Field-level inference and new observables for the next era of galaxy surveys

or why I am excited about large-scale structure cosmology

Nguyễn Nhật Minh (Minh Nguyen) - Kavli IPMU, University of Tokyo
Email: nhat.minh.nguyen@ipmu.jp.

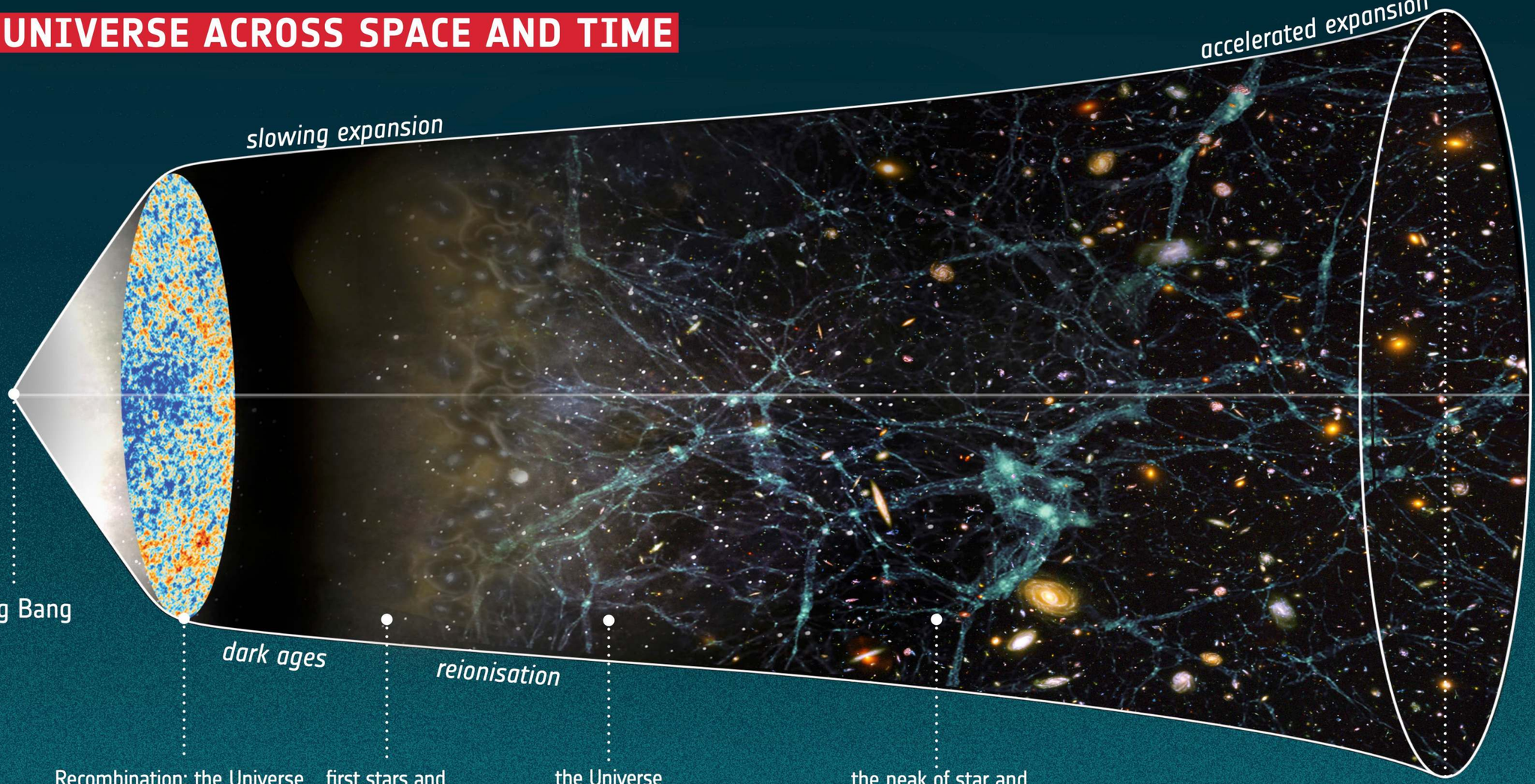
Credit: DESI/NOIRLab

LeCosPa seminar - NTU, Taipei, Taiwan - 2026 June 1

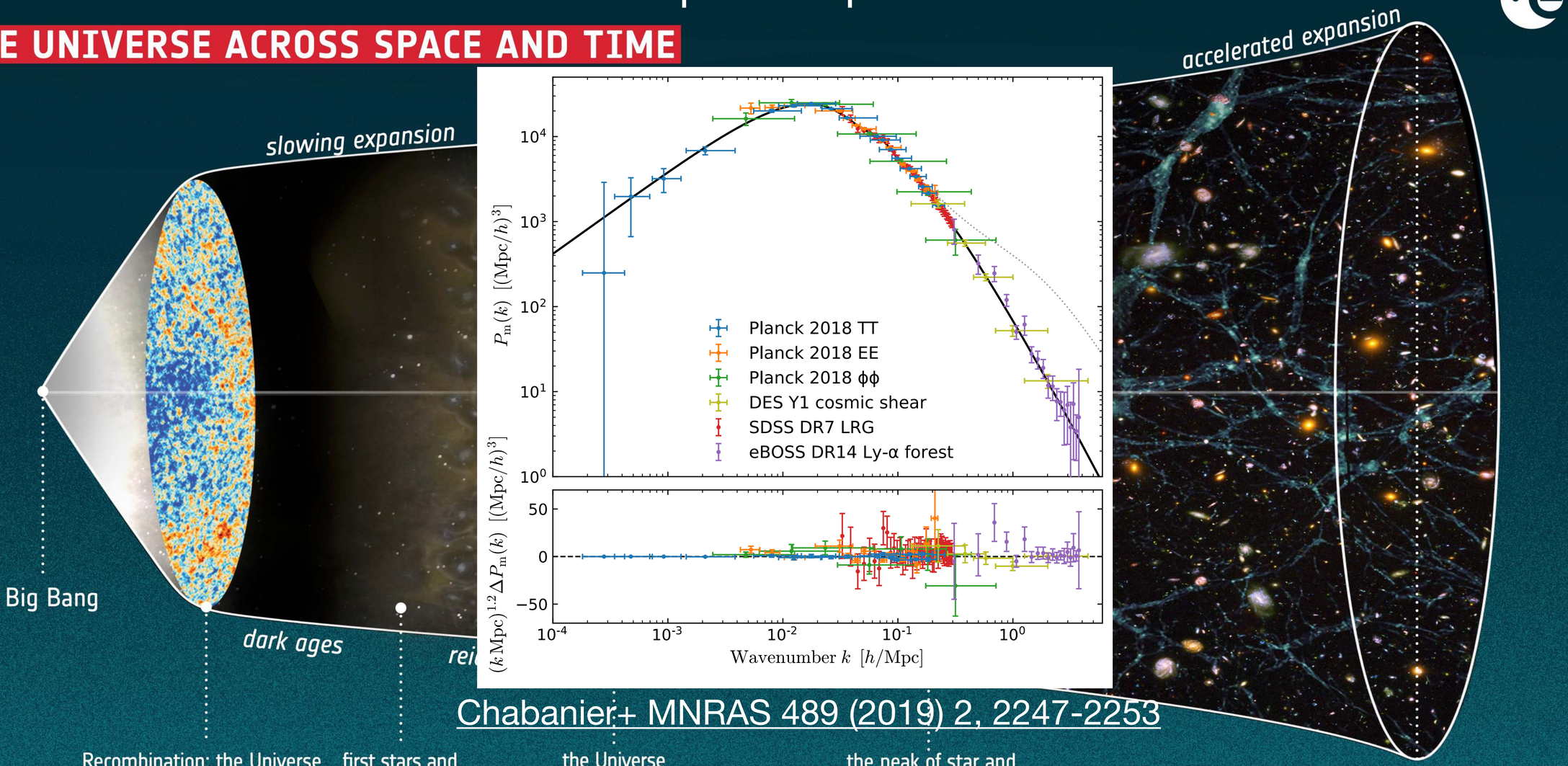
Why do we do cosmology (from LSS)?

Credit: Dark Energy Spectroscopic Instrument

THE UNIVERSE ACROSS SPACE AND TIME



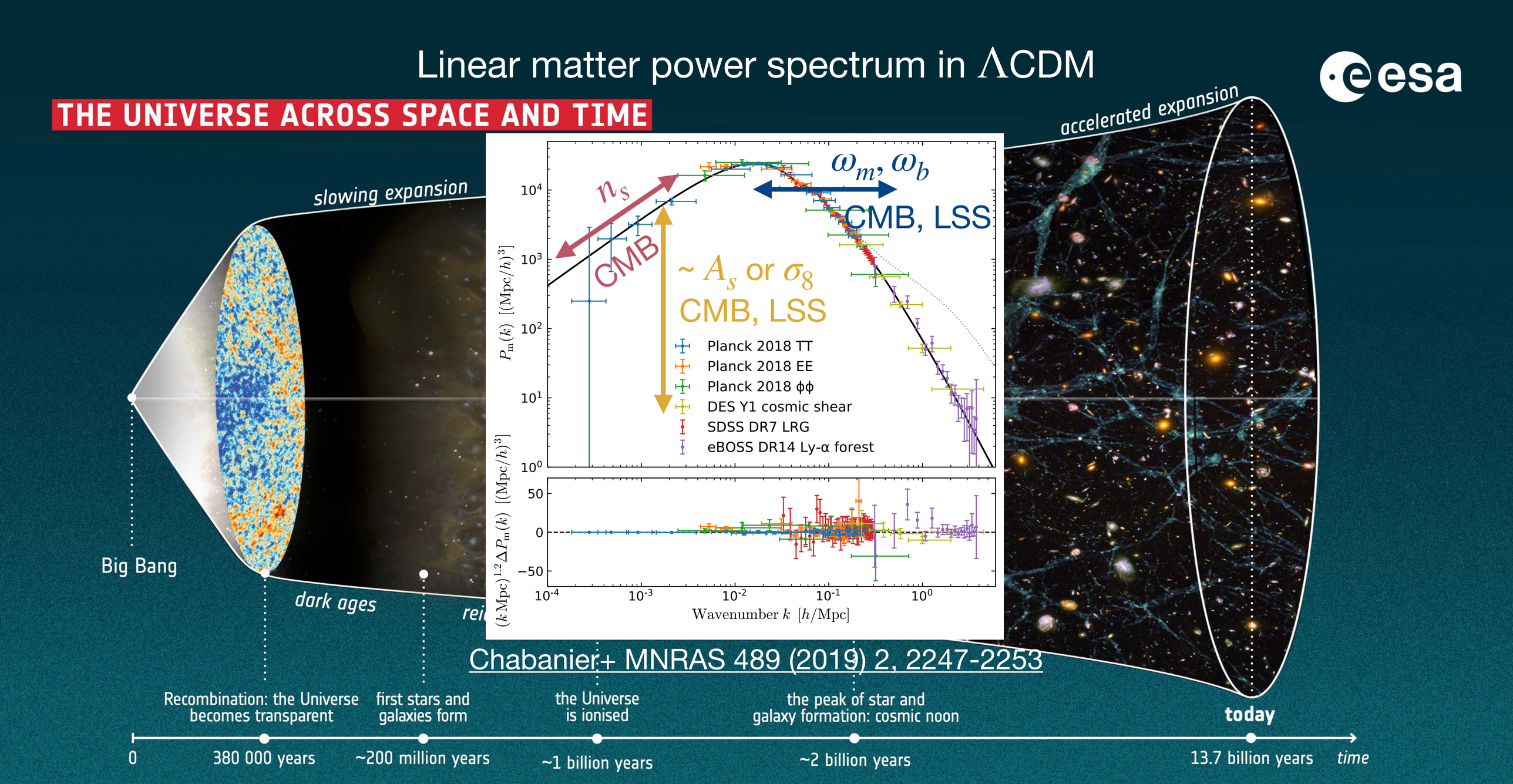
Linear matter power spectrum in ΛCDM



Chabanier+ MNRAS 489 (2019) 2, 2247-2253

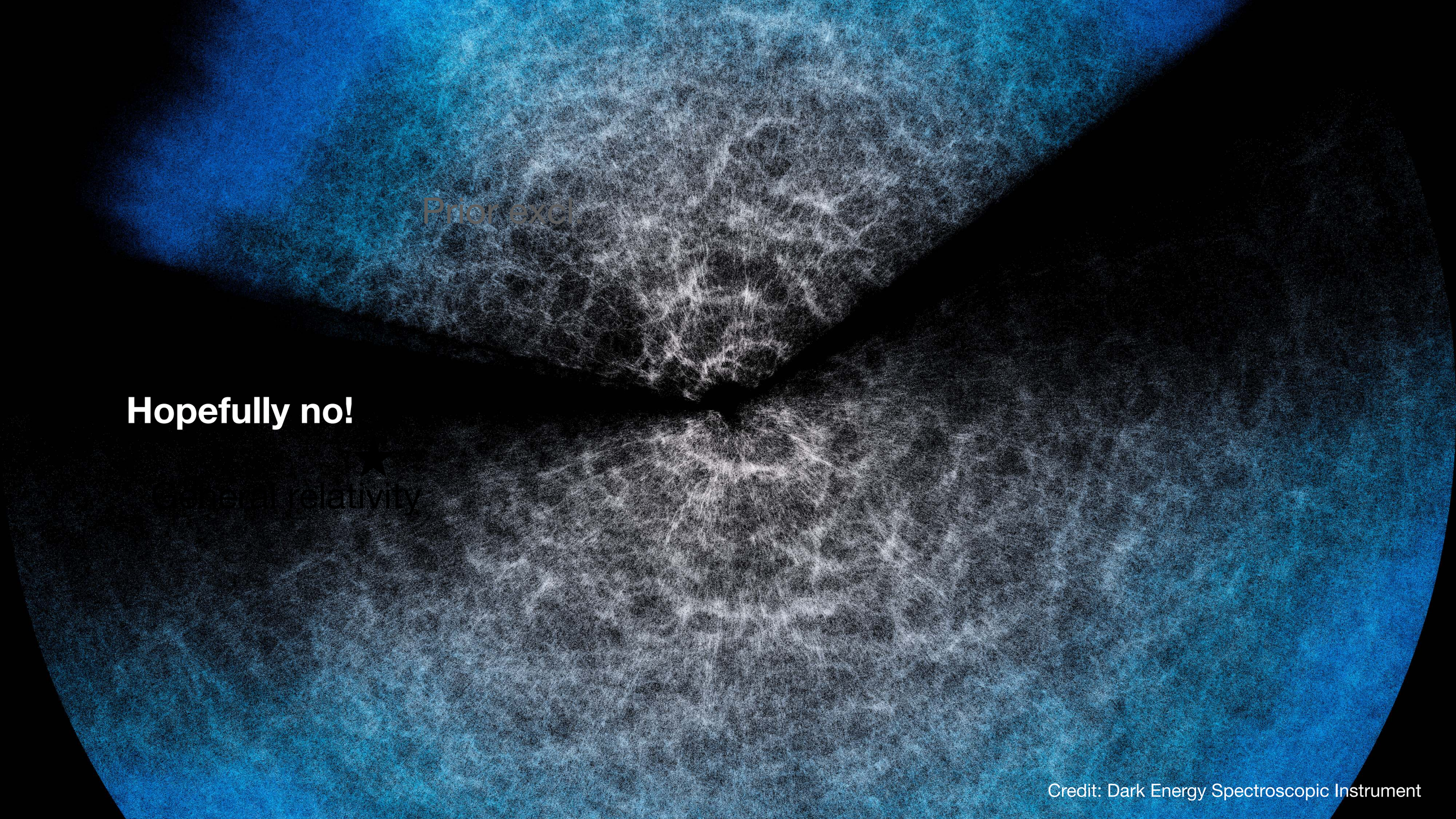
Cosmology from LSS aims to find the model that describes both CMB and LSS

Linear matter power spectrum in ΛCDM



Chabanier+ MNRAS 489 (2019) 2, 2247-2253

The standard model ΛCDM is boringly robust. Is it the end of cosmology (from LSS)?



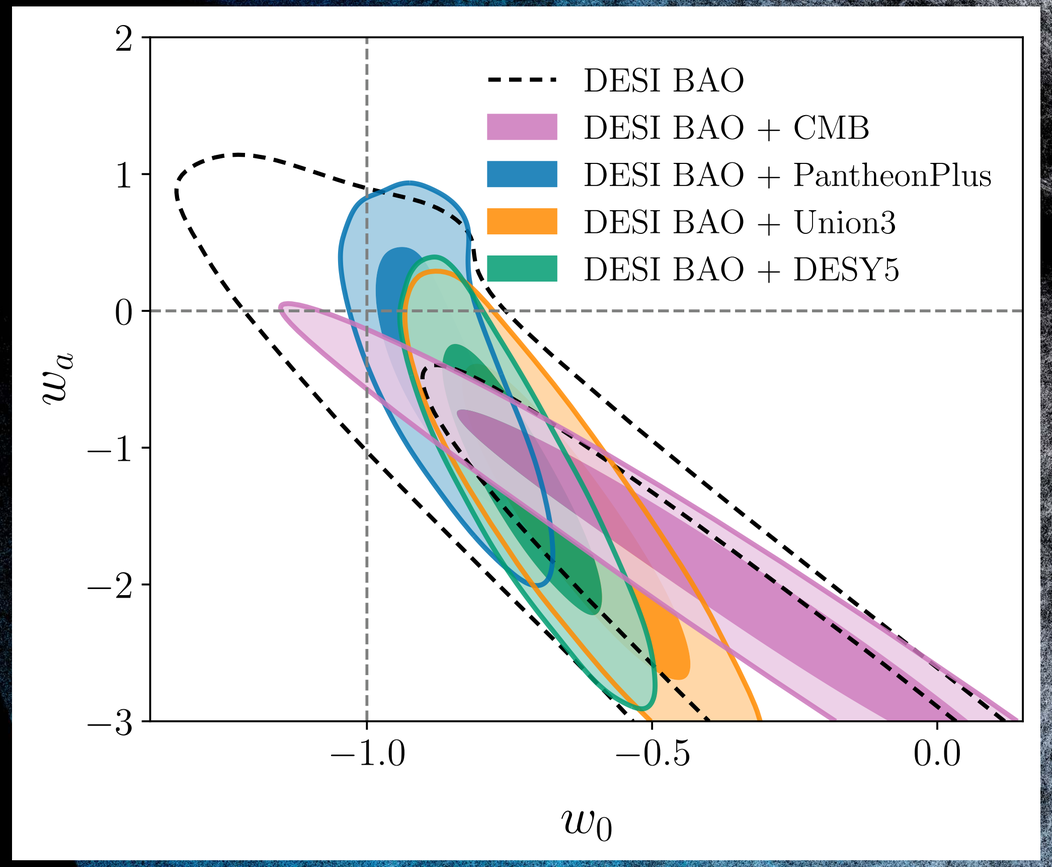
Hopefully no!

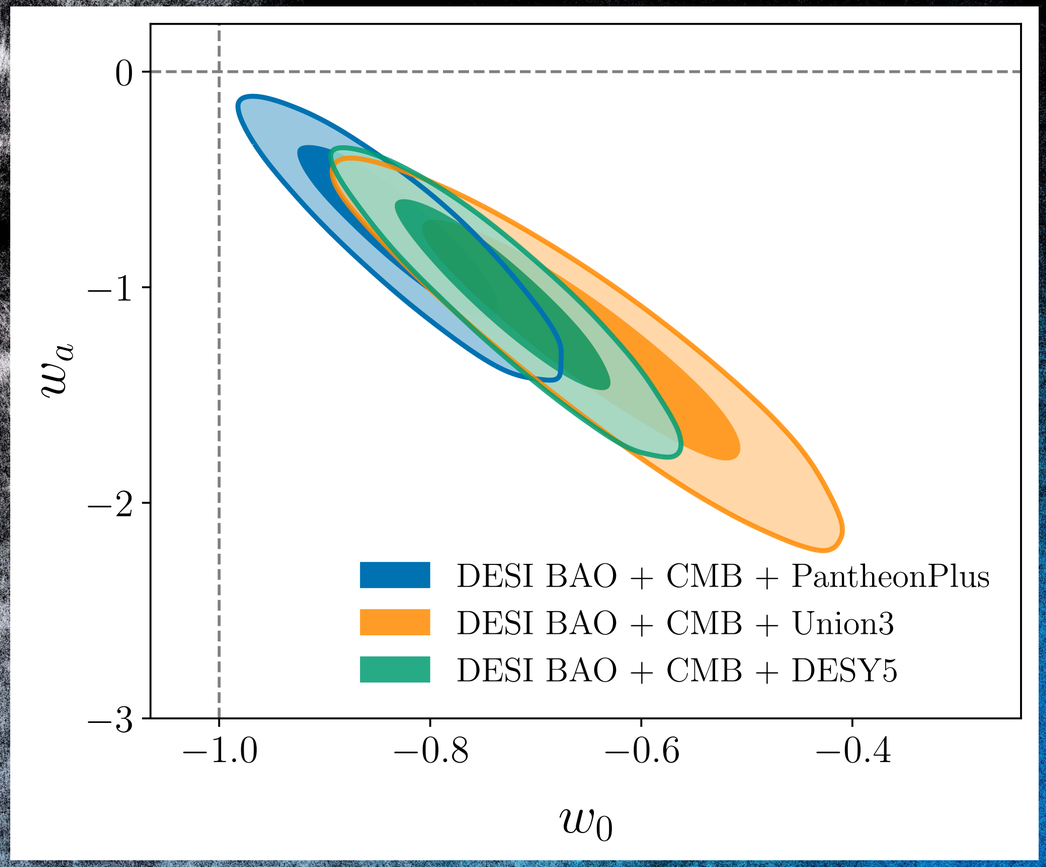
Credit: Dark Energy Spectroscopic Instrument

What I am generally excited about - Cosmic Acceleration and Dynamical Dark Energy

Dark energy appears to be evolving with time as $w = w_0 + w_a(1 - a)$

DESI 2024 VI, JCAP02(2025) incl. MN

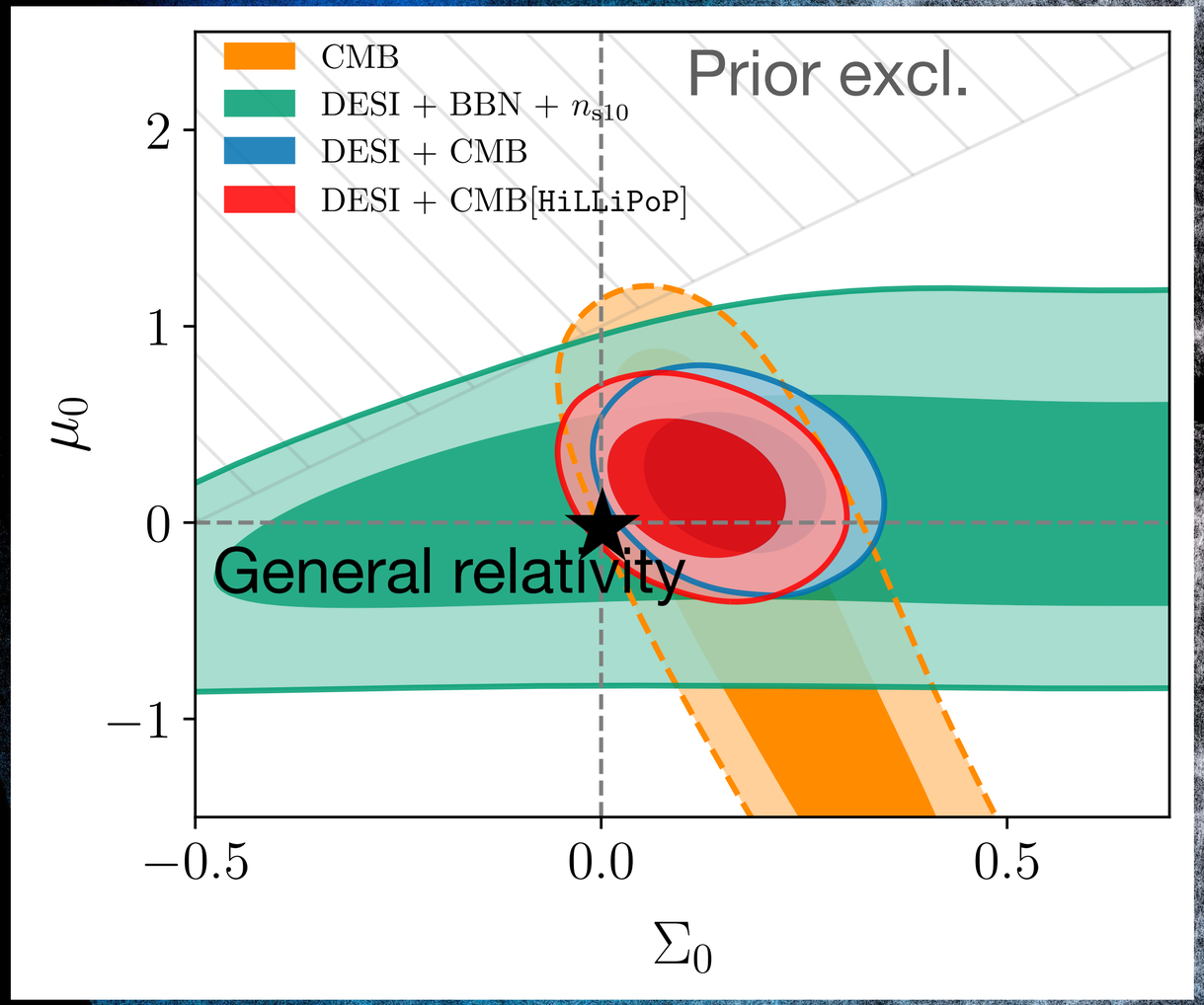


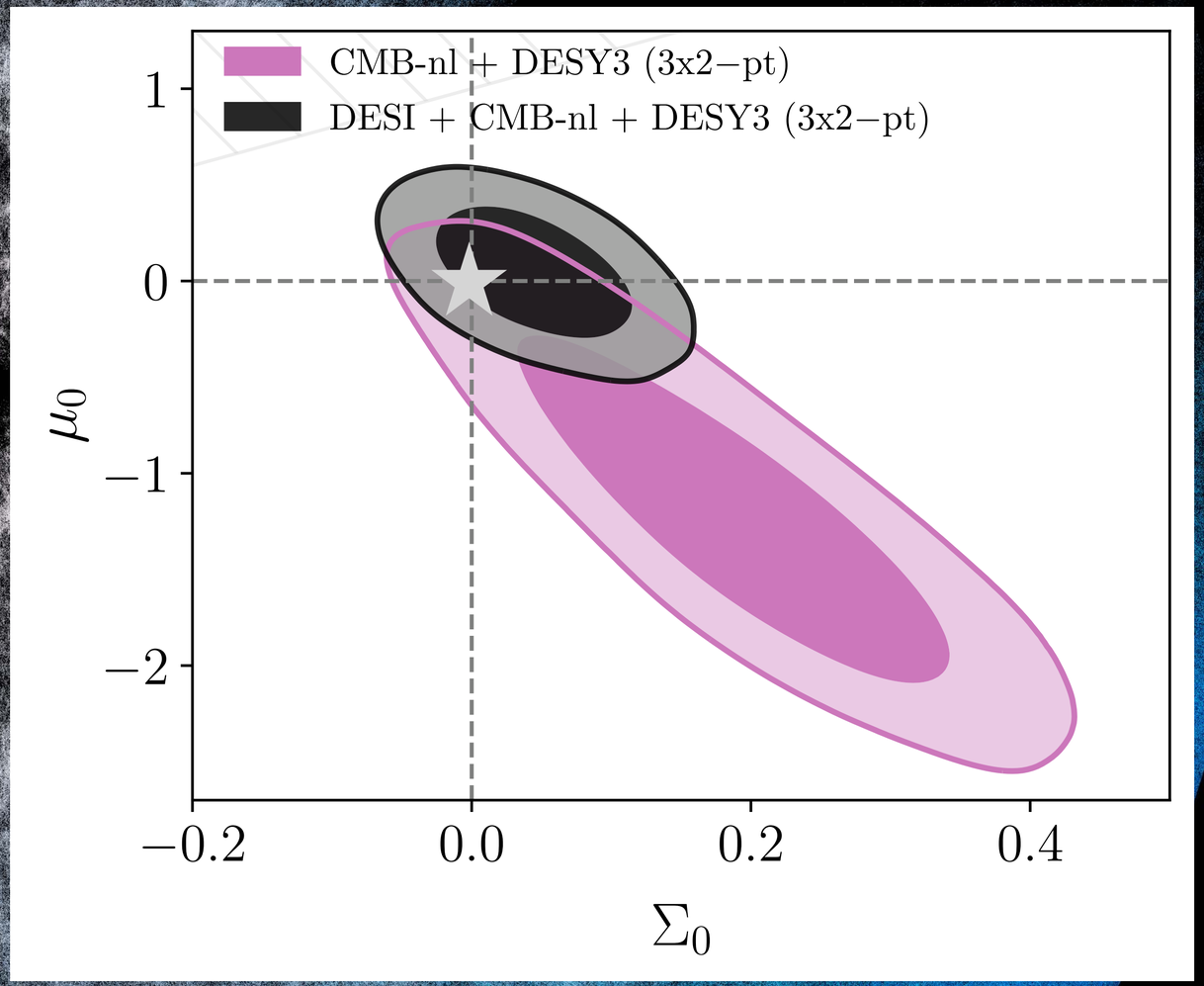


Same conclusion at higher significance in DESI DR2 Results II, PRD 112, 083515

What I am generally excited about - Growth of Structure and Modified Gravity

No evidence for modified gravity beyond GR in (μ_0, Σ_0) extension



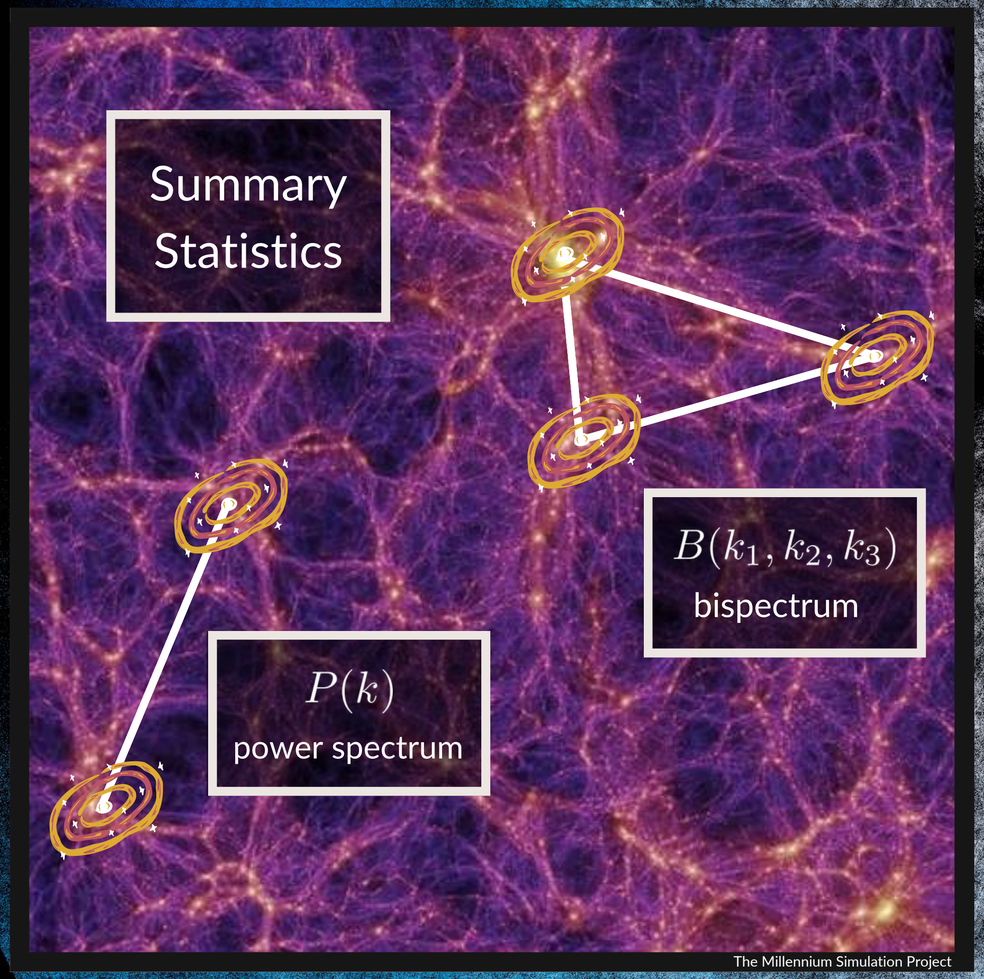


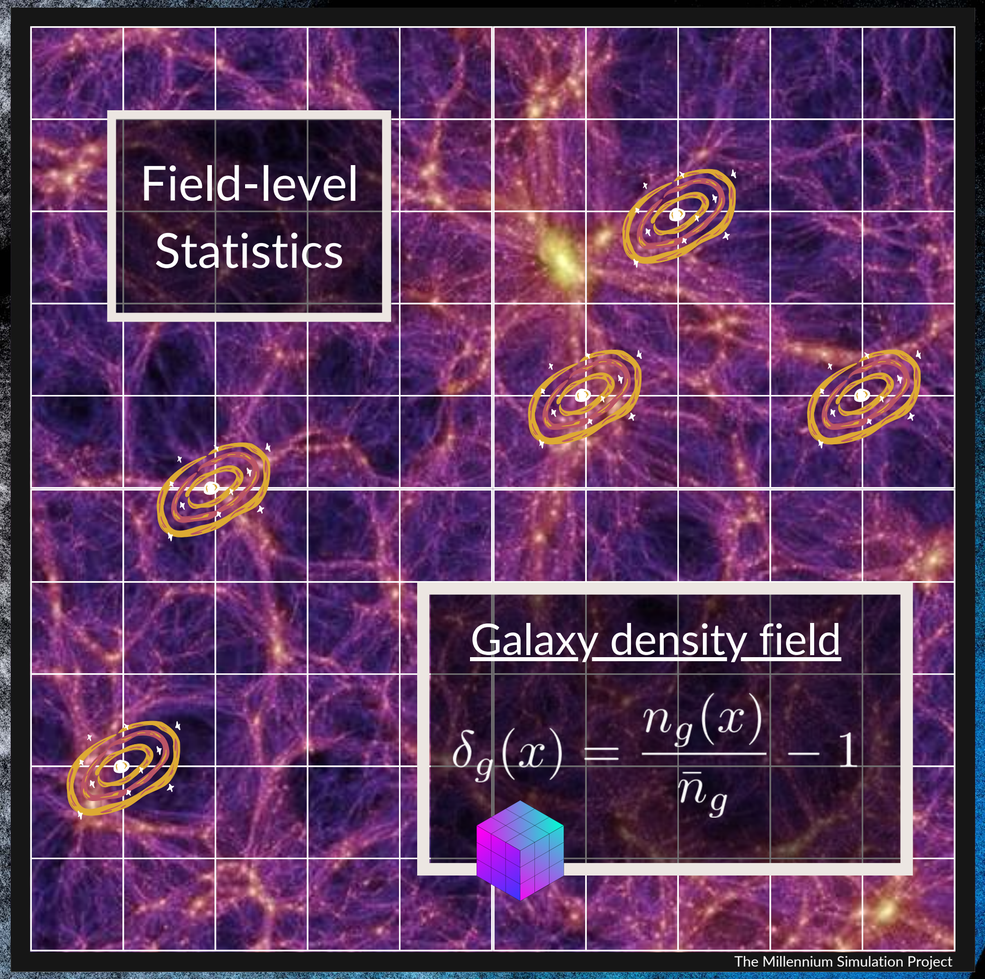
DESI 2024 VII, JCAP07(2025) incl. MN

Credit: Dark Energy Spectroscopic Instrument

What I am generally excited about - Field-level inference

How much information can be extracted from galaxy distribution?

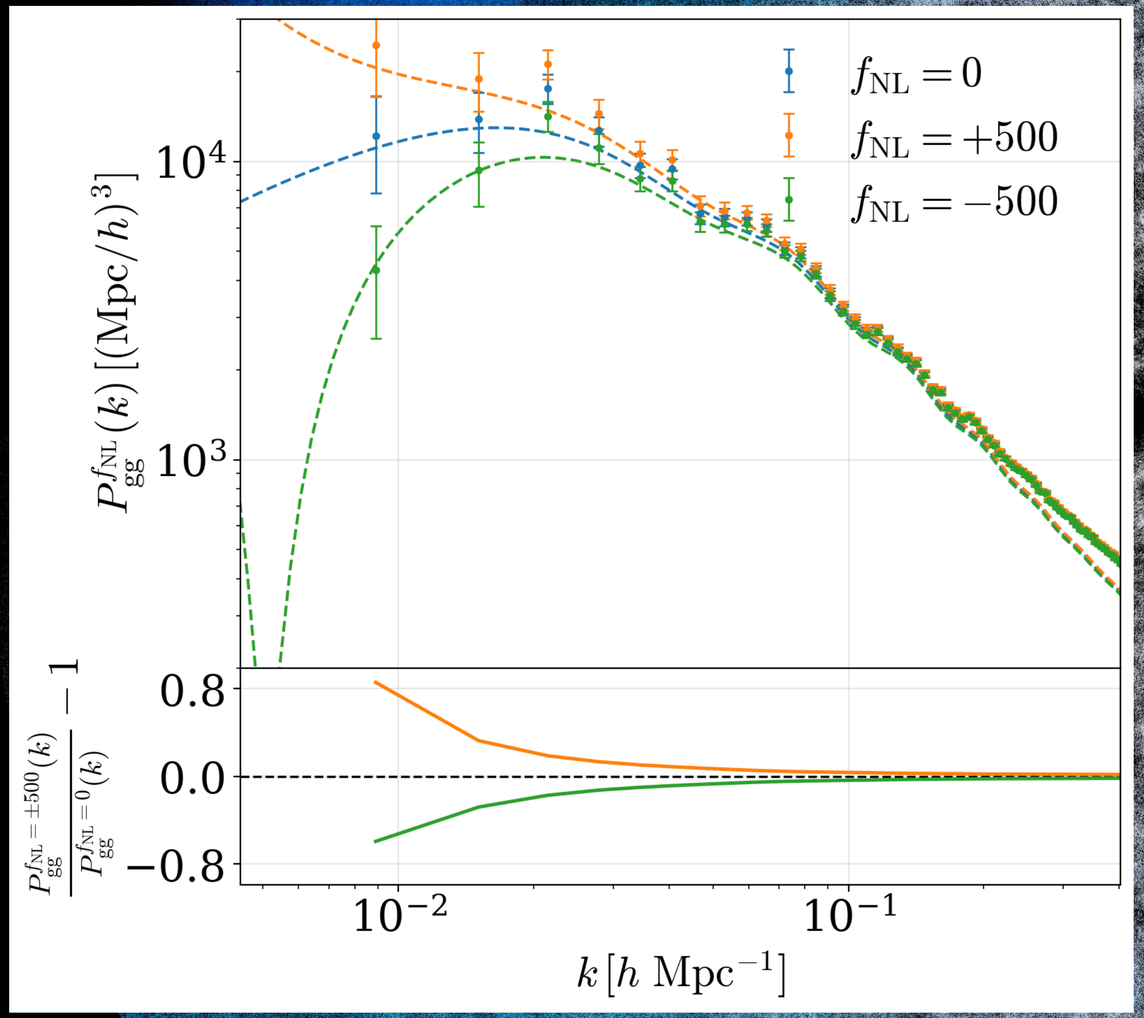


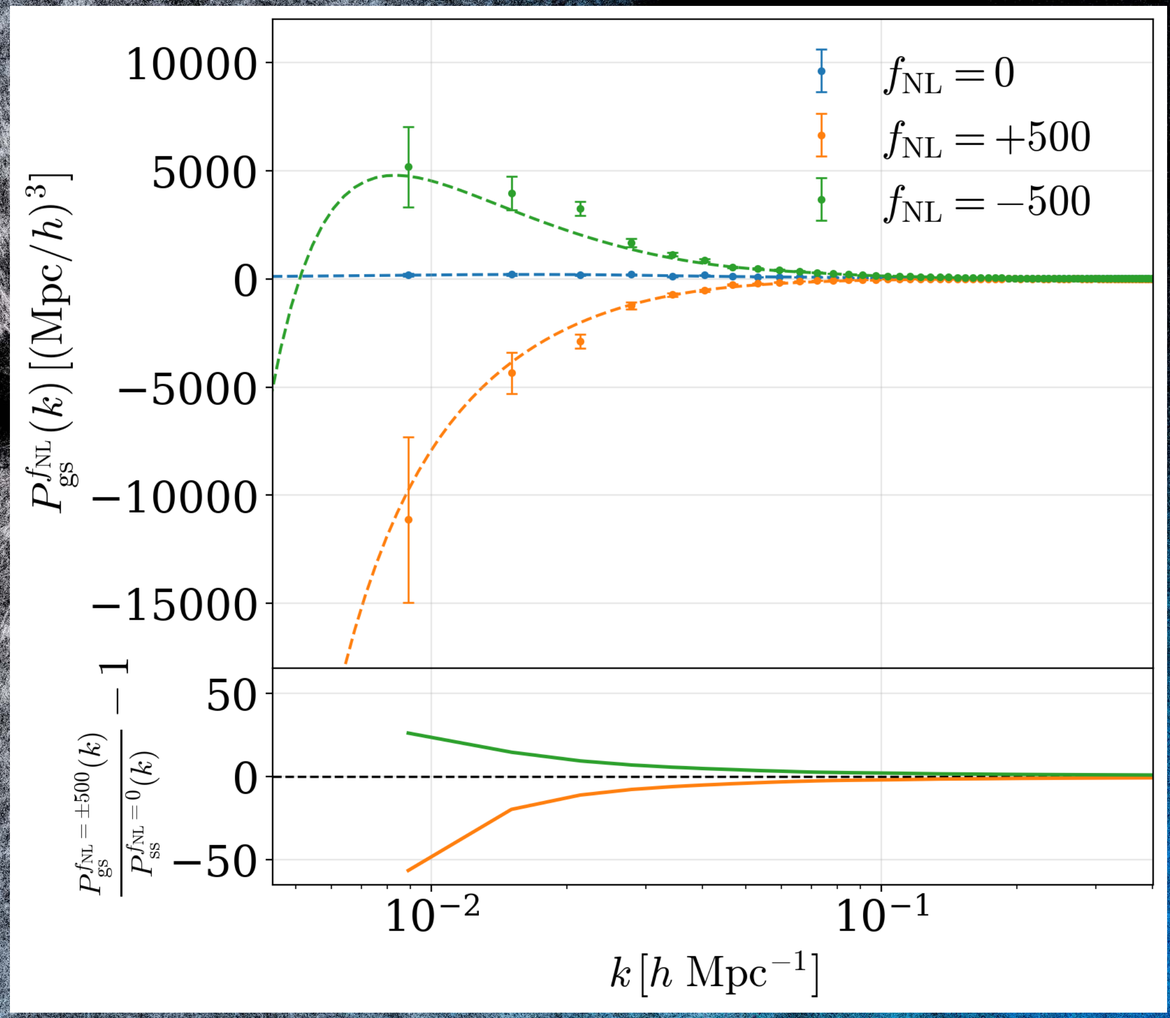


MN+2024 PRL 133, 221006 Beyond-2pt Collaboration, ApJ 990 99 incl. MN Credit: Dark Energy Spectroscopic Instrument

What I am recently excited about - tiny Primordial non-Gaussianity in the early universe

The galaxy distribution on large scales is modulated by local PNG

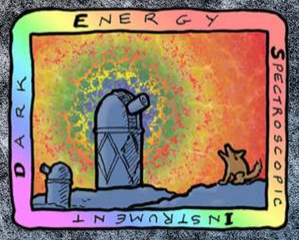




MN, Akitsu, Taruya 2026, PRD

Credit: Dark Energy Spectroscopic Instrument

The current and next era of galaxy surveys



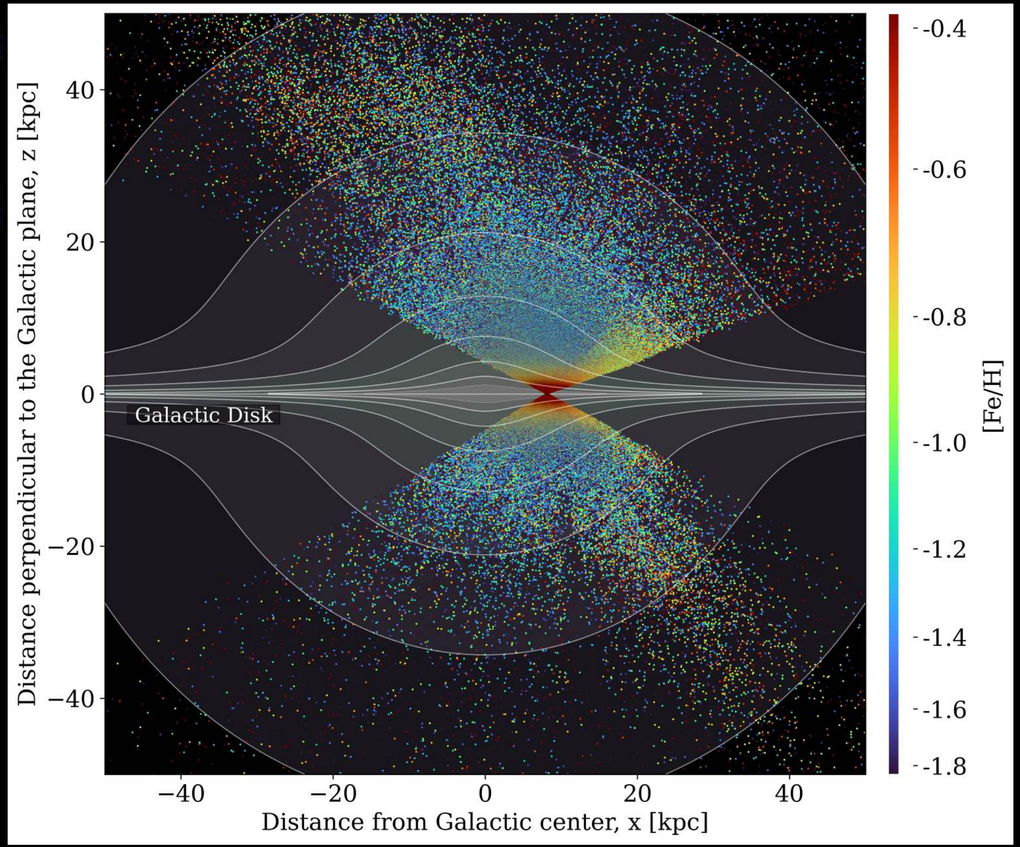
DARK ENERGY SPECTROSCOPIC INSTRUMENT

U.S. Department of Energy Office of Science

UNIVERSITY OF MICHIGAN

Credit: Dark Energy Spectroscopic Instrument

~15M galaxies in DESI DR2



DESI DR1

Credit: DESI/NOIRLab

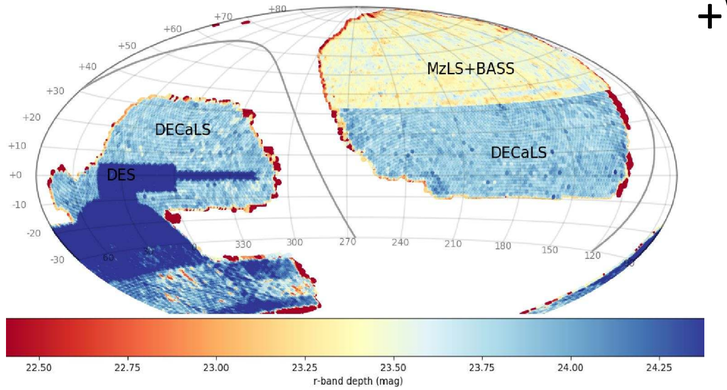
Wide-field imaging survey(s) —> target selection —> spectroscopic survey

DARK ENERGY SPECTROSCOPIC INSTRUMENT

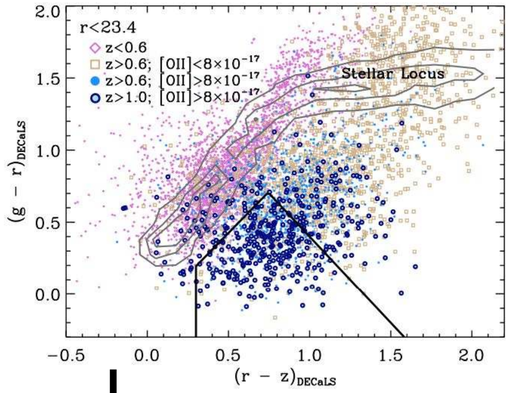
U.S. Department of Energy Office of Science

Imaging Surveys (2014-2019)

+WISE IR sat



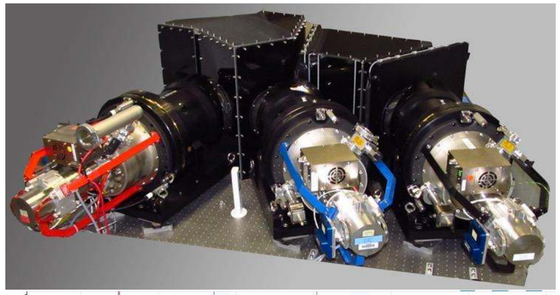
Target Selection

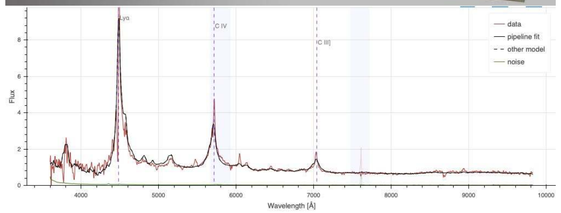


Observation…

…of 5000 objects every ~20mins…

…and measure their redshift

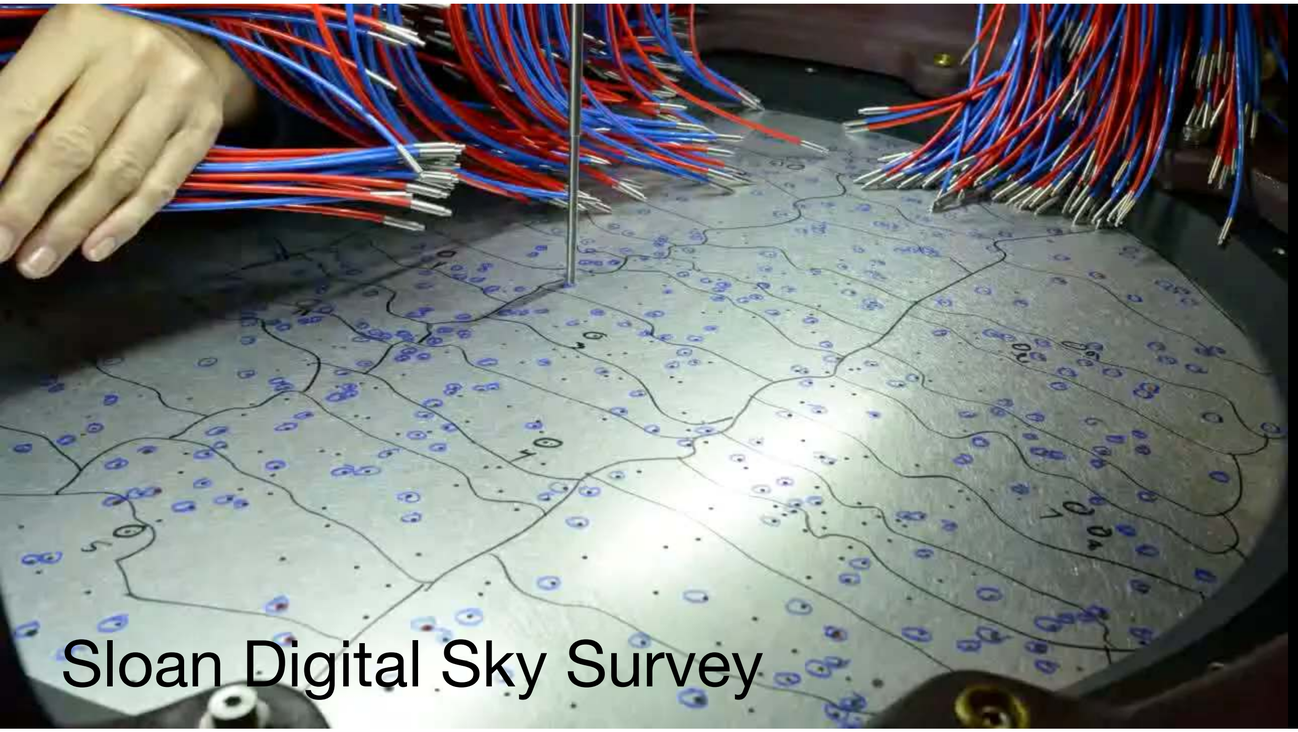




Advances in multi-object spectrograph really push the limit LSS survey number density

Plug plate

Robotic positioners

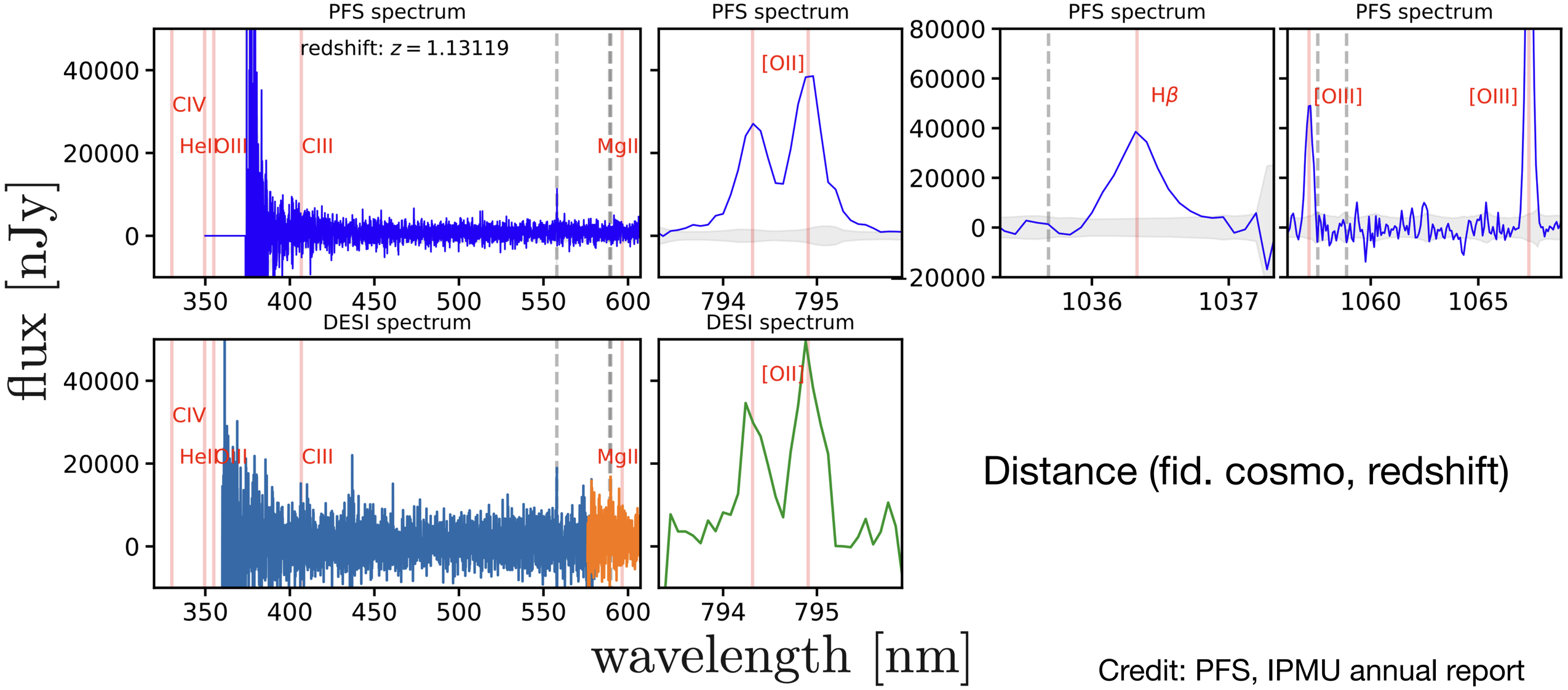


Sloan Digital Sky Survey

Dark Energy Spectroscopic Instrument

5000 objects in every ~20 minutes

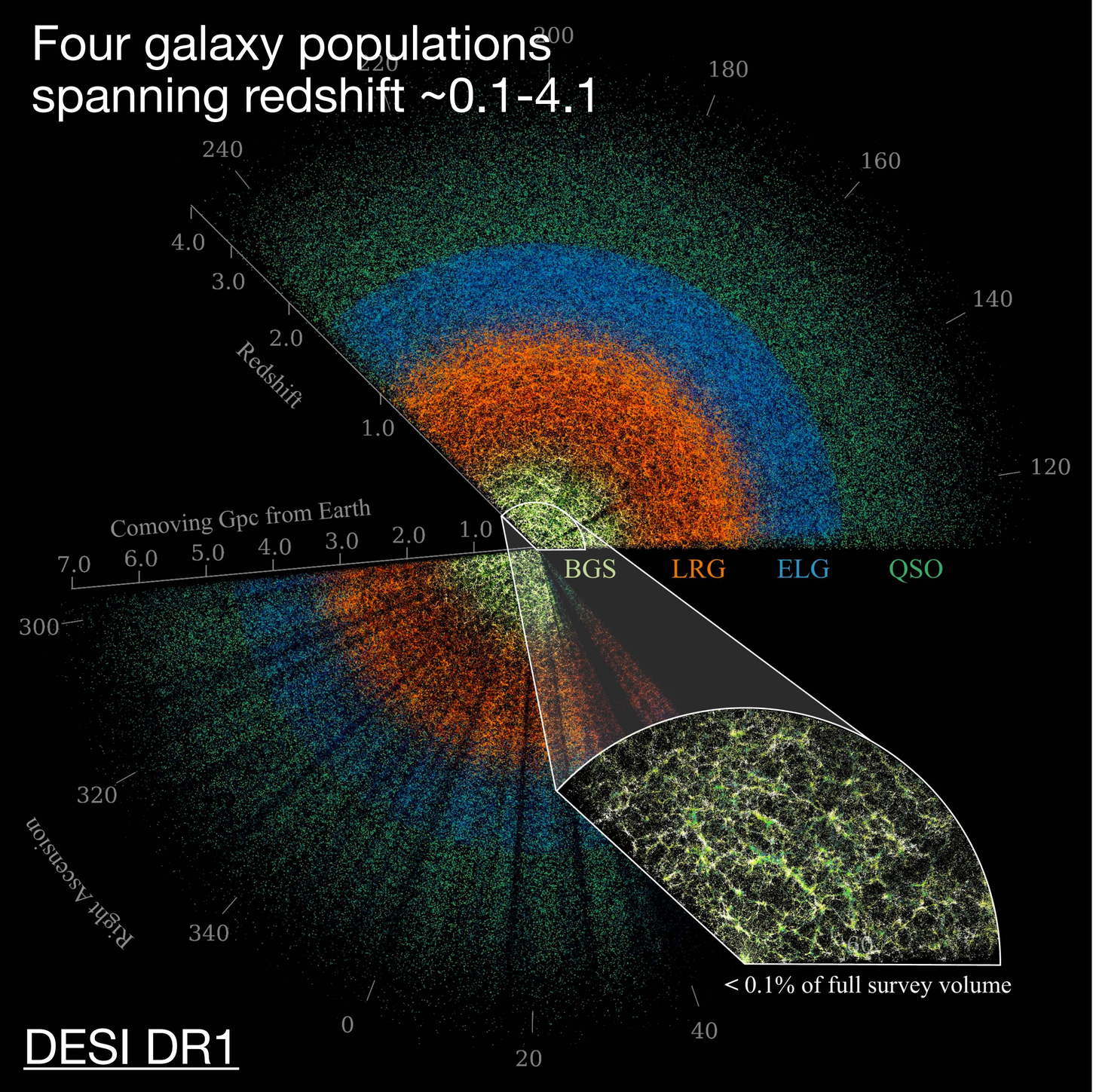
Assuming a fiducial cosmology, spectral redshifts can be converted to distances



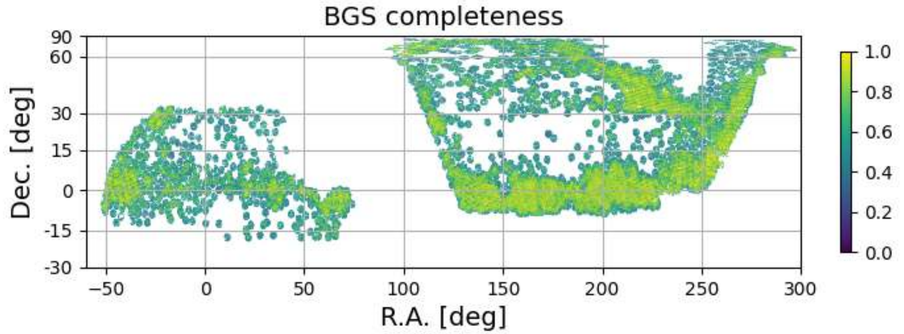
Distance (fid. cosmo, redshift)

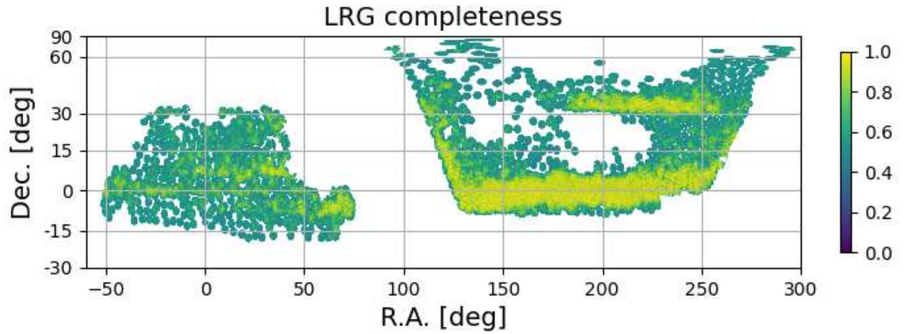
Credit: PFS, IPMU annual report

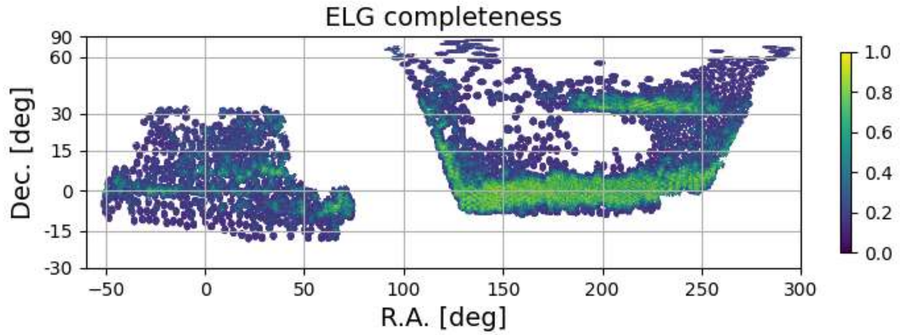
Four galaxy populations spanning redshift ~0.1-4.1

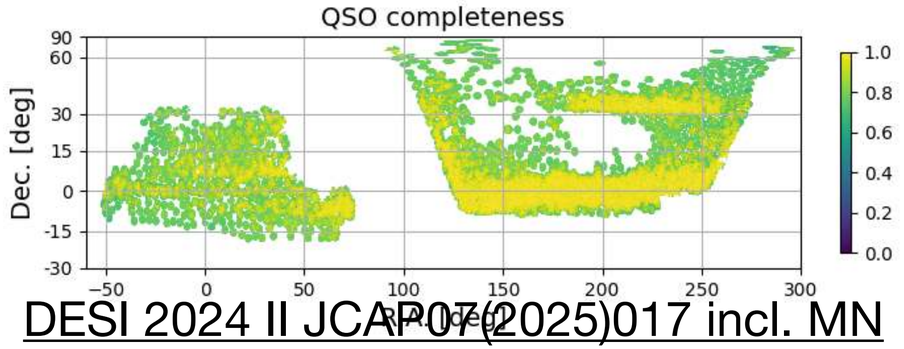


DESI DR1









DESI 2024 II JCAP07(2025)017 incl. MN

The current analysis of galaxy surveys

Two-point statistics of galaxy clustering

DARK ENERGY SPECTROSCOPIC INSTRUMENT

U.S. Department of Energy Office of Science

UNIVERSITY OF MICHIGAN

Credit: Dark Energy Spectroscopic Instrument

(Hierarchical) Bayesian inference and two-point statistics of galaxy maps

$$P(\theta \mid d) = \frac{P(d \mid \theta)\, P(\theta)}{P(d)} \propto P(d \mid \theta)\, P(\theta)$$

What is the cosmology θ given the data d?

(Hierarchical) Bayesian inference and two-point statistics of galaxy maps

$$P(\theta \mid d) = \frac{P(d \mid \theta)\, P(\theta)}{P(d)} \propto P(d \mid \theta)\, P(\theta)$$

What is the cosmology θ given the data d?

$$n_g(\mathbf{x}, \tau) \longrightarrow n_g(\mathbf{x}) \quad \longrightarrow \quad \delta_g(\mathbf{x}) = \frac{n_g(\mathbf{x}) - \bar{n}_g}{\bar{n}_g}$$

Full 3D galaxy density field

(Hierarchical) Bayesian inference and two-point statistics of galaxy maps

$$P(\theta \mid d) = \frac{P(d \mid \theta)\, P(\theta)}{P(d)} \propto P(d \mid \theta)\, P(\theta)$$

What is the cosmology θ given the data d?

$$n_g(\mathbf{x}, \tau) \longrightarrow n_g(\mathbf{x}) \quad \longrightarrow \quad \delta_g(\mathbf{x}) = \frac{n_g(\mathbf{x}) - \bar{n}_g}{\bar{n}_g}$$

Full 3D galaxy density field

$$P\Big(\theta \mid \delta_g(\mathbf{x})\Big) \propto P(\theta) \int \mathscr{D}\delta_{\rm ini}\, P\Big(\delta_g(\mathbf{x}) \mid \delta_{\rm ini}, \theta\Big)\, P(\delta_{\rm ini} \mid \theta)$$

Field-level inference analysis

(Hierarchical) Bayesian inference and two-point statistics of galaxy maps

$$P(\theta \mid d) = \frac{P(d \mid \theta)\, P(\theta)}{P(d)} \propto P(d \mid \theta)\, P(\theta)$$

What is the cosmology θ given the data d?

$$n_g(\mathbf{x}, \tau) \longrightarrow n_g(\mathbf{x}) \quad \longrightarrow \quad \delta_g(\mathbf{x}) = \frac{n_g(\mathbf{x}) - \bar{n}_g}{\bar{n}_g}$$

Full 3D galaxy density field

$$\xi_g(\mathbf{r}) \equiv \Big\langle \delta_g(\mathbf{x})\, \delta_g(\mathbf{x}+\mathbf{r}) \Big\rangle$$

Compressed two-point statistics

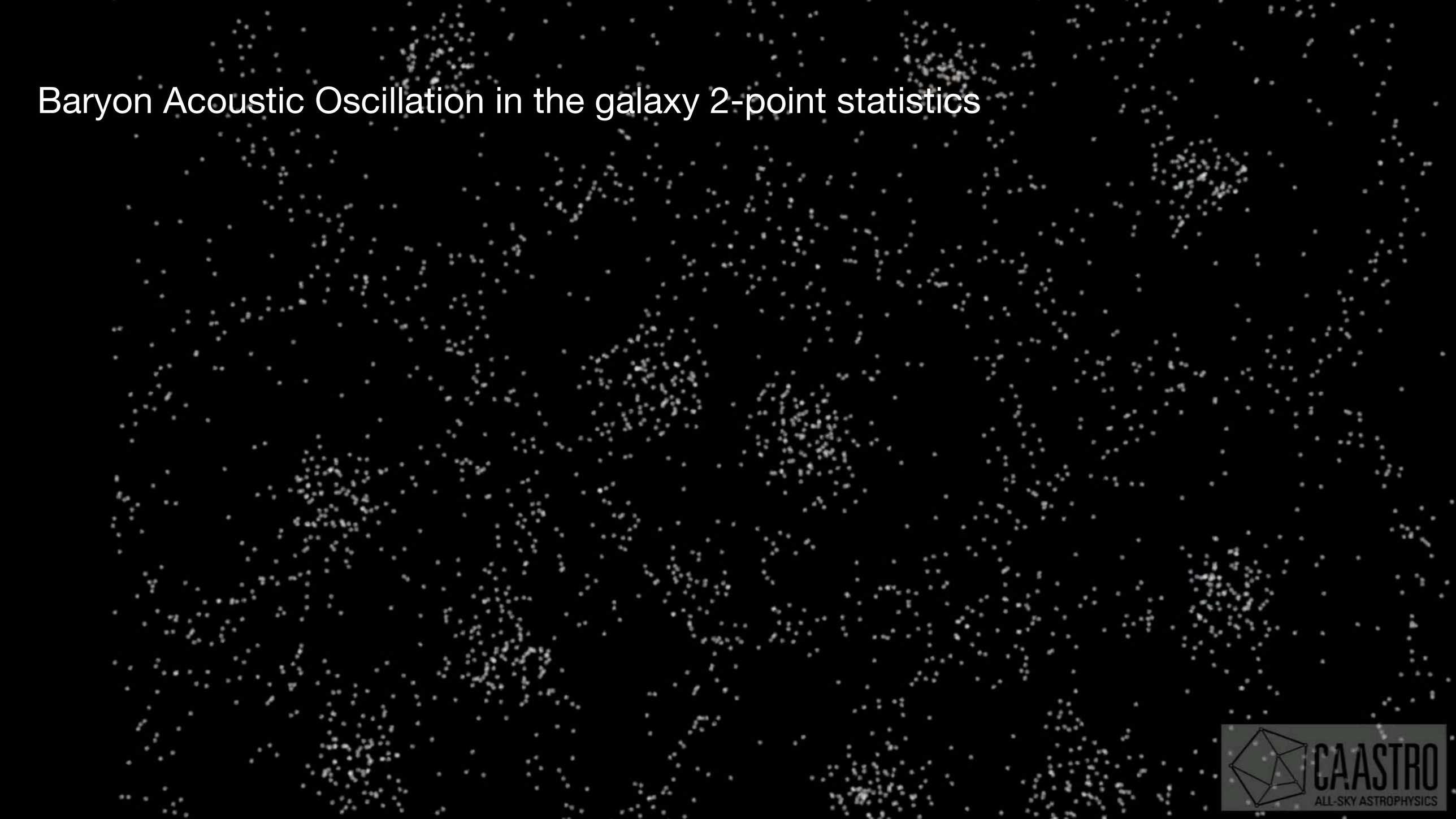
$$\xi_g(r) = \int_0^\infty \frac{k^2 dk}{2\pi^2} P_g(k) \frac{\sin kr}{kr}$$

$$P^{\rm th}_{\rm lin}(k;\theta) \equiv \Big\langle \, |\, \delta_{\rm ini}(\mathbf{k})\, |^2 \Big\rangle_\theta$$

$$P\Big(\theta \mid \widehat{P}_g(k)\Big) \propto P(\theta)\, P\Big(\widehat{P}_g(k) \mid P^{\rm th}_{\rm lin}(k;\theta), \theta\Big)$$

Power-spectrum analysis

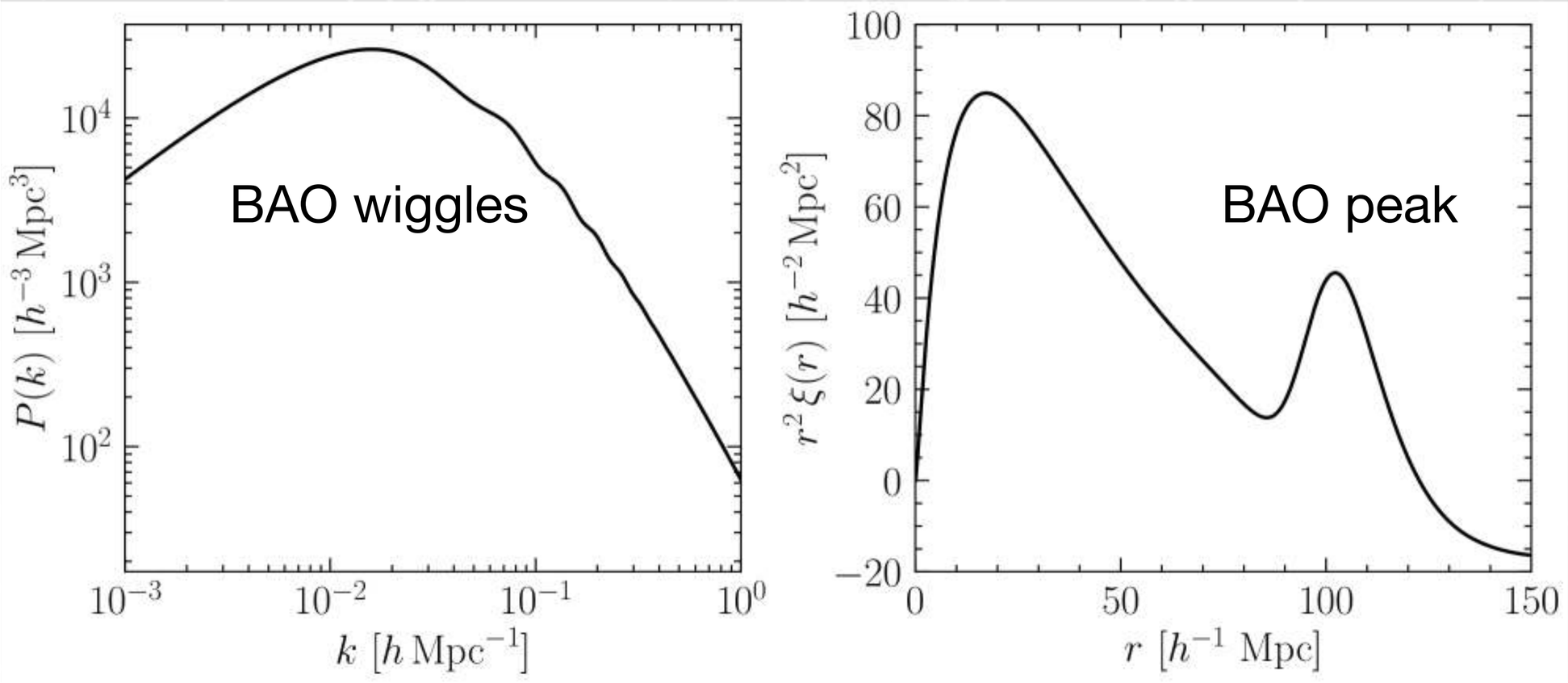
Baryon Acoustic Oscillation in the galaxy 2-point statistics



BAO imprints in the galaxy 2-point statistics

Fourier transform

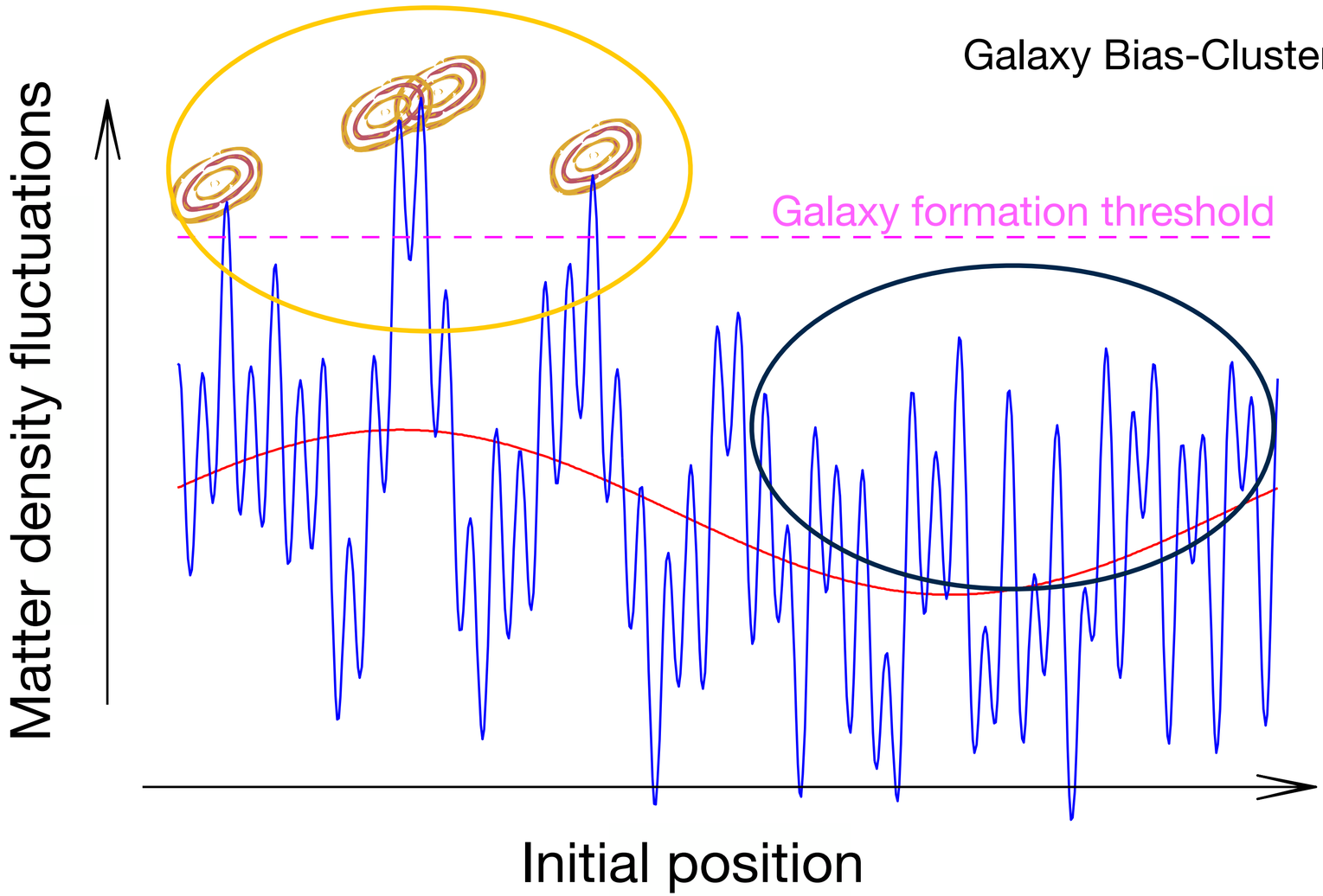
Power spectrum ⟷ 2-point correlation function

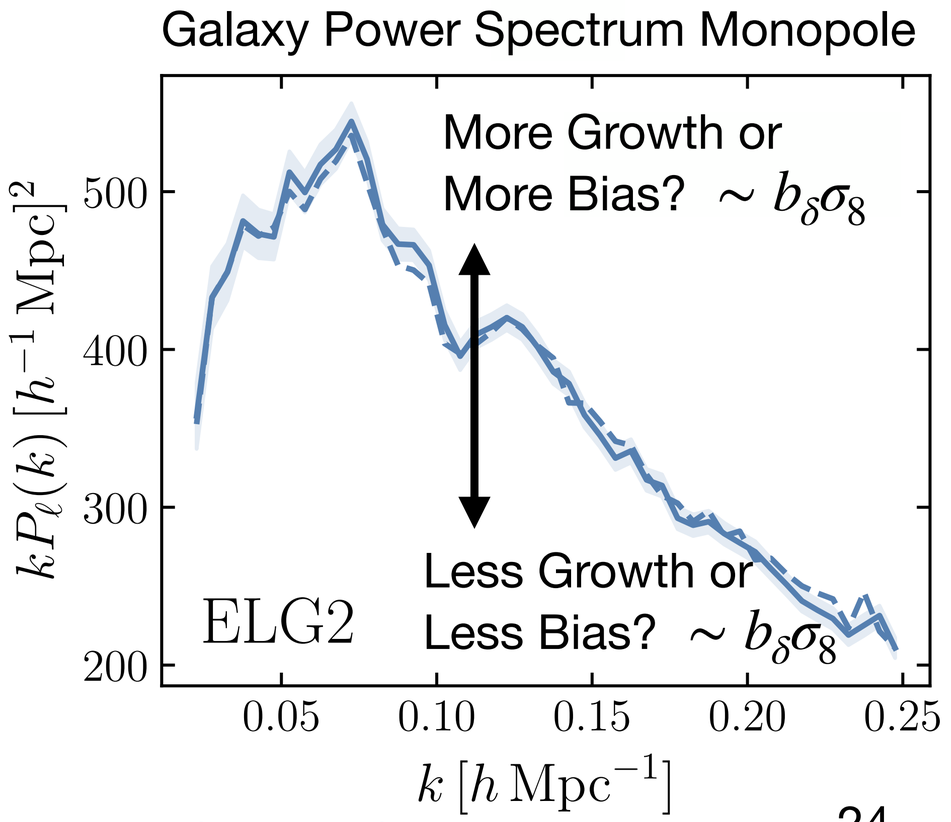


CAASTRO
ALL-SKY ASTROPHYSICS

First complication: galaxies are rare peaks in matter fluctuations hence biased tracers

Galaxy Bias-Clustering Degeneracy $P_g \propto (b_\delta \sigma_8)^2$

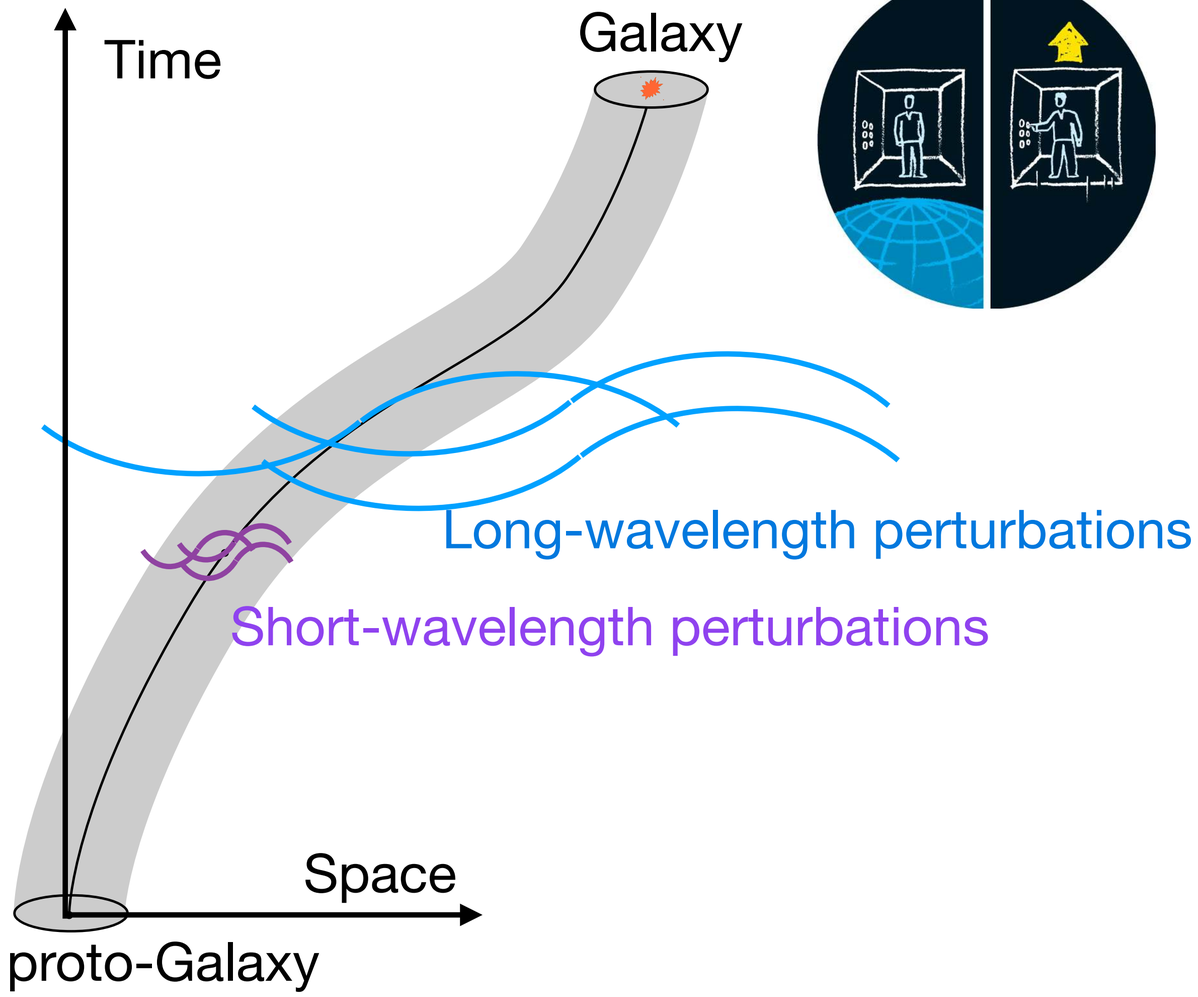




Adapted from Desjacques+ 2016

DESI 2024 V (incl. MN)

Still we can describe the galaxy-matter bias relation with Effective Field Theory of LSS



Equivalence principle

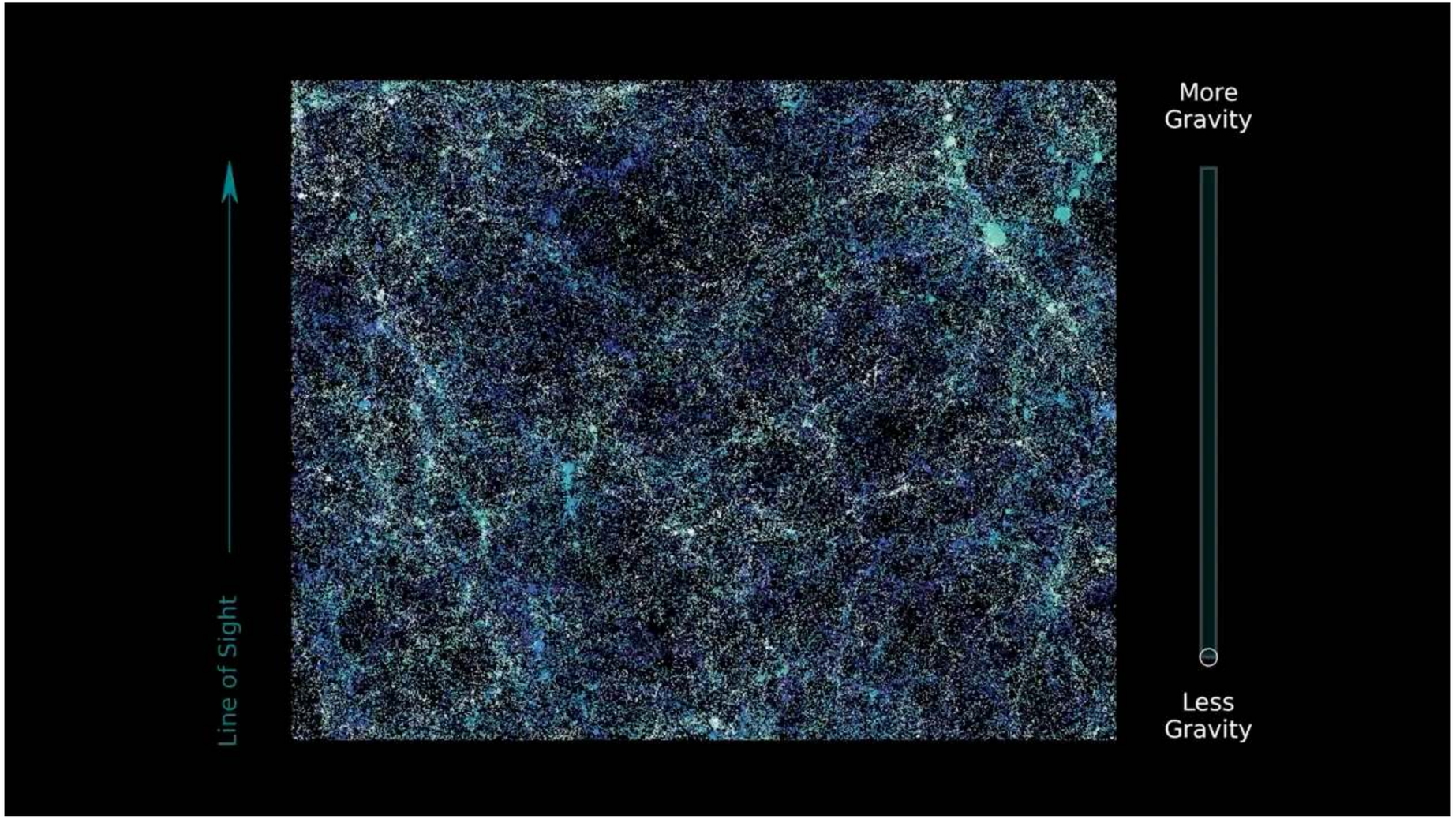
EFT galaxy bias expansion

$$\delta_g(\theta, \hat{s}) = \sum_O b_O O(\theta, \hat{s}) + \epsilon$$

O: bias expansion; ϵ: stochasticity

Adapted from Desjacques+ 2016

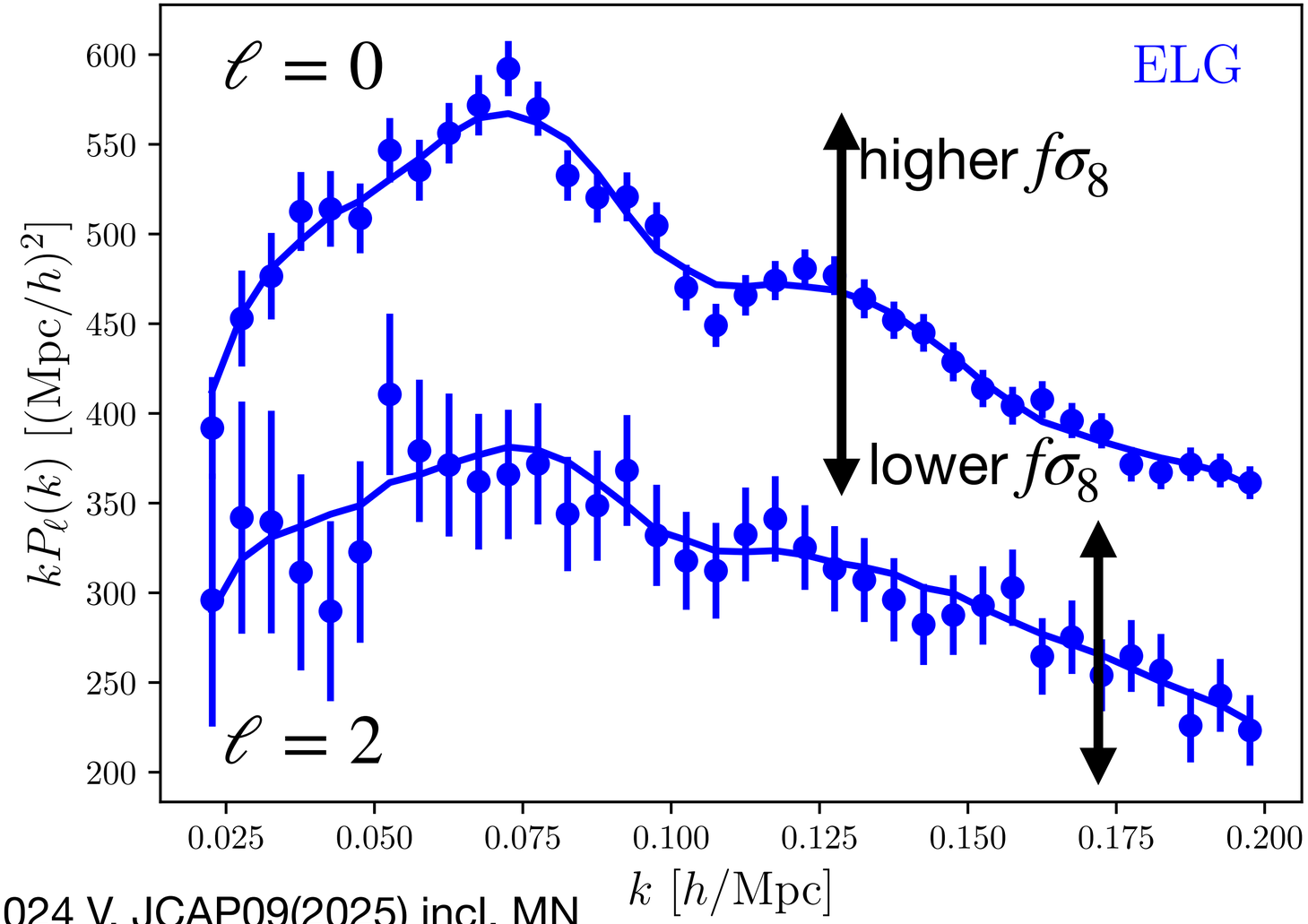
Second complication (or additional source of information): Redshift-space distortion



Credit: DESI, Claire Lamman and Michael Rashkovetskyi

Putting everything together, we can predict the power spectrum in redshift space

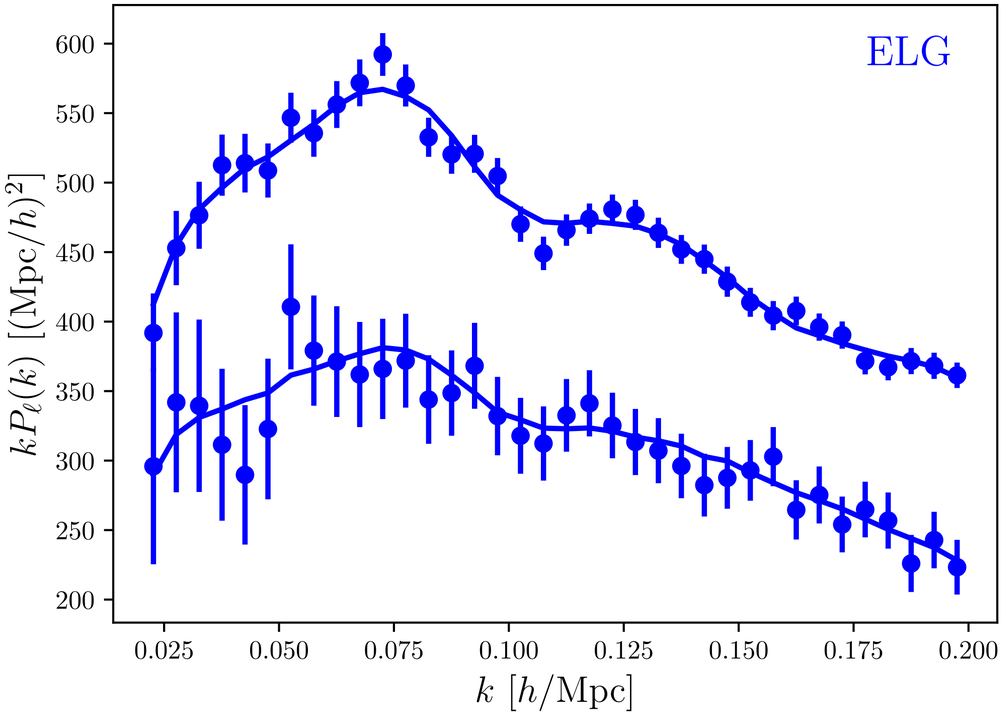
$$P^s_{gg}(k,\mu) = P_{\ell=0}(k)L_0(\mu) + P_{\ell=2}(k)L_2(\mu) + P_{\ell=4}(k)L_4(\mu) + \cdots$$

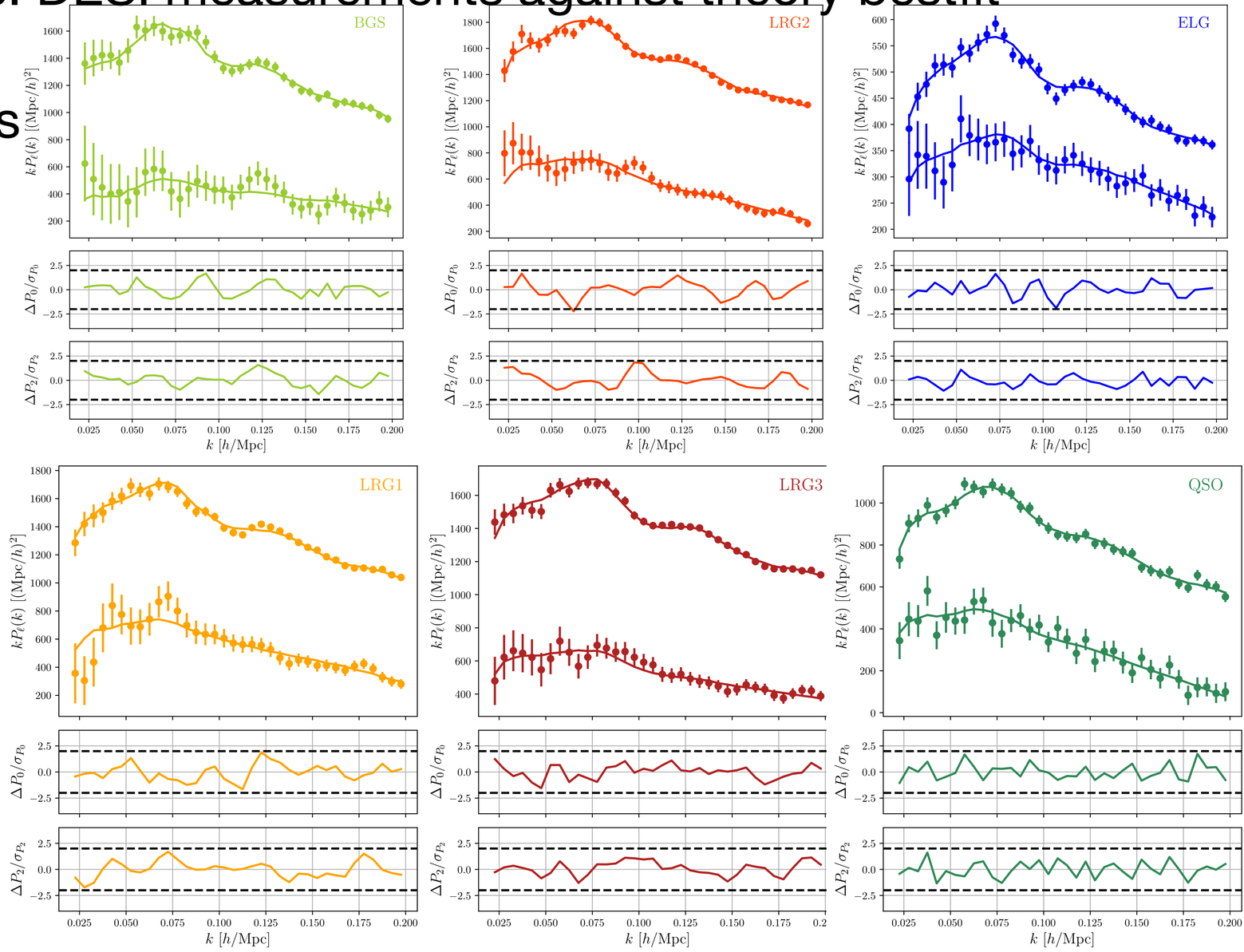


DESI 2024 V, JCAP09(2025) incl. MN

Power spectrum in redshift space: DESI measurements against theory bestfit

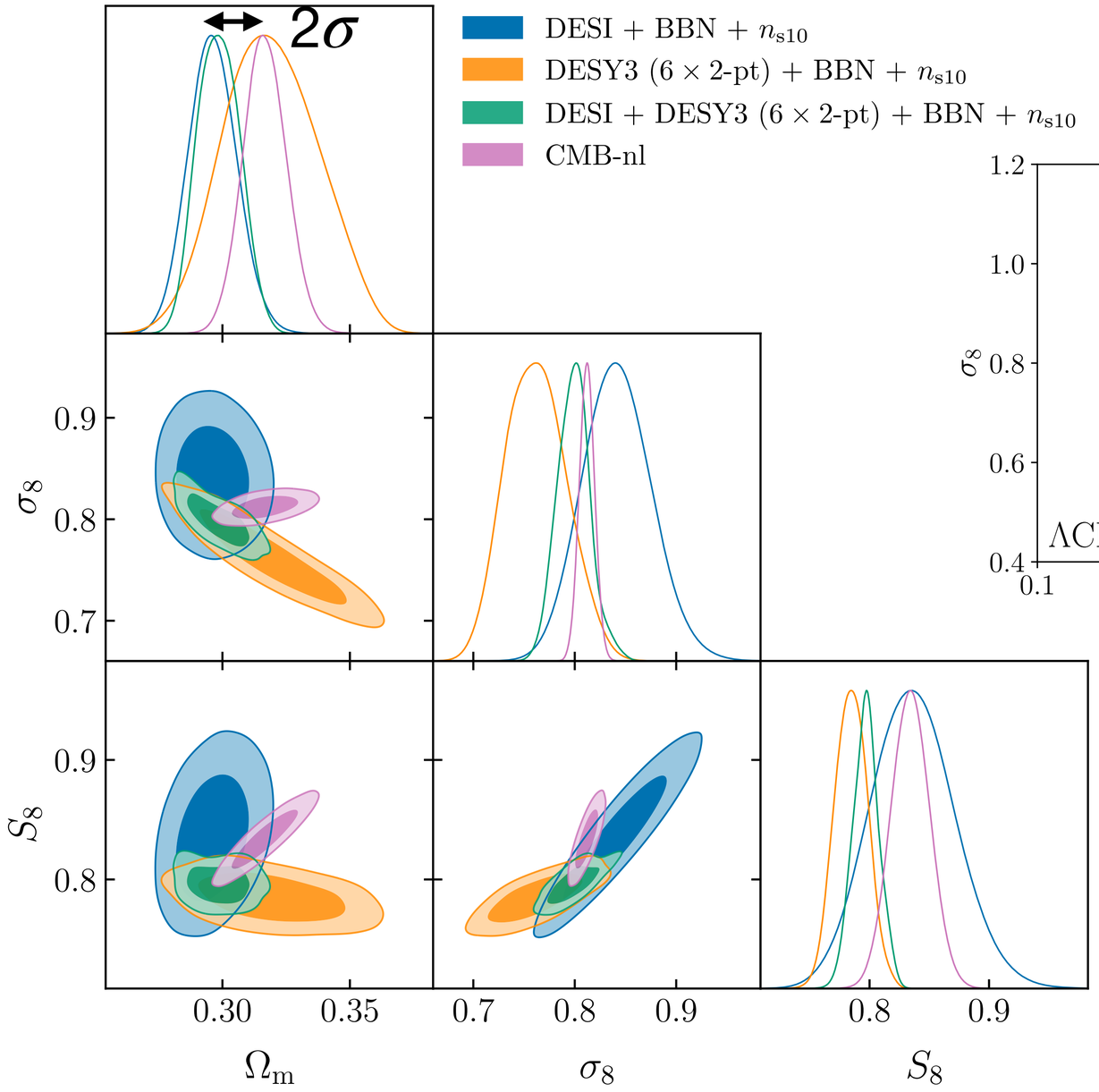
Galaxy power spectrum multipoles





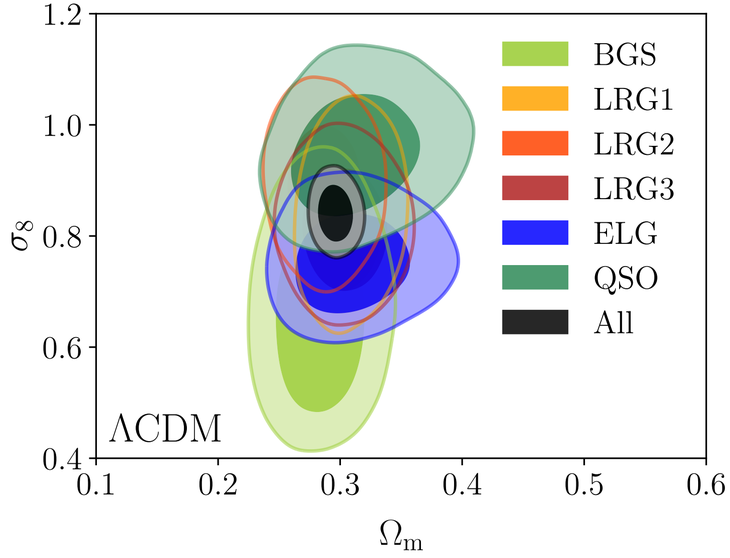
DESI 2024 V, JCAP09(2025) incl. MN

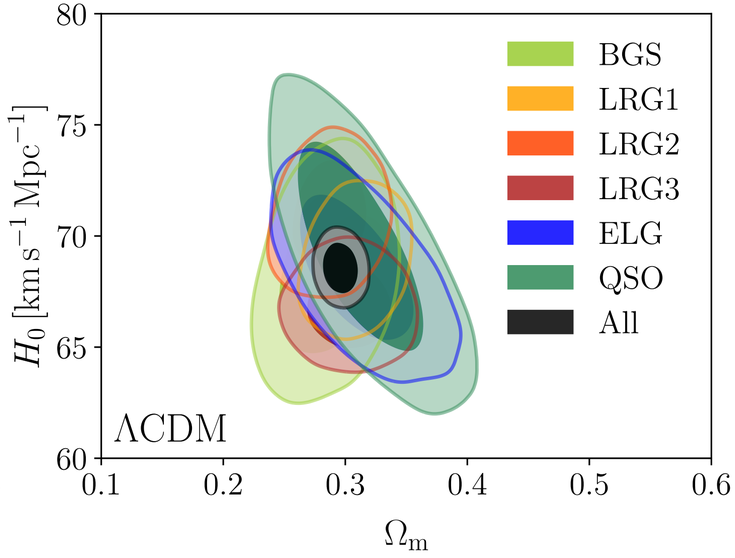
DESI DR1 results for the standard model ΛCDM



DESI 2024 V, JCAP09(2025) incl. MN

DESI 2024 VII, JCAP07(2025) incl. MN





DESI data alone is consistent with ΛCDM.

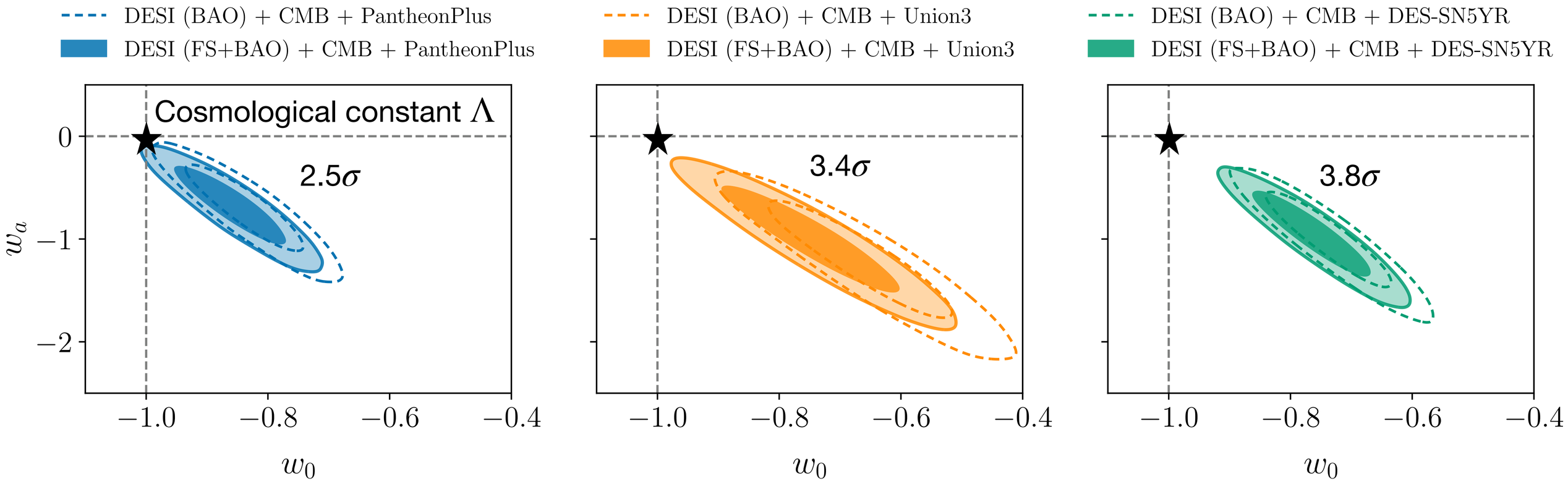
There is a mild ω_m tension with Planck CMB and SNe observations*.

*not shown, but please ask!

DESI DR1 results for beyond ΛCDM - dark energy

Dark energy appears to be evolving with time as $w = w_0 + w_a(1-a)$

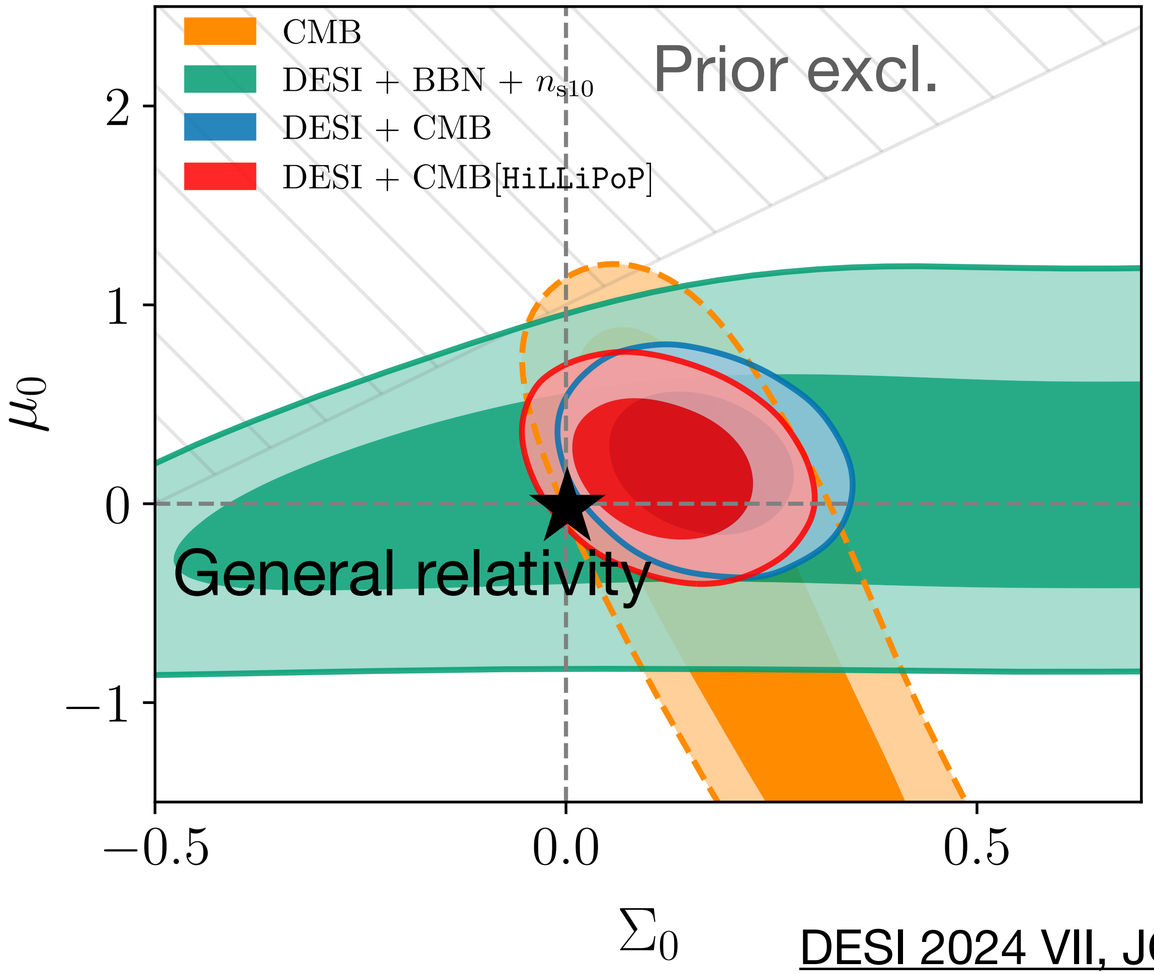
DESI 2024 VII, JCAP07(2025) incl. MN

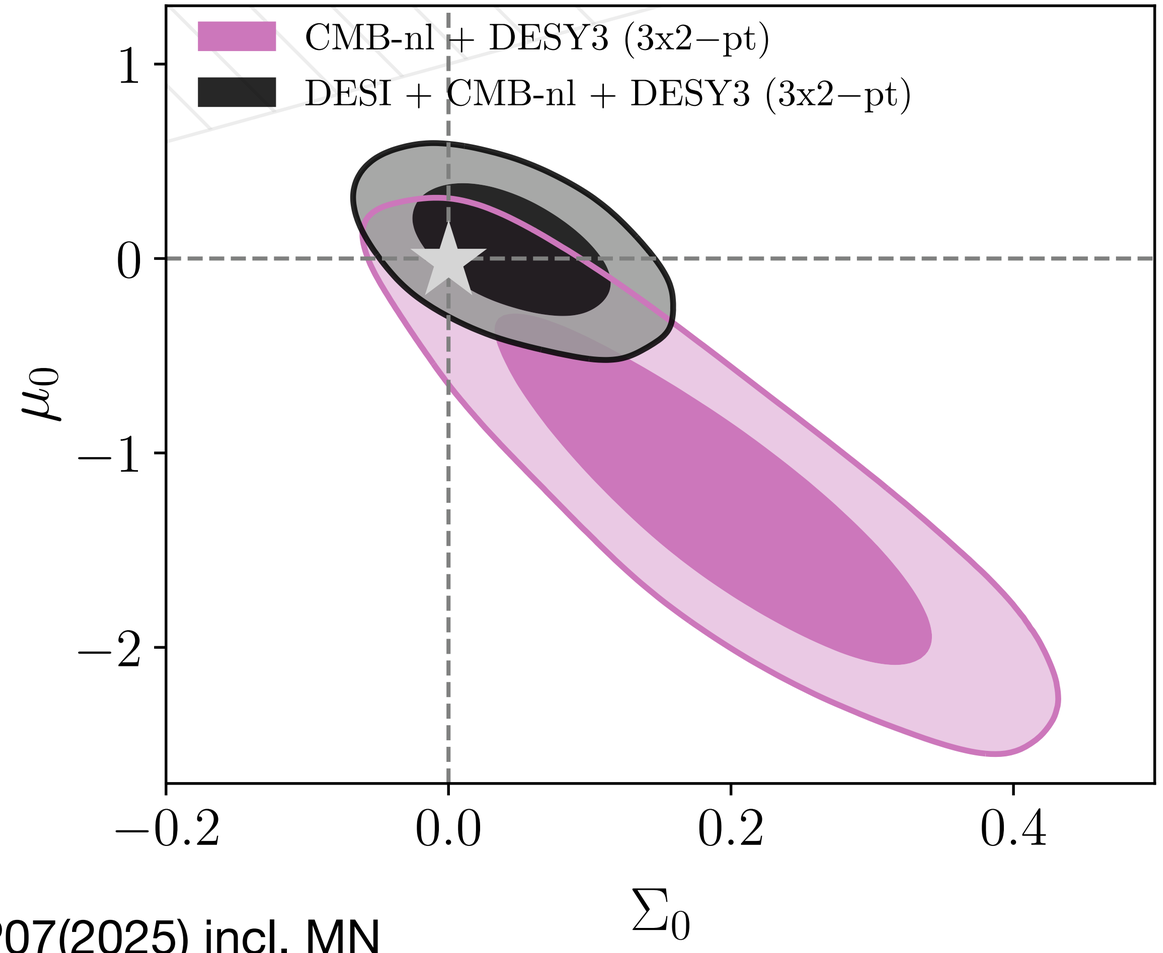


Same conclusion at higher significance in DESI DR2 Results II, PRD 112, 083515

DESI DR1 results for beyond ΛCDM - modified gravity

No evidence for modified gravity beyond GR in (μ_0, Σ_0) extension





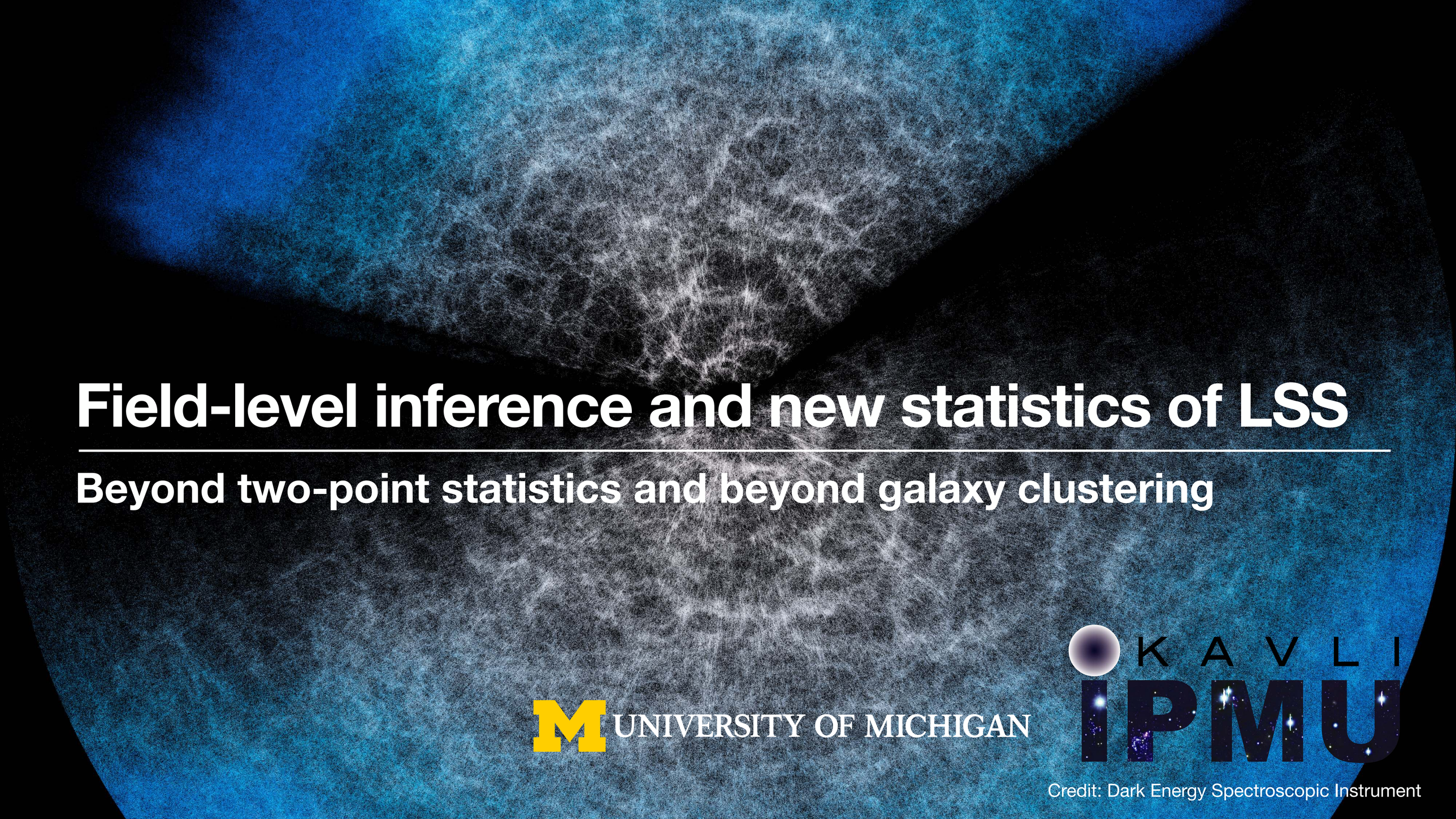
DESI 2024 VII, JCAP07(2025) incl. MN

Field-level inference and new statistics of LSS

Beyond two-point statistics and beyond galaxy clustering

UNIVERSITY OF MICHIGAN

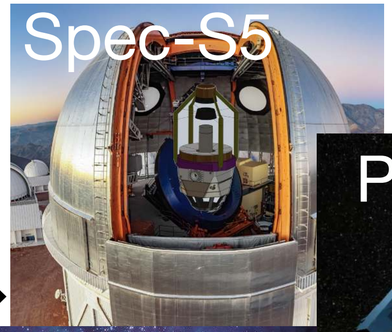
KAVLI IPMU

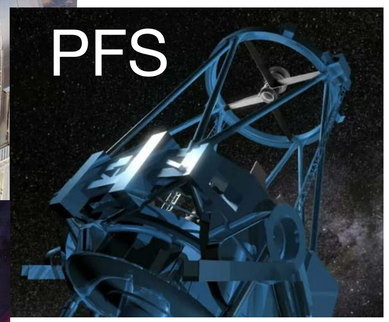


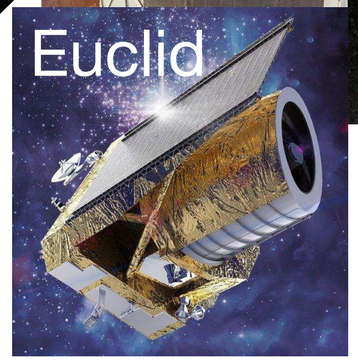
Credit: Dark Energy Spectroscopic Instrument

What else can we do without bigger telescopes and bigger sensors?

Diminishing returns

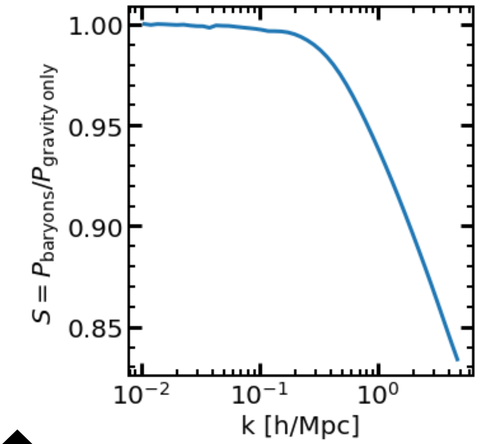


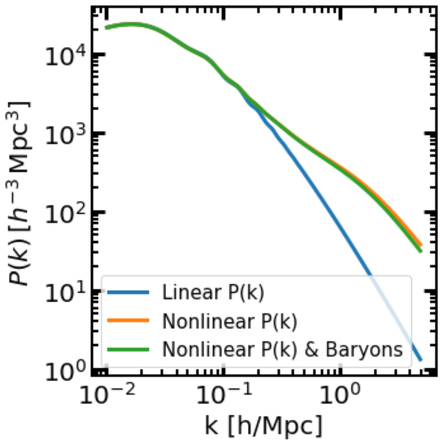




We are here

LSS survey volume



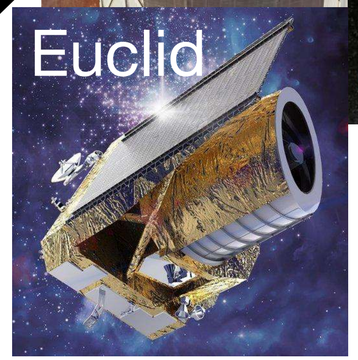


Credit: BACCOemu

LSS analysis cutoff scale

Astro complications

Diminishing returns

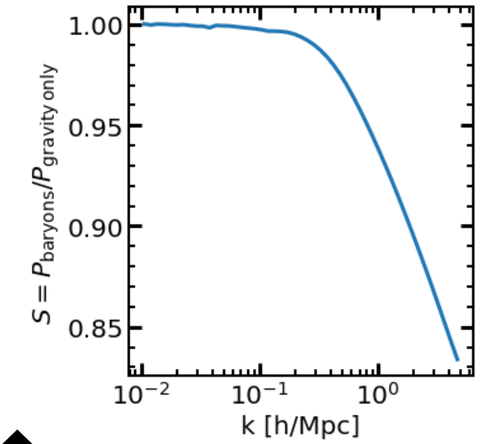


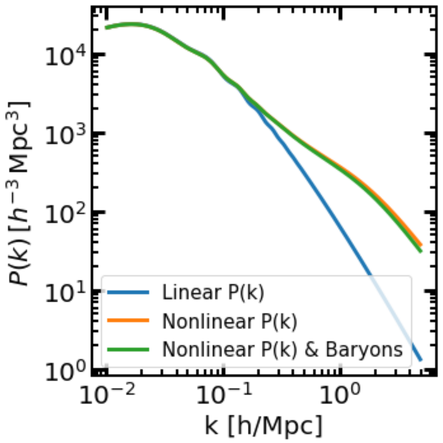
We are here

Under-explored dimension

LSS survey volume

LSS statistics





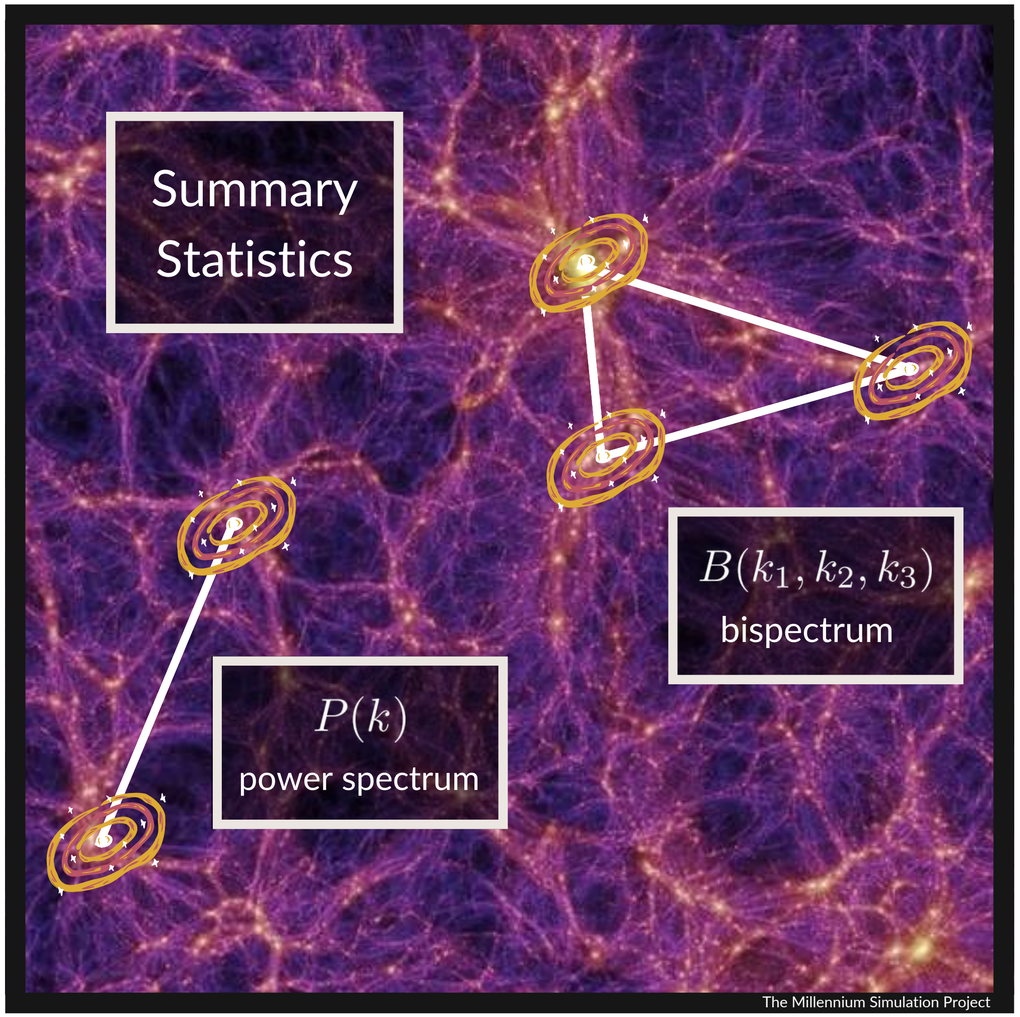
Credit: BACCOemu

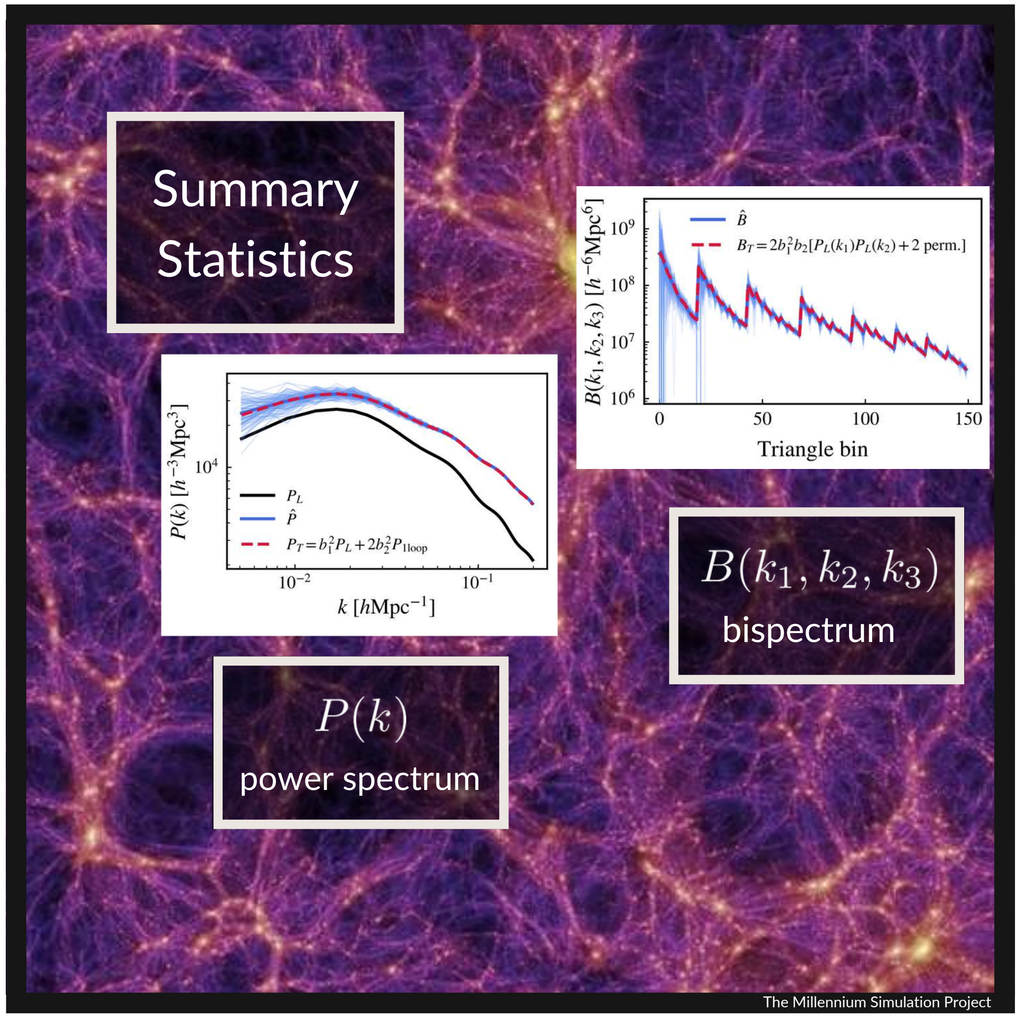
LSS analysis cutoff scale

Astro complications

Beyond-two-point statistics = Three-point statistics?

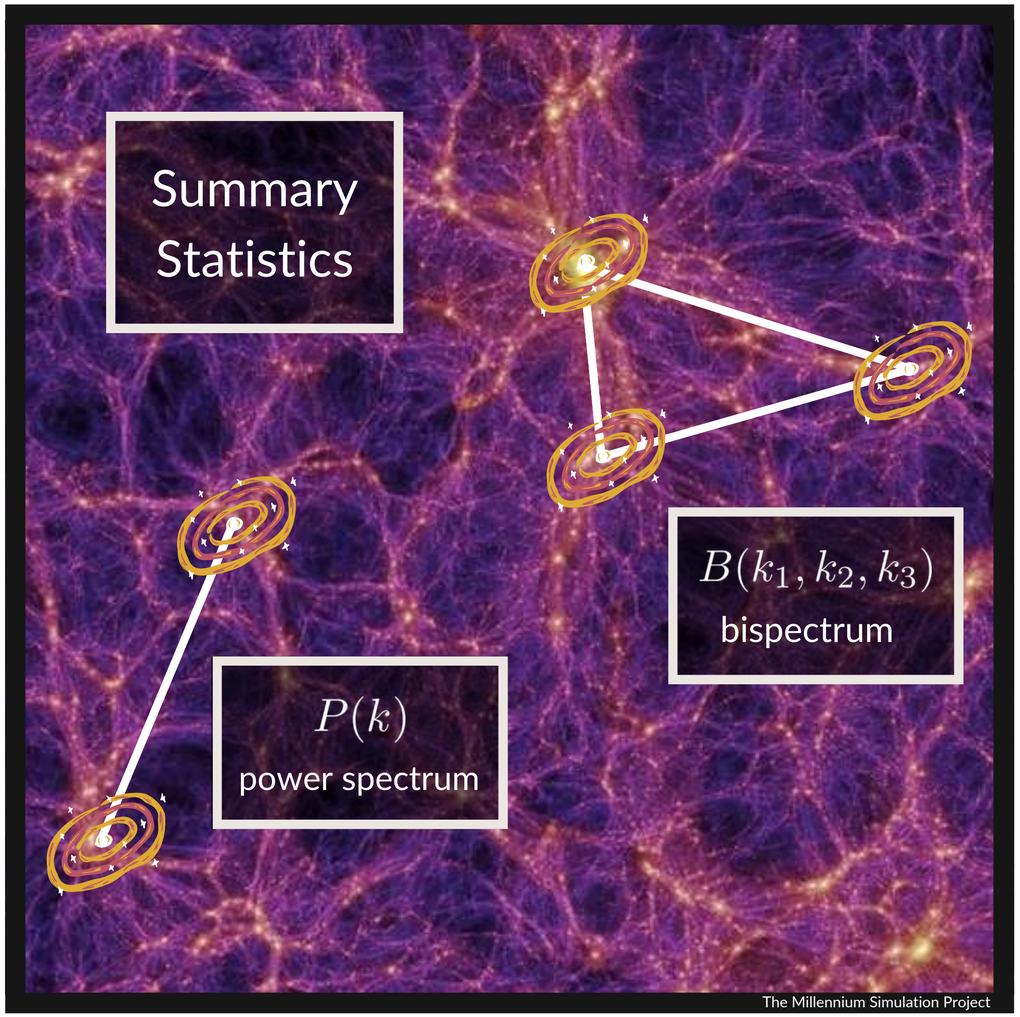
Bispectrum probes triplets/triangles of galaxies adding higher-order information

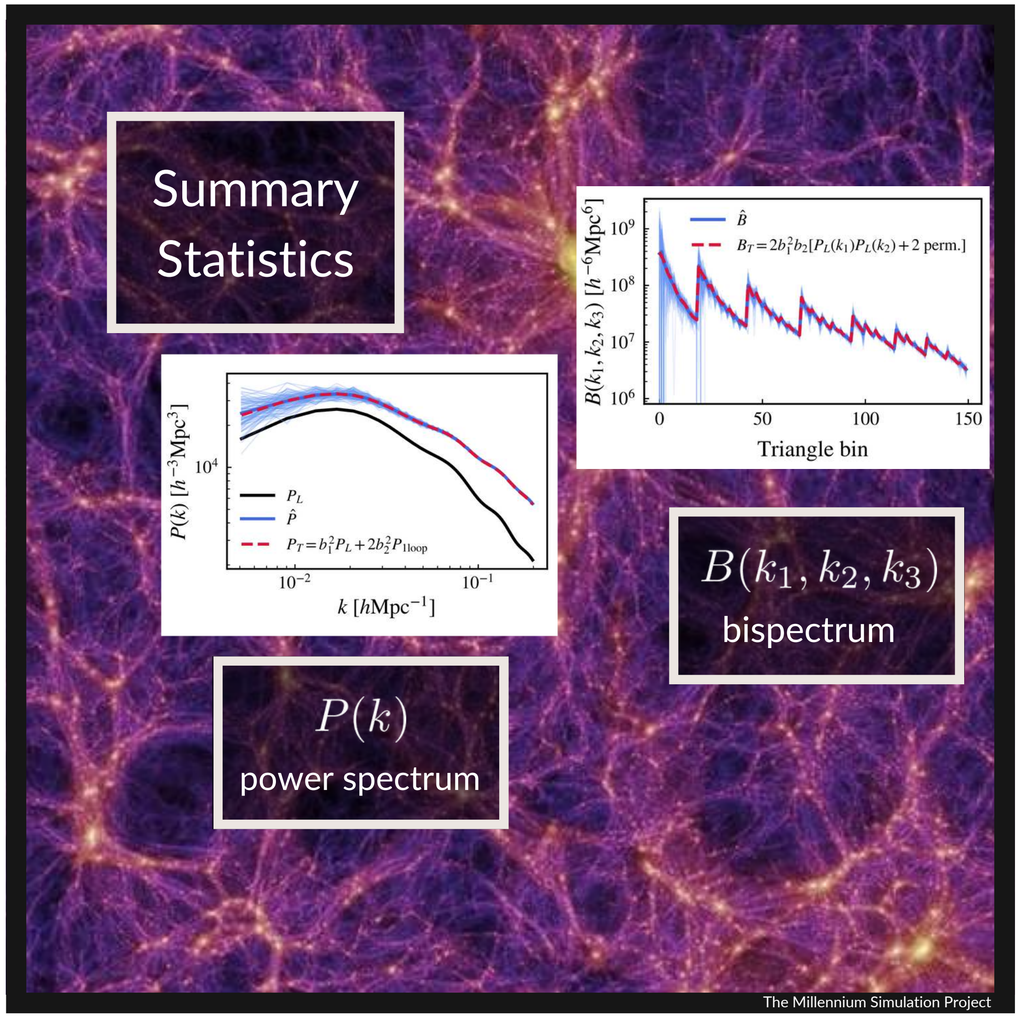




Beyond-two-point statistics = Three-point statistics?

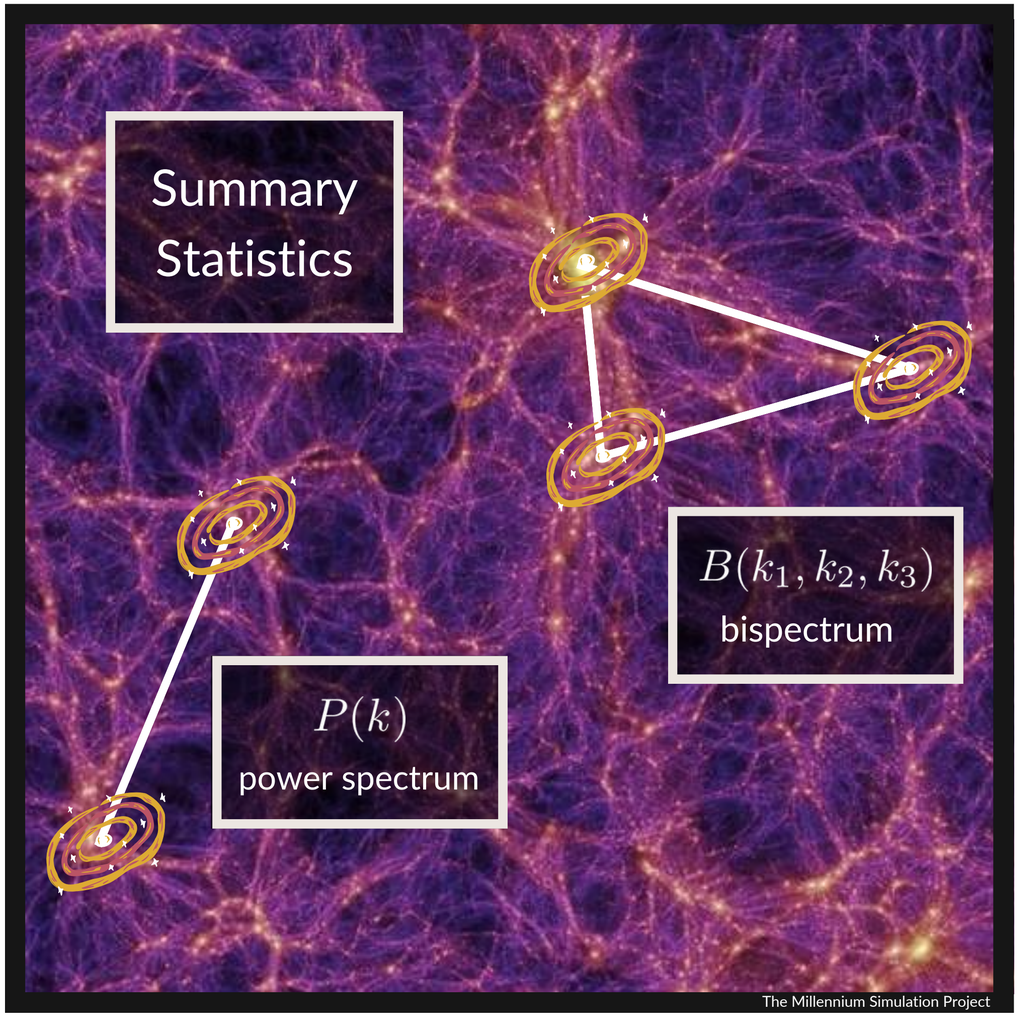
Bispectrum probes triplets/triangles of galaxies adding higher-order information

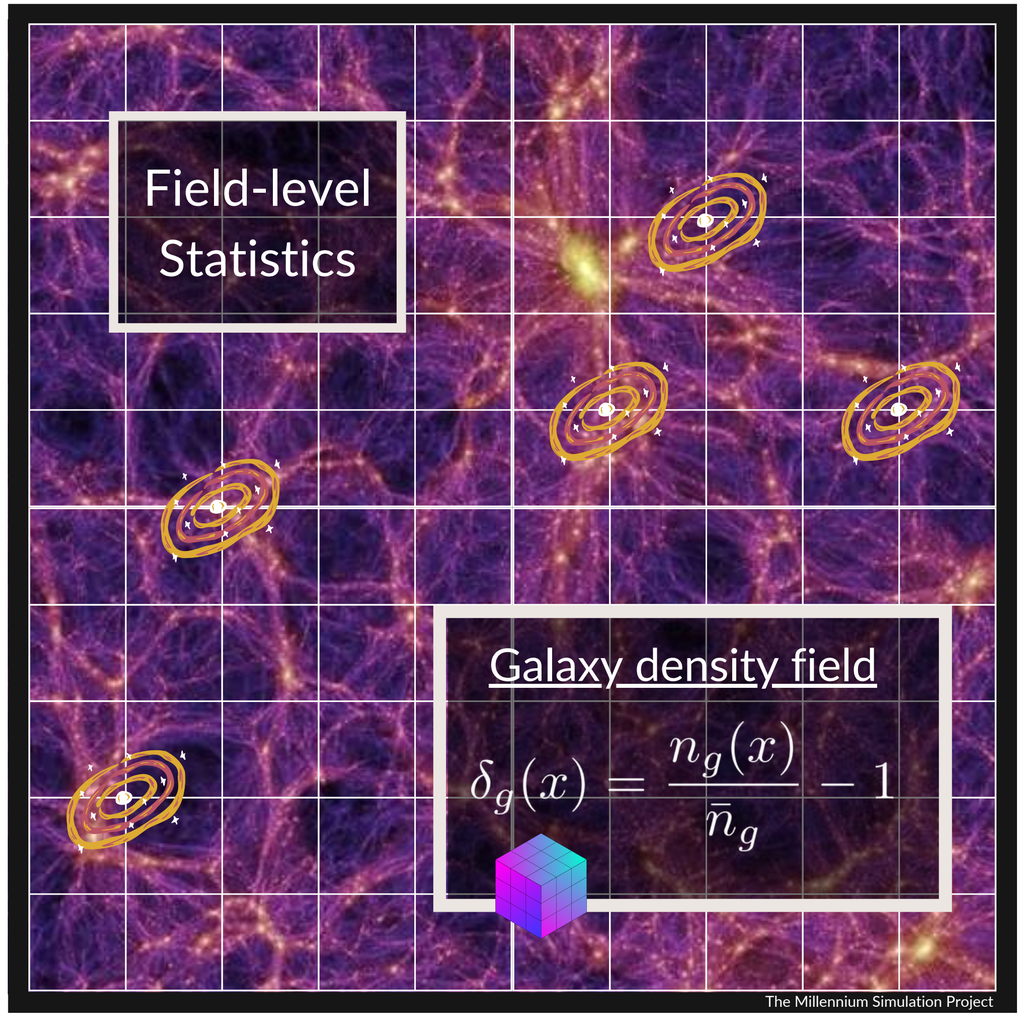




Field-level inference - rethinking how we analyze the data

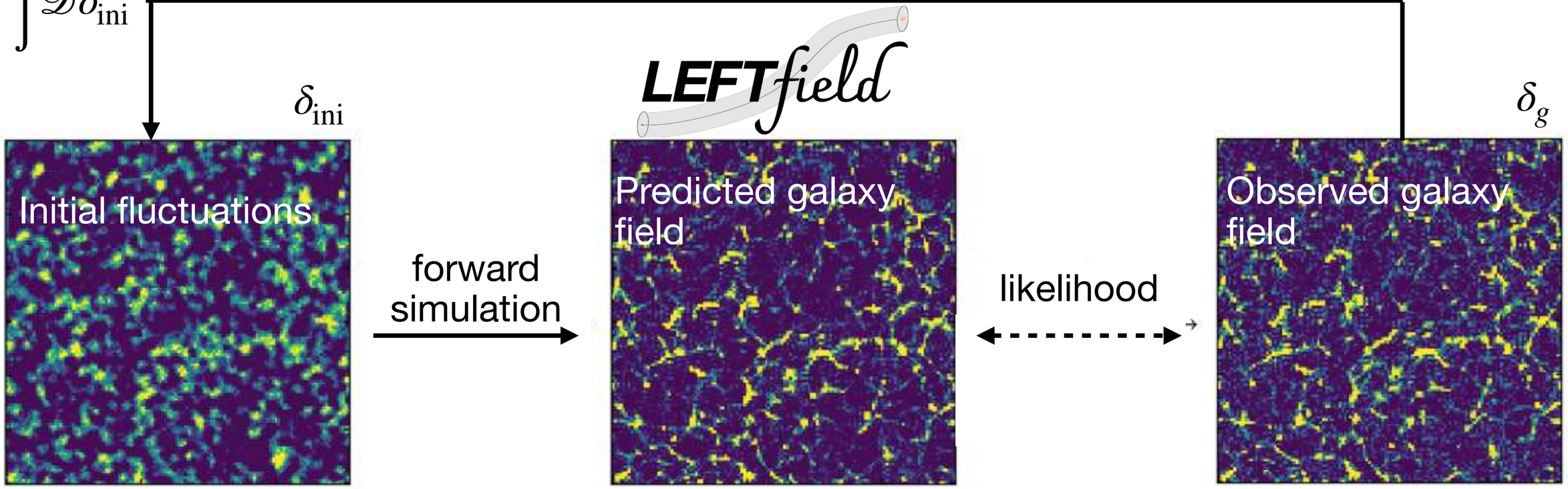
Summary statistics can be lossy, can we directly analyze the data at the field level?

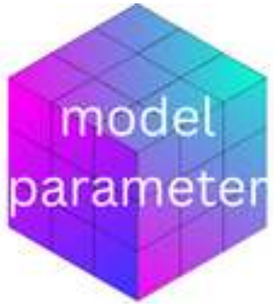




Field-level inference - rethinking how we analyze the data

$\int \mathscr{D}\delta_{\rm ini}$ Explicit sampling and marginalization

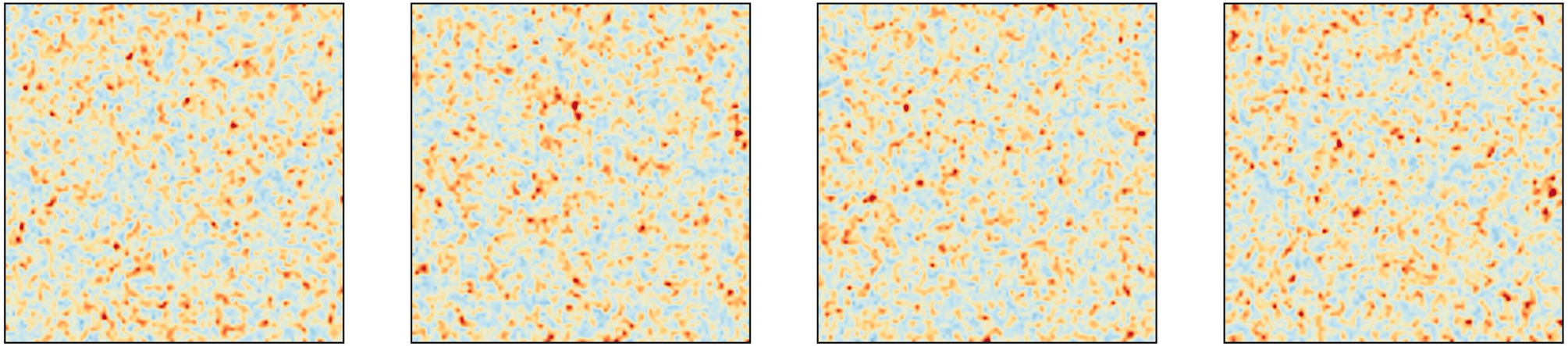




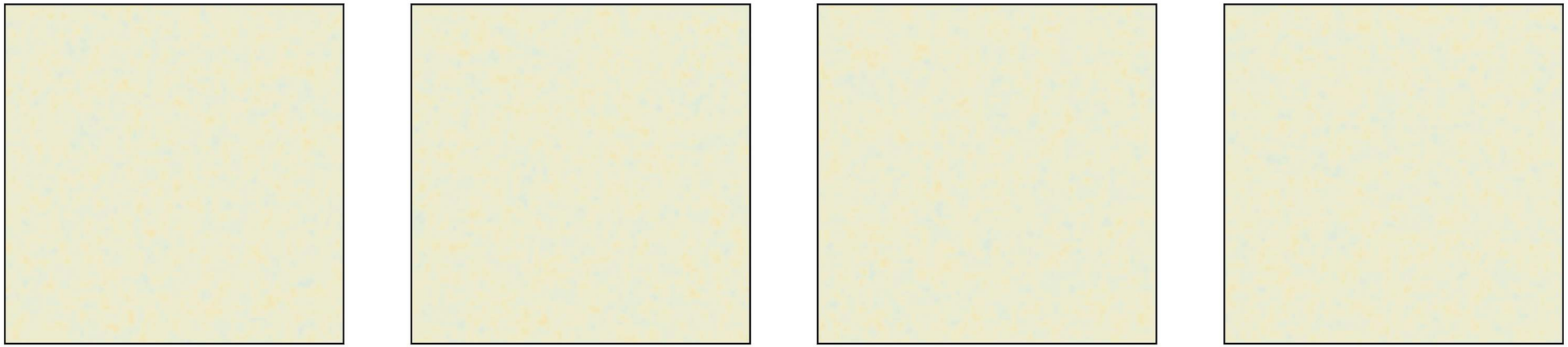
likelihood

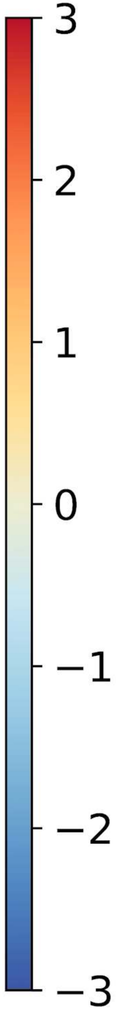
$\log \mathcal{L} = -1.184e+08$

data



model prediction





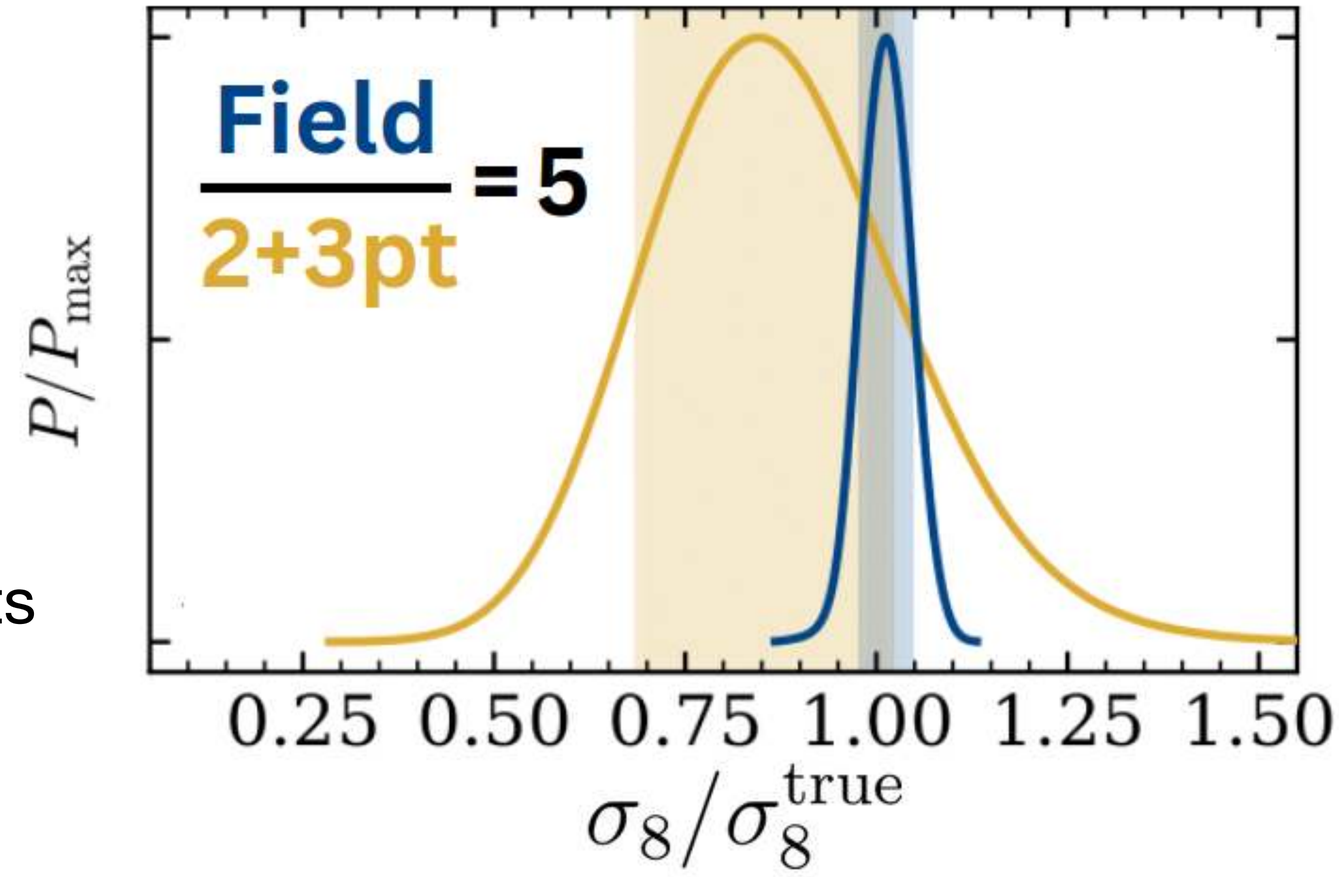
2+3-point (P+B) vs. Field-level inference comparison

Apple-to-apple comparison on

N-body halo catalogs

no redshift space distortion

conservative analysis scale cuts



Effectively x25 survey volume!

MN+2024 PRL 133, 221006

but see also Akitsu+2025

Public data challenge for novel beyond-two-point statistics to recover true cosmology from blind analyses

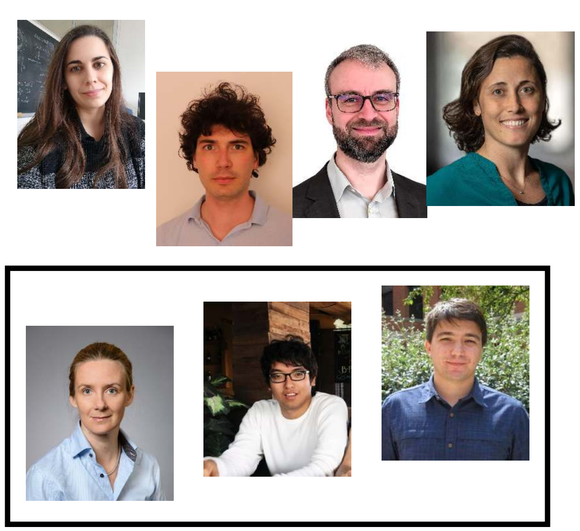
Beyond-2pt Collaboration, ApJ 990 99

incl. MN

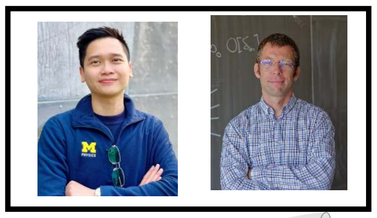
A Parameter-Masked Mock Data Challenge for Beyond-Two-Point Galaxy Clustering Statistics*

THE BEYOND-2PT COLLABORATION

ELISABETH KRAUSE,[1] YOSUKE KOBAYASHI,[1,2] ANDRÉS N. SALCEDO,[1] MIKHAIL M. IVANOV,[3] TOM ABEL,[4,5,6] KAZUYUKI AKITSU,[7] RAUL E. ANGULO,[8,9] GIOVANNI CABASS,[10] SOFIA CONTARINI,[11,12,13] CAROLINA CUESTA-LAZARO,[14,15,16] CHANGHOON HAHN,[17] NICO HAMAUS,[18,19] DONGHUI JEONG,[20,21] CHIRAG MODI,[22,23] NHAT-MINH NGUYEN,[24,25] TAKAHIRO NISHIMICHI,[2,26,27] ENRIQUE PAILLAS,[28,29] MARCOS PELLEJERO IBAÑEZ,[30] OLIVER H. E. PHILCOX,[31,32] ALICE PISANI,[33,22,34,17] FABIAN SCHMIDT,[35] SATOSHI TANAKA,[26] GIOVANNI VERZA,[36,22] SIHAN YUAN,[4,6] MATTEO ZENNARO,[37]



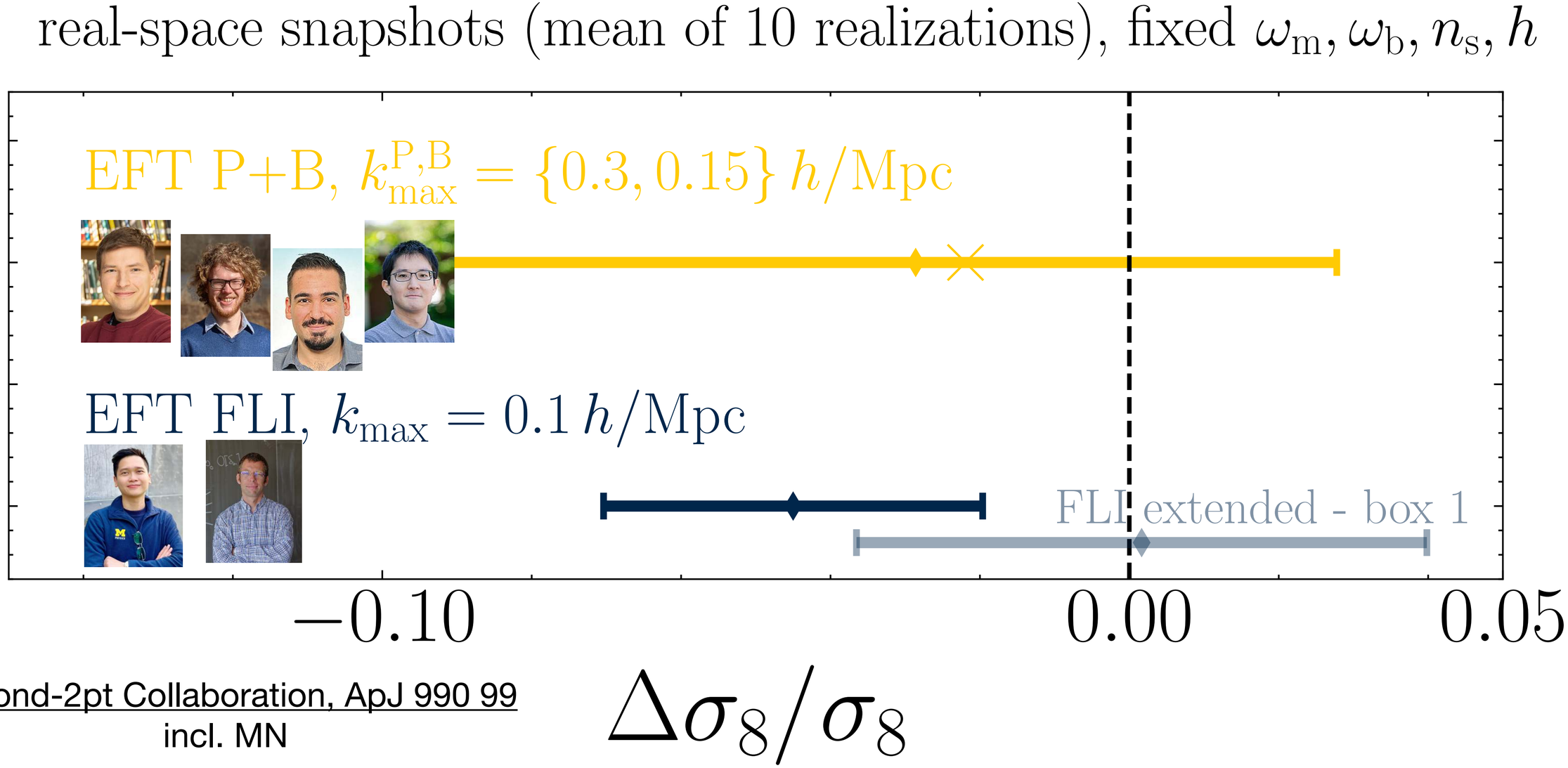
Challenge organizers



LEFTfield

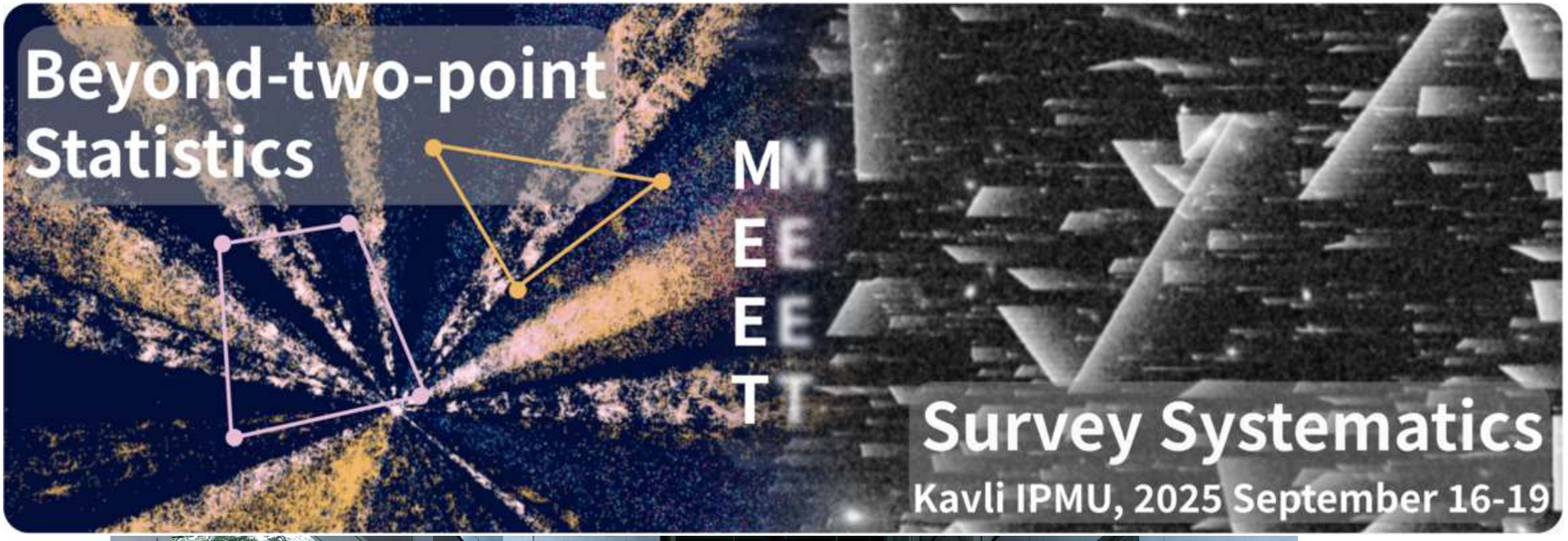
Our team

P+B vs FLI on HOD mocks in the beyond-2point blind data challenge



Beyond-2pt Collaboration, ApJ 990 99
incl. MN

New physics or systematics? The beyond-2pt community will be ready to answer that



~70 participants, mostly overseas

postdocs > PhDs > faculties

→

New maps of systematics

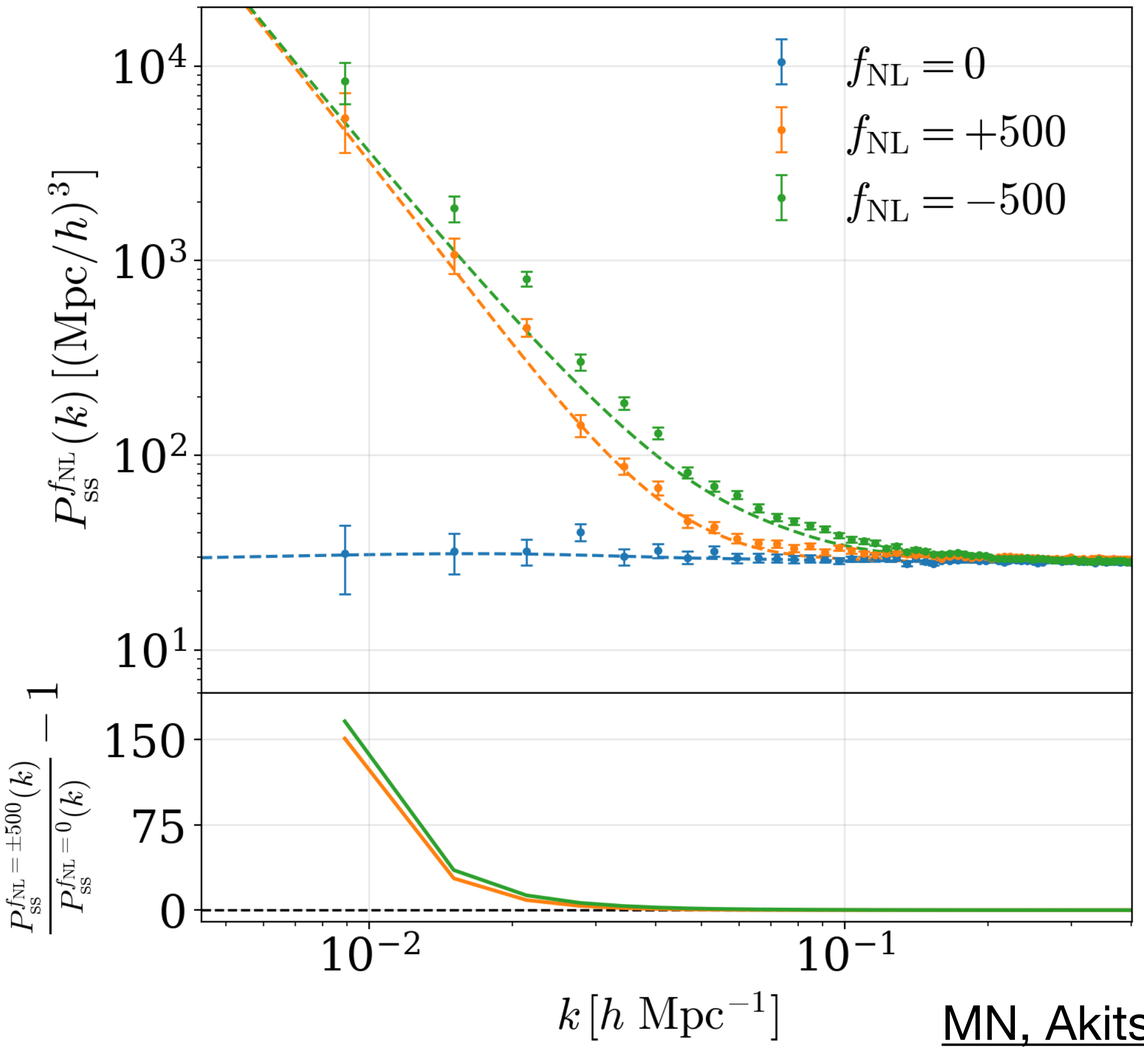
New models for systematics

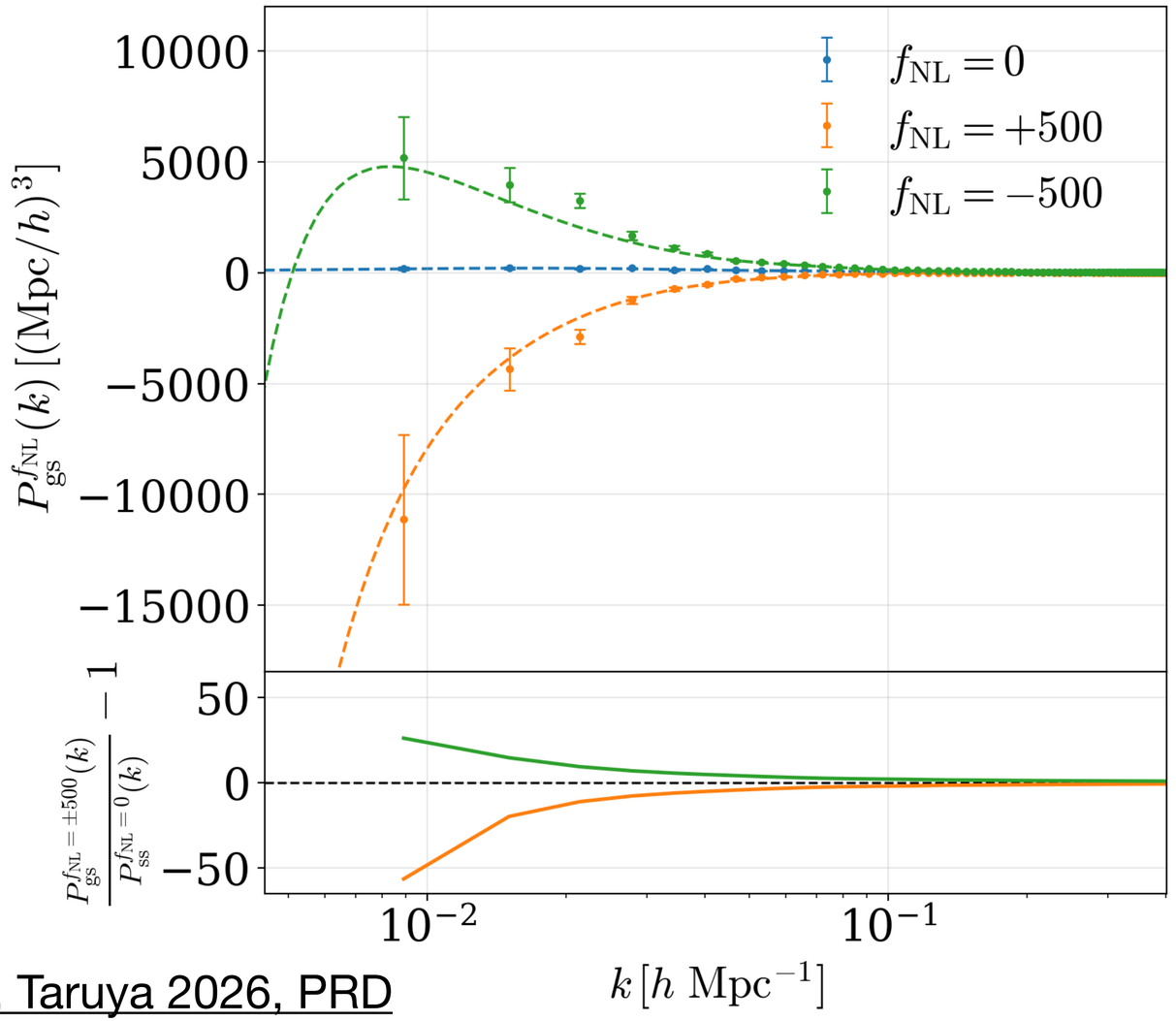
Standardize validation schemes

New data challenges!

Galaxy sizes as a new statistics to constrain local PNG

$$\Phi(\mathbf{x}) = \phi_G(\mathbf{x}) + f_{\rm NL}^{\rm local}\left[\phi_G^2(\mathbf{x}) - \left\langle \phi_G^2(\mathbf{x})\right\rangle\right]$$

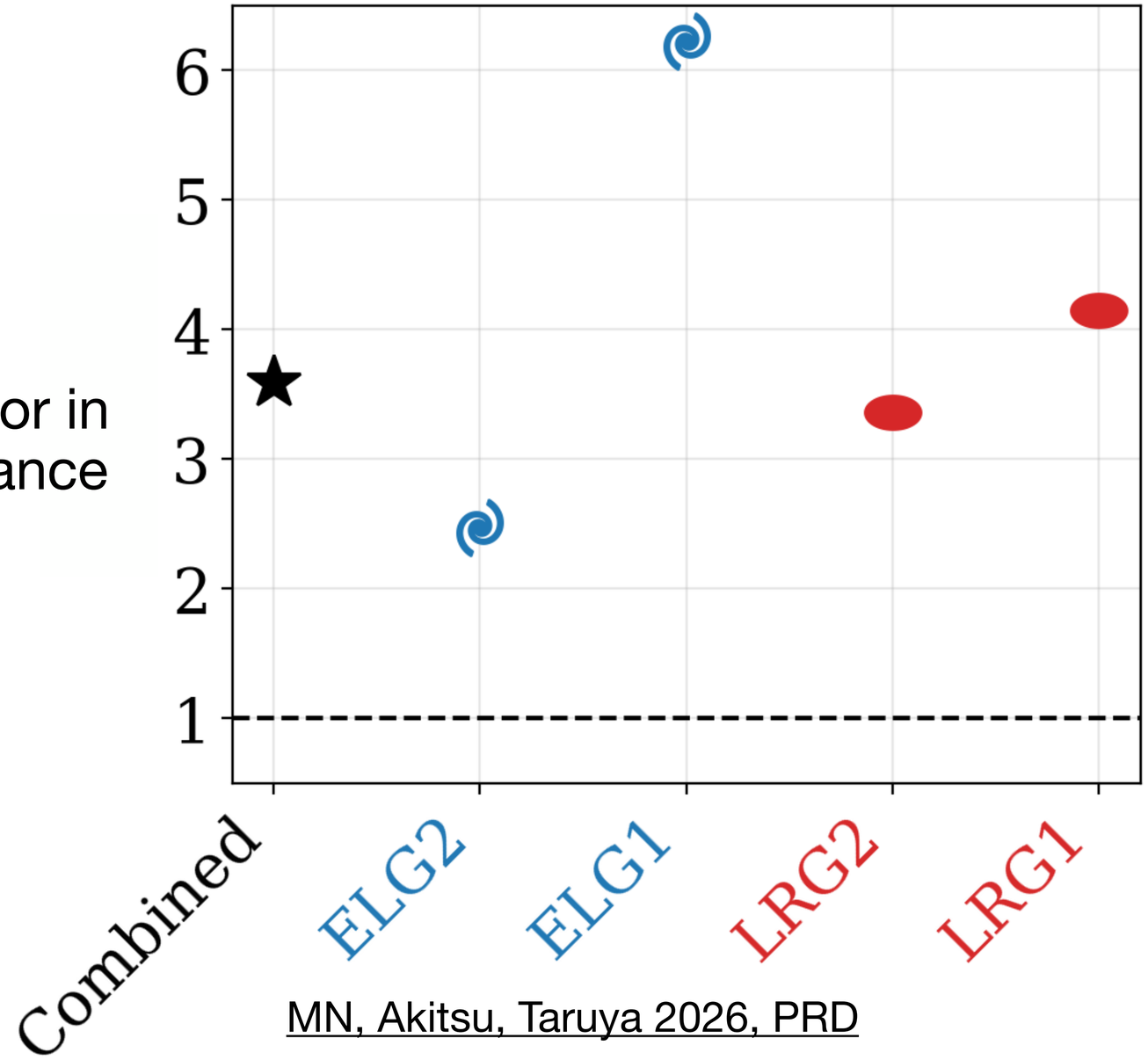




MN, Akitsu, Taruya 2026, PRD

Galaxy sizes as a natural multitracer of local PNG

Improvement factor in detection significance

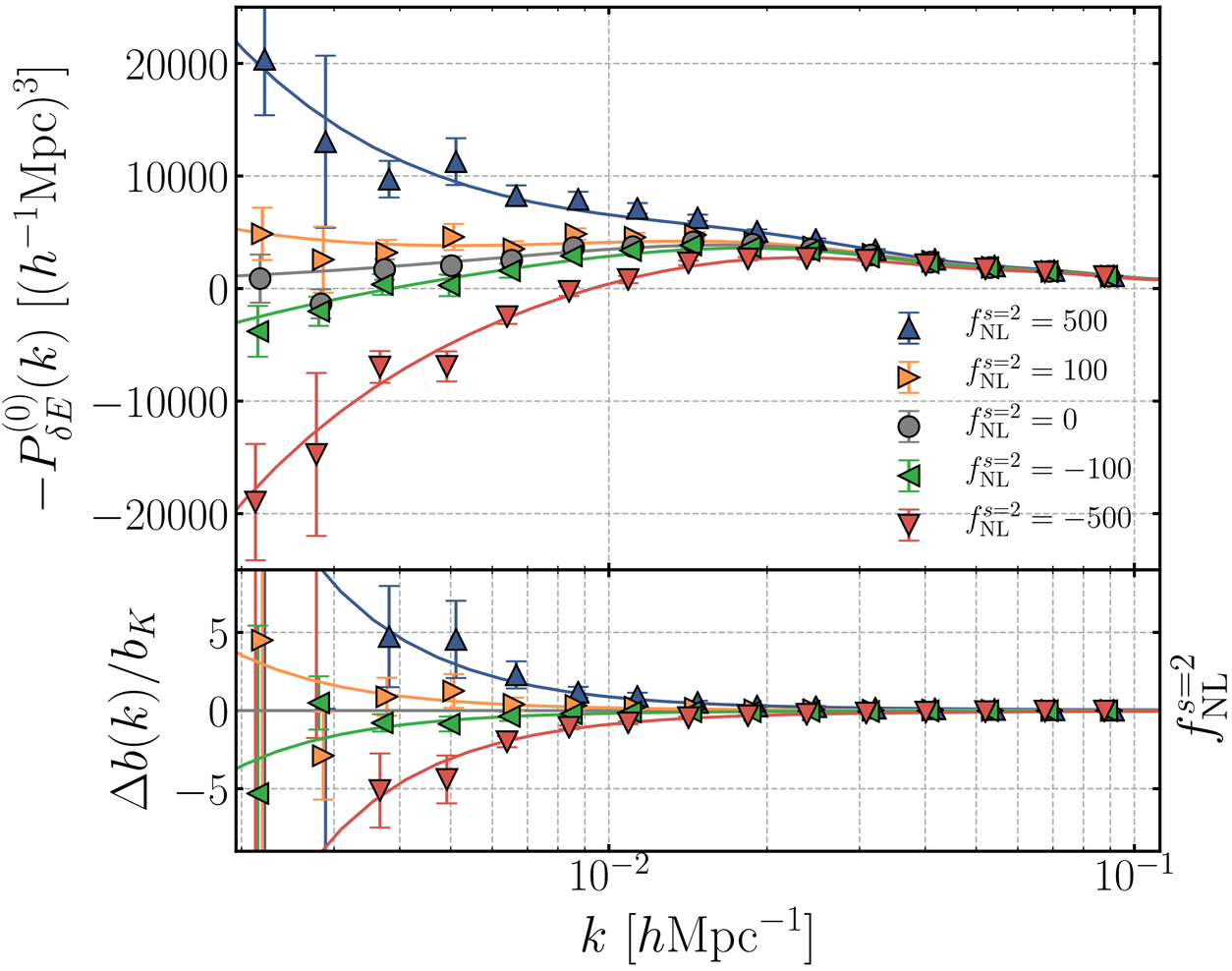


DESI full-sample forecast

MN, Akitsu, Taruya 2026, PRD

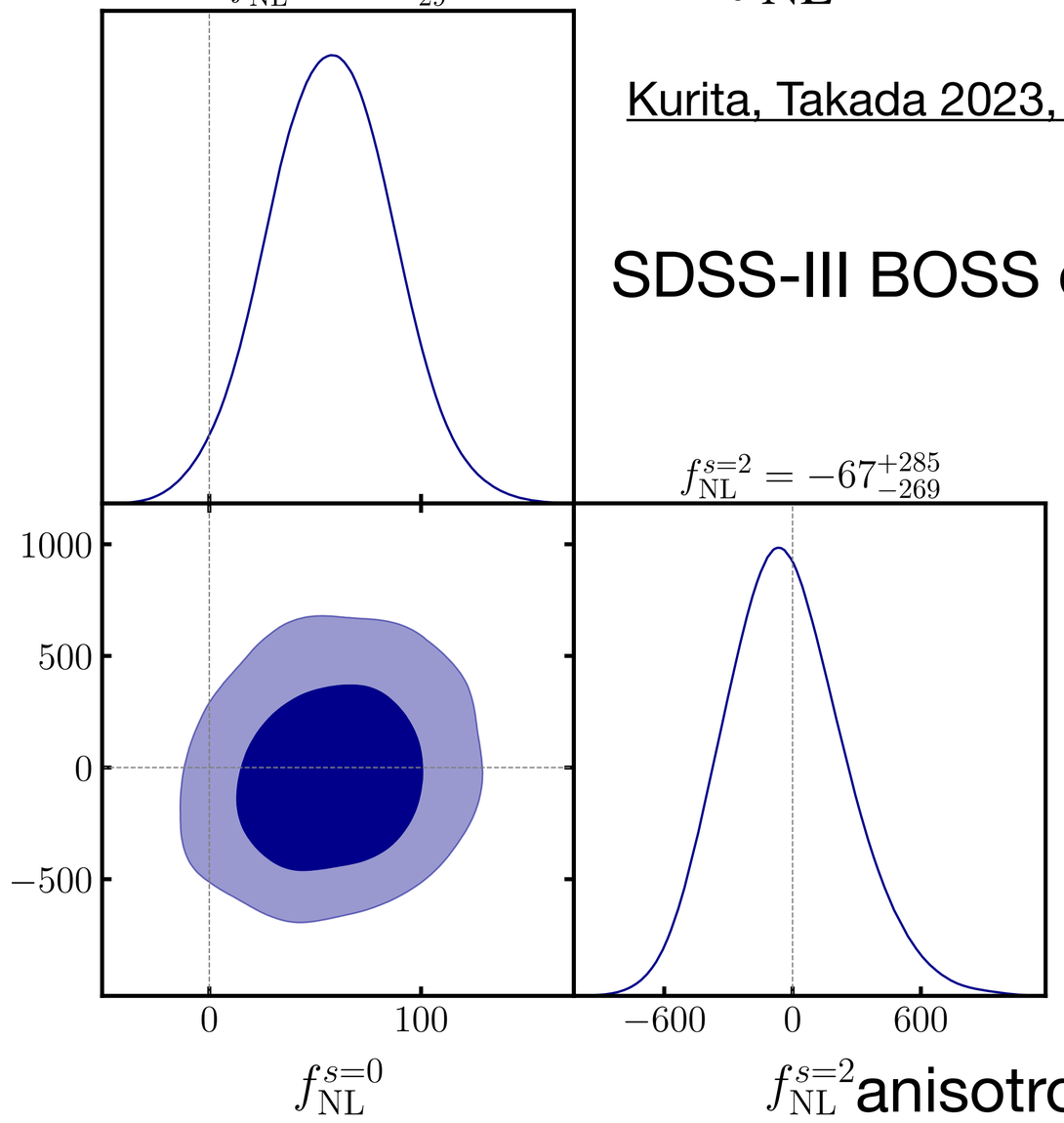
Galaxy shapes similarly probe *anisotropic* local PNG

Akitsu+ 2020, PRD



High-resolution N-body simulation

isotropic f_{NL}



Kurita, Takada 2023, PRD

SDSS-III BOSS data

anisotropic f_{NL}

The next era of LSS surveys and anlysis

Precision tests of fundamental physics at cosmological scales

Search for new physics in both the late and early universe

New methods and statistics will maximize discovery potential

KAVLI IPMU

Credit: VisualCinnamon