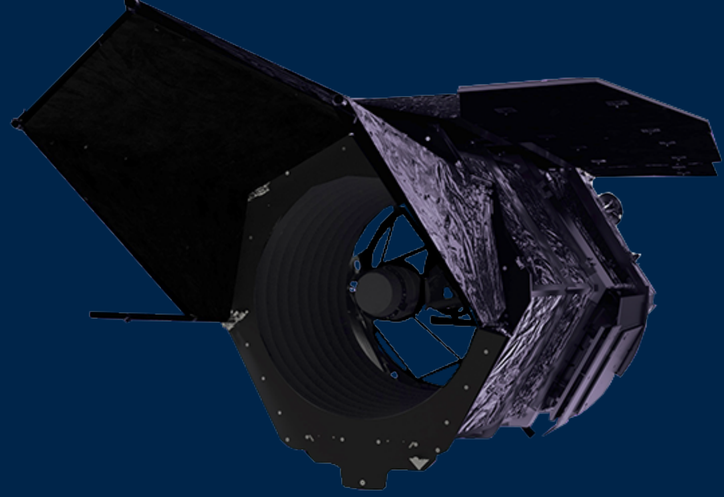
UIUC database



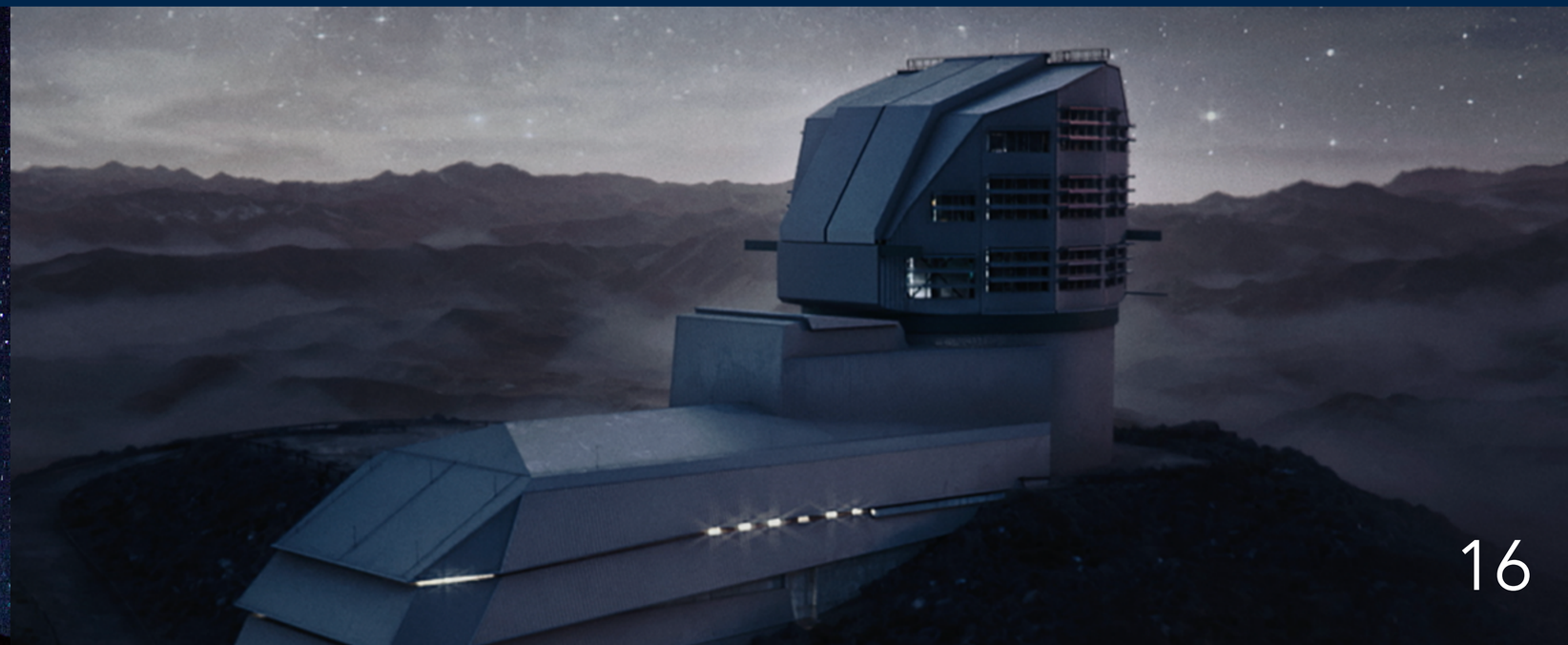
Sifre & Mallat 2013

weak lensing cosmology

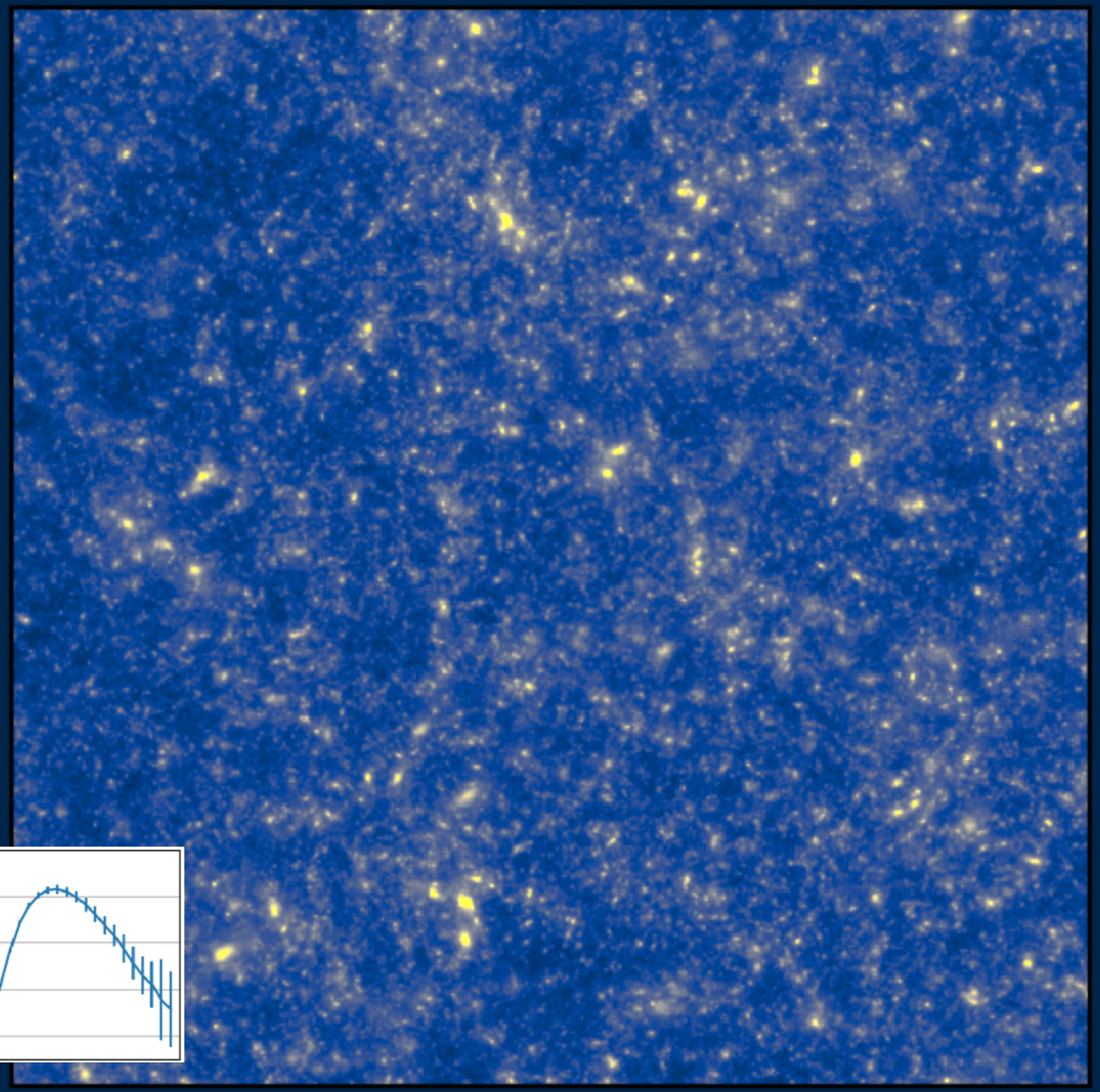


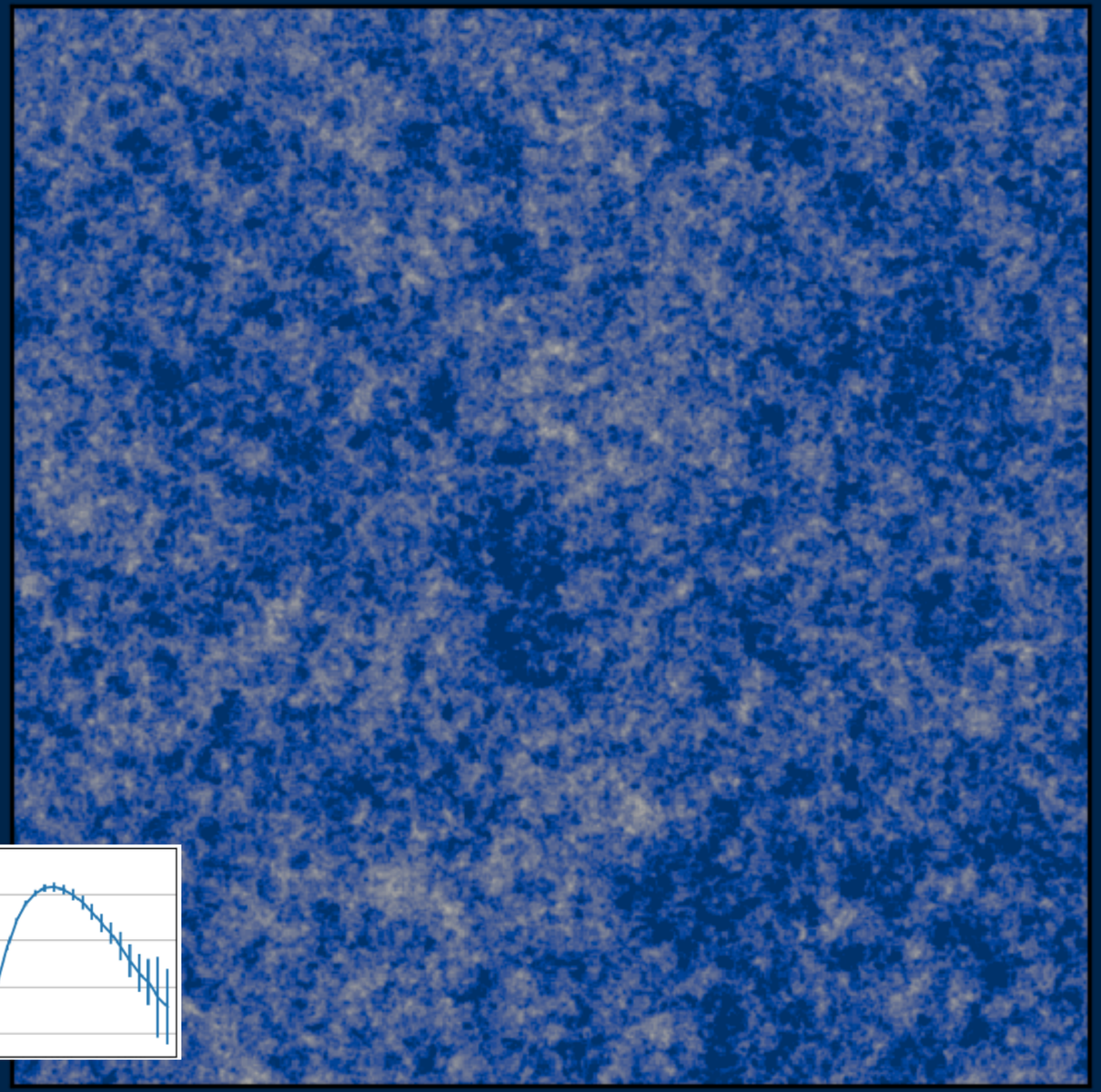


raw data → galaxy catalog → mass map → $\sigma_8, \Omega_m, w, M_\nu$



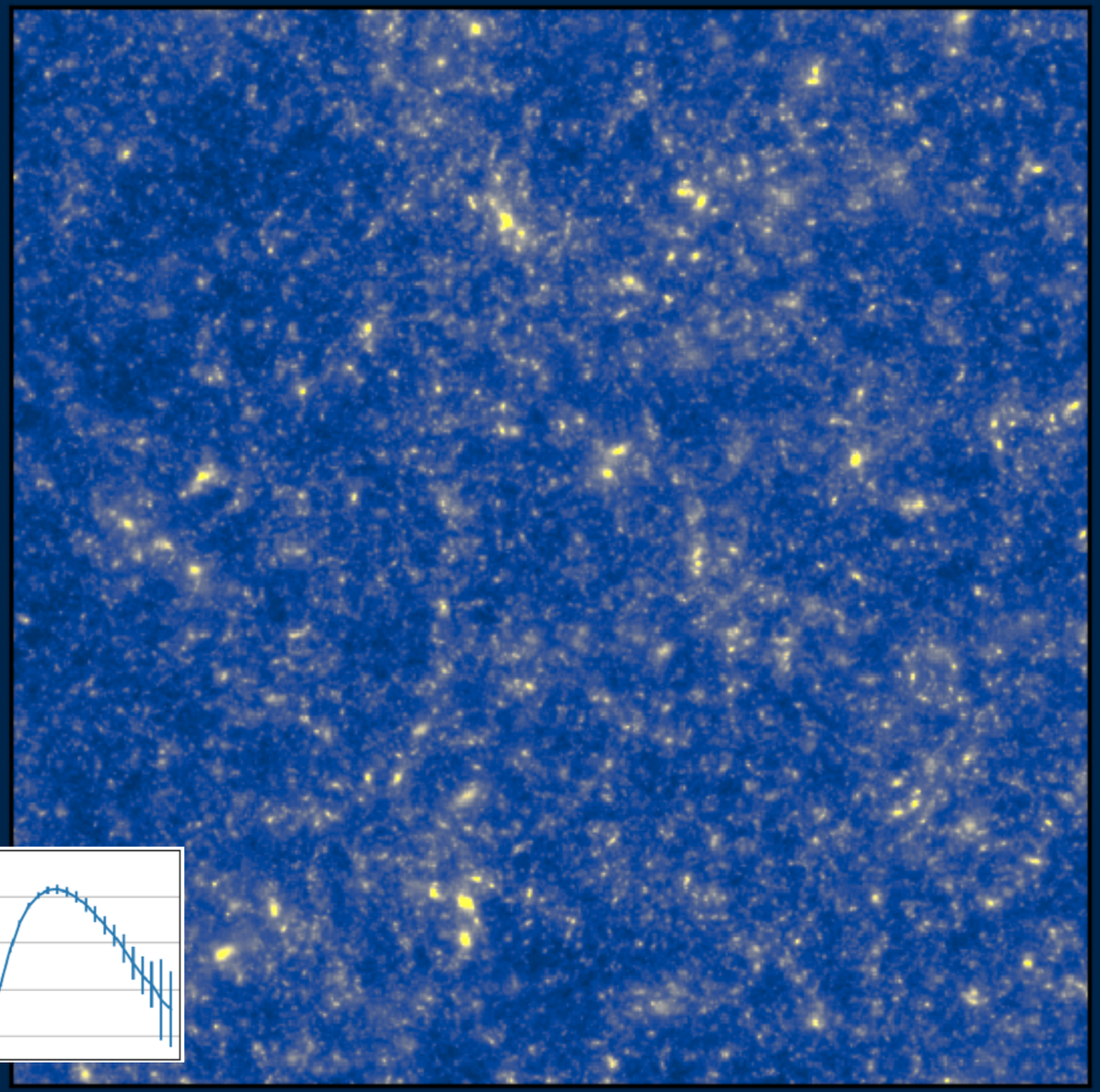
simulated mass maps (from Columbia lensing group)

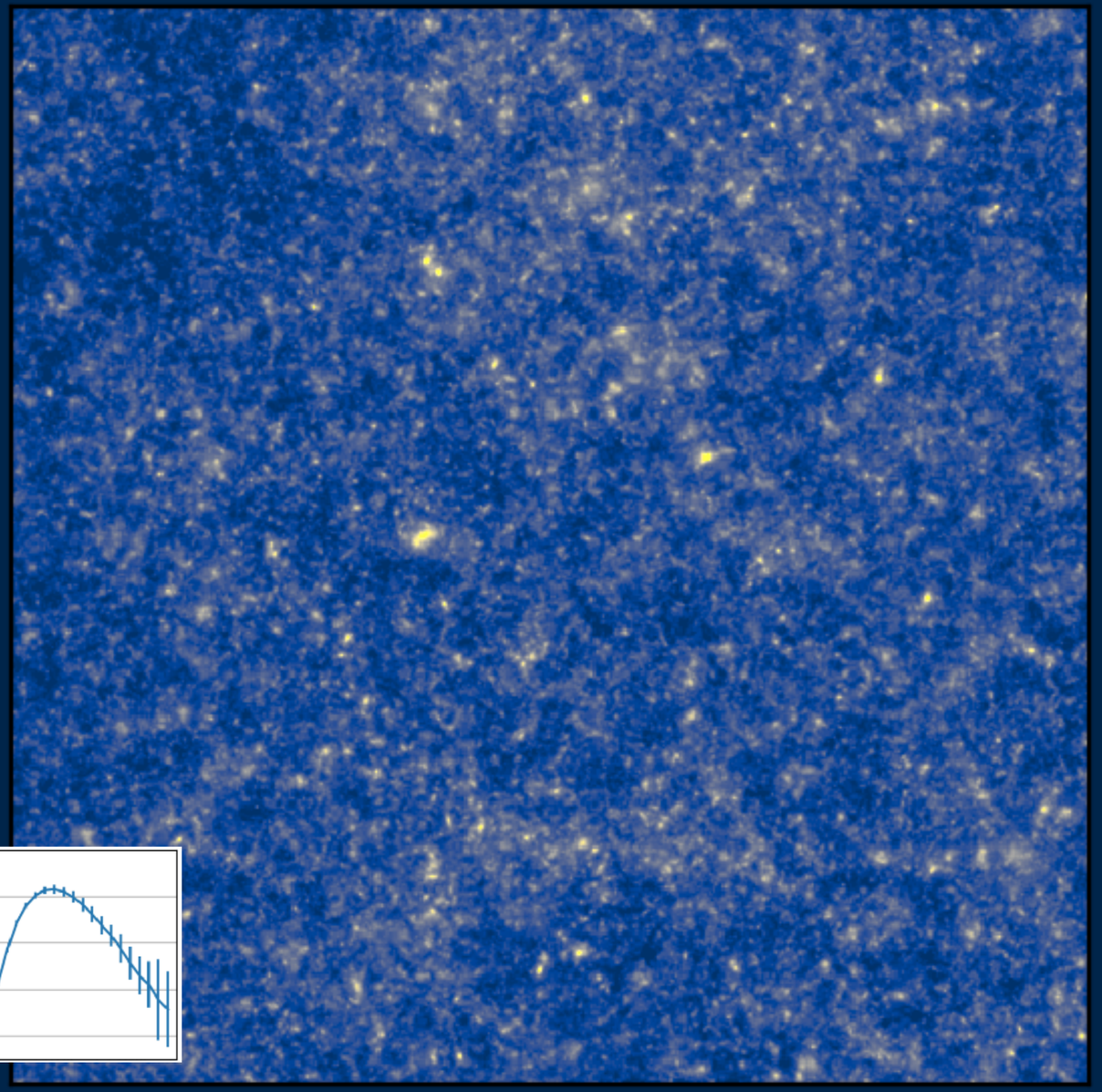




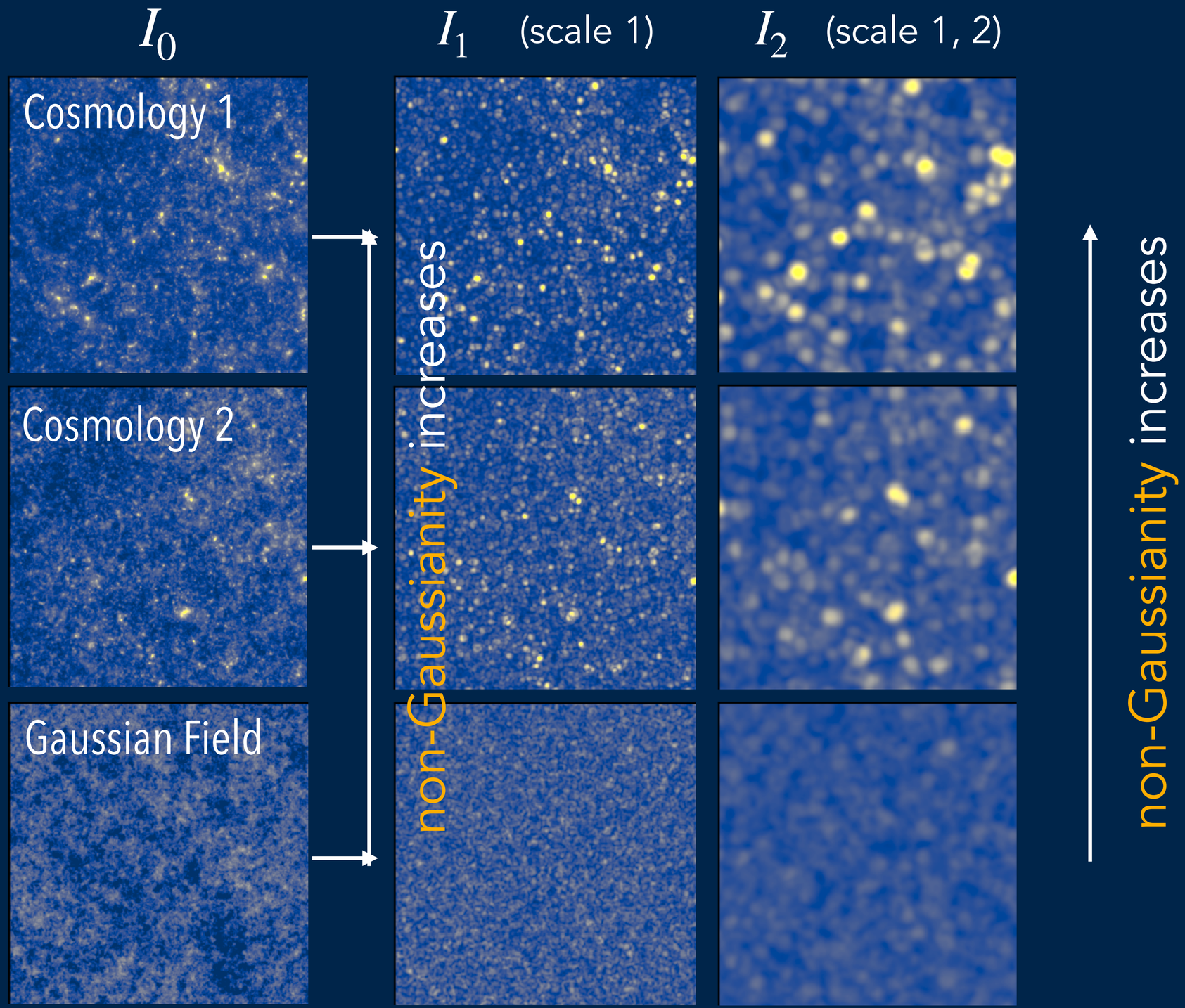
Matilla Zorrilla et al. 2016, Gupta et al. 2018

simulated mass maps (from Columbia lensing group)

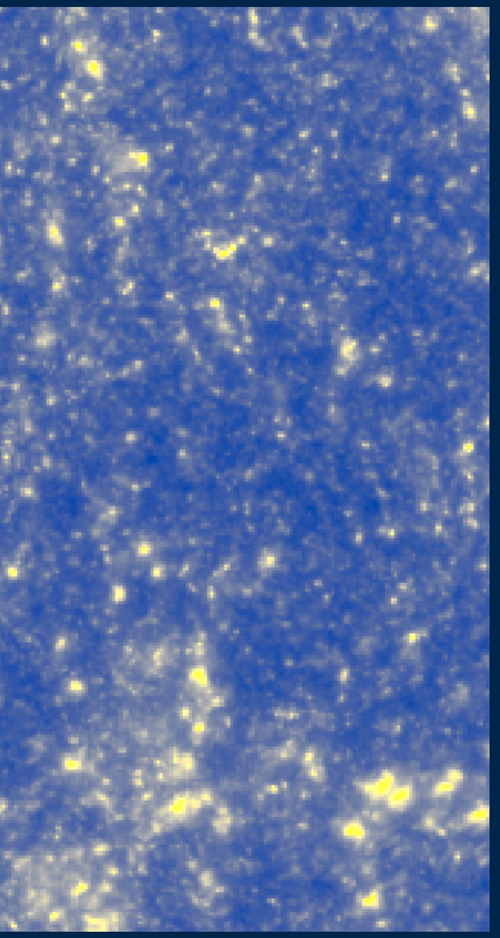




Matilla Zorrilla et al. 2016, Gupta et al. 2018


I_0
I_1 (scale 1)
I_2 (scale 1, 2)
Cosmology 1
Cosmology 2
Gaussian Field
non-Gaussianity increases
non-Gaussianity increases

weak lensing cosmology



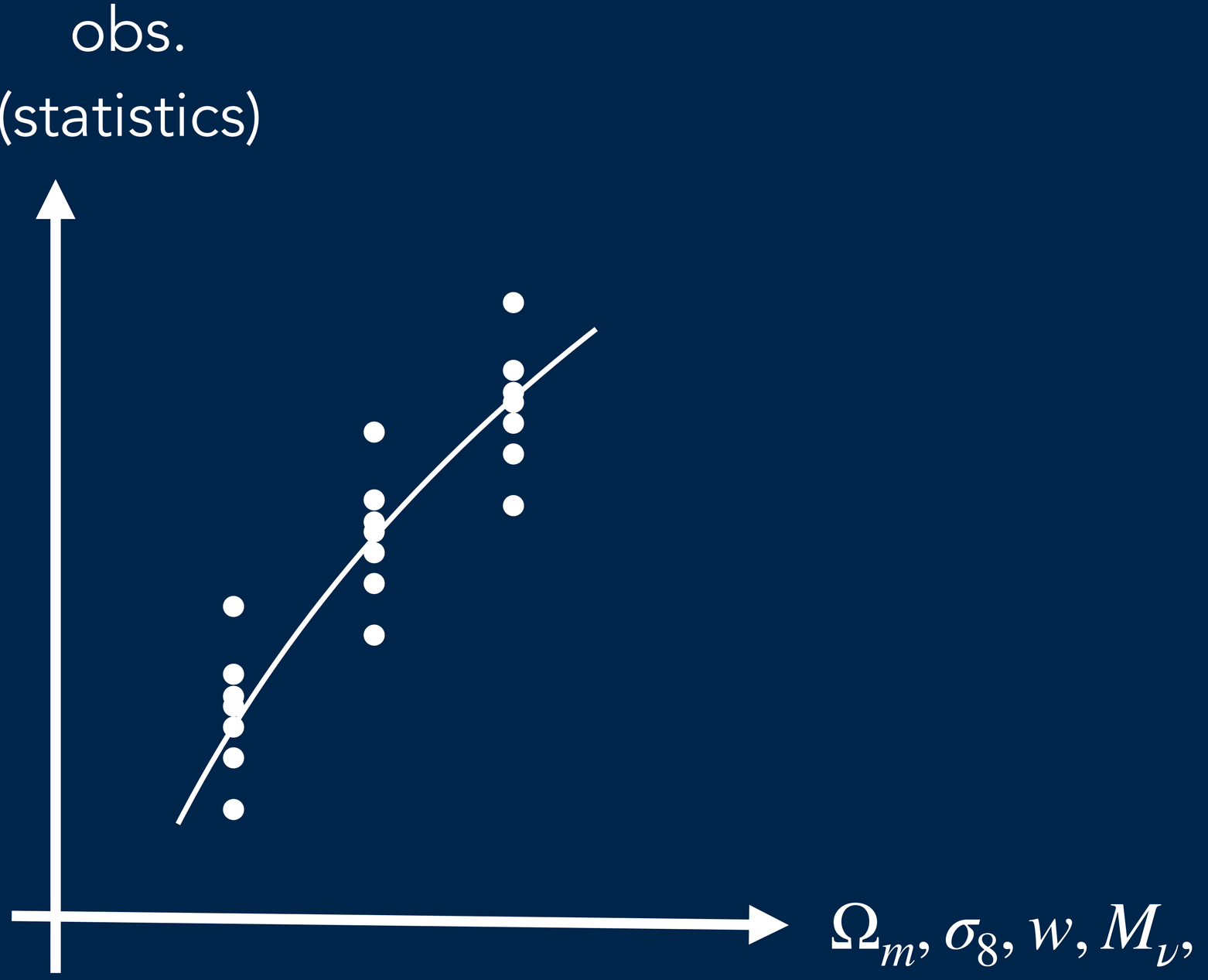
$P(k)$

scattering coefficients

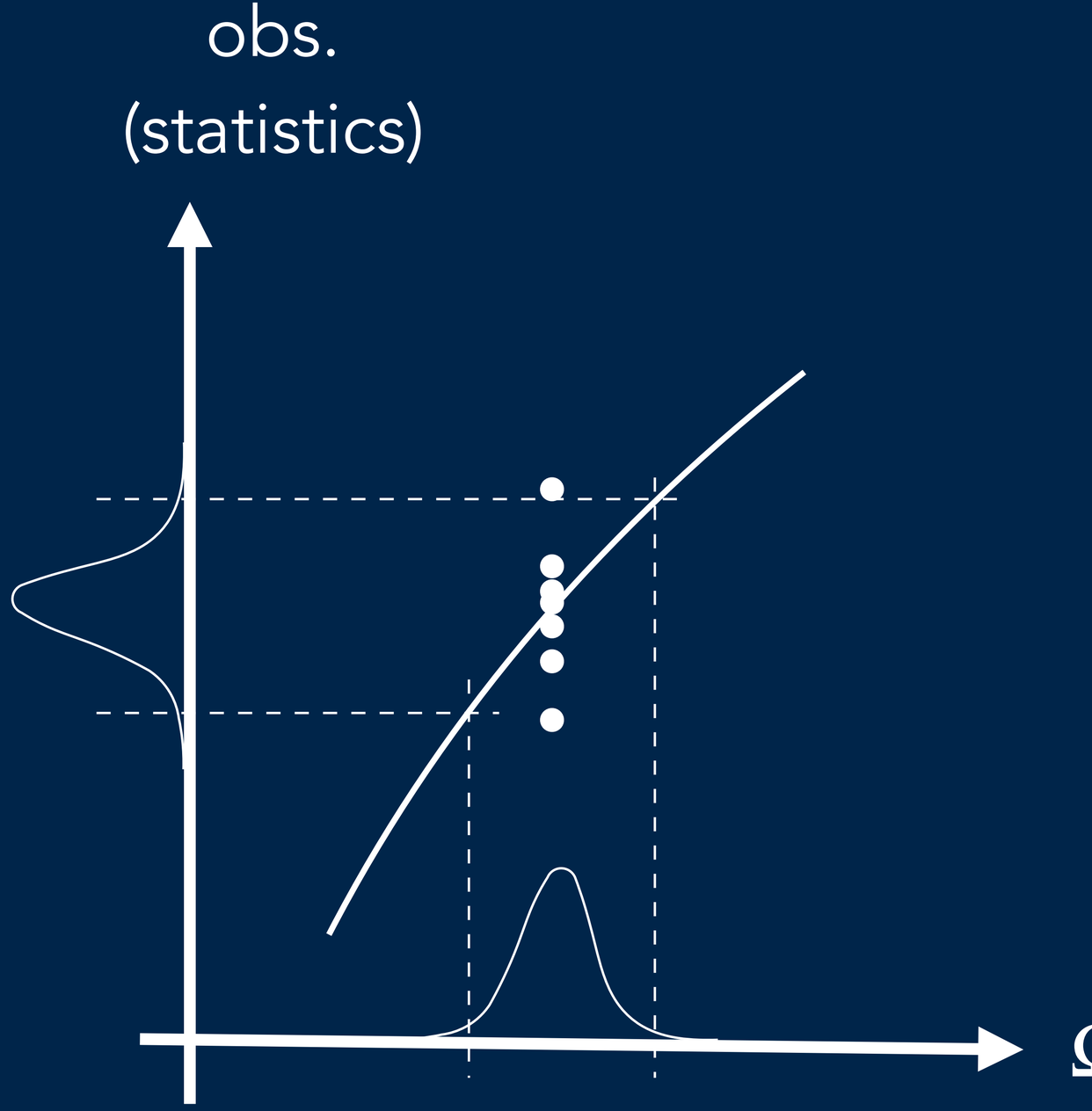
CNN

$\Omega_m, \sigma_8, w, M_\nu,$
etc.

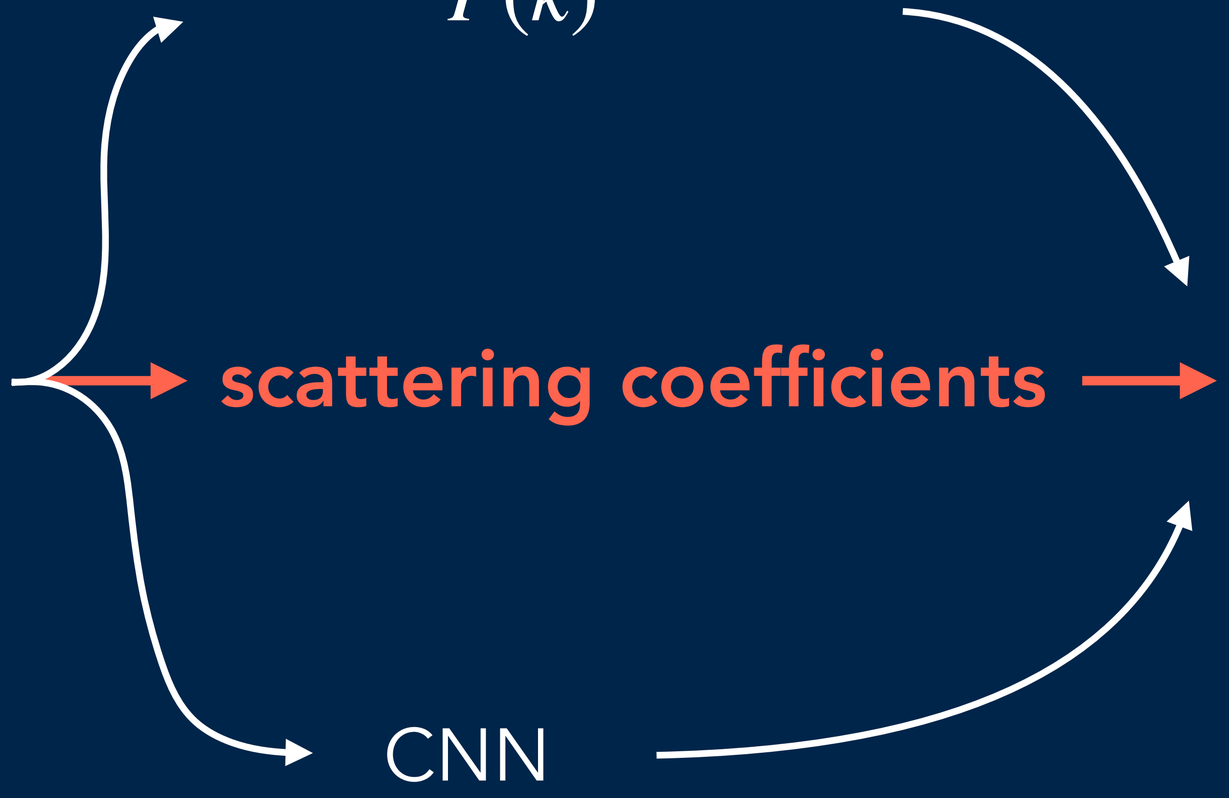
Fisher forecast



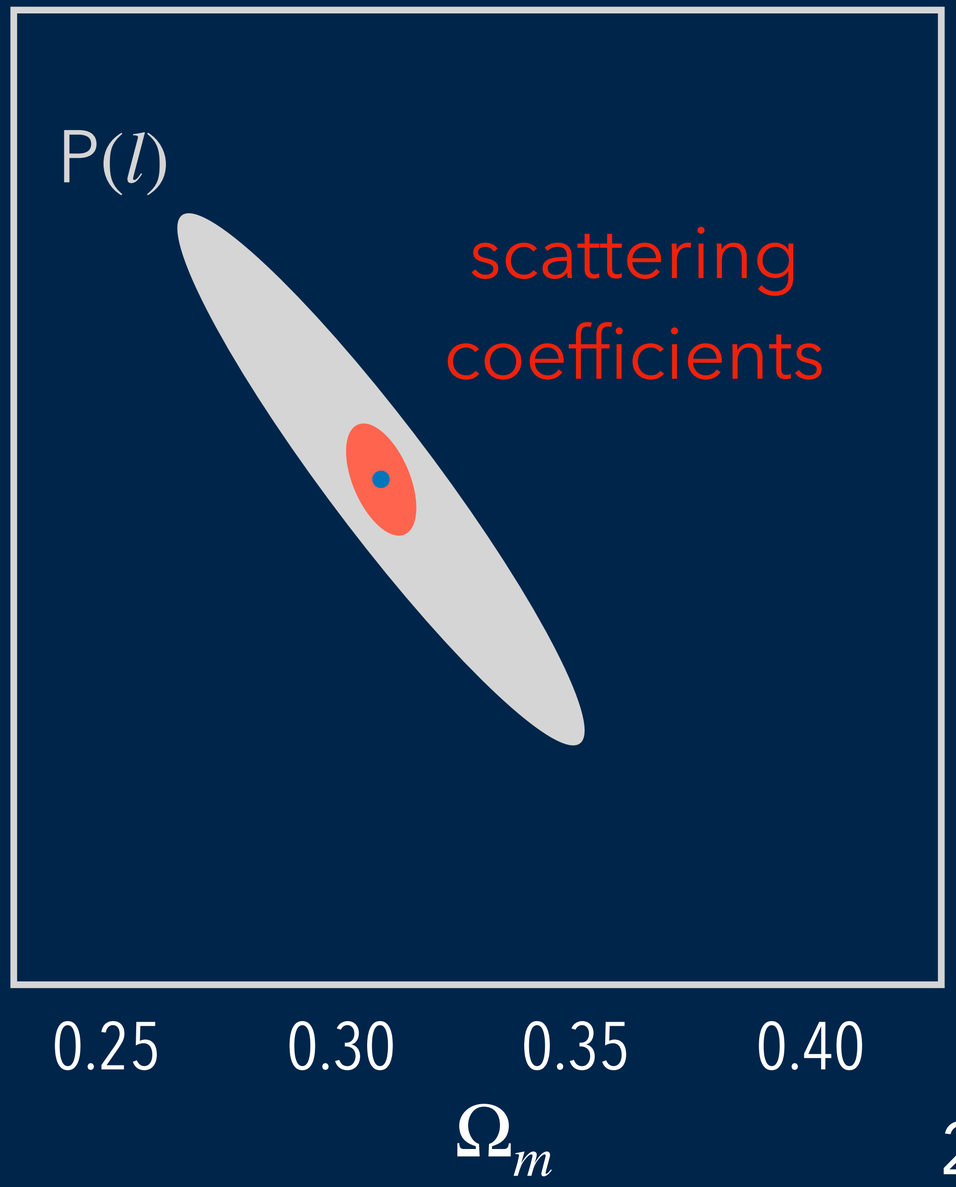
Fisher forecast



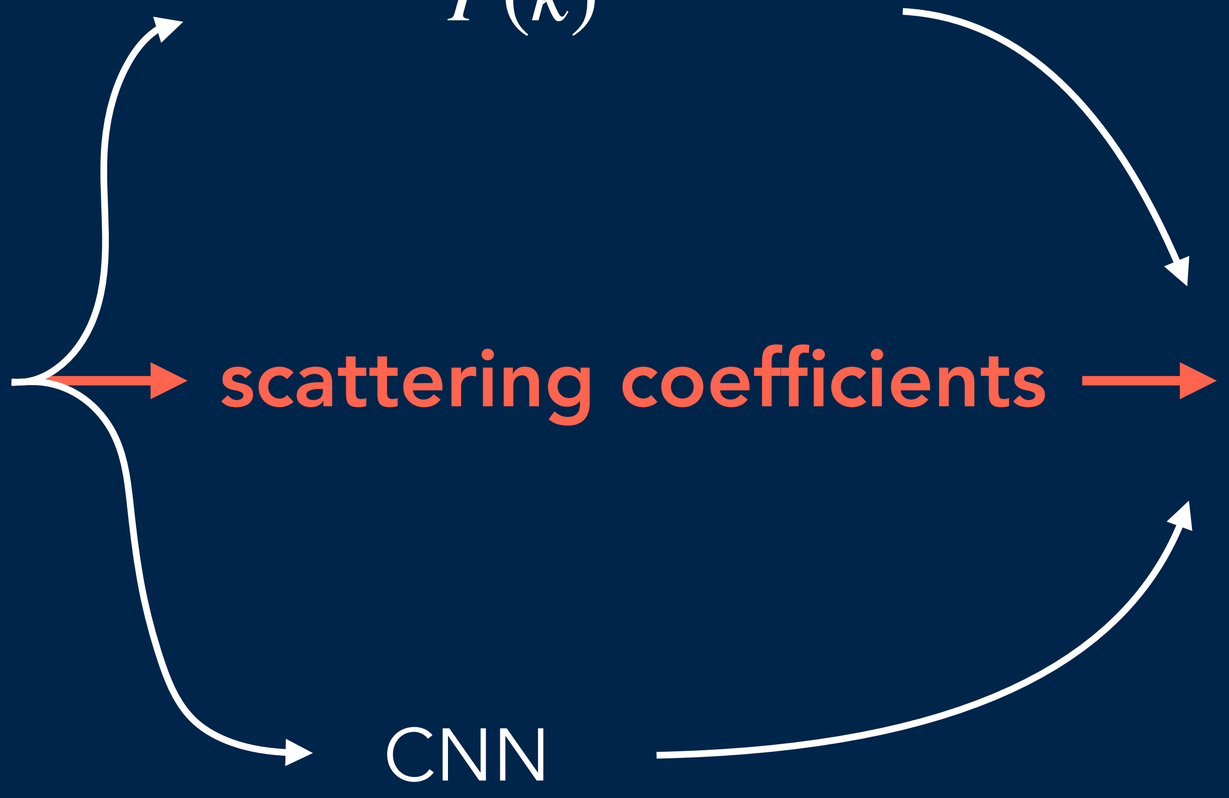
weak lensing cosmology



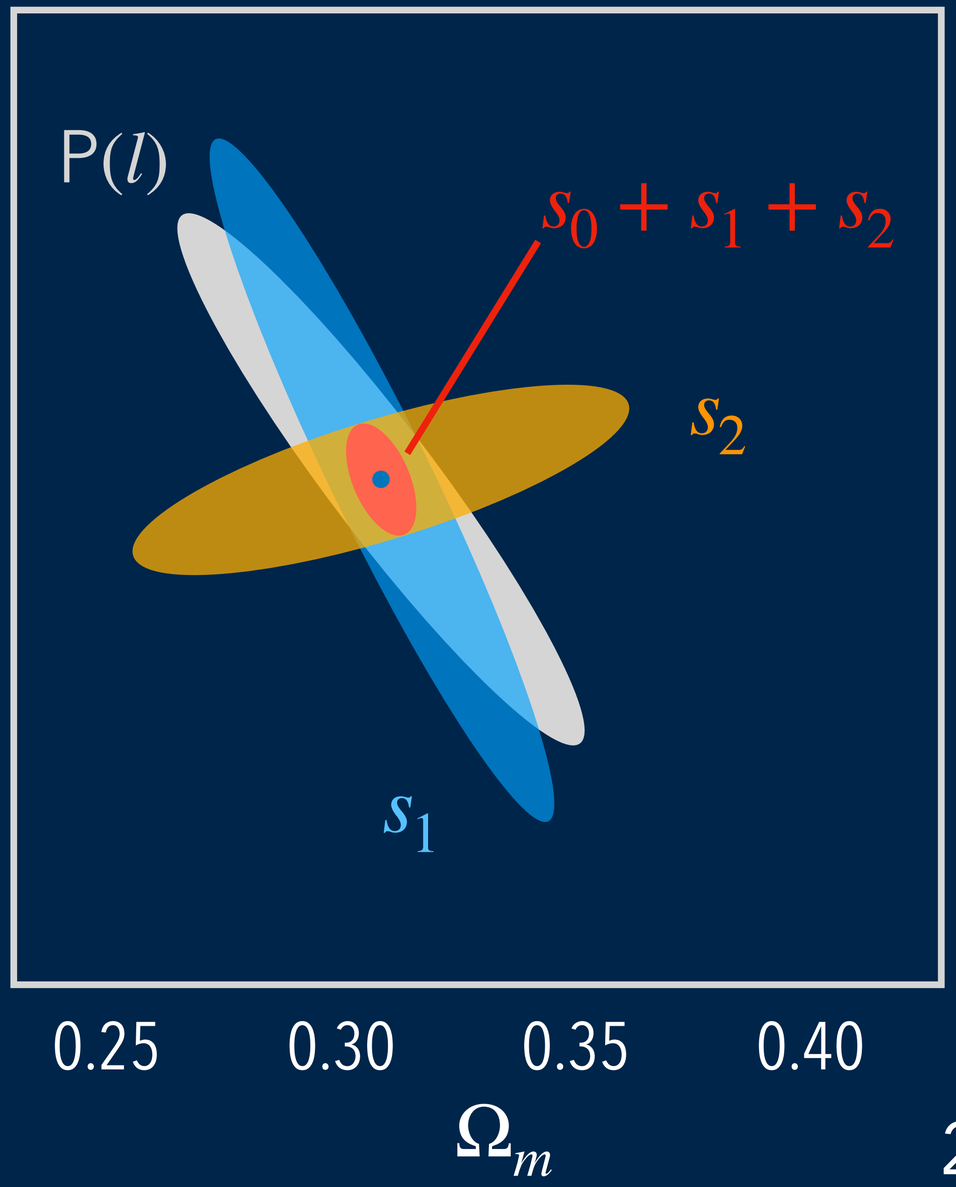
3.5x3.5 deg^2 noiseless map
scale range: 1 arcmin to 3.5 deg



weak lensing cosmology



3.5x3.5 deg² noiseless map
scale range: 1 arcmin to 3.5 deg



weak lensing cosmology

scale range: 1 arcmin to 3.5 deg

$P(k)$

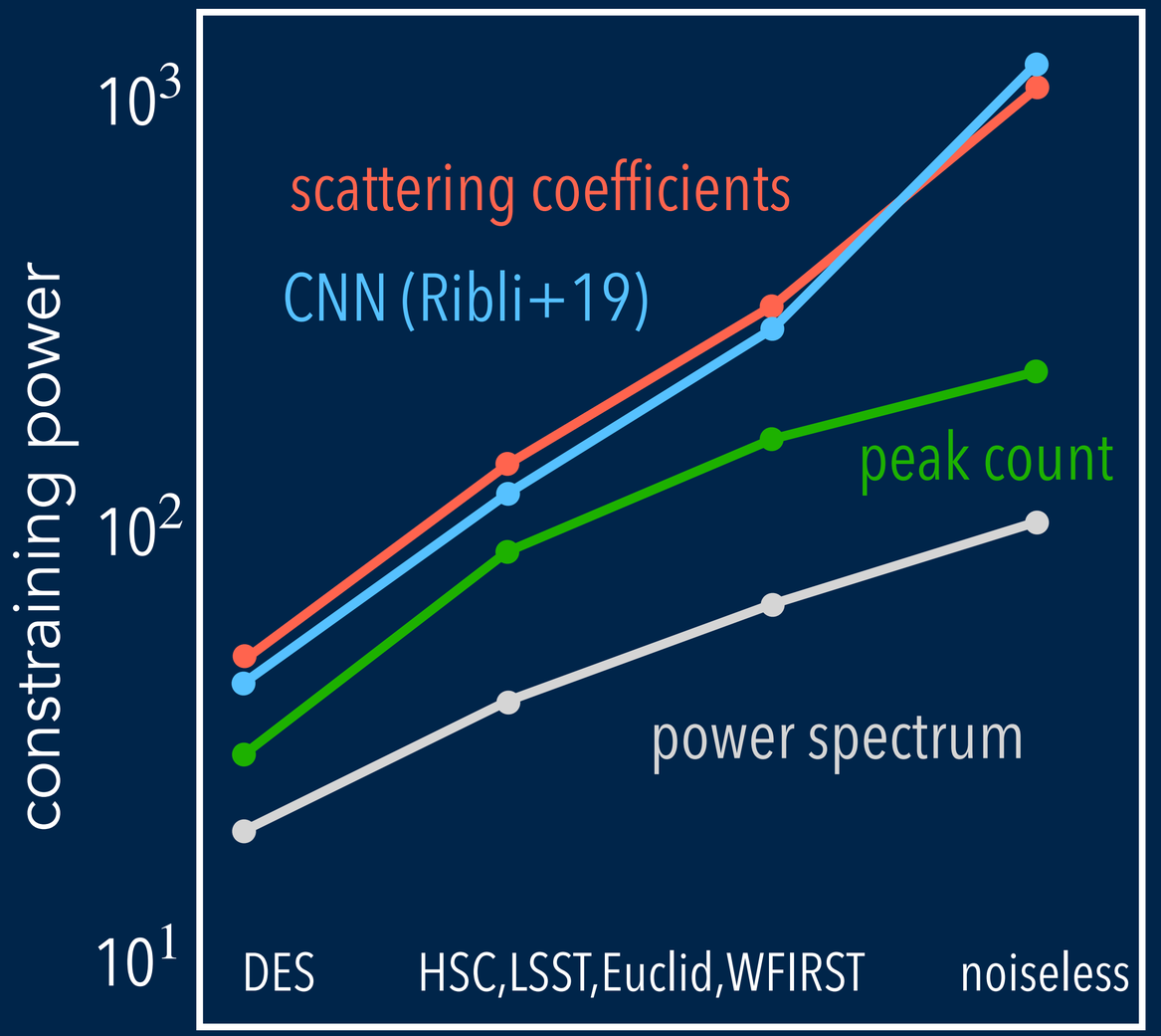
~20 coefficients

scattering coefficients

37 coefficients

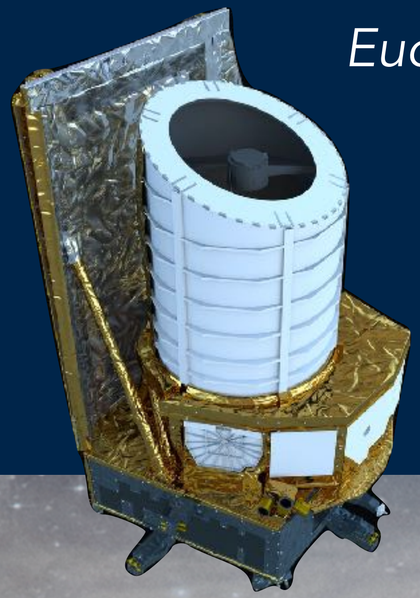
CNN

millions



dark energy & neutrino mass

on real data …



Euclid

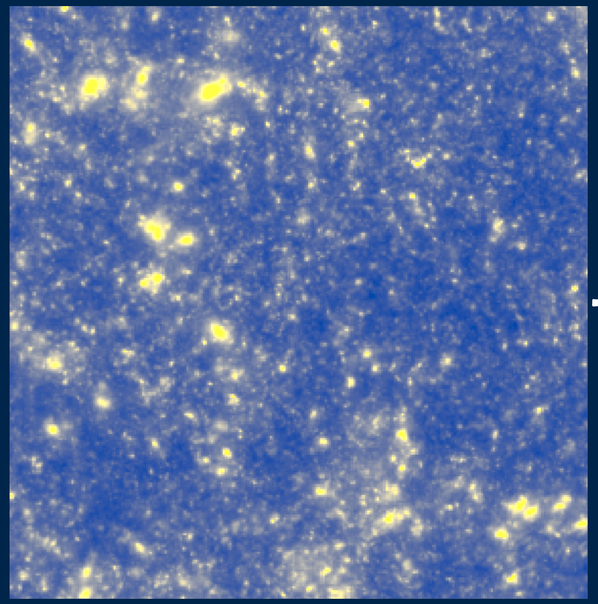
Roman Space Telescope (WFIRST)

HSC

Rubin Observatory (LSST)

How do we extract information from a field?

scattering transform:
similar to both N-point function
and CNN, but different



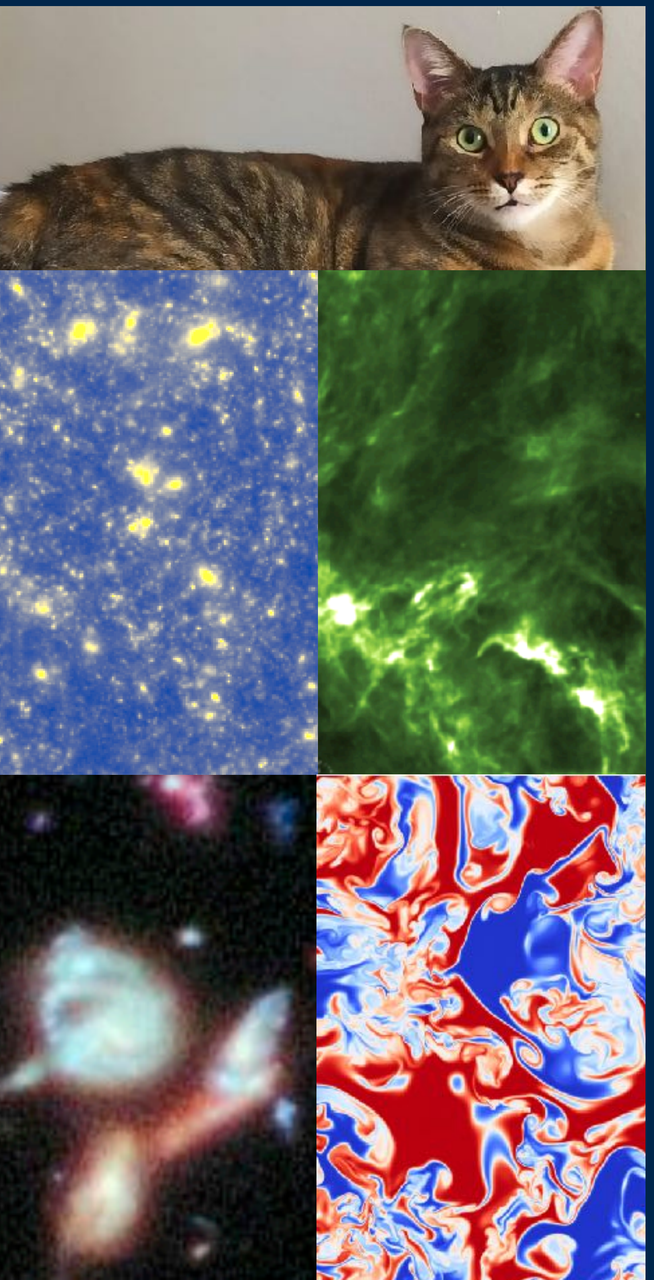
$P(k)$ plus
classic statistics

low order, wavelet binning
scattering transform
~ CNN without training

CNN

physical
parameters

How do we extract information from a field?



$P(k)$ plus
classic statistics

scattering transform

CNN

physical
parameters