Cosmic fields beyond 2pt:
practical challenges and opportunities

Chihway Chang (UChicago/KICP)

with Dhayaa Anbajagane, Yuuki Omori and many others in DES/DESC

Outline

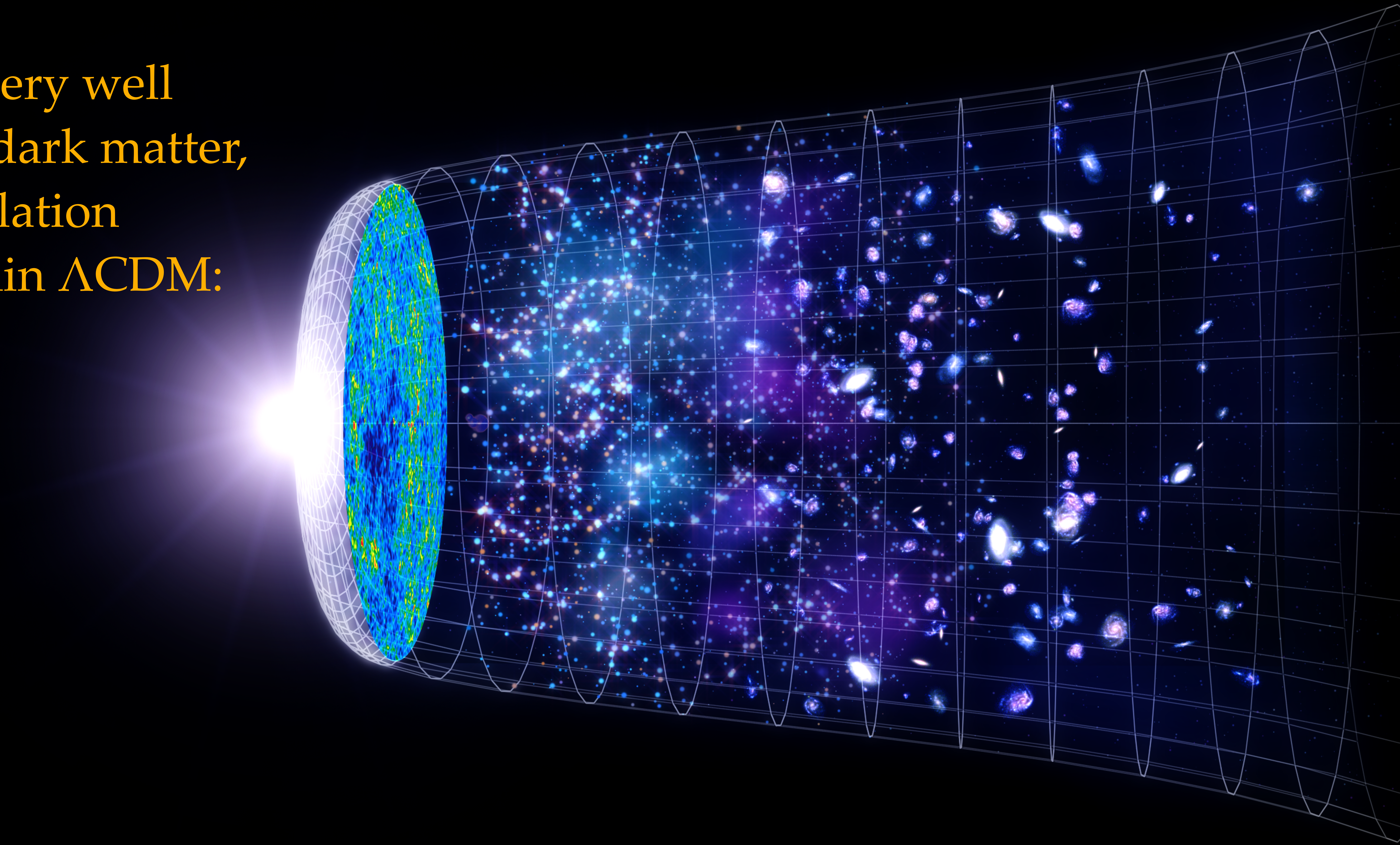
- The ΛCDM paradigm and extracting information beyond 2pt
- Practical challenges: beyond 2pt systematics
- Opportunities: primordial non-Gaussianity
- Towards field-level inference
- Summary & outlook

Outline

- **The ΛCDM paradigm and extracting information beyond 2pt**
- Practical challenges: beyond 2pt systematics
- Opportunities: primordial non-Gaussianity
- Towards field-level inference
- Summary & outlook

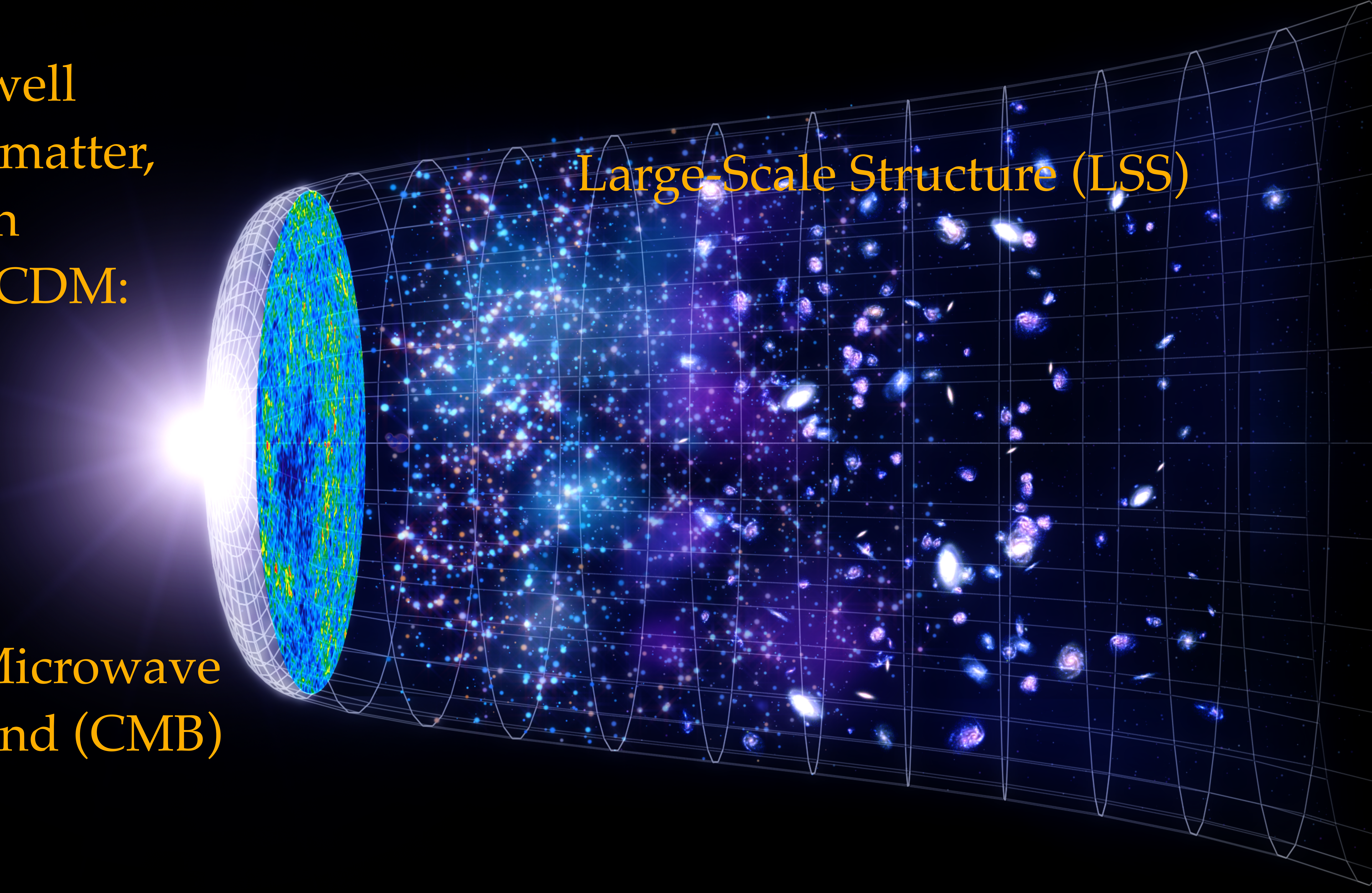
Cosmology as we know it today

- ΛCDM works very well
- Key questions: dark matter, dark energy, inflation
- Curiosities within ΛCDM: H0, S8, w0wa



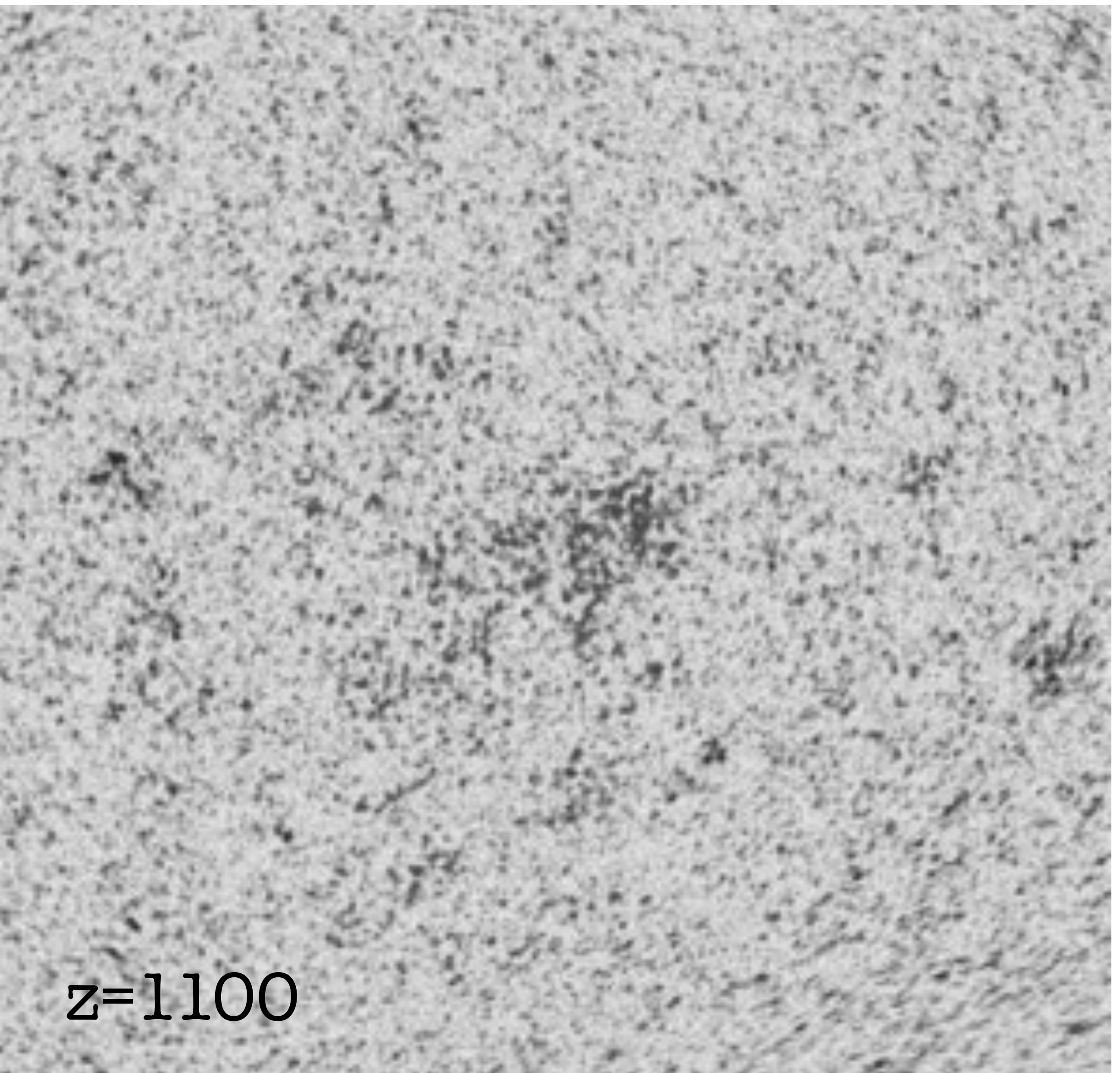
Cosmology as we know it today

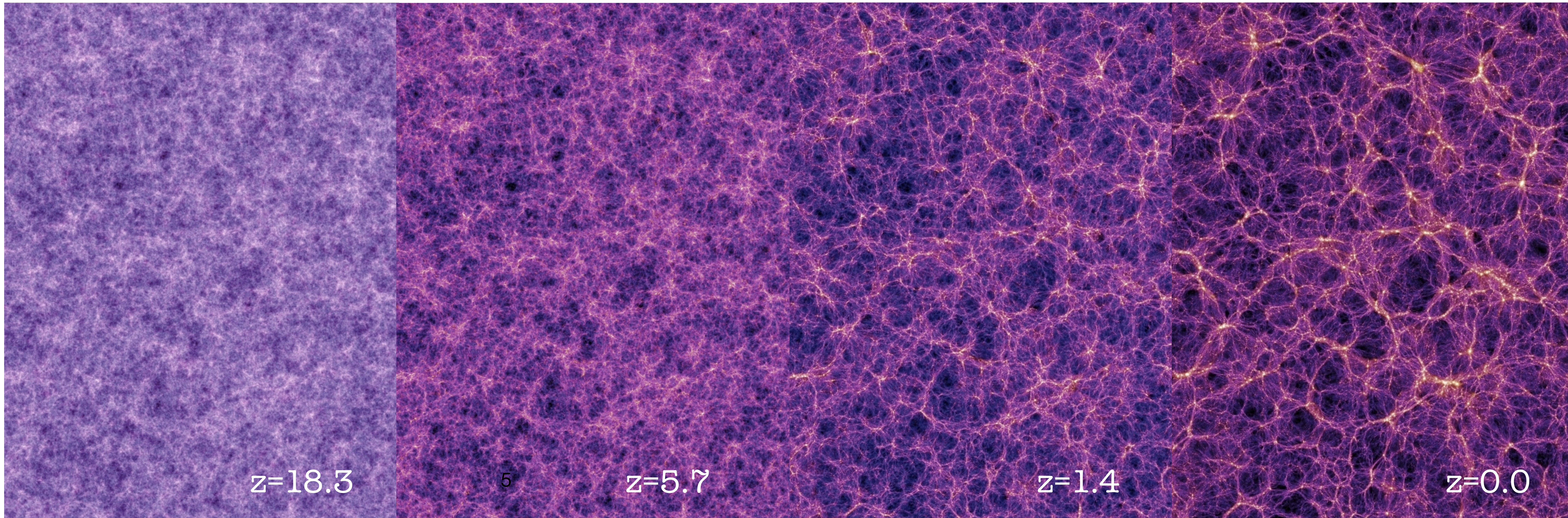
- ΛCDM works very well
- Key questions: dark matter, dark energy, inflation
- Curiosities within ΛCDM: H0, S8, w0wa



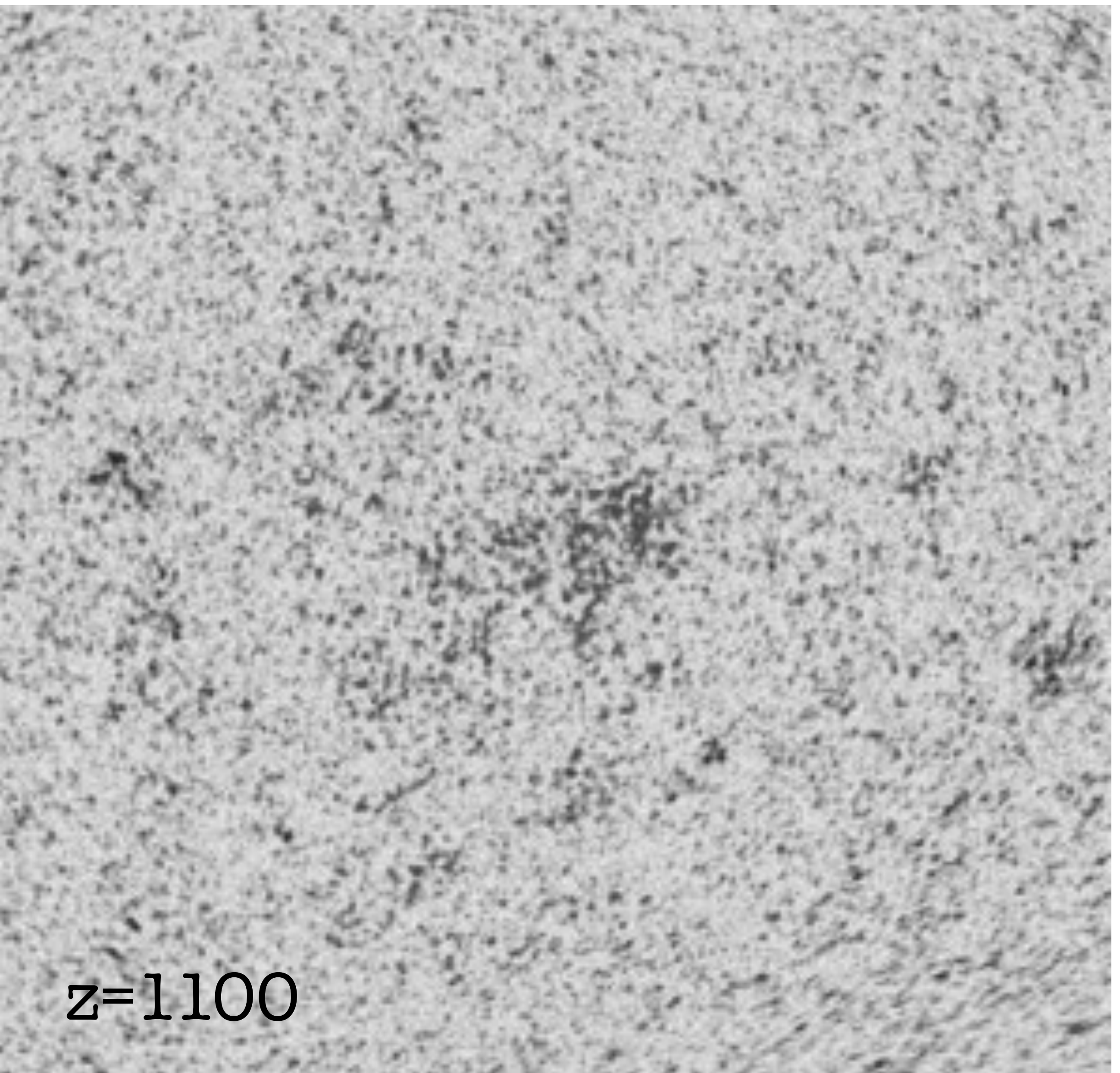
Early-universe physics

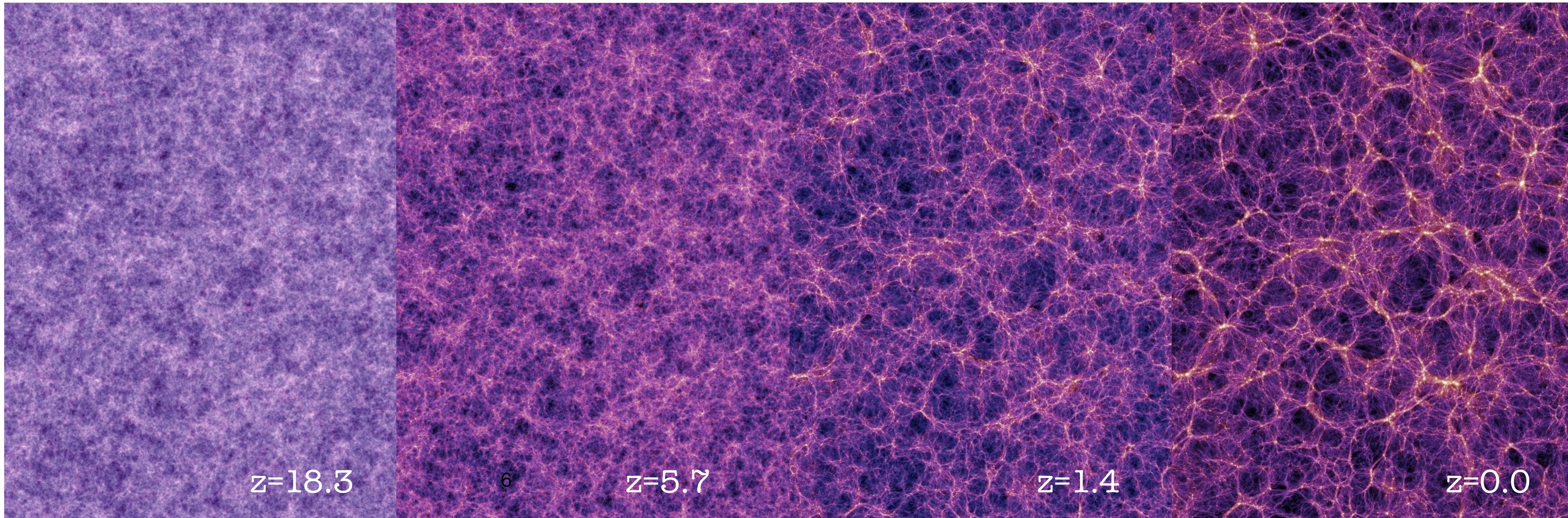
Gravity, dark matter, dark energy



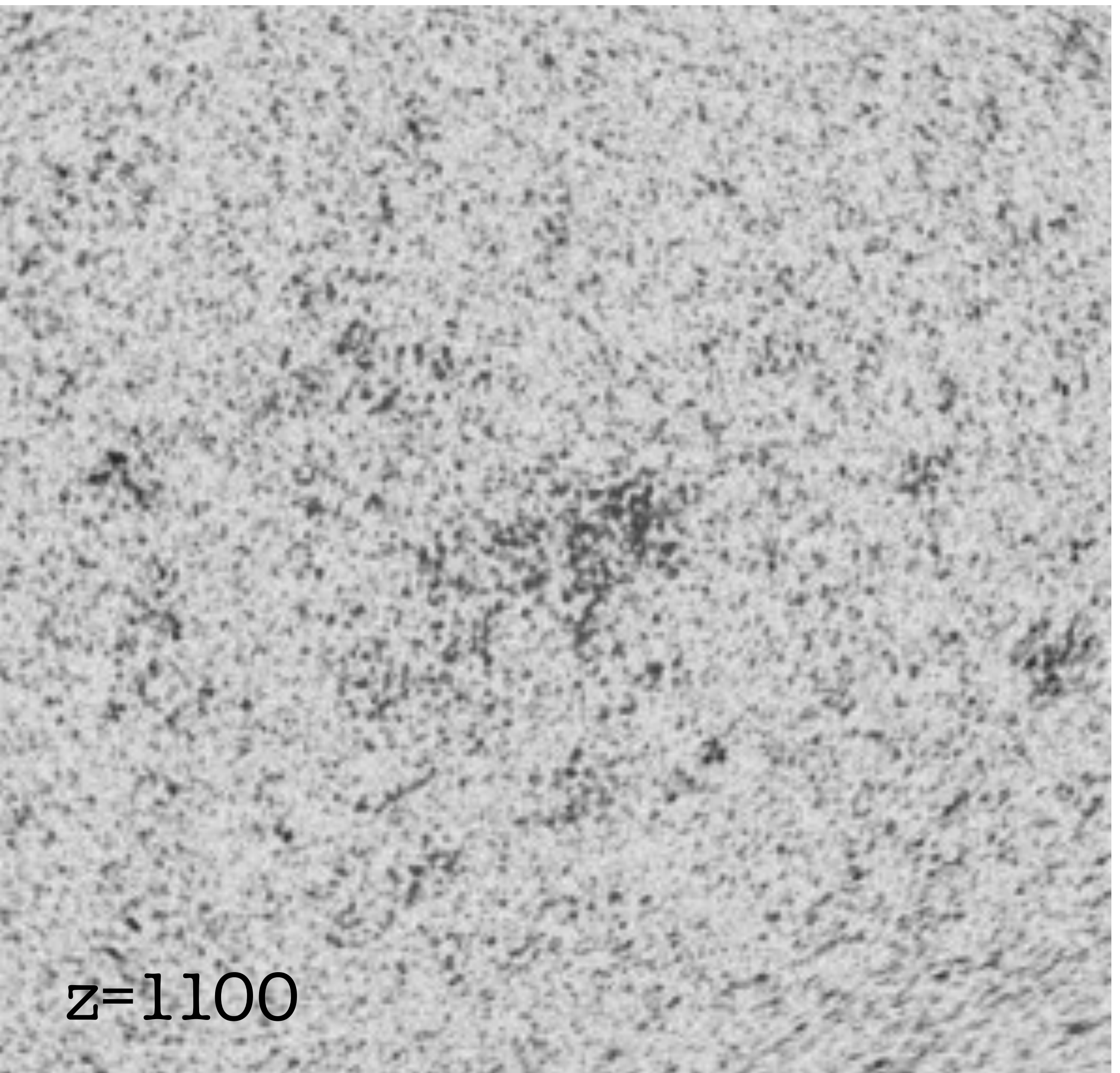


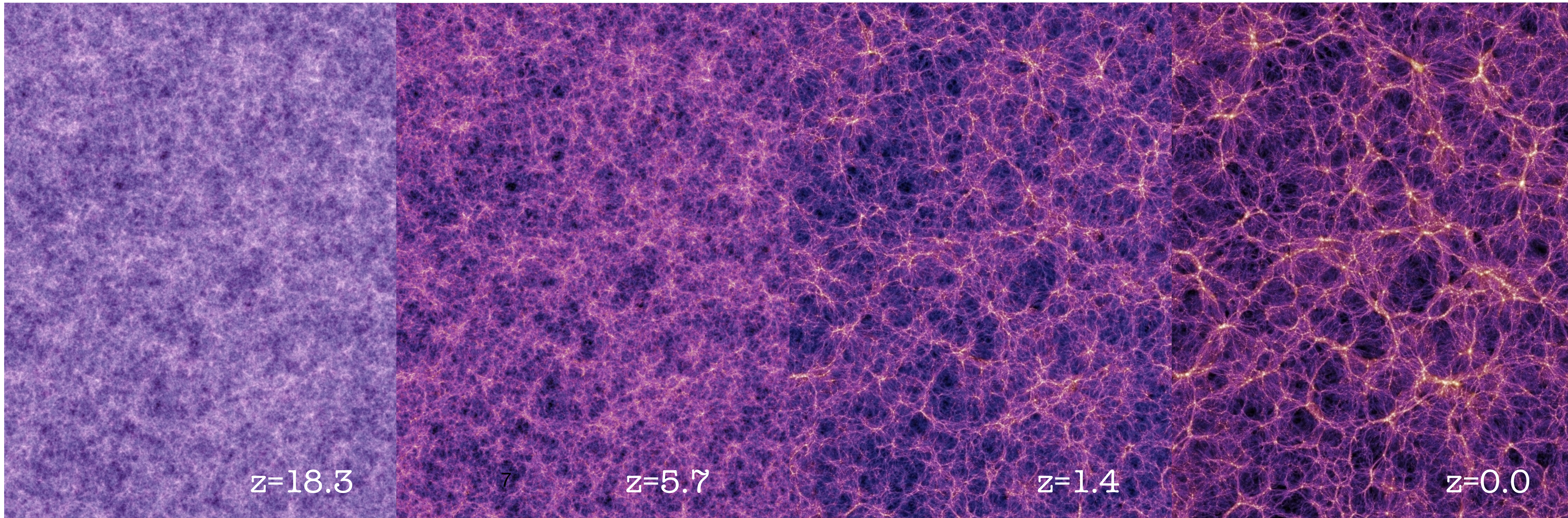
Astrophysics





Astrophysics





The LSS is a powerful probe for

Early-universe physics

ΛCDM (gravity, dark energy, dark matter)

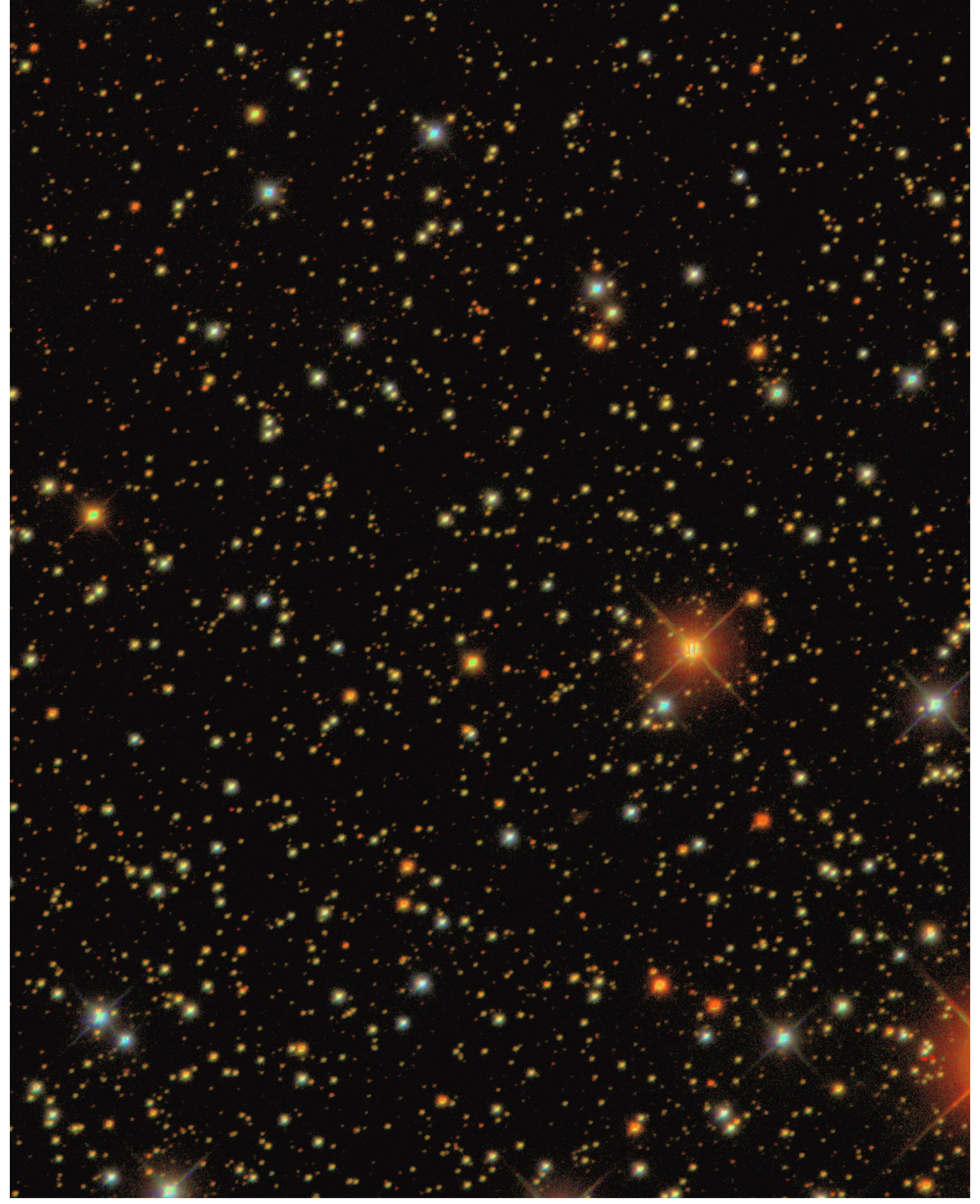
Astrophysics (galaxy evolution)

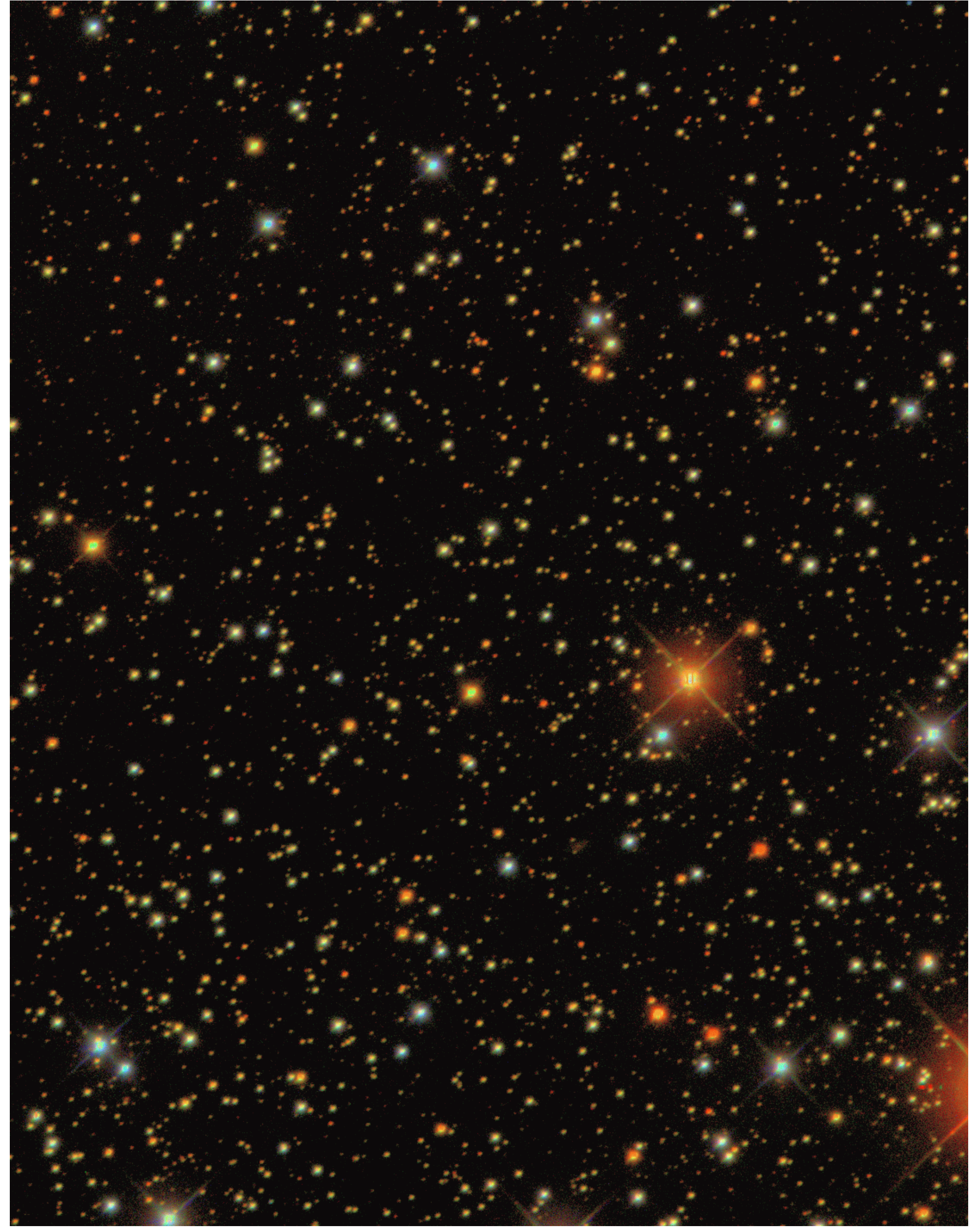
Compared to CMB

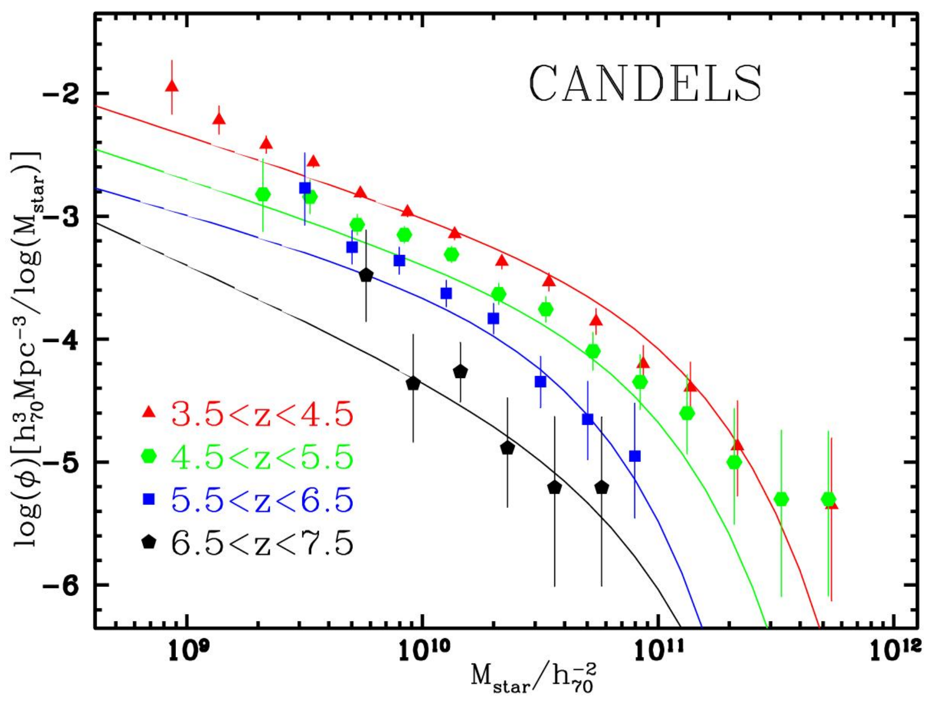
More overall information

The information is more complex

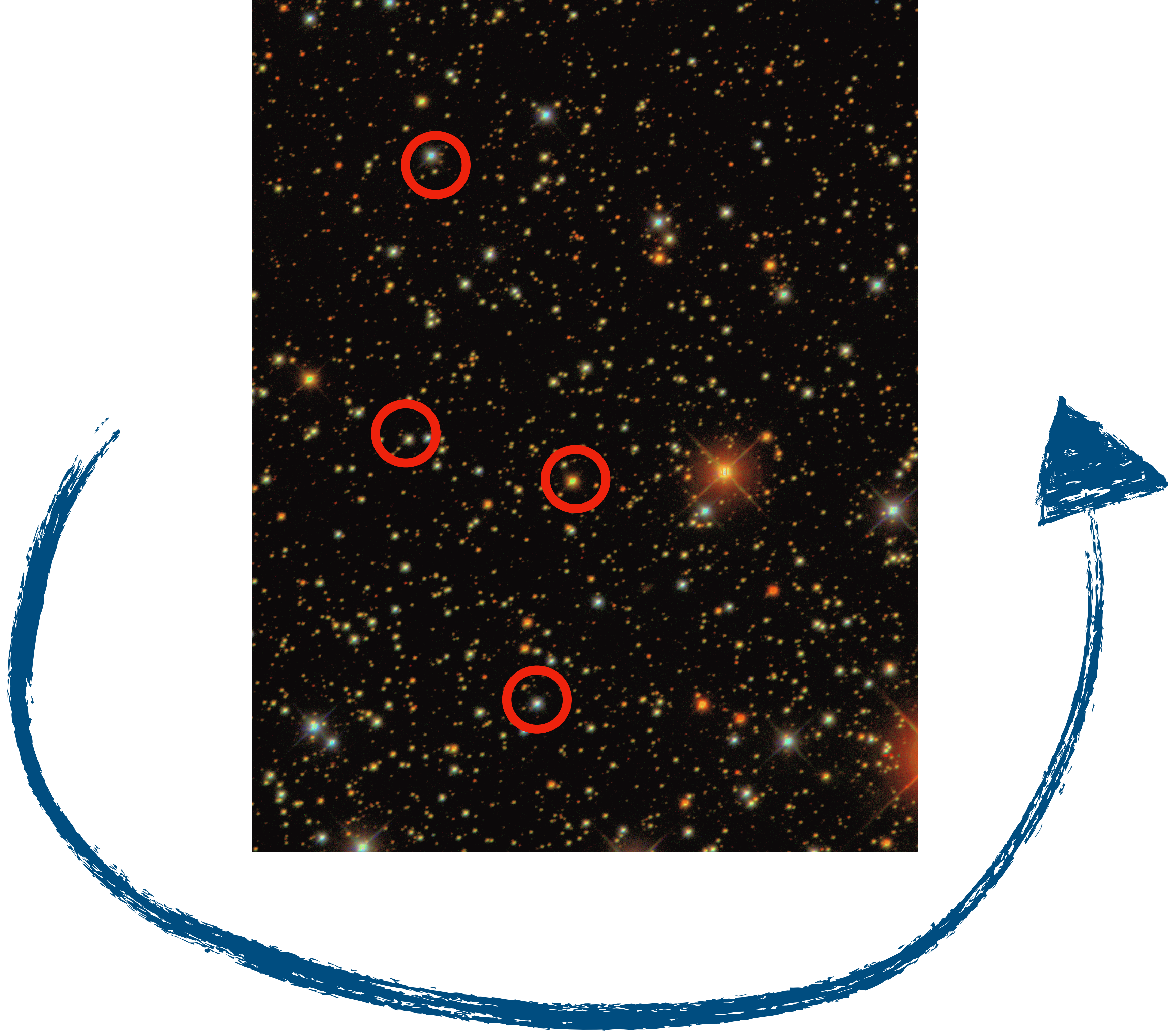
Non-Gaussianity

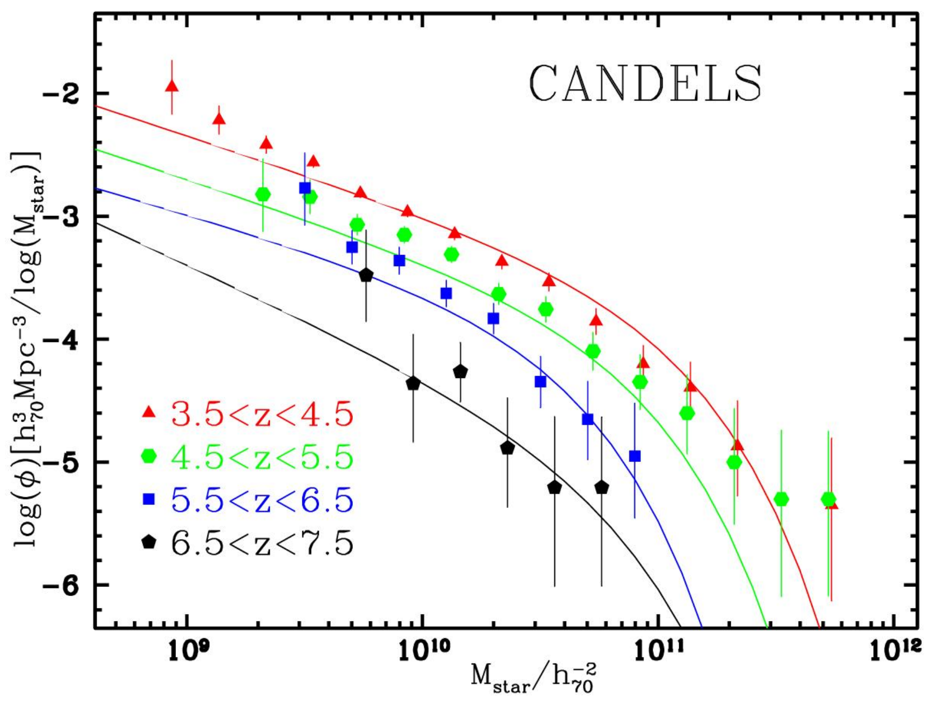




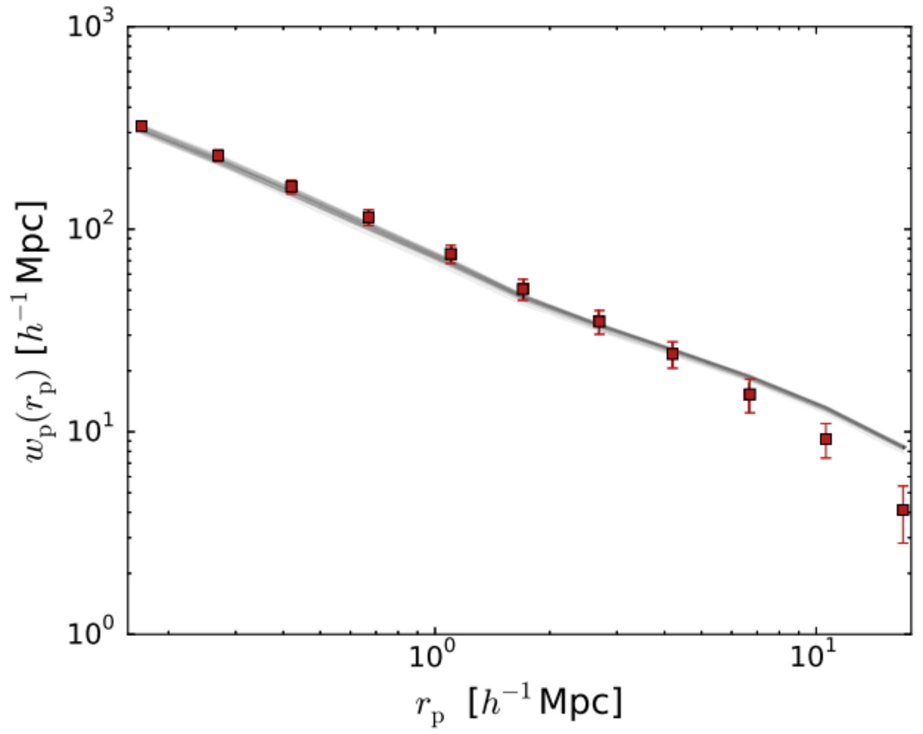


Number counts

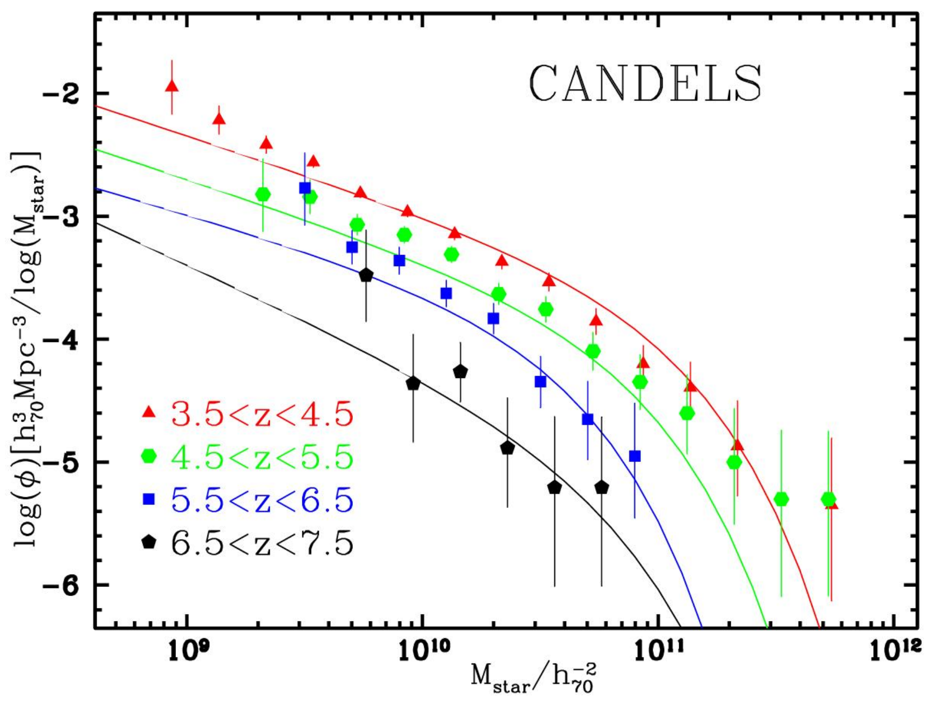




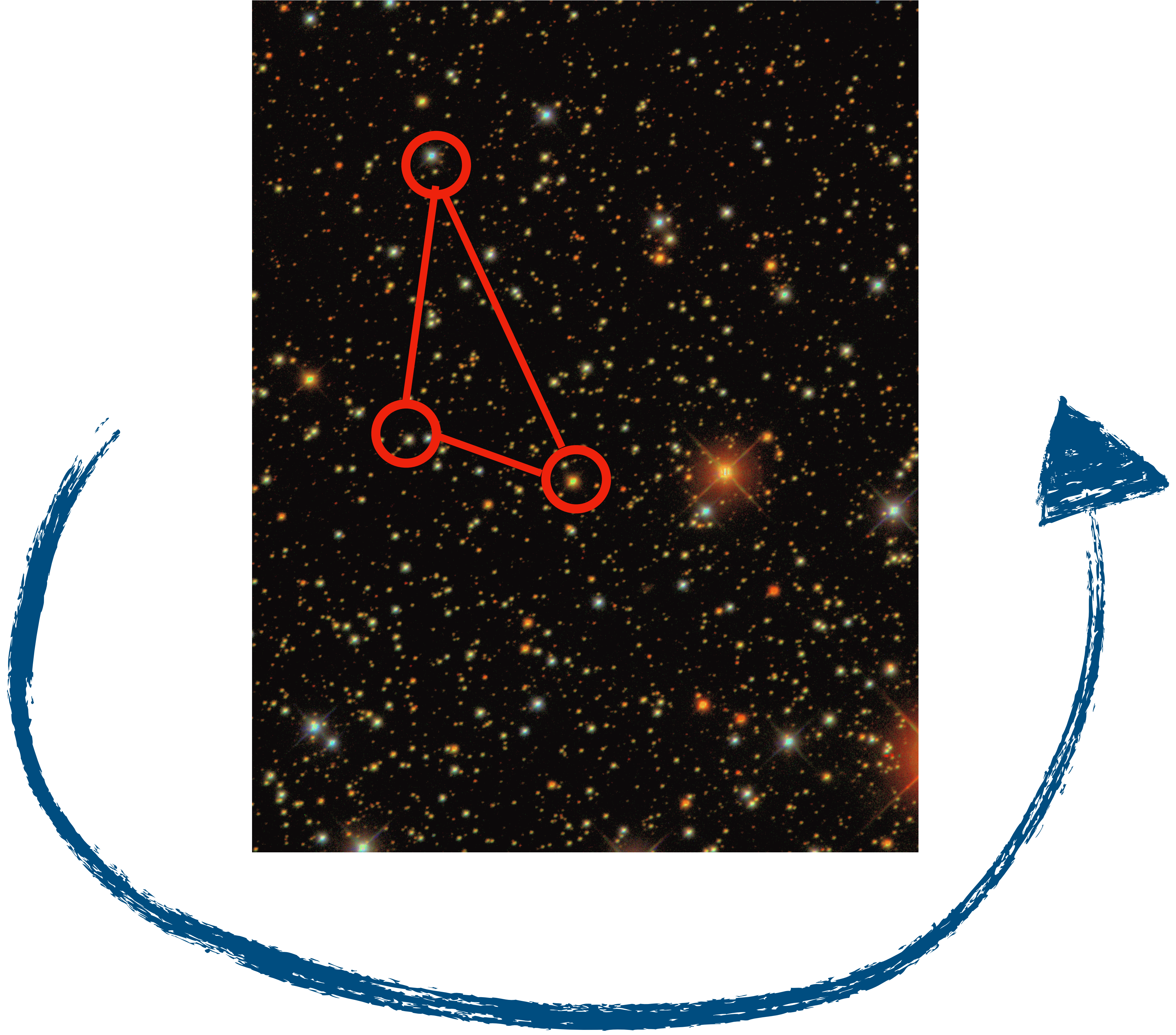
Number counts

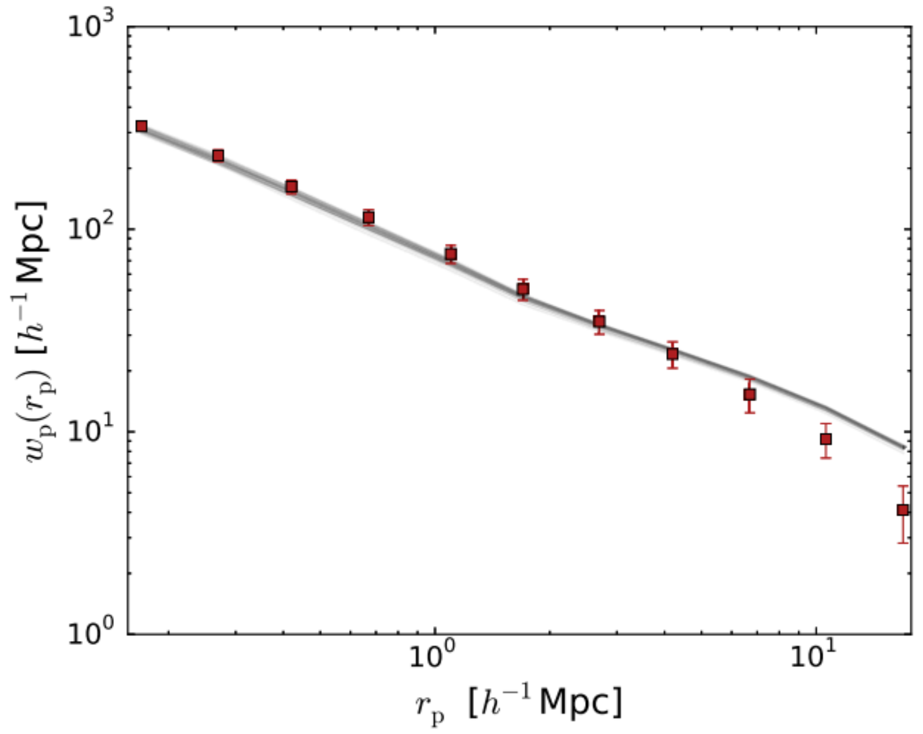


Two-point correlation

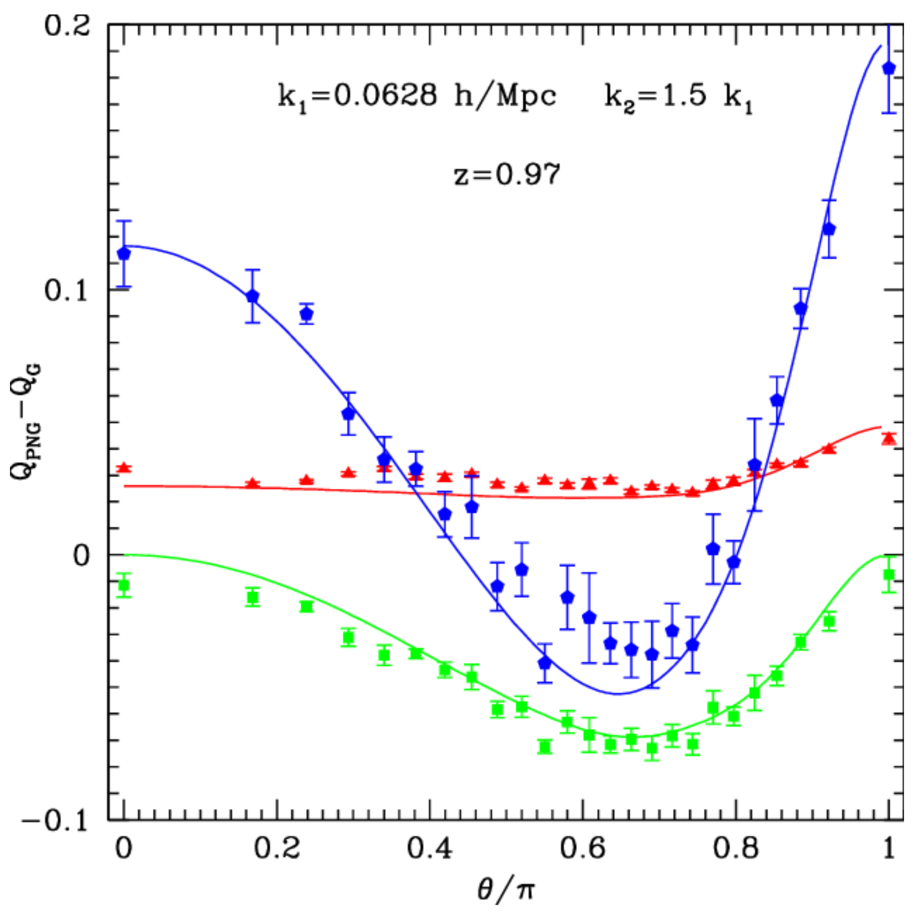


Number counts

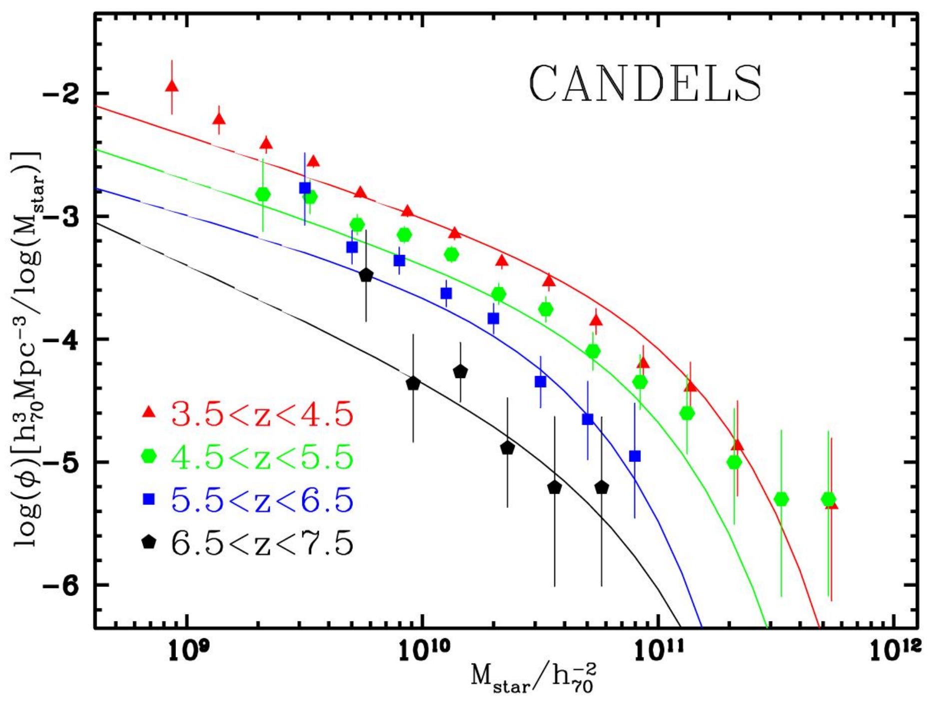




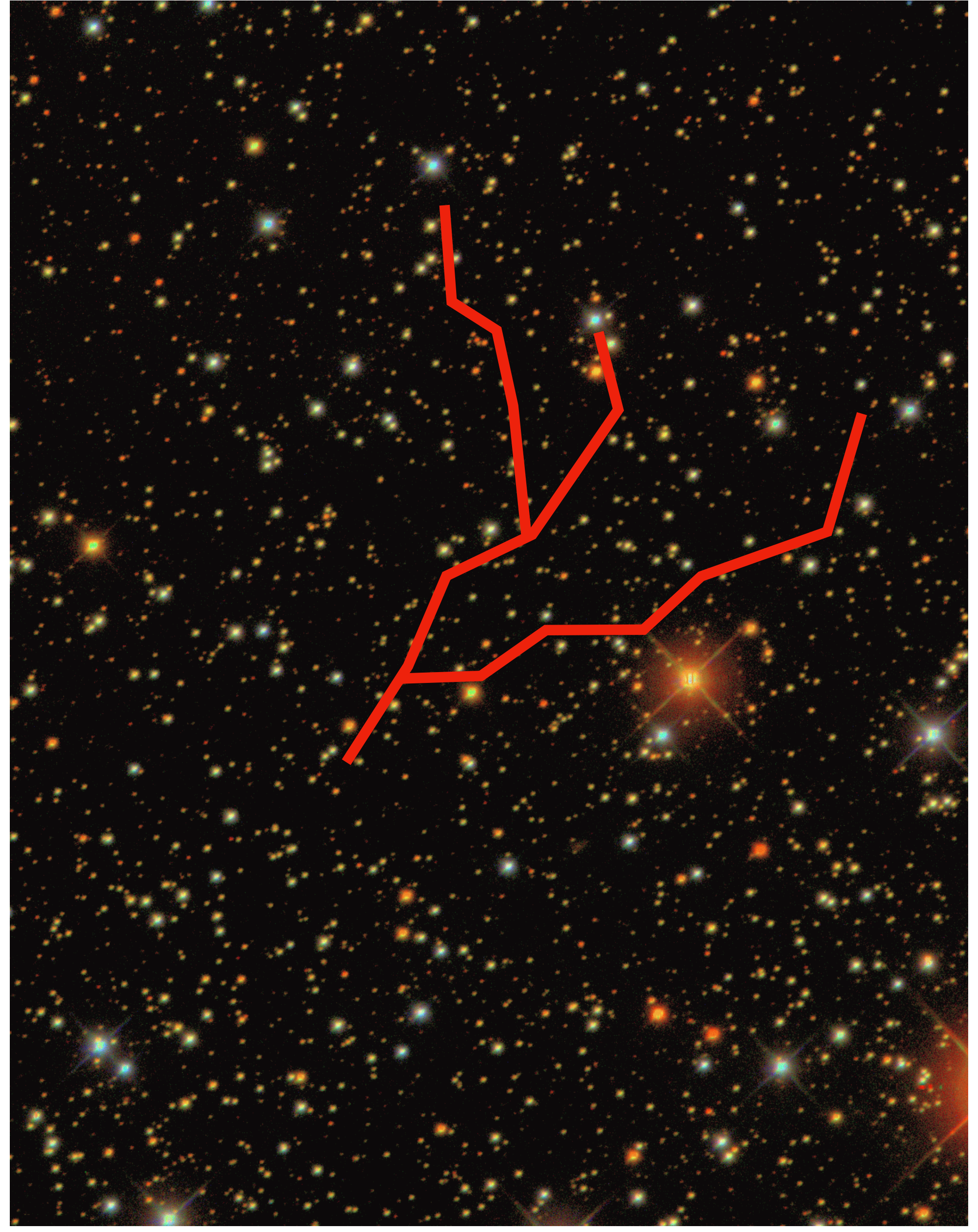
Two-point correlation

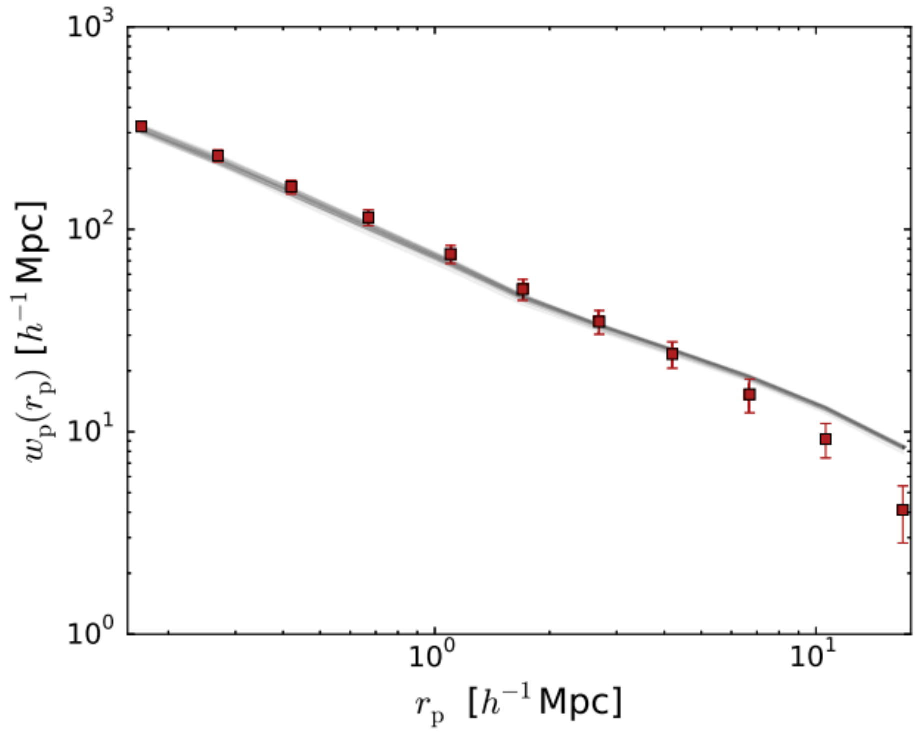


Three-point correlation

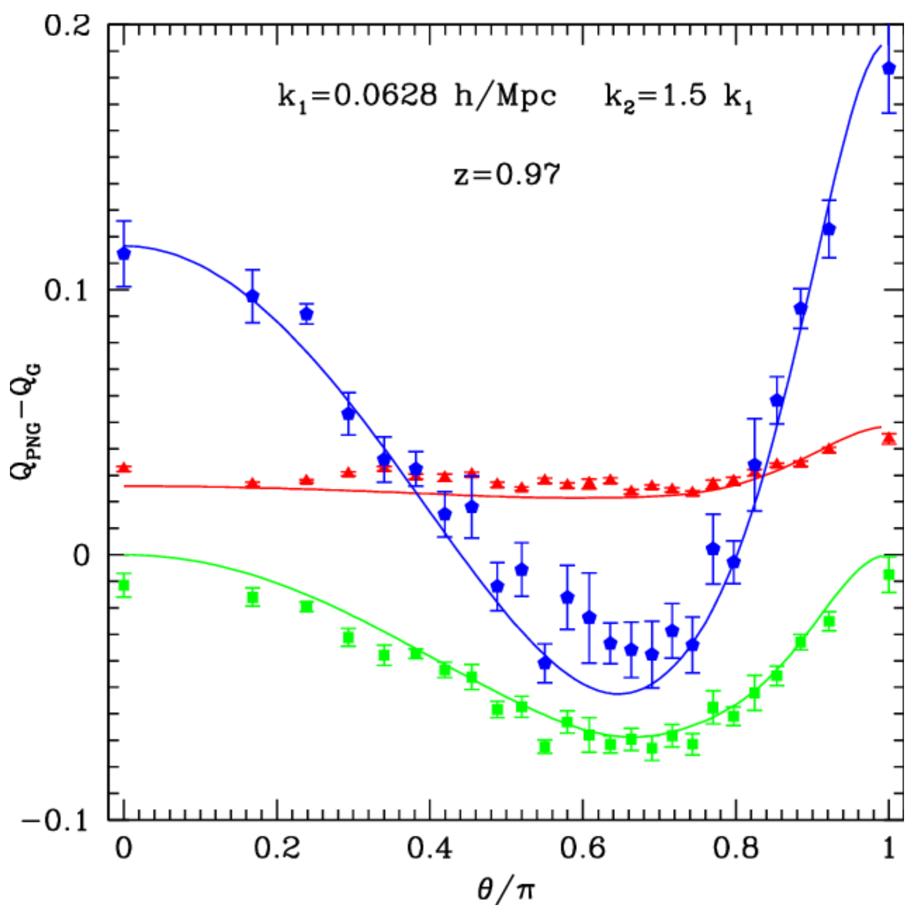


Number counts

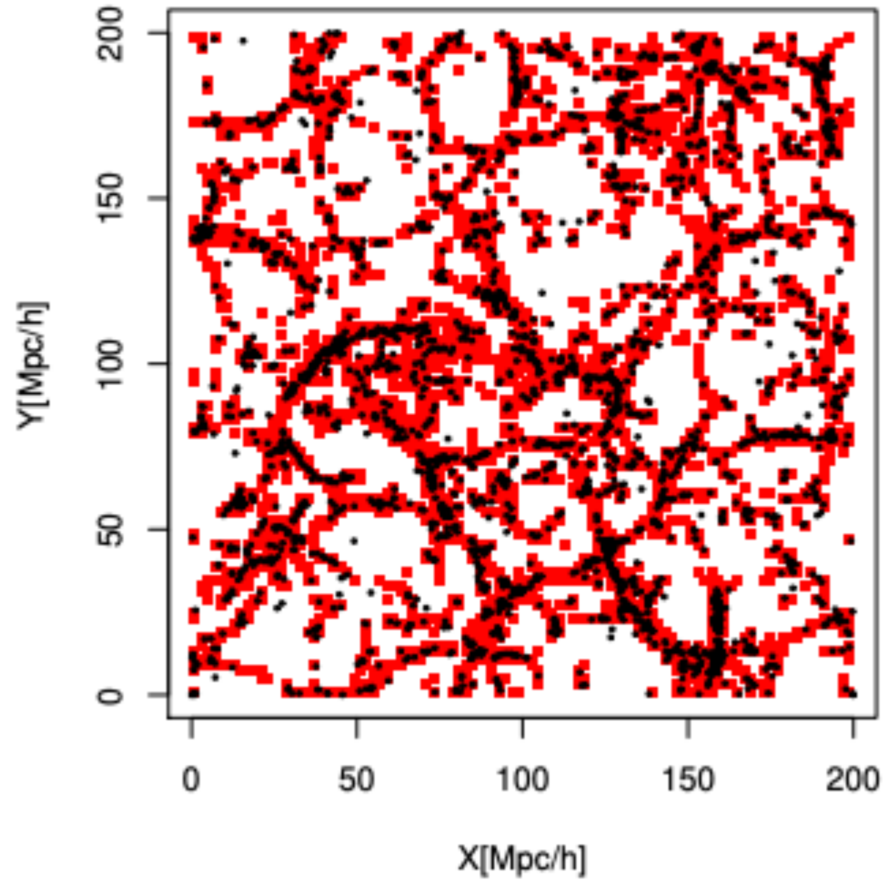




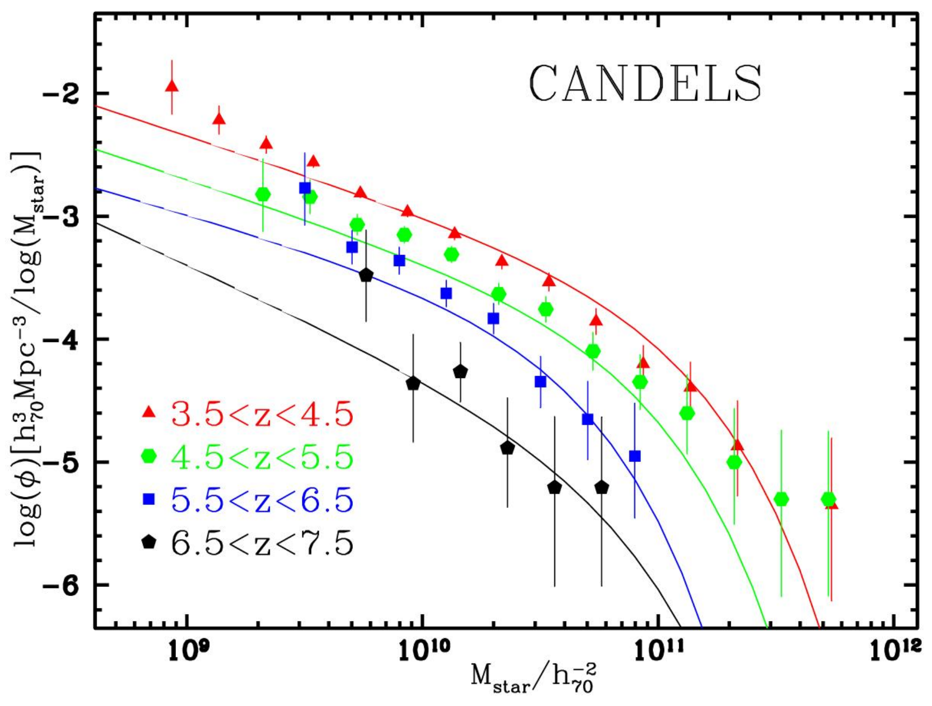
Two-point correlation



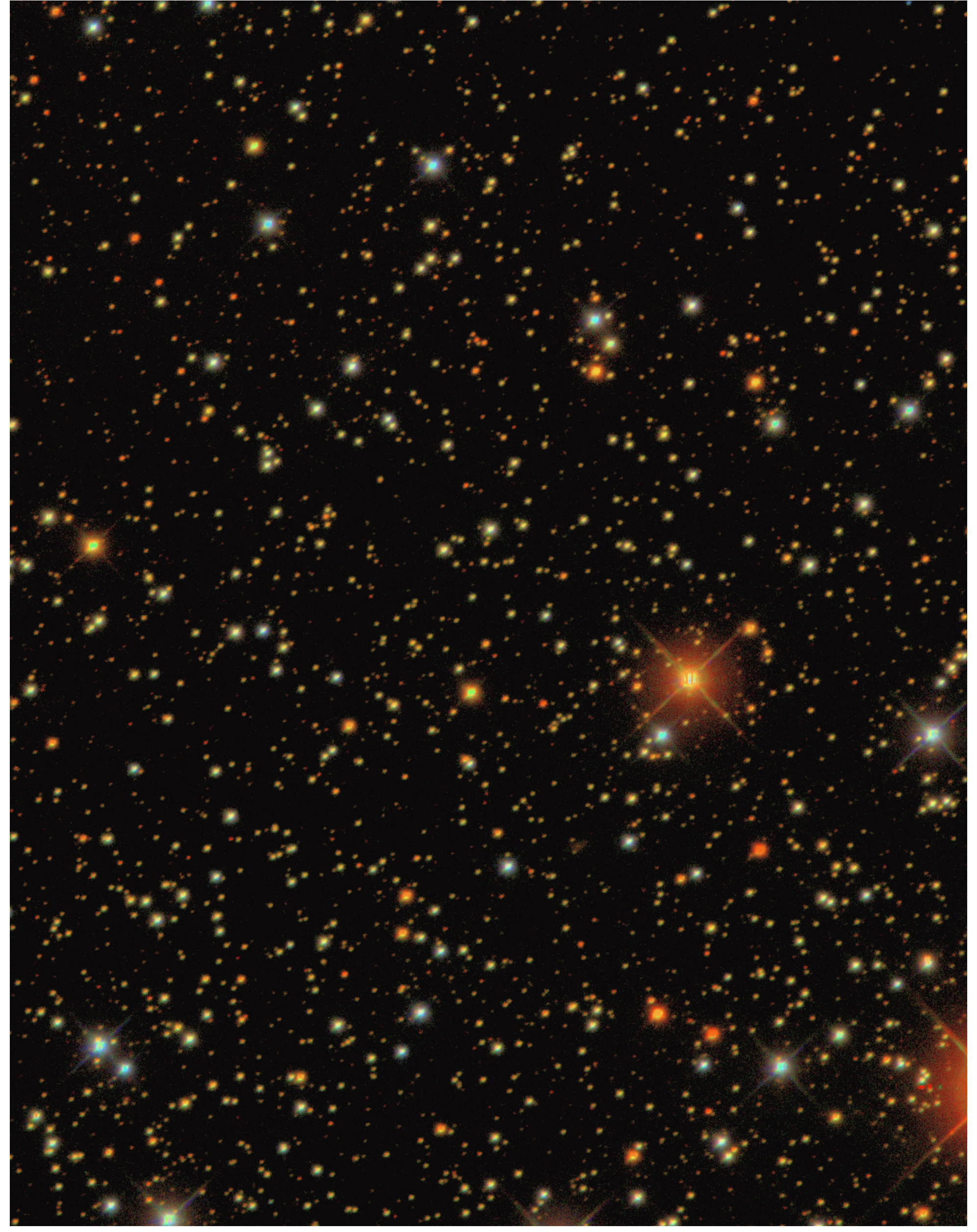
Three-point correlation

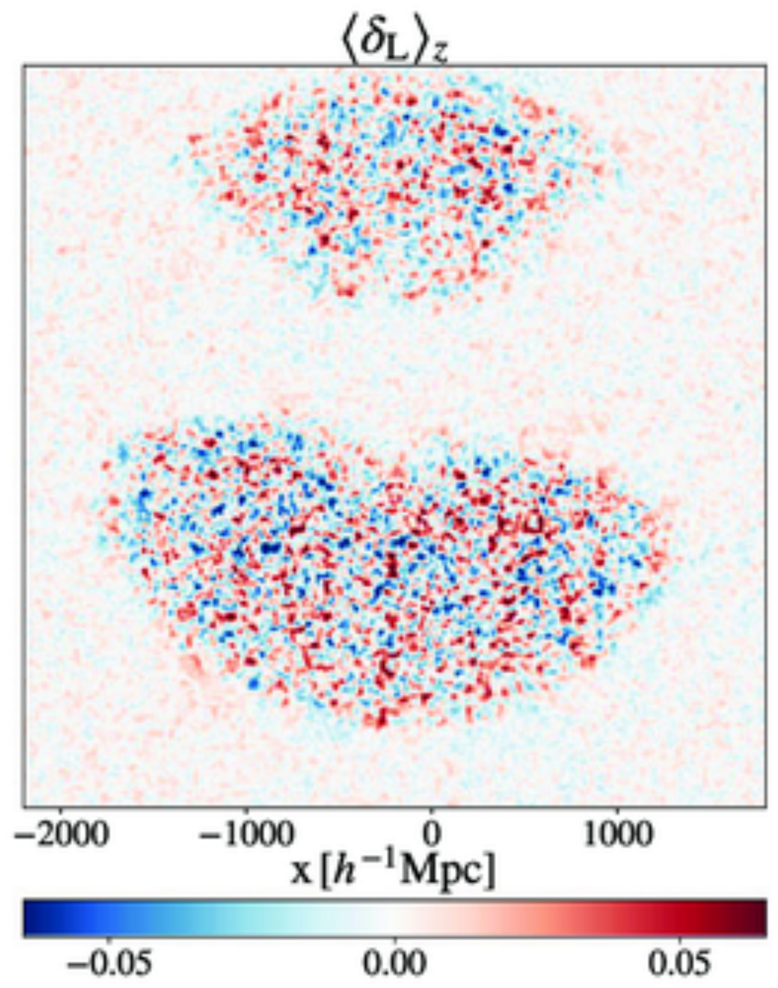


Complex structures (filaments / voids)

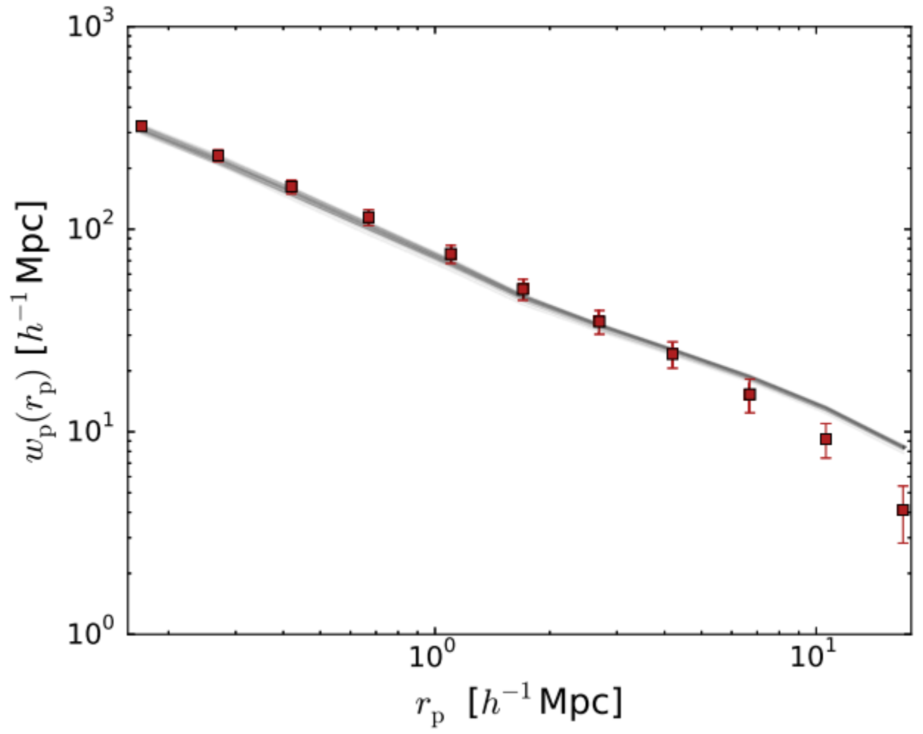


Number counts

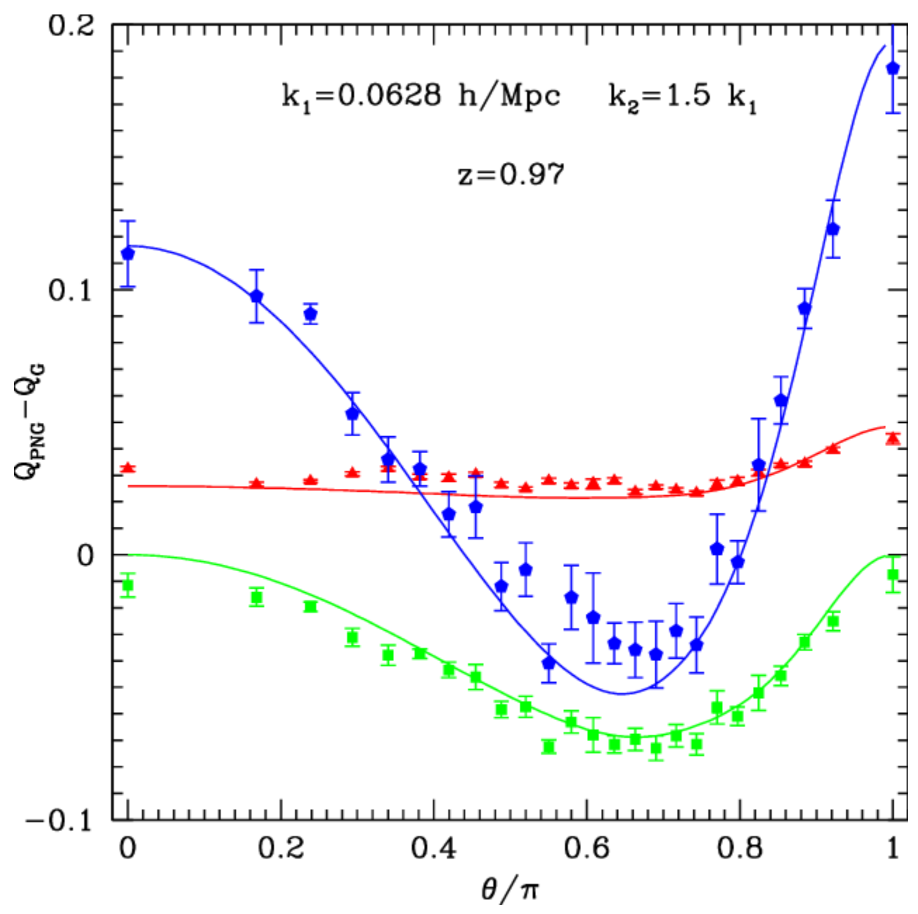




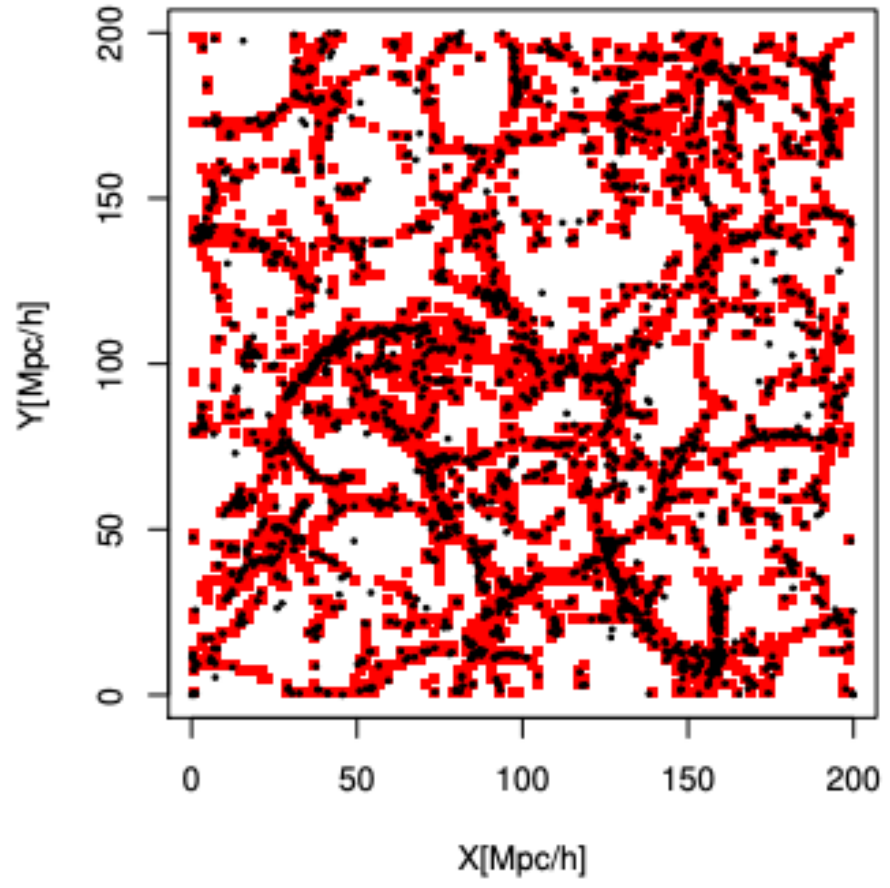
Full-field



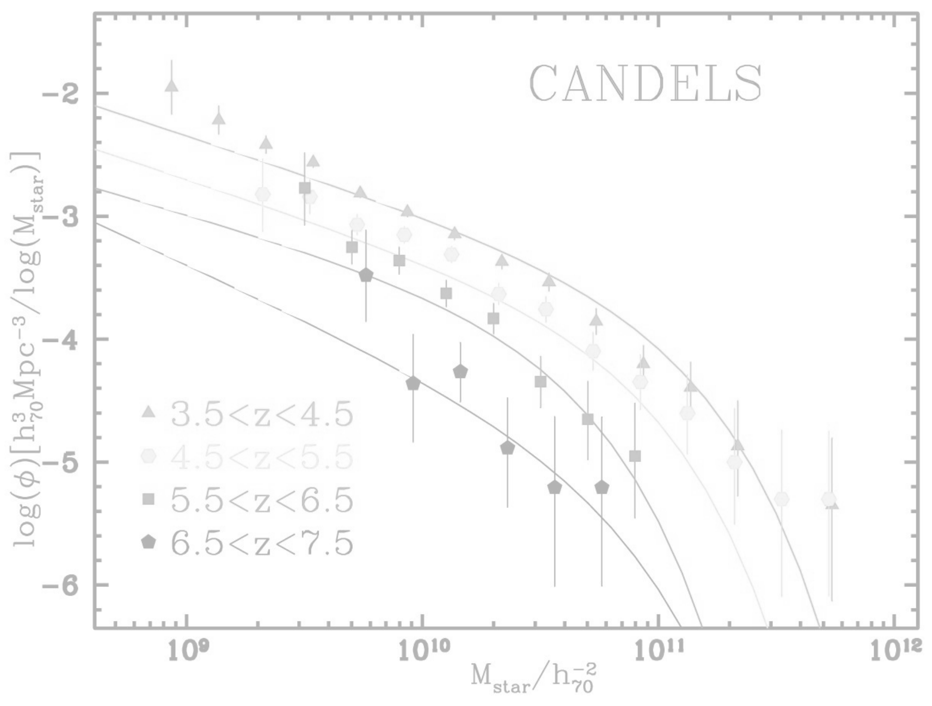
Two-point correlation

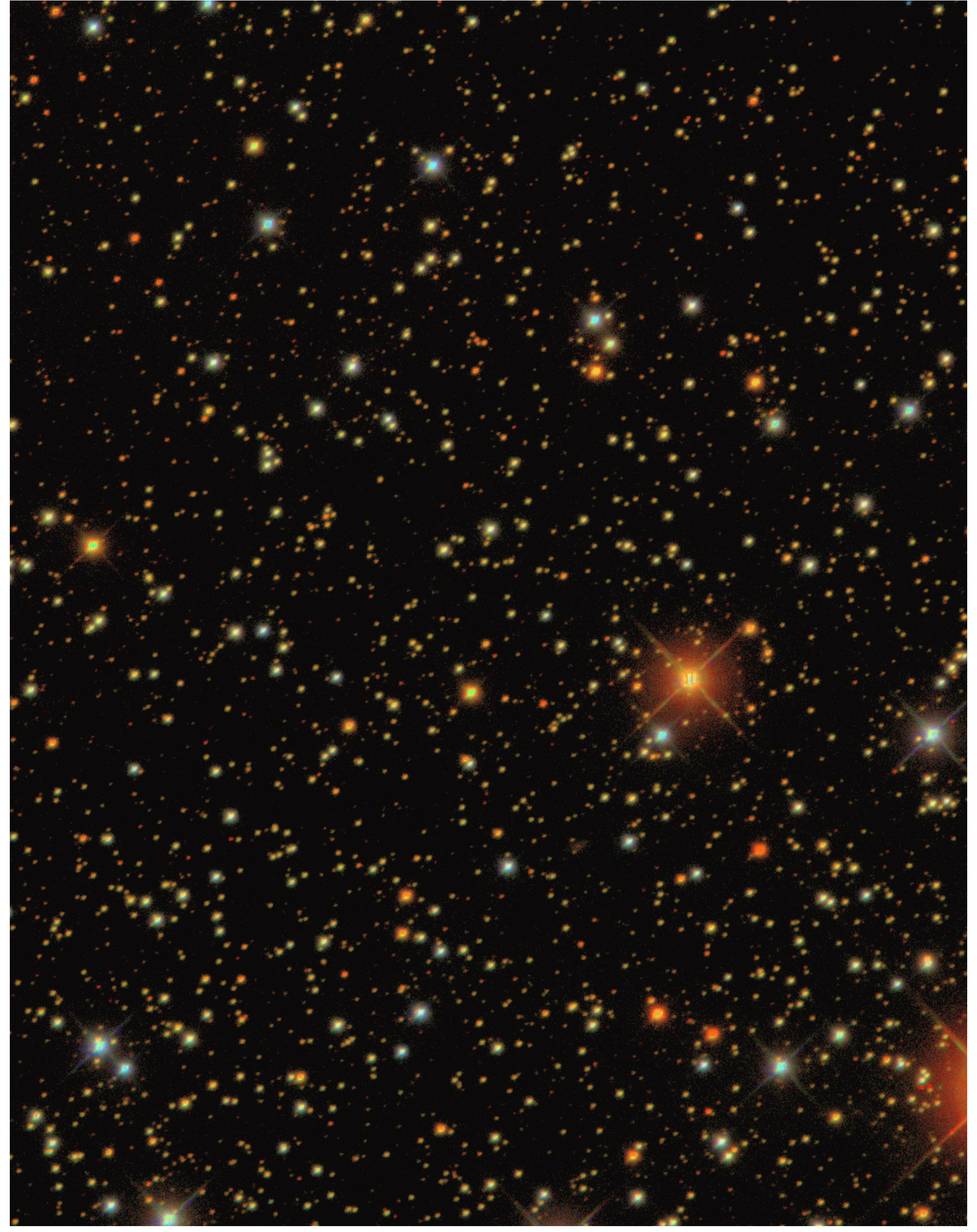


Three-point correlation



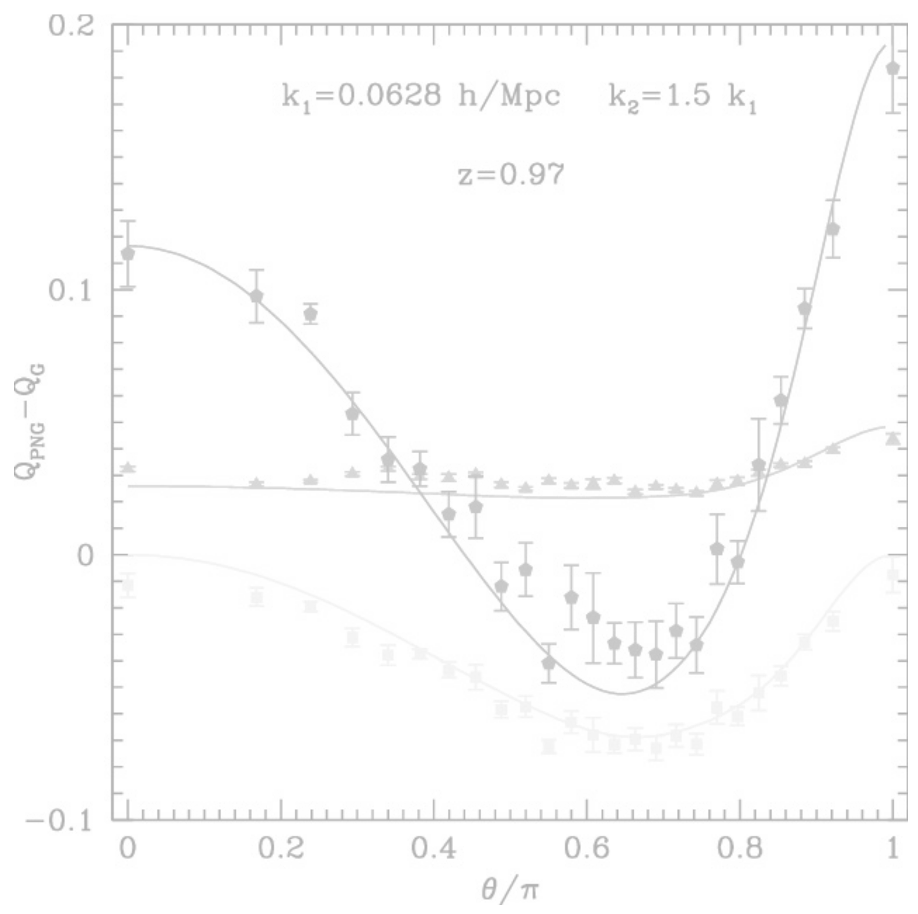
Complex structures (filaments / voids)

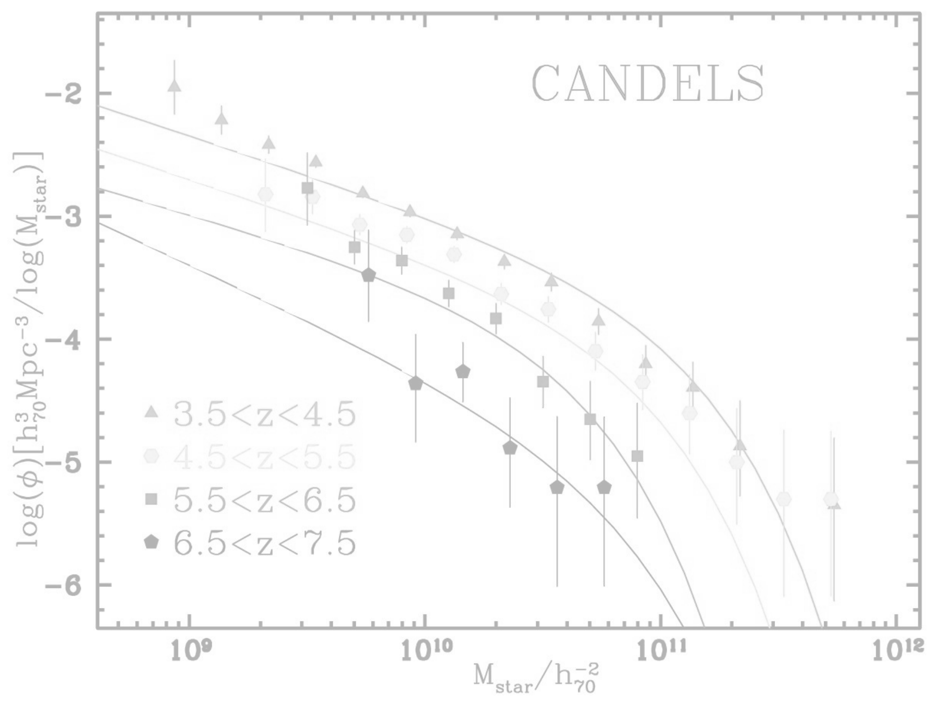


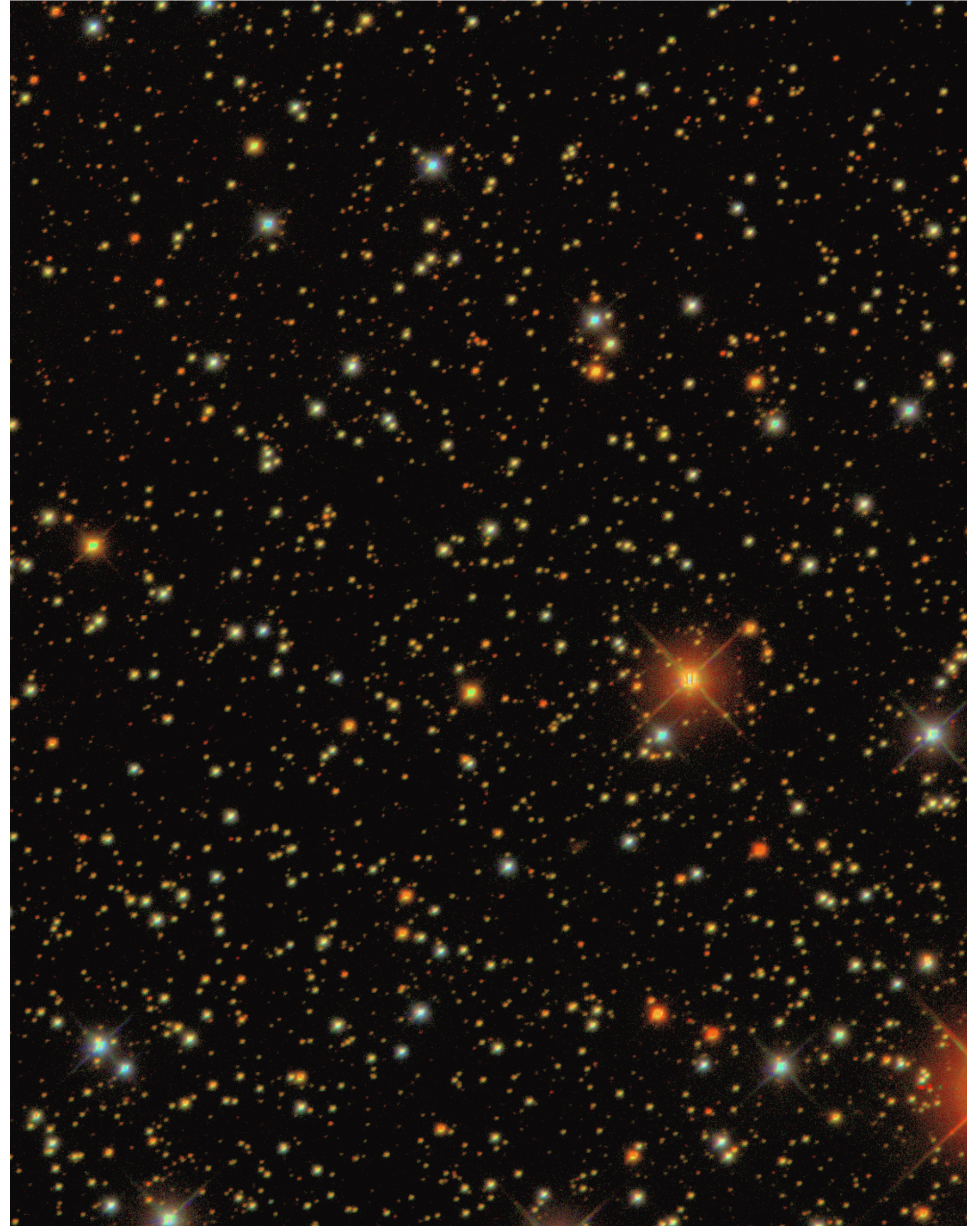


- Information rich
- Complete description
- Complex measurement and modeling
- Require lots of data to constrain lots of model parameters

- Simple measurement and modeling
- Require less data and more robust to nuisance parameters
- Less information
- Can only test limited range of models



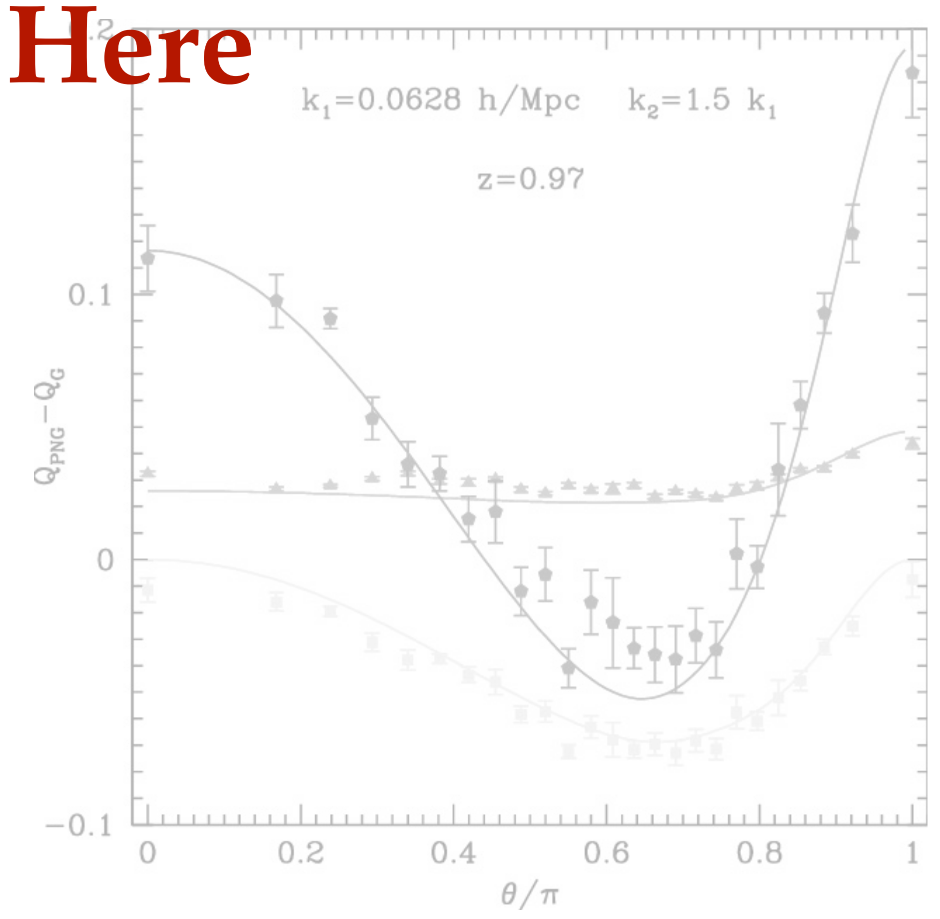


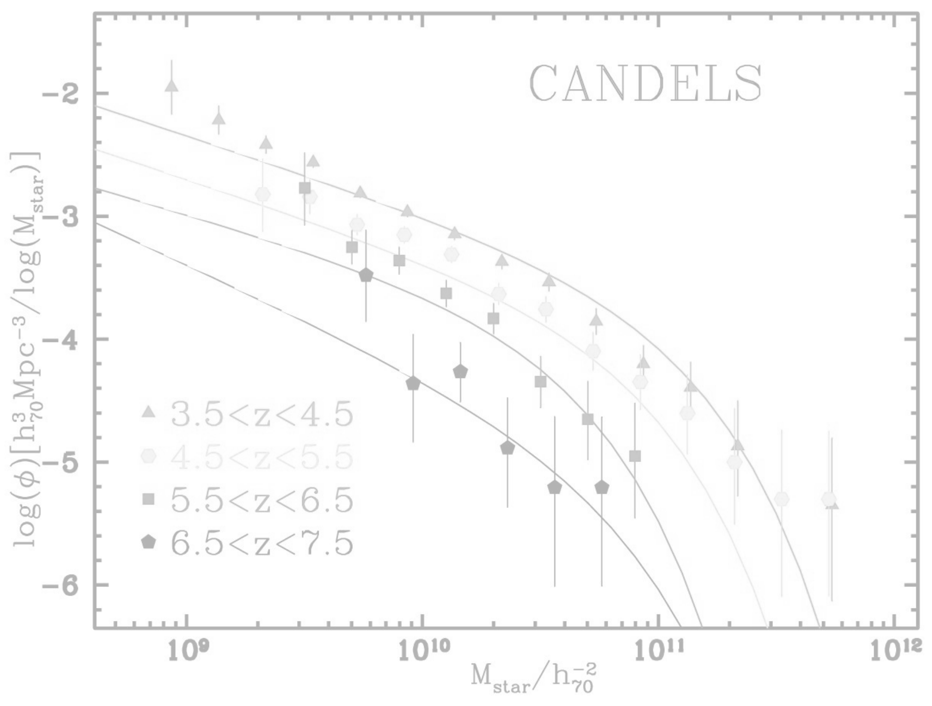


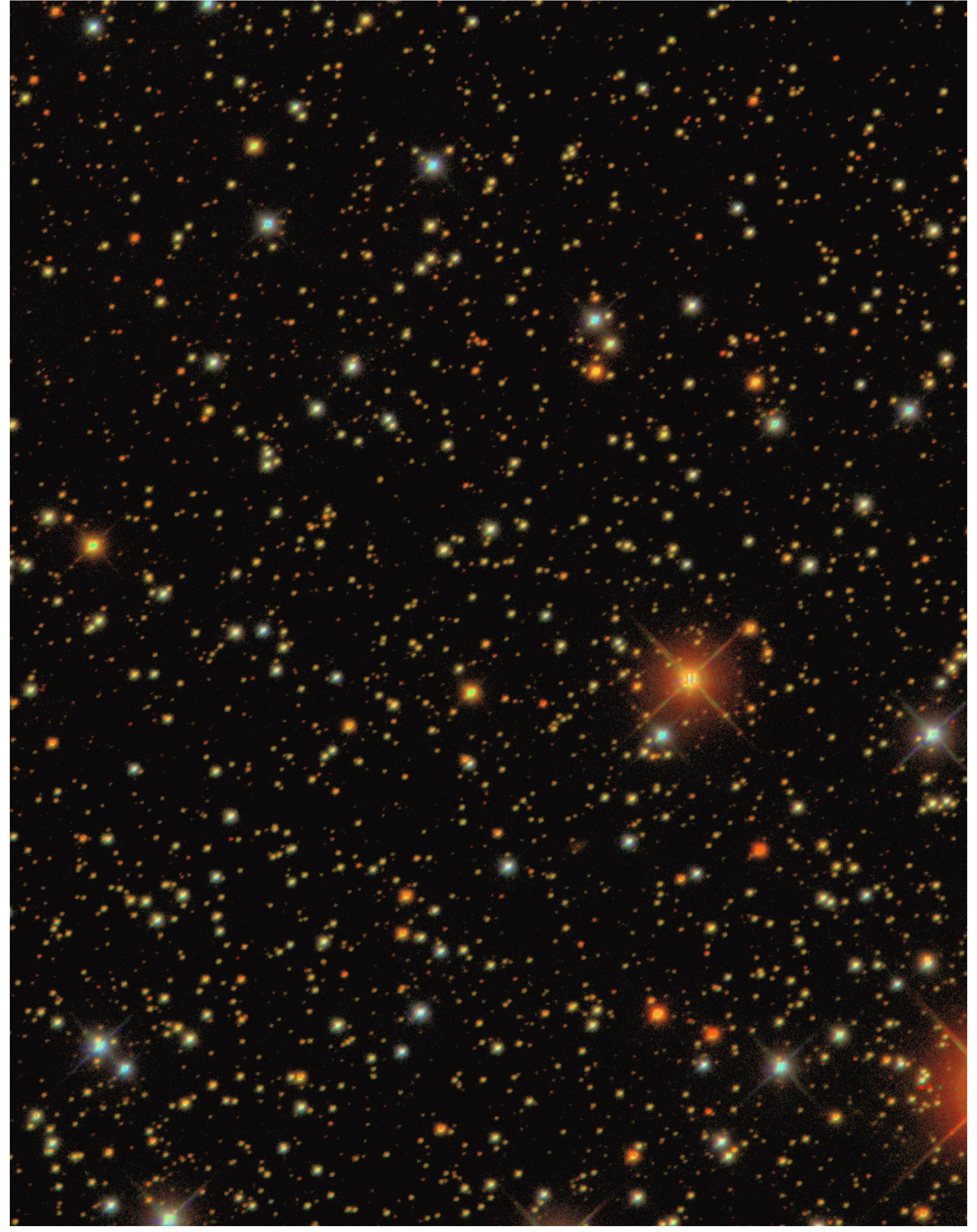
- Information rich
- Complete description
- Complex measurement and modeling
- Require lots of data to constrain lots of model parameters

- Simple measurement and modeling
- Require less data and more robust to nuisance parameters
- Less information
- Can only test limited range of models

We Are Here





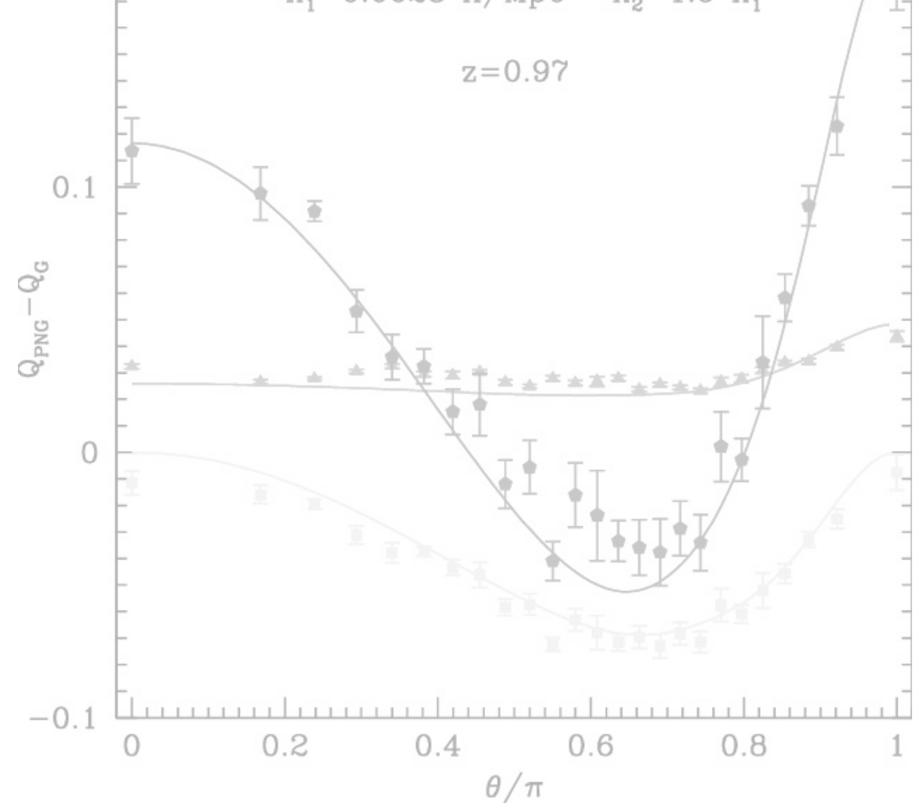


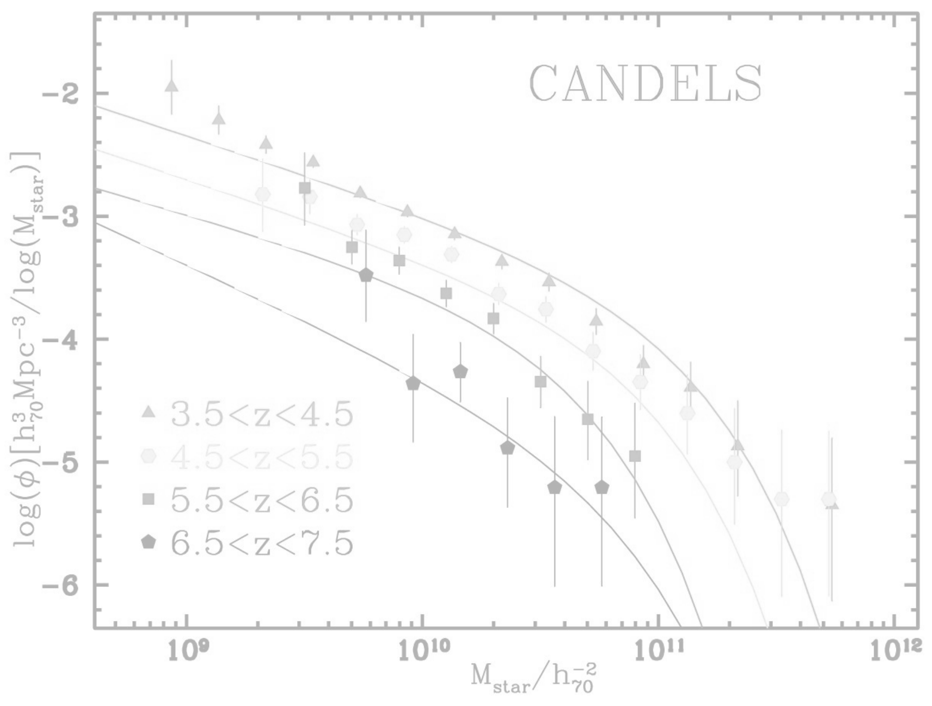
- Information rich
- Complete description
- Complex measurement and modeling
- Require lots of data to constrain lots of model parameters

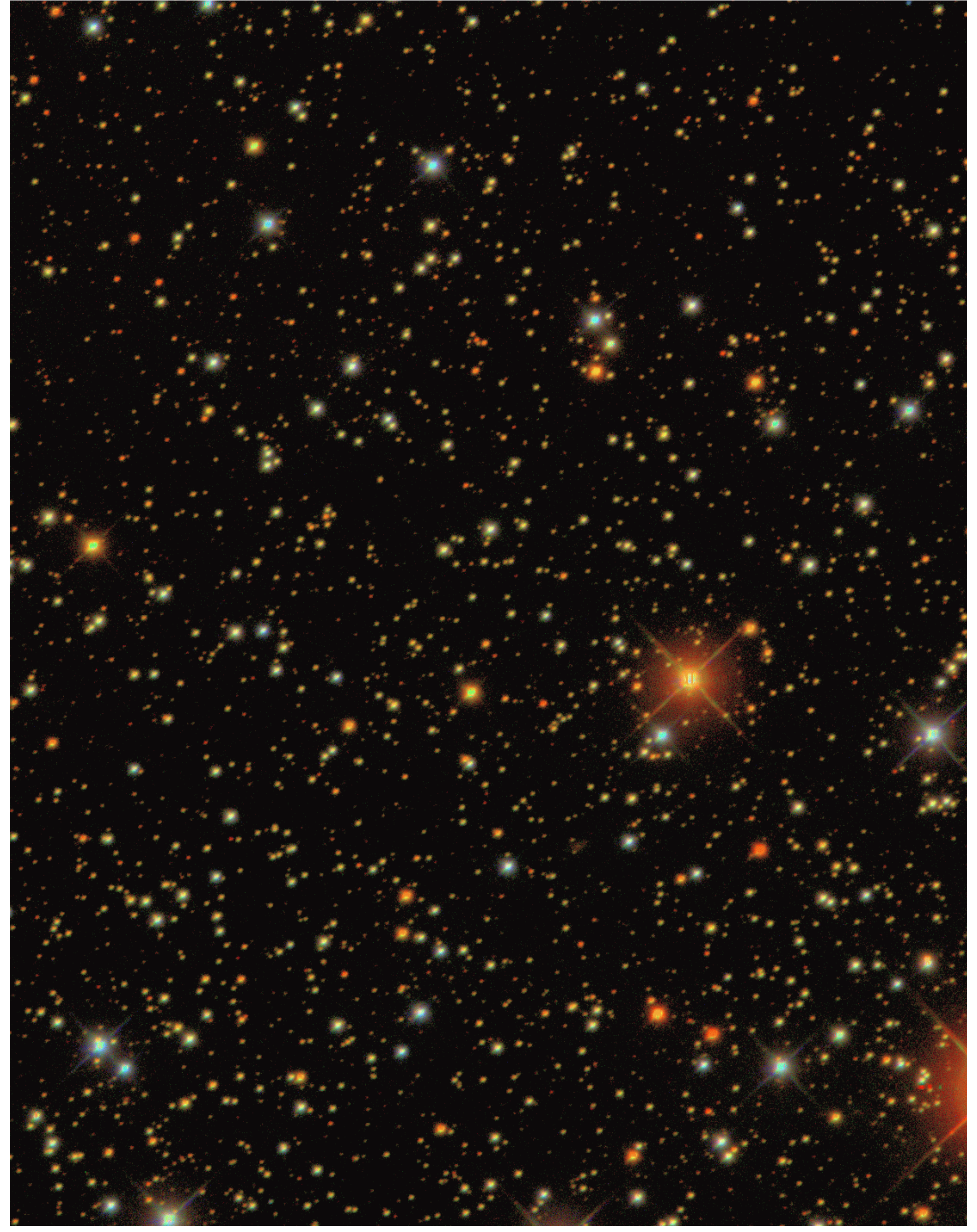
10 years ago

We Are Here

- Simple measurement and modeling
- Require less data and more robust to nuisance parameters
- Less information
- Can only test limited range of models







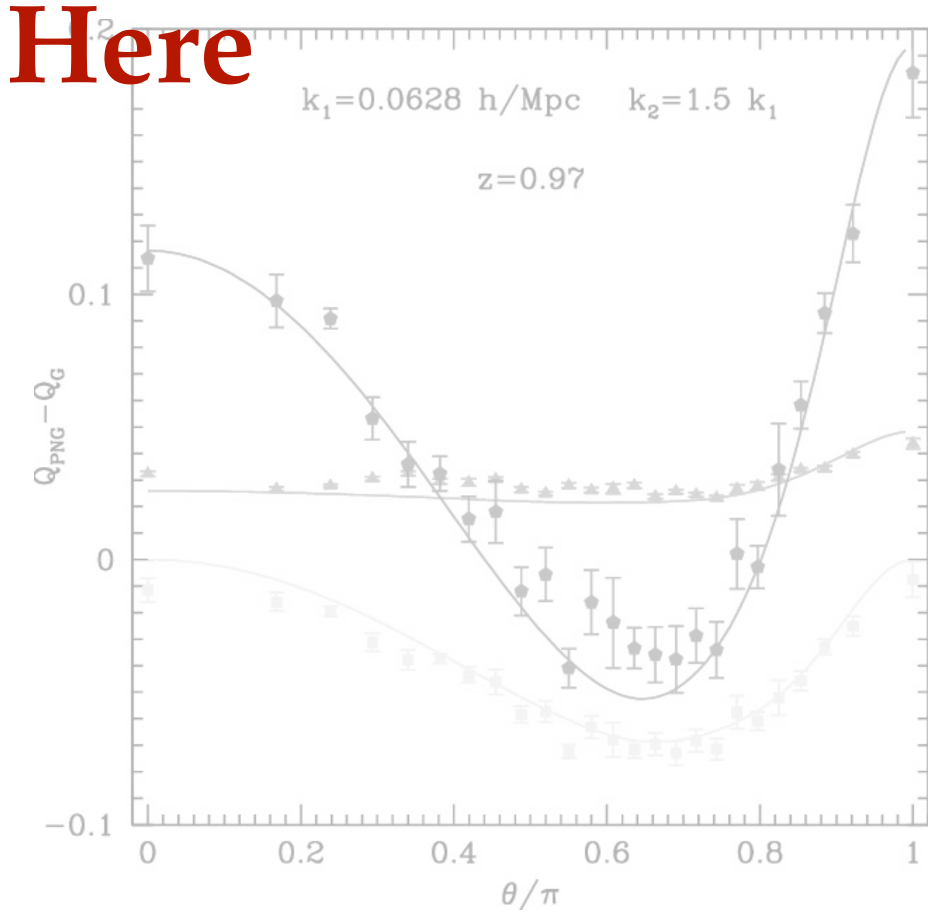
- Information rich
- Complete description
- Complex measurement and modeling
- Require lots of data to constrain lots of model parameters

10 years ago

10 years from now?

We Are Here

- Simple measurement and modeling
- Require less data and more robust to nuisance parameters
- Less information
- Can only test limited range of models



Lensing Higher-Order Statistics (HOS)

N-point correlations and related

N-point correlation function / N-spectrum

Mass aperture statistics and their moments

Moments / Cumulative Distribution Functions

Integrated N-point correlation function

Density-split statistics

Field-level inference

Bayesian hierarchical modelling

Deep learnings, e.g. CNN

Summary statistics

Peaks / voids

Topological statistics (Minkowski, Betti #, persistent homology)

Wavelets / scattering transform

Petri et al. (2015), Gruen et al. (2018), Allys et al. (2020), Halder et al. (2021), Zücher et al. (2021), Banerjee & Abel (2021), Fluri et al. (2022), Lanzieri et al. (2023)…

Lensing Higher-Order Statistics (HOS)

See also talk (and references) from

Jia
Bhuv
Joachim
Daniela
Niall
Judit
Lucas
Jason
Eleni
Arthur
Supranta
Adrian
Sihao
Luisa
…

Summary statistics

N-point correlations and related

N-point correlation function / N-spectrum

Mass aperture statistics and their moments

Moments / Cumulative Distribution Functions

Integrated N-point correlation function

Density-split statistics

Peaks / voids

Topological statistics (Minkowski, Betti #, persistent homology)

Wavelets / scattering transform

Field-level inference

Bayesian hierarchical modelling

Deep learnings, e.g. CNN

Petri et al. (2015), Gruen et al. (2018), Allys et al. (2020), Halder et al. (2021), Zücher et al. (2021), Banerjee & Abel (2021), Fluri et al. (2022), Lanzieri et al. (2023)…

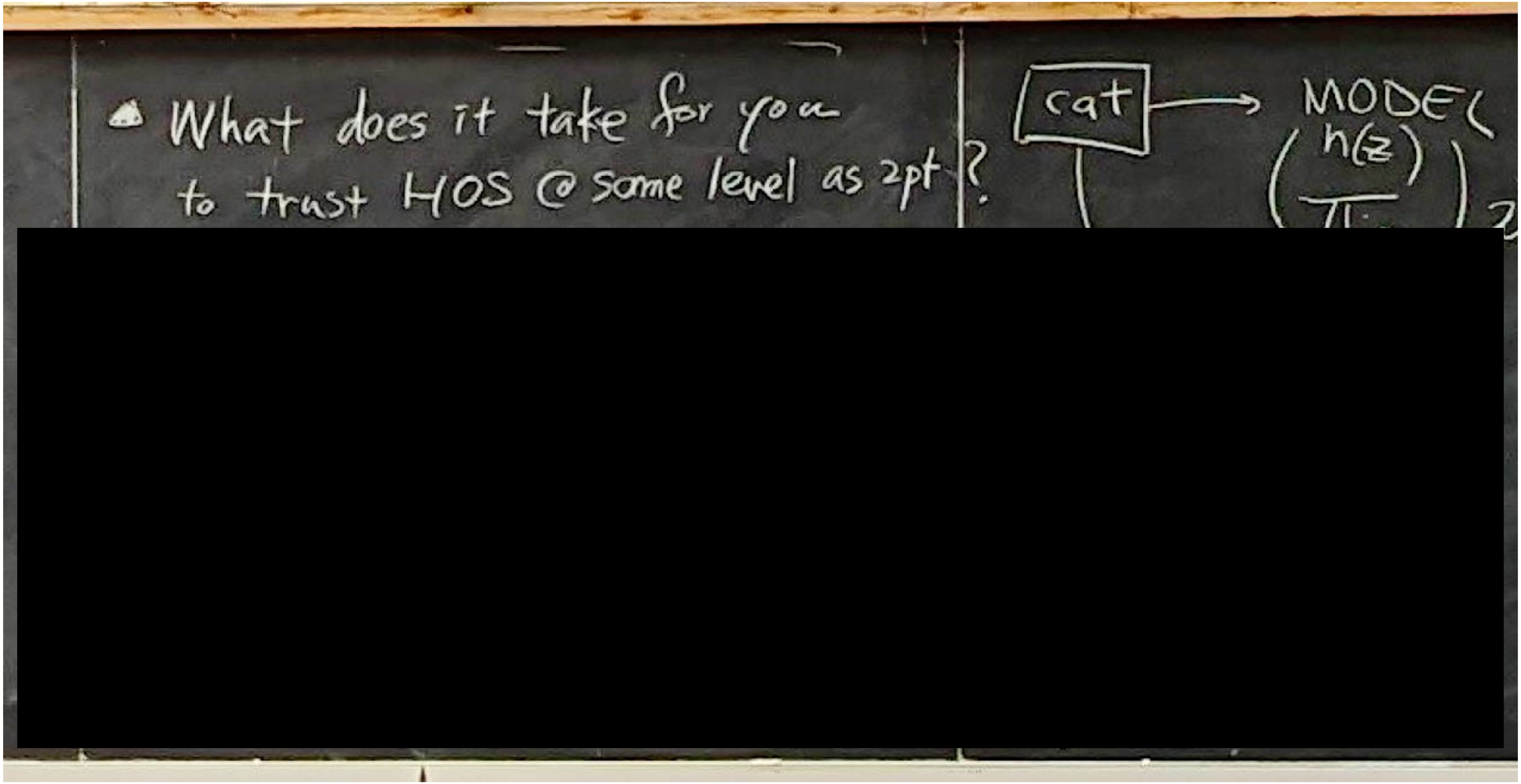
Outline

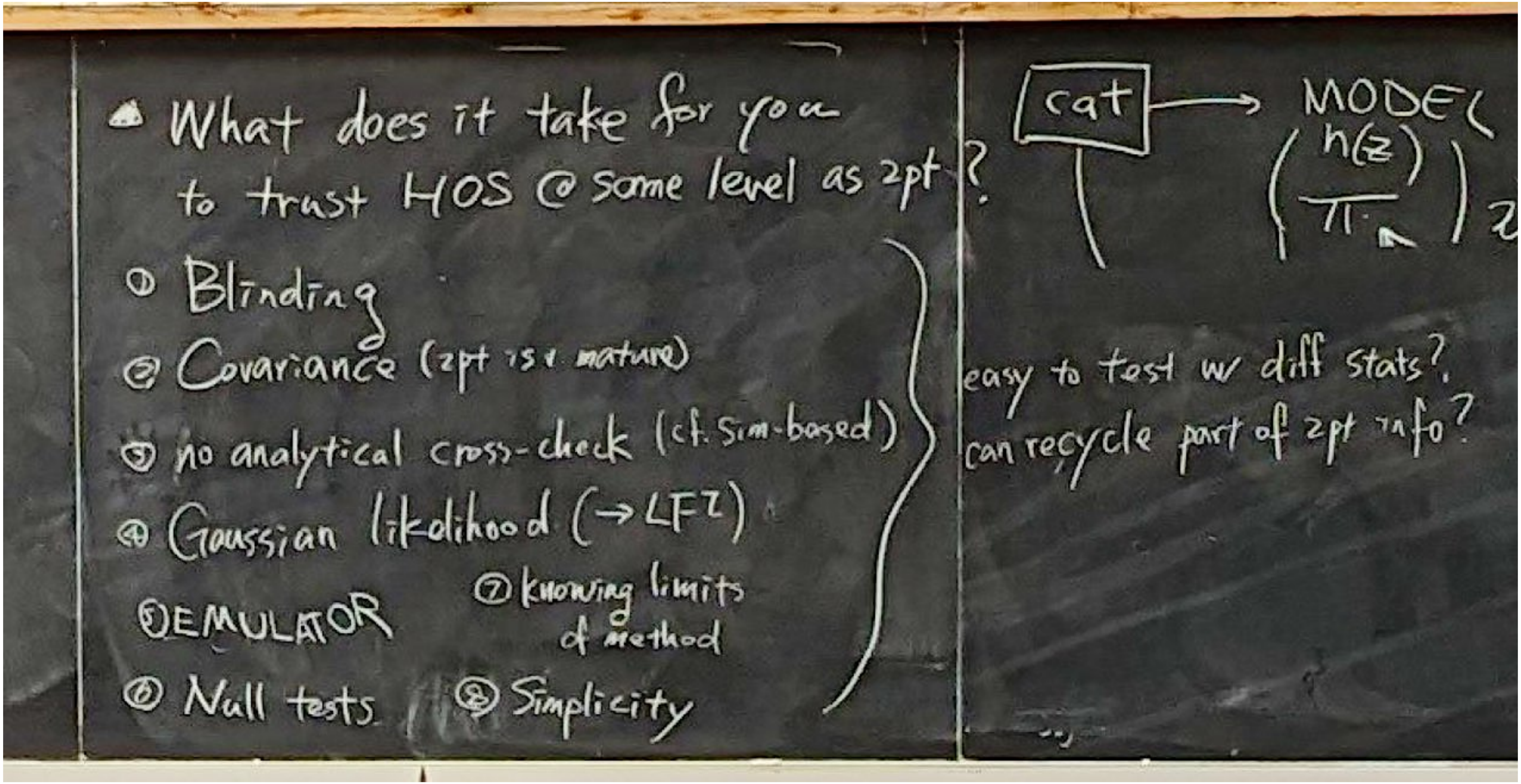
- The ΛCDM paradigm and extracting information beyond 2pt
- **Practical challenges: beyond 2pt systematics**
- Opportunities: primordial non-Gaussianity
- Towards field-level inference
- Summary & outlook

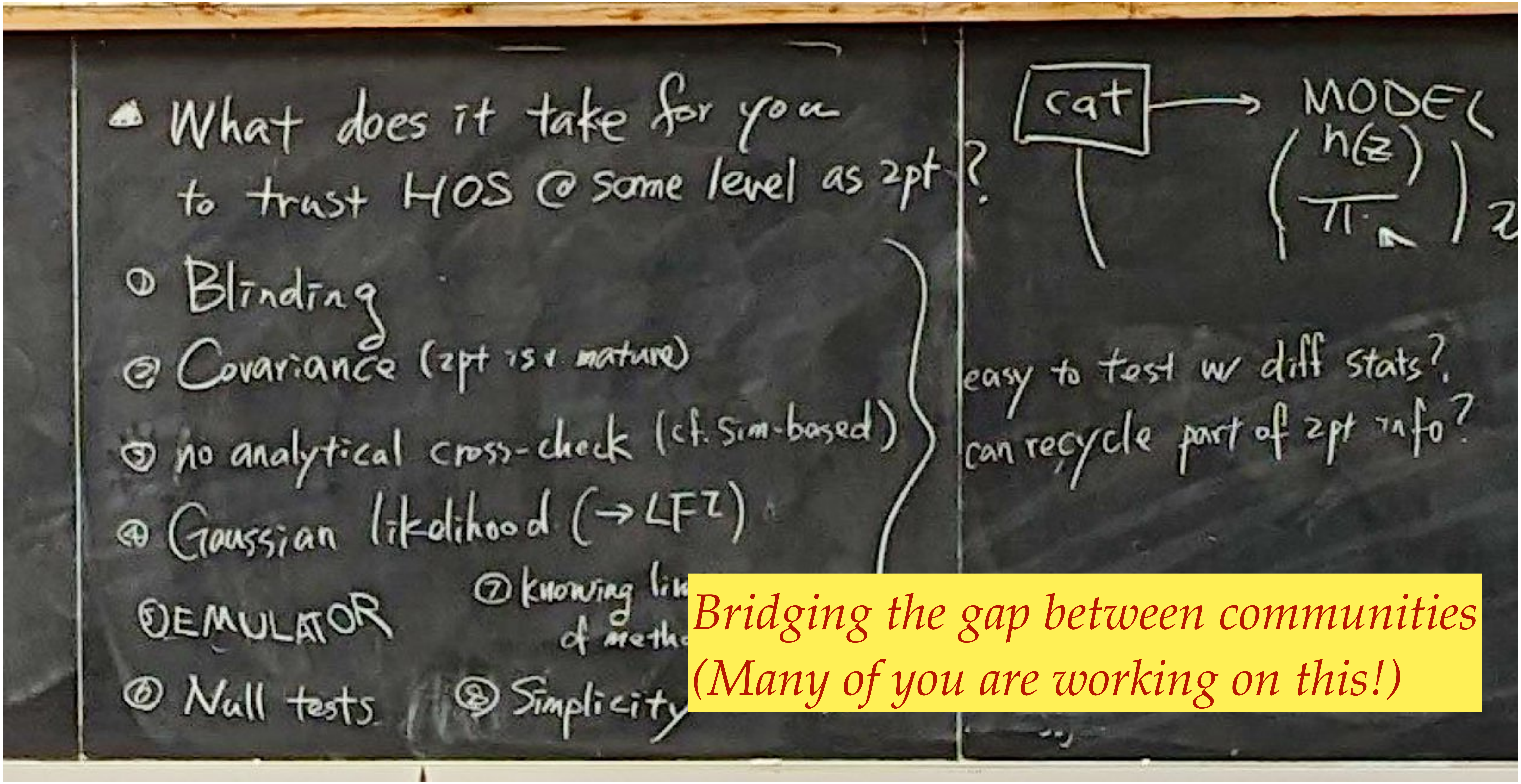
Anbajagane, CC, Banerjee et al. (2024)

Dhayaa Anbajagane

Arka Banerjee

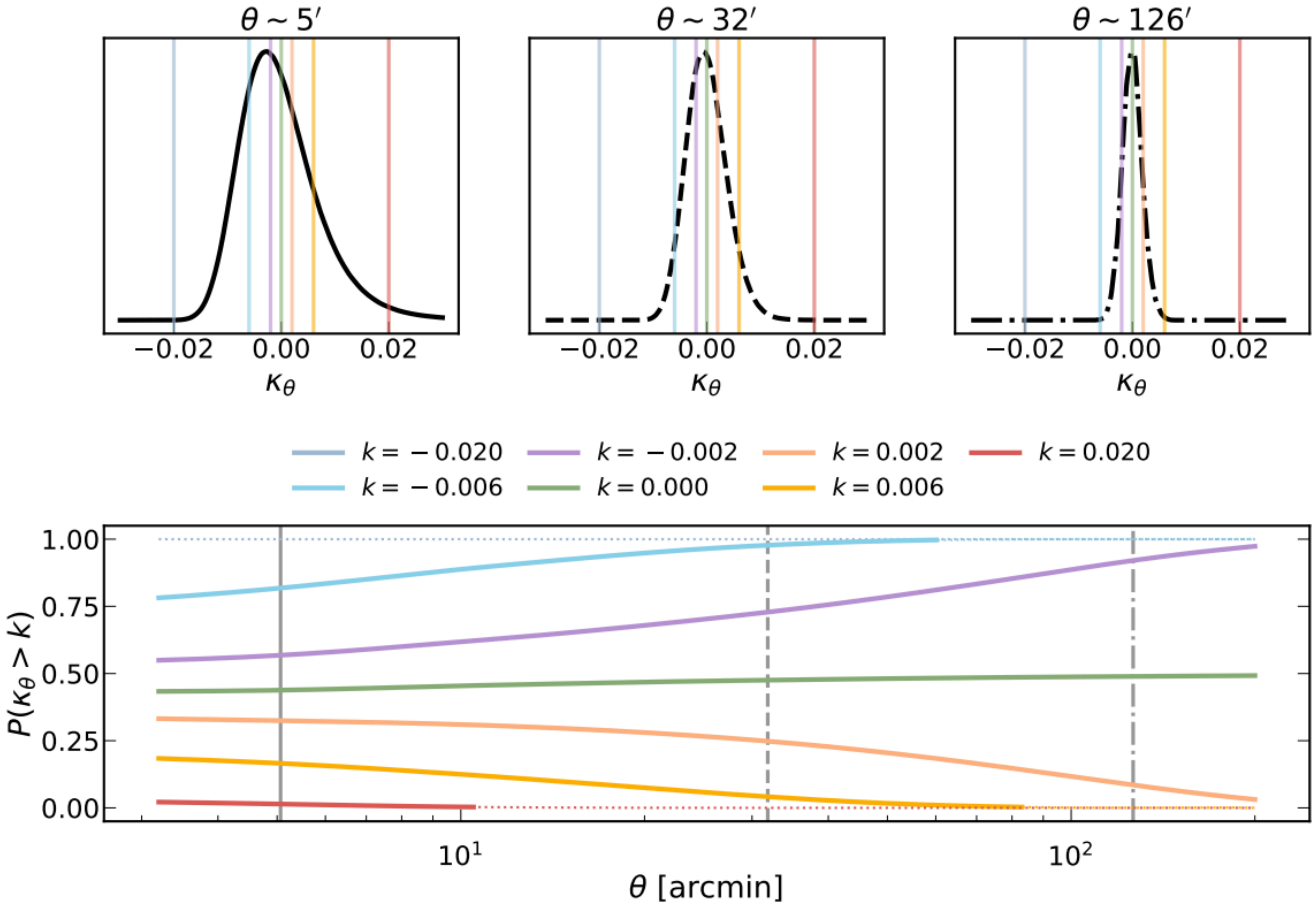

What does it take for you
to trust HOS @ some level as 2pt?
cat
MODEL
h(z)
[redacted]


What does it take for you
to trust HOS @ same level as 2pt ?
① Blinding
② Covariance (2pt is + mature)
③ no analytical cross-check (cf. sim-based)
④ Gaussian likelihood (→ LFI)
⑤ EMULATOR
⑦ knowing limits
of method
⑥ Null tests
⑧ Simplicity
cat
MODEL
n(z)
π
easy to test w/ diff stats?
can recycle part of 2pt info?



Bridging the gap between communities
(Many of you are working on this!)

Towards robust HOS: scale-dependent CDFs



The first question to ask

- How much information is there in the field? And how much can we practically extract considering both systematic effects and computational limitations?

The first question to ask

- How much information is there in the field? And how much can we practically extract considering both systematic effects and computational limitations?

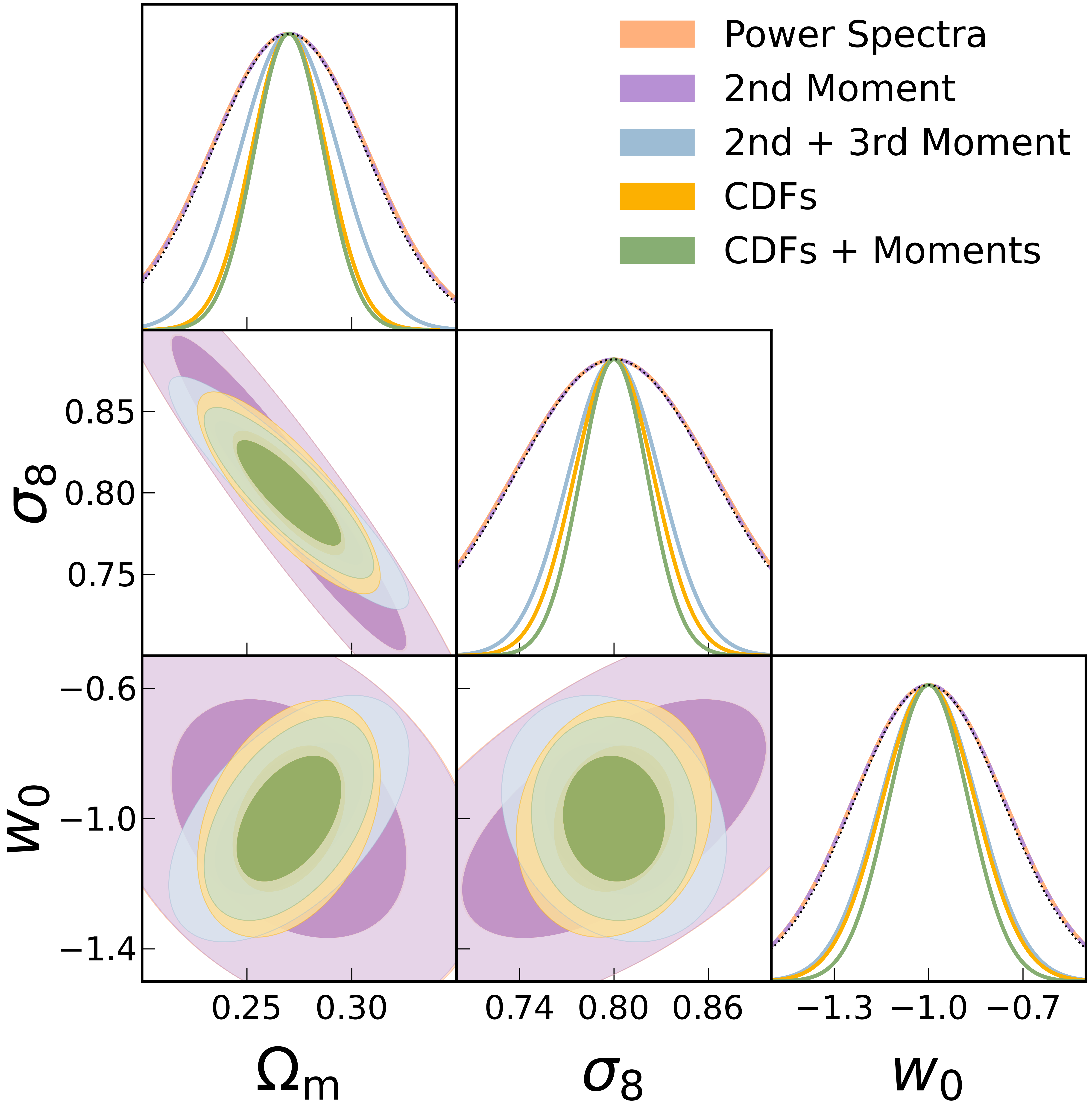
- The answer to these questions depends on
  - The science question of interest (w/LCDM, EDE, fNL, non-CDM)
  - The dataset (density/galaxy, redshift, noise, systematics)
  - The implementation (statistics, inference, sampling)

A simple exploration

- Assuming
  - wCDM
  - DES Y3-like lensing
  - Fixed scale
  - Moments and CDFs*

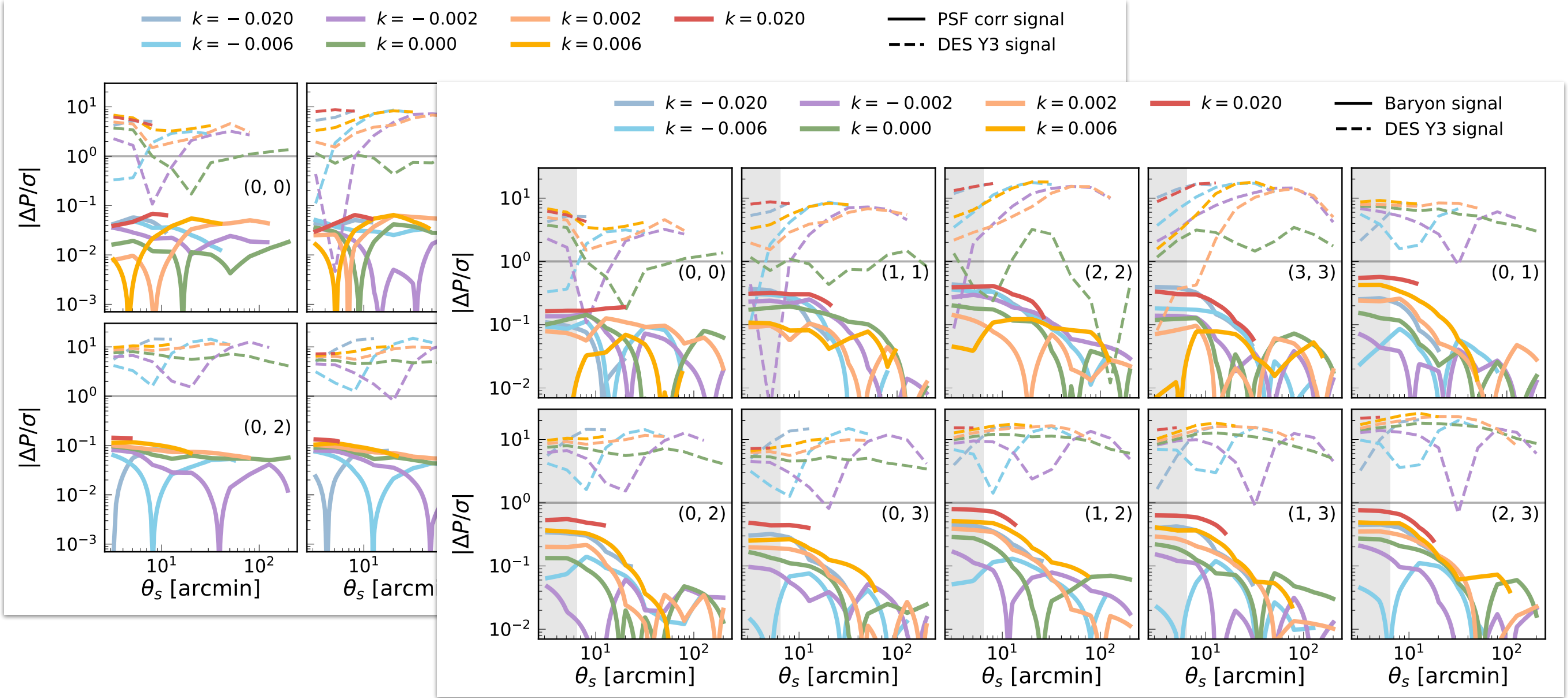
- A simple analysis says that we gain significantly (3x in area) going from 2pt to 3pt, and much less after that

Recall Bhuv and Sihao's talk



* no directional information

Next, redo all the 2pt tests…

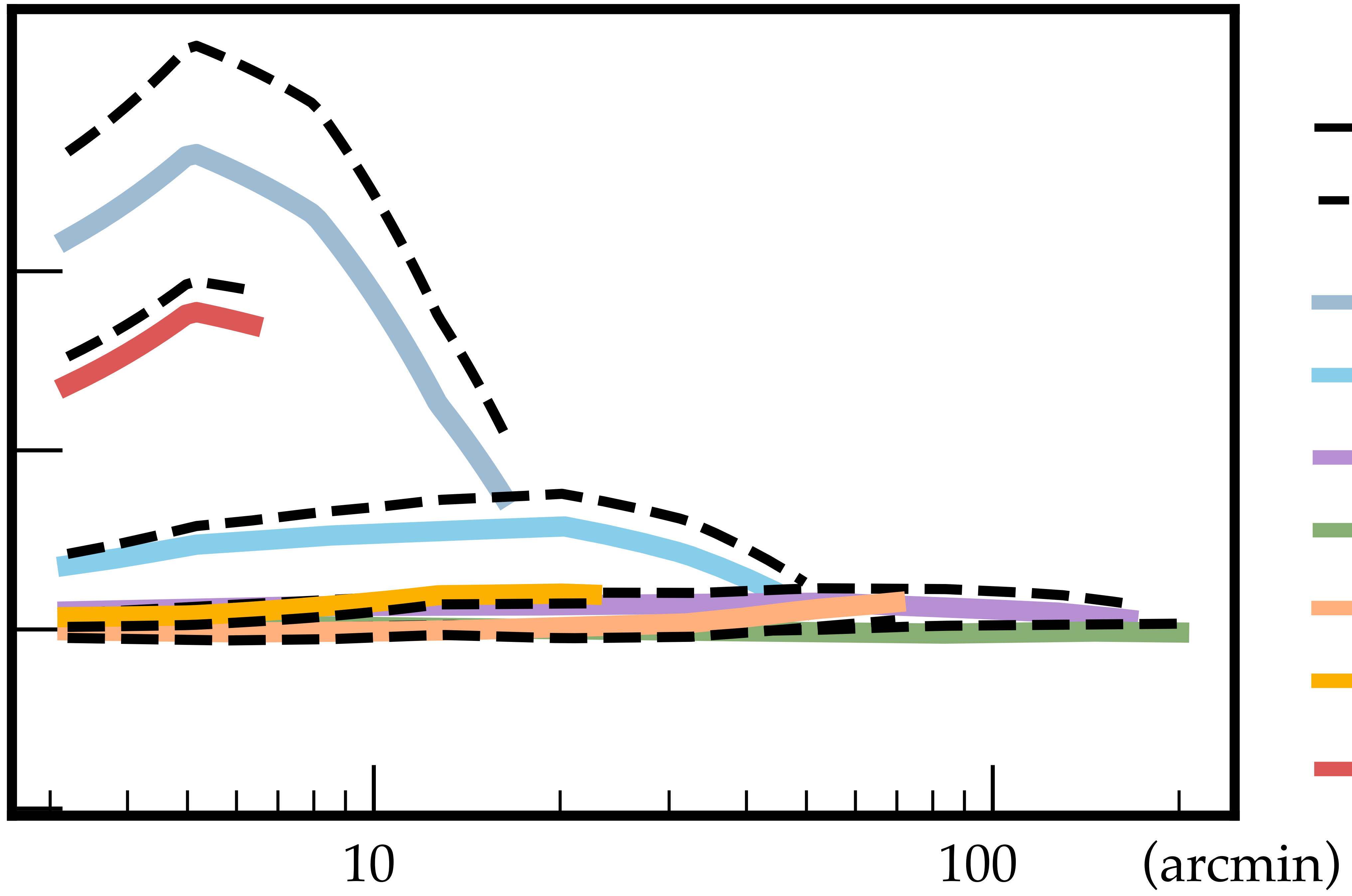


But that may not be sufficient!

- New systematic effects could appear at higher order

Correlation of signal and noise at all orders

Gatti, Jeffrey et al. (2023)



Robust HOS

- For data in the near future and wCDM/LCDM, having a 3pt-level pipeline could bring us a long way way to extracting all the accessible non-Gaussian information

- Ensuring that the analysis is robust to all sources of systematic effects (at least at the 3pt level) is crucial for having HOS be trusted to the same level as 2pt

- It is encouraging that people are working on these (sometimes tedious) tasks and working together with people familiar with 2pt analyses

Outline

- The ΛCDM paradigm and extracting information beyond 2pt
- Practical challenges: beyond 2pt systematics
- **Opportunities: primordial non-Gaussianity**
- Towards field-level inference
- Summary & outlook

Anbajagane, CC, Lee, Gatti (2024)

Dhayaa Anbajagane

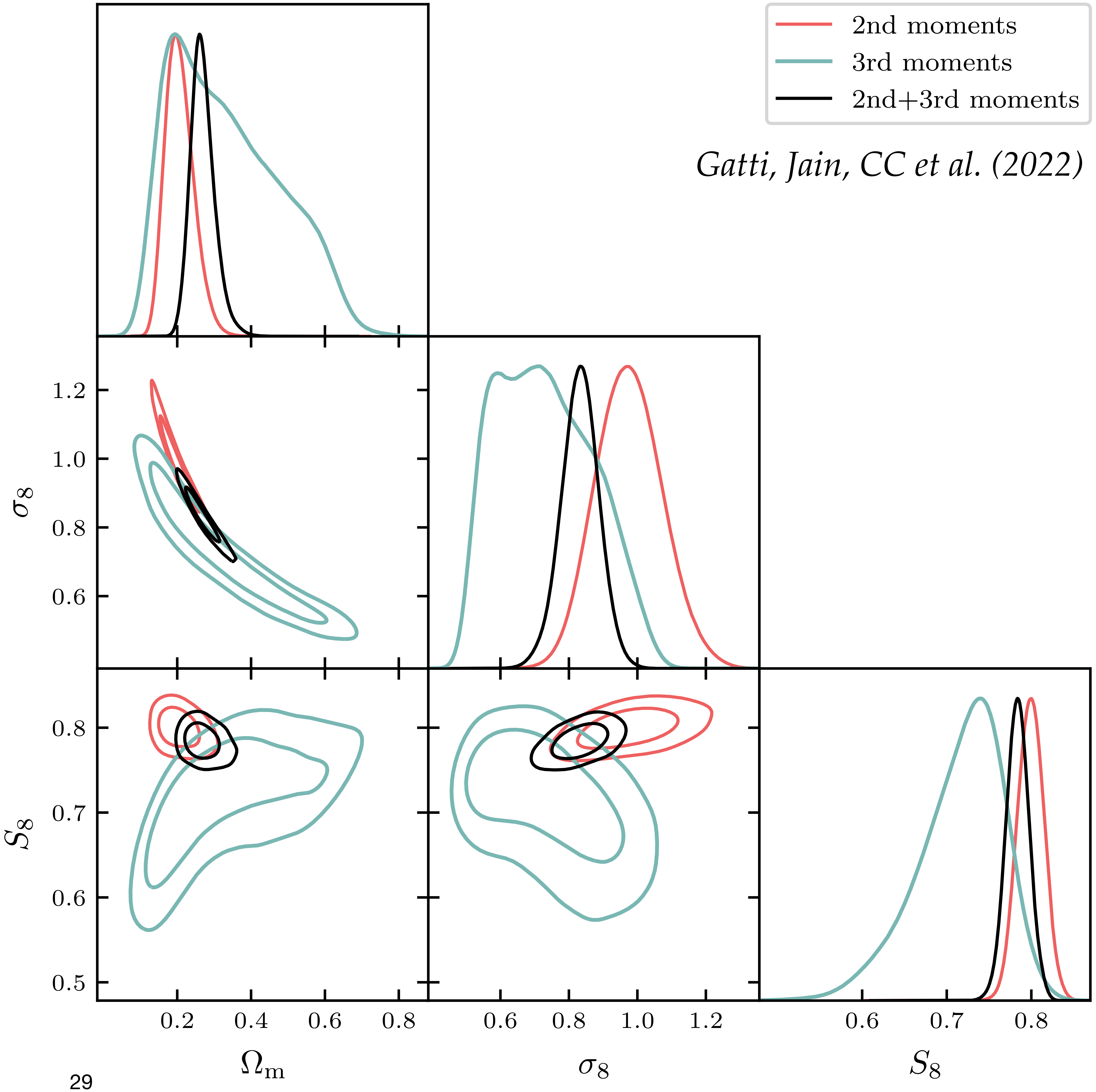
Hayden Lee

Marco Gatti

wCDM/LCMD

- Most literature today use HOS to extract information beyond 2pt in the field to constrain wCDM/LCDM
- This makes sense since gravity is highly non-Gaussian, there's a lot of information we are leaving behind when we only do 2pt

See talks from Bhuv, Joachim, Daniela, Niall, Judit, Supranta…



Gatti, Jain, CC et al. (2022)

What else can we do with this information?

- We can try to test different models of the early universe (e.g. primordial non-Gaussianity, or PNG)

What else can we do with this information?

- However, gravity is highly non-Gaussian, the low-redshift observables contain both primordial and late-time non-Gaussianity (cf. CMB)

What else can we do with this information?

- But, perhaps not all hope is lost, we do know fairly well how to forward-model gravity. Can we disentangle the primordial and late-time non-Gaussianity via simulations?

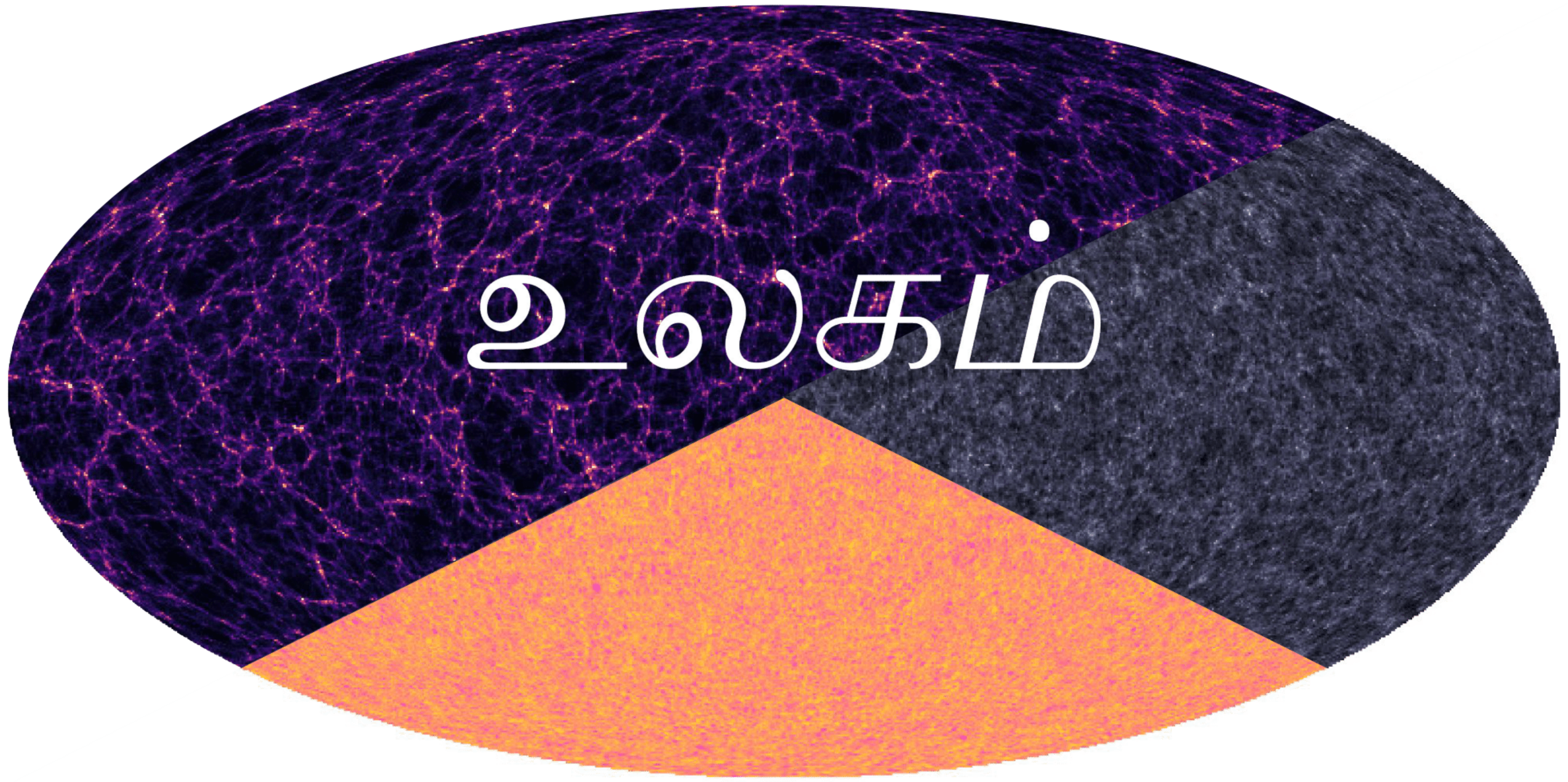
What else can we do with this information?

- But, perhaps not all hope is lost, we do know fairly well how to forward-model gravity. Can we disentangle the primordial and late-time non-Gaussianity via simulations?

Inject bispectrum templates $(f_{\rm NL}^{\rm eq},\ f_{\rm NL}^{\rm or,lss},\ f_{\rm NL}^{\rm loc},\ f_{\rm NL}^{\rm or,cmb})$

“The World”

The Ulagam Simulations



https://ulagam-simulations.readthedocs.io/

- N-body simulation suite designed for full-sky analyses of wide-field surveys for fNL
- Initial conditions from **Quijote** (Villaescusa-Navarro et al. 2020) and **Quijote-PNG** (Coulton et al. 2022)

$$N = 512^3$$

$$L = 1\ \mathrm{Gpc}/h$$

$$\{\Omega_\mathrm{m}, \sigma_8, n_s, w, f_\mathrm{NL}^\mathrm{X}\}$$

Accurate to k < 1 [h/Mpc], $\ell < 1000$

Gravitational potential

$$\Phi(\mathbf{k}) = \phi(\mathbf{k}) + \int f_{\rm NL}[\delta_D] K(\mathbf{k}_1, \mathbf{k}_2) \phi(\mathbf{k}_1)\phi(\mathbf{k}_2) d^3k_1 d^3k_2$$

Gaussian

$$B_\Phi = 2 f_{\rm NL} K(\mathbf{k}_1, \mathbf{k}_2) P_{\Phi,1} P_{\Phi,2} + \text{cyc.}$$

$f_{\rm NL}^{\rm loc}$ Presence of second scalar field

$f_{\rm NL}^{\rm eq}$ Presence of “non-canonical” kinetic terms

$f_{\rm NL}^{\rm or,cmb}$ Approx. orthogonal to local and equilateral

$f_{\rm NL}^{\rm or,lss}$ Similar to above, but better approximation at squeezed

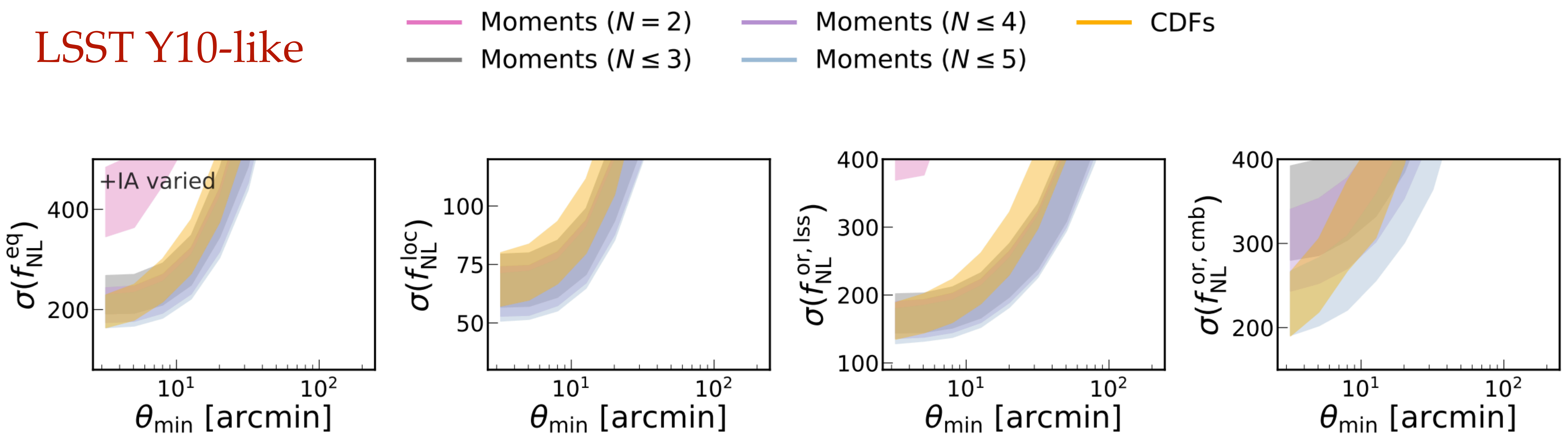
Analysis setup

- Weak lensing convergence maps
- Moments (N=2-5), CDFs
- Fisher forecast on constraints on fNL
- Simulation covariance at fiducial cosmology
- Compare DES Y3, DES Y6, LSST Y1, LSST Y10

Run	$\mathbf{P}_{\rm fid} \pm \Delta P$	$N_{\rm sim}$
Fiducial	—	**2000**
Local PNG, $f_{\rm NL}^{\rm loc}$	$\mathbf{0} \pm 100$	100
Equilateral PNG, $f_{\rm NL}^{\rm eq}$	$\mathbf{0} \pm 100$	100
LSS Orthogonal PNG, $f_{\rm NL}^{\rm or,\,lss}$	$\mathbf{0} \pm 100$	100
CMB Orthogonal PNG, $f_{\rm NL}^{\rm or,\,cmb}$	$\mathbf{0} \pm 100$	100
Matter density, $\Omega_{\rm m}$	$\mathbf{0.3175} \pm 0.01$	100
Density fluctuations amplitude. σ_8	$\mathbf{0.834} \pm 0.015$	100
Dark energy EoS w_0	$-\mathbf{1} \pm 0.05$	100
Spectral index n_s	$\mathbf{0.9624} \pm 0.02$	100

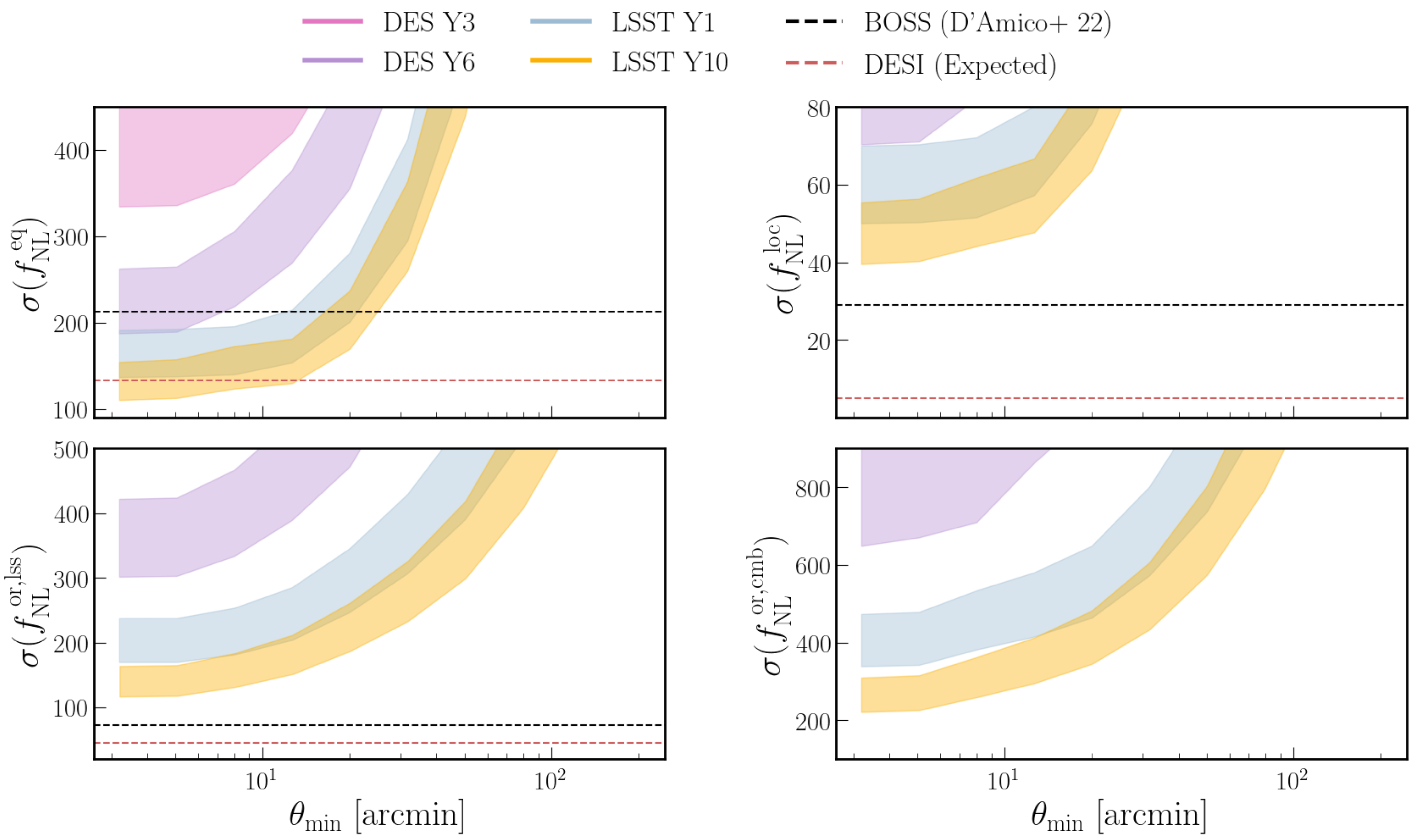
In which order is the information stored?

LSST Y10-like

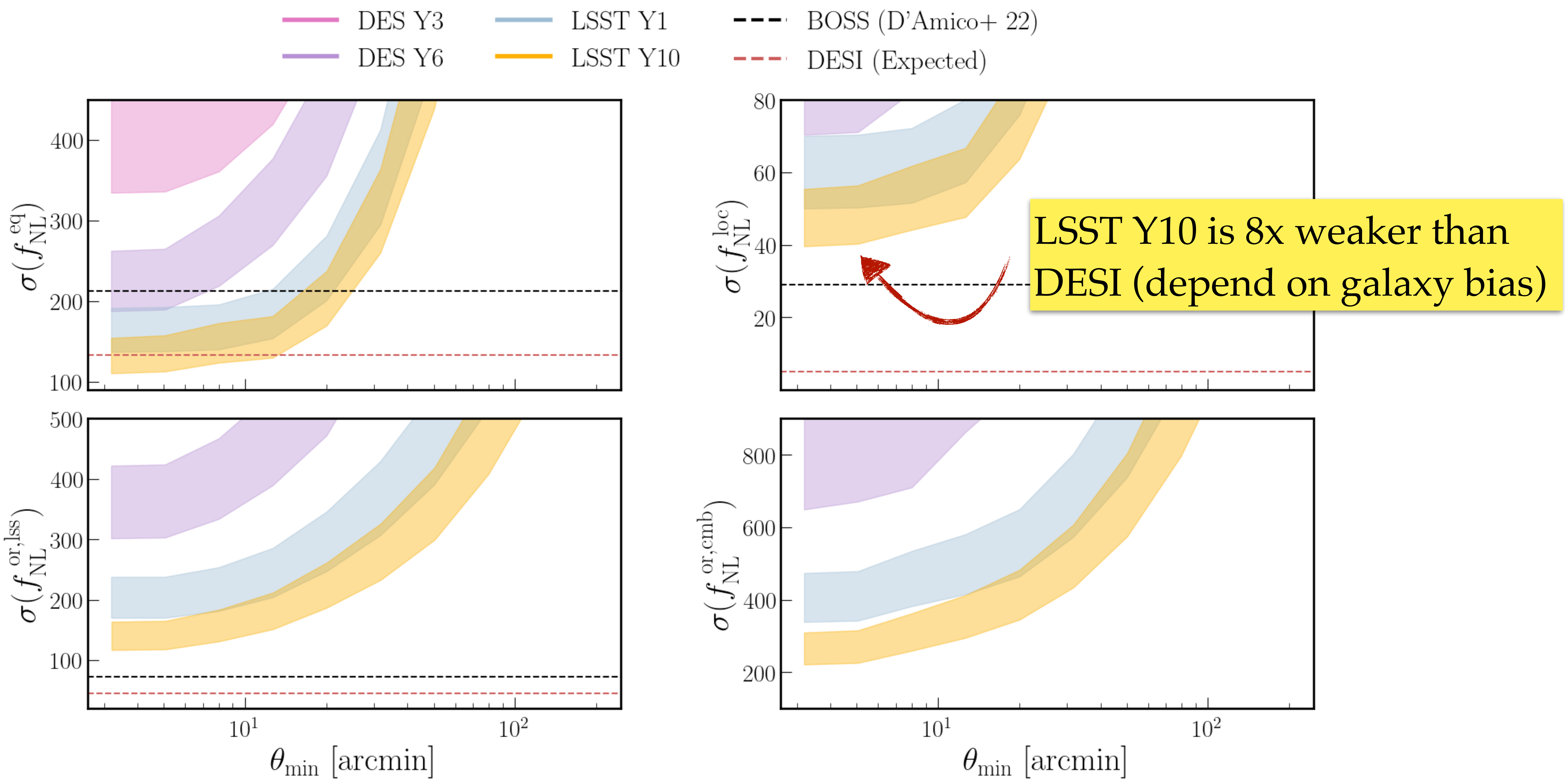


A reasonable scale to look at is ~10′ (5-15 Mpc)

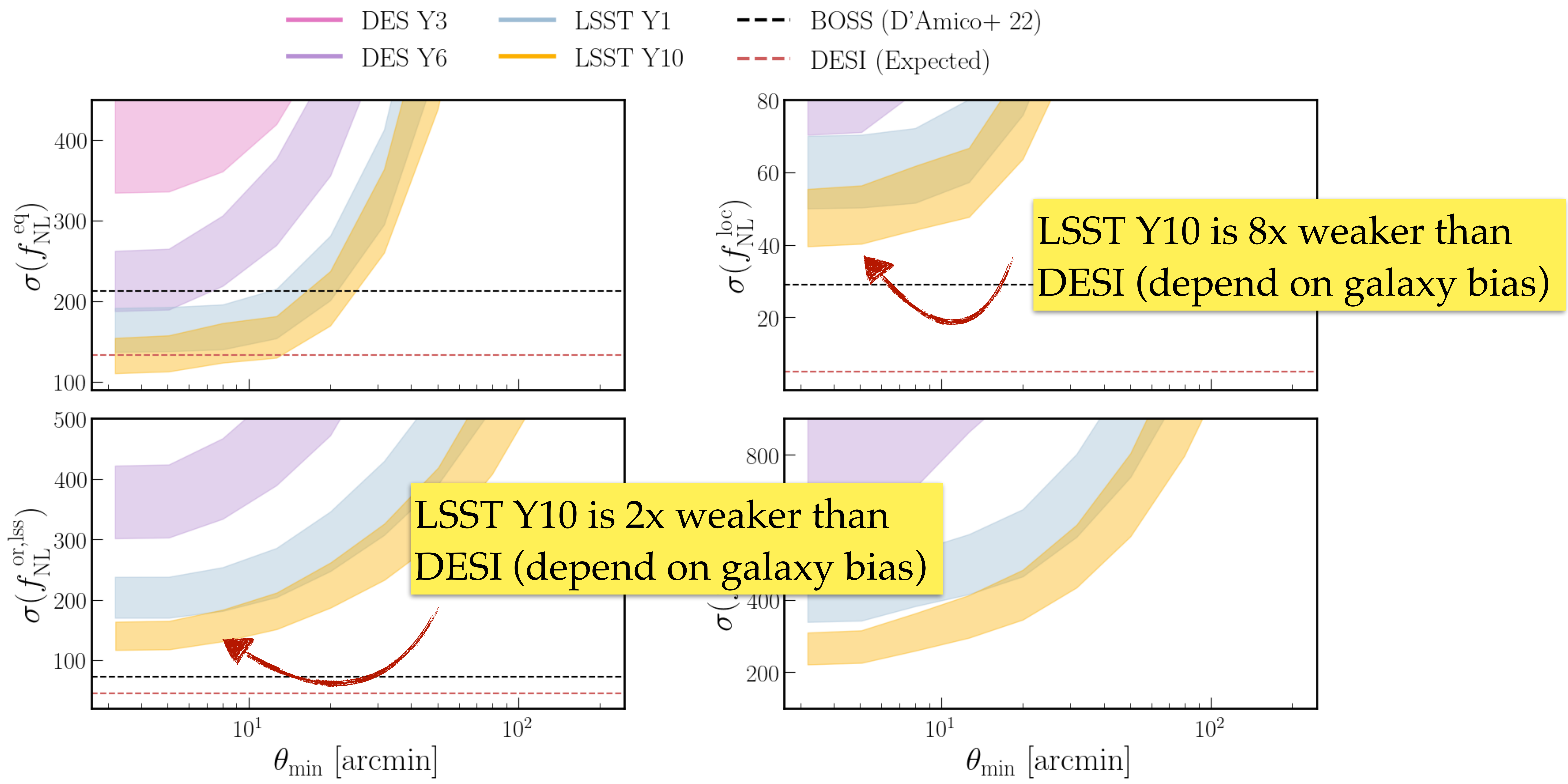
Comparison between datasets



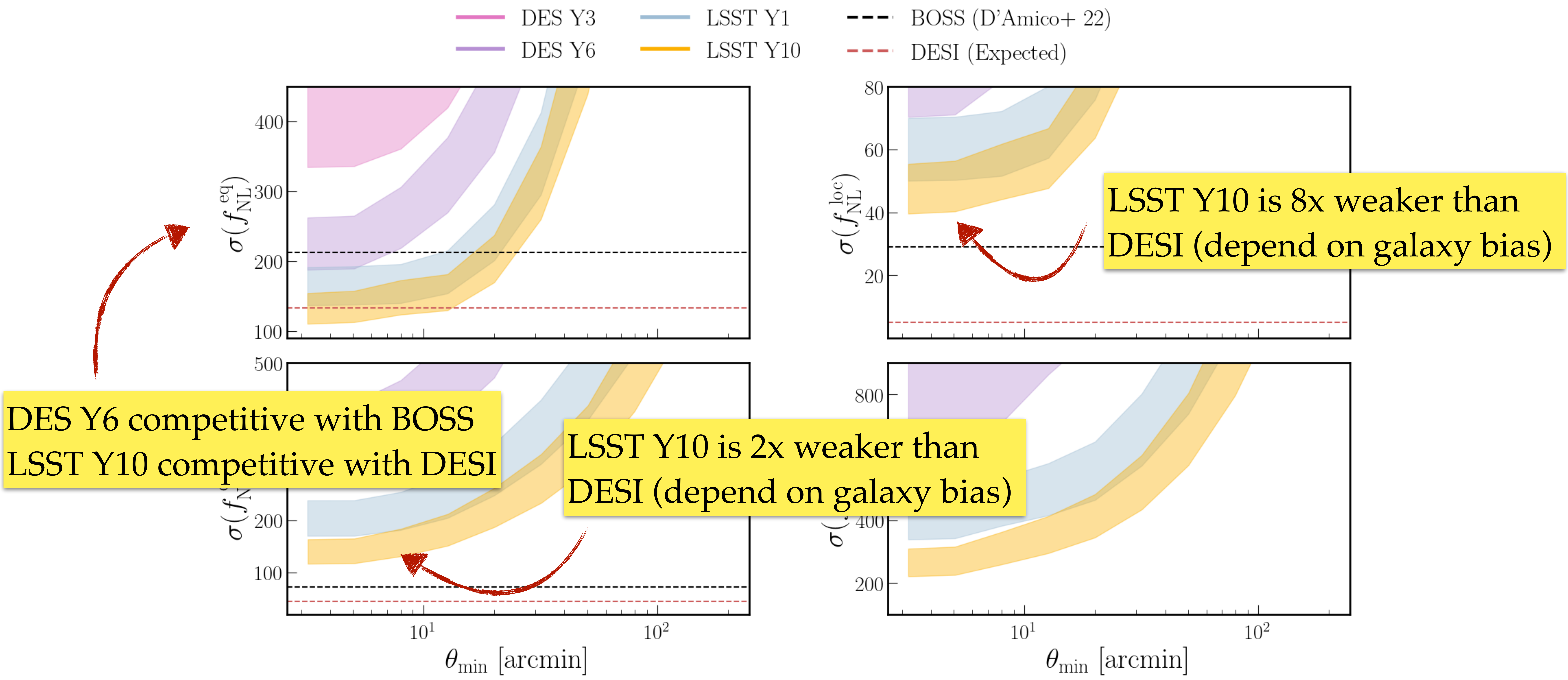
Comparison between datasets



Comparison between datasets



Comparison between datasets



Lensing HOS for early-universe physics

- We expect lensing HOS could contribute meaningfully to the constraints on fNL

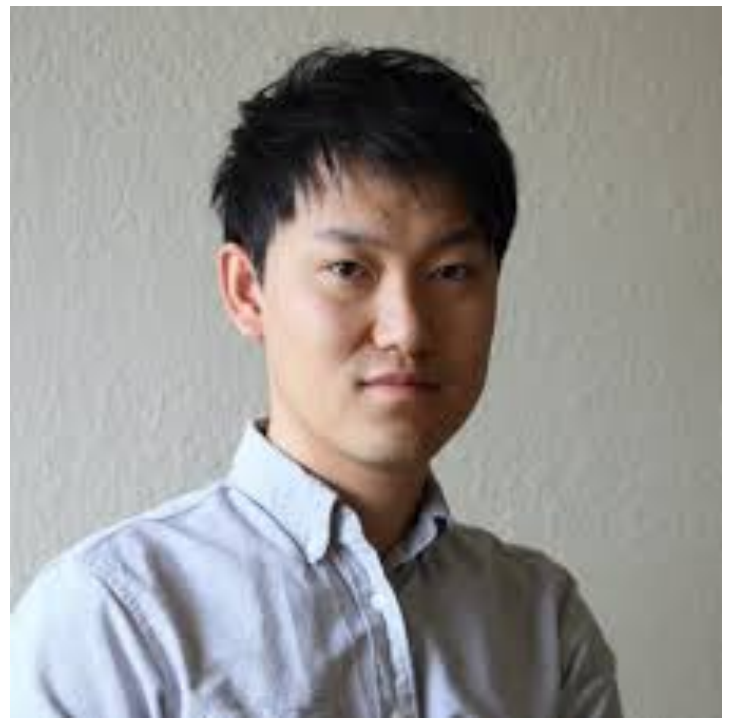
- Advantages of doing this:
  - Independent cross-check with scale-dependent galaxy bias measurements
  - Offer potential detection of scale-dependent PNG
  - Lensing simulations are faster than galaxies
  - Combined with galaxies could potentially calibrate galaxy bias

- Be imaginative in new things we can already test!

Outline

- The ΛCDM paradigm and extracting information beyond 2pt
- Practical challenges: beyond 2pt systematics
- Opportunities: primordial non-Gaussianity
- **Towards field-level inference**
- Summary & outlook

Omori, Zeghal, Lanusse, CC et al. (in prep)


Yuuki Omori


Justine Zehgal


François Lanusse

Full-field inference

- The ultimate level of modeling the LSS is when we have a field-level model of what we see — maps of initial conditions + cosmological and nuisance parameters
- Individually we know how to do these steps, the most non-trivial step is evolving gravity over cosmic time

Forward-model structure formation

There is a spectrum of implementations to forward-model structure formation

N-body — FastPM — LPT — Lognormal / Gaussian

Full physics
Computationally expensive

Approximate physics
Computationally cheap

Forward-model structure formation

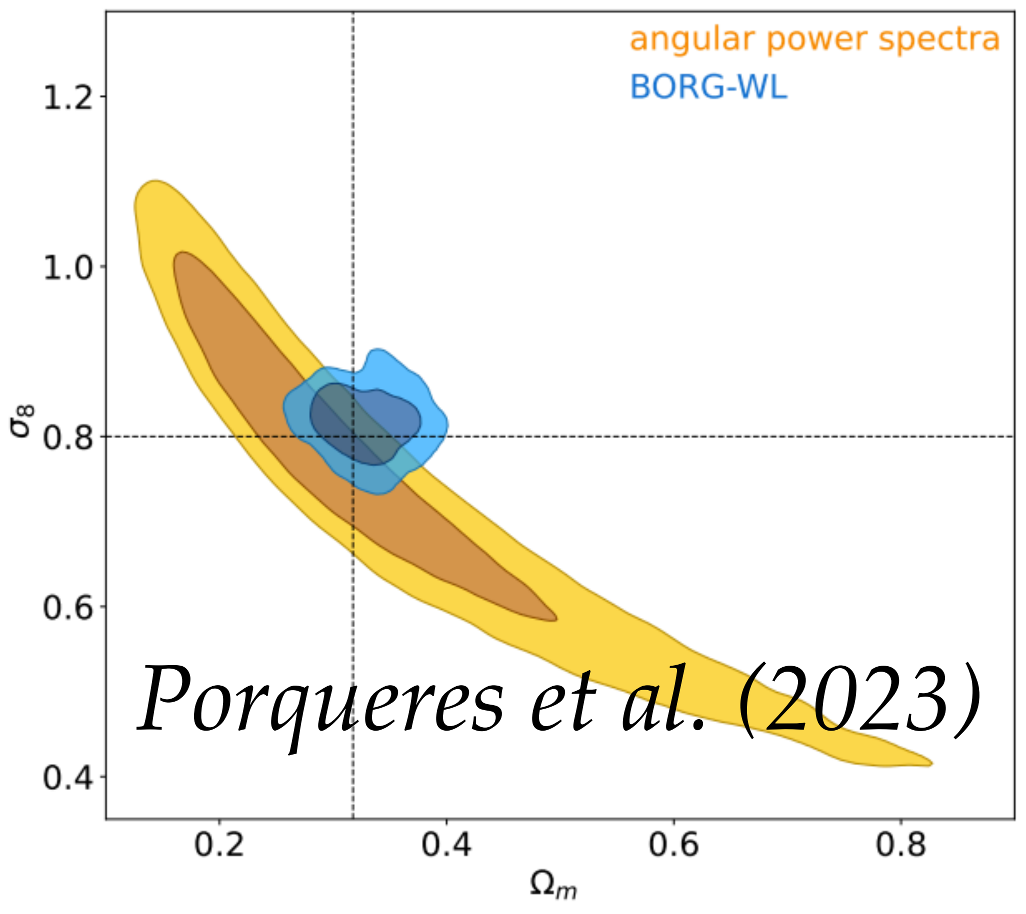
There is a spectrum of implementations to forward-model structure formation

N-body　　FastPM　　LPT　　Lognormal / Gaussian

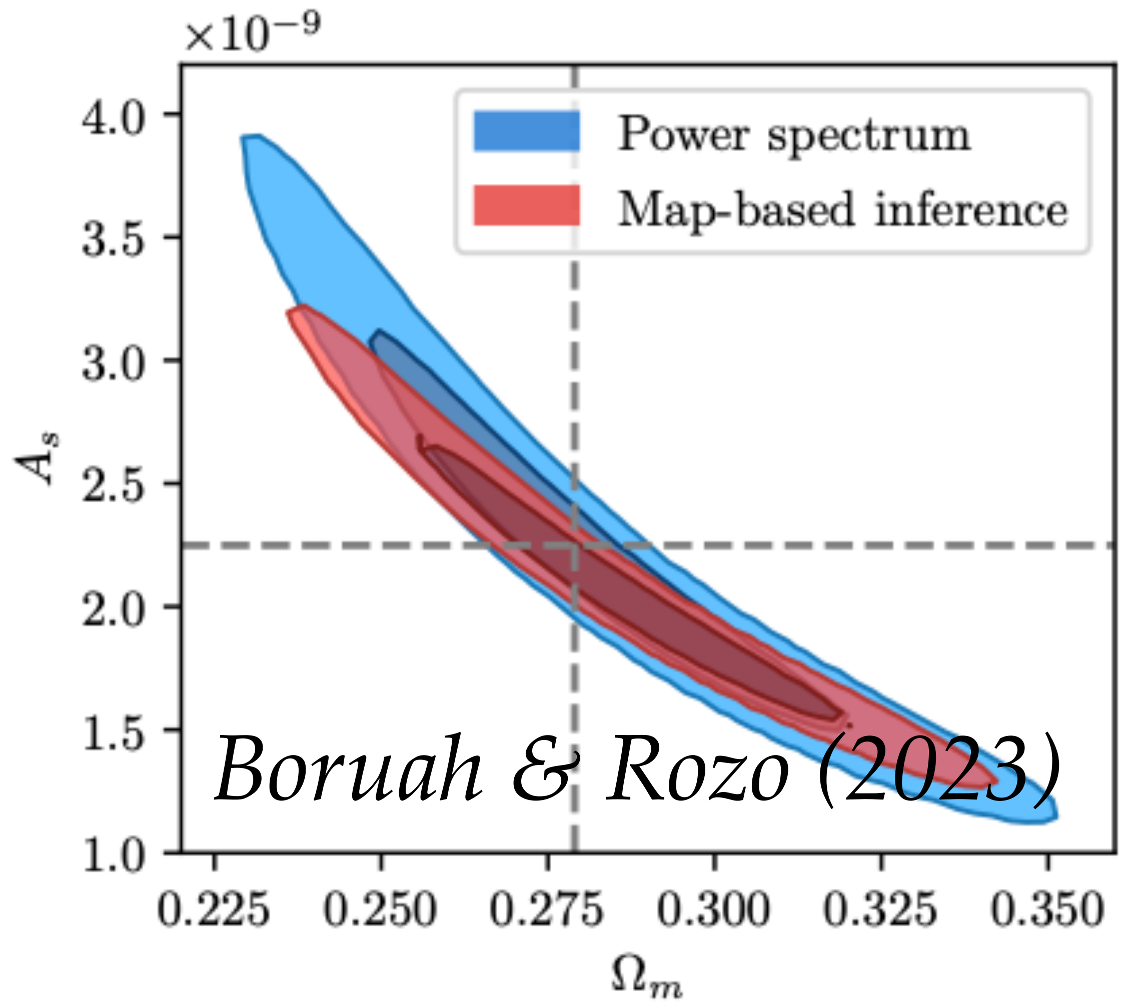
Full physics
Computationally expensive

Approximate physics
Computationally cheap

See talks from Florent, Eleni, Supranta, Arthur, Adrian…



Porqueres et al. (2023)



Boruah & Rozo (2023)

Forward-model structure formation

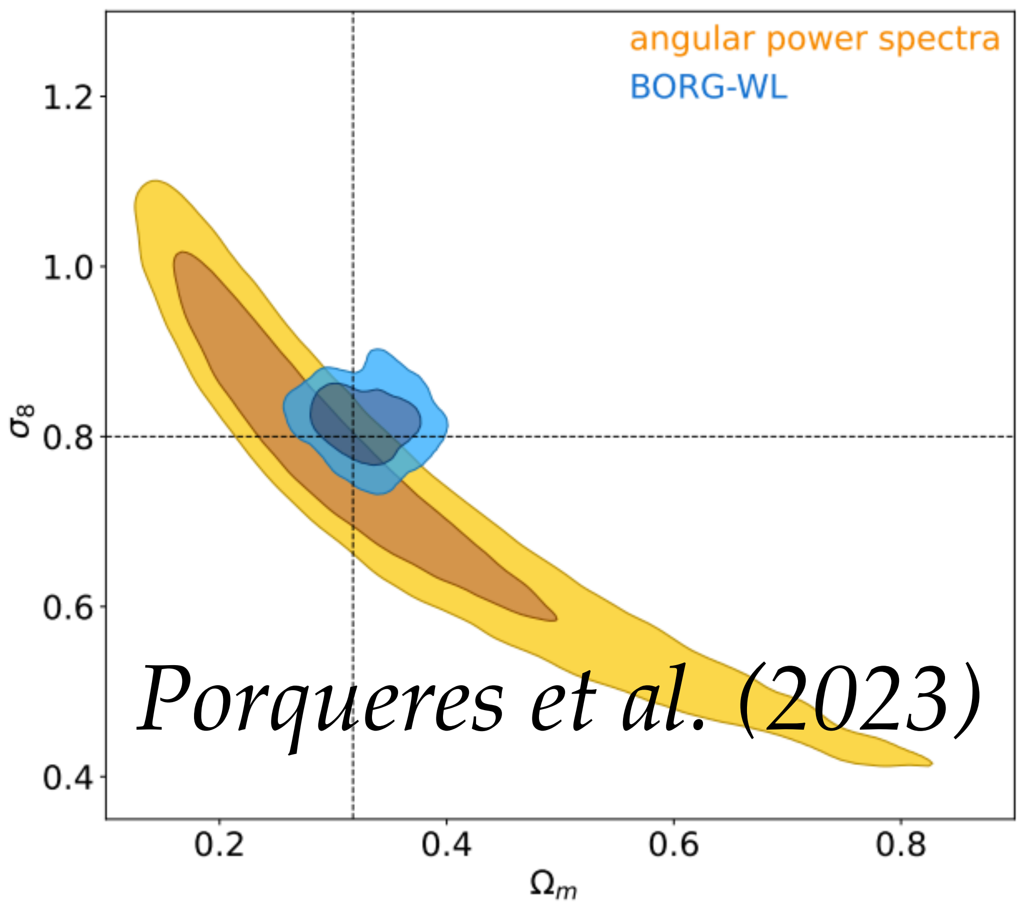
There is a spectrum of implementations to forward-model structure formation

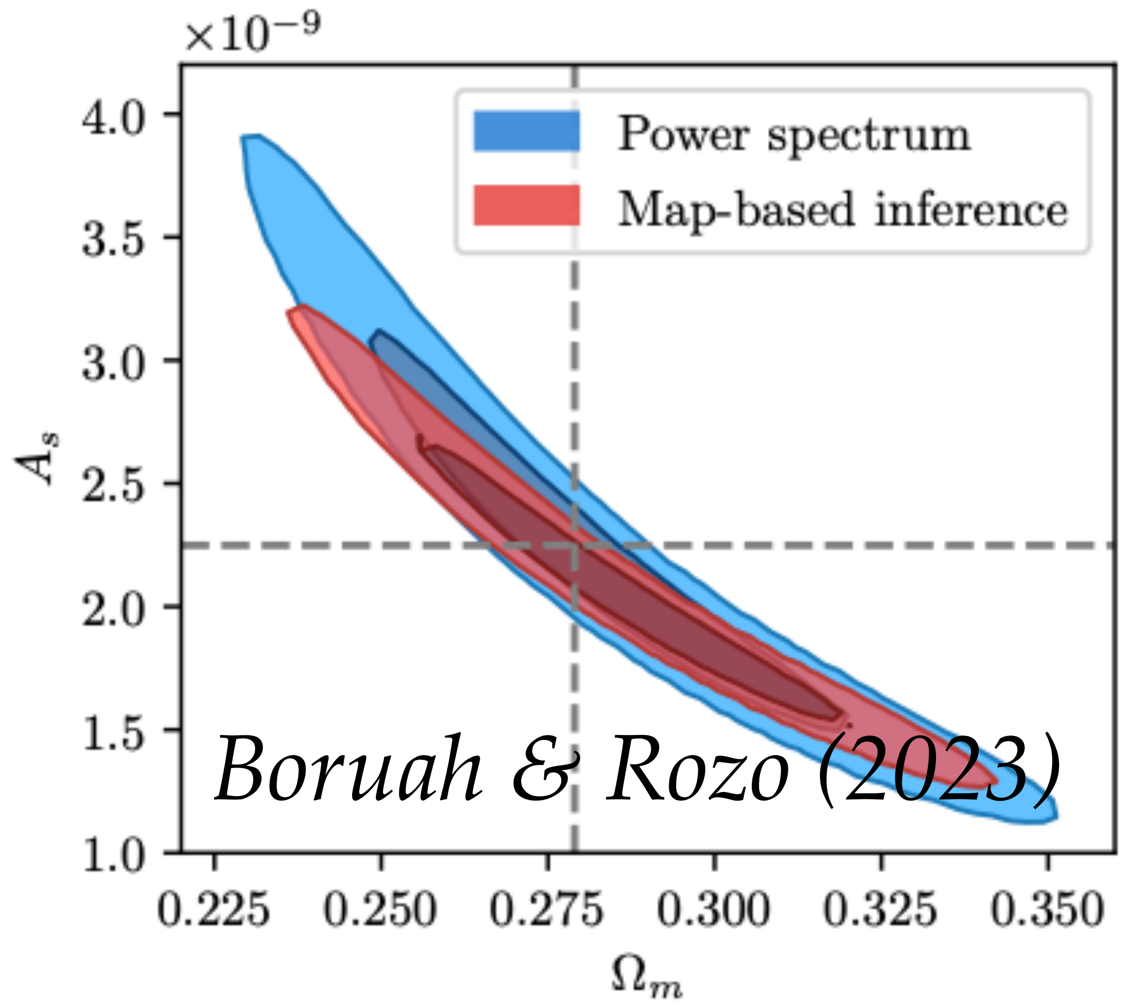
N-body | JaxPM | LPT | Lognormal Gaussian

Full physics
Computationally expensive

Approximate physics
Computationally cheap

See talks from Florent, Eleni, Supranta, Arthur, Adrian…





JaxPM-based field-level inference framework

- https://github.com/DifferentiableUniverseInitiative/JaxPM

Hybrid Physical-Neural ODEs for Fast N-body Simulations

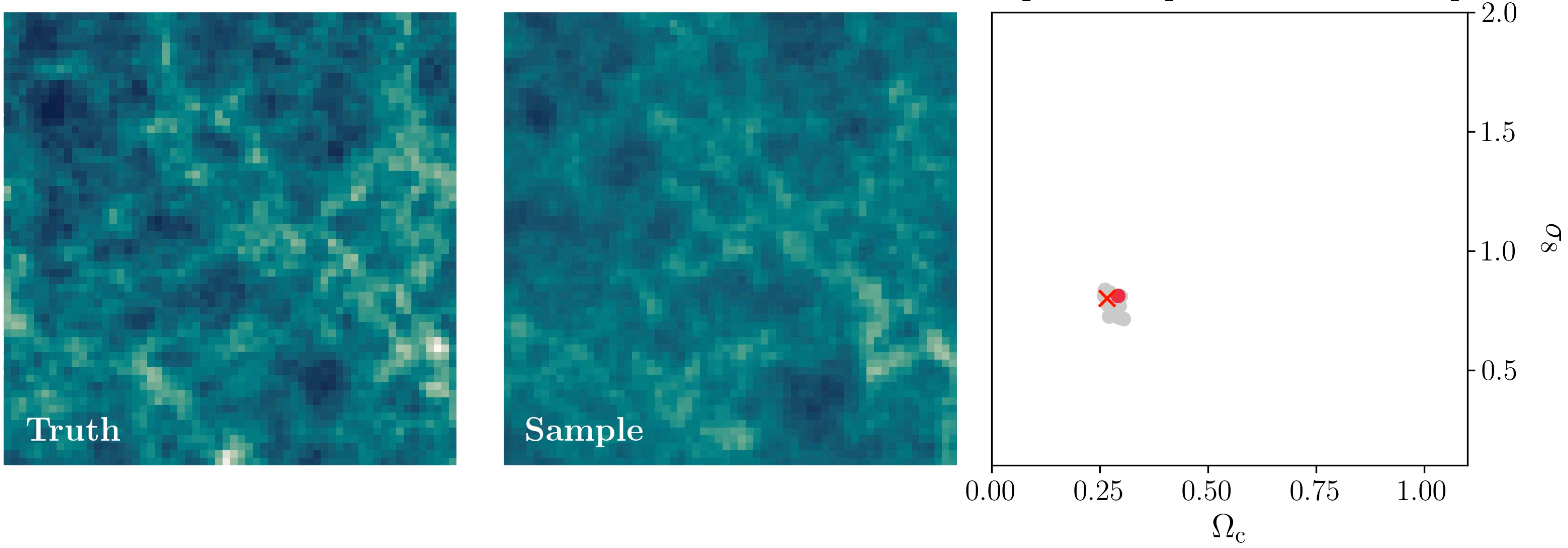
Denise Lanzieri [* 1] **François Lanusse** [* 2] **Jean-Luc Starck** [2]

(2023)

- Example: 2s (LPT), 90s (JaxPM) for single HMC step
  - LSST Y10-like, 5 redshift bins
  - 5x5 deg^2, 400x400x4600 Mpc/h, 200x200x128 pixels
- Require sampling ~5M parameters (!!)

The challenge of sampling

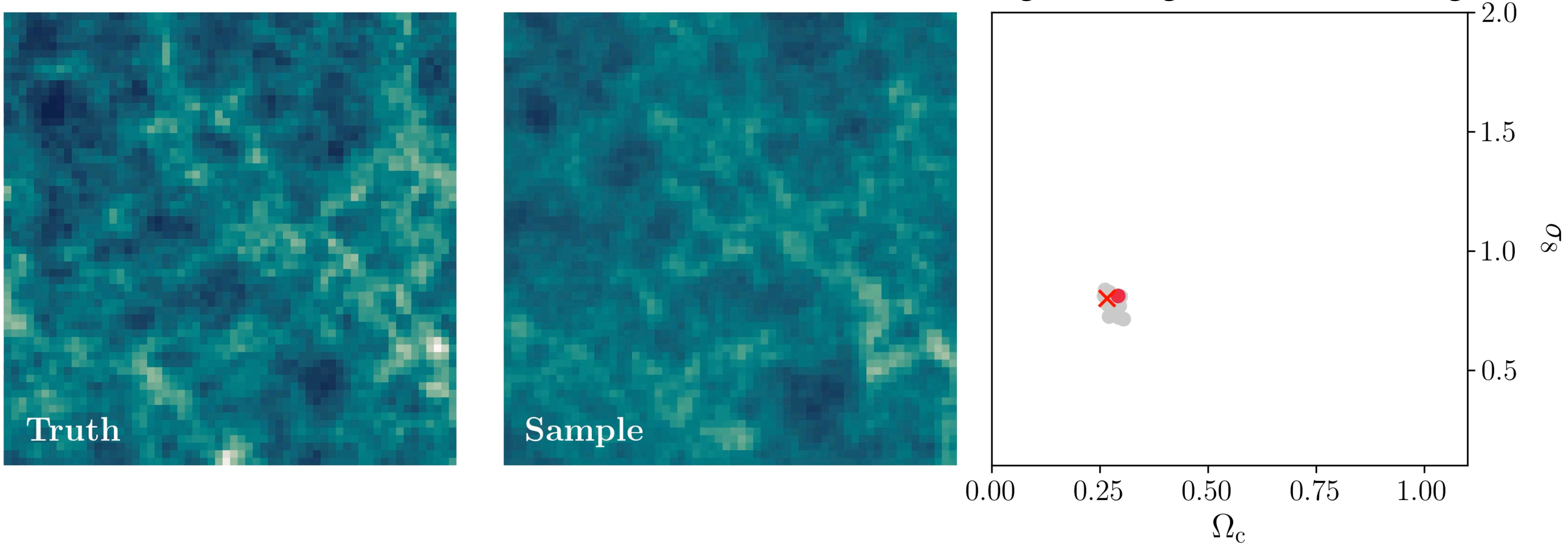
LPT, weak lensing convergence field, 25 deg²



Lots of fine-tuning is still needed to recover contours that make sense (ongoing work…)
Focus on being constantly grounded by the 2pt analysis we think we know how to do, and fair comparison between different methods

The challenge of sampling

LPT, weak lensing convergence field, 25 deg²



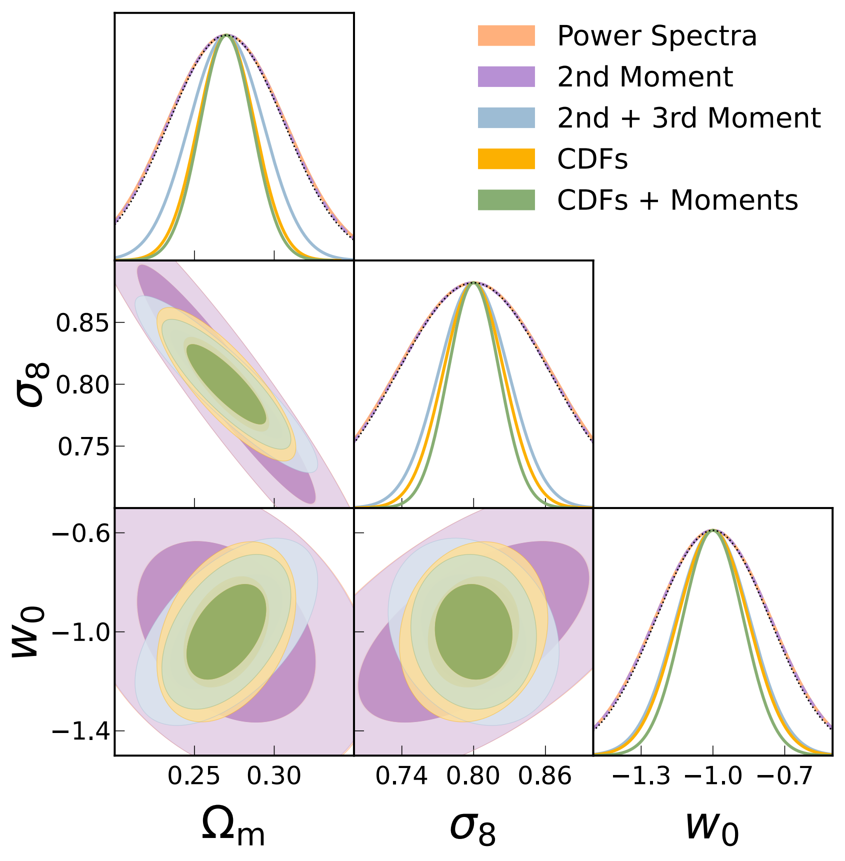
Lots of fine-tuning is still needed to recover contours that make sense (ongoing work…)
Focus on being constantly grounded by the 2pt analysis we think we know how to do, and fair comparison between different methods

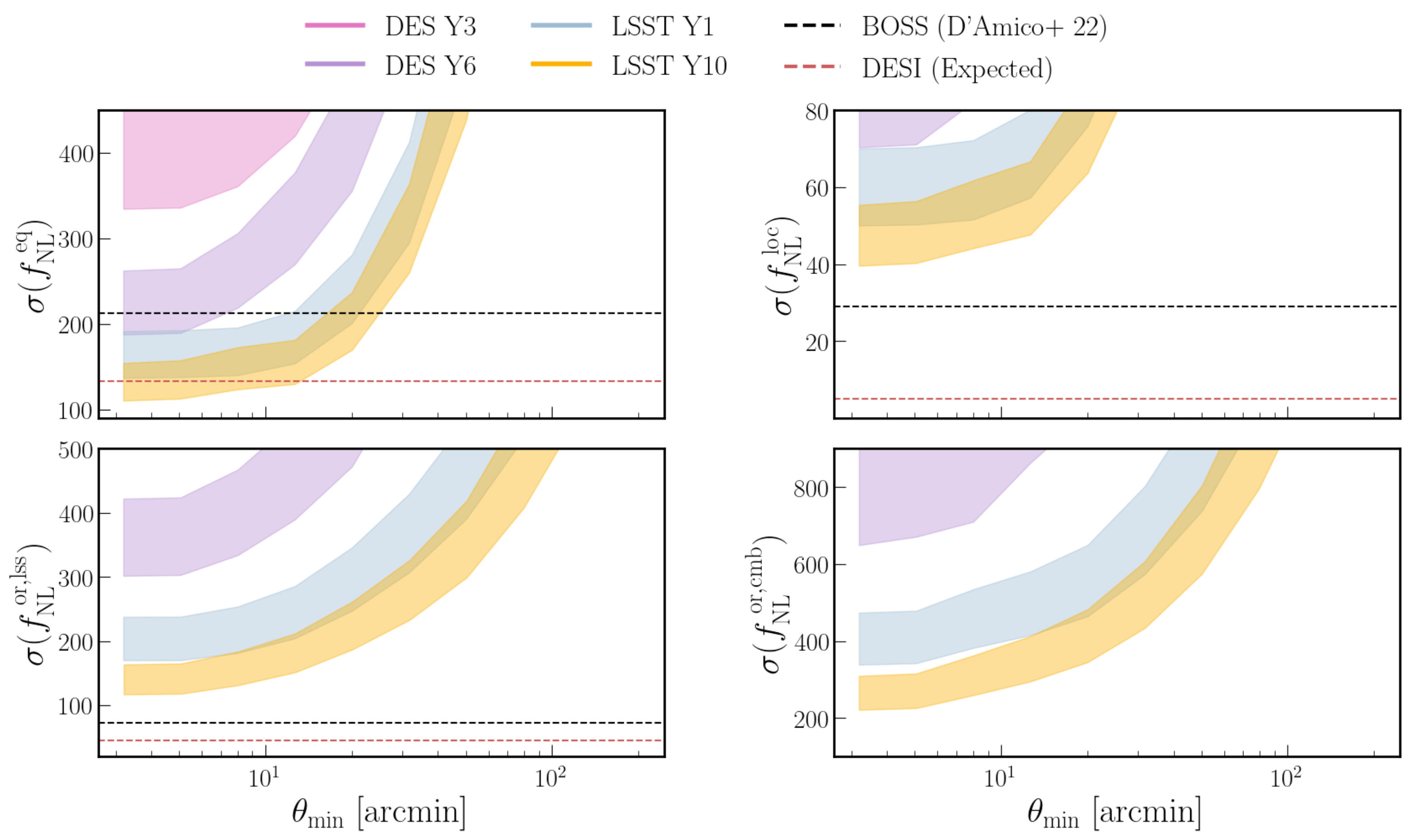
Outline

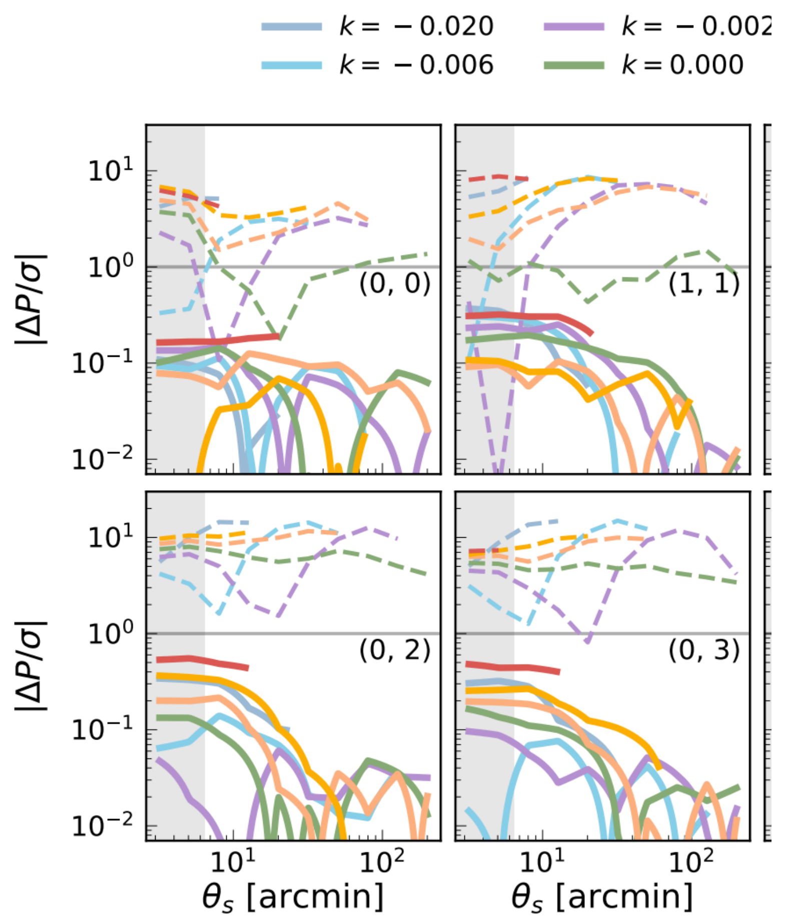
- The ΛCDM paradigm and extracting information beyond 2pt
- Opportunities: primordial non-Gaussianity
- Practical challenges: beyond 2pt systematics
- Towards field-level inference
- **Summary & outlook**

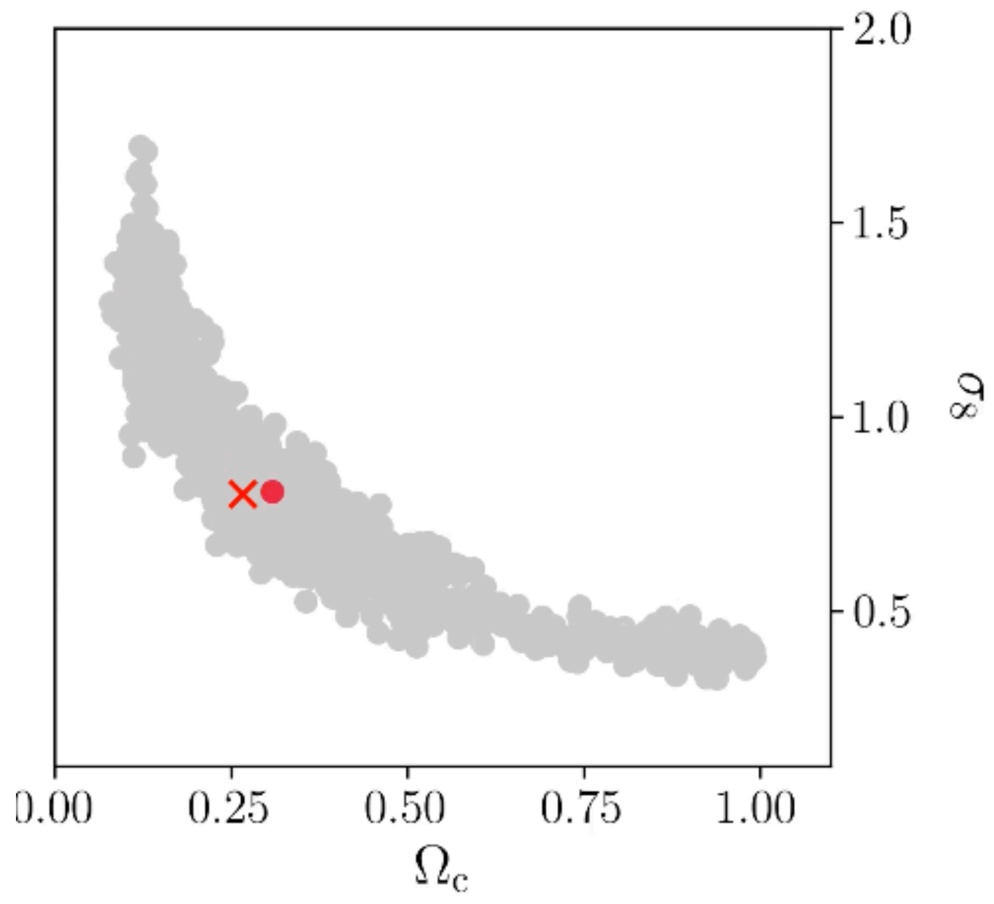
Summary & Outlook

- The gain in beyond 2pt statistics depends on **the science, the dataset, and how we extract the signal** — it may not always make sense to go to the highest order
- There are practical challenges that we are working on to make HOS **more robust**
- There is great opportunities for learning about the **new physics** via lensing HOS
- **Field-level inference** is the final frontier — lots of work needed to do full-physics

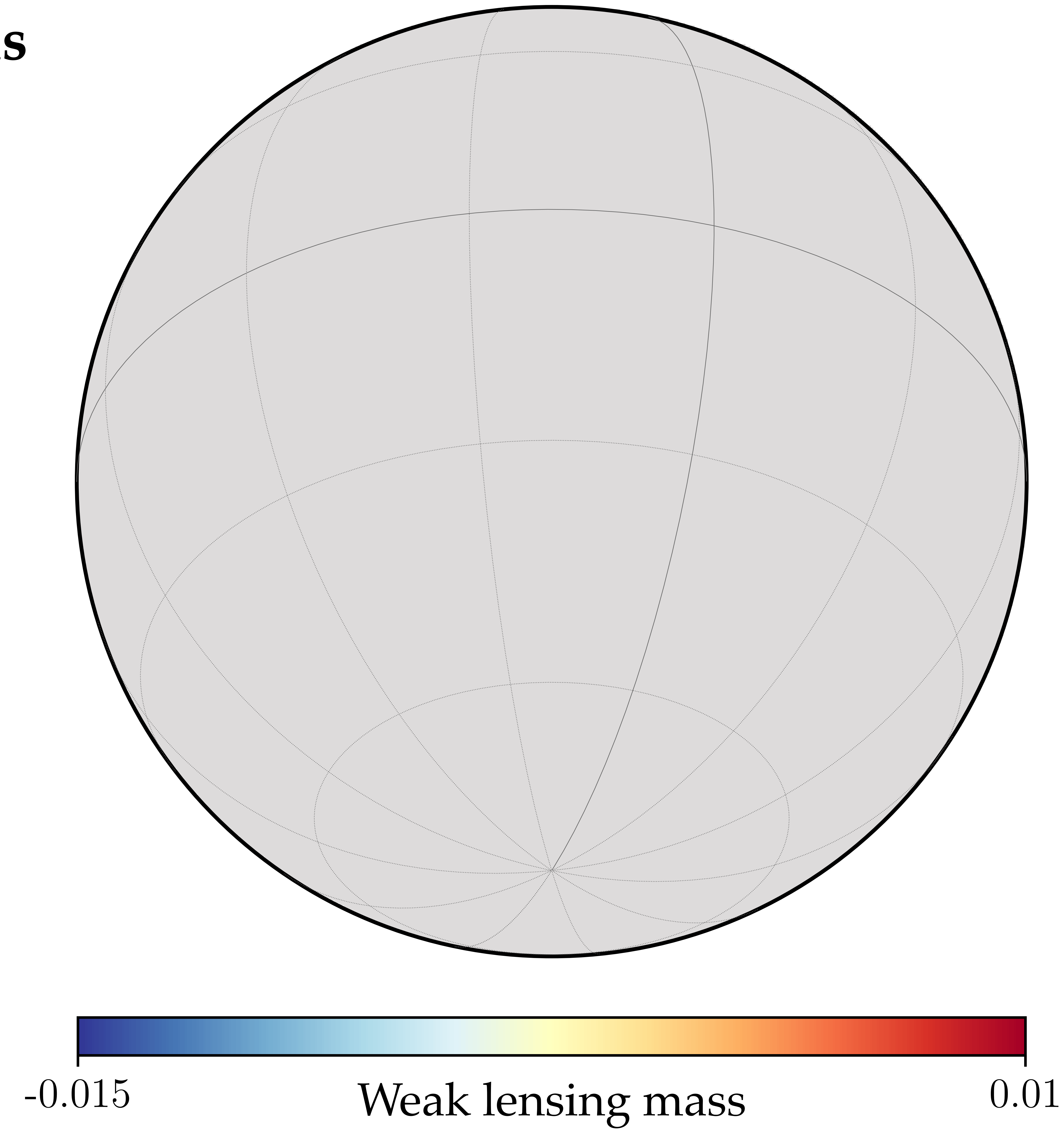




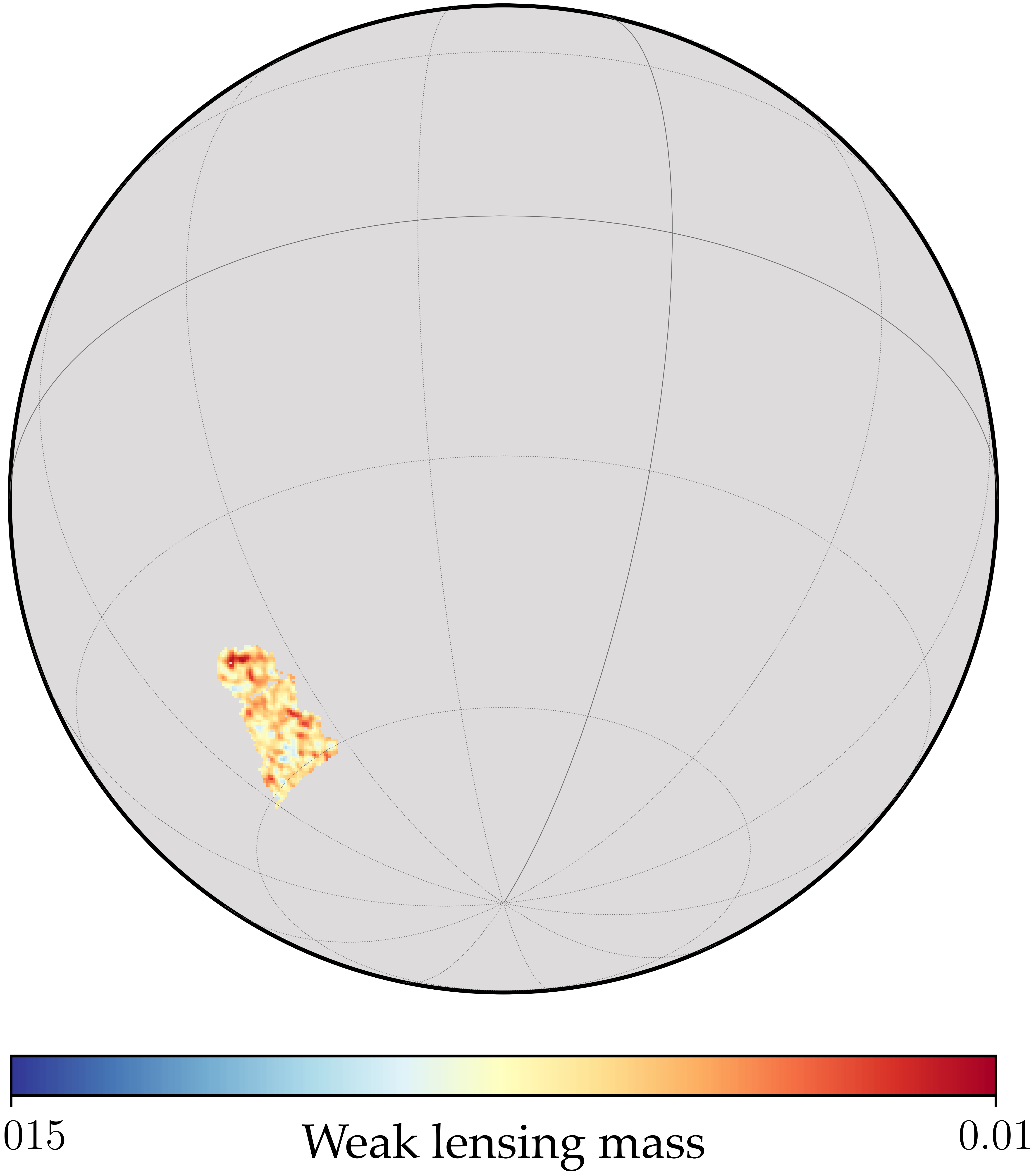




The community has made a lot of progress making these maps!



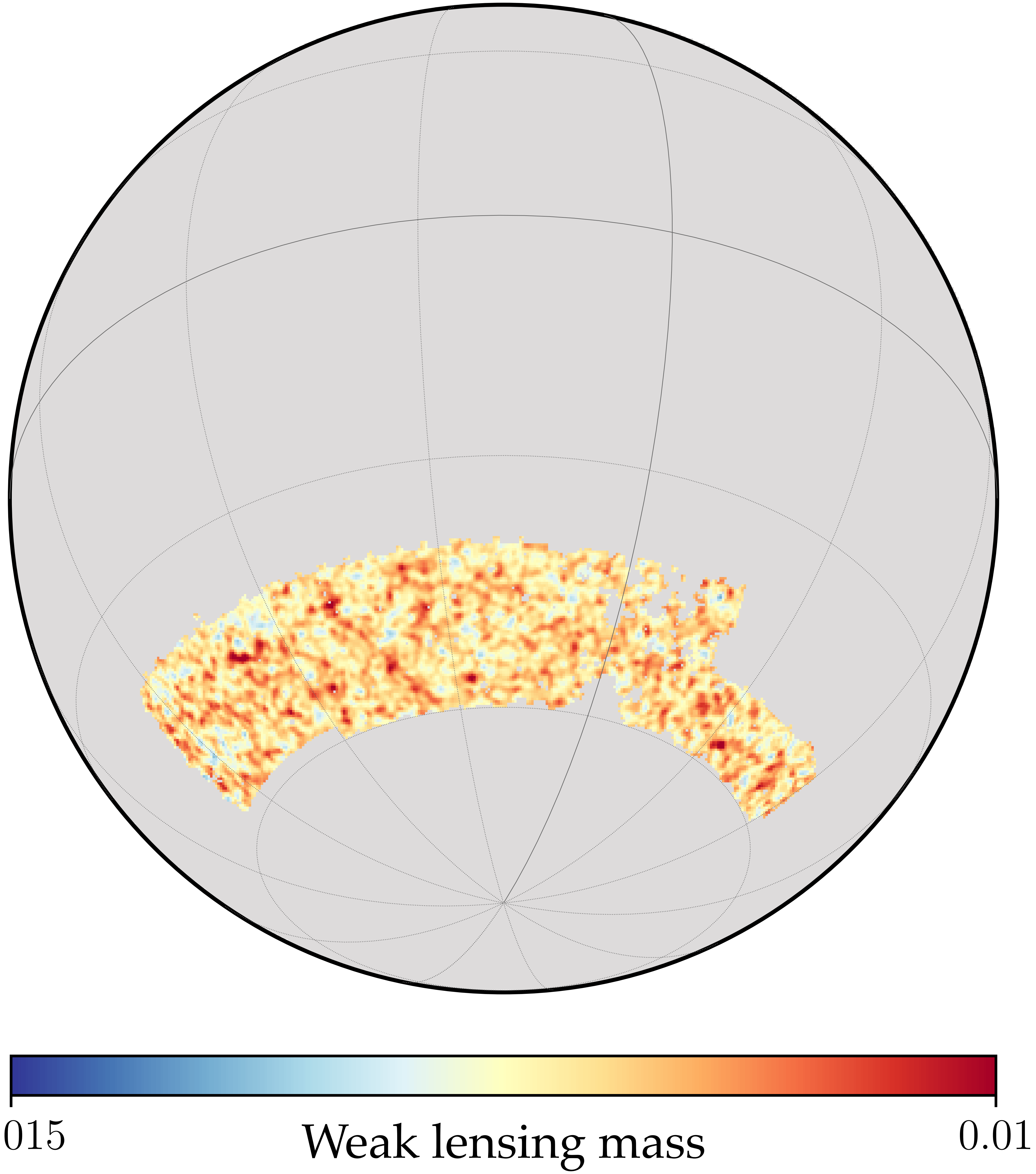
SV (2015)



150 deg^2
3.4 M galaxies

Vikram, CC, Jain et al. (2015)
CC, Vikram, Jain et al. (2015)
DES Collaboration

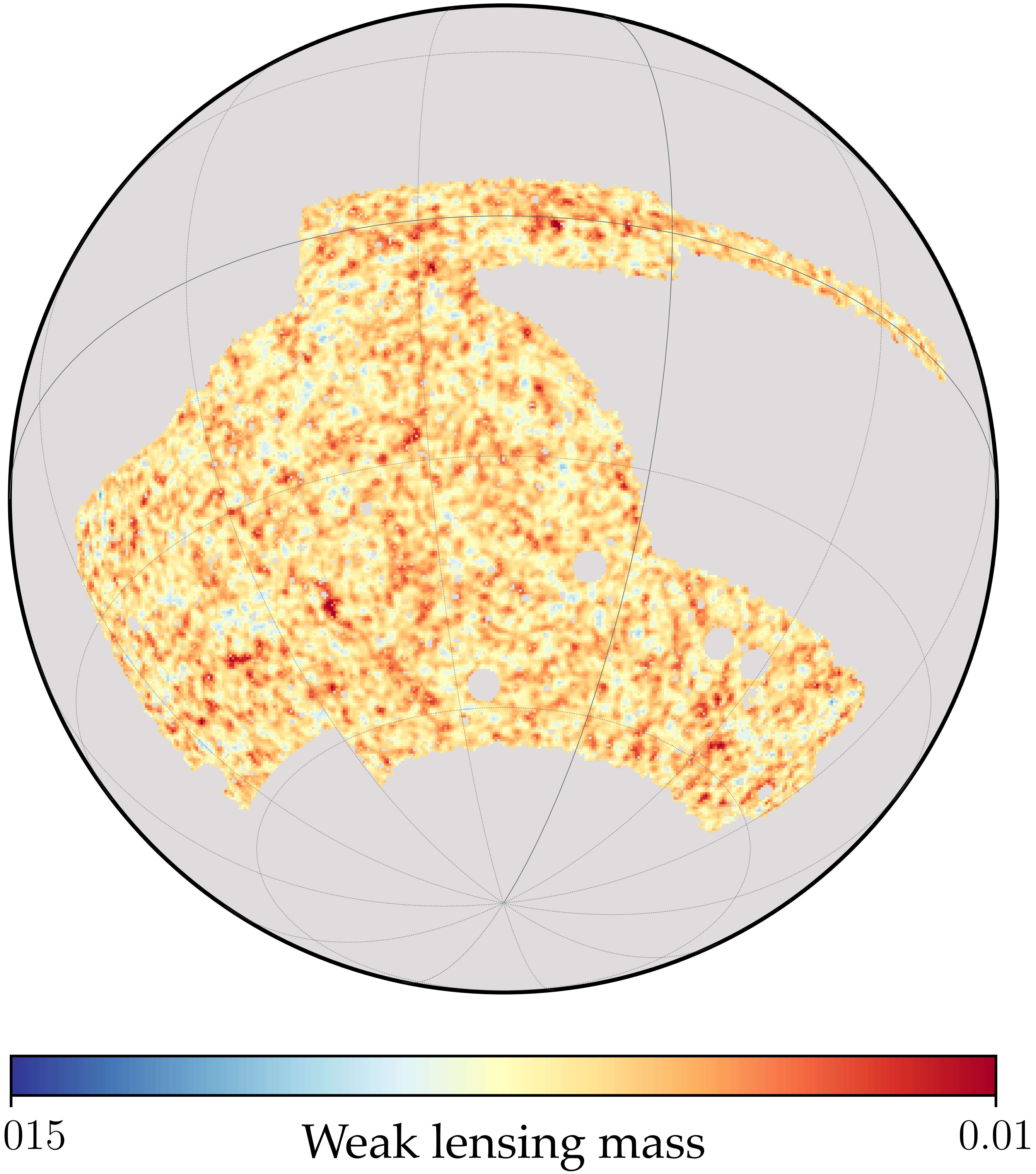
Y1 (2018)



1,300 deg^2
35 M galaxies

CC et al. (2018)
DES Collaboration

Y3 (2021)

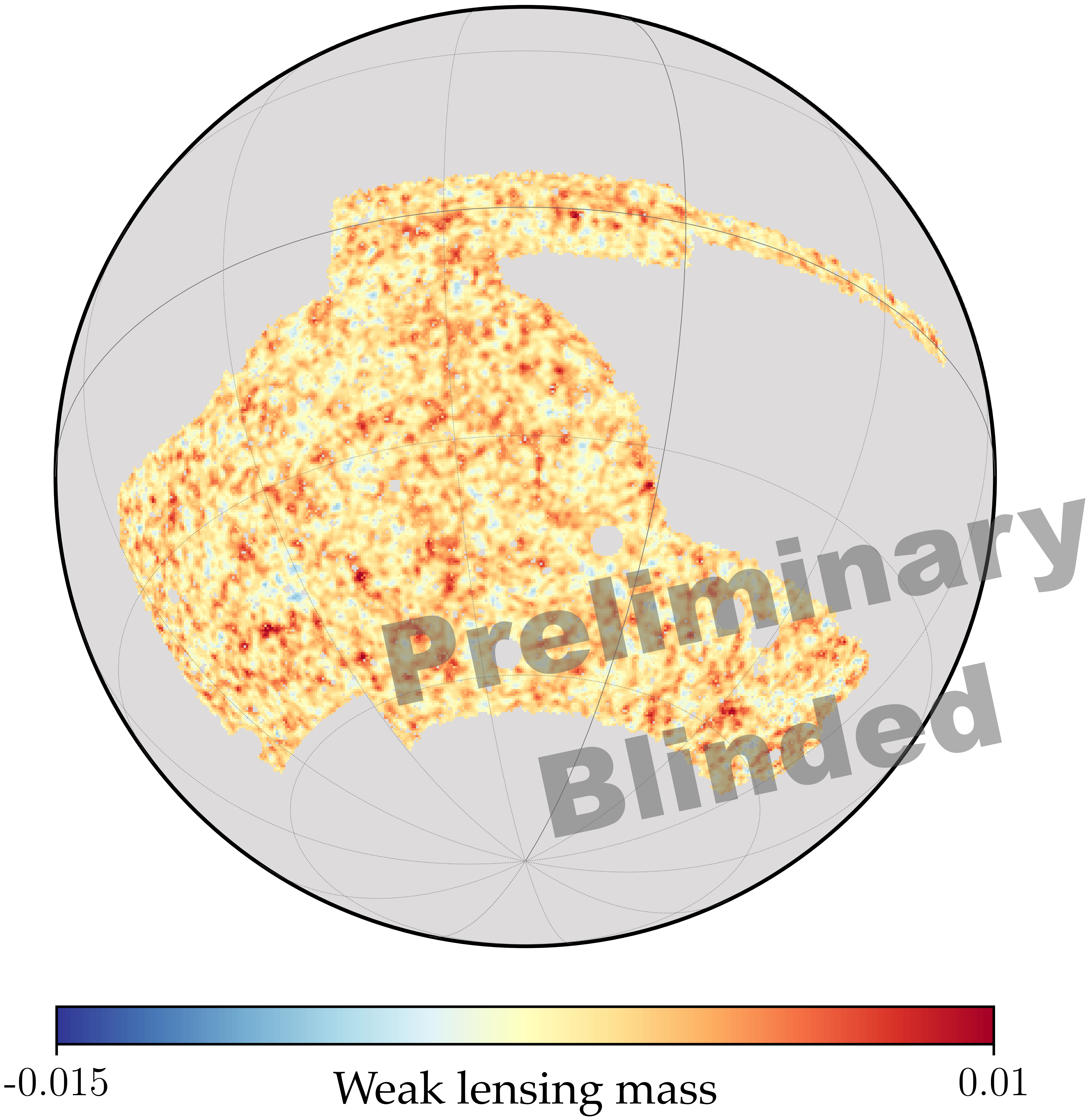


4,300 deg^2
100 M galaxies

Jeffrey, Gatti, CC et al. (2021)
DES Collaboration

Y6 (2024)

expected



~4,300 deg^2

~150 M galaxies

DES Collaboration

May 16


ASC INDUSTRIES
CAP. 4 TON
VERA C. RUBIN
OBSERVATORY
SLAC
CPWU 298401 9
FLAMMABLE
KEEP FIRE AWAY

Summary & Outlook

- The gain in beyond 2pt statistics depends on **the science, the dataset, and how we extract the signal** — it may not always make sense to go to the highest order
- There are practical challenges that we are working on to make HOS **more robust**
- There is great opportunities for learning about the **new physics** via lensing HOS
- **Field-level inference** is the final frontier — lots of work needed to do full-physics