The scattering transform in cosmology, or, a CNN without training

Sihao Cheng (程思浩)
Johns Hopkins University

arXiv: 2006.08561
arXiv: 2103.09247

Berkeley Cosmology Seminar
September 14th, 2021

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

image credit: Sloan Digital Sky Survey

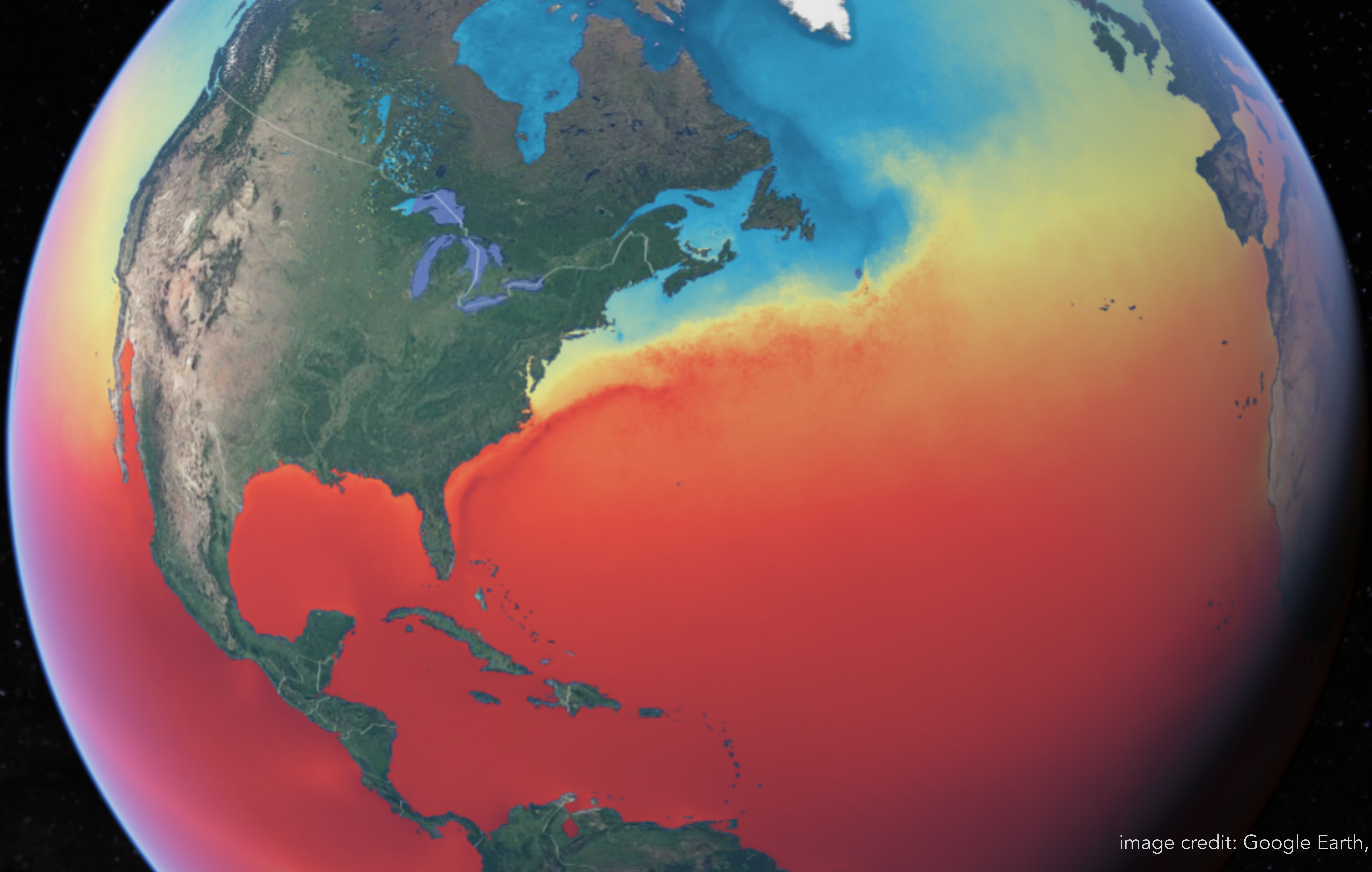
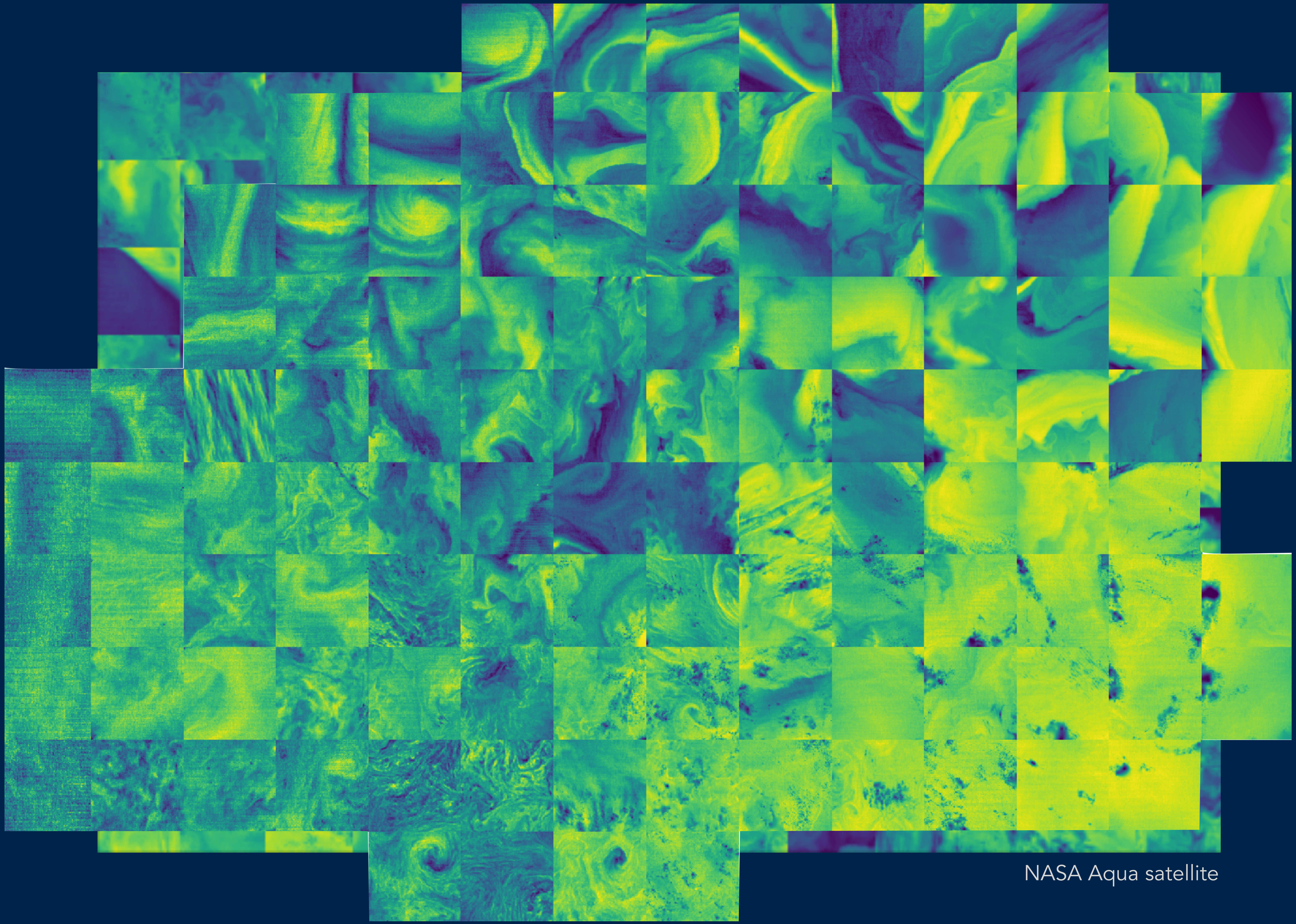


image credit: Google Earth, NOAA

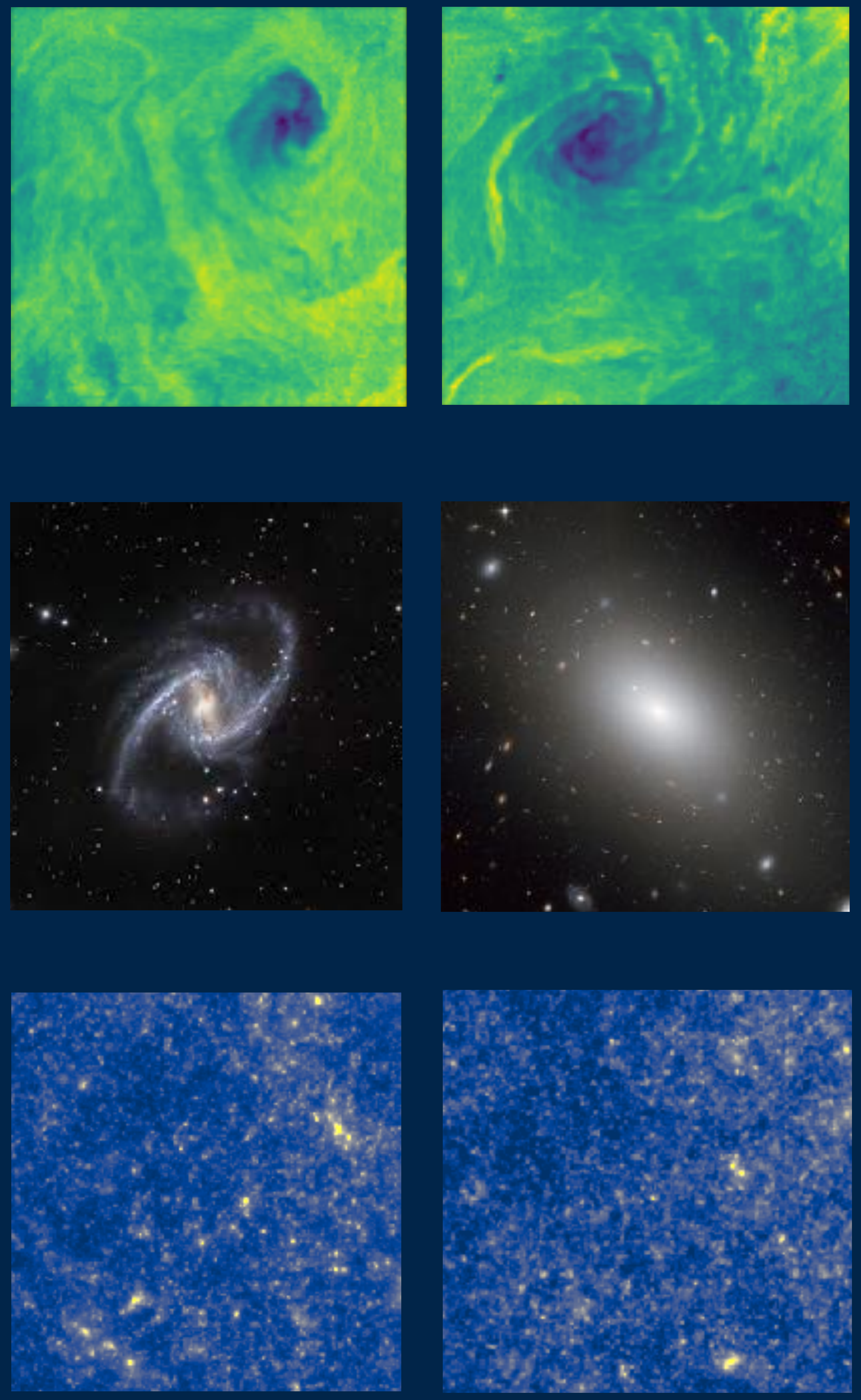


NASA Aqua satellite

complex
data

power spectrum P(k)

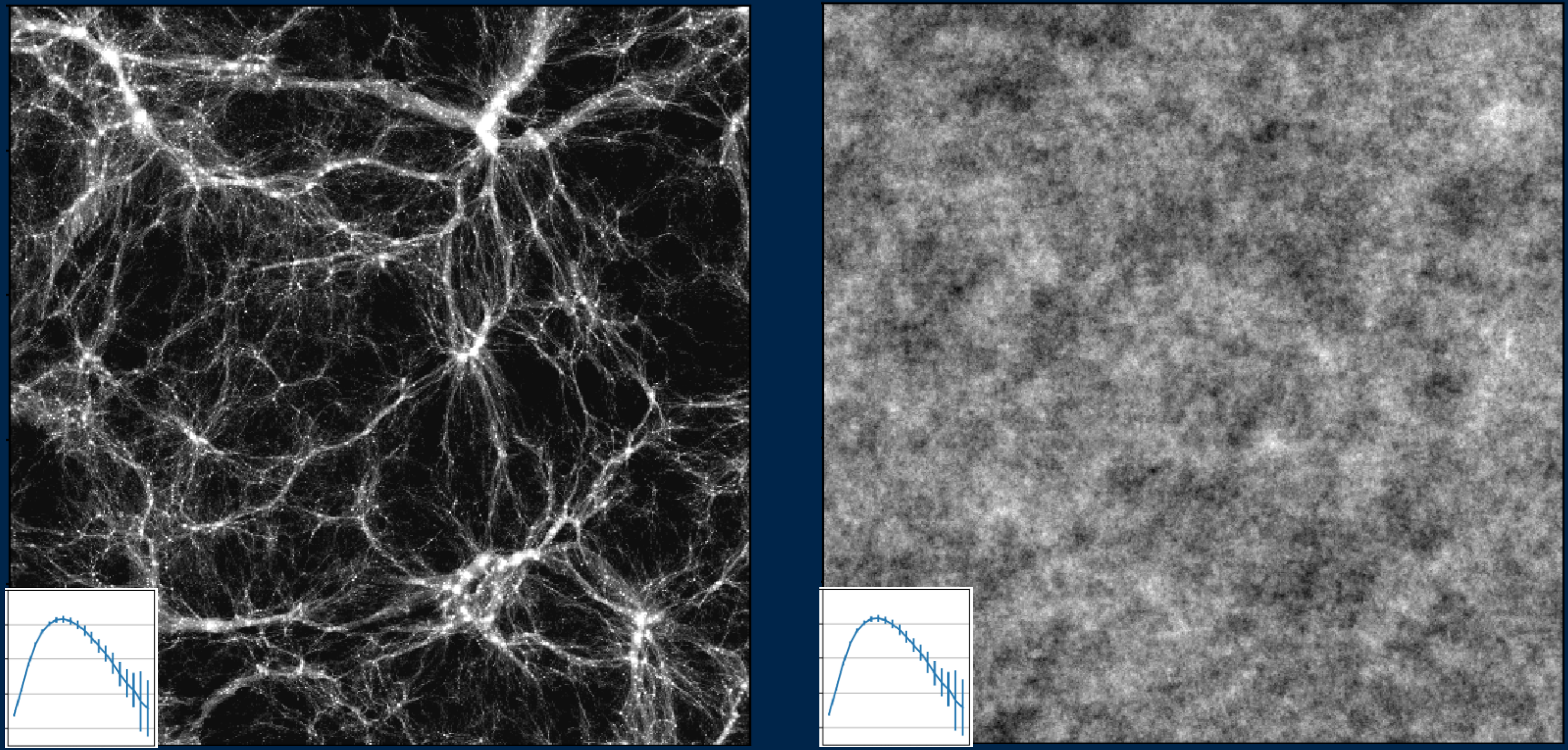
simple
information



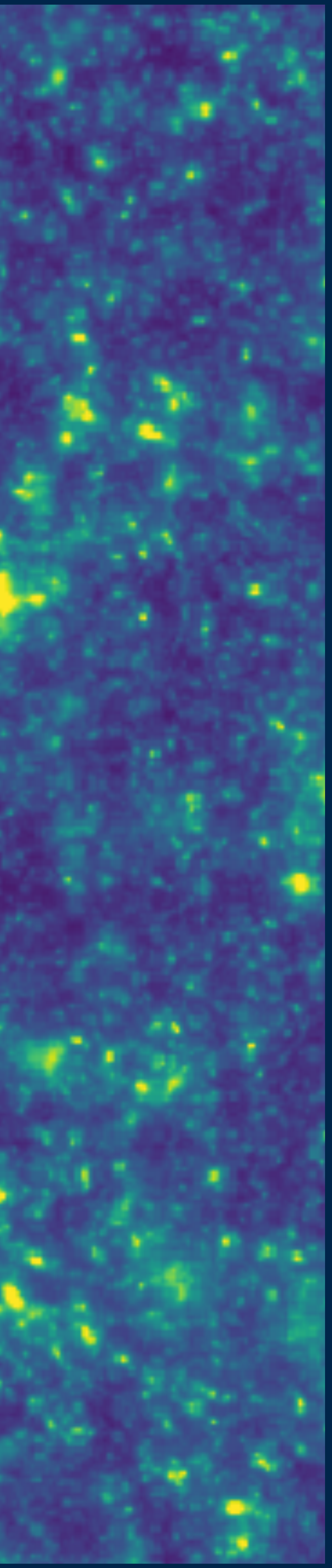
data exploration
no model

classification
discrete model

parameter inference
continuous model



How do we characterize a field?



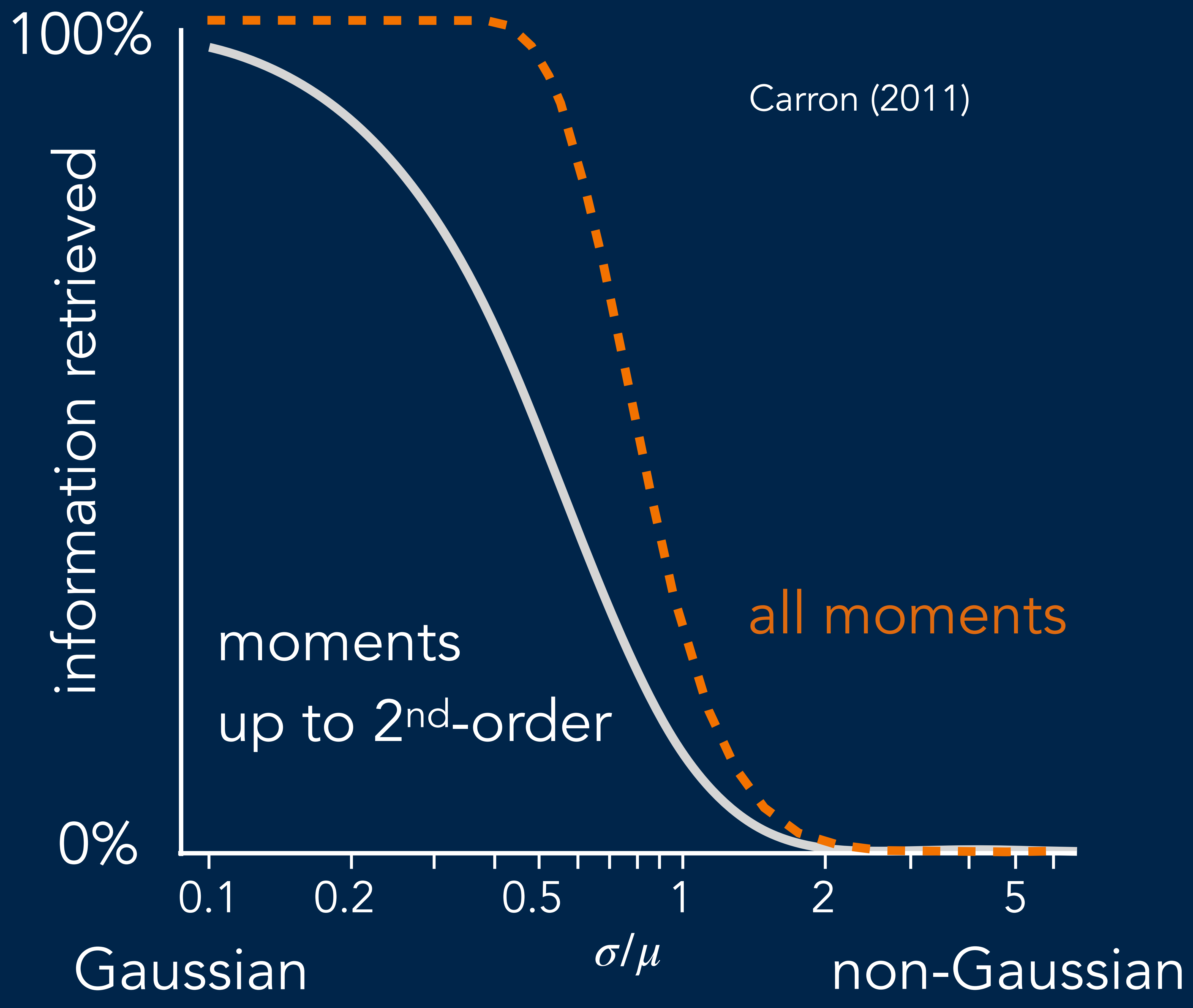
$$\langle \delta_1 \delta_2 \ldots \delta_n \rangle$$

power spectrum

plus classic statistics

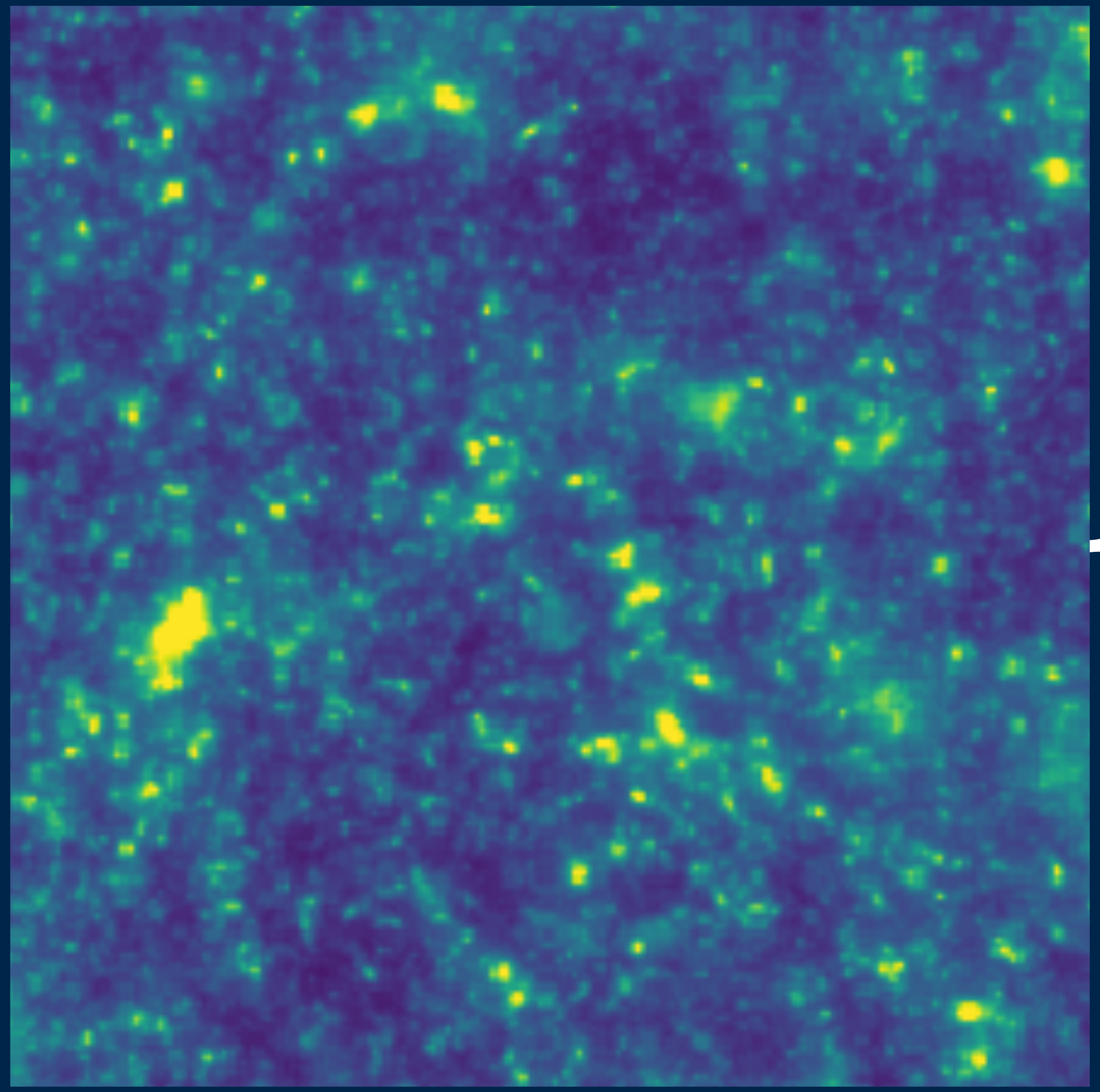
→ physical information

log-normal variable



- lose information
- too many coefficients

How do we characterize a field?



a number of limitations

power spectrum

plus classic statistics

scattering transform

(Mallat 2012)

physical information

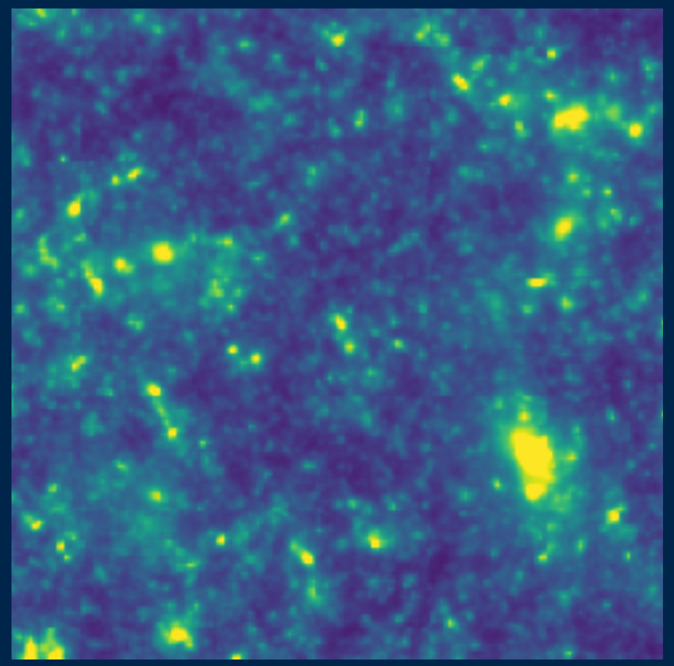
CNN

C = hierarchical convolutions

NN = learning ability, but a black box

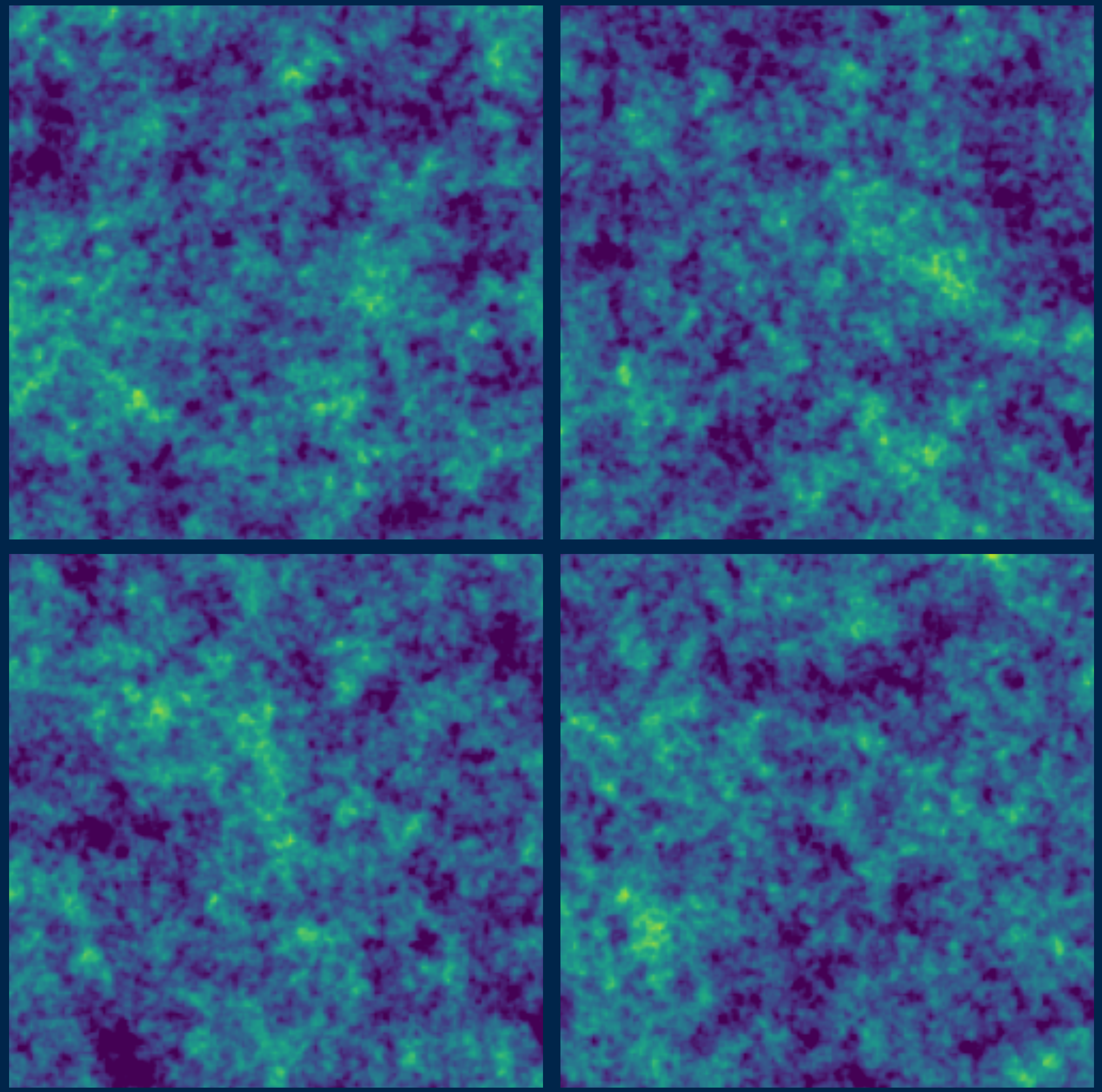
What do the statistics see?

arXiv: 2103.09247

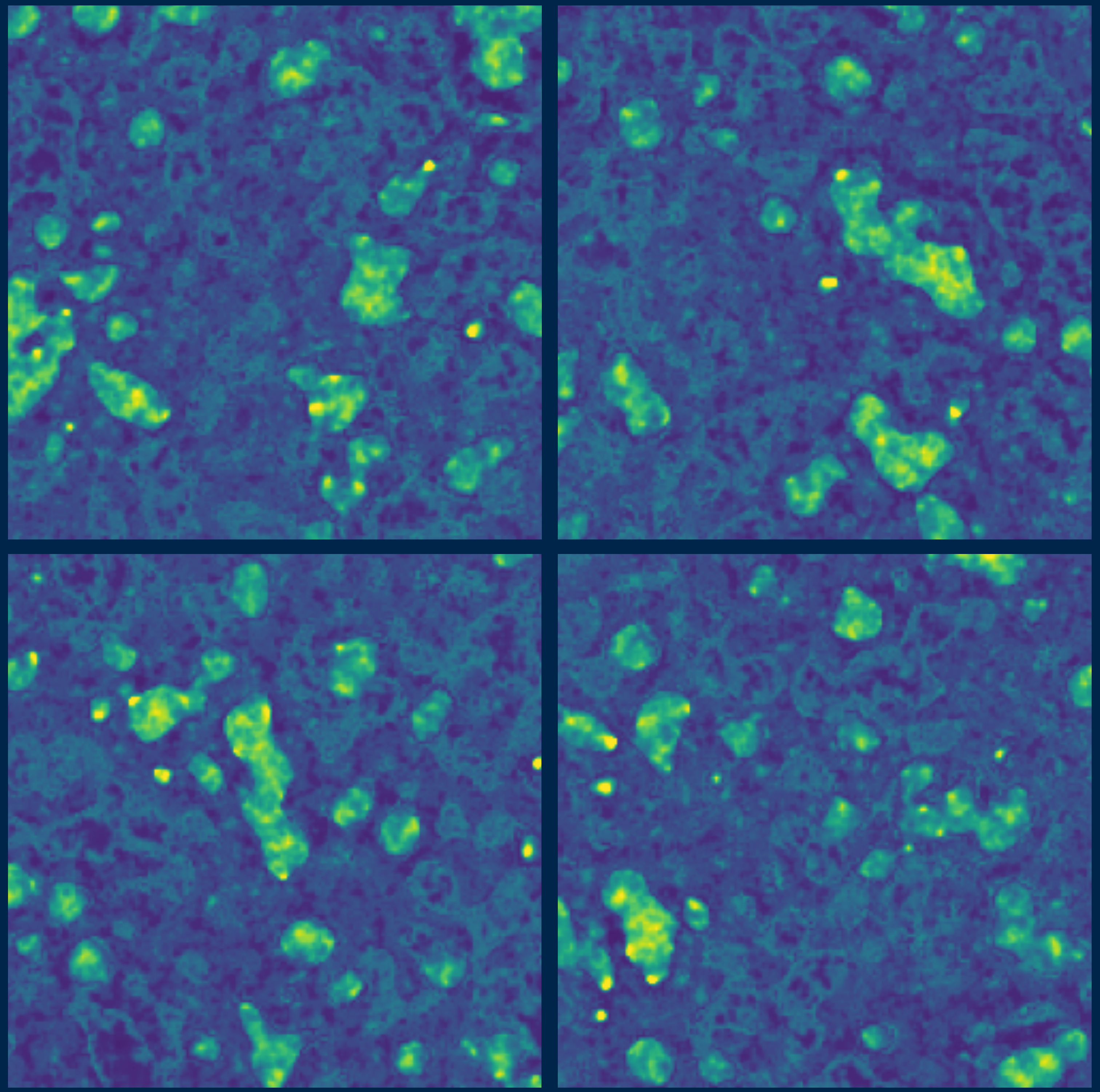


input map

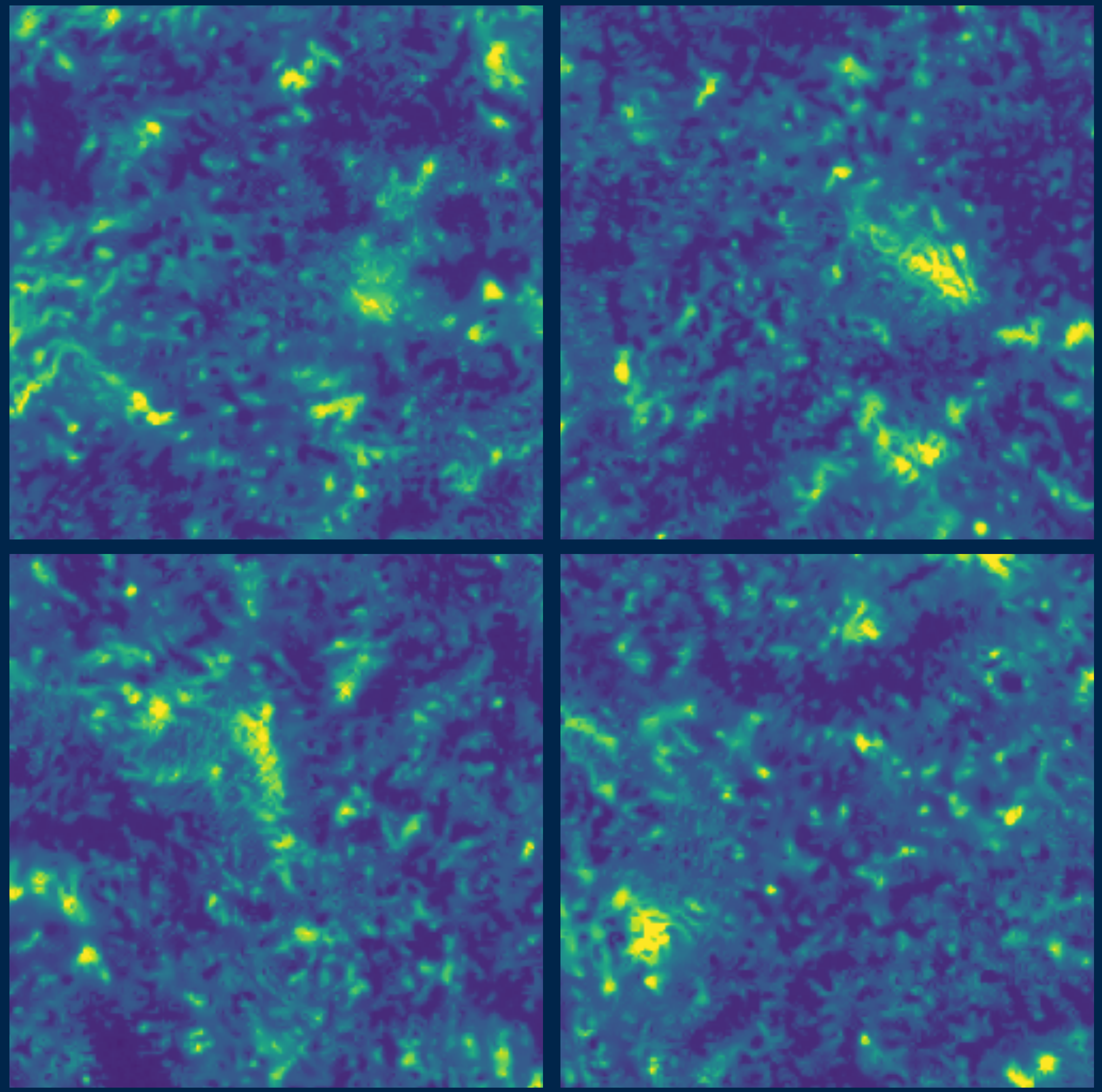
with power spectrum P(k)



with P(k) and bispectrum



with scattering statistics

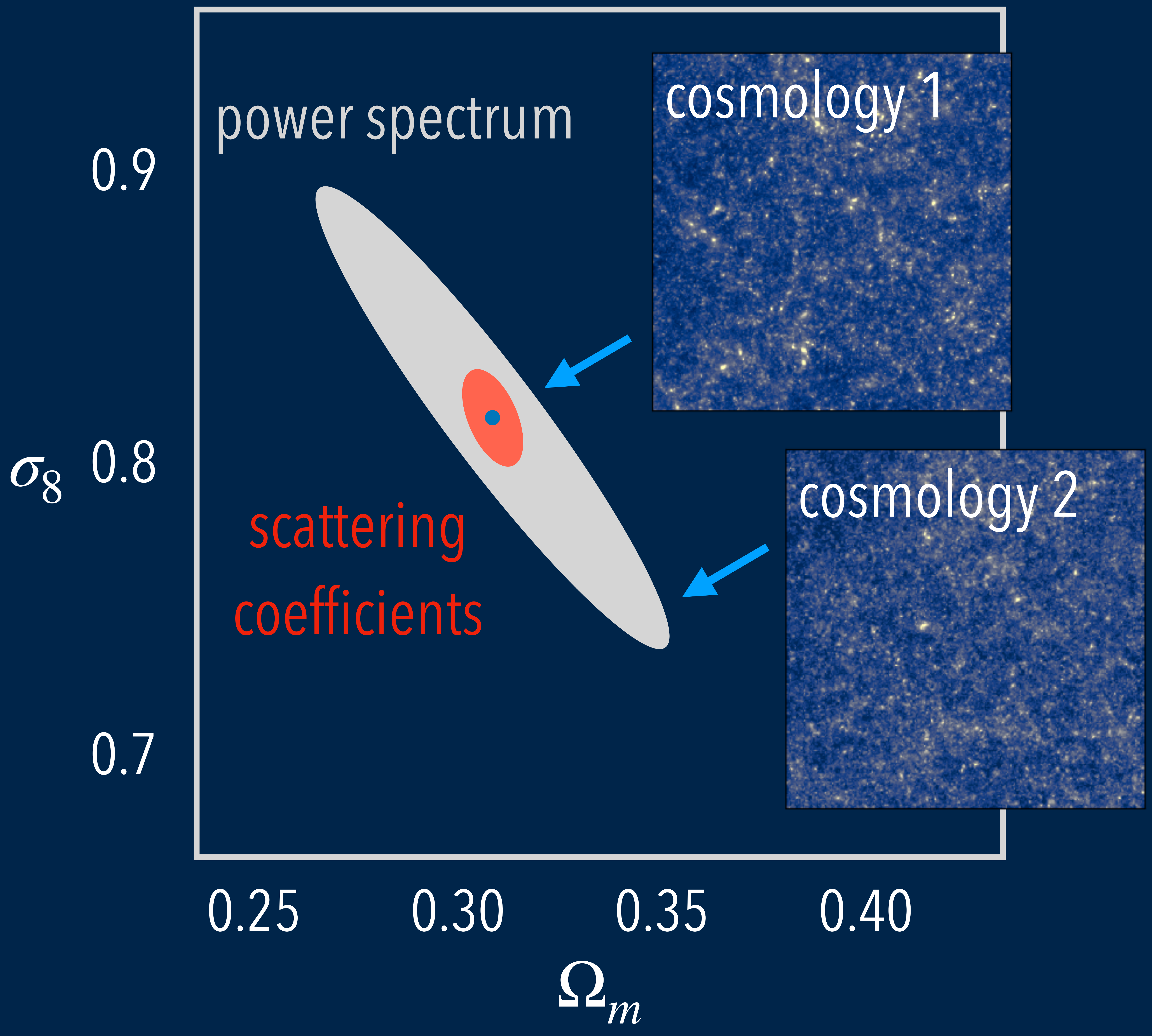


Cheng & Menard 2021

Turing pattern | Ising model | sea temperature | solar UV image | cosmic matter

with scattering statistics (translation invariant)

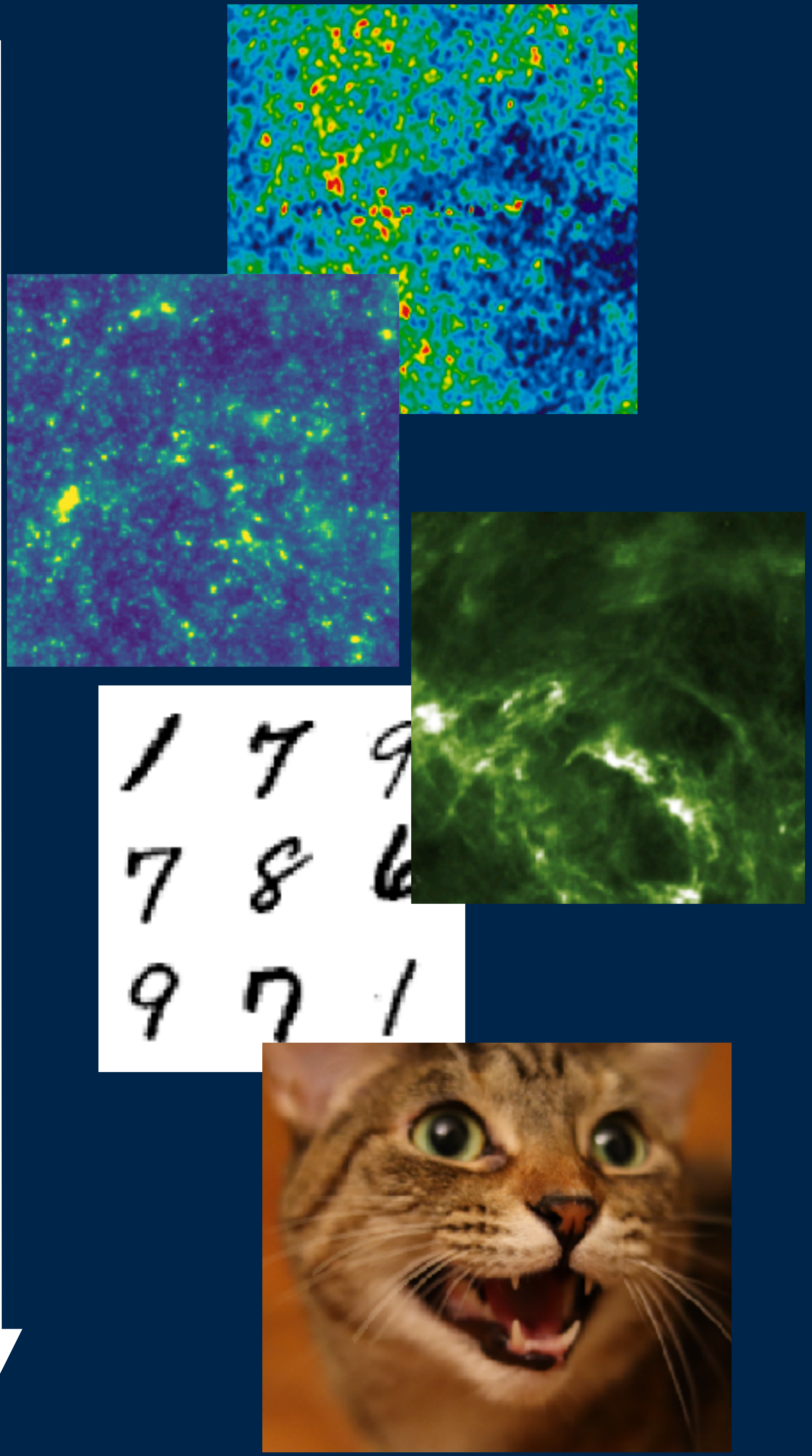
Cheng & Menard, in prep



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

How do we characterize a field?

apparent "complexity"



power spectrum

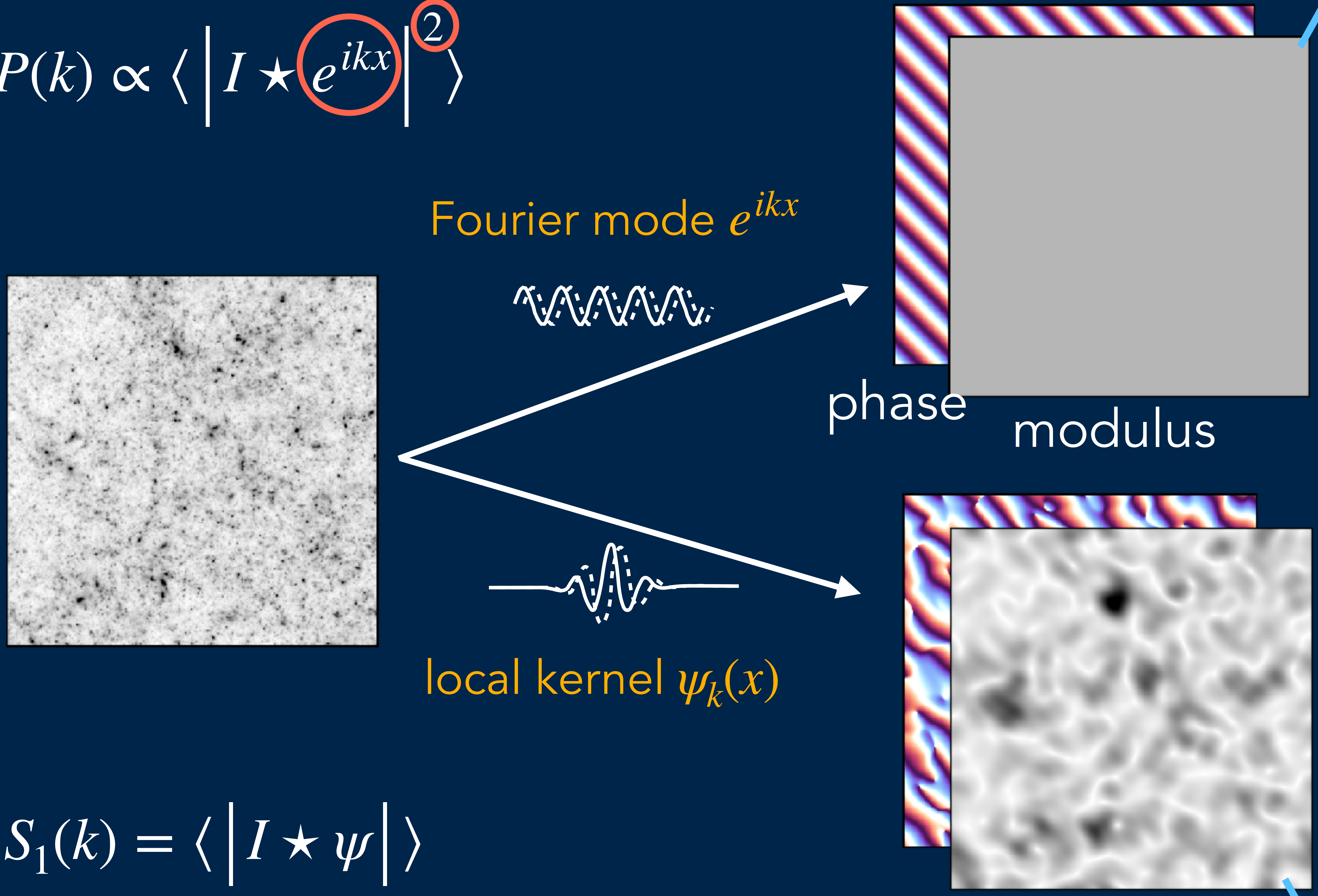
scattering transform

CNN

from power spectrum to scattering transform

$$\langle \cdot^2 \rangle = P(k)$$

$$P(k) \propto \langle \left| I \star e^{ikx} \right|^2 \rangle$$



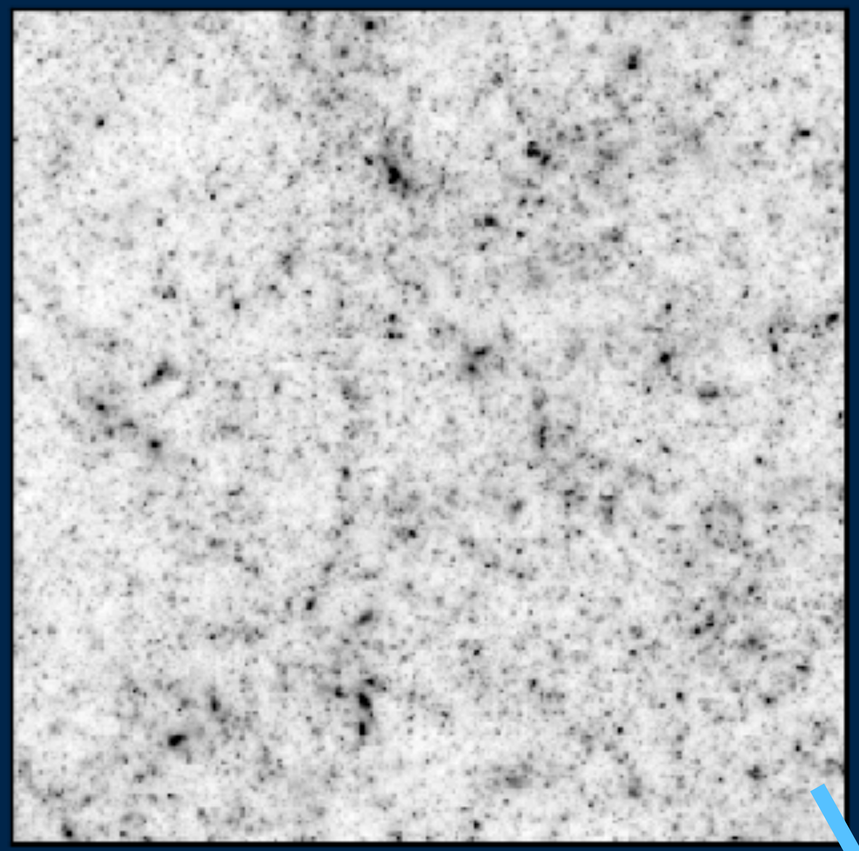
$$S_1(k) = \langle \left| I \star \psi \right| \rangle$$

$$\langle \cdot \rangle = S_1(k)$$

from power spectrum to scattering transform

$$P(k) \propto \left\langle \left| I \star e^{ikx} \right|^2 \right\rangle$$

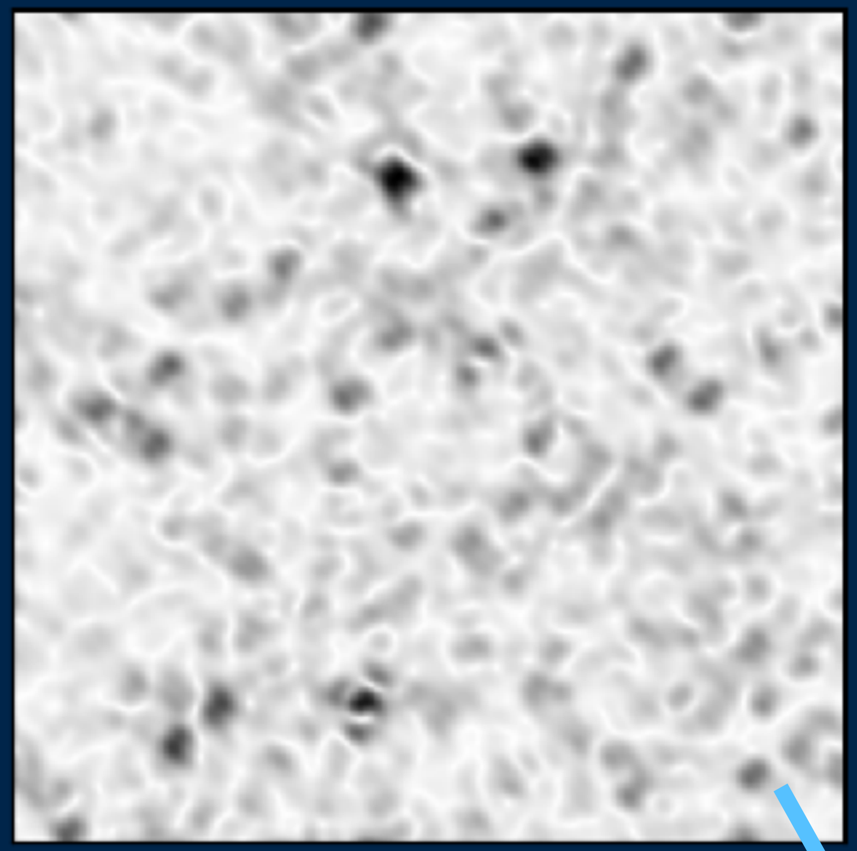
I



scale 1

$S_0 = \langle I \rangle$

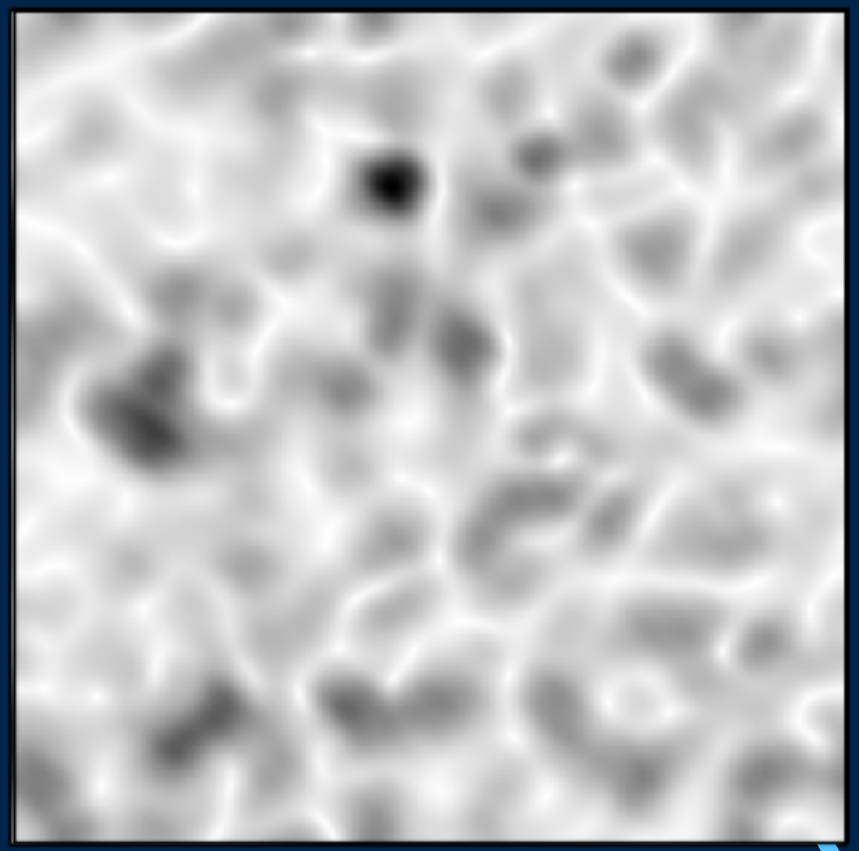
$I_1 = \left| I \star \psi \right|$



scale 2

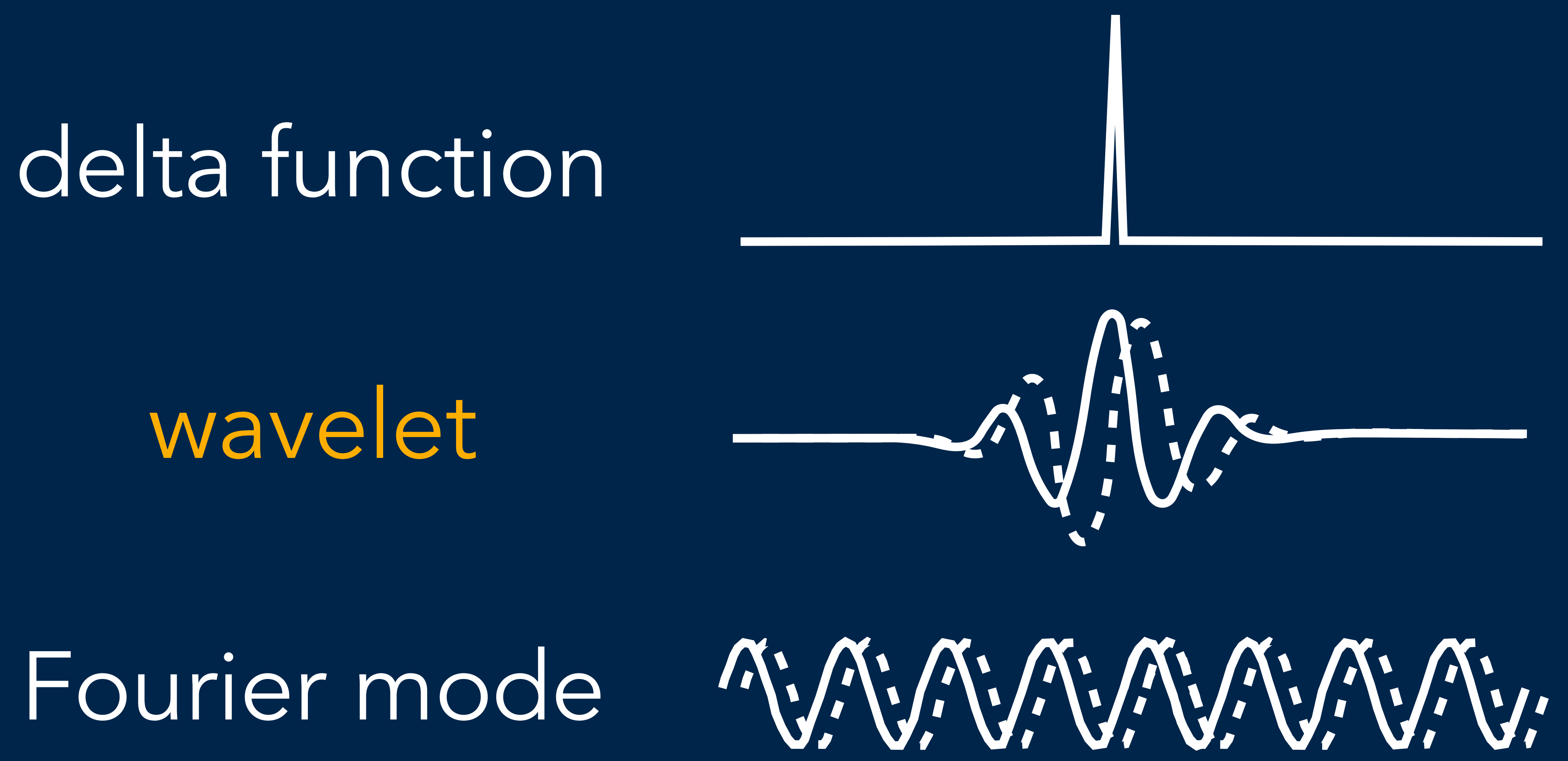
$S_1(k) = \langle I_1 \rangle$

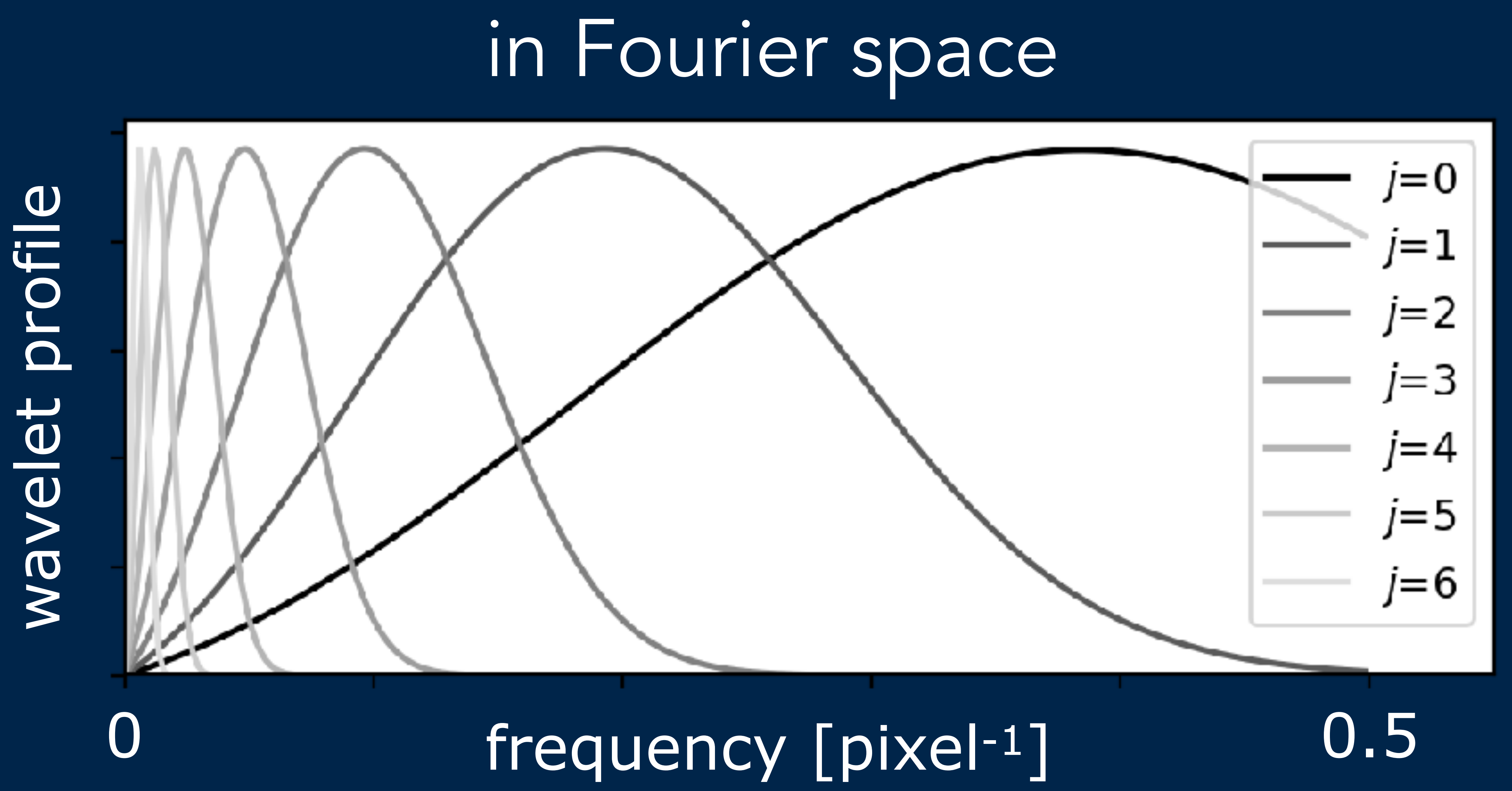
$I_2 = \left| \left| I \star \psi_1 \right| \star \psi_2 \right|$



$S_2(k_1, k_2) = \langle I_2 \rangle$

wavelets: logarithmic binning is efficient and stable





$$\psi^{j,l}(x)$$

j: logarithmic sampling of scales

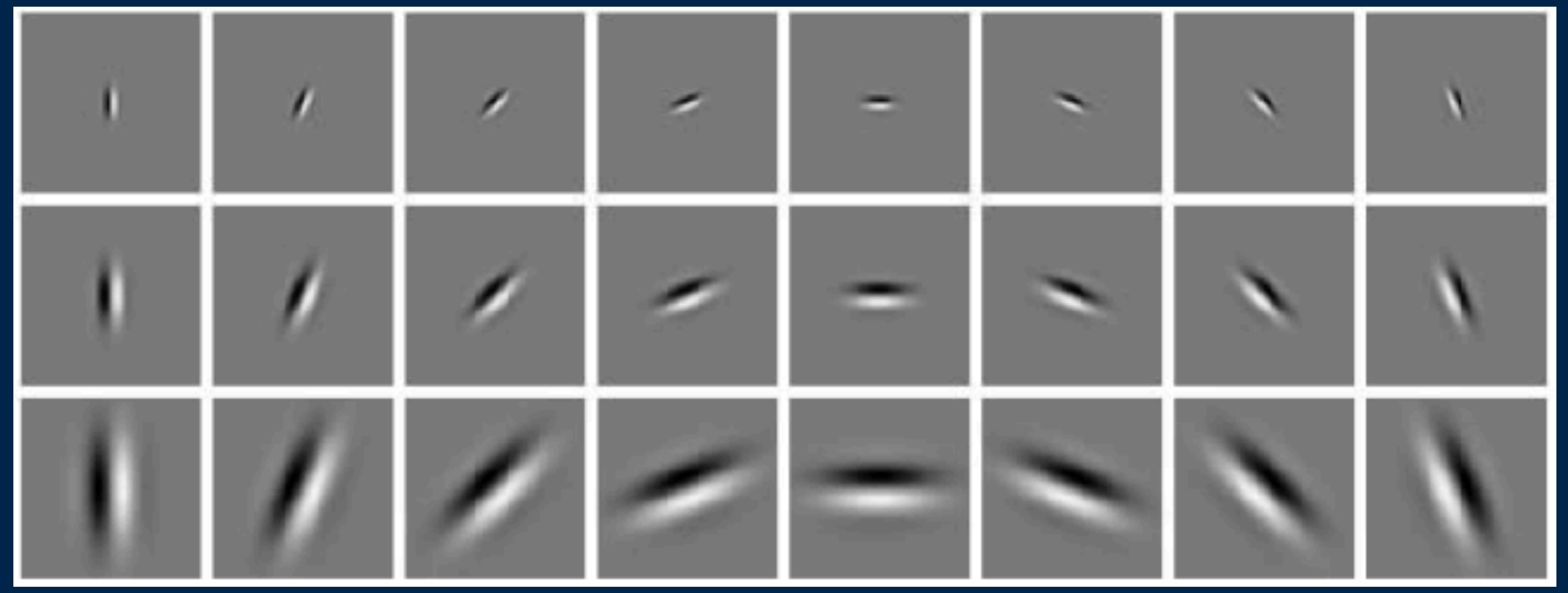
l: orientations

wavelets: an efficient decomposition

close to Gabor wavelets

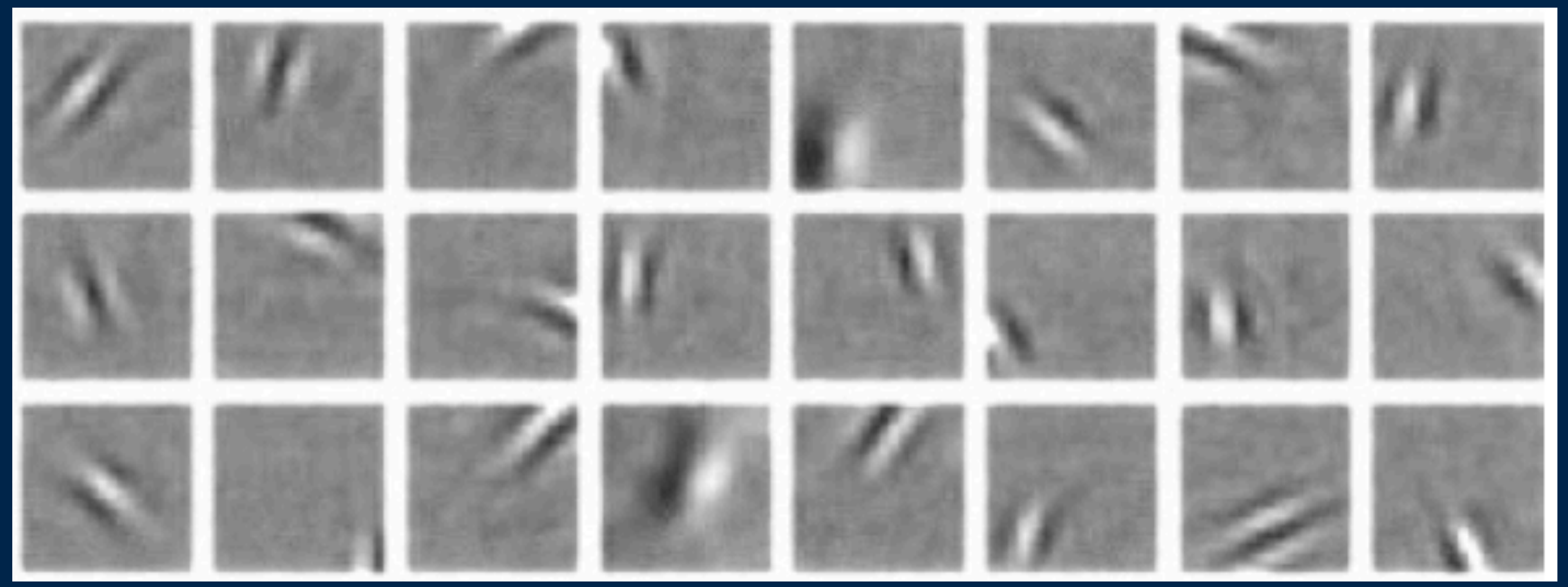
receptive fields of mammal vision

(Hubel & Wiesel 1968)



sparse representation of natural images

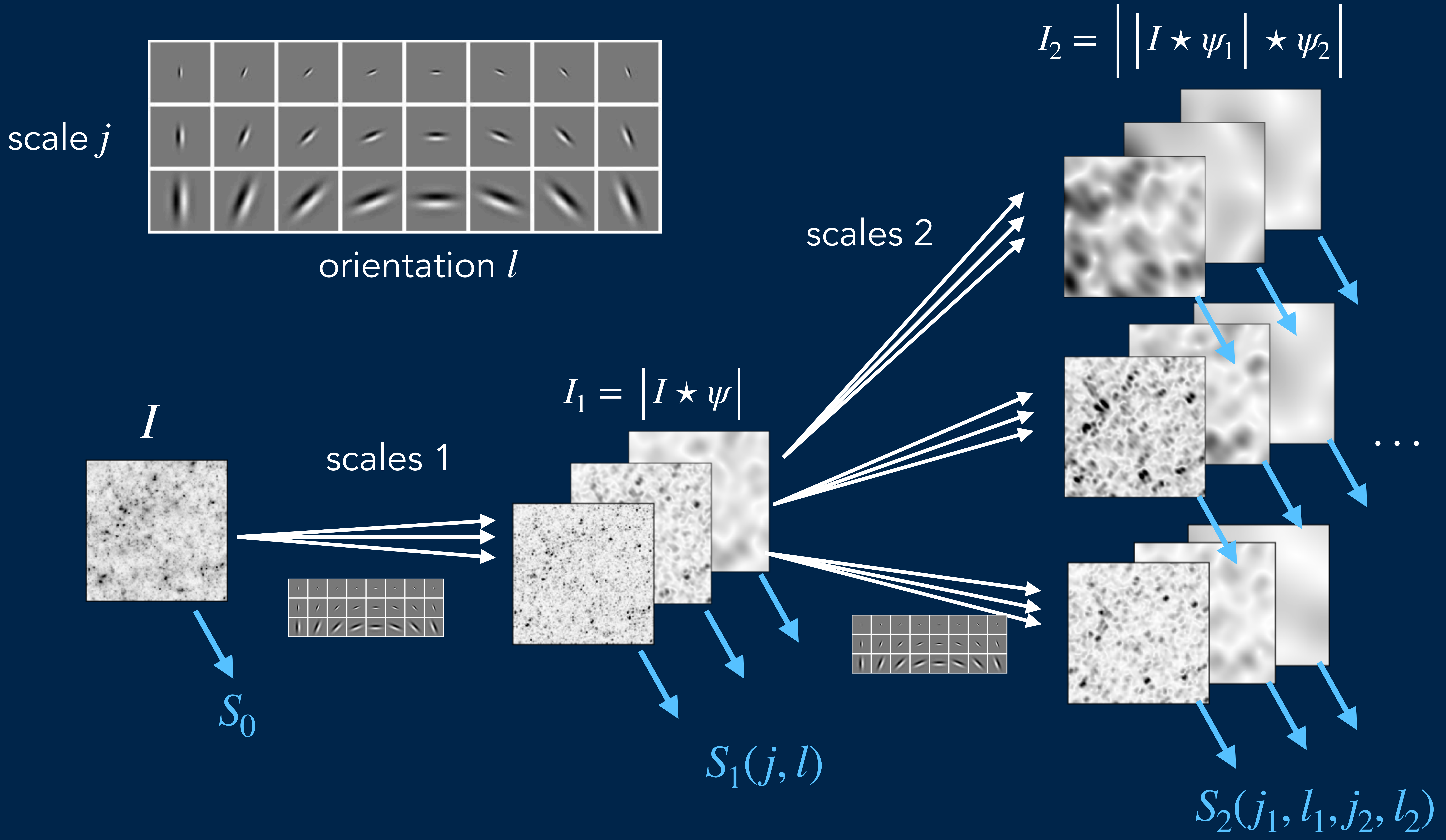
(Olshausen & Field 1996)



kernels learned in AlexNet

(Krizhevsky, Sutskever, & Hinton 2012)

convolutional network



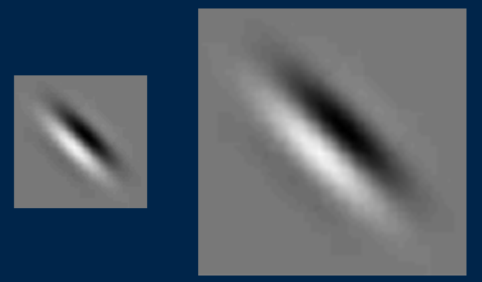
interpretation

$$S_0 = \langle I \rangle:$$ mean

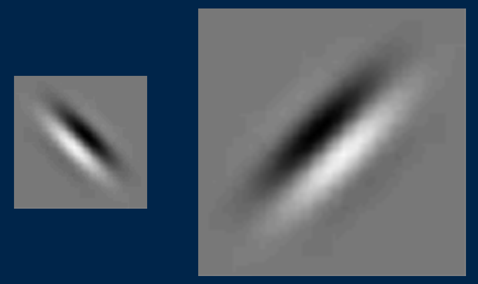
$$S_1(j,l) = \langle\, |I \star \psi|\, \rangle:$$ $\sim P(k)$

$$S_2(j_1,l_1,j_2,l_2) = \langle\, \Big|\, |I \star \psi_1| \star \psi_2 \Big|\, \rangle:$$ non-Gaussianity. $\sim P(k)$ of $P(k)$

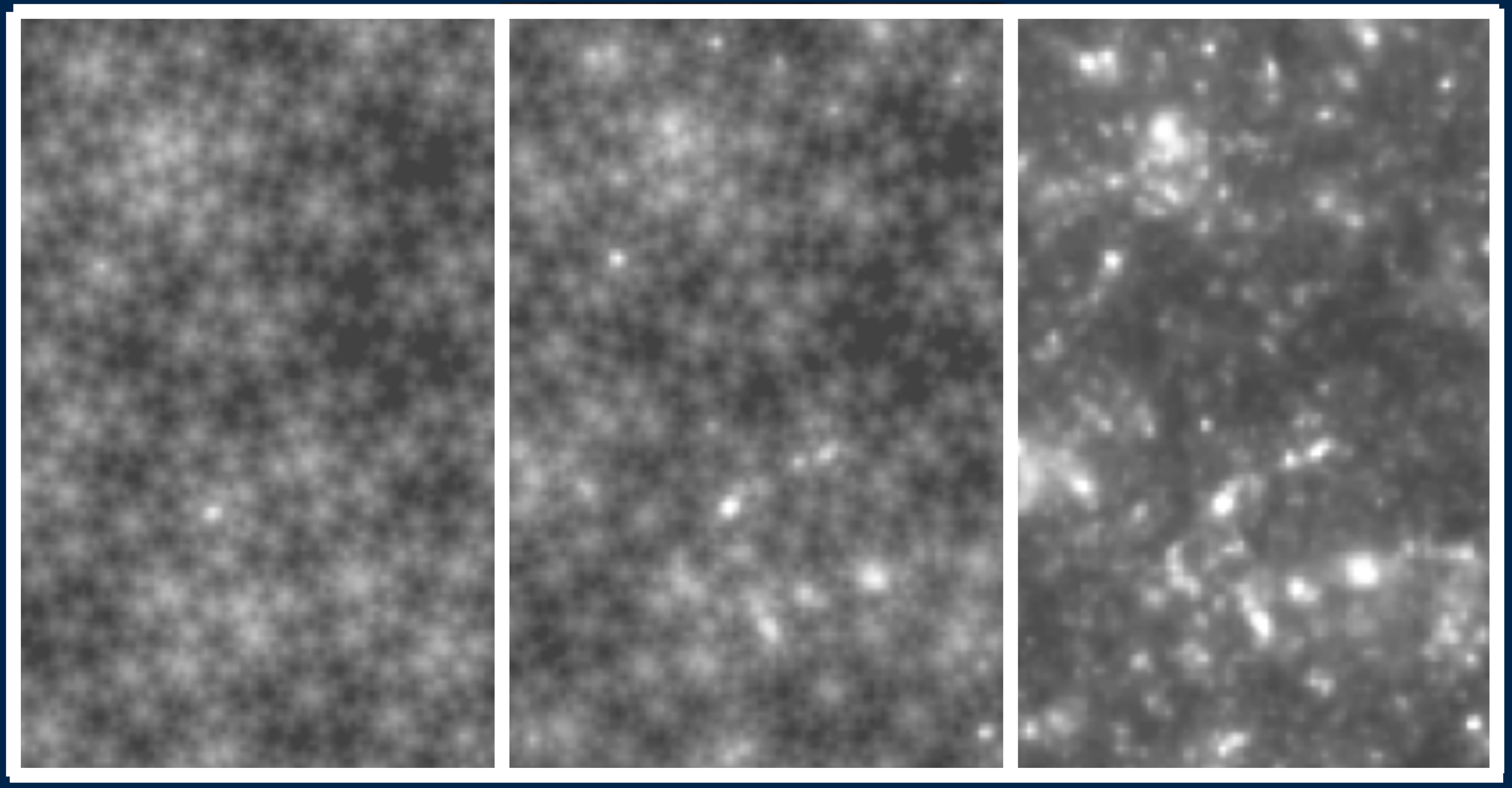
$l_1 \parallel l_2$:



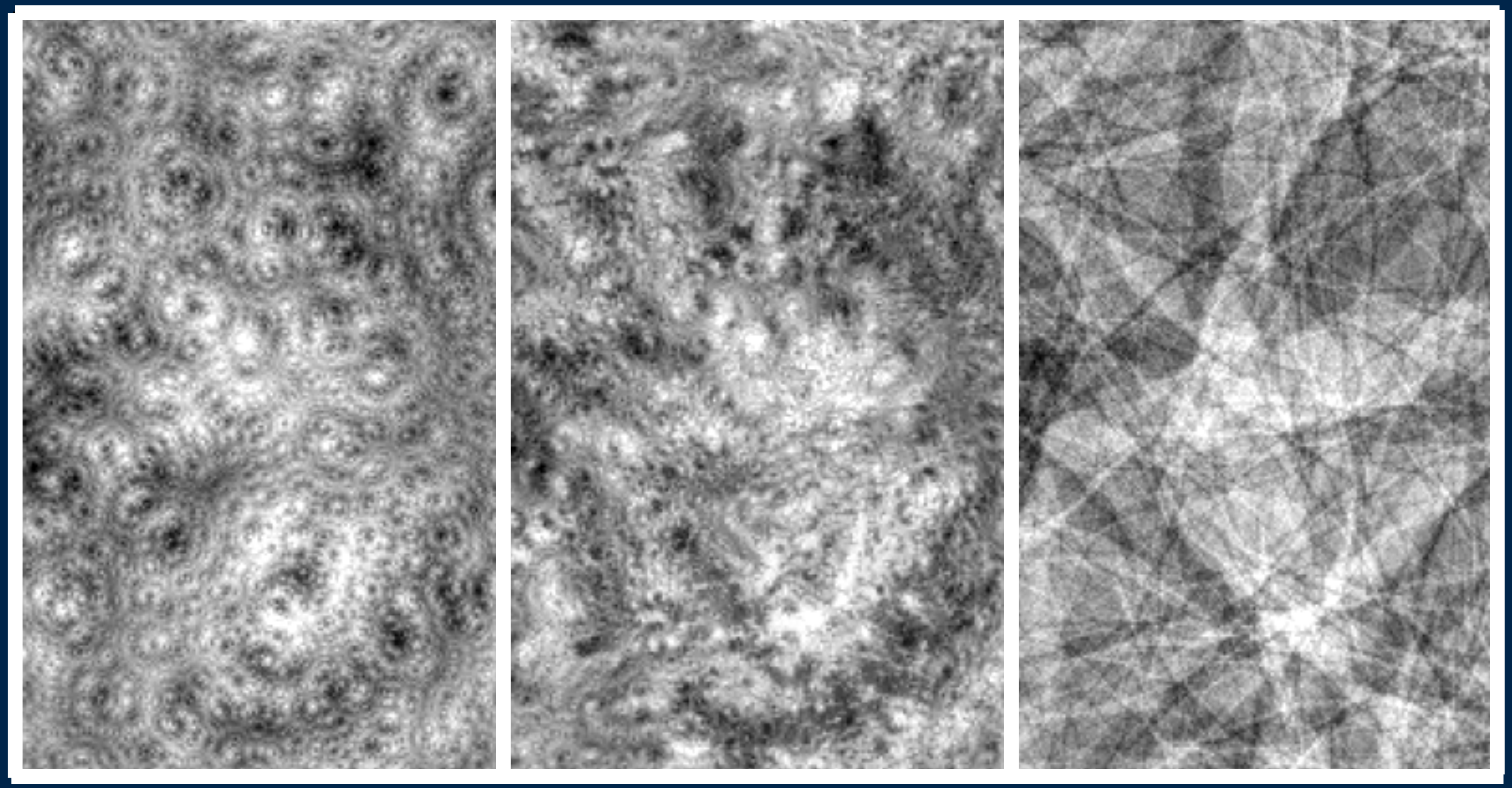
$l_1 \perp l_2$:



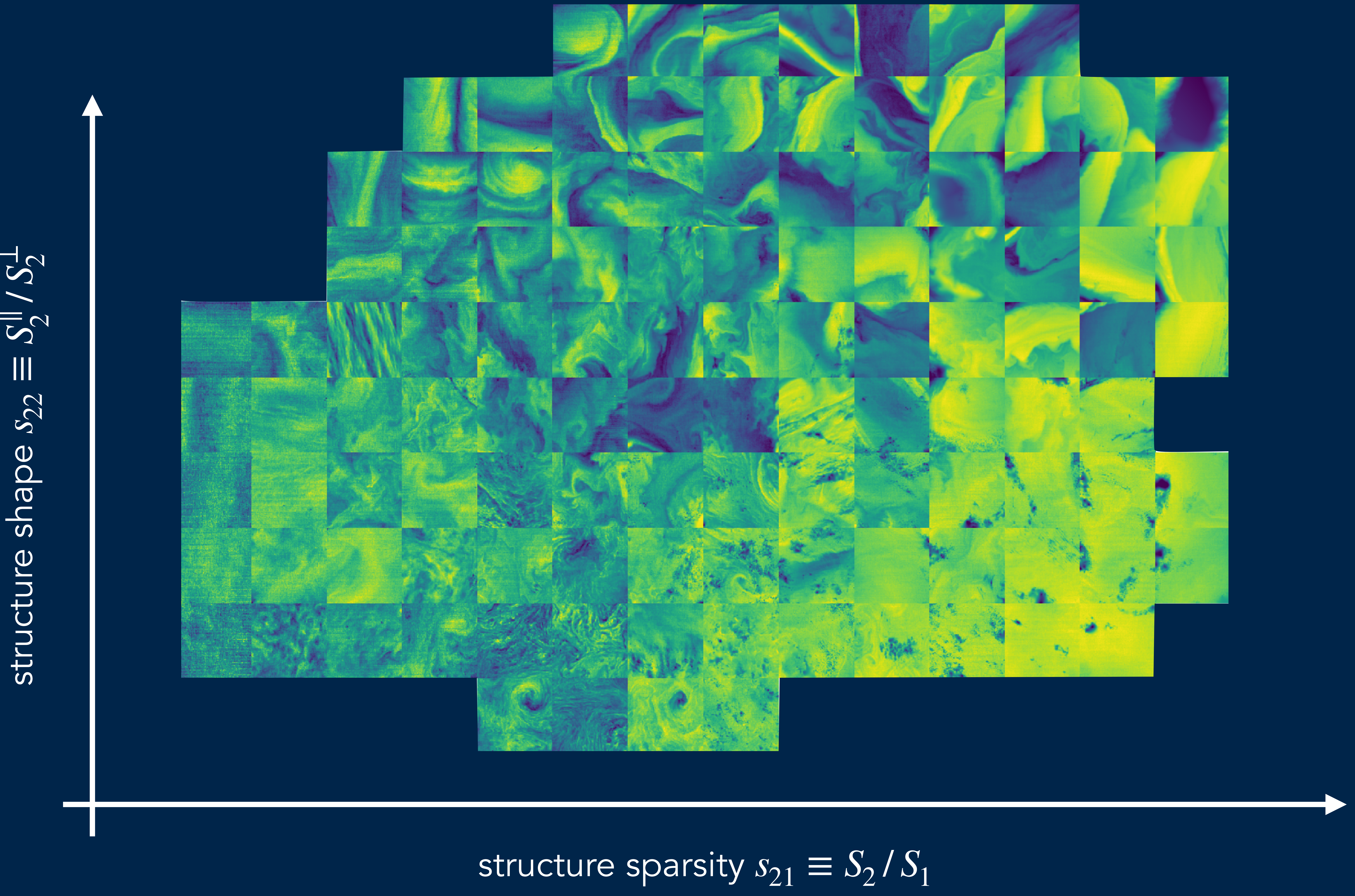
structure sparsity $s_{21} \equiv \overline{S_2} / S_1$



structure shape $s_{22} \equiv S_2^{\parallel} / S_2^{\perp}$



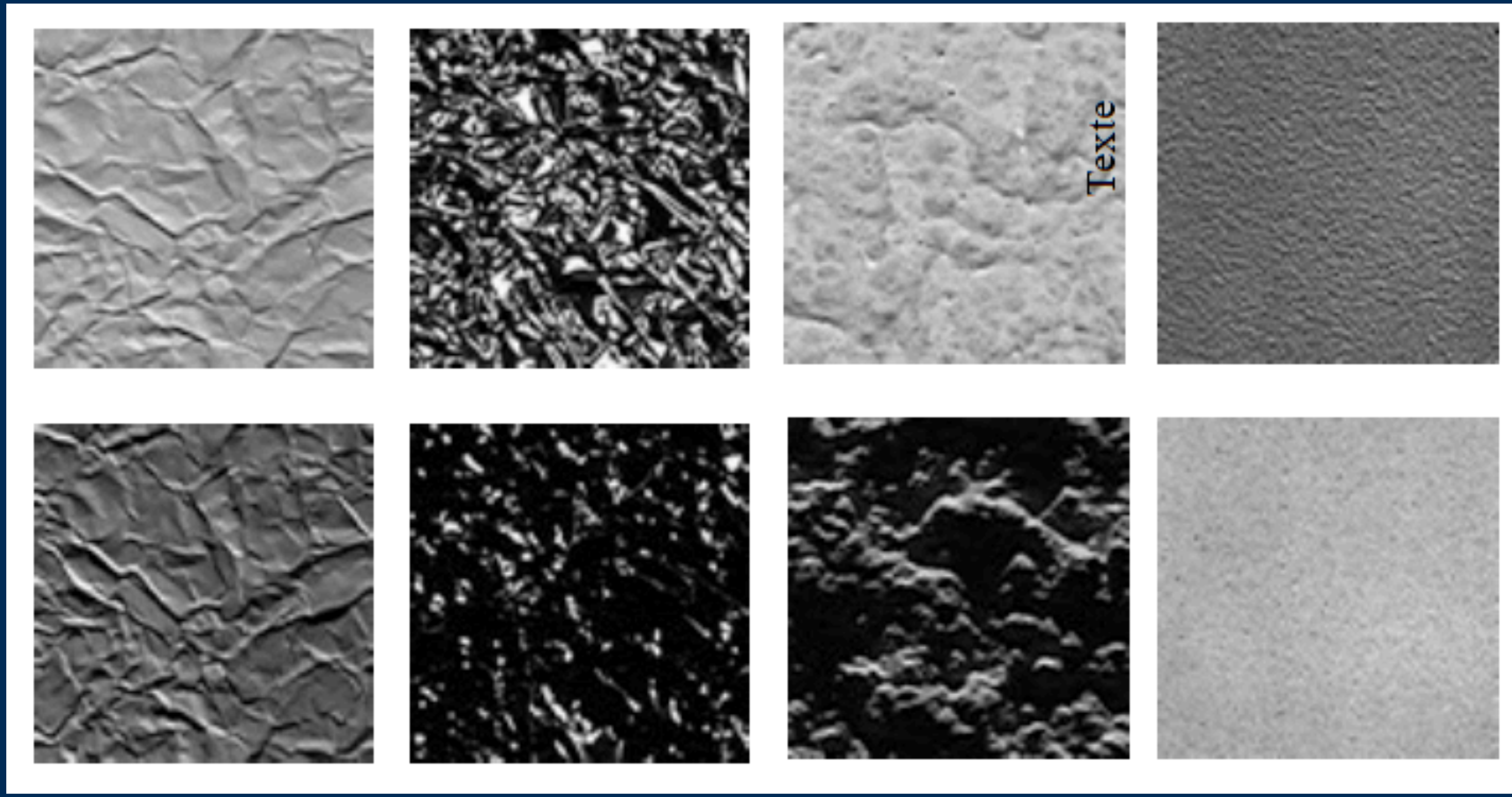
Cheng & Menard, in prep



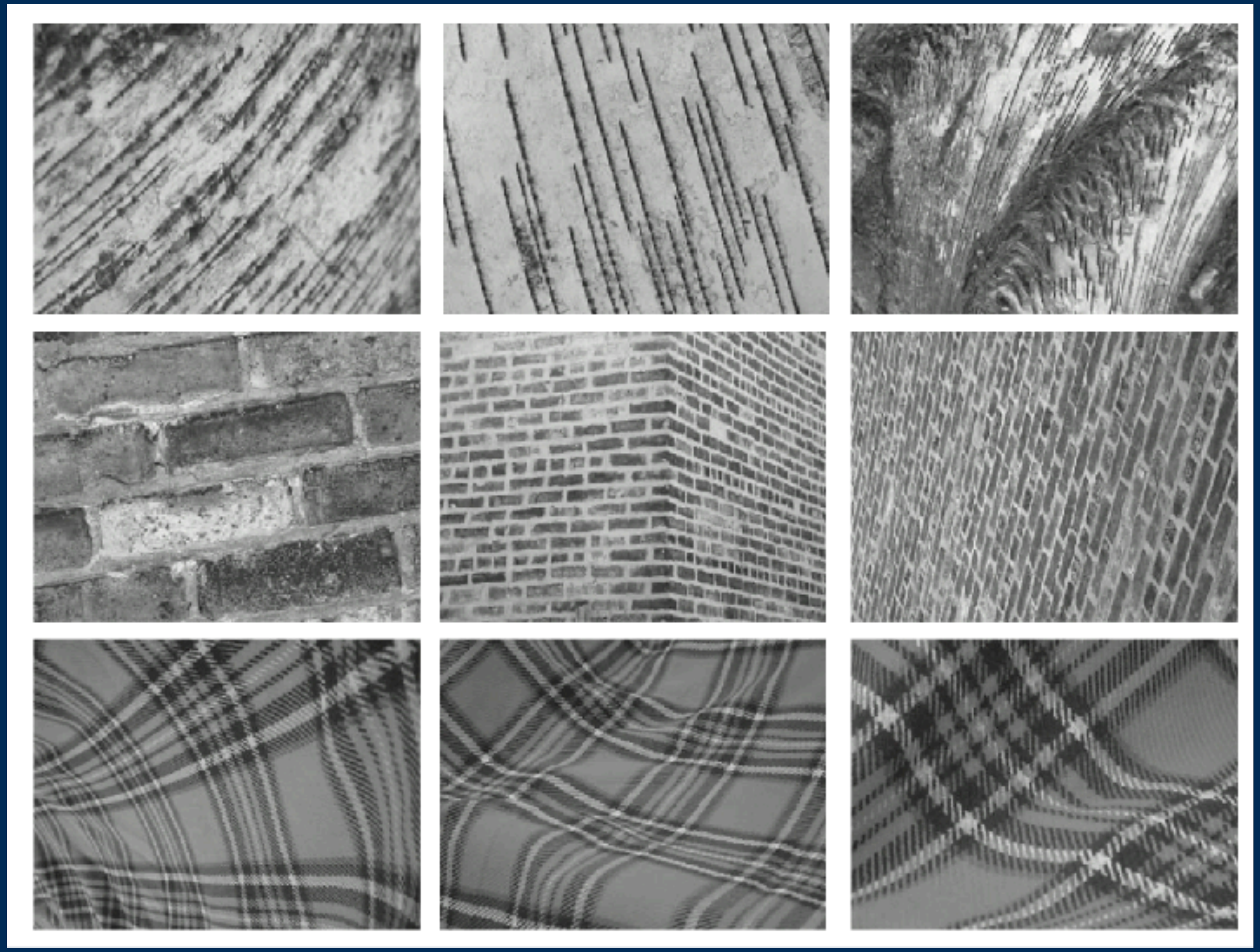
structure shape $s_{22} \equiv S_2^{\parallel} / S_2^{\perp}$
structure sparsity $s_{21} \equiv S_2 / S_1$

scattering transform in computer vision

CUREt database



UIUC database



Bruna & Mallat 2013

Sifre & Mallat 2013

gravitational lensing

image credit: NASA, ESA, and J. Lotz and the HFF Team (STScI)

lensing cosmology

Euclid space mission

Roman space telescope

Kilo-Degree Survey

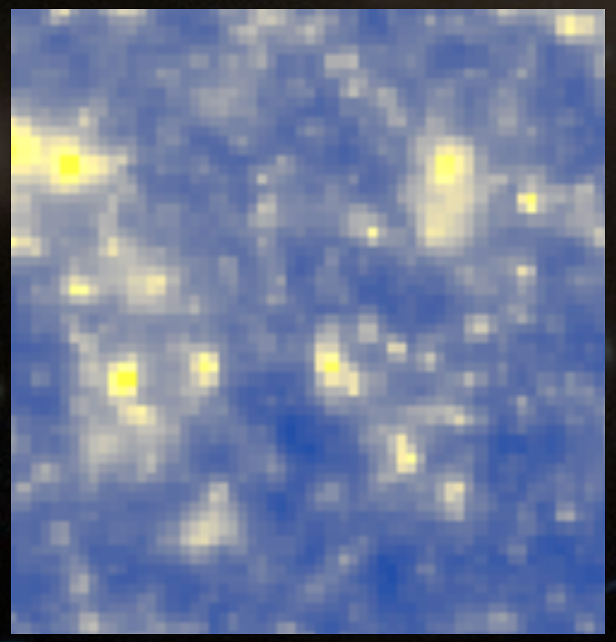
Dark Energy Survey

Subaru HSC Survey

Rubin observatory

raw data → galaxy catalog → mass map → properties of dark matter, dark energy, etc.

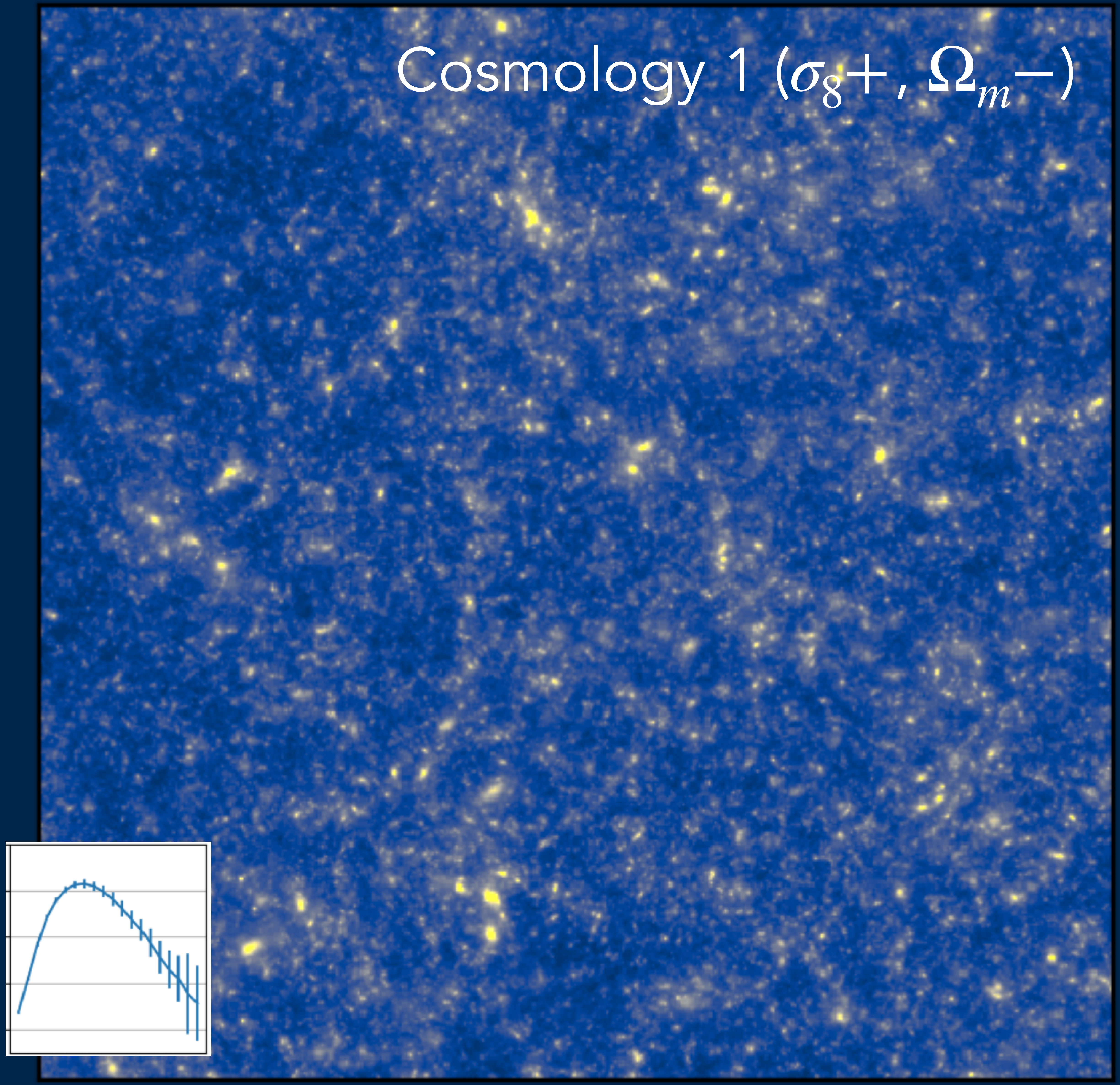
$$\begin{pmatrix} x_1, y_1, \epsilon_1 \\ x_2, y_2, \epsilon_2 \\ x_3, y_3, \epsilon_3 \\ x_4, y_4, \epsilon_4 \\ \cdots \end{pmatrix}$$

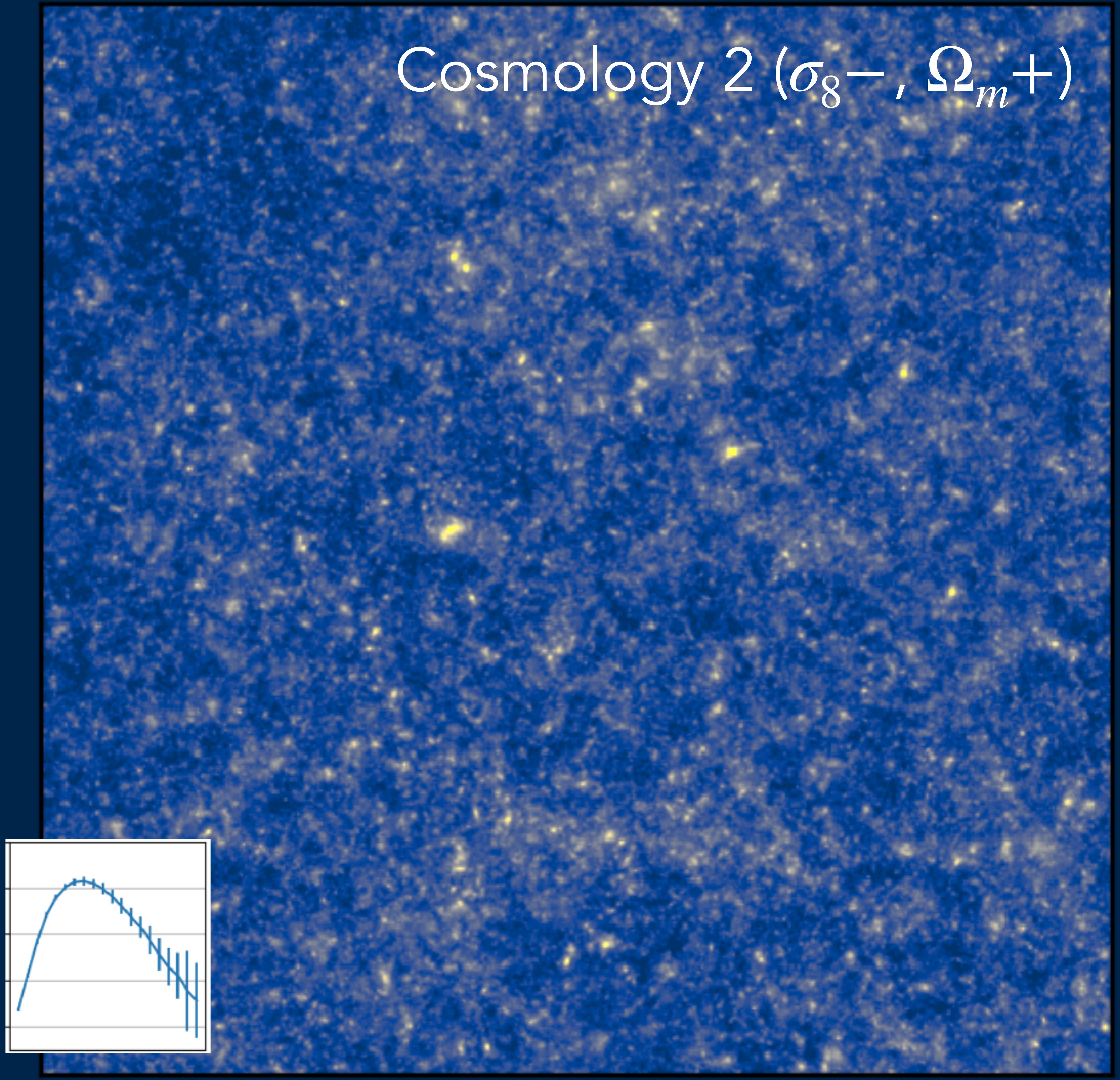


$(\sigma_8, \Omega_m, w, M_\nu)$

image credit: NASA, ESA, and J. Lotz and the HFF Team (STScI)

simulations (from Columbia lensing group)



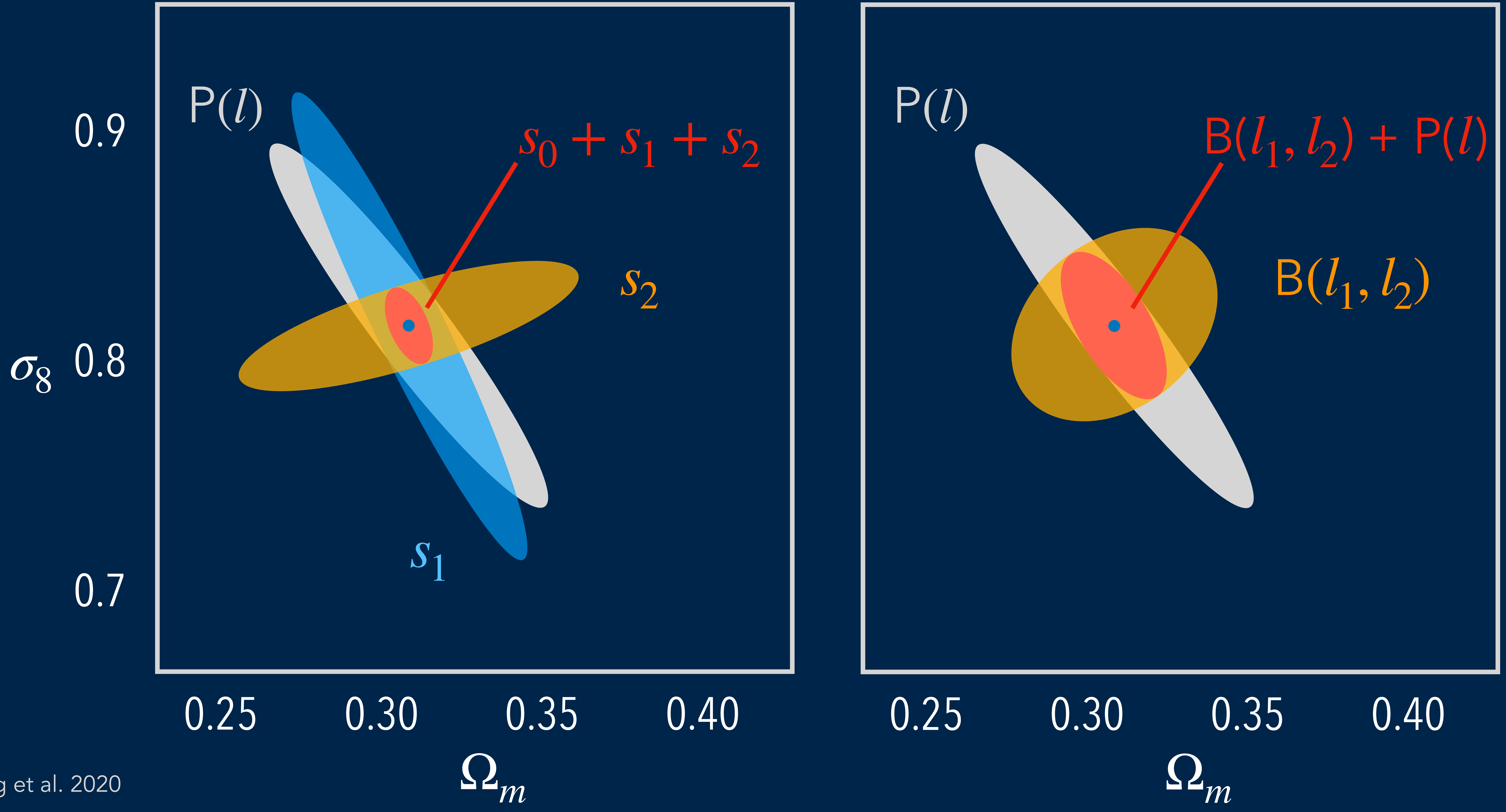


Matilla Zorrilla et al. 2016, Gupta et al. 2018 (Columbia lensing group)

inferring cosmological parameters

arXiv: 2006.08561

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

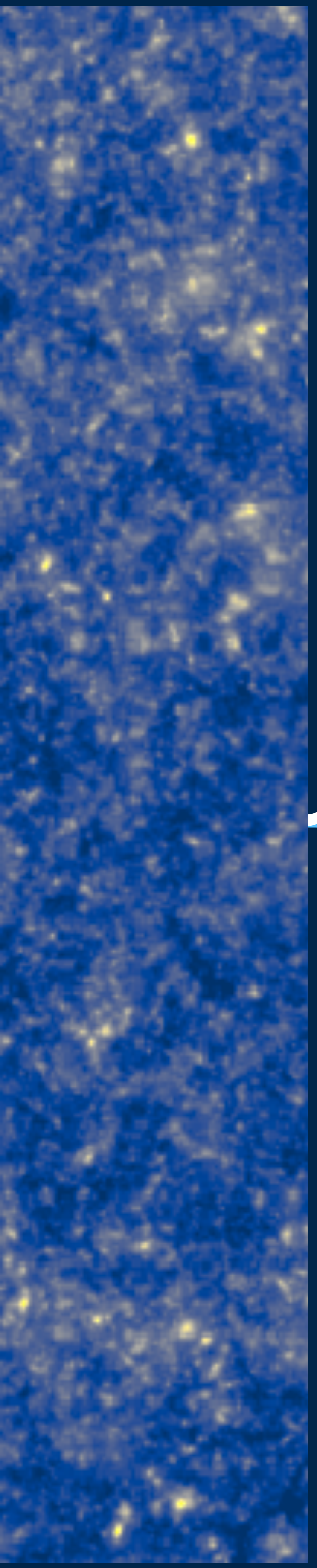


Cheng et al. 2020

inferring cosmological parameters

arXiv: 2006.08561

scale range: 1 arcmin to 3.5 deg



$P(l)$

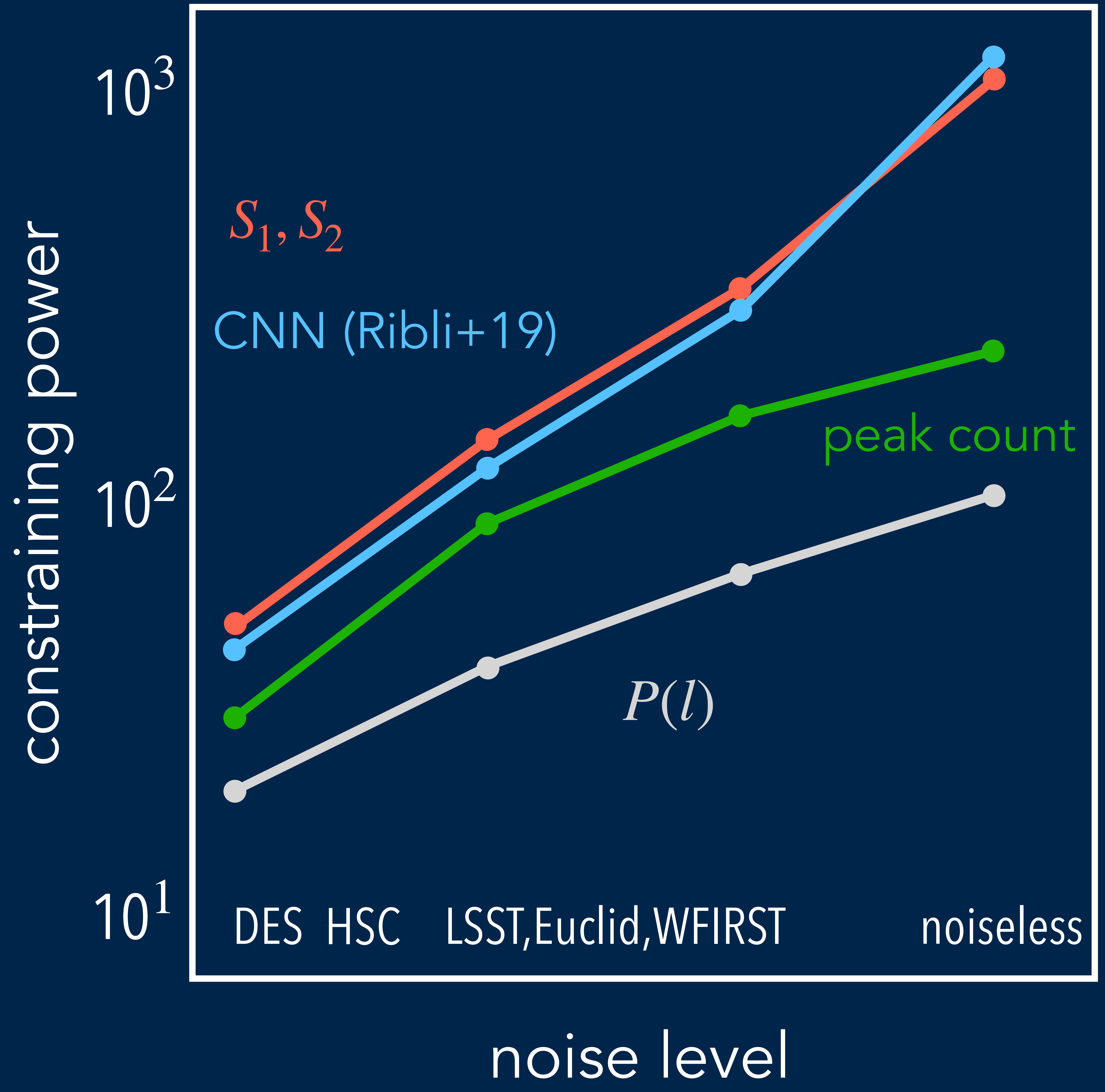
~20 coefficients

scattering coefficients

37 coefficients

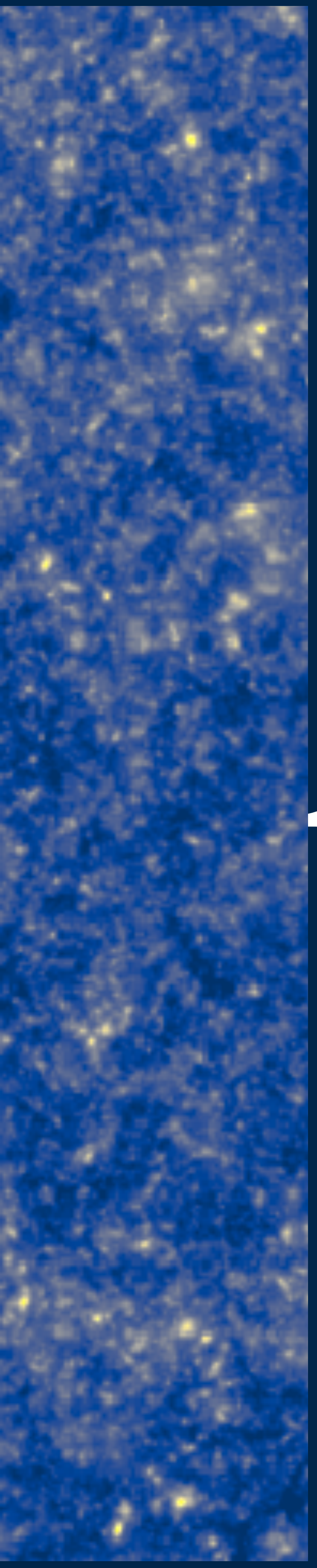
CNN

millions



Cheng et al. 2020

inferring cosmological parameters



$P(l)$

~20 coefficients

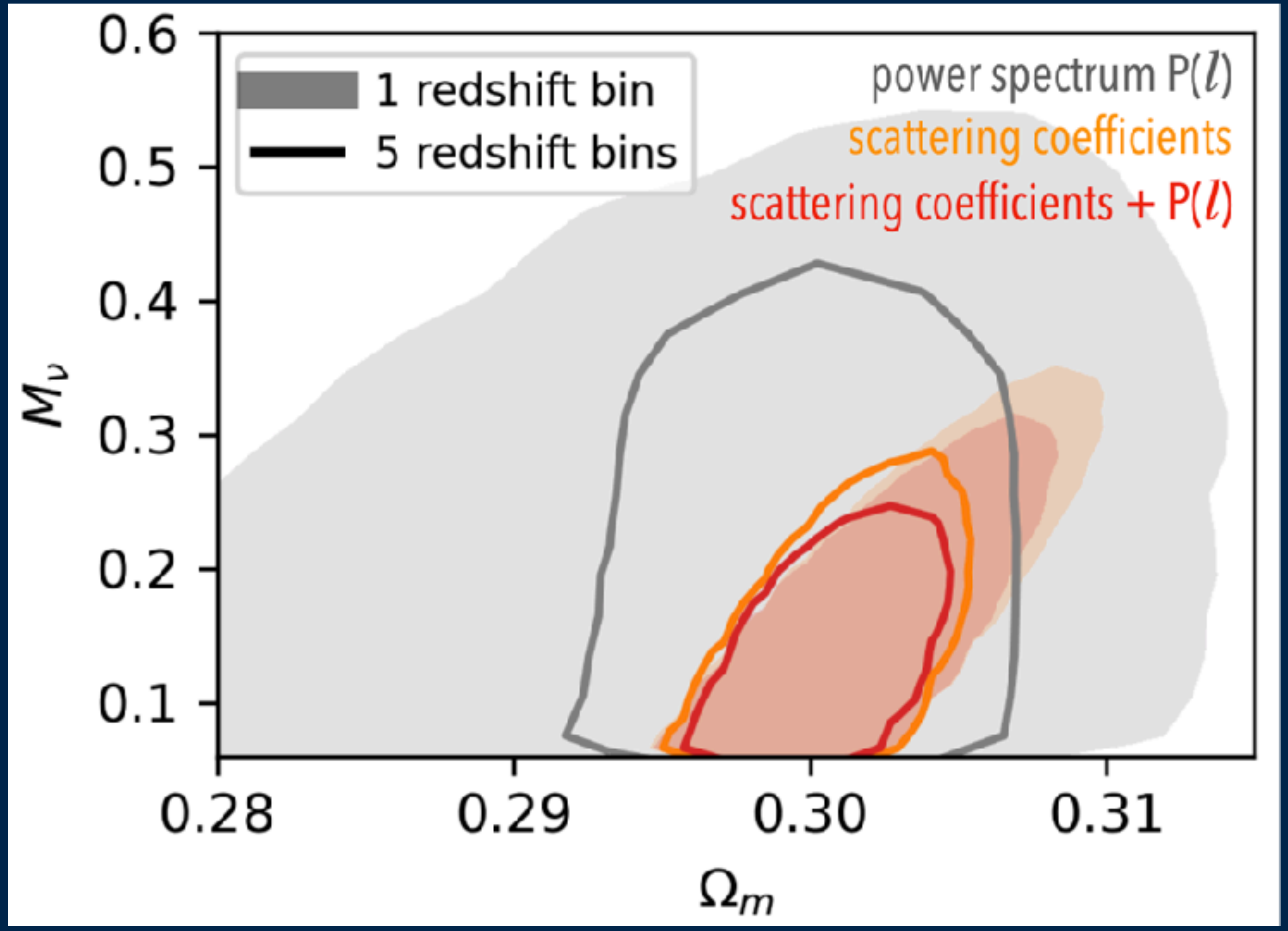
scattering coefficients

37 coefficients

CNN

millions

Cheng & Menard 2021

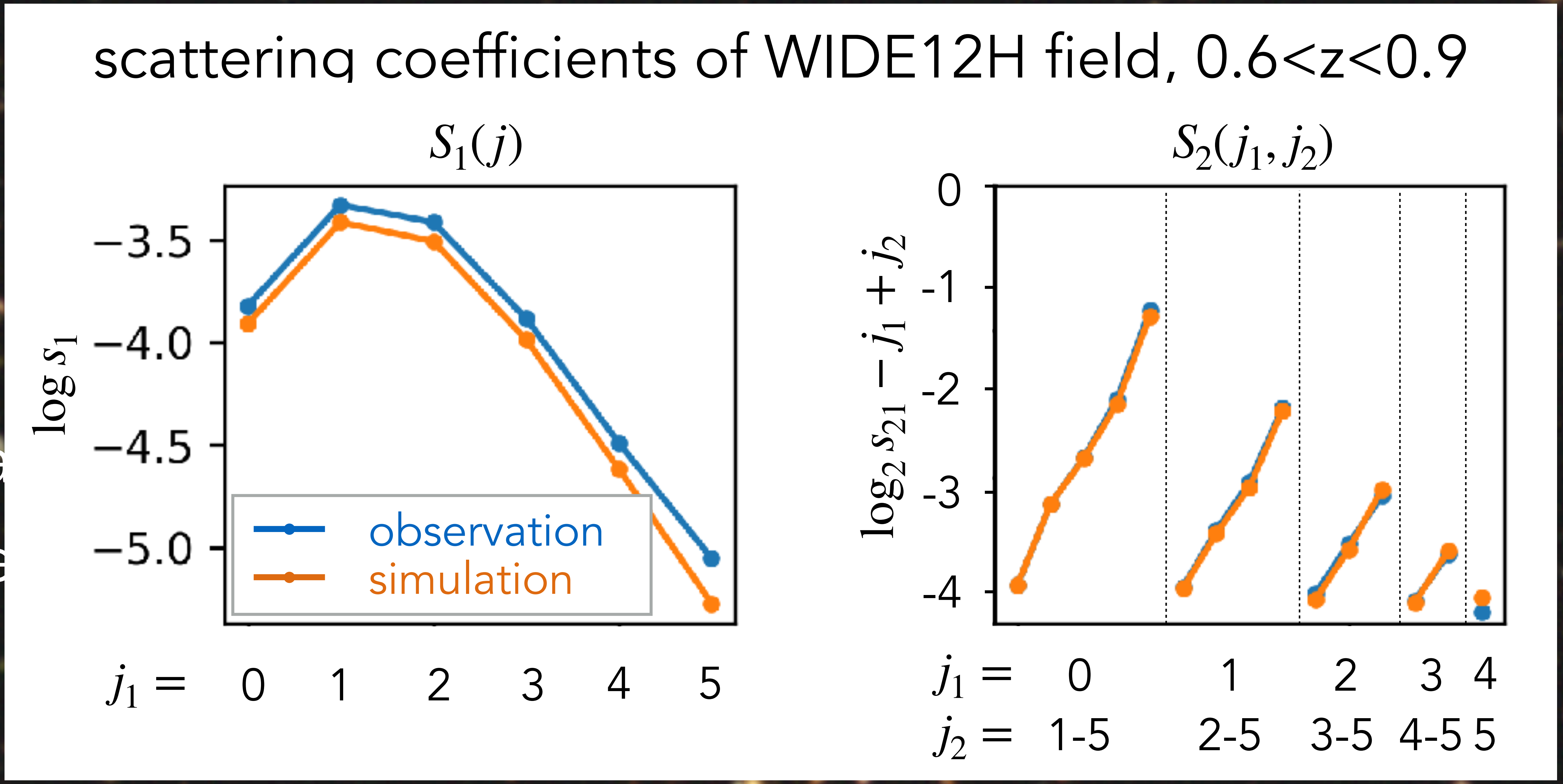


similar results for
other cosmological parameters
(w_0, w_a, M_ν)

towards real data

HSC survey (data release 1)

- Subaru 8.2m, ~0”.6 seeing
- 26 mag (1 part of 100 million of huma
- 137 deg^2 , 90 nights, 9 million galaxie
- 108 full-sky + 100 simulations
- blinding

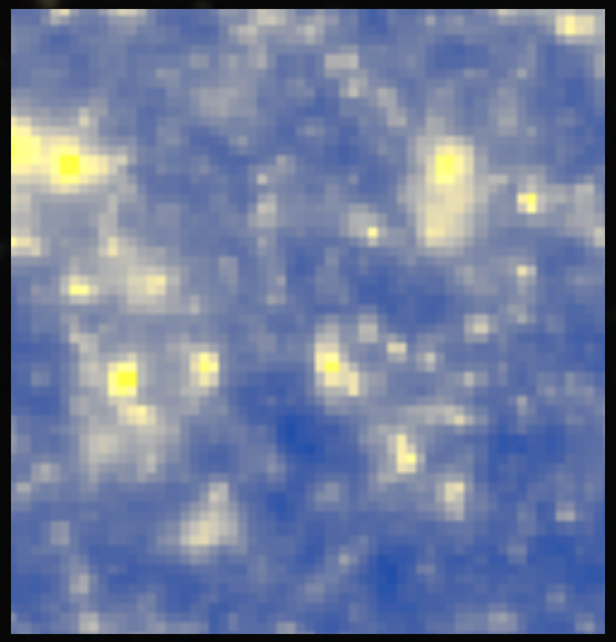


Cheng et al. in prep

first real constraint in a few months!

raw data ⟶ galaxy catalog ⟶ mass map ⟶ properties of dark matter, dark energy, etc. $(\sigma_8, \Omega_m, w, M_\nu)$

$$\begin{pmatrix} x_1, y_1, \epsilon_1 \\ x_2, y_2, \epsilon_2 \\ x_3, y_3, \epsilon_3 \\ x_4, y_4, \epsilon_4 \\ \cdots \end{pmatrix}$$



Euclid space mission

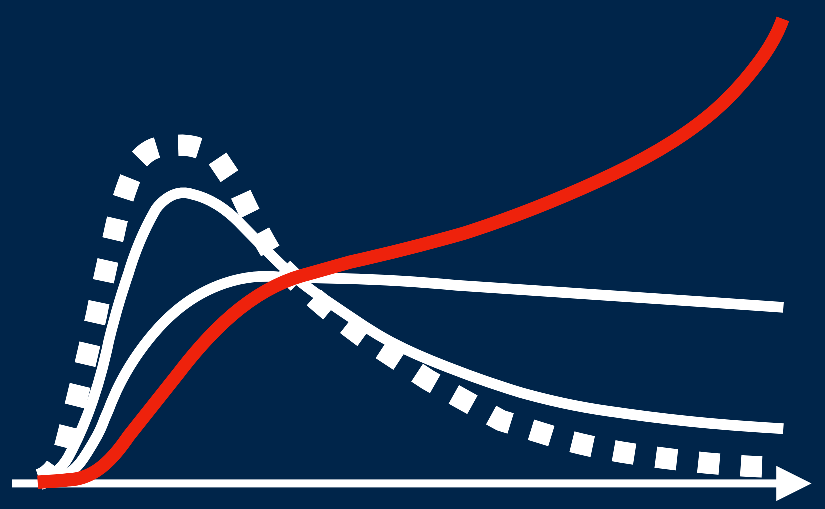
- half of the sky
- billions of galaxies
- Hubble-quality images

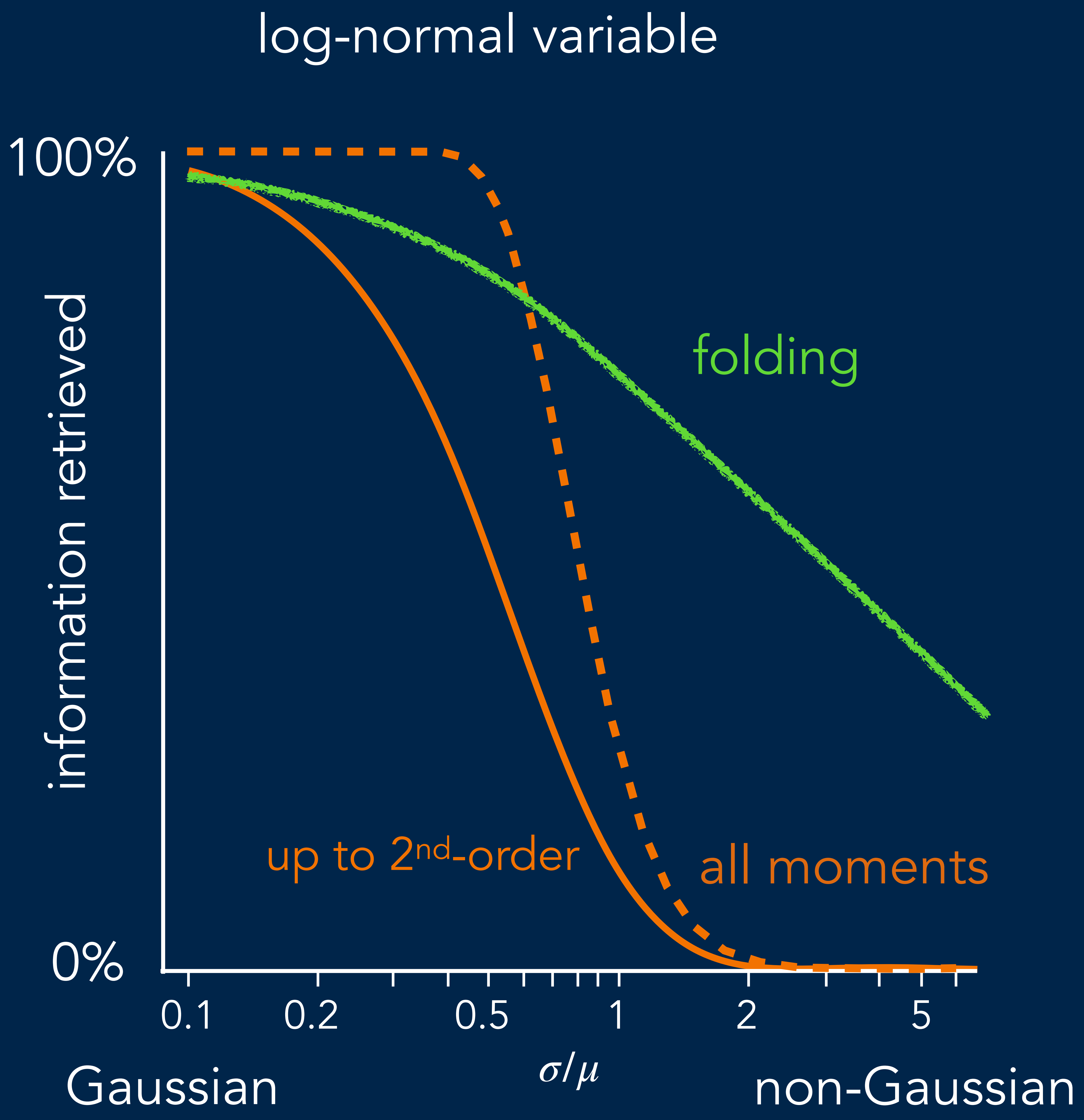
image credit: NAOJ, HSC-SSP

one-variable illustration

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

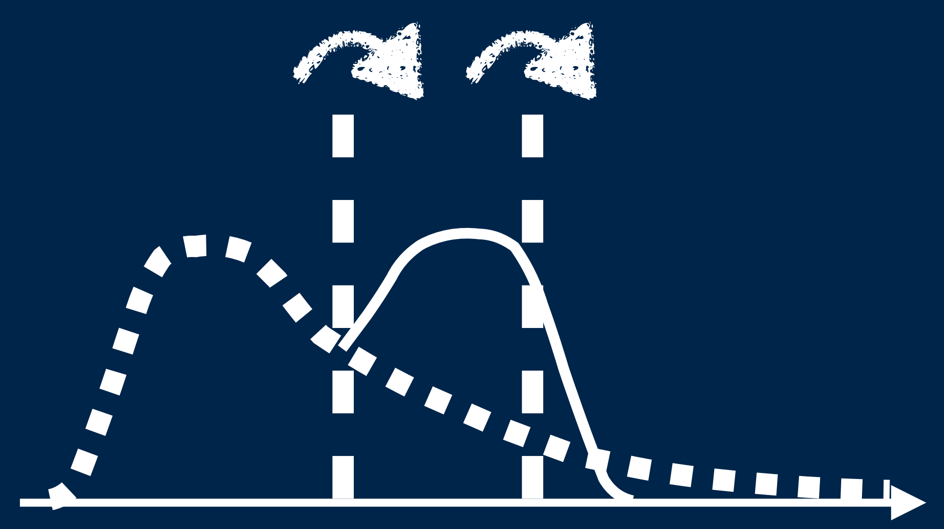
amplifying the tail





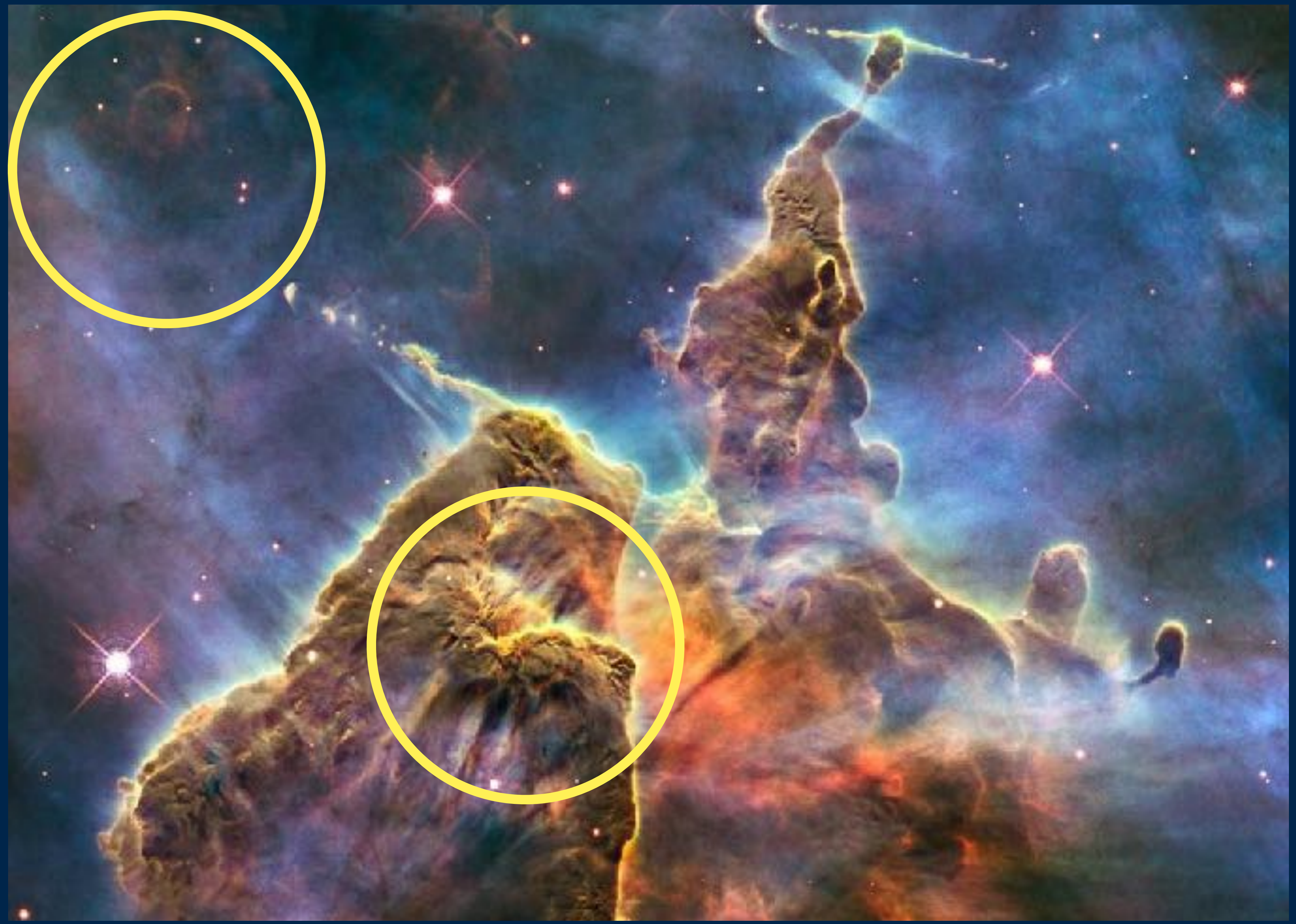
$\langle || I \star \psi | \star \psi | \rangle$

folding the core



Cheng & Menard, in prep

extensions of the scattering transform



local translation invariance $S_1(x) = \left| I \star \psi \right| \star \phi$

rotation invariance

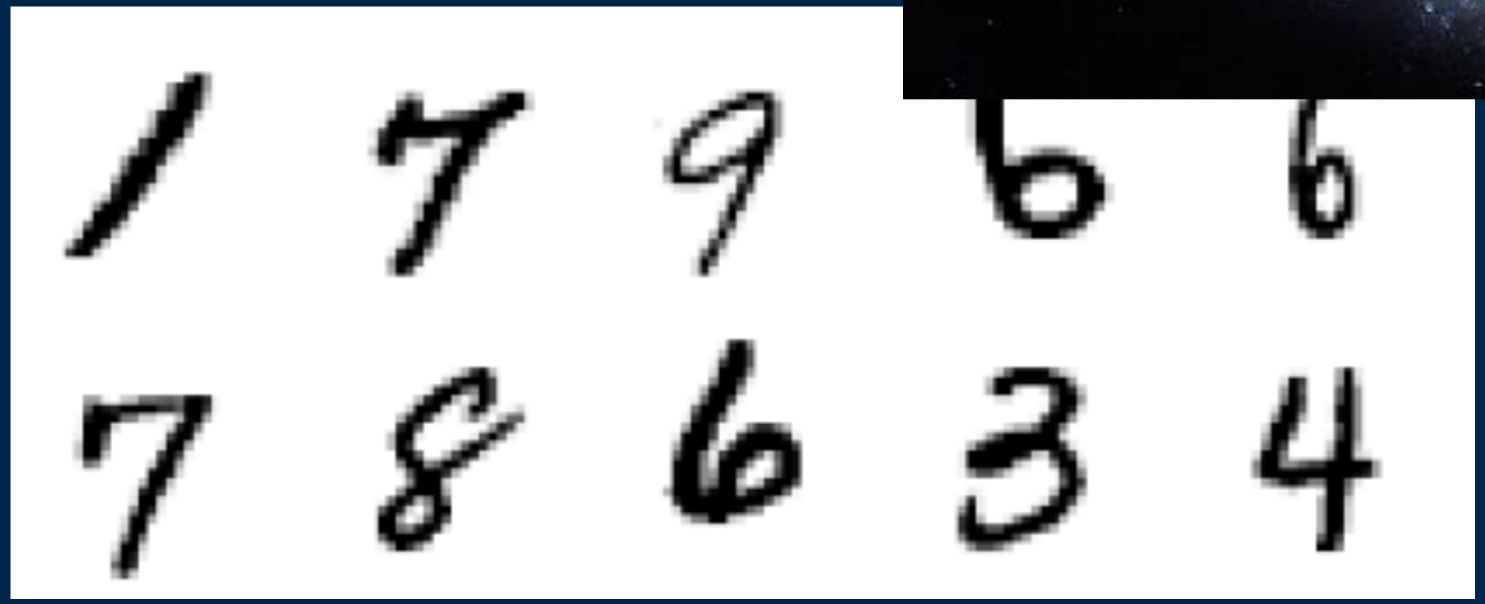
Sifre & Mallat 2013

learning ability

e.g., Zarka, Guth, & Mallat 2020

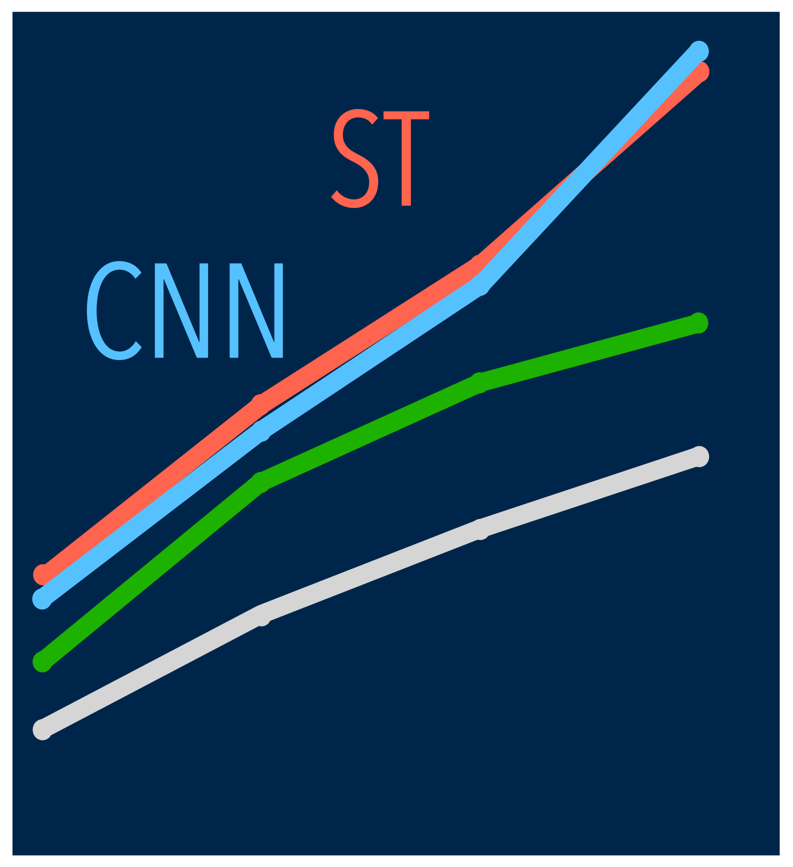
cross correlation: phase harmonics

Mallat, Zhang, & Rochette 2018, Allay et al 2020



Bruna & Mallat 2013

How do we characterize a field?



apparent "complexity"

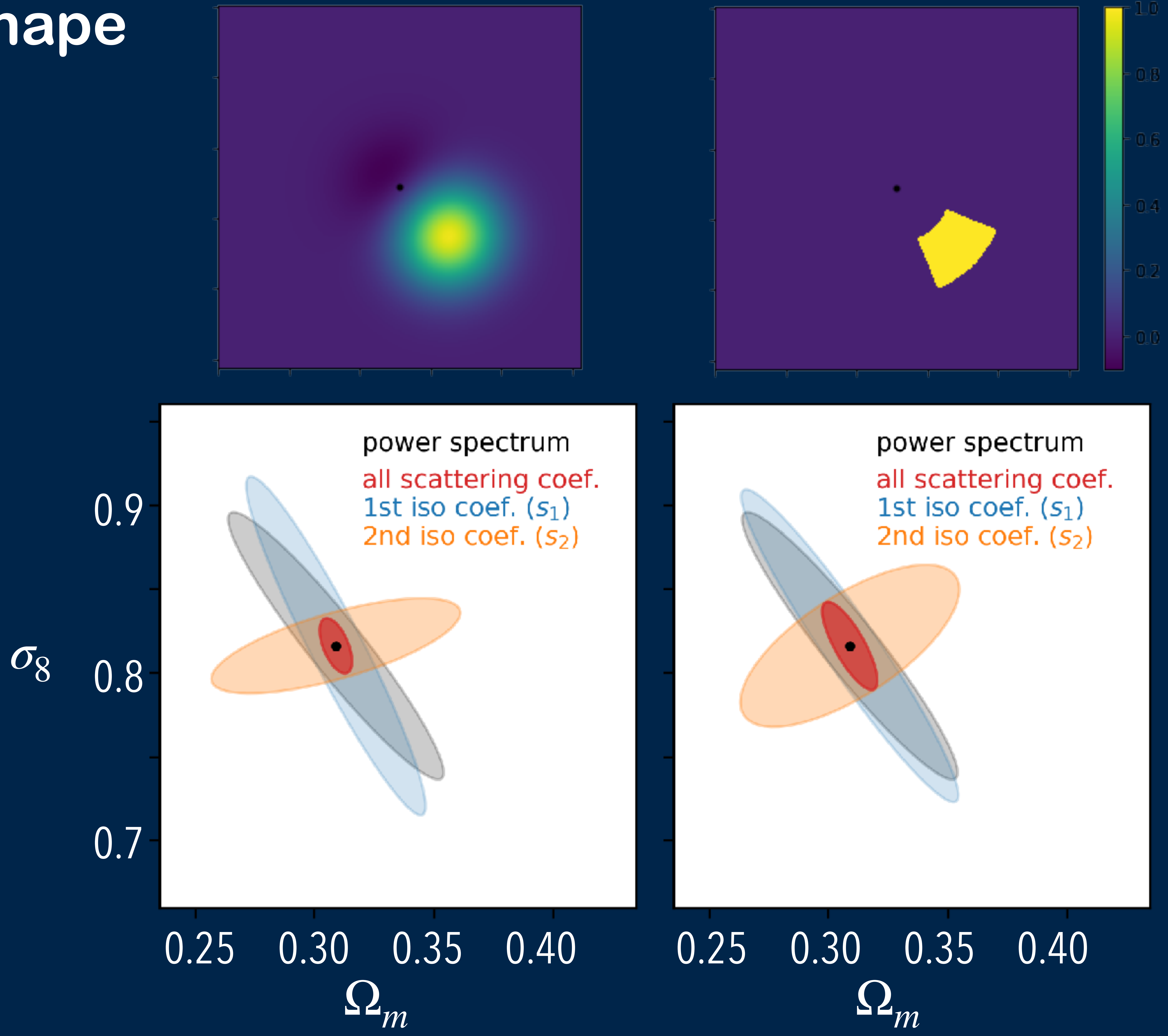
power spectrum

scattering transform

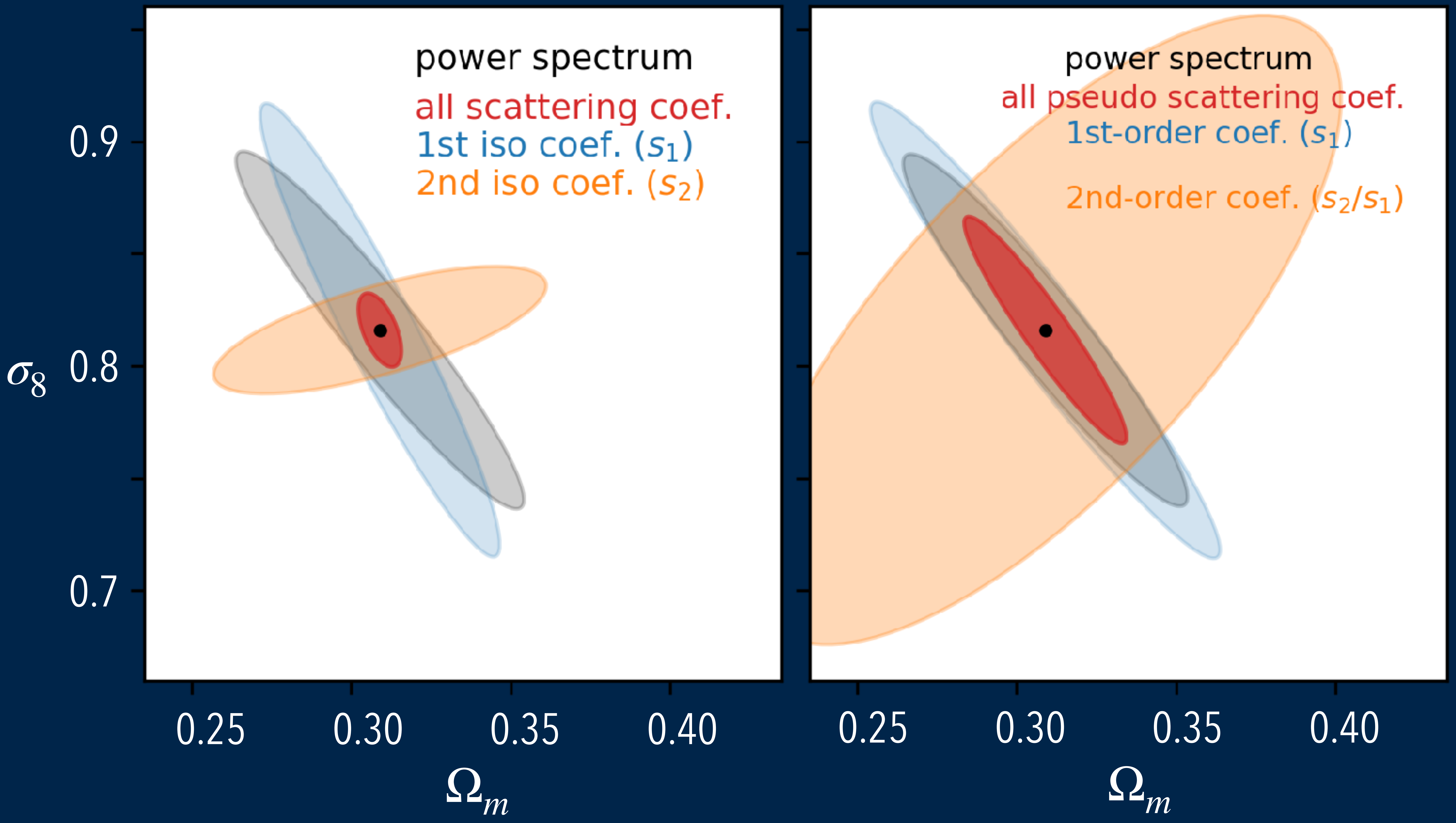
CNN

information

wavelet shape



tri-spectrum



modulus —> modulus squared: ST becomes tri-spectrum