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

Speakers: Sihao Cheng

Series: Cosmology & Gravitation

Date: December 13, 2021 - 11:00 AM

URL: https://pirsa.org/21120020

Abstract: Patterns and complex textures are ubiquitous in astronomical data but challenging to quantify. I will present a new powerful statistic called the "scattering transform". It borrows ideas from convolutional neural nets (CNNs) while retaining the advantages of traditional statistics. As an example, I will show its application to weak lensing cosmology, where it outperforms classic statistics and is on a par with CNNs. I will also show interesting visual interpretations of the scattering statistics and possible extensions of this "mathematical neural network" idea. I argue that the scattering transform provides a powerful new approach in cosmology and beyond.

Related papers:

https://arxiv.org/abs/2112.01288

https://arxiv.org/abs/2103.09247

https://arxiv.org/abs/2006.08561

Zoom Link: https://pitp.zoom.us/j/91612161747?pwd=bnQrVmo4ZjBjaUdQMDBNZGhFS2NPQT09

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

Sihao Cheng (程思浩)
Johns Hopkins University
École Normale Supérieure

arXiv: 2006.08561
arXiv: 2103.09247
arXiv: 2112.01288

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

Perimeter Institute
December 13th, 2021

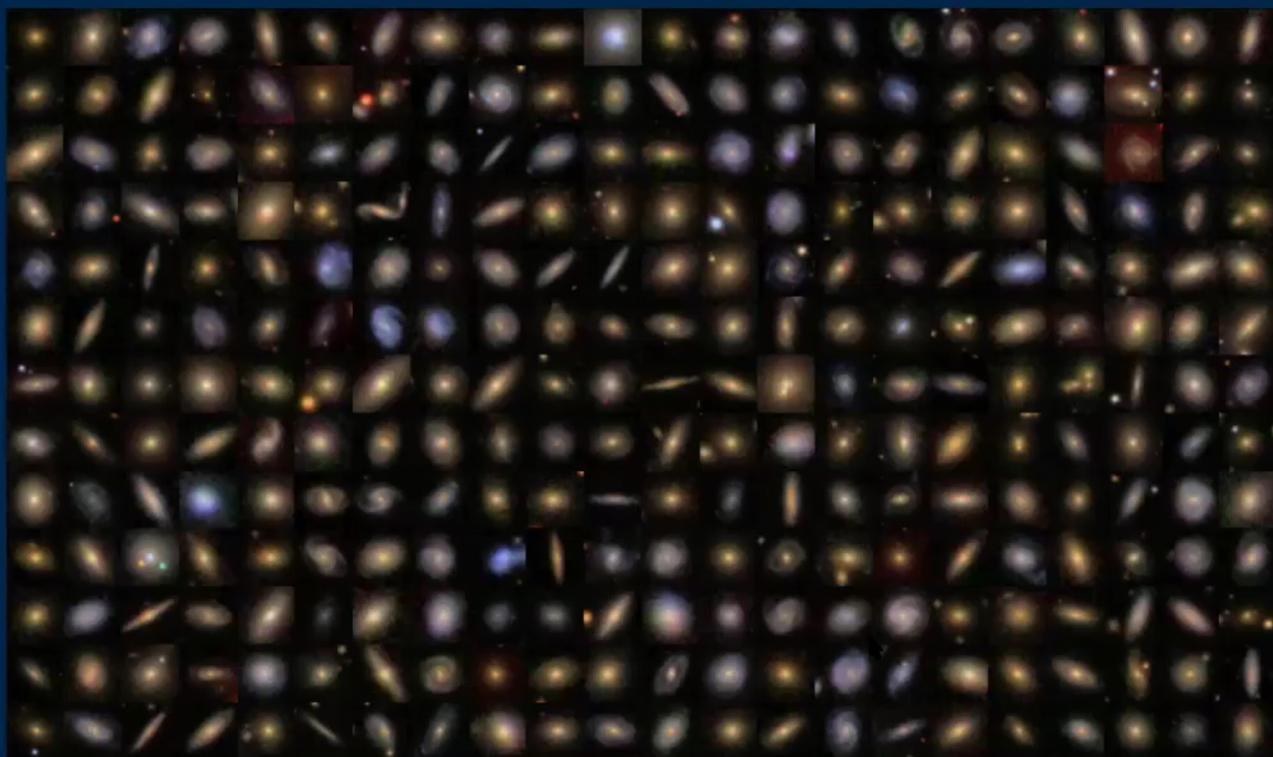
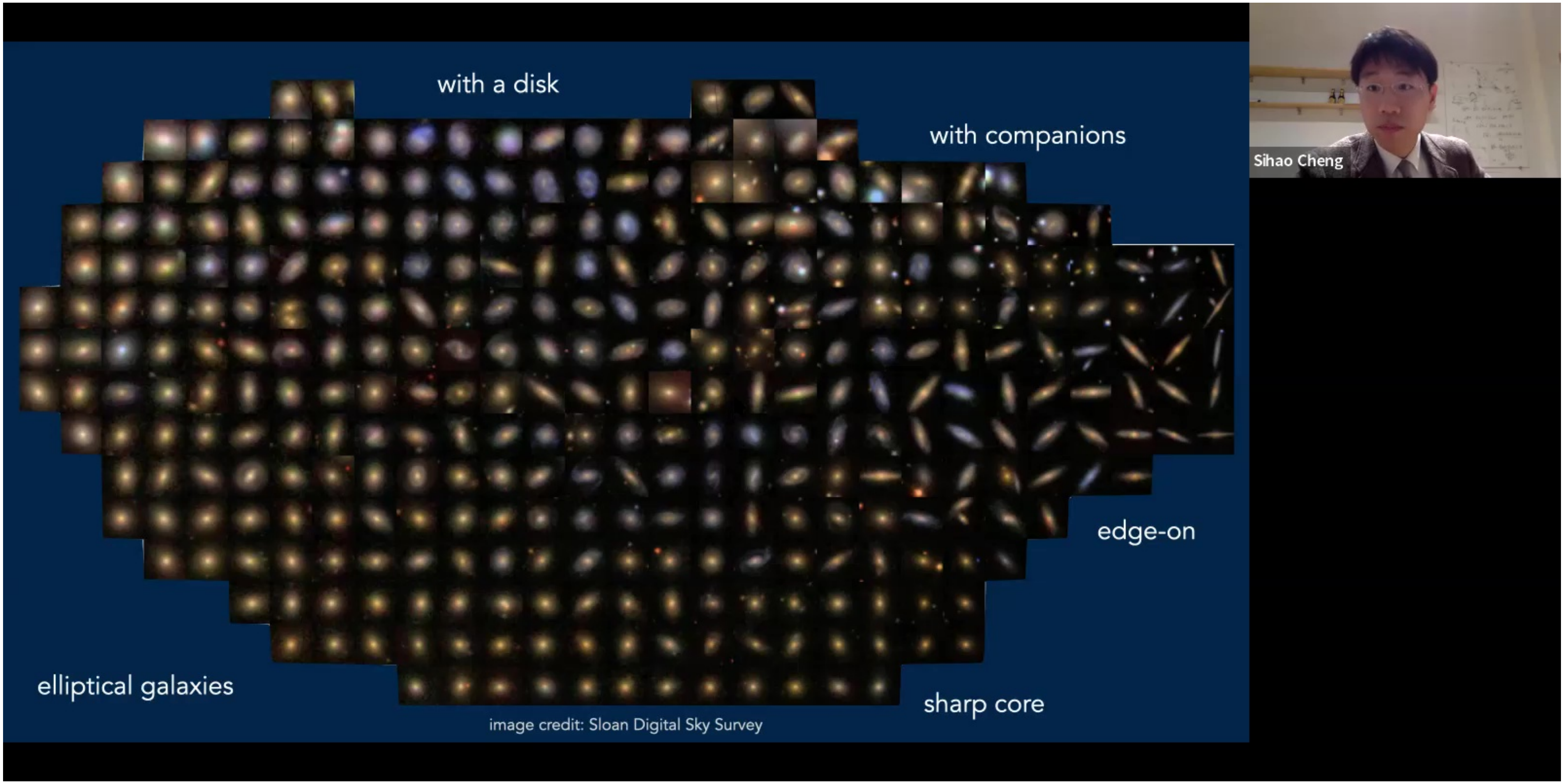
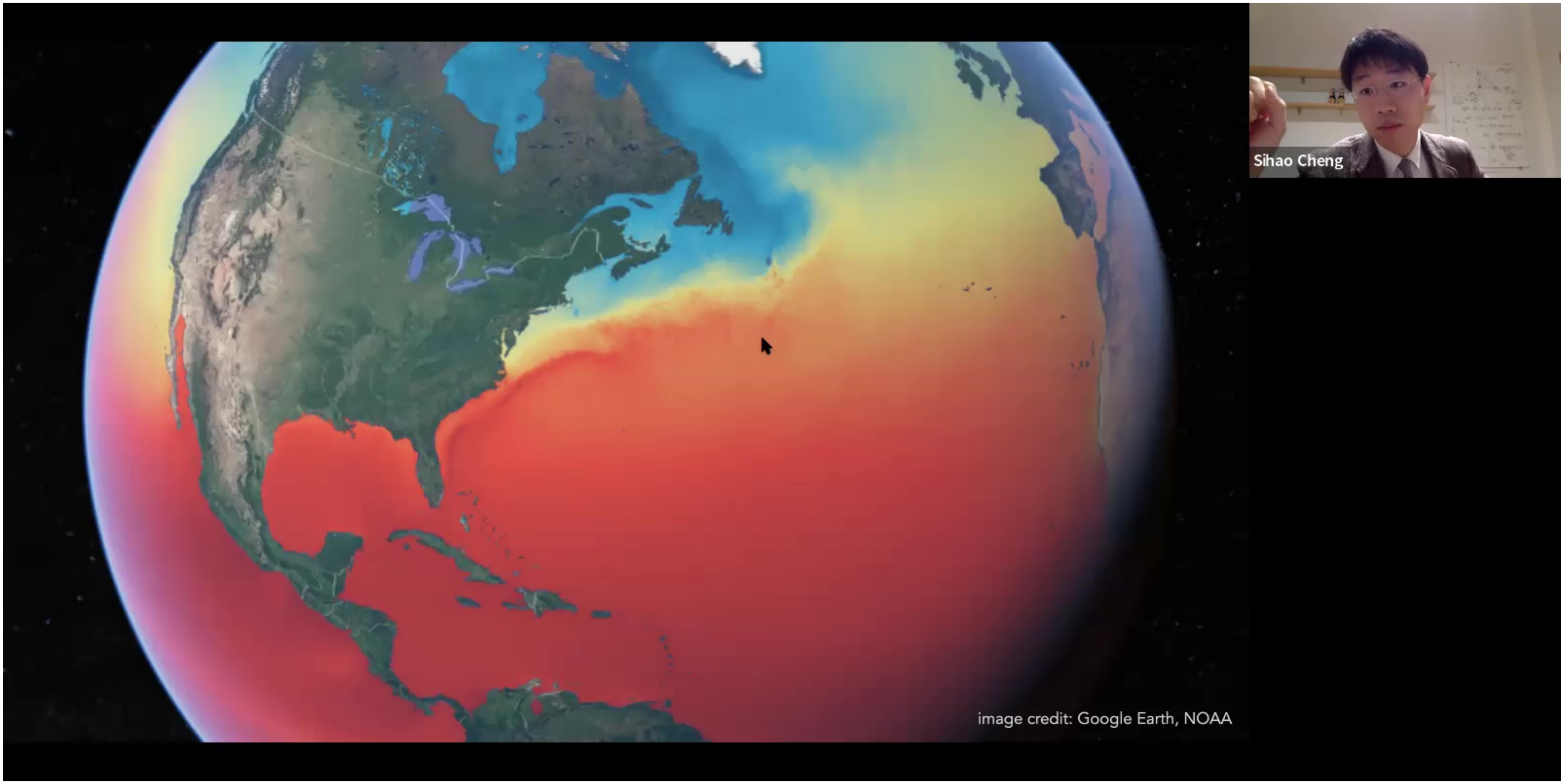
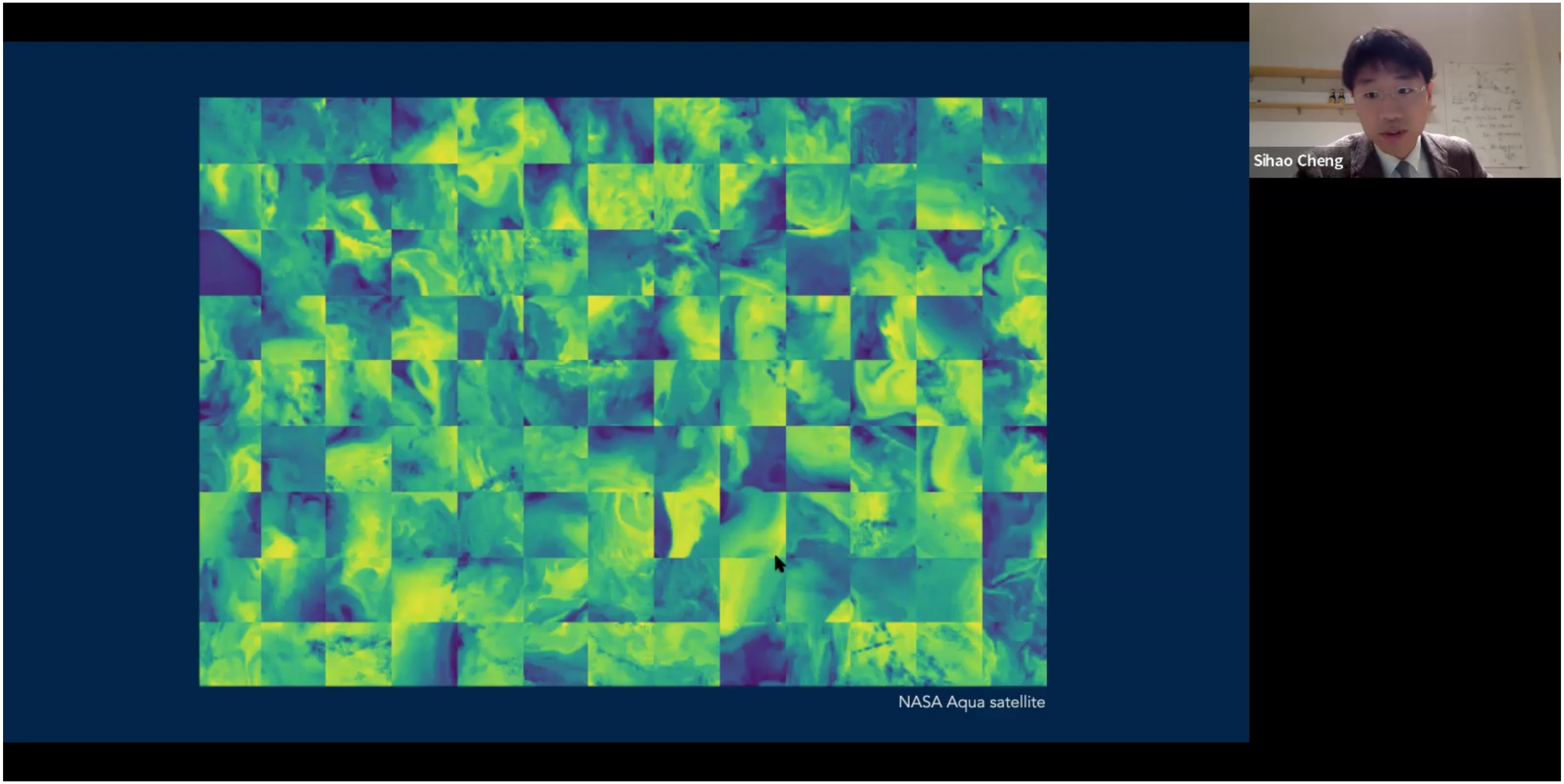
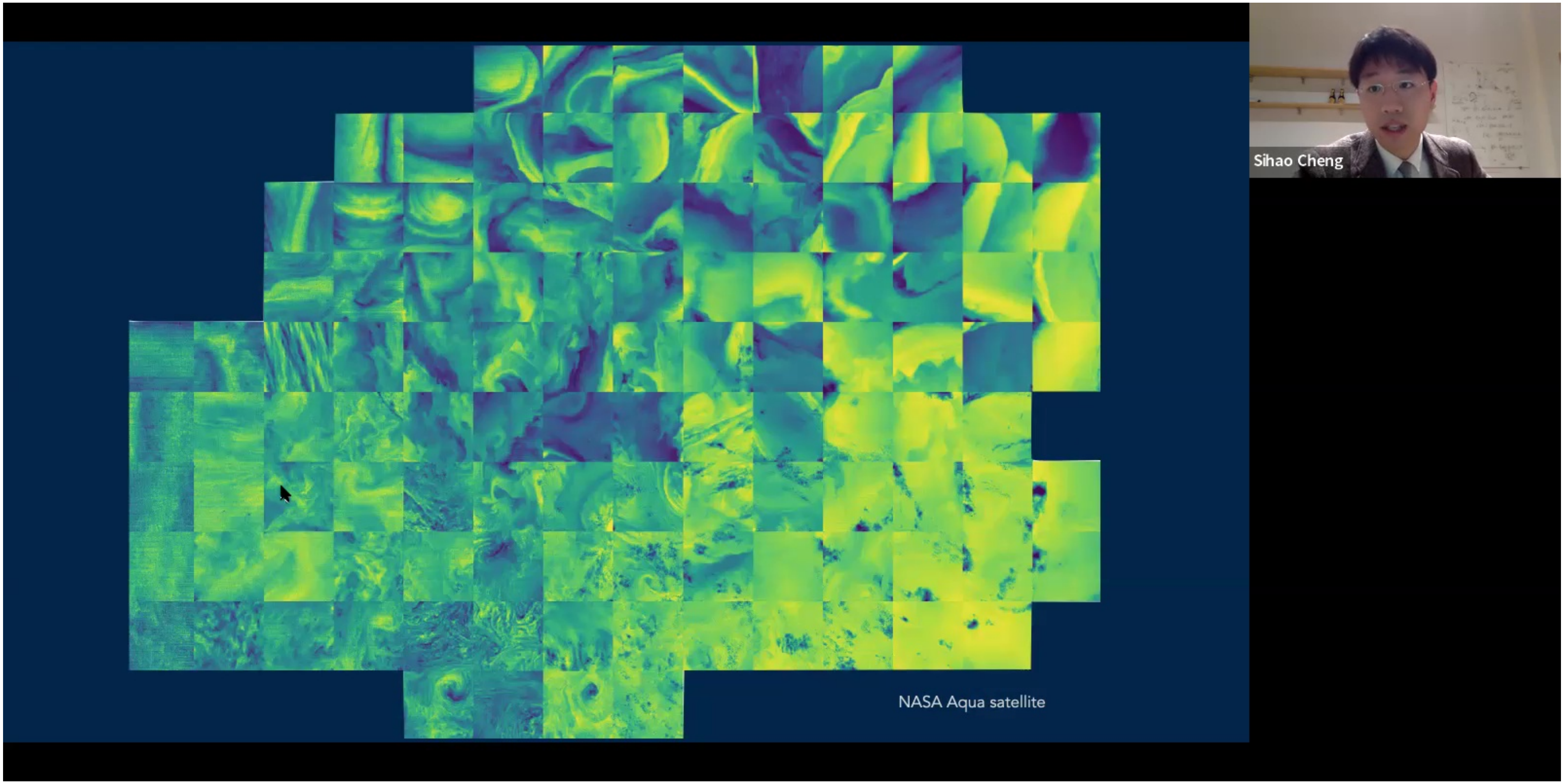


image credit: Sloan Digital Sky Survey

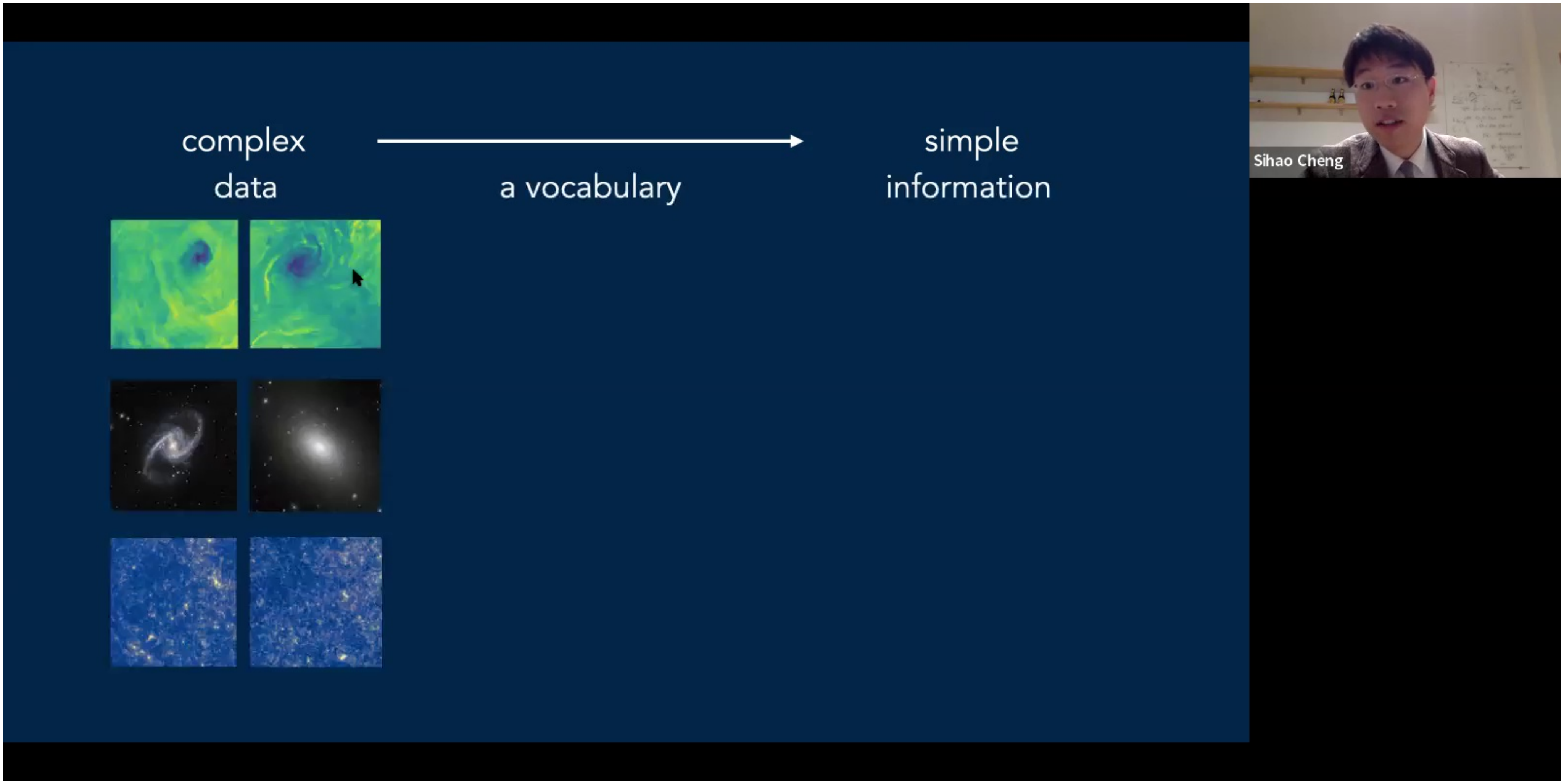

with a disk
with companions
edge-on
sharp core
elliptical galaxies
image credit: Sloan Digital Sky Survey
Sihao Cheng


Sihao Cheng
image credit: Google Earth, NOAA

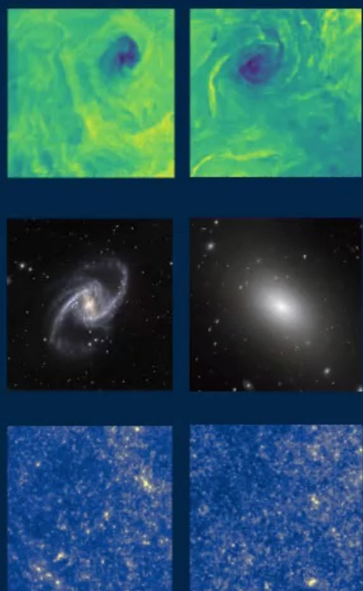

NASA Aqua satellite
Sihao Cheng


NASA Aqua satellite
Sihao Cheng


complex
simple
Sihao Cheng


complex
data
a vocabulary
simple
information
Sihao Cheng

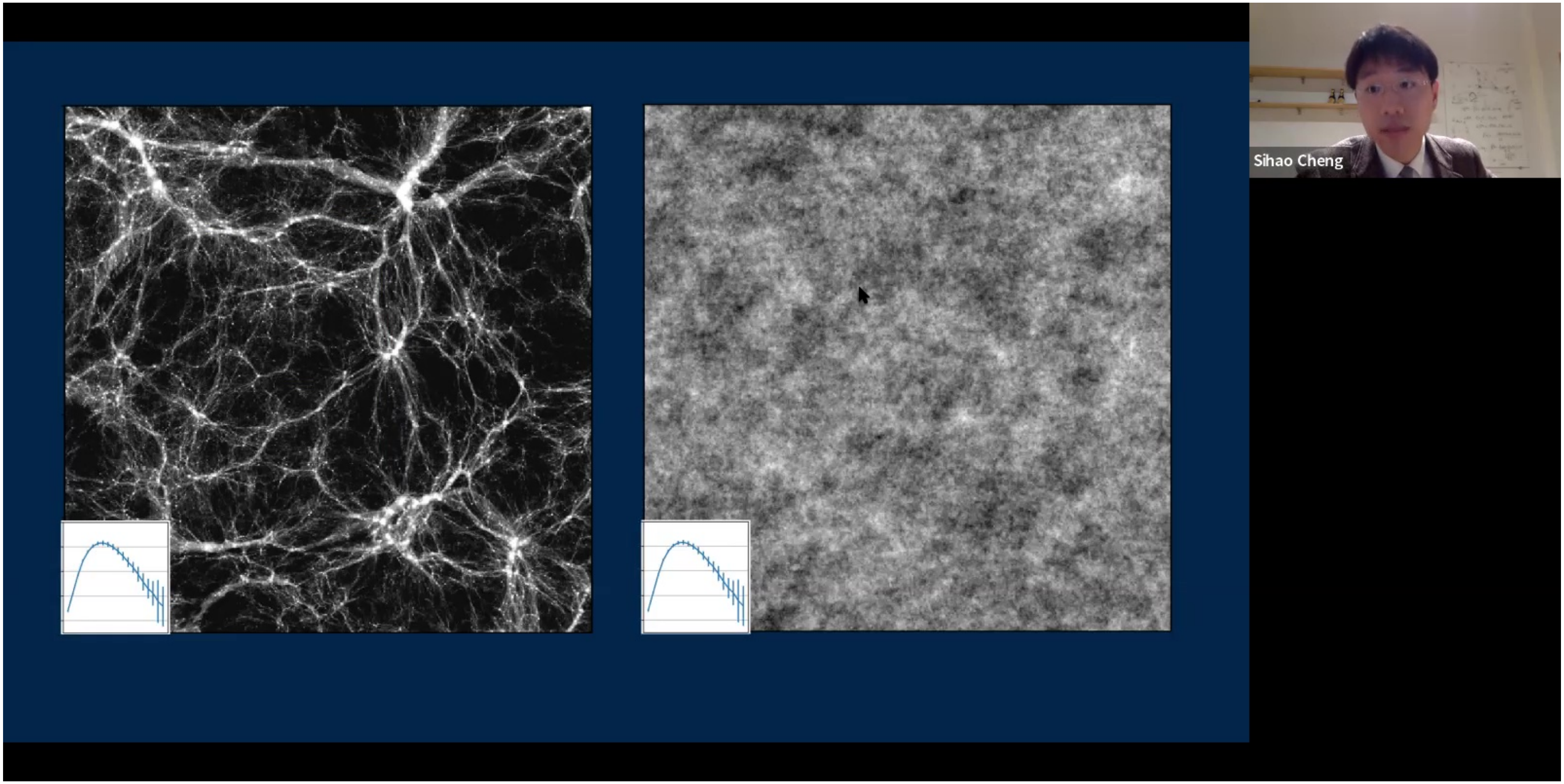
complex data → a vocabulary → simple information

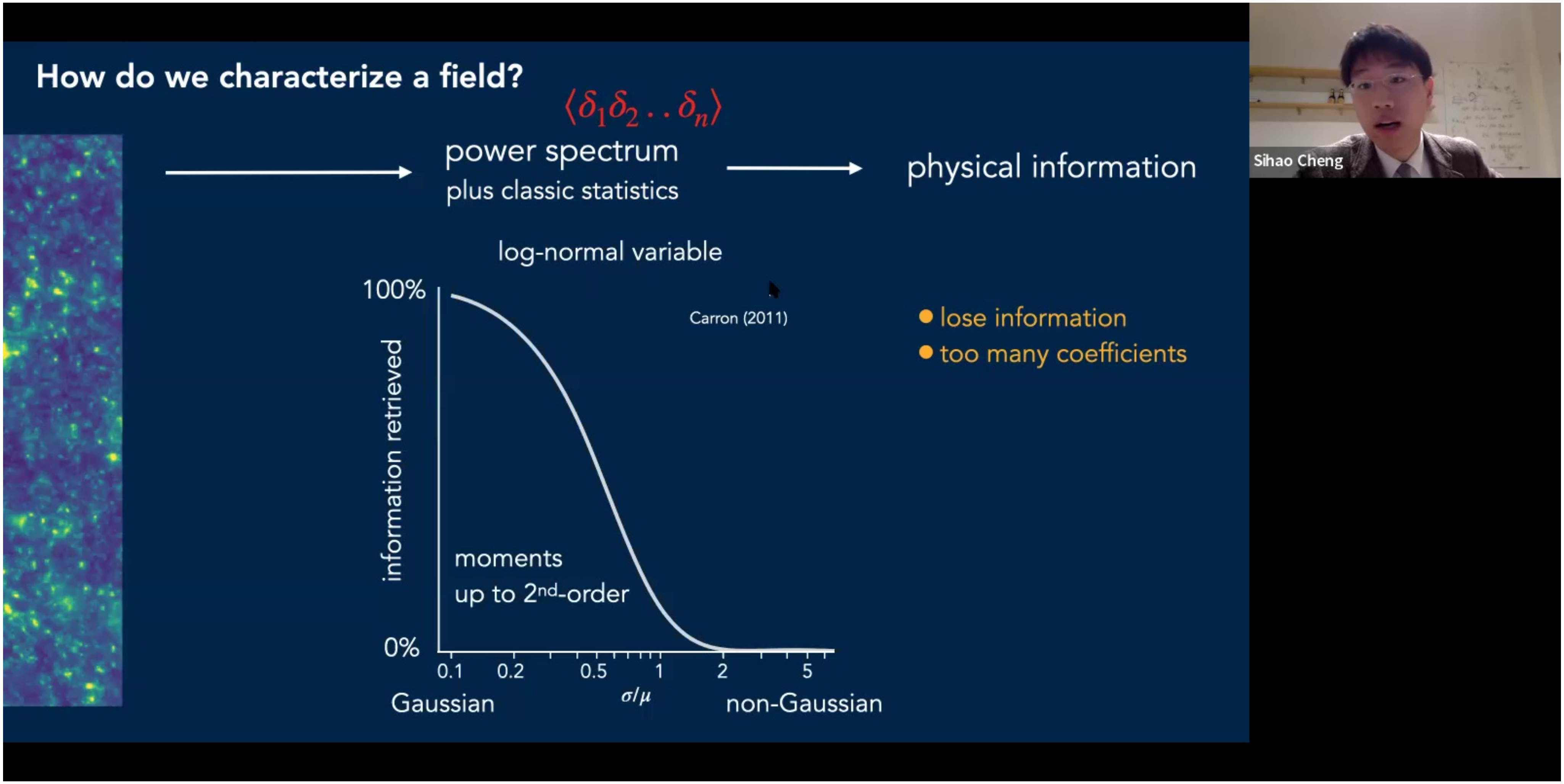


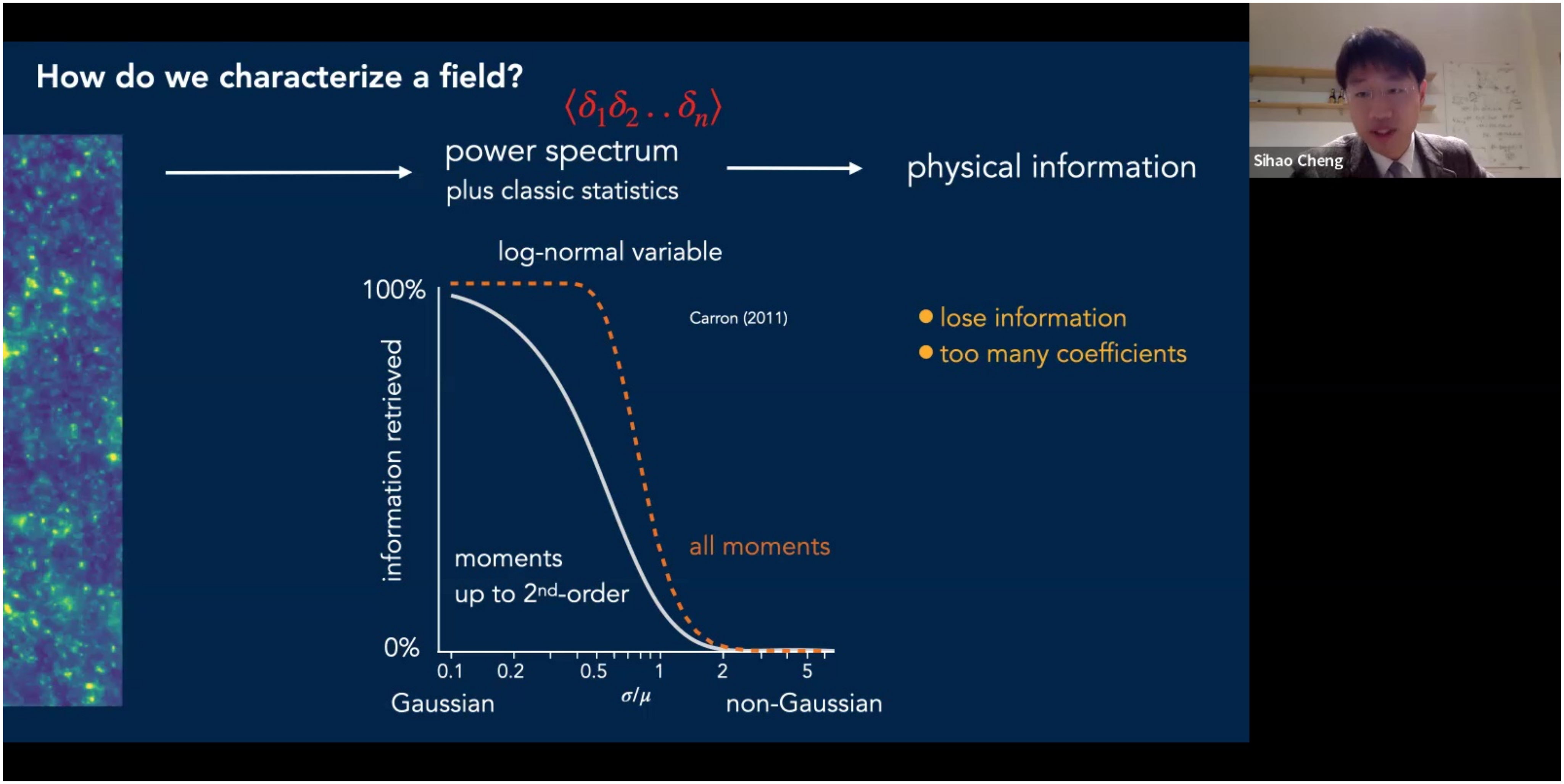
data exploration
no model

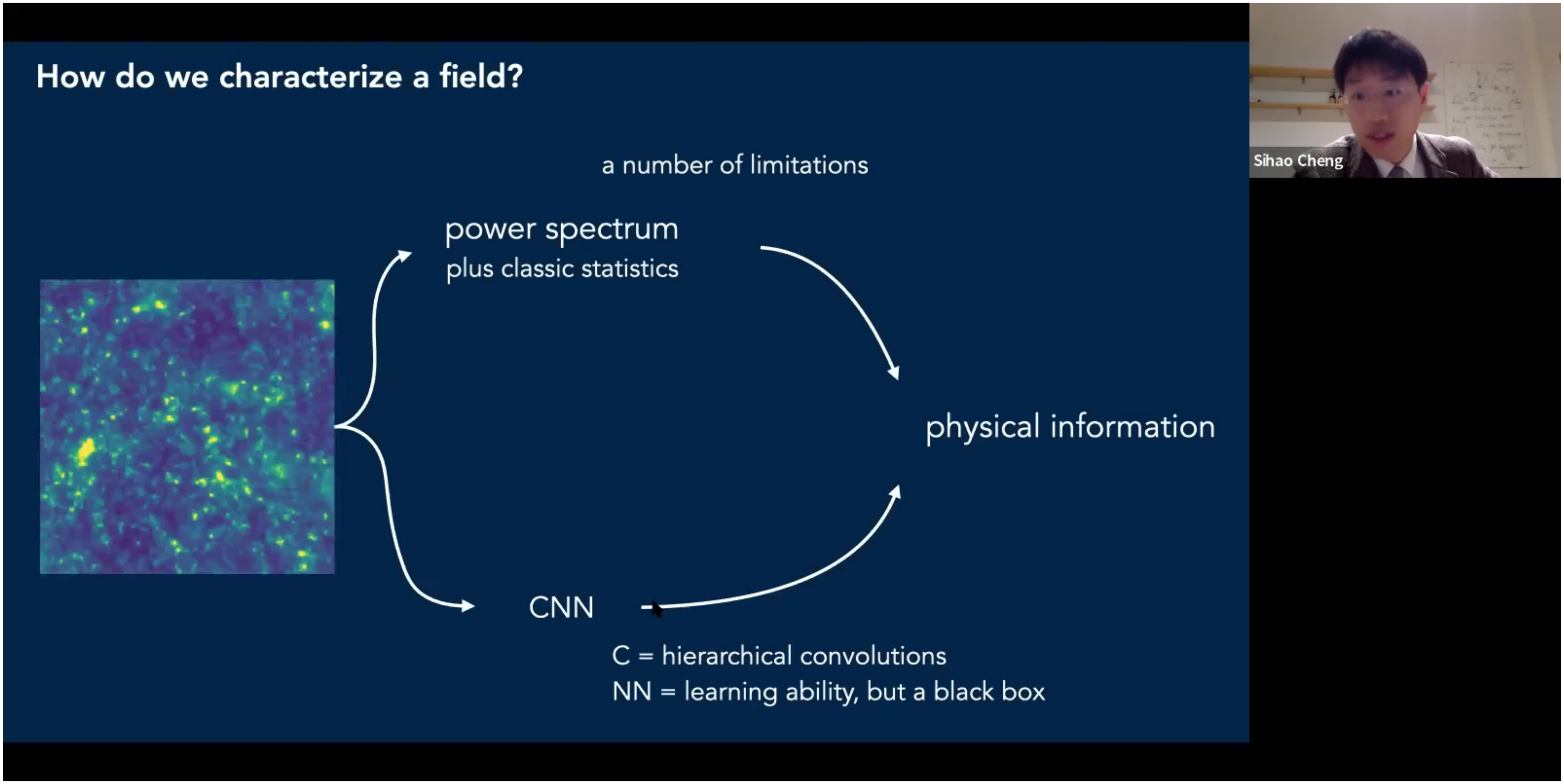
classification
discrete model

parameter inference
continuous model

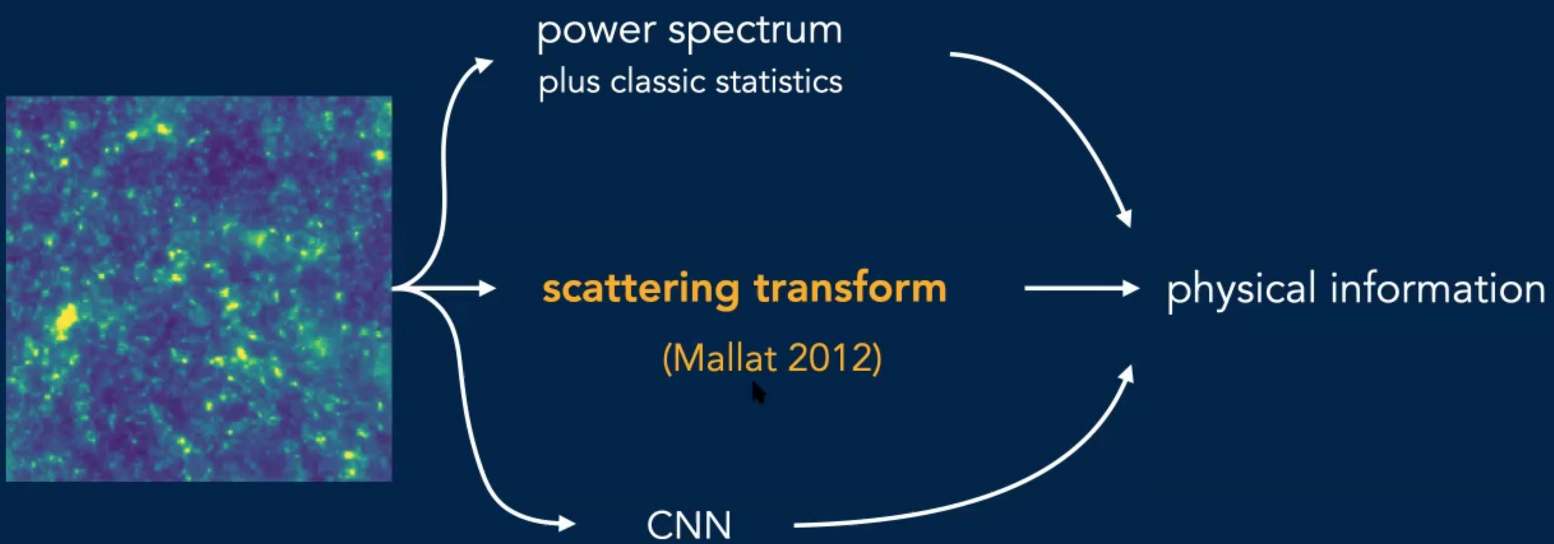

Sihao Cheng


How do we characterize a field?
⟨δ1δ2..δn⟩
power spectrum
plus classic statistics
physical information
log-normal variable
100%
Carron (2011)
information retrieved
moments
up to 2nd-order
0%
0.1
0.2
0.5
1
2
5
Gaussian
σ/μ
non-Gaussian
lose information
too many coefficients
Sihao Cheng

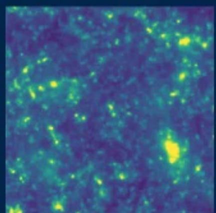

How do we characterize a field?
$\langle \delta_1 \delta_2 .. \delta_n \rangle$
power spectrum
plus classic statistics
physical information
log-normal variable
100%
0%
information retrieved
Carron (2011)
moments
up to 2nd-order
all moments
0.1
0.2
0.5
1
2
5
Gaussian
σ/μ
non-Gaussian
lose information
too many coefficients
Sihao Cheng


How do we characterize a field?
a number of limitations
power spectrum
plus classic statistics
physical information
CNN
C = hierarchical convolutions
NN = learning ability, but a black box
Sihao Cheng

How do we characterize a field?

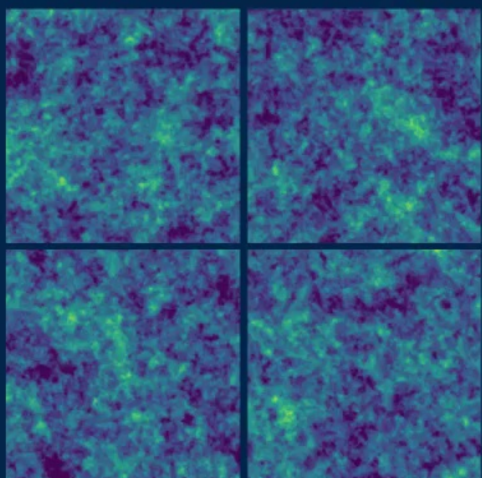


What do the statistics see?



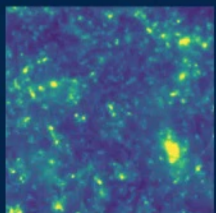
input map

with power spectrum P(k)



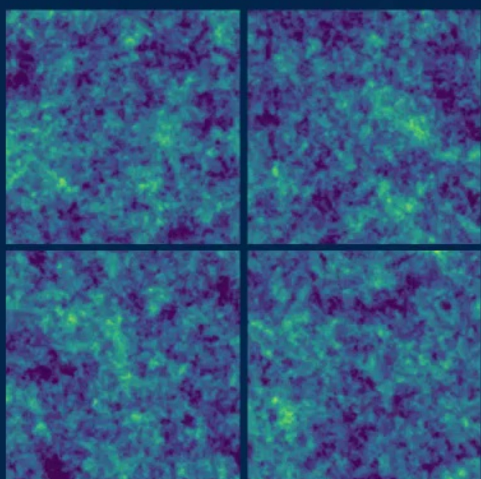
Cheng & Menard 2021a

What do the statistics see?

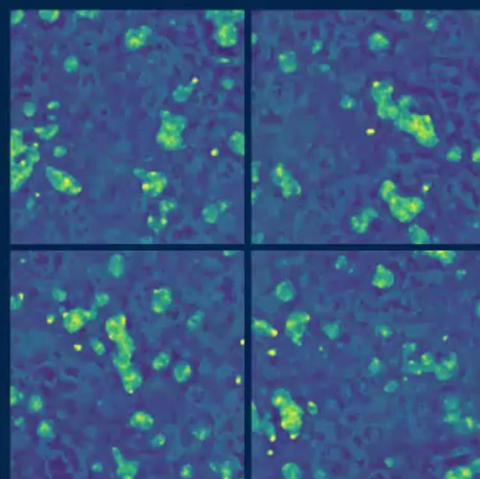


input map

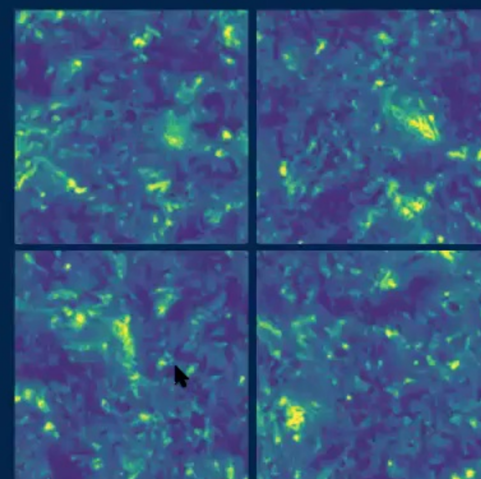
with power spectrum P(k)



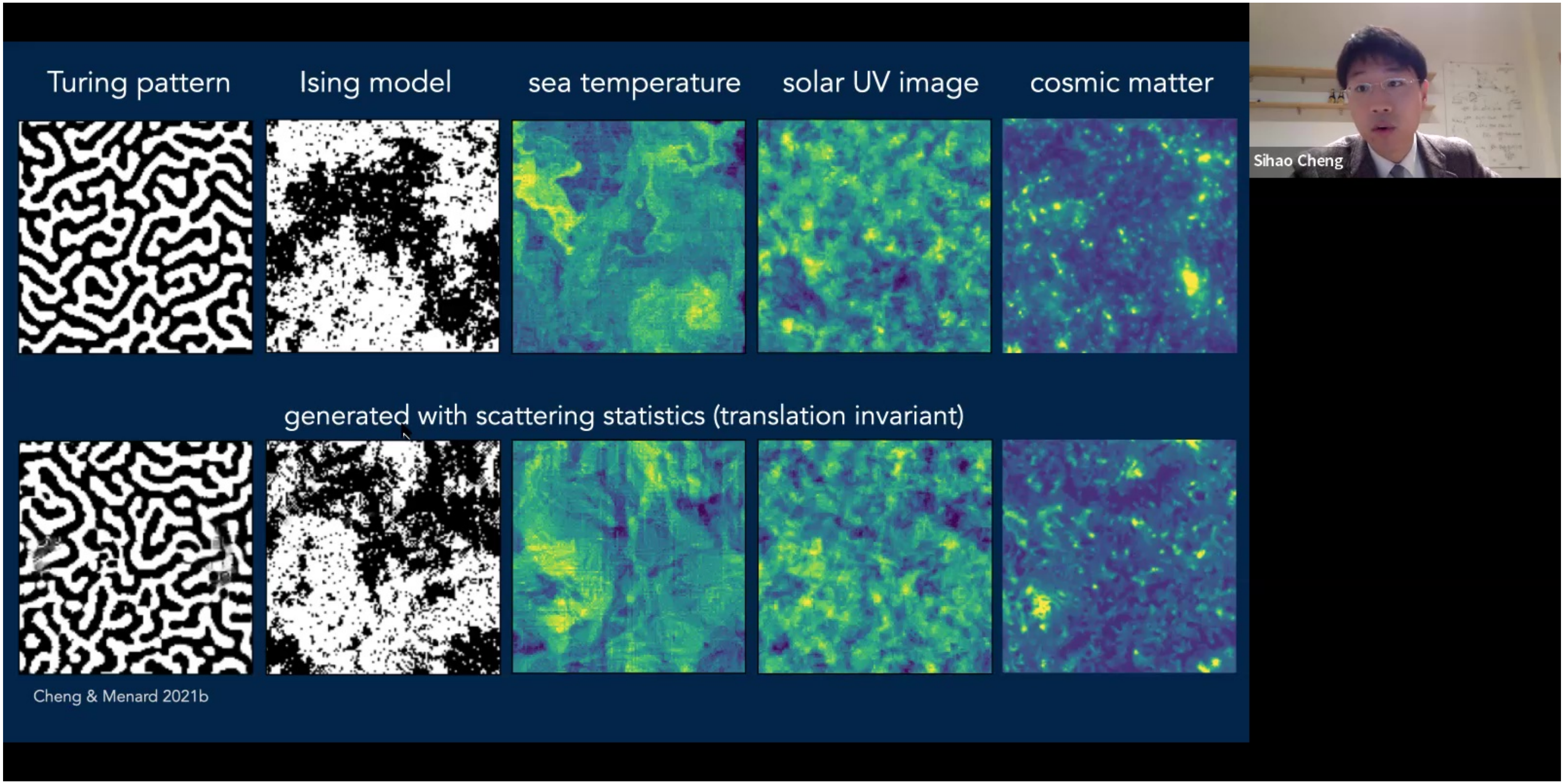
with P(k) and bispectrum

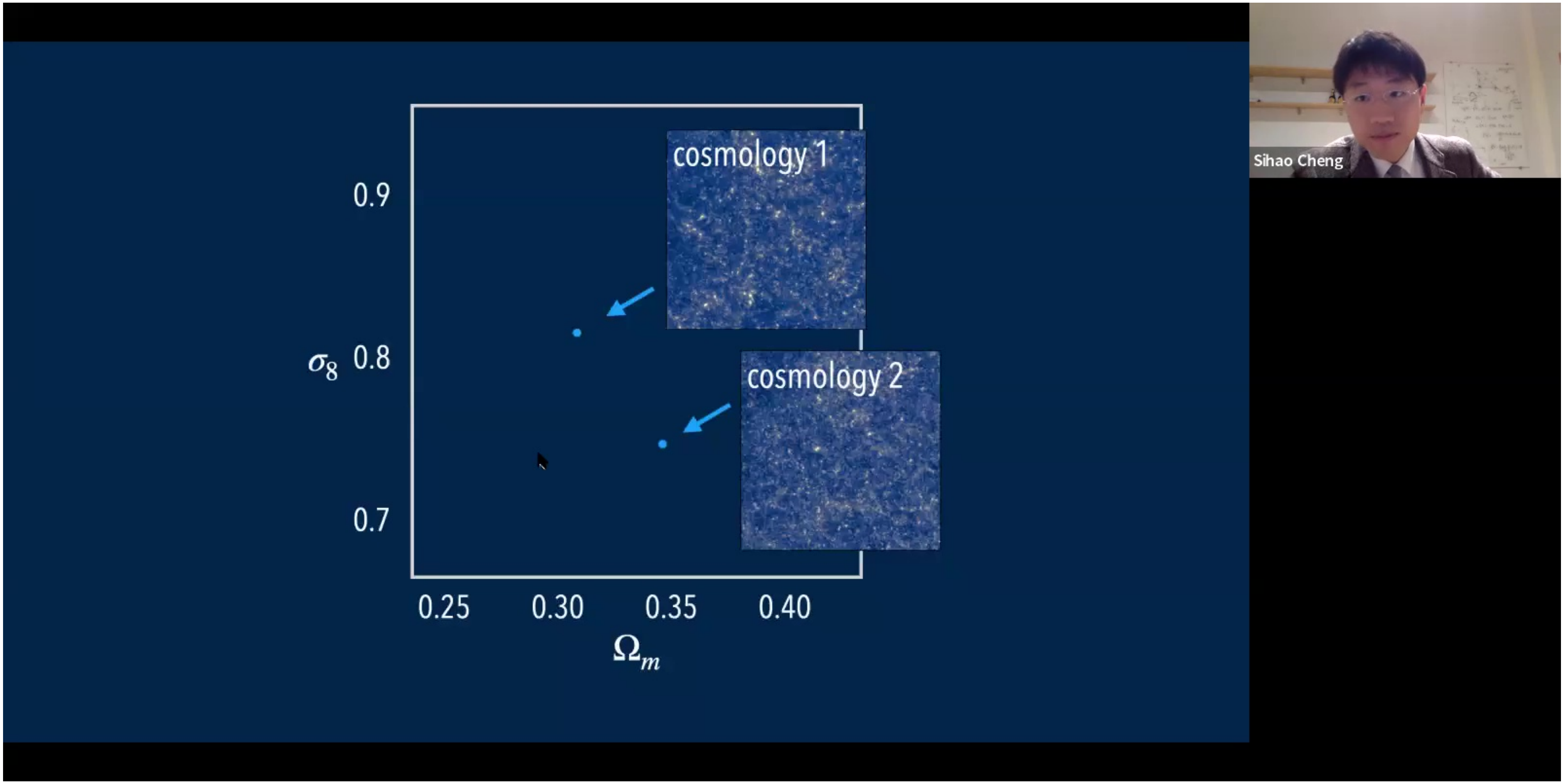


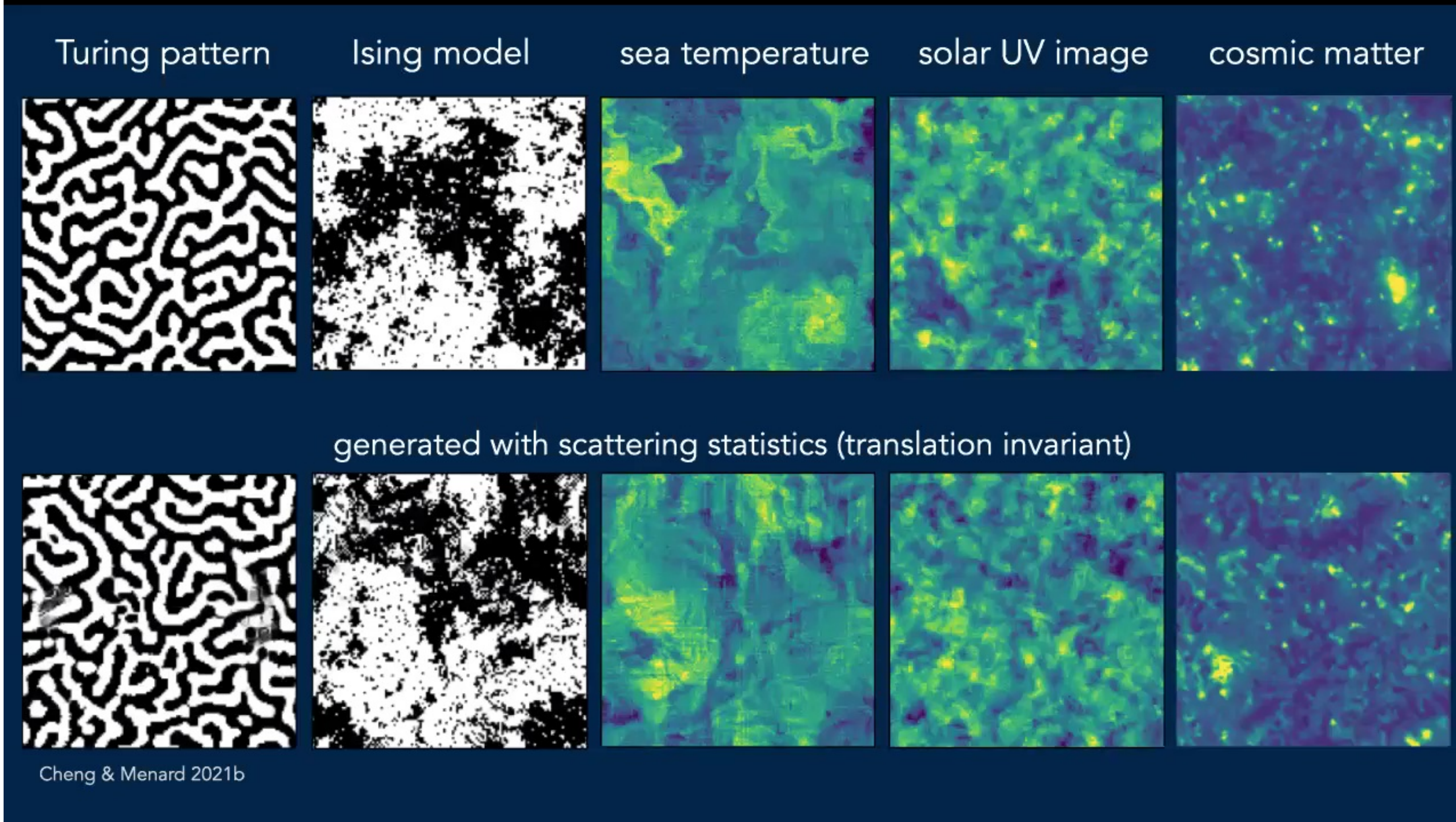
with scattering statistics

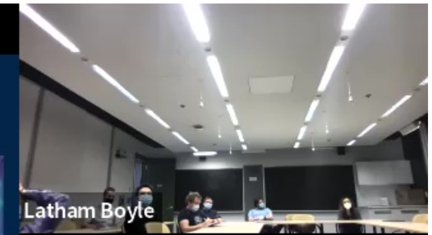


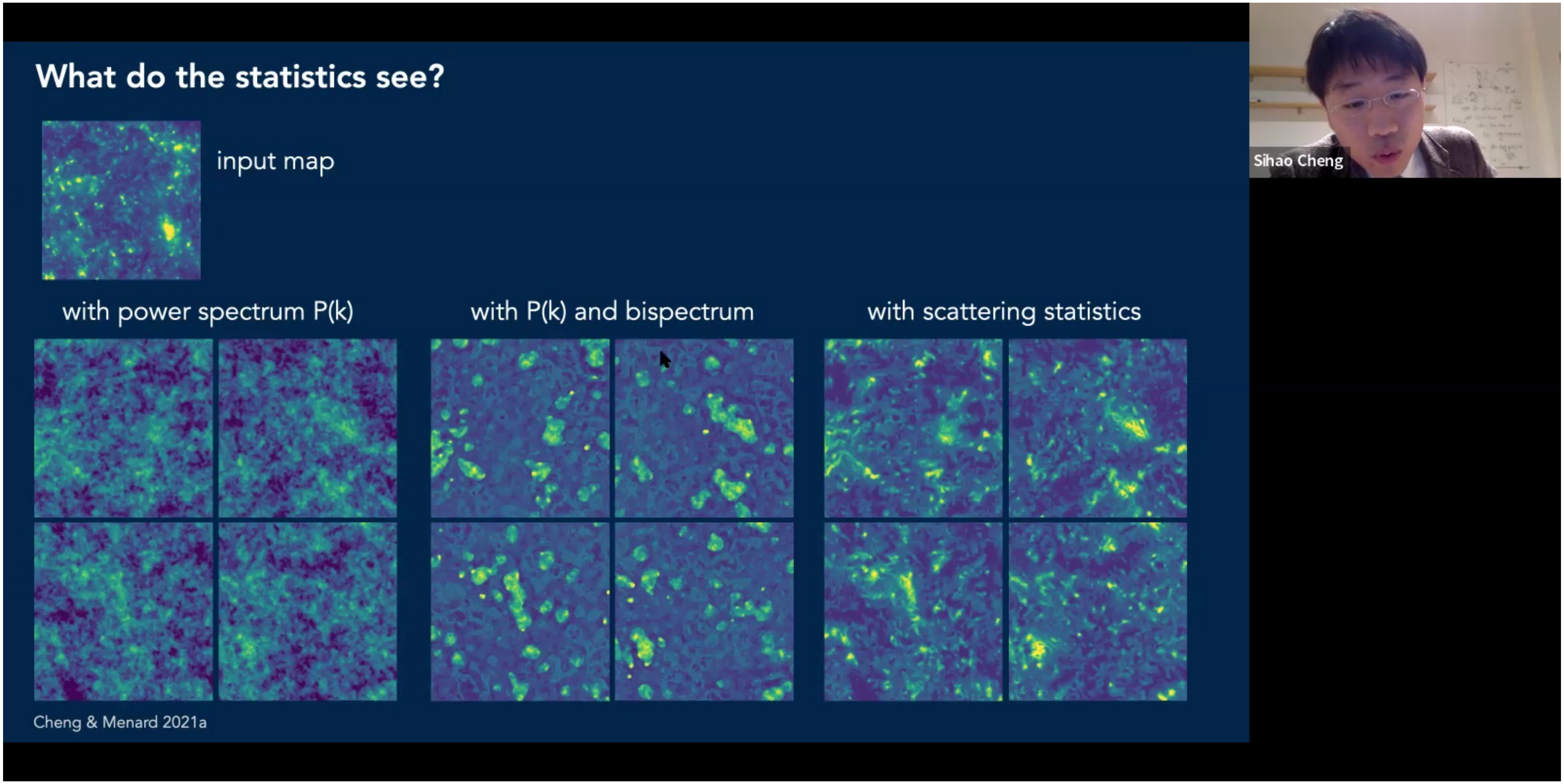
Cheng & Menard 2021a

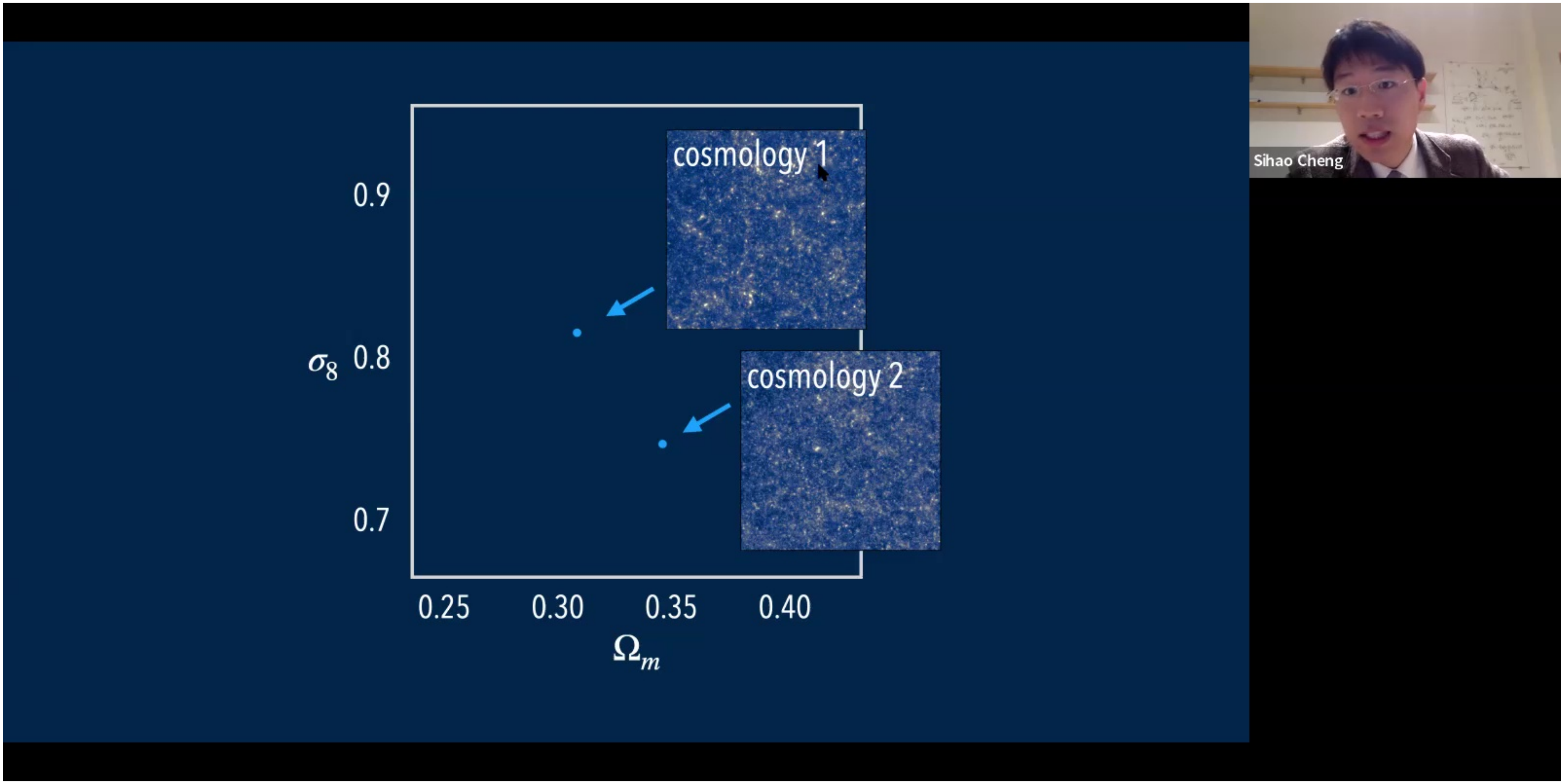

Turing pattern
Ising model
sea temperature
solar UV image
cosmic matter
generated with scattering statistics (translation invariant)
Cheng & Menard 2021b
Sihao Cheng

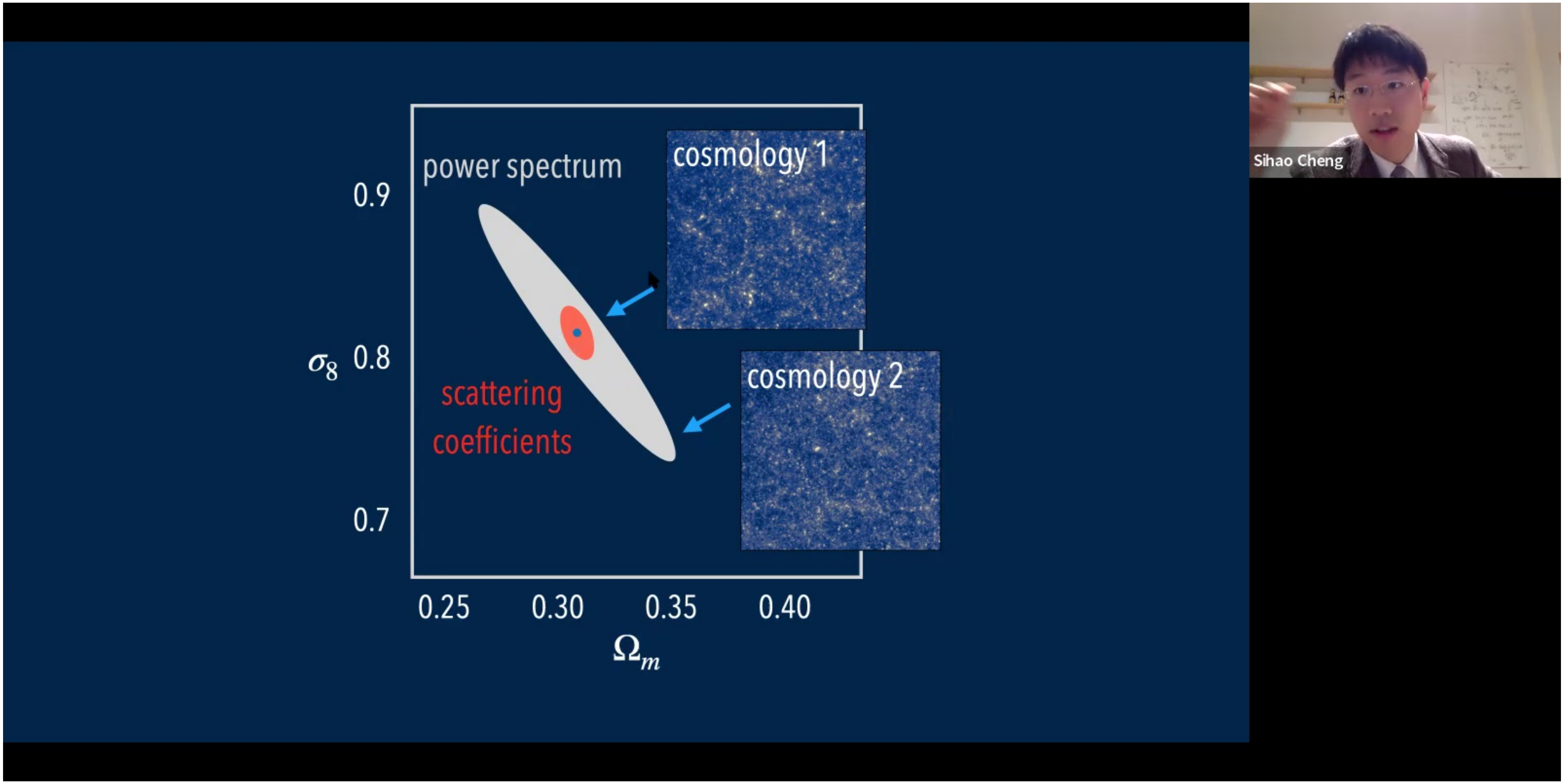

cosmology 1
cosmology 2
0.9
σ_8 0.8
0.7
0.25
0.30
0.35
0.40
Ω_m
Sihao Cheng

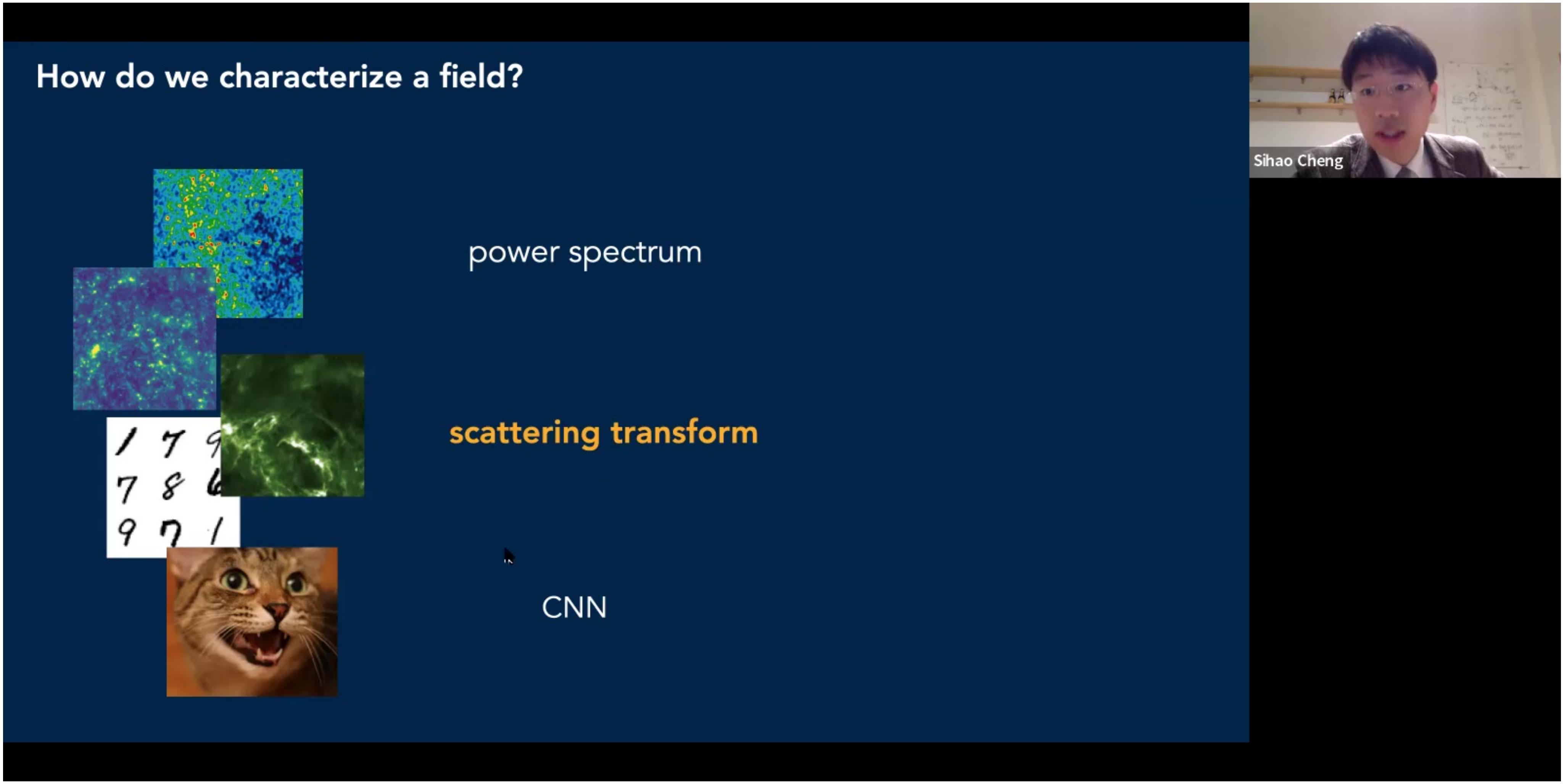

Turing pattern
Ising model
sea temperature
solar UV image
cosmic matter
generated with scattering statistics (translation invariant)
Cheng & Menard 2021b

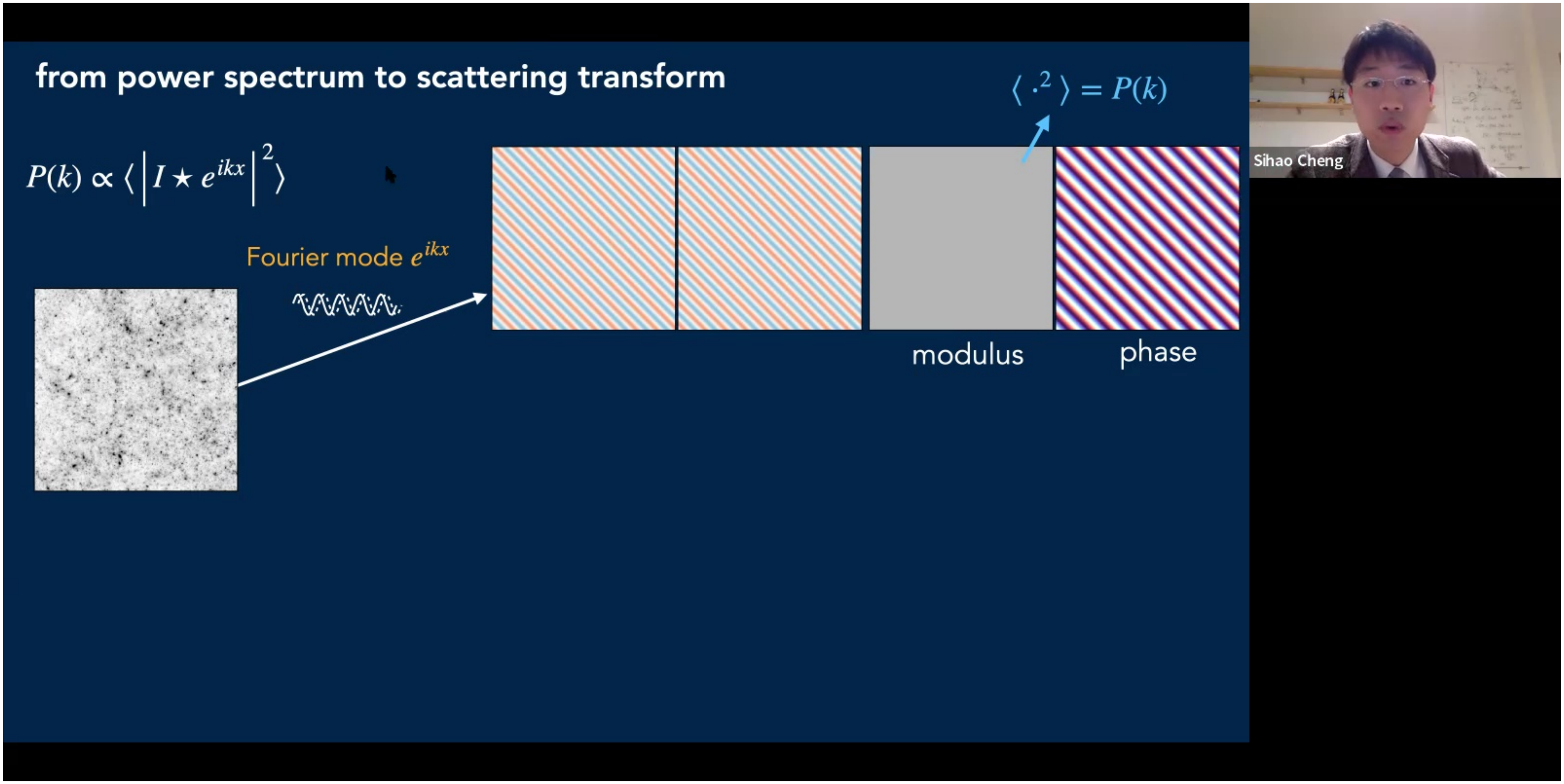

Latham Boyle

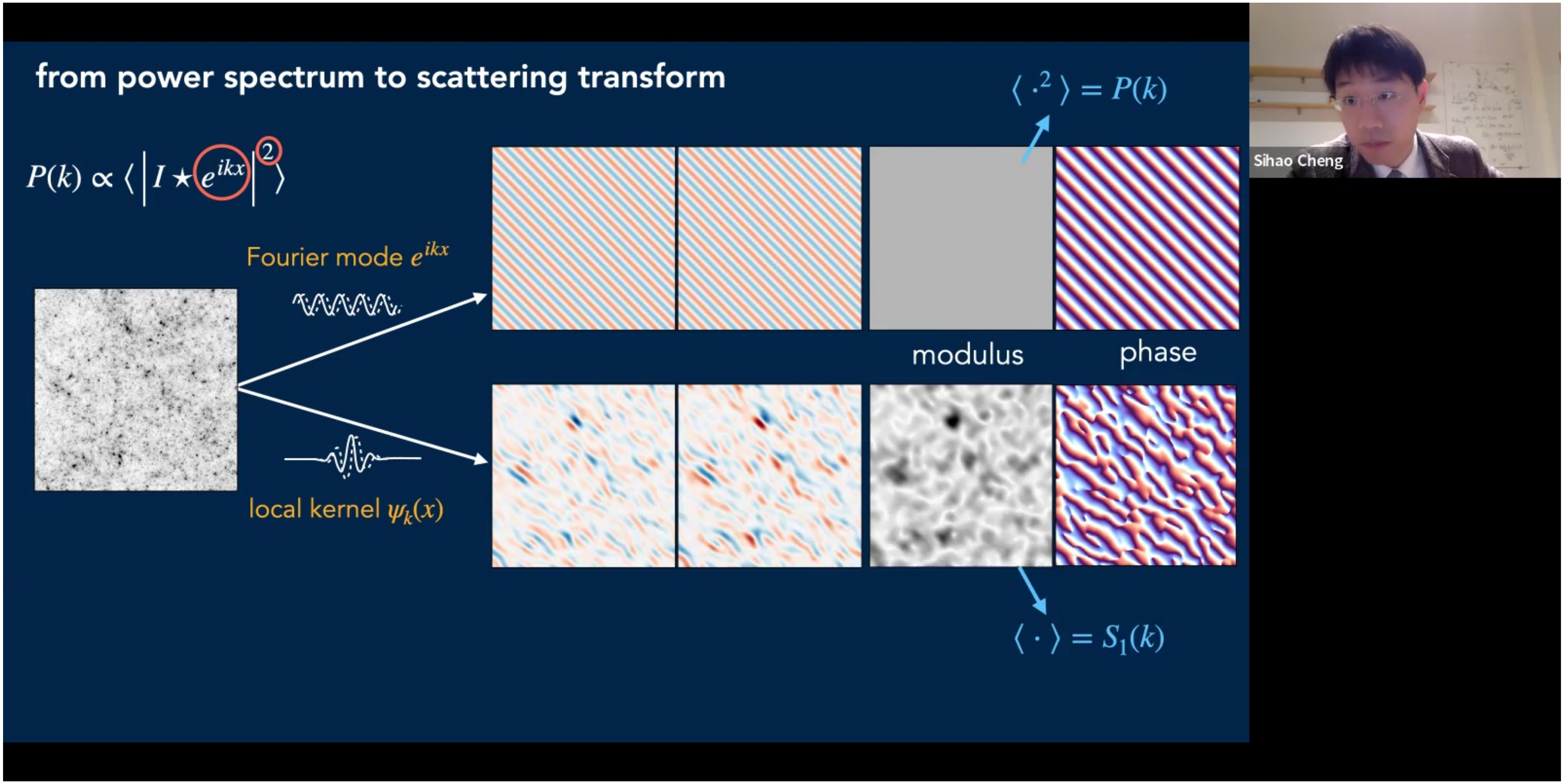

What do the statistics see?
input map
with power spectrum P(k)
with P(k) and bispectrum
with scattering statistics
Cheng & Menard 2021a
Sihao Cheng


cosmology 1
cosmology 2
0.9
0.8
0.7
σ_8
0.25
0.30
0.35
0.40
Ω_m
Sihao Cheng

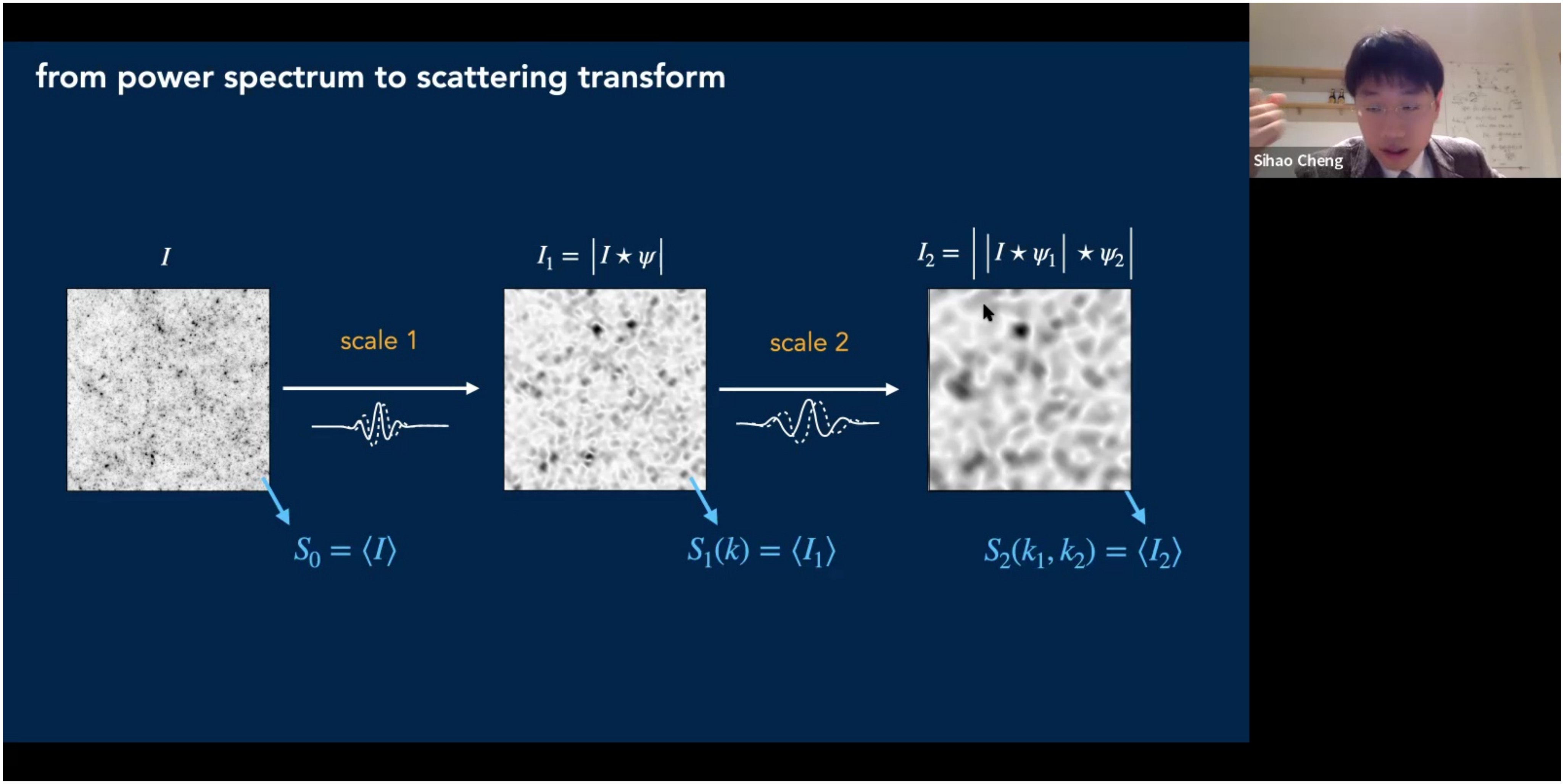

power spectrum
cosmology 1
cosmology 2
scattering
coefficients
0.9
0.8
0.7
σ_8
0.25
0.30
0.35
0.40
Ω_m
Sihao Cheng


How do we characterize a field?
power spectrum
scattering transform
CNN
Sihao Cheng


from power spectrum to scattering transform
$P(k) \propto \langle \left| I \star e^{ikx} \right|^2 \rangle$
Fourier mode e^{ikx}
$\langle \cdot^2 \rangle = P(k)$
modulus
phase
Sihao Cheng

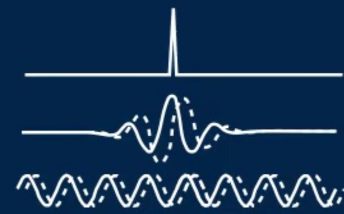


from power spectrum to scattering transform
$P(k) \propto \langle \left| I \star e^{ikx} \right|^2 \rangle$
Fourier mode e^{ikx}
local kernel $\psi_k(x)$
modulus
phase
$\langle \cdot^2 \rangle = P(k)$
$\langle \cdot \rangle = S_1(k)$
Sihao Cheng


from power spectrum to scattering transform
I
$I_1 = |I \star \psi|$
$I_2 = ||I \star \psi_1| \star \psi_2|$
scale 1
scale 2
$S_0 = \langle I \rangle$
$S_1(k) = \langle I_1 \rangle$
$S_2(k_1, k_2) = \langle I_2 \rangle$
Sihao Cheng

wavelets: an efficient decomposition

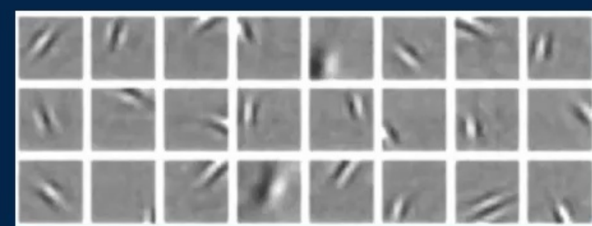
Dirac
wavelet
Fourier



receptive fields of mammal vision

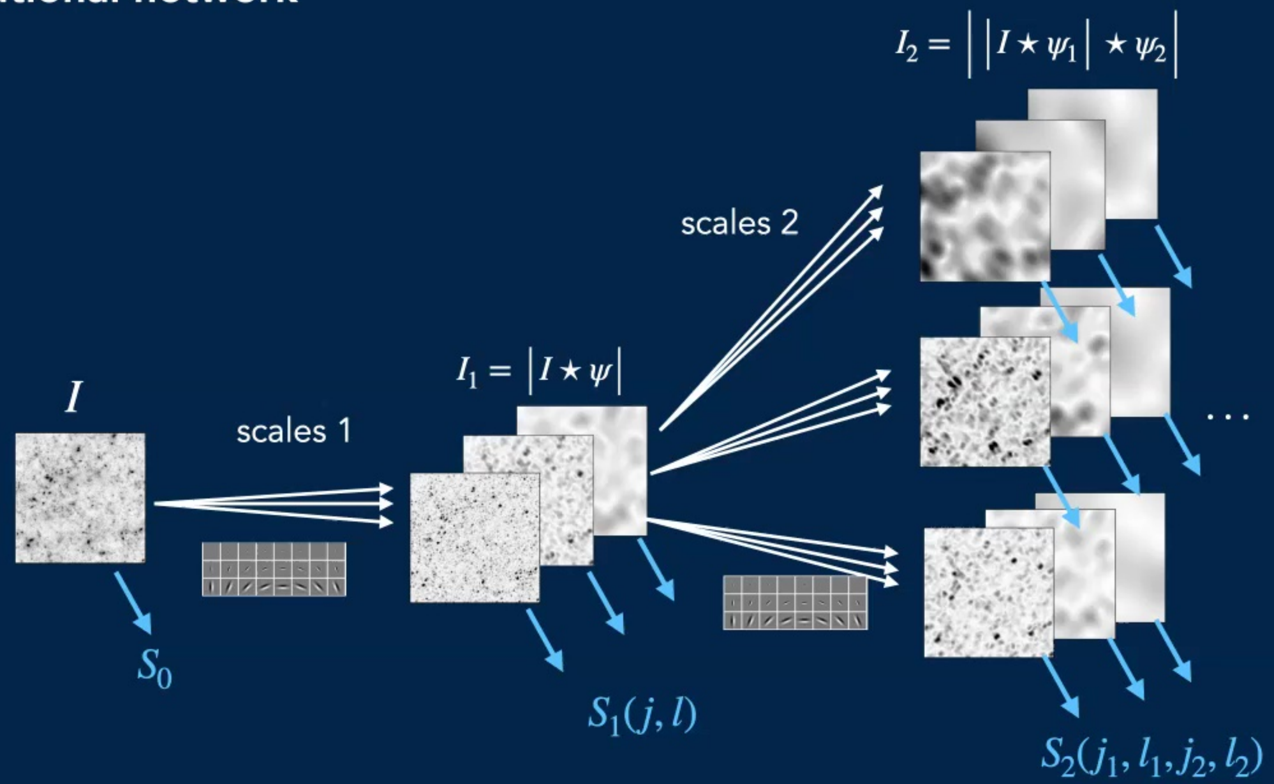
close to Gabor wavelets

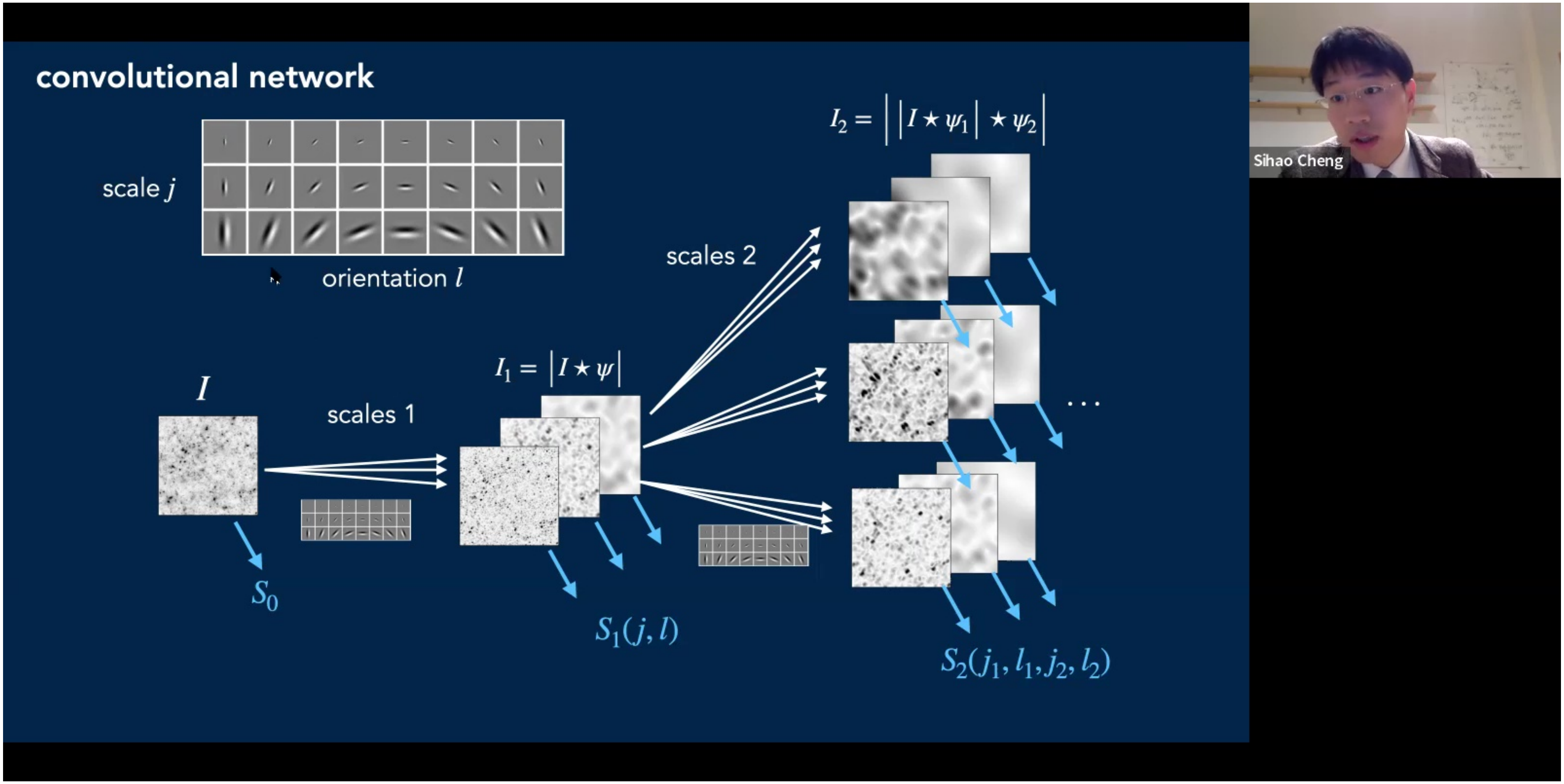
sparse representation of natural images



kernels learned in AlexNet

convolutional network




convolutional network
scale j
orientation l
$I_2 = \left| \left| I \star \psi_1 \right| \star \psi_2 \right|$
scales 2
$I_1 = \left| I \star \psi \right|$
I
scales 1
S_0
$S_1(j, l)$
$S_2(j_1, l_1, j_2, l_2)$
Sihao Cheng

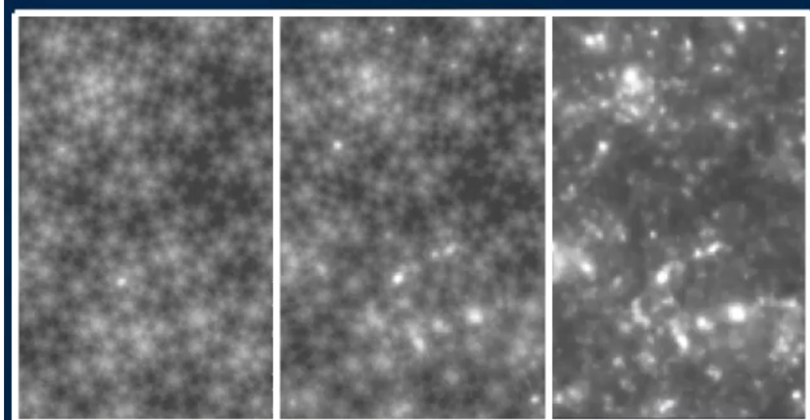
interpretation

$S_0 = \langle I \rangle$: mean

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

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

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



interpretation

$S_0 = \langle I \rangle$: mean

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

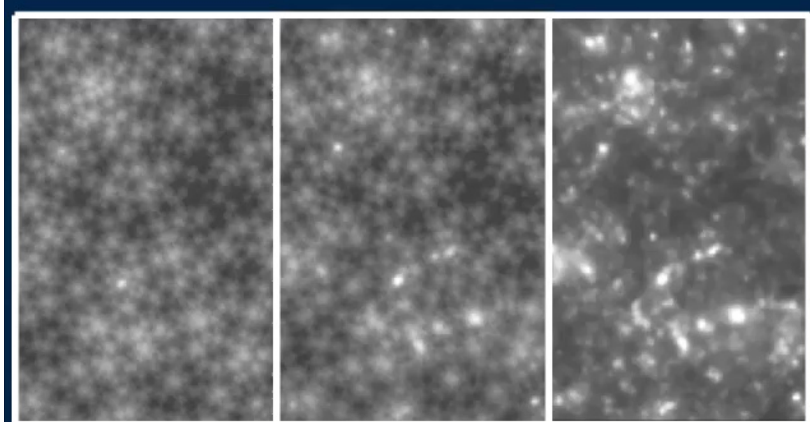
$S_2(j_1, l_1, j_2, l_2) = \langle \big| |I \star \psi_1| \star \psi_2 \big| \rangle$: non-Gaussianity. $\sim P(k)$ of $P(k)$

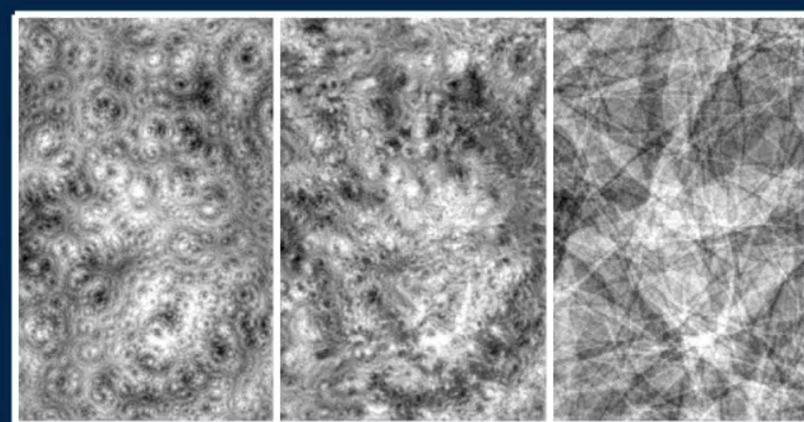
$l_1 \parallel l_2$:

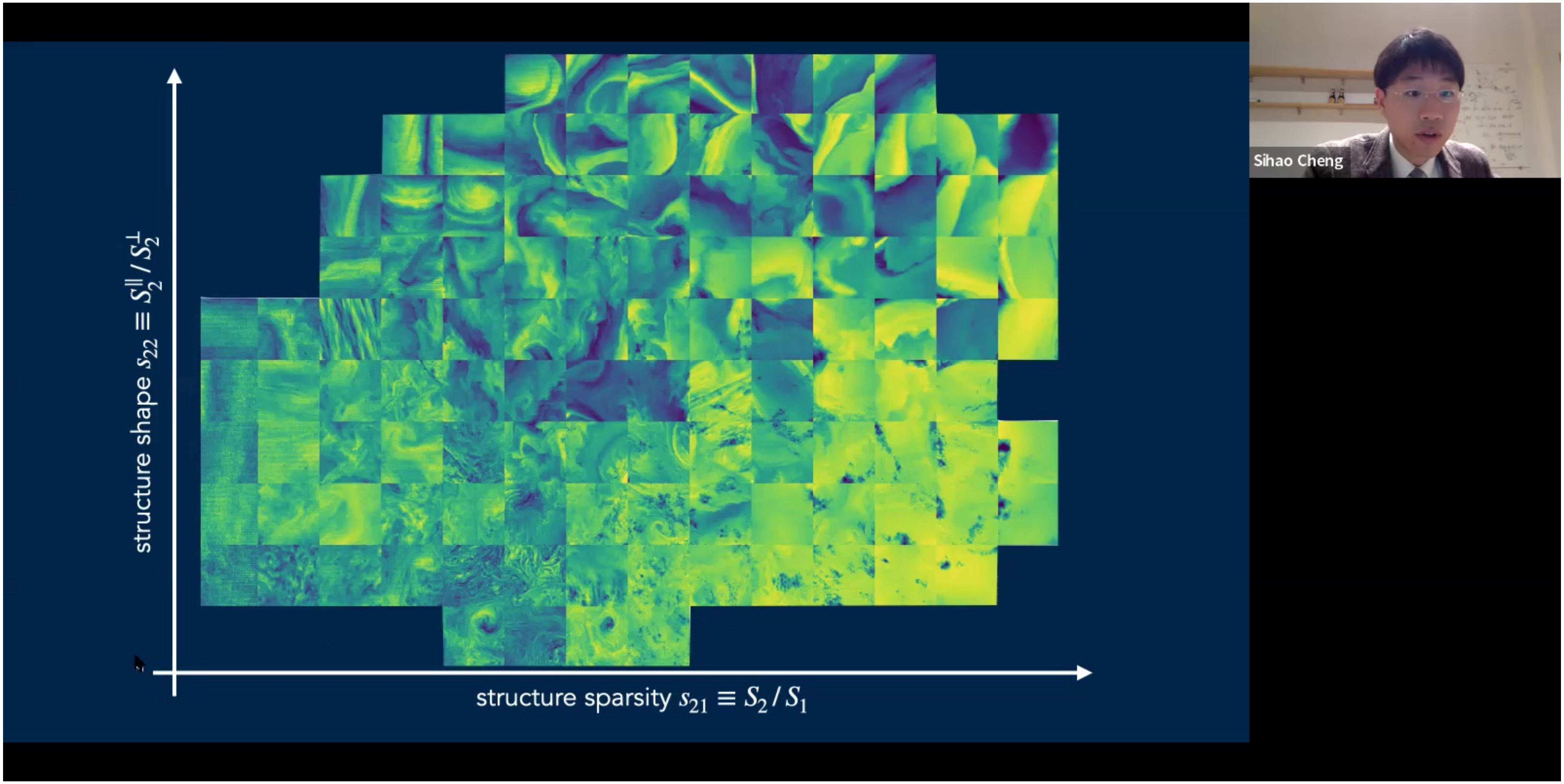
$l_1 \perp l_2$:

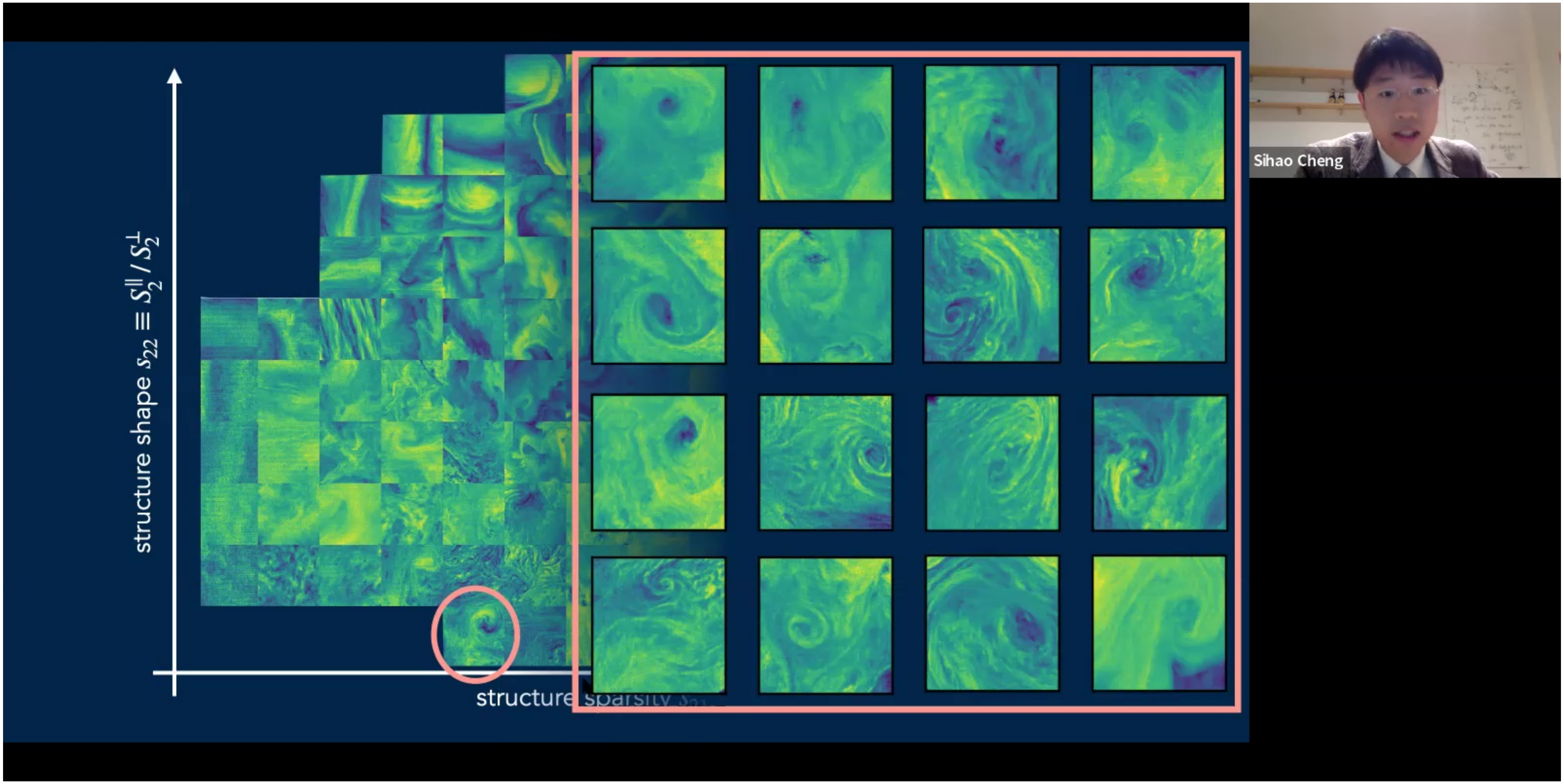
structure sparsity $s_{21} \equiv S_2 / S_1$

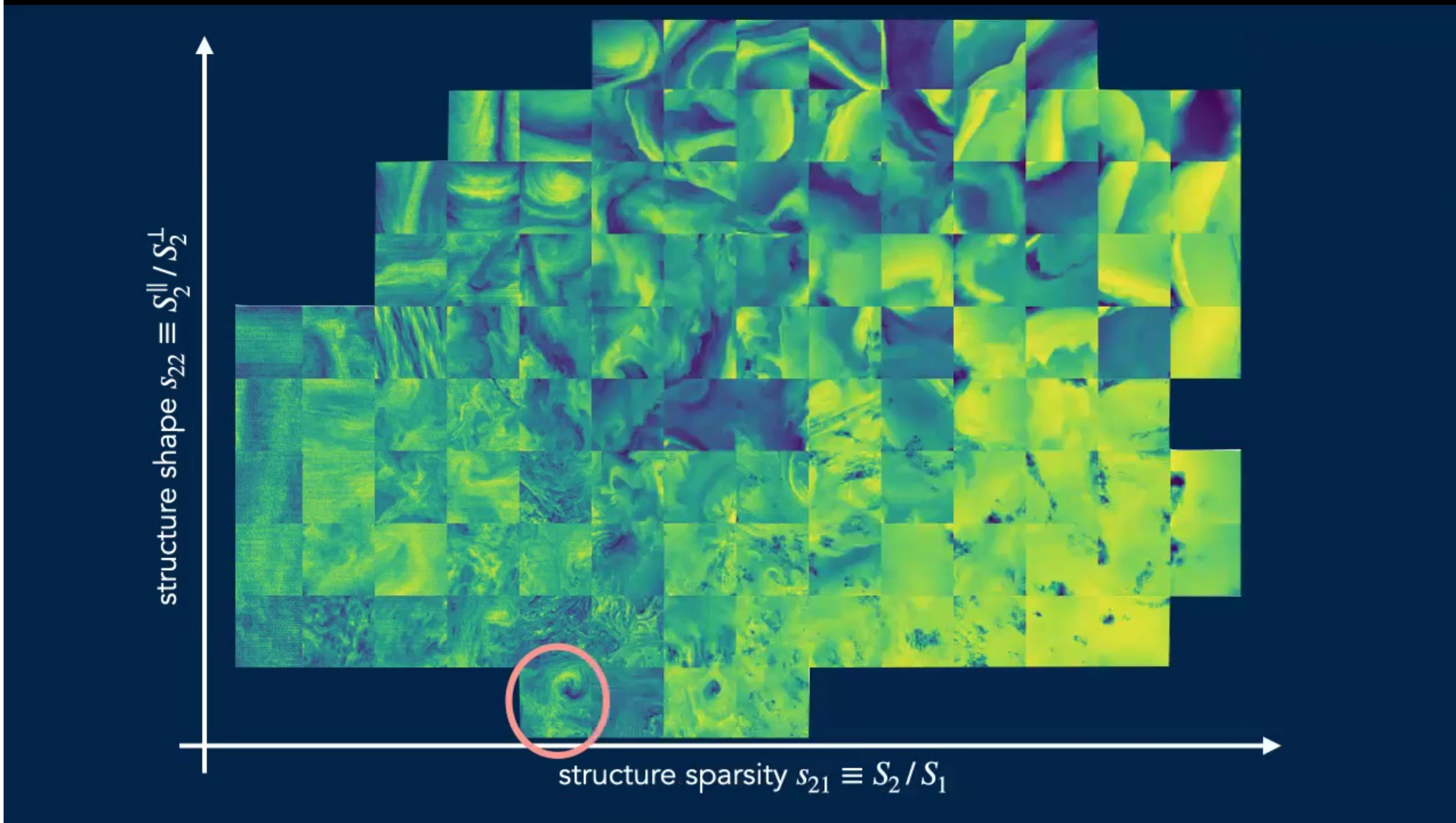
structure shape $s_{22} \equiv S_2^{\parallel} / S_2^{\perp}$

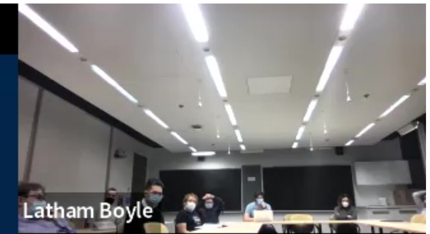


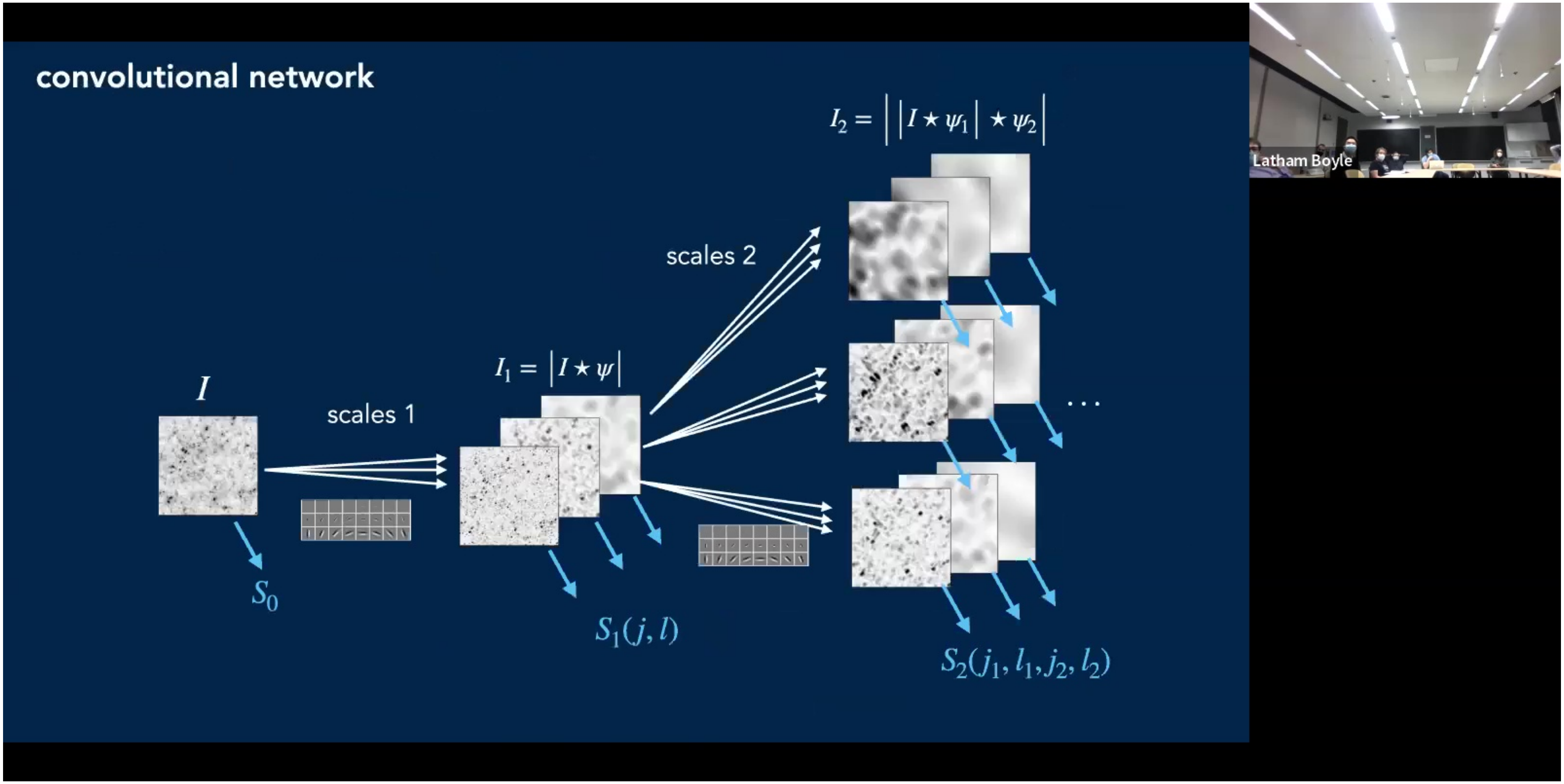



structure shape $s_{22} \equiv S_2^{\parallel} / S_2^{\perp}$
structure sparsity $s_{21} \equiv S_2 / S_1$
Sihao Cheng

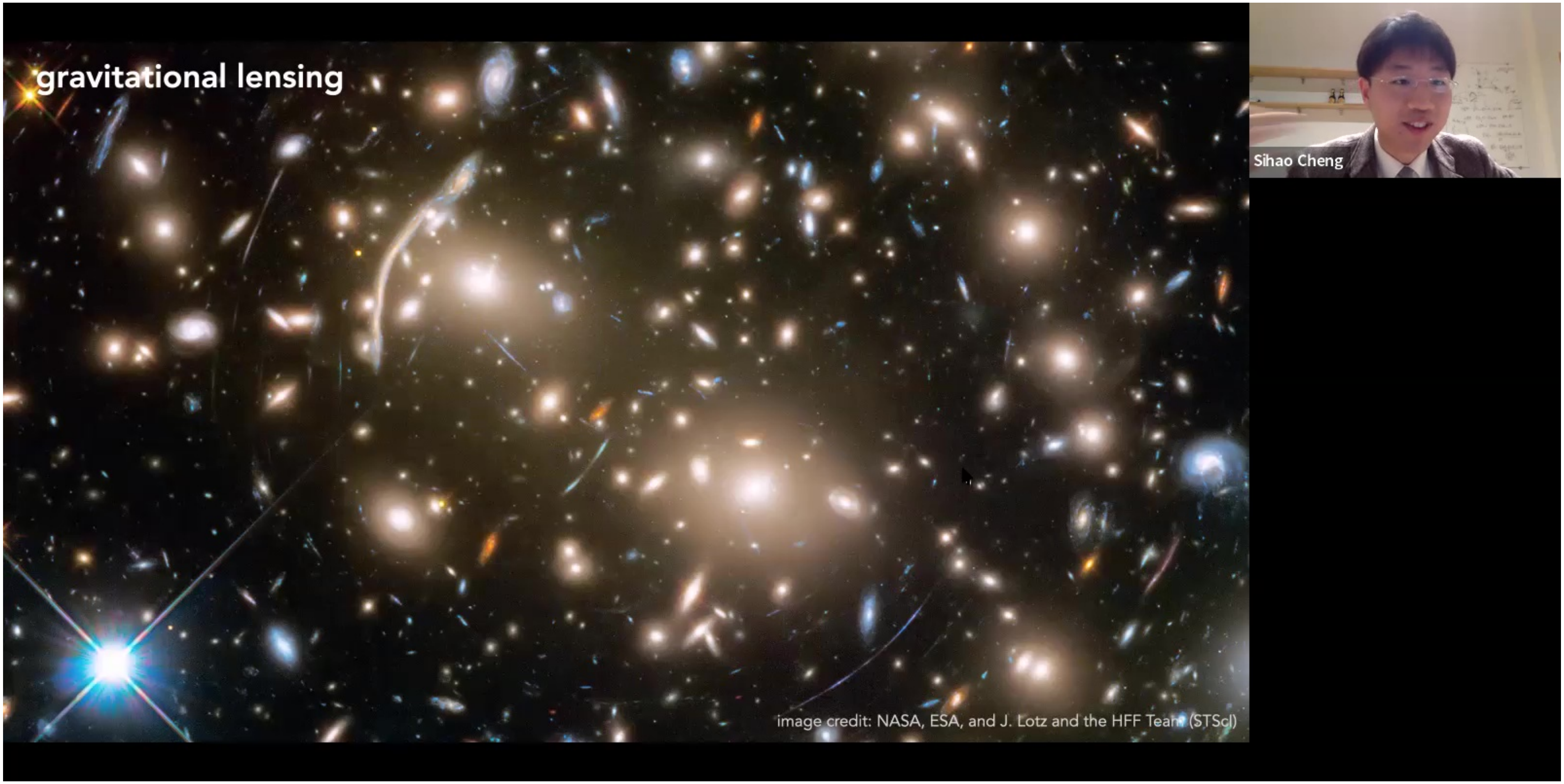

structure shape $s_{22} \equiv S_2^{\parallel} / S_2^{\perp}$
structure sparsity
Sihao Cheng

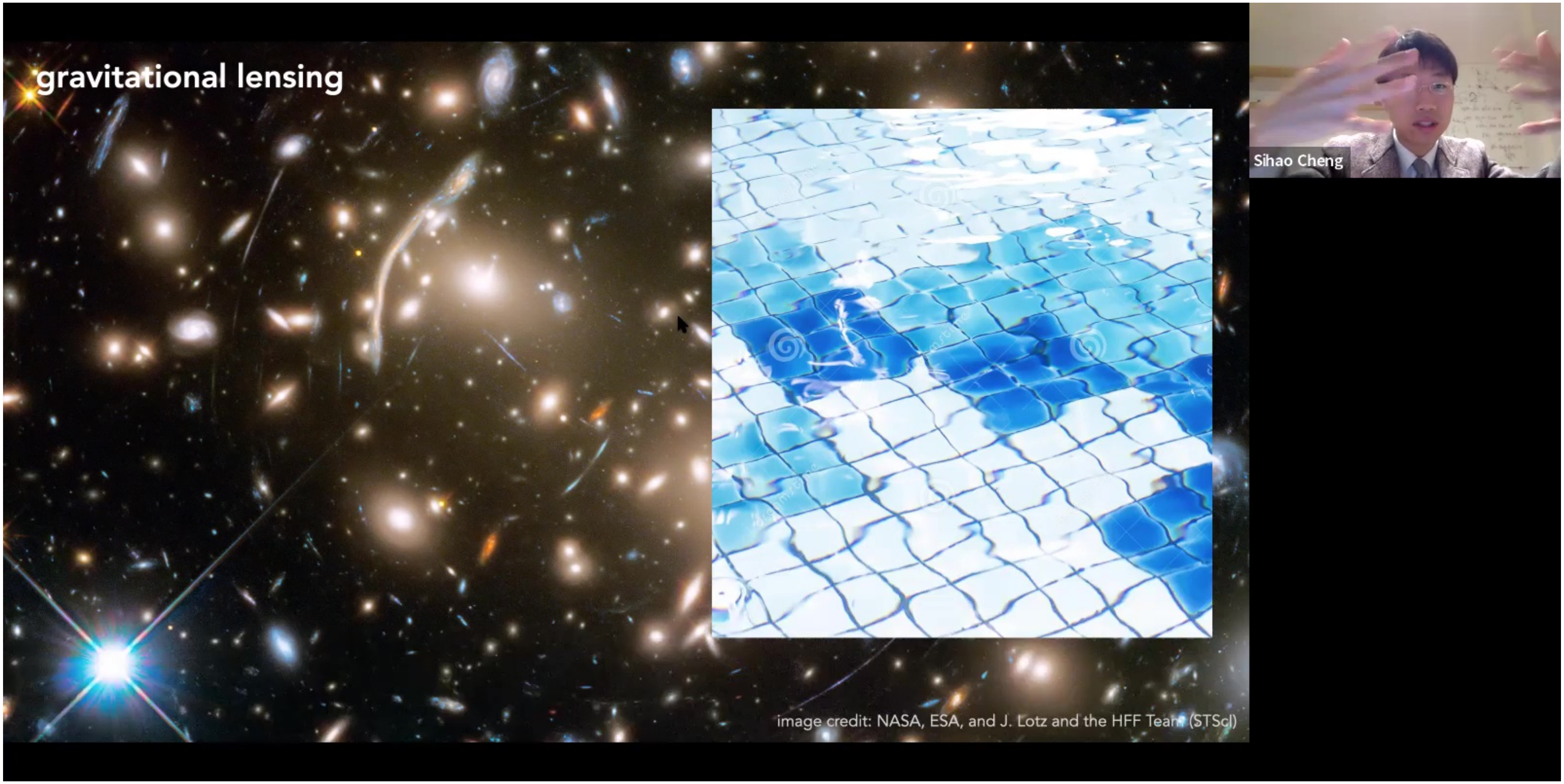

structure shape $s_{22} \equiv S_2^{\parallel} / S_2^{\perp}$
structure sparsity $s_{21} \equiv S_2 / S_1$

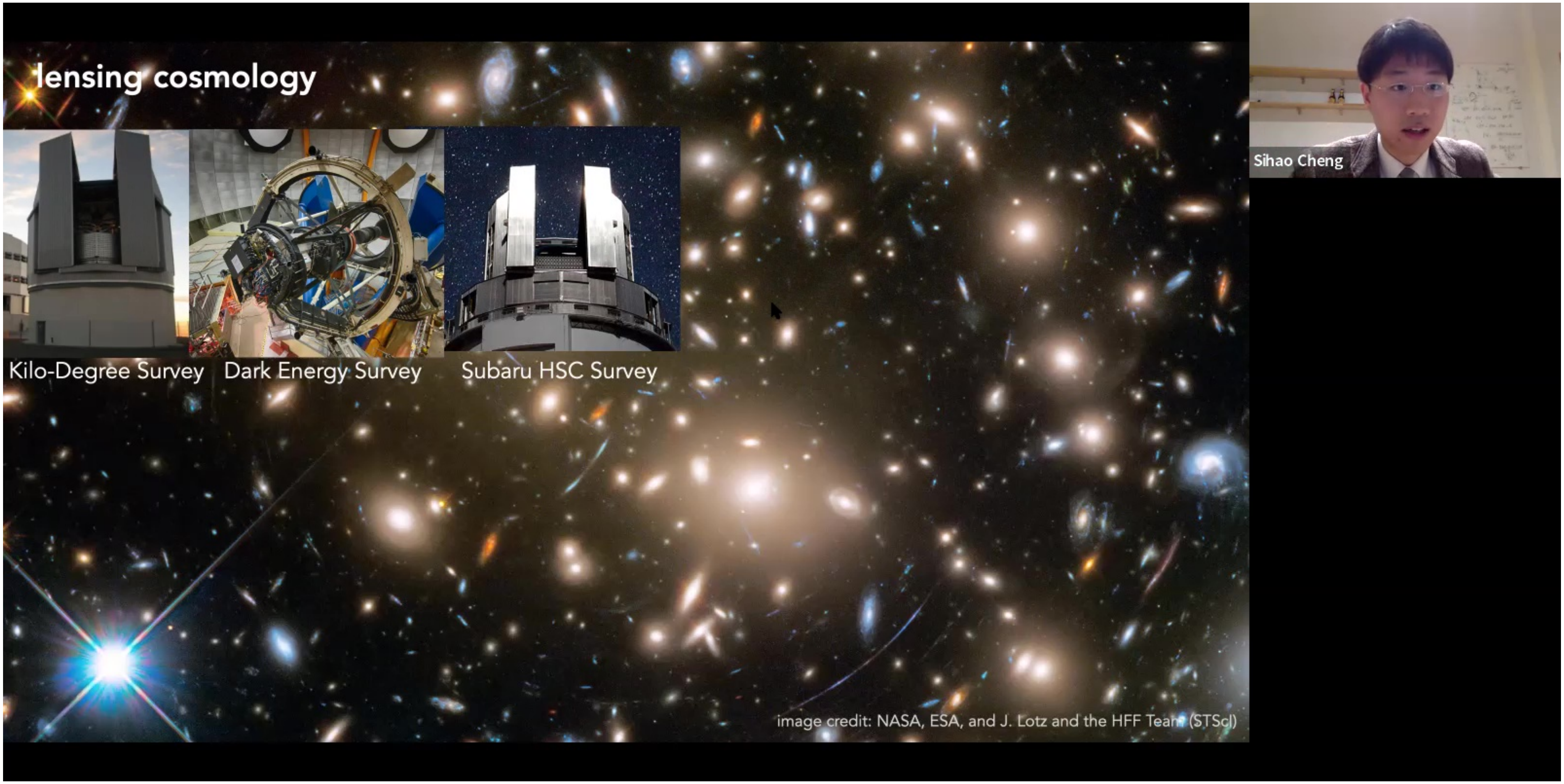

Latham Boyle

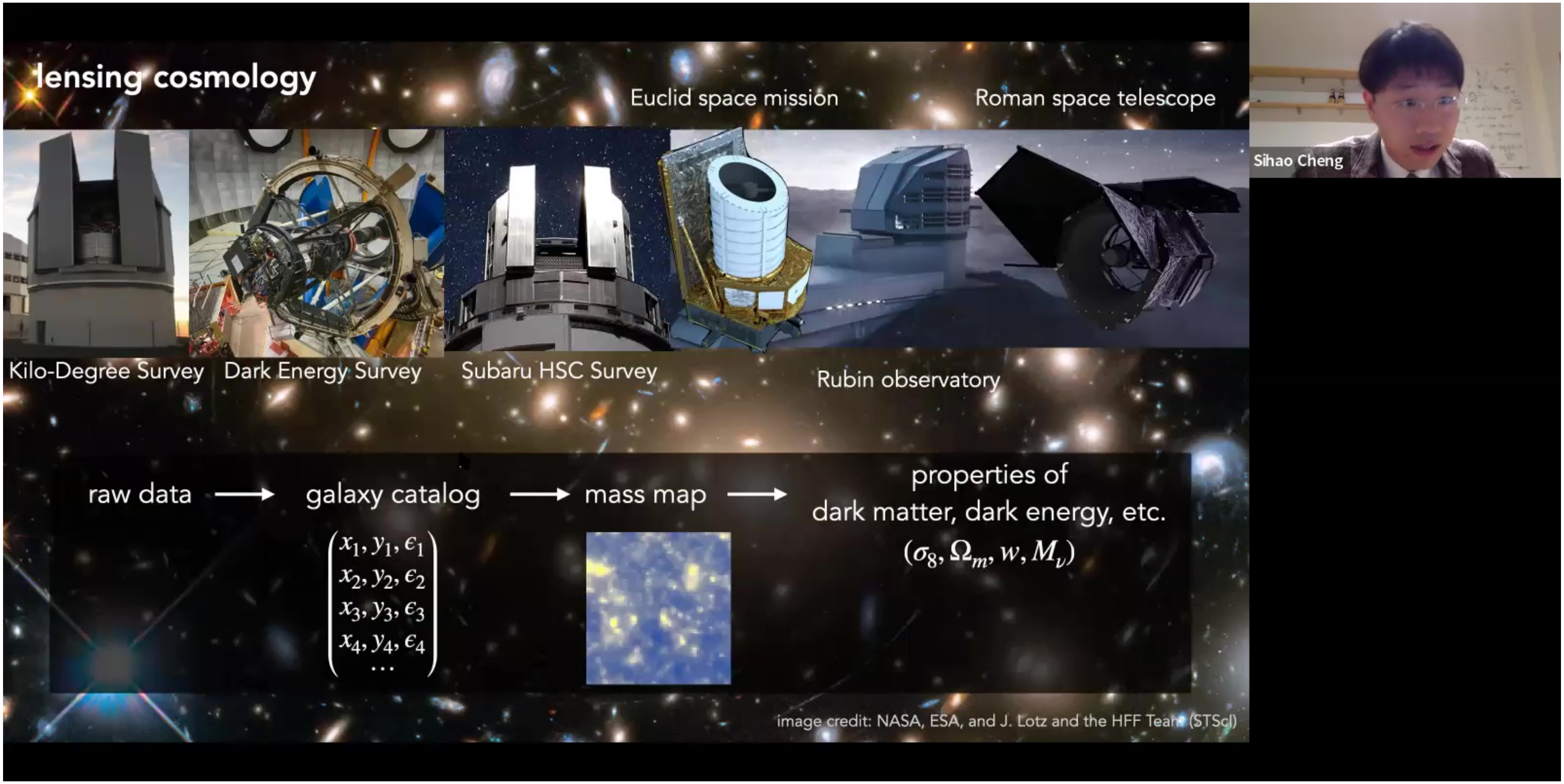


convolutional network
$I_2 = \left| \left| I \star \psi_1 \right| \star \psi_2 \right|$
scales 2
$I_1 = \left| I \star \psi \right|$
I
scales 1
. . .
S_0
$S_1(j, l)$
$S_2(j_1, l_1, j_2, l_2)$
Latham Boyle

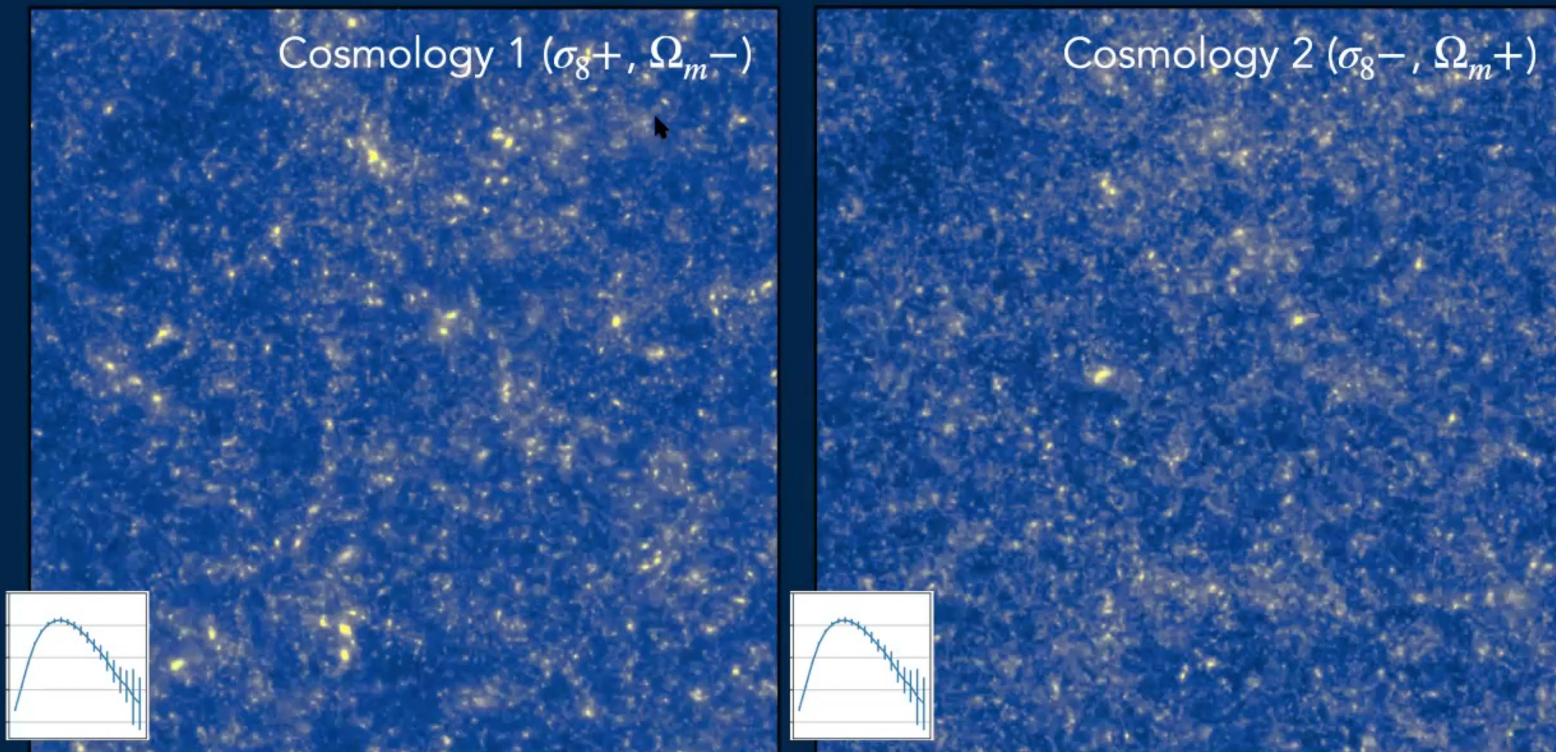

gravitational lensing
image credit: NASA, ESA, and J. Lotz and the HFF Team (STScI)
Sihao Cheng


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

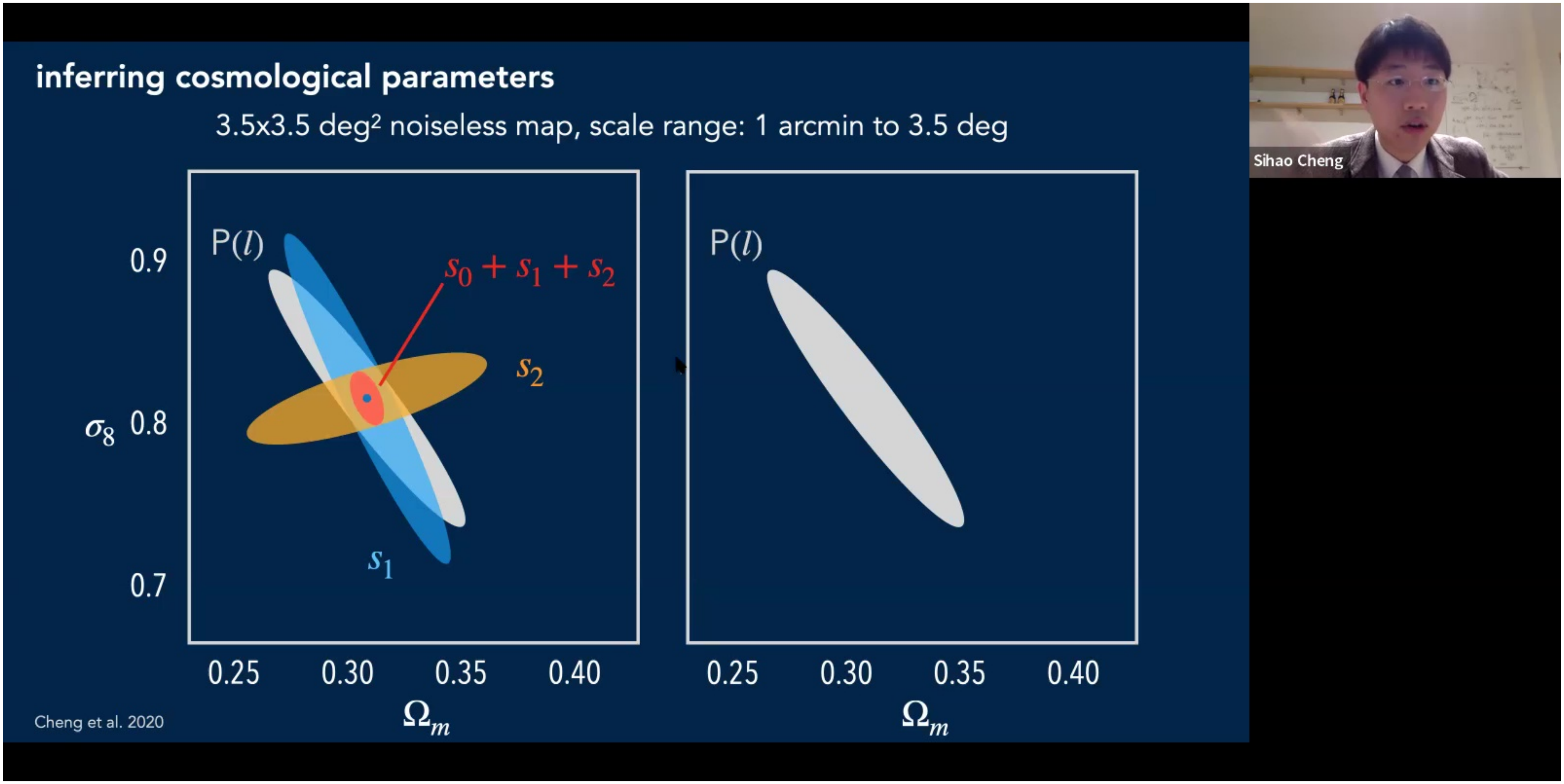

lensing cosmology
Kilo-Degree Survey
Dark Energy Survey
Subaru HSC Survey
image credit: NASA, ESA, and J. Lotz and the HFF Team (STScI)
Sihao Cheng


lensing cosmology
Euclid space mission
Roman space telescope
Sihao Cheng
Kilo-Degree Survey
Dark Energy Survey
Subaru HSC Survey
Rubin observatory
raw data → galaxy catalog → mass map → properties of dark matter, dark energy, etc.
$(x_1, y_1, \epsilon_1;\ x_2, y_2, \epsilon_2;\ x_3, y_3, \epsilon_3;\ x_4, y_4, \epsilon_4;\ \ldots)$
$(\sigma_8, \Omega_m, w, M_\nu)$
image credit: NASA, ESA, and J. Lotz and the HFF Team (STScI)

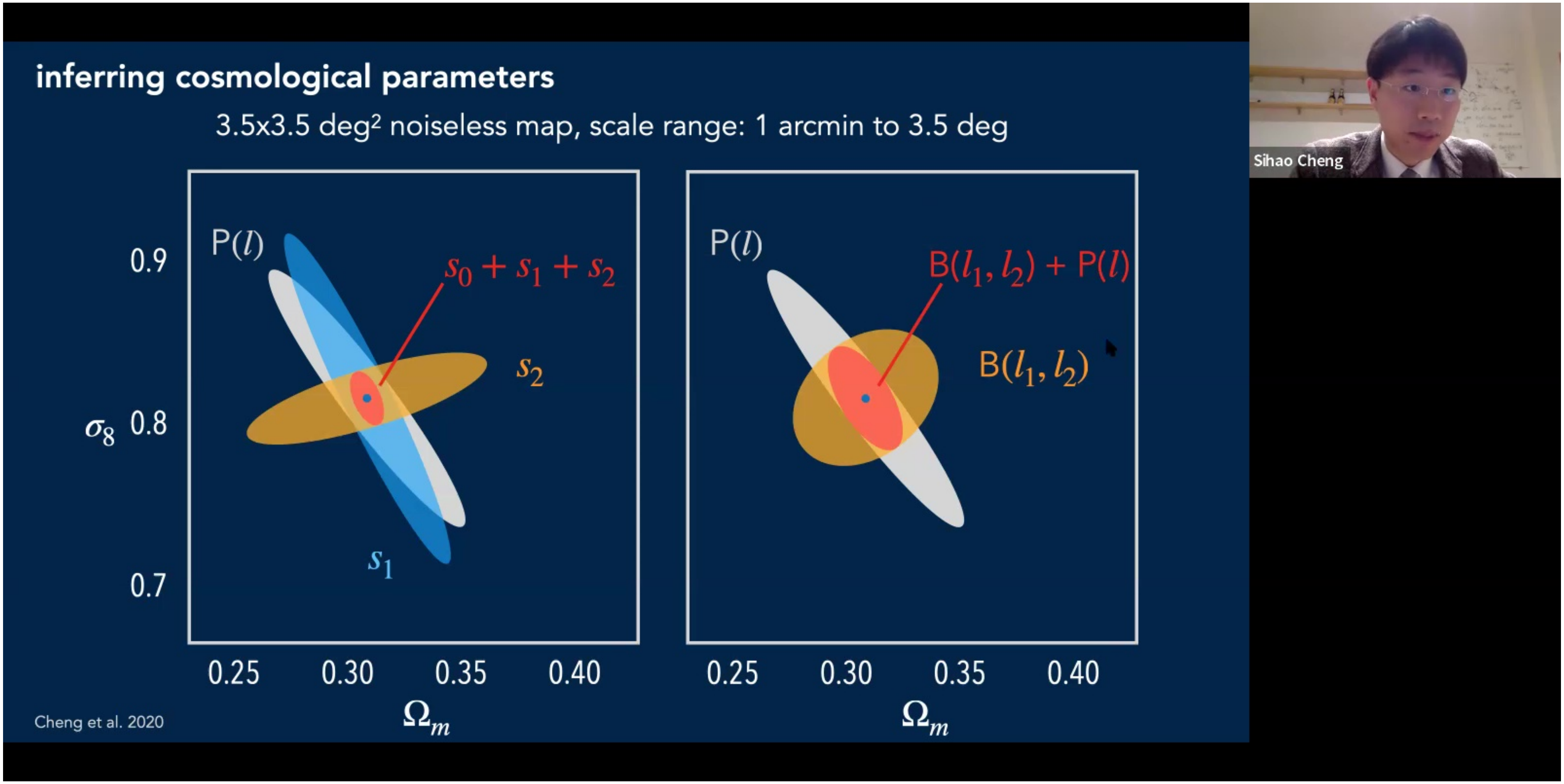
simulations (from Columbia lensing group)

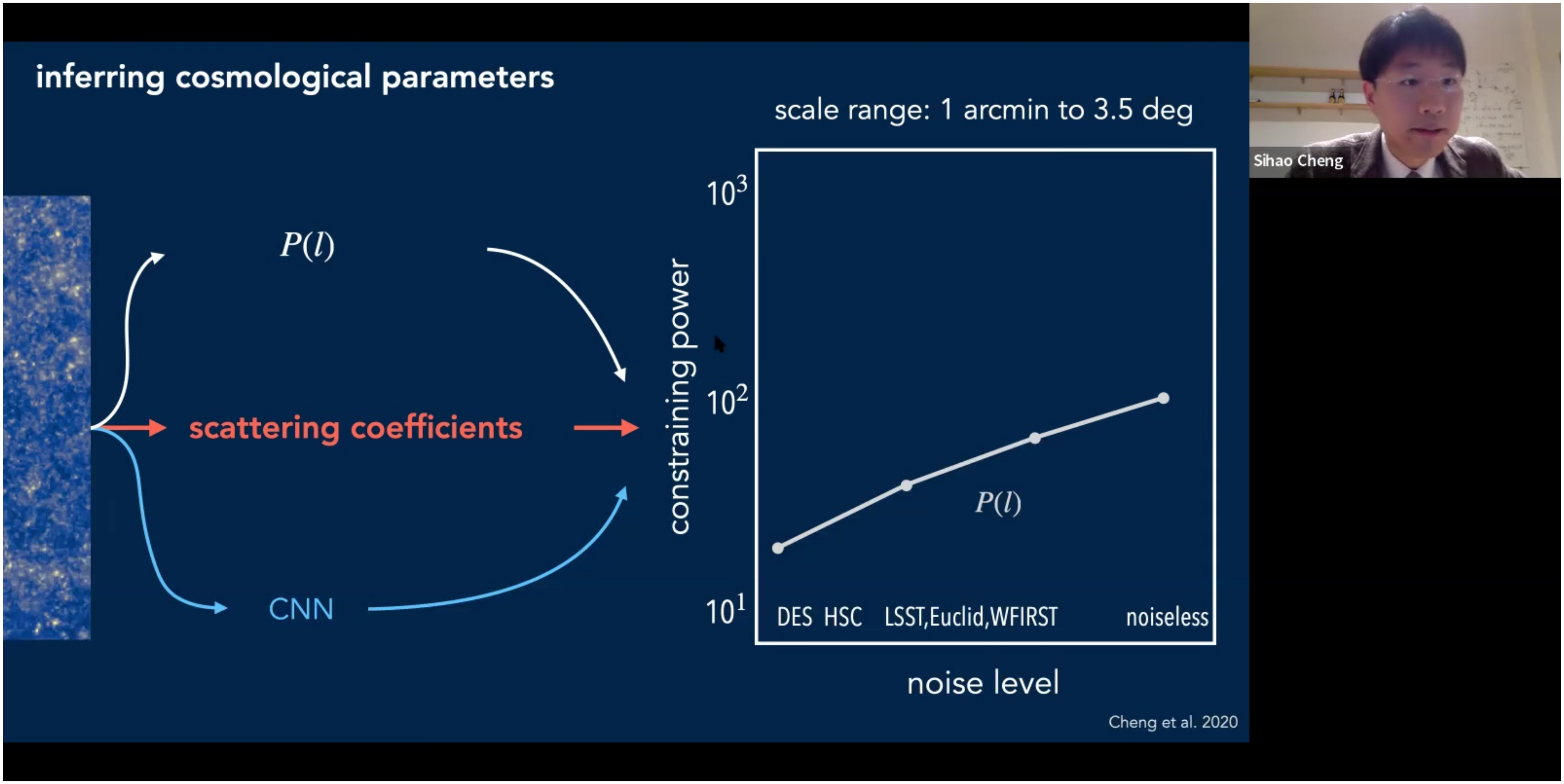


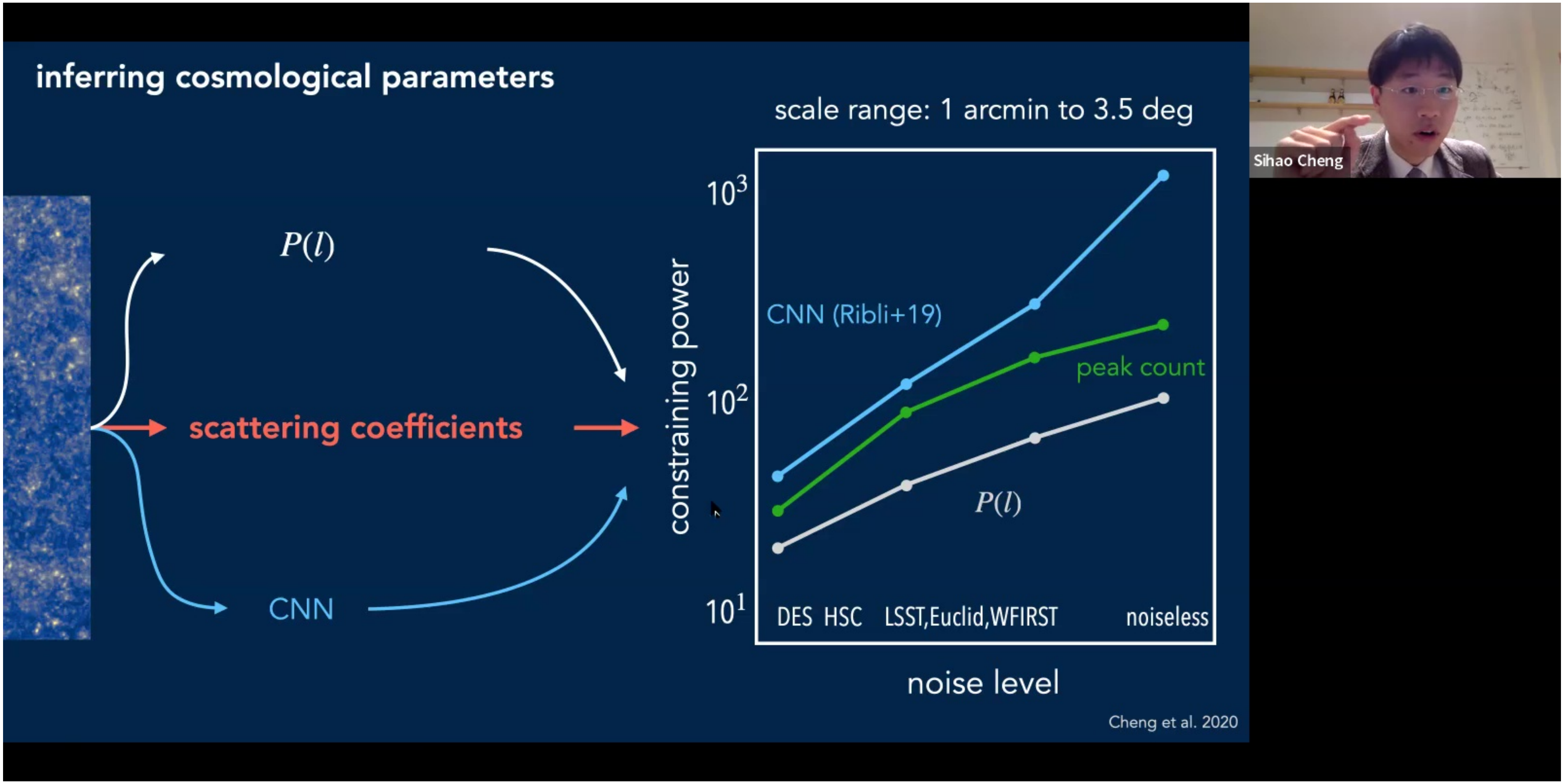
Matilla Zorrilla et al. 2016, Gupta et al. 2018 (Columbia lensing group)

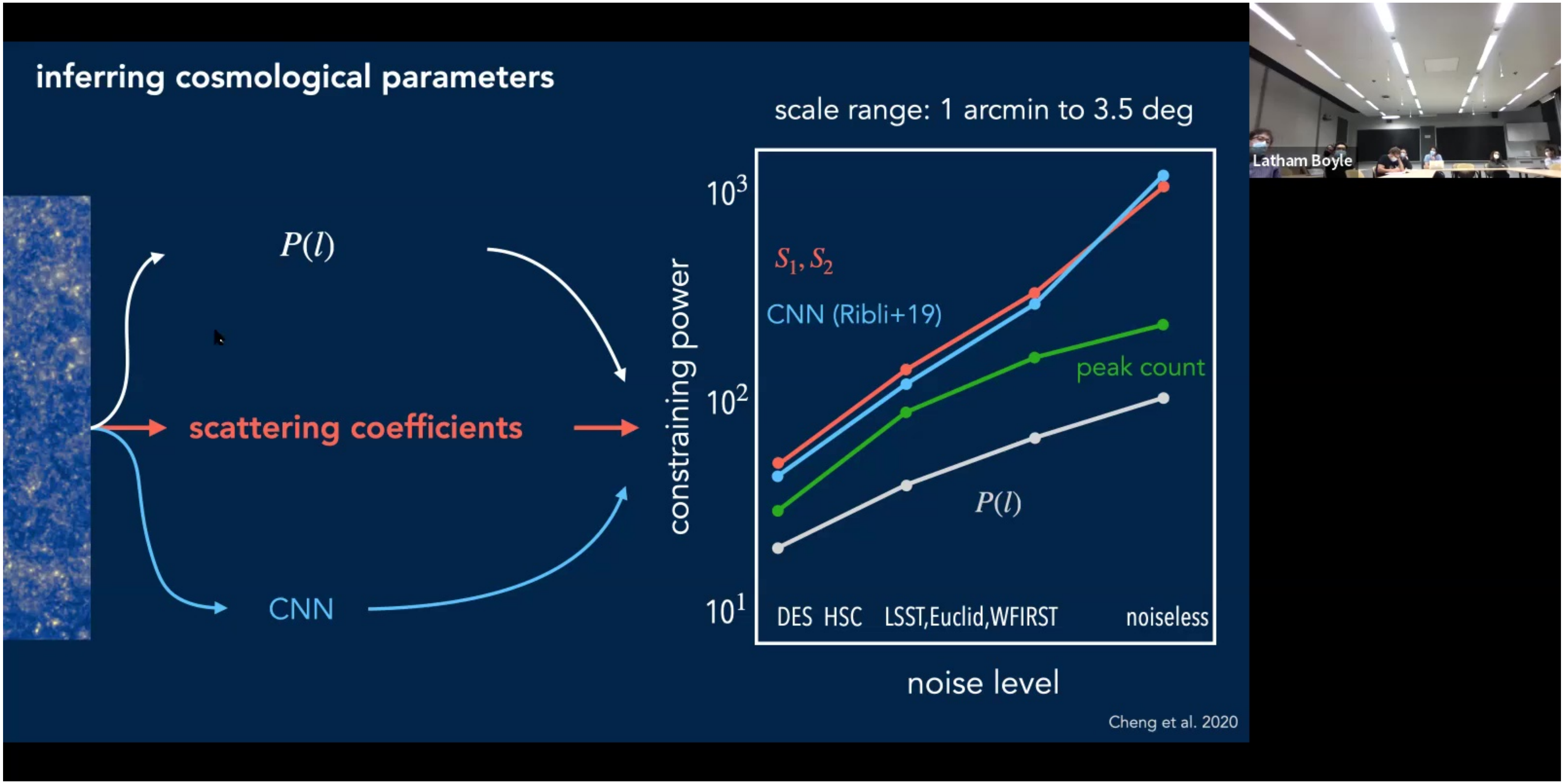


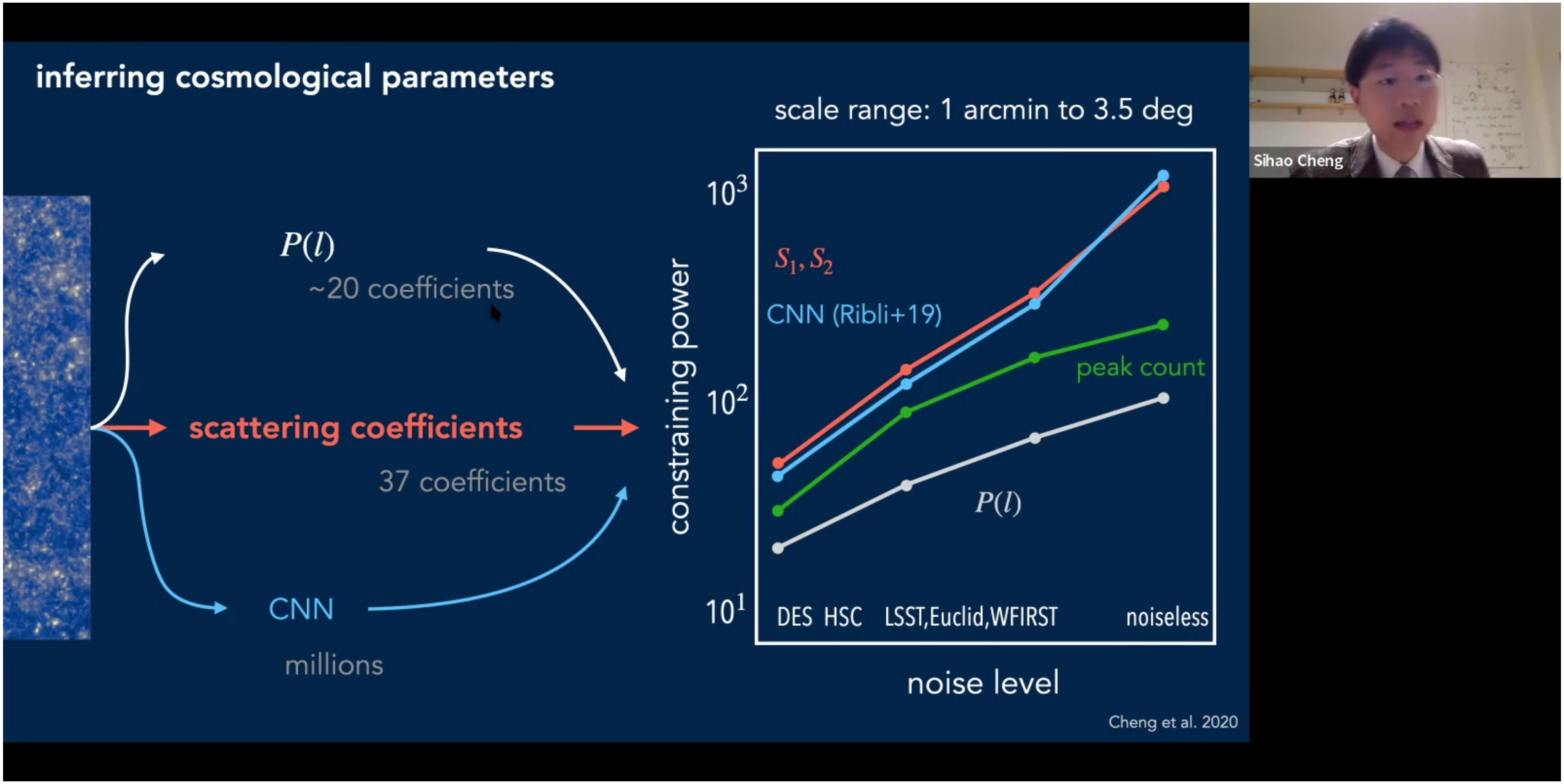
inferring cosmological parameters
3.5x3.5 deg² noiseless map, scale range: 1 arcmin to 3.5 deg
P(l)
$s_0 + s_1 + s_2$
s_2
s_1
σ_8
0.9
0.8
0.7
0.25
0.30
0.35
0.40
Ω_m
P(l)
0.25
0.30
0.35
0.40
Ω_m
Cheng et al. 2020
Sihao Cheng

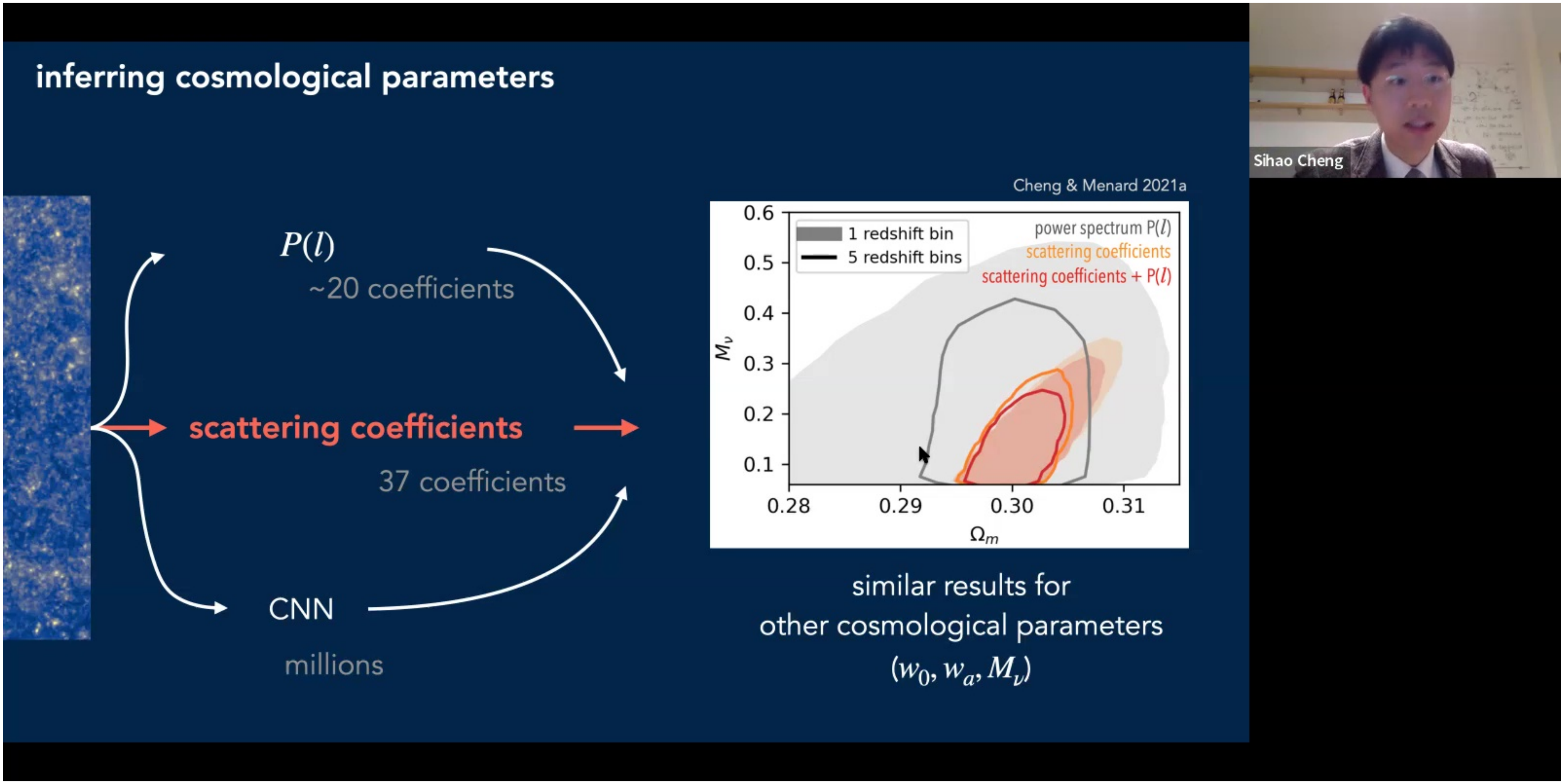

inferring cosmological parameters
3.5x3.5 deg² noiseless map, scale range: 1 arcmin to 3.5 deg
P(l)
s0 + s1 + s2
s2
s1
σ8
0.9
0.8
0.7
0.25
0.30
0.35
0.40
Ωm
P(l)
B(l1, l2) + P(l)
B(l1, l2)
0.25
0.30
0.35
0.40
Ωm
Cheng et al. 2020
Sihao Cheng


inferring cosmological parameters
scale range: 1 arcmin to 3.5 deg
P(l)
scattering coefficients
CNN
constraining power
10^3
10^2
10^1
P(l)
DES HSC LSST,Euclid,WFIRST noiseless
noise level
Cheng et al. 2020
Sihao Cheng


inferring cosmological parameters
scale range: 1 arcmin to 3.5 deg
P(l)
scattering coefficients
CNN
constraining power
10^3
10^2
10^1
CNN (Ribli+19)
peak count
P(l)
DES HSC LSST,Euclid,WFIRST noiseless
noise level
Cheng et al. 2020
Sihao Cheng

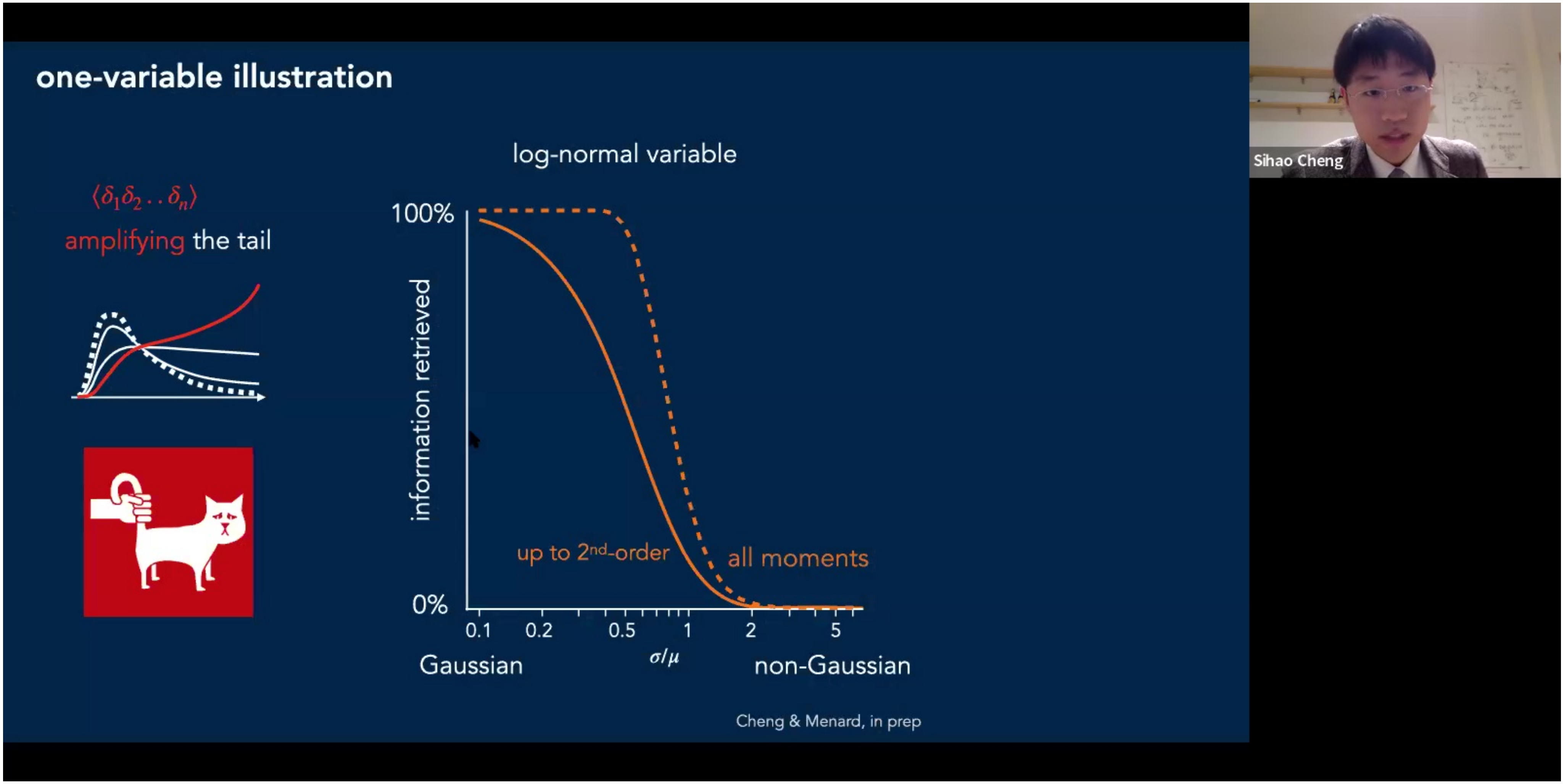

inferring cosmological parameters
P(l)
scattering coefficients
CNN
scale range: 1 arcmin to 3.5 deg
constraining power
10^3
10^2
10^1
S1, S2
CNN (Ribli+19)
peak count
P(l)
DES HSC
LSST,Euclid,WFIRST
noiseless
noise level
Cheng et al. 2020
Latham Boyle

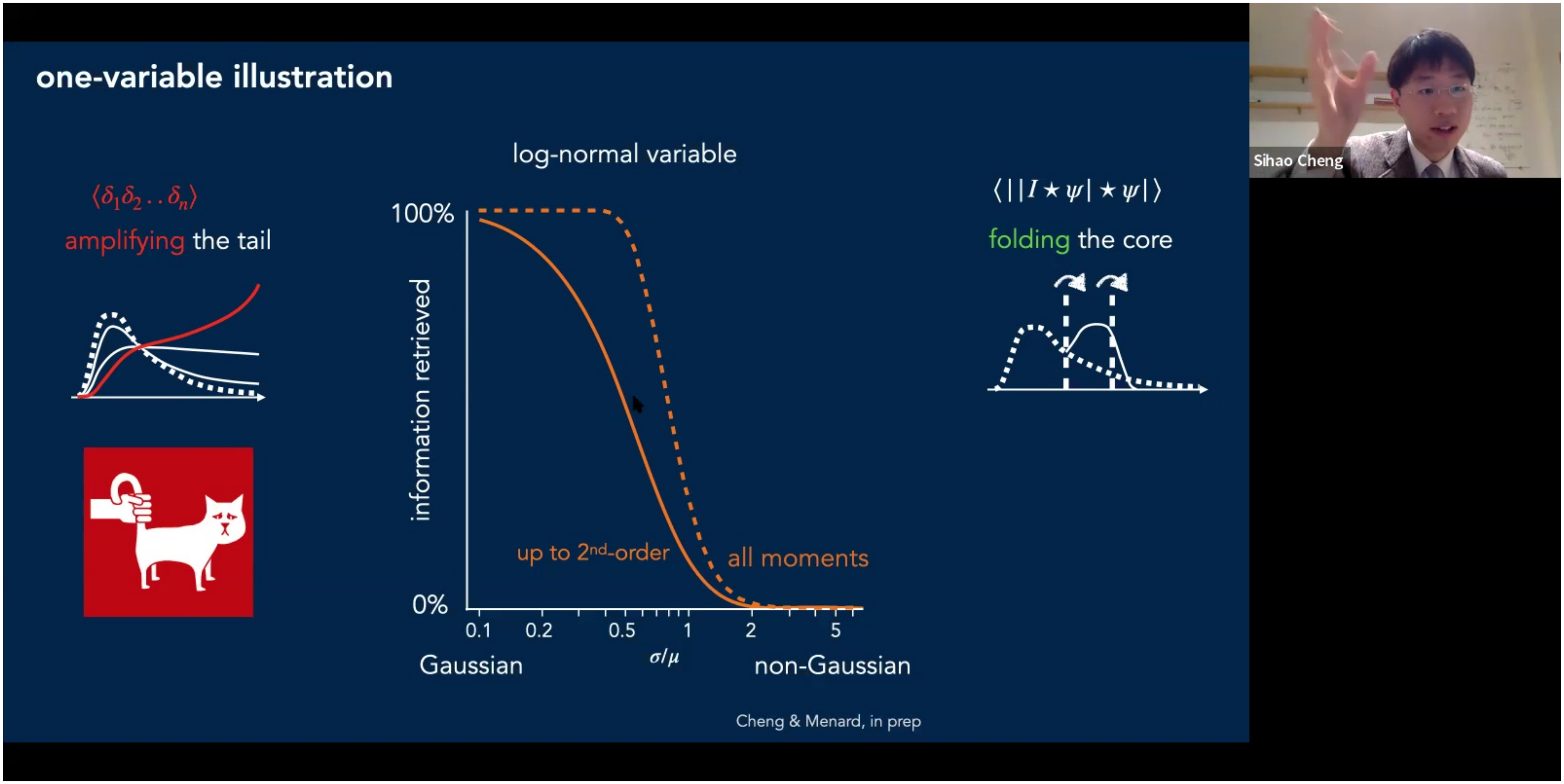

inferring cosmological parameters
scale range: 1 arcmin to 3.5 deg
P(l)
~20 coefficients
scattering coefficients
37 coefficients
CNN
millions
constraining power
10^3
10^2
10^1
S1, S2
CNN (Ribli+19)
peak count
P(l)
DES HSC LSST,Euclid,WFIRST noiseless
noise level
Cheng et al. 2020
Sihao Cheng


inferring cosmological parameters
P(l)
~20 coefficients
scattering coefficients
37 coefficients
CNN
millions
Cheng & Menard 2021a
1 redshift bin
5 redshift bins
power spectrum P(l)
scattering coefficients
scattering coefficients + P(l)
0.6
0.5
0.4
0.3
0.2
0.1
0.28
0.29
0.30
0.31
M_ν
Ω_m
similar results for
other cosmological parameters
(w_0, w_a, M_ν)
Sihao Cheng

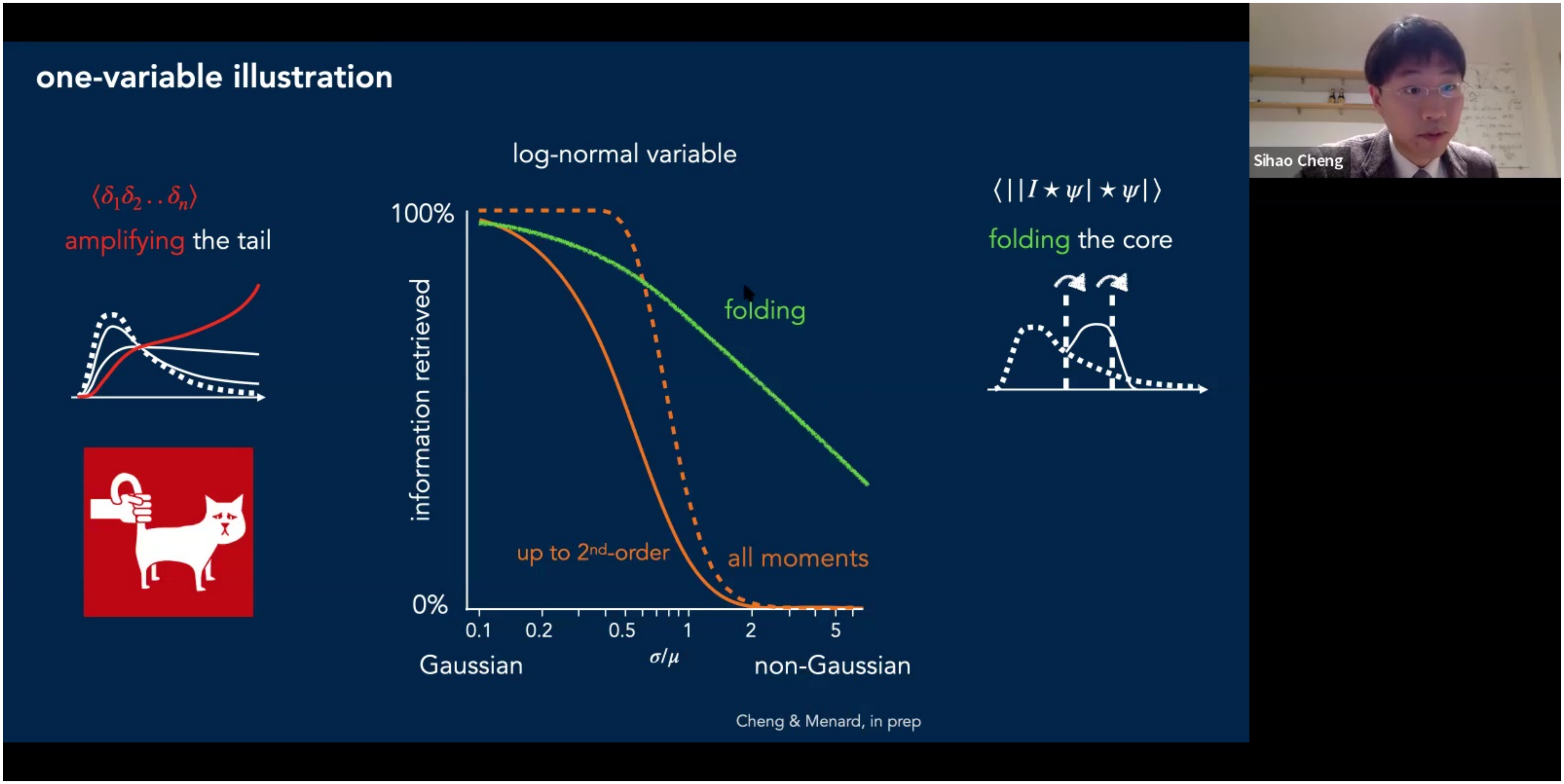

towards real data
Subaru telescope survey
Subaru 8.2m
1 part of 100 million of human eye limit)
1400 deg^2 , 330 nights, 100 million galaxies with shape
image credit: NAOJ, HSC-SSP
Sihao Cheng


towards real data
Subaru telescope survey
Subaru 8.2m
1 part of 100 million of human eye limit)
1400 deg^2 , 330 nights, 100 million galaxies with shape
first real constraint in a few months!
image credit: NAOJ, HSC-SSP
Sihao Cheng


one-variable illustration
$\langle \delta_1 \delta_2 .. \delta_n \rangle$
amplifying the tail
log-normal variable
100%
0%
information retrieved
up to 2nd-order
all moments
0.1
0.2
0.5
1
2
5
Gaussian
σ/μ
non-Gaussian
Cheng & Menard, in prep
Sihao Cheng



one-variable illustration
log-normal variable
$\langle \delta_1 \delta_2 .. \delta_n \rangle$
amplifying the tail
100%
0%
information retrieved
up to 2nd-order
all moments
0.1
0.2
0.5
1
2
5
Gaussian
σ/μ
non-Gaussian
$\langle | | I \star \psi | \star \psi | \rangle$
folding the core
Cheng & Menard, in prep
Sihao Cheng


one-variable illustration
log-normal variable
$\langle \delta_1 \delta_2 .. \delta_n \rangle$
amplifying the tail
100%
0%
information retrieved
folding
up to 2nd-order
all moments
0.1 0.2 0.5 1 2 5
σ/μ
Gaussian
non-Gaussian
$\langle | | I \star \psi | \star \psi | \rangle$
folding the core
Cheng & Menard, in prep
Sihao Cheng

interesting directions in astrophysics

1D, 2D, 3D fields

extensions of the idea

$$I \qquad \left| I \star \psi_1 \right| \qquad \left| \left| I \star \psi_1 \right| \star \psi_2 \right| \quad \dots$$

$$\left| I \star \psi_2 \right|$$

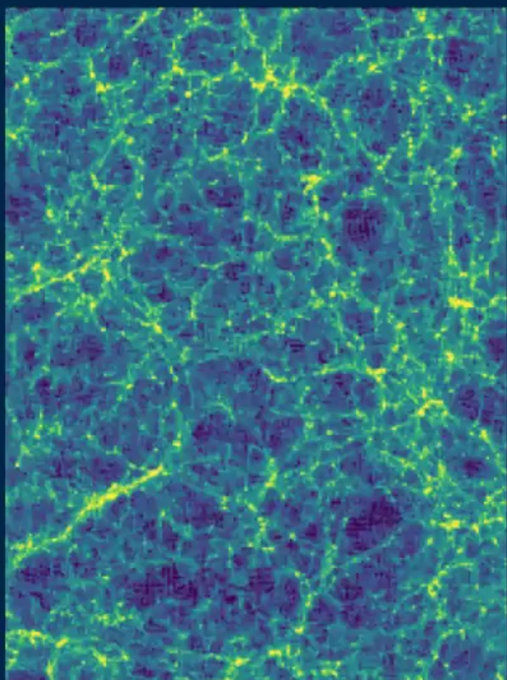
cross correlation: phase harmonics

Mallat, Zhang, & Rochette 2018

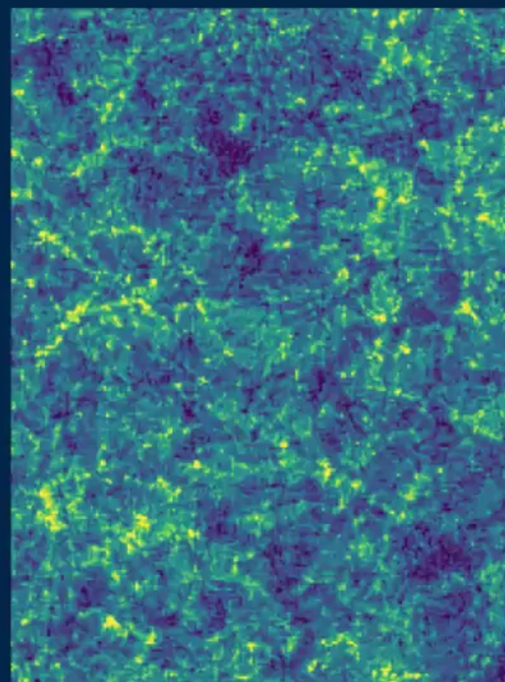
learning ability

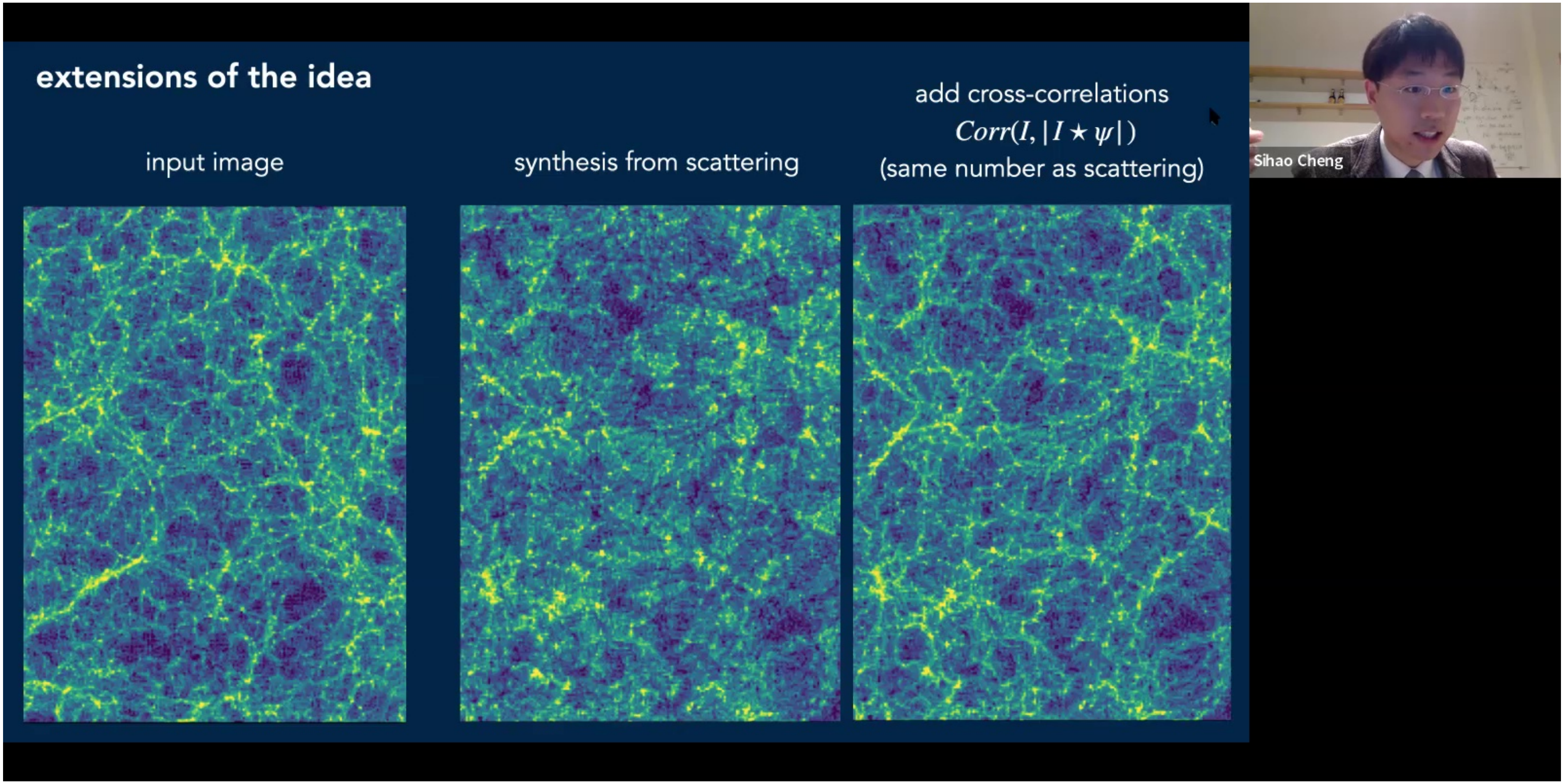
extensions of the idea

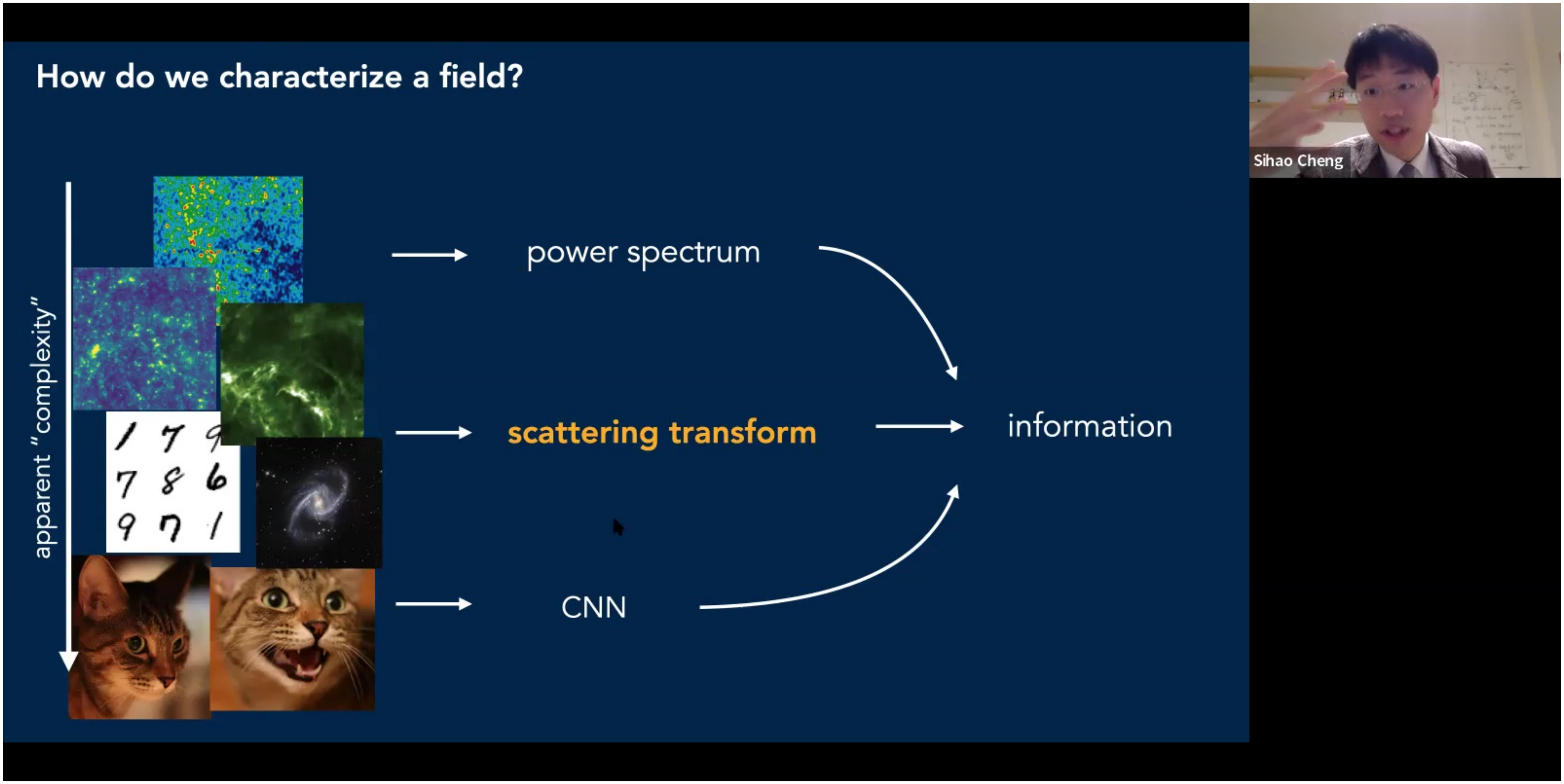
input image

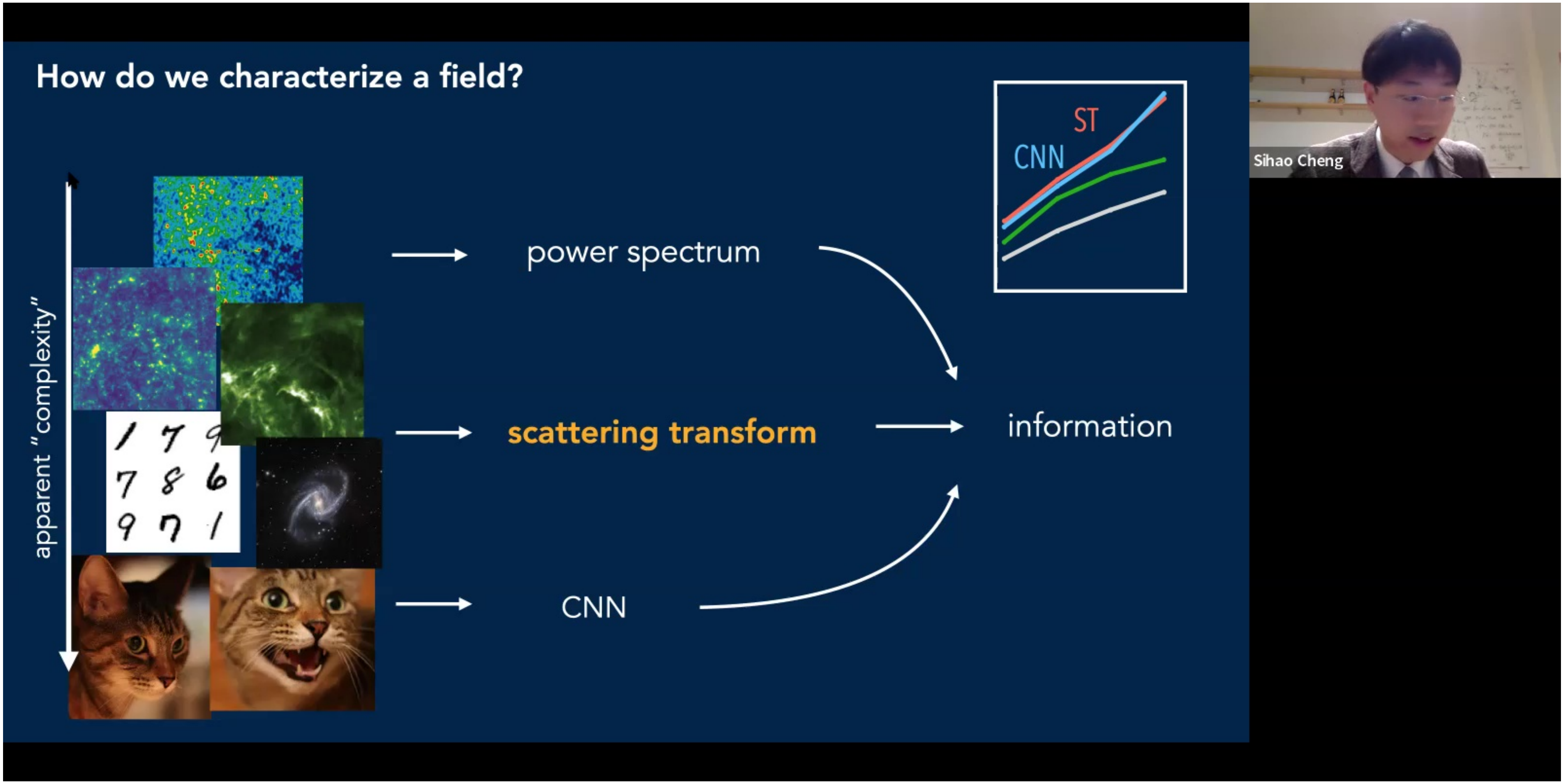


synthesis from scattering




extensions of the idea
input image
synthesis from scattering
add cross-correlations
$Corr(I, |I \star \psi|)$
(same number as scattering)
Sihao Cheng


How do we characterize a field?
apparent "complexity"
power spectrum
scattering transform
CNN
information
Sihao Cheng


How do we characterize a field?
apparent "complexity"
power spectrum
scattering transform
CNN
information
ST
CNN
Sihao Cheng