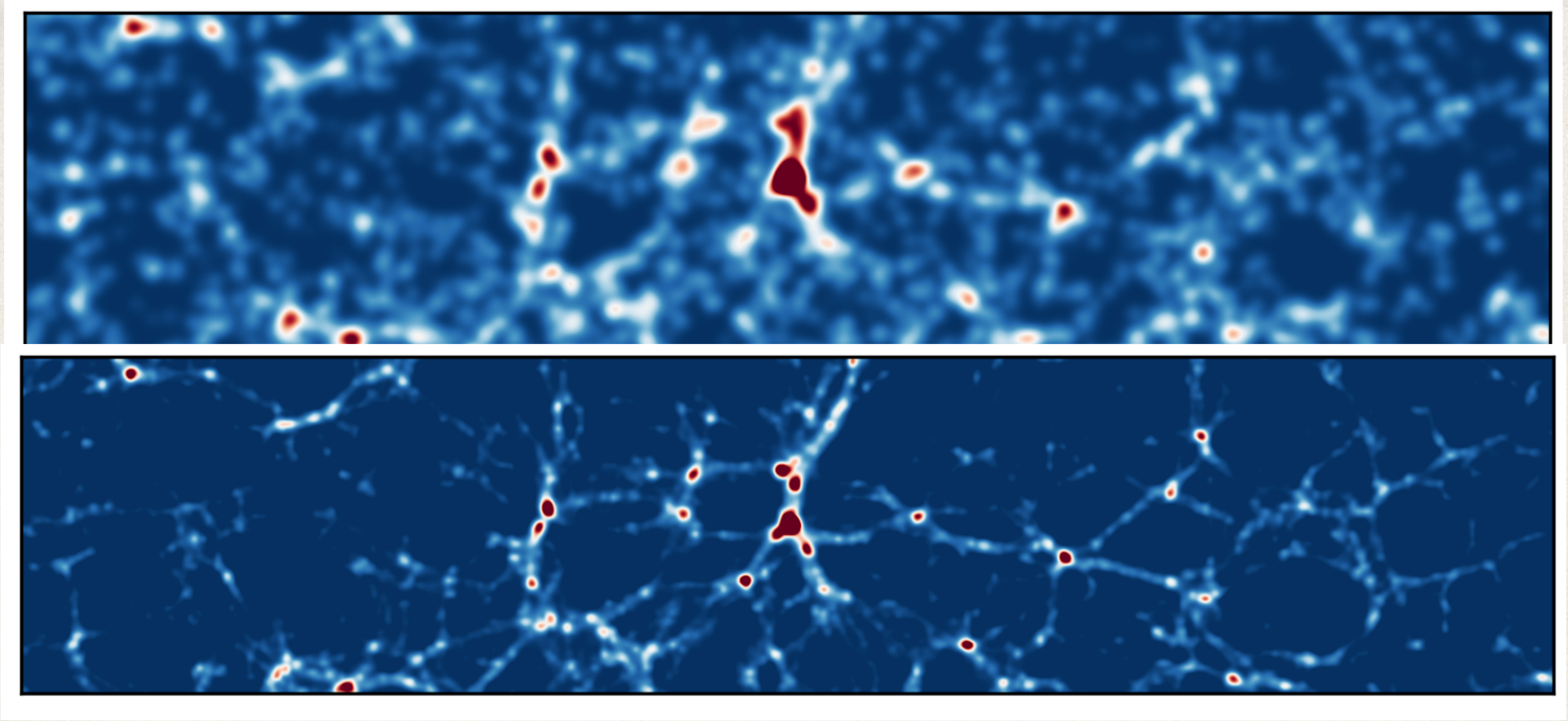


Cross-Correlation Science in the Era of Precision Cosmology

Joe DeRose (UC Santa Cruz & UC Berkeley)

1/29/21 — LBNL RPM/INPA

Outline

A roadmap to obtaining accurate and precise cosmological constraints from cross-correlation analyses

- **Motivation**
- **Part I**: Analysis robustness tests with synthetic sky catalogs
  - Case study: DES & DESI
- **Part II**: Accurate and precise models for lensing and clustering cross-correlations
  - Combining perturbation theory & simulation

The Standard Model of Cosmology

We are converging on a standard picture of how the universe evolved over the last 13 billion years:

- Quantum fluctuations made macroscopic by inflation seed large scale structure imprinted on CMB
- Structure growth driven by the presence of dark matter
- Late time accelerated expansion driven by dark energy

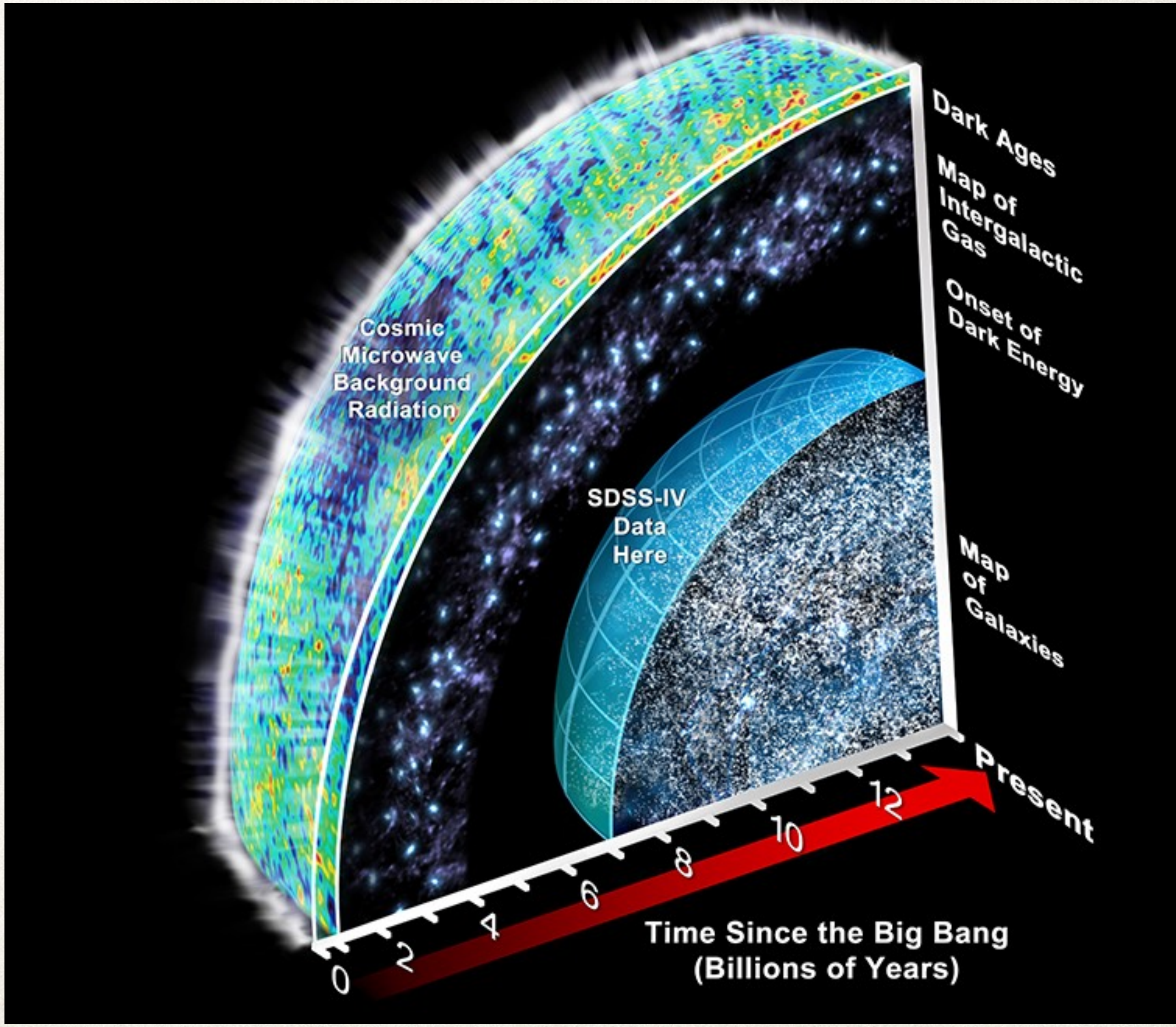
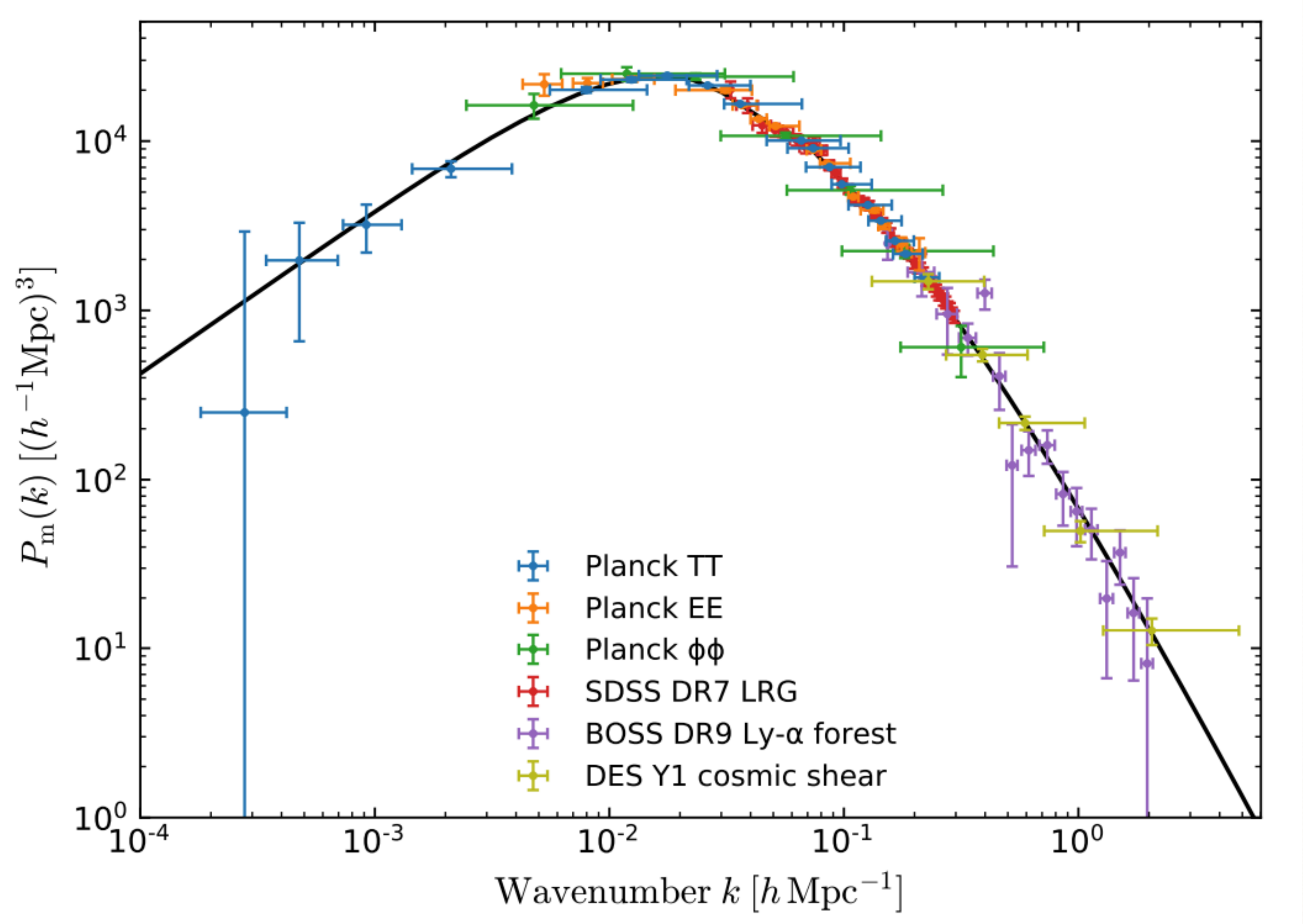


figure courtesy of eBOSS/SDSS

The Standard Model of Cosmology

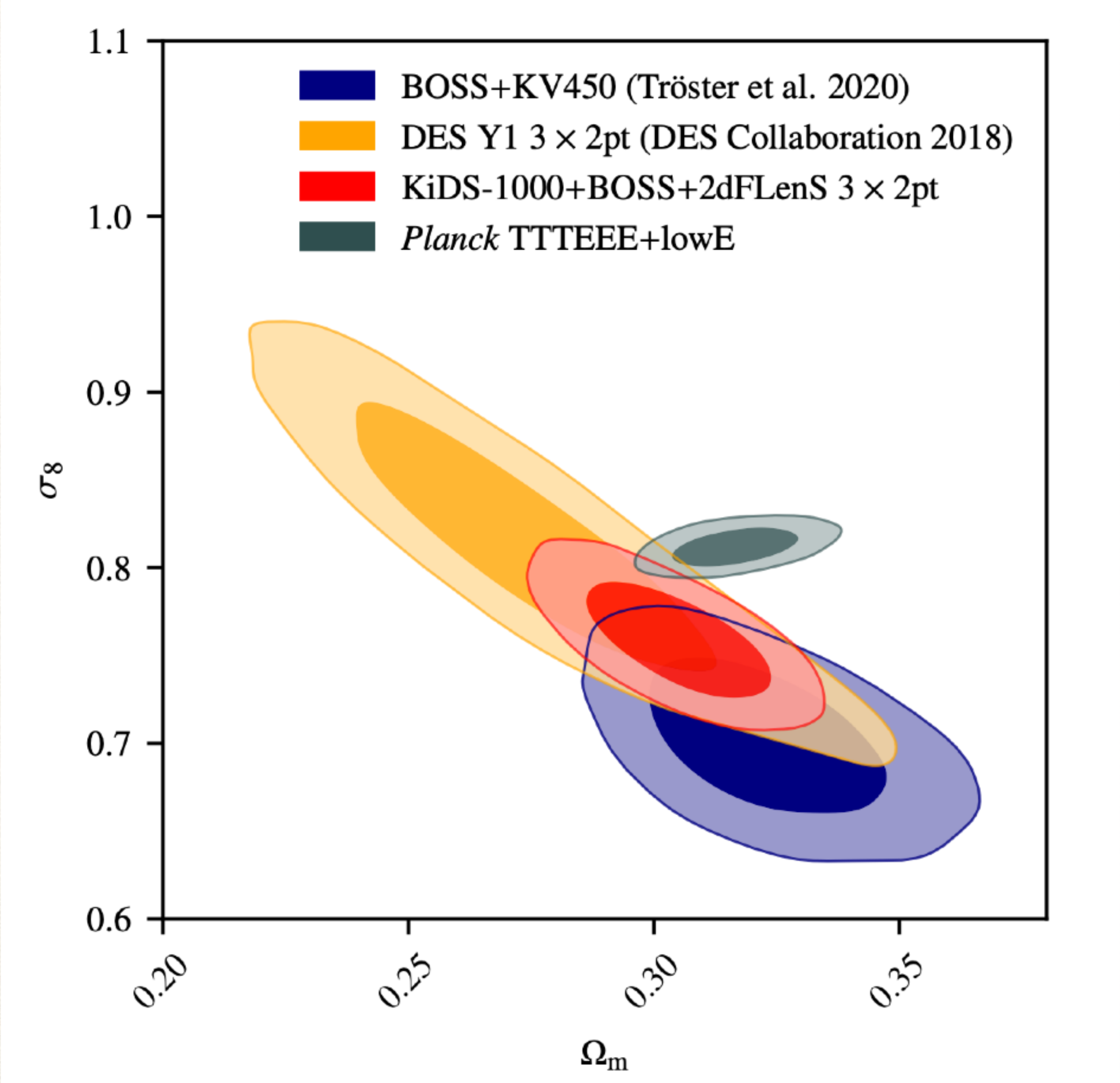
Nearly all cosmological observations agree on this standard model: LCDM



Planck Collaboration 2018

Low Redshift Universe Tests of LCDM

Heymans & Troster et al. 2020



Upcoming surveys will put LCDM to the test, but only if we can get a handle on systematics (both observational and modeling)

The next generation is here

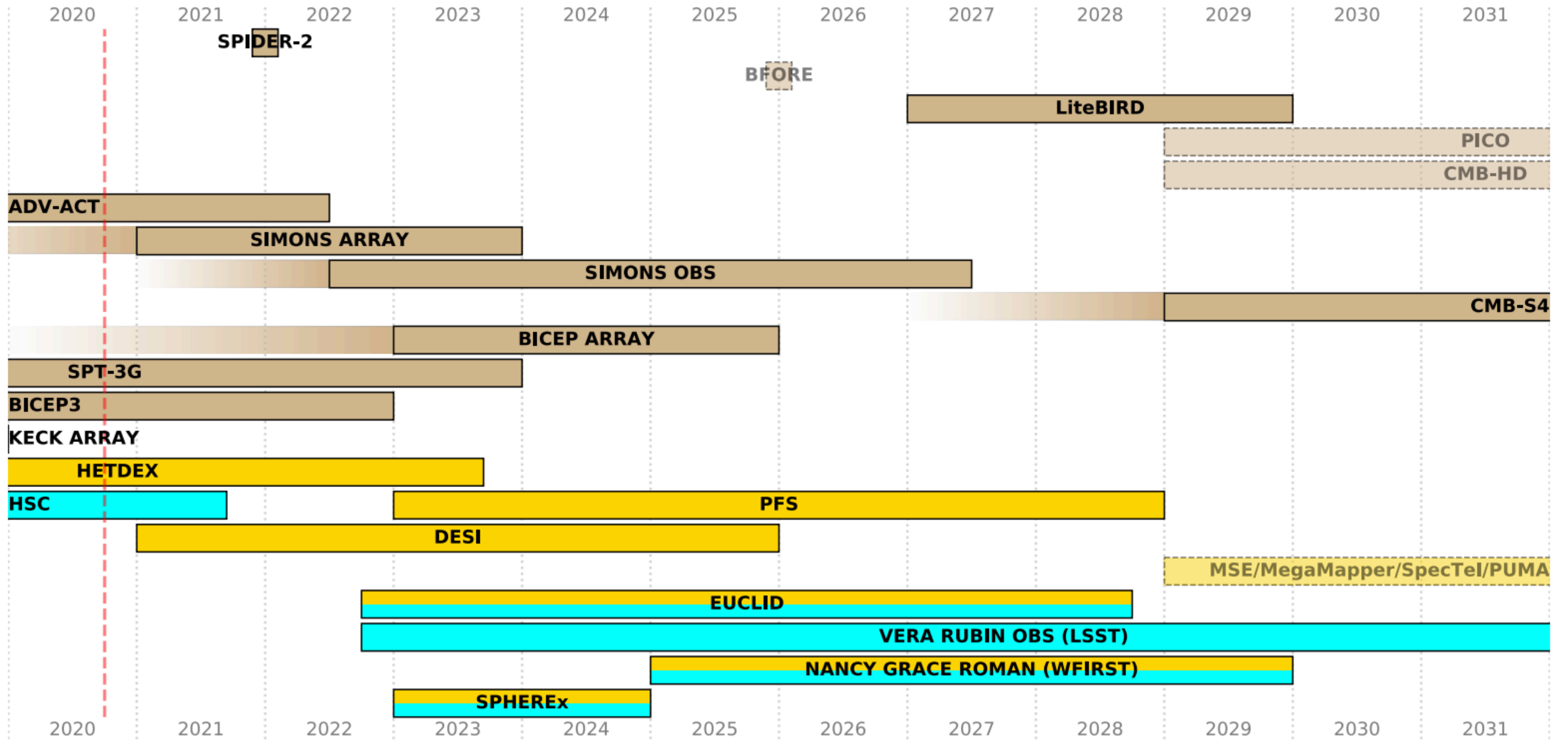
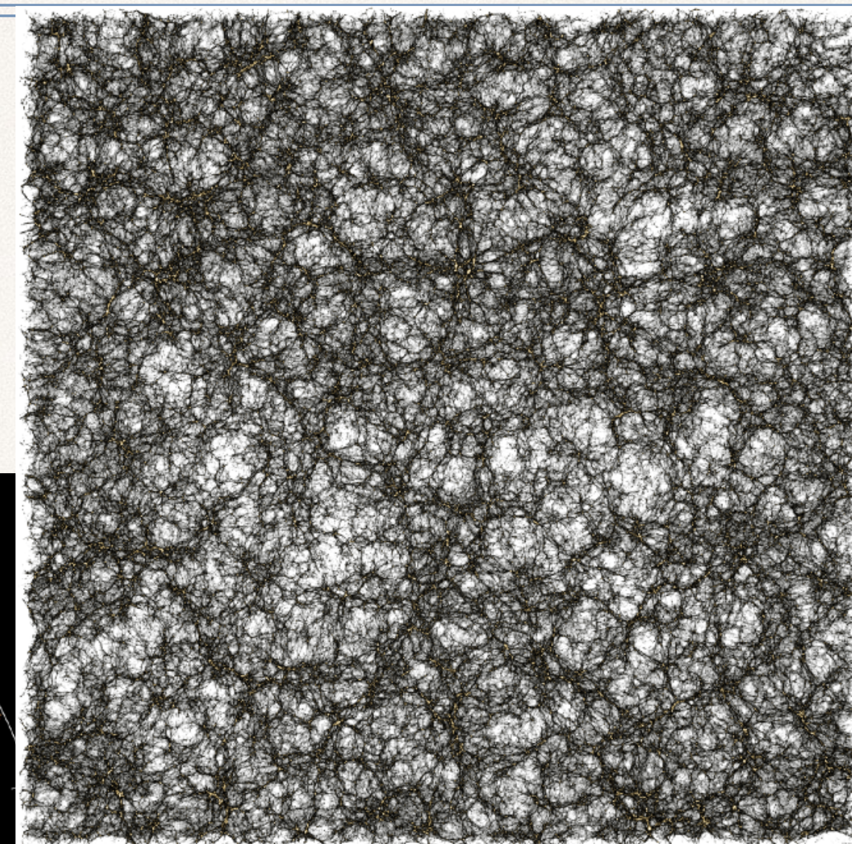
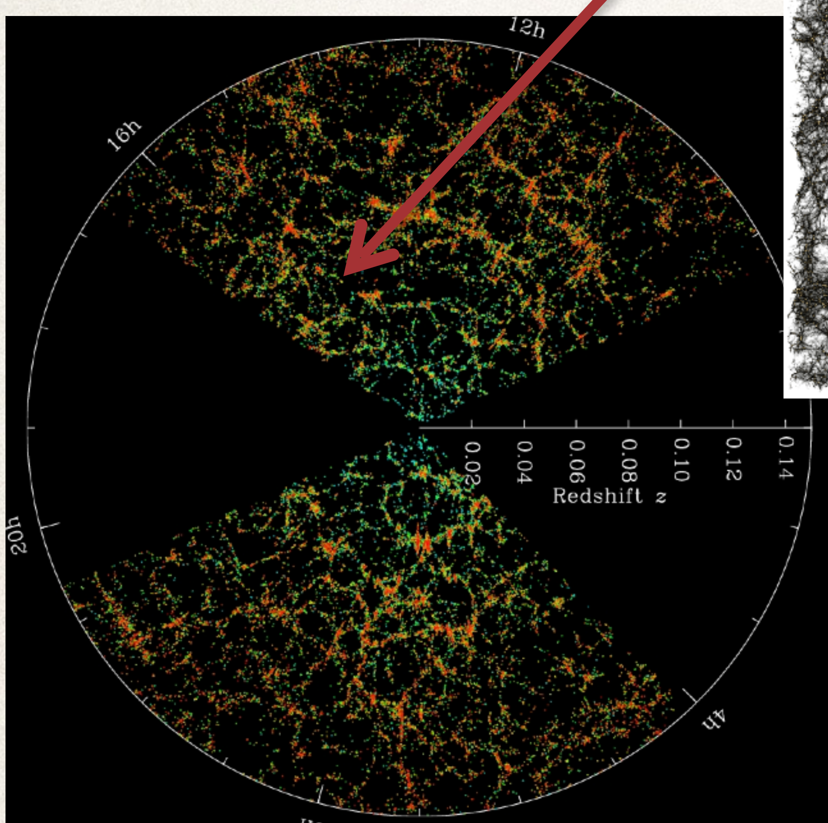
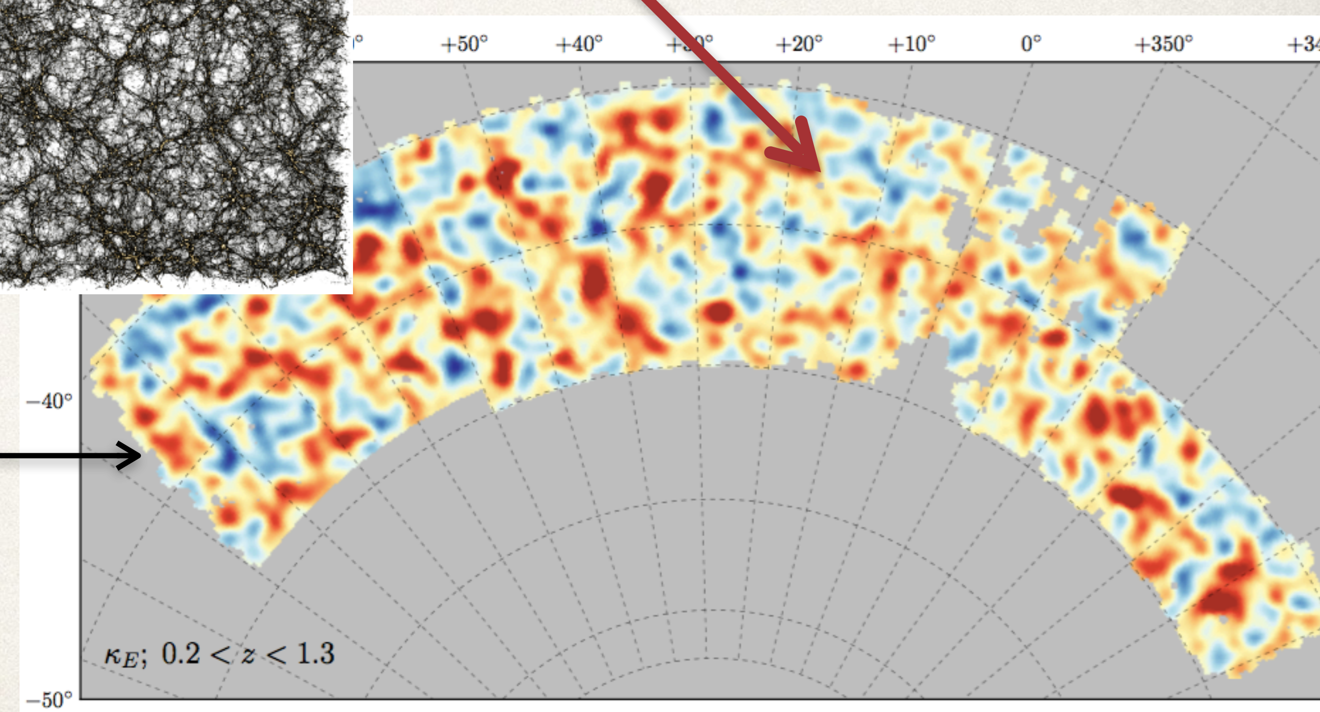


fig. courtesy of D. Kirkby

Probes of low-redshift structure: 3x2-point







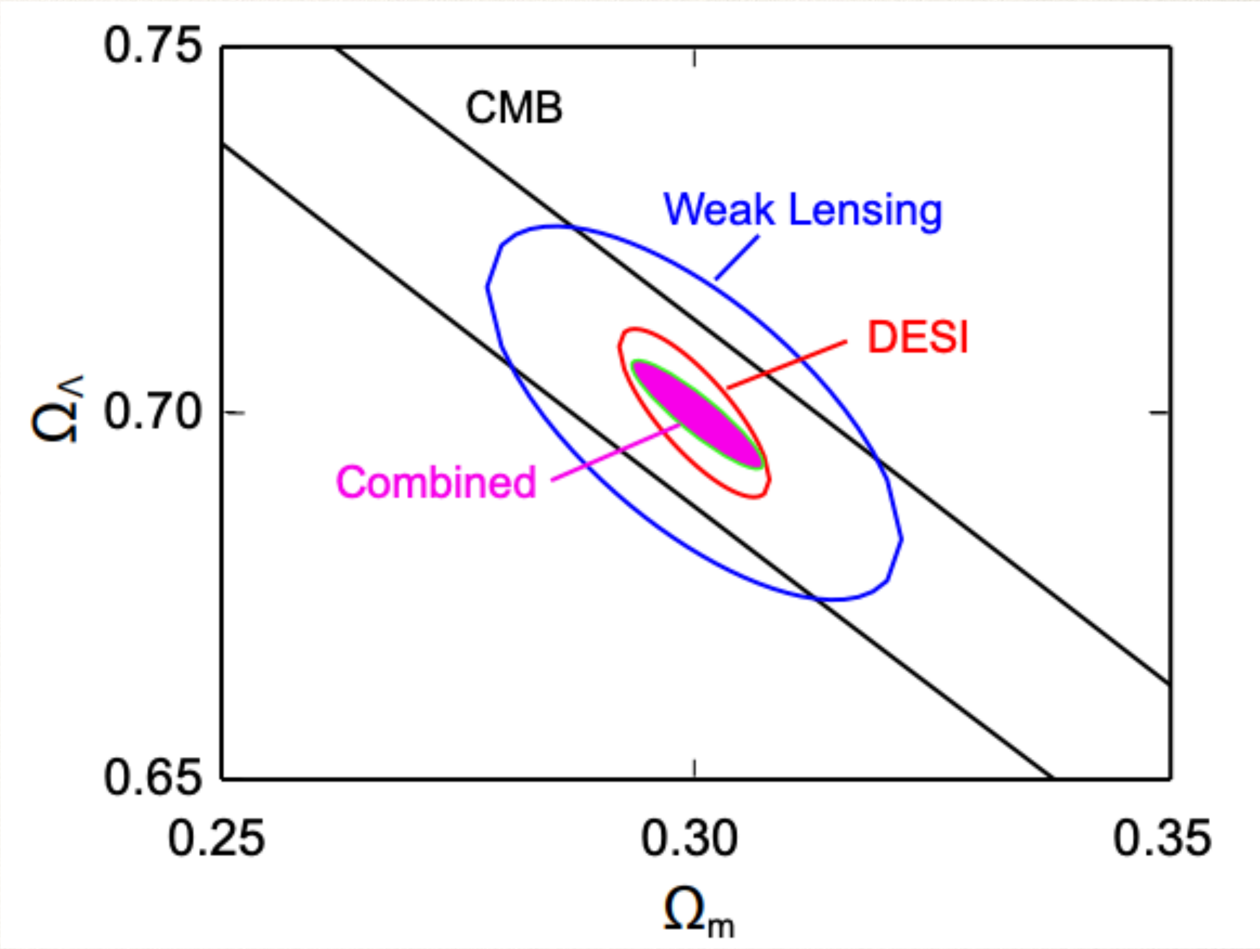
galaxies x *galaxies*
galaxy clustering

galaxies x *lensing*
galaxy-galaxy lensing

lensing x *lensing*
cosmic shear

The power of combined CMB/Galaxy clustering/WL

We need to exploit cross-correlations between DESI / LSST / SO / S4 in order to make the most out of all experiments.



DESI collaboration 2016

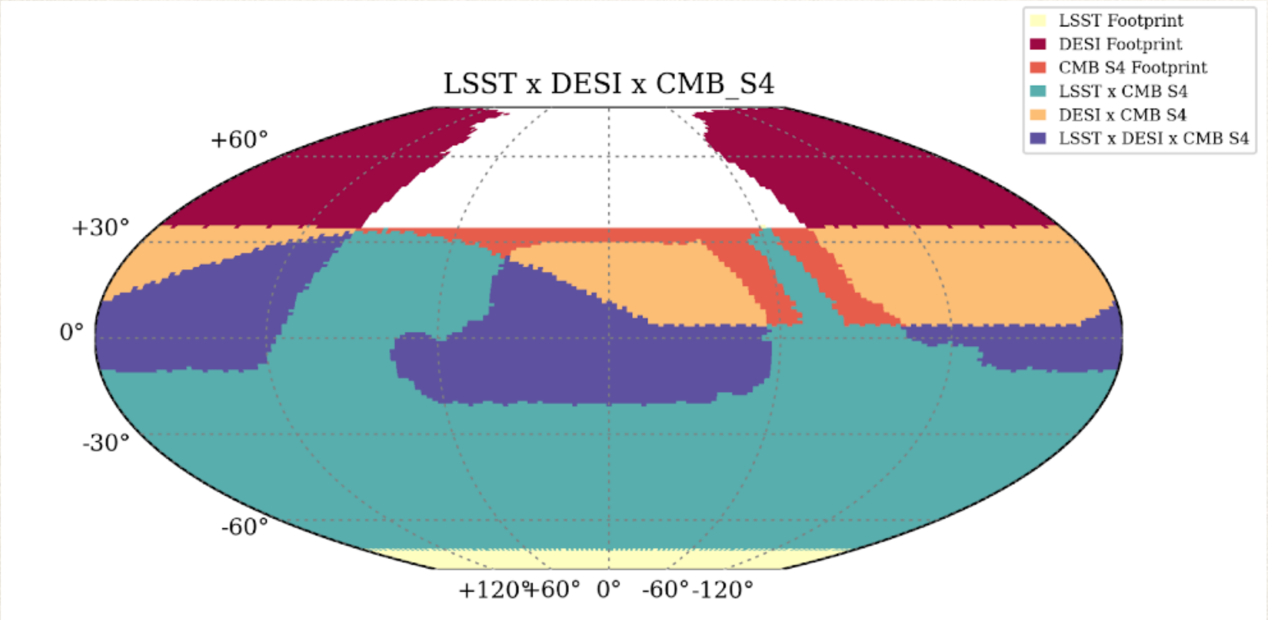


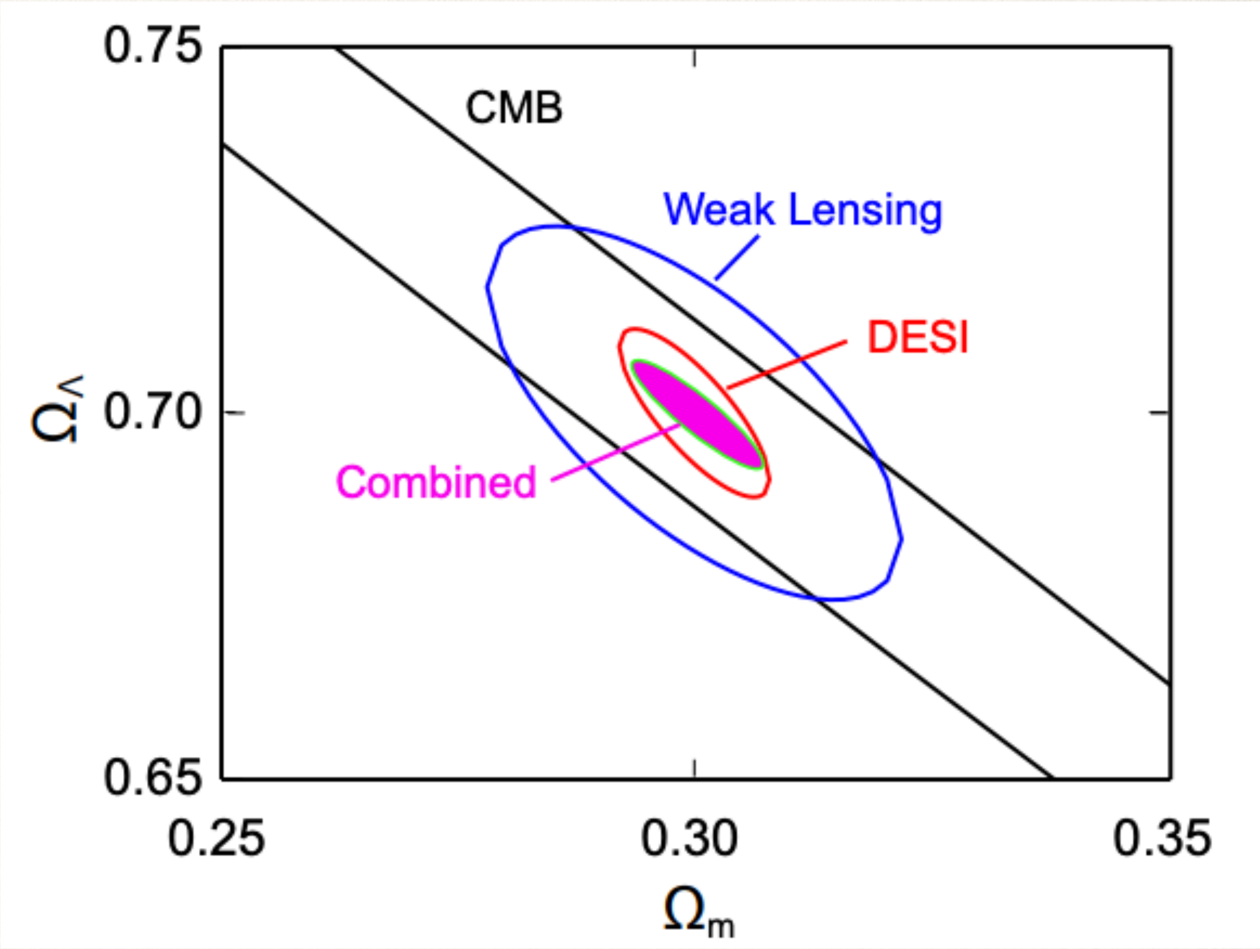
fig. courtesy of C. Zhou and A. Leauthaud

The power of combined CMB/Galaxy clustering/WL

We need:

- Simulations that include x-corr to develop and test analysis methodology.
- Models that are accurate enough to confront this powerful data!

Lots of work ahead



DESI collaboration 2016

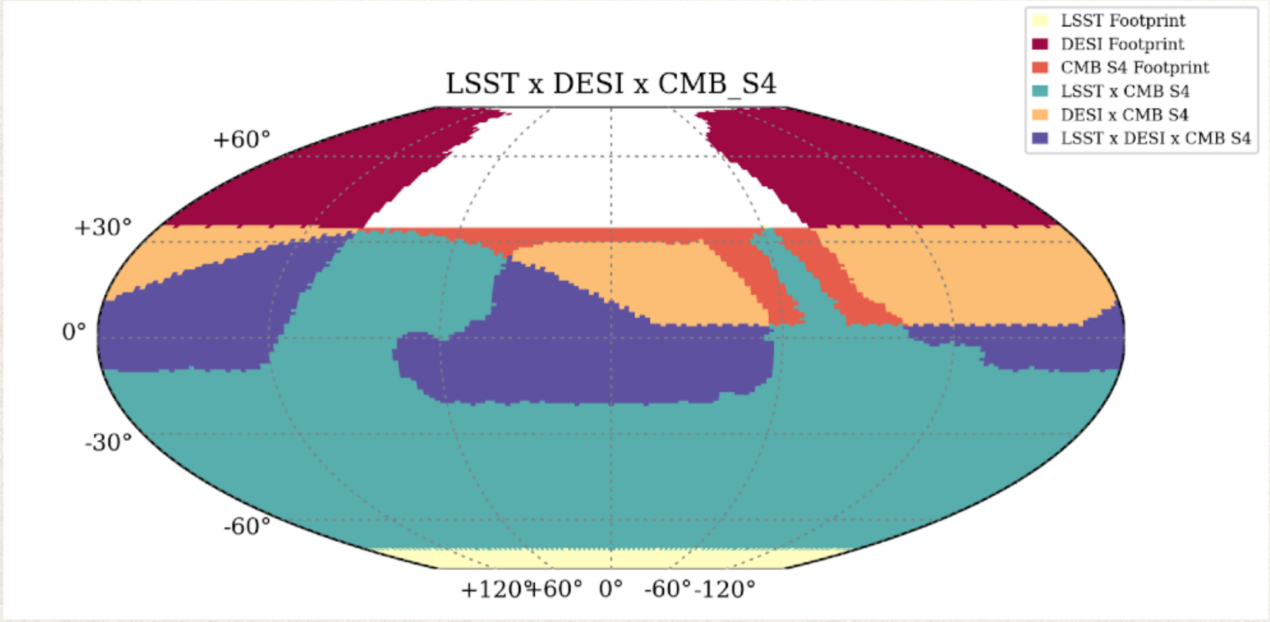
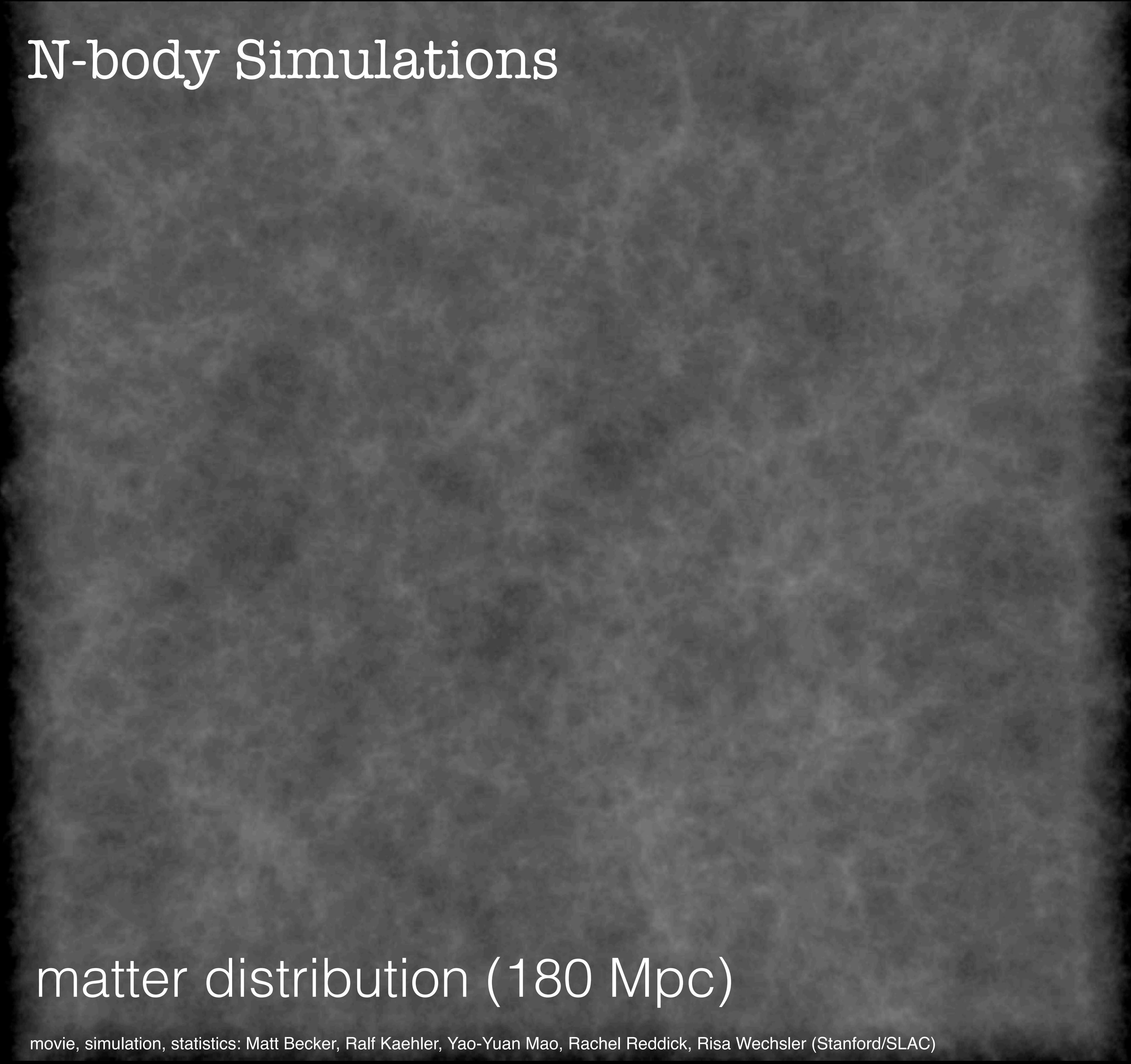


fig. courtesy of C. Zhou and A. Leauthaud

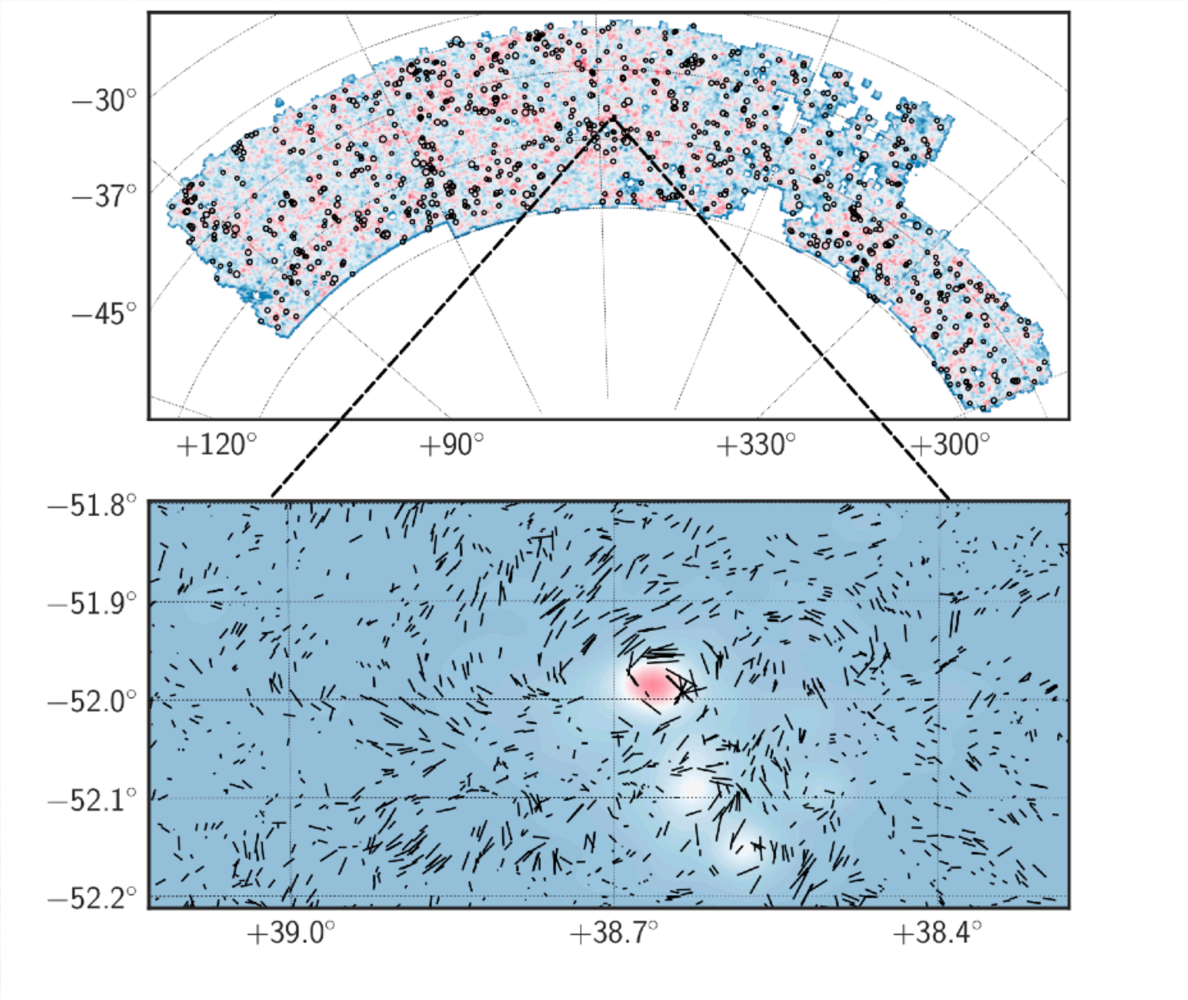
N-body Simulations



matter distribution (180 Mpc)

movie, simulation, statistics: Matt Becker, Ralf Kaehler, Yao-Yuan Mao, Rachel Reddick, Risa Wechsler (Stanford/SLAC)

Part I: Testing x-corr analyses with synthetic sky catalogs



Testing x-corr analyses with synthetic skies

- **Goal:** Perform end-to-end tests going from simulated galaxy catalogs contaminated with realistic systematics to cosmological parameters.
  - Robustness to modeling assumptions/observational systematics

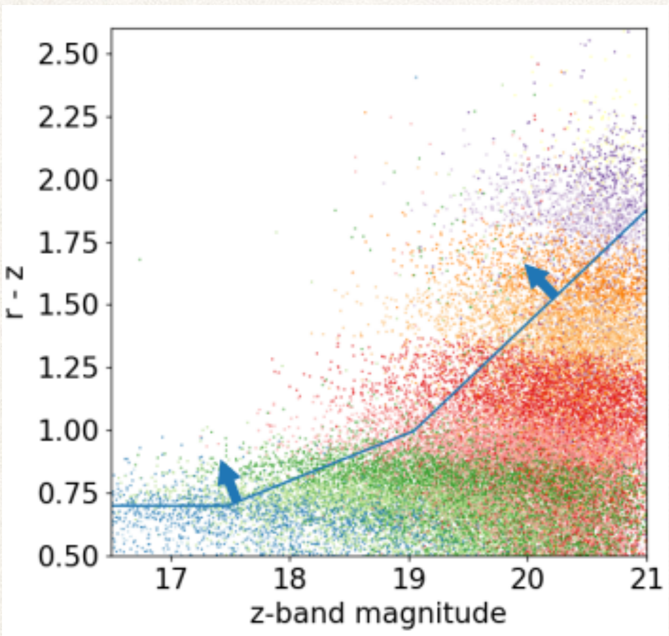
Testing x-corr analyses with synthetic skies

- **Goal:** Perform end-to-end tests going from simulated galaxy catalogs contaminated with realistic systematics to cosmological parameters.
  - Robustness to modeling assumptions/observational systematics
- **Requirements:**
  - Need realistic correlations between cosmological observables and quantities used for sample selection

Example: DESI target selection

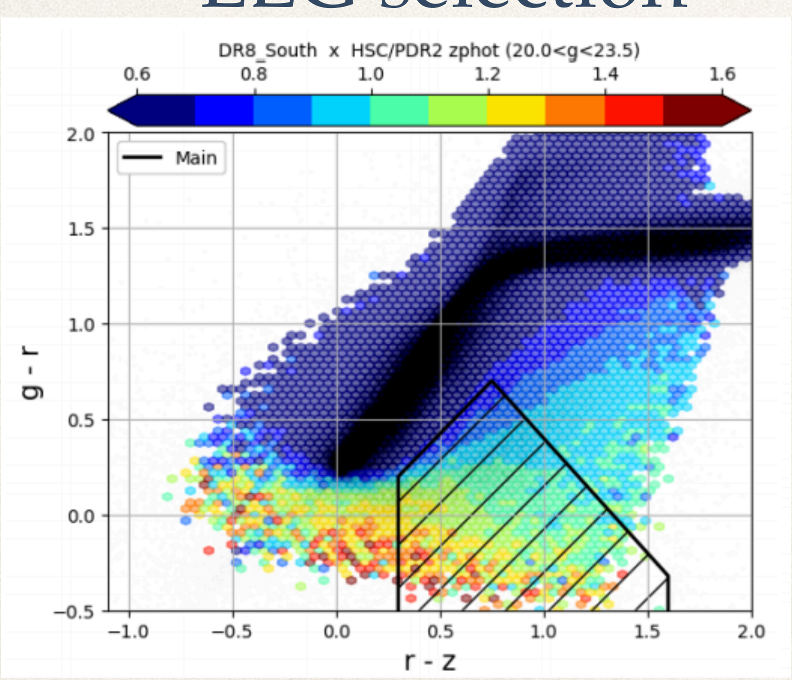
LRG selection

Zhou et al 2020



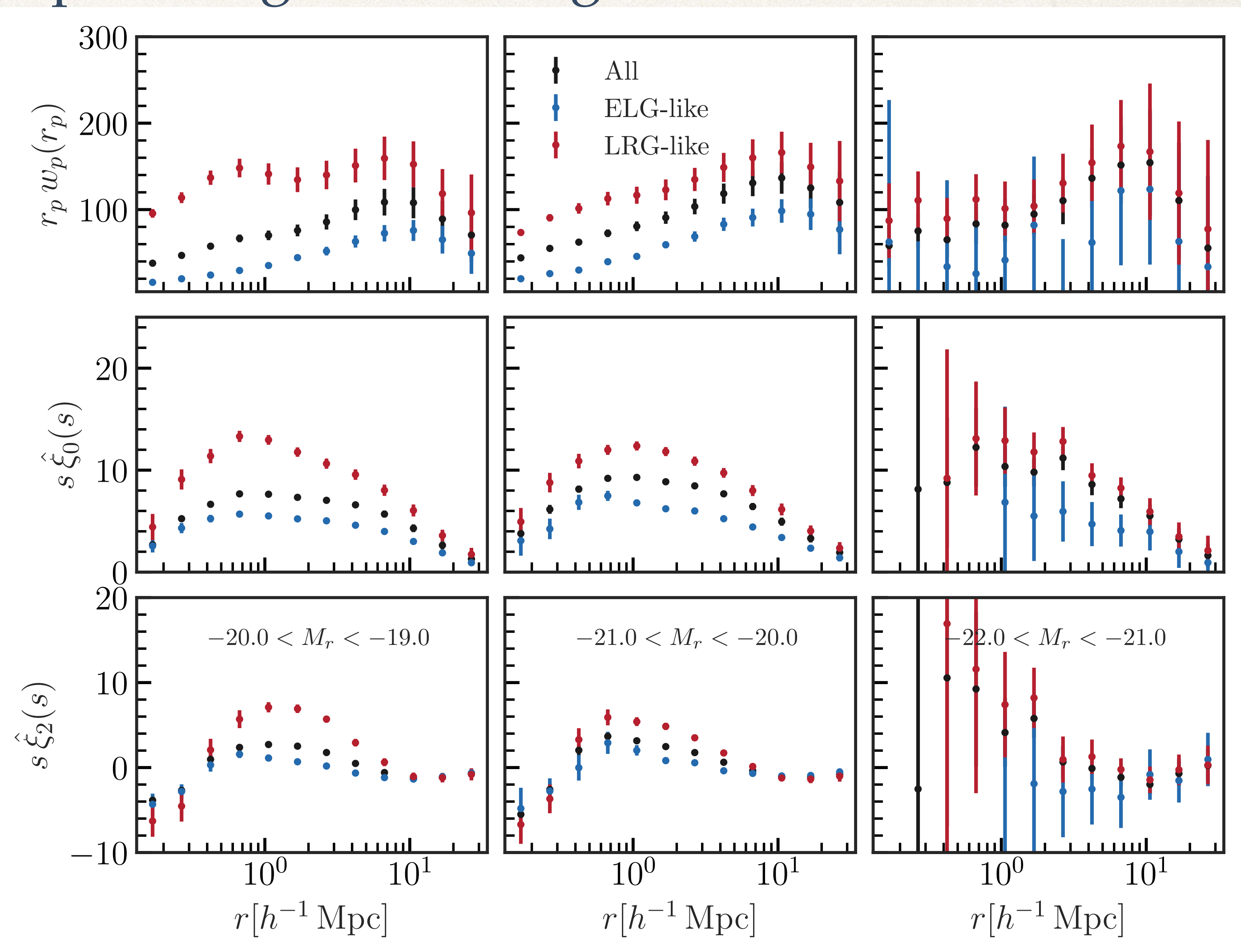
ELG selection

Raichoor et al 2020



Variables used in target selection correlate strongly with clustering amplitude.

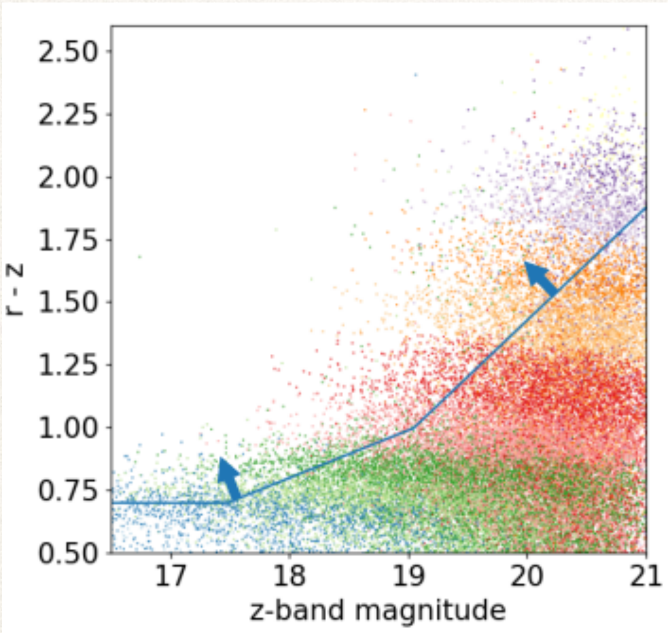
impacts, e.g., fiber assignment / fiber collisions

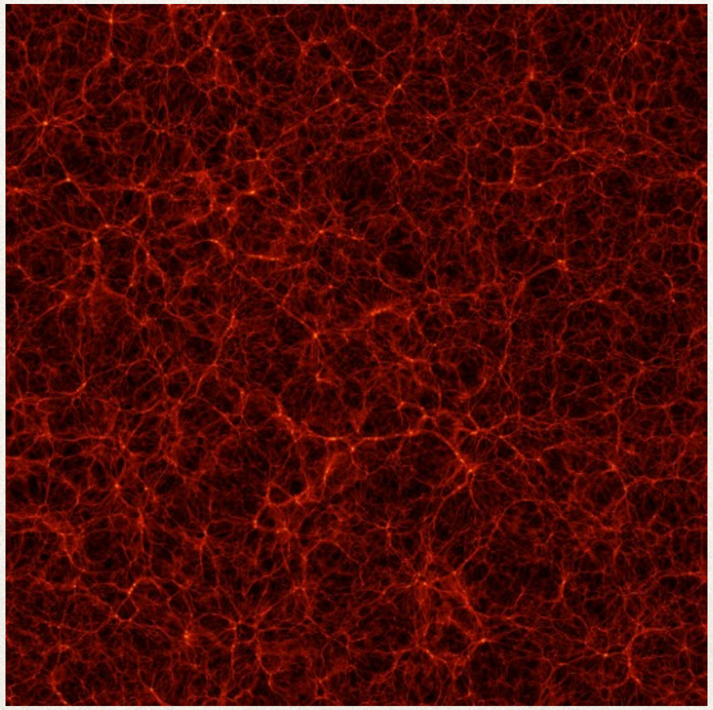


Example: DESI target selection

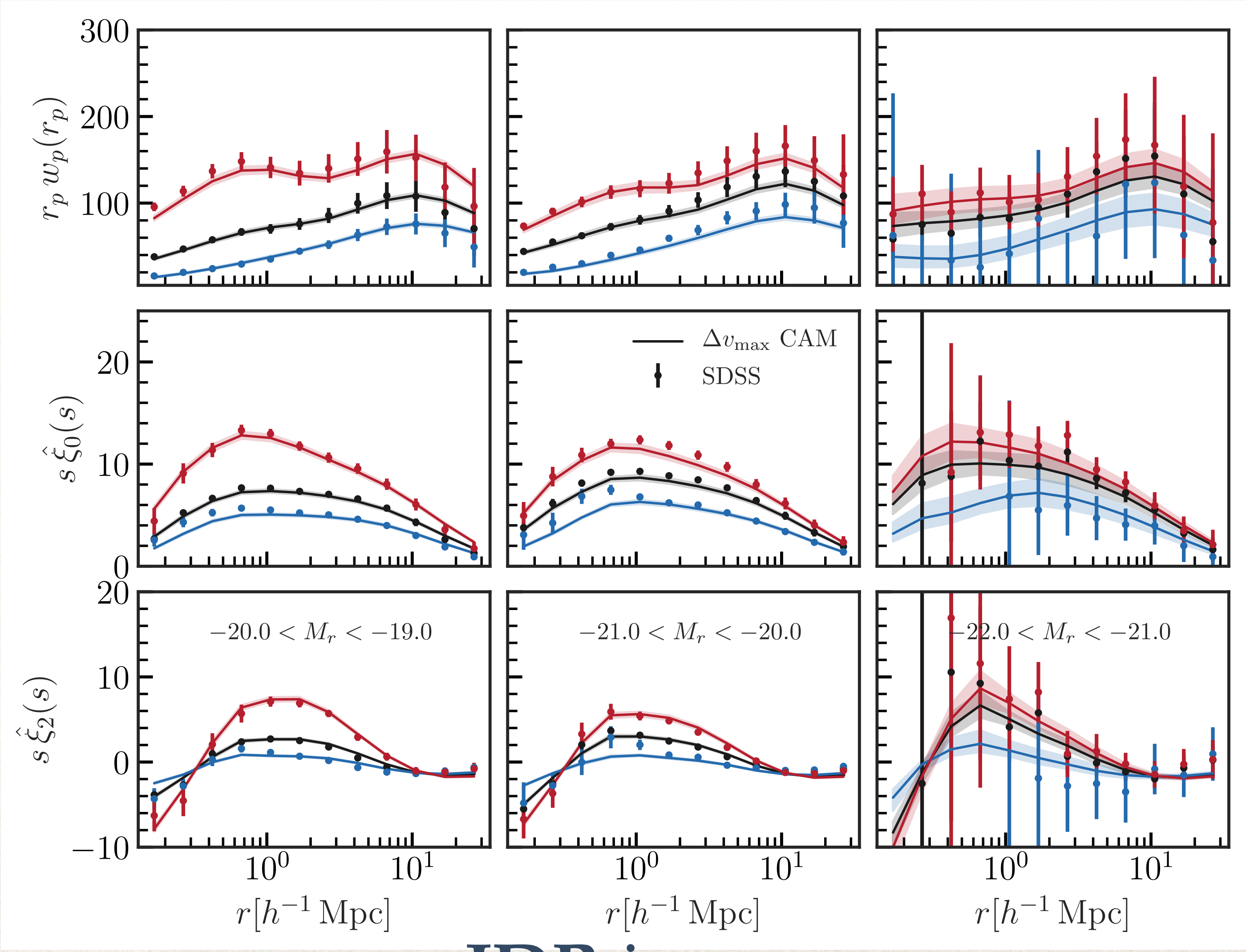
LRG selection

Zhou et al 2020





High-resolution N-body simulations combined with prescriptions for empirically matching color and magnitude distributions can reproduce these trends (SHAM)



JDR in prep.

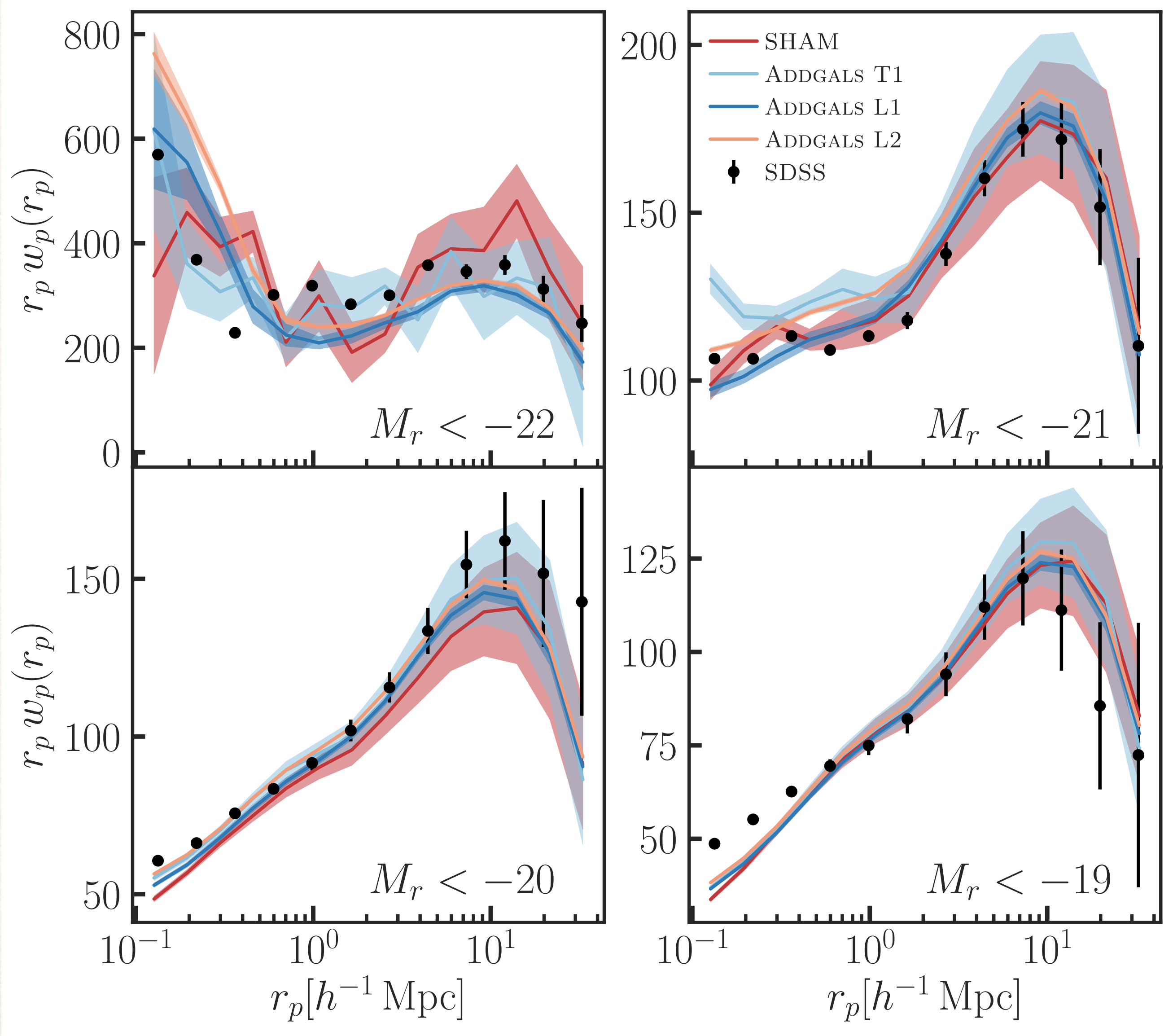
Testing x-corr analyses with synthetic skies

- **Goal:** Perform end-to-end tests going from simulated galaxy catalogs contaminated with realistic systematics to cosmological parameters.
  - Robustness to modeling assumptions/observational systematics
- **Requirements:**
  - Need realistic correlations between cosmological observables and quantities used for sample selection
  - High enough resolution to model all measurements accurately, e.g. clustering and lensing
  - **Many times the volume of the survey (must be inexpensive)**

ADDGALS

Adding Density Determined Galaxies to Lightcone Simulations

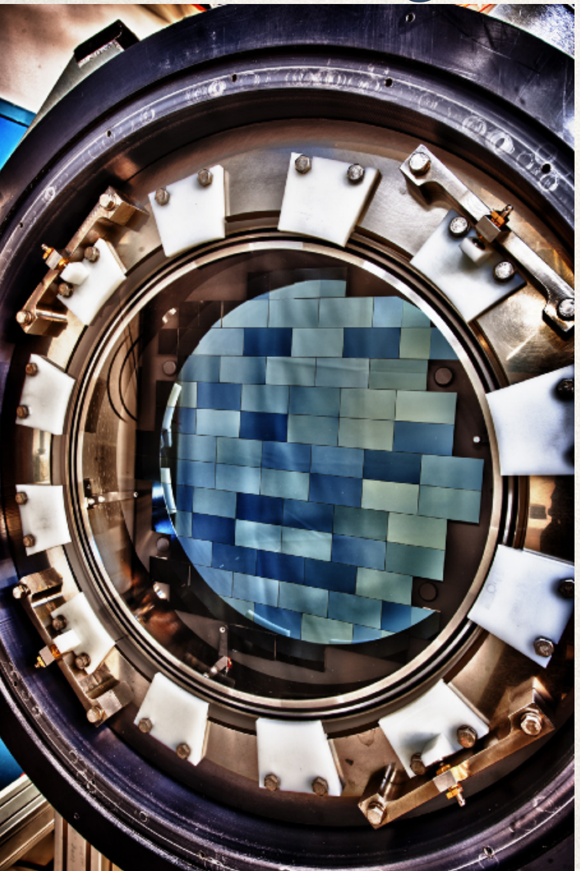
- Run abundance matching model on high resolution simulation
- Use machine learning to bootstrap information from high-res simulation into moderate resolution light-cones.

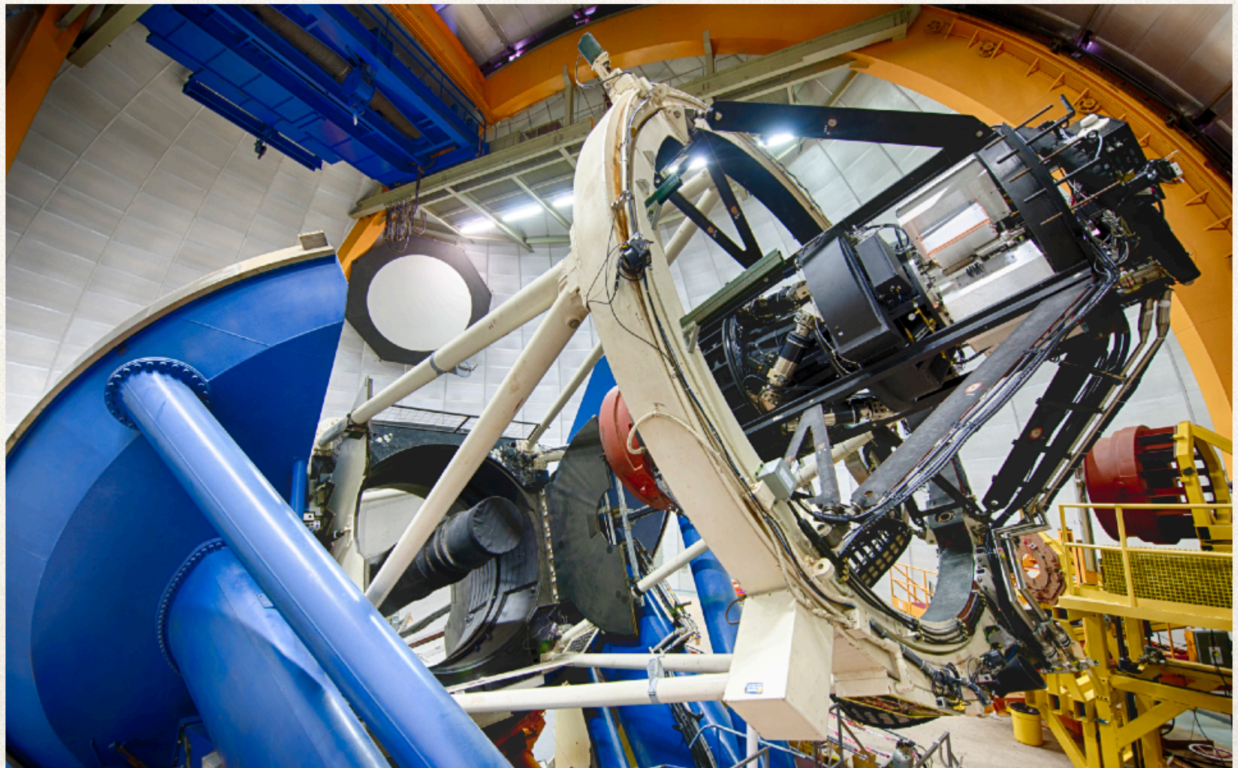


Wechsler, **JDR** in prep.

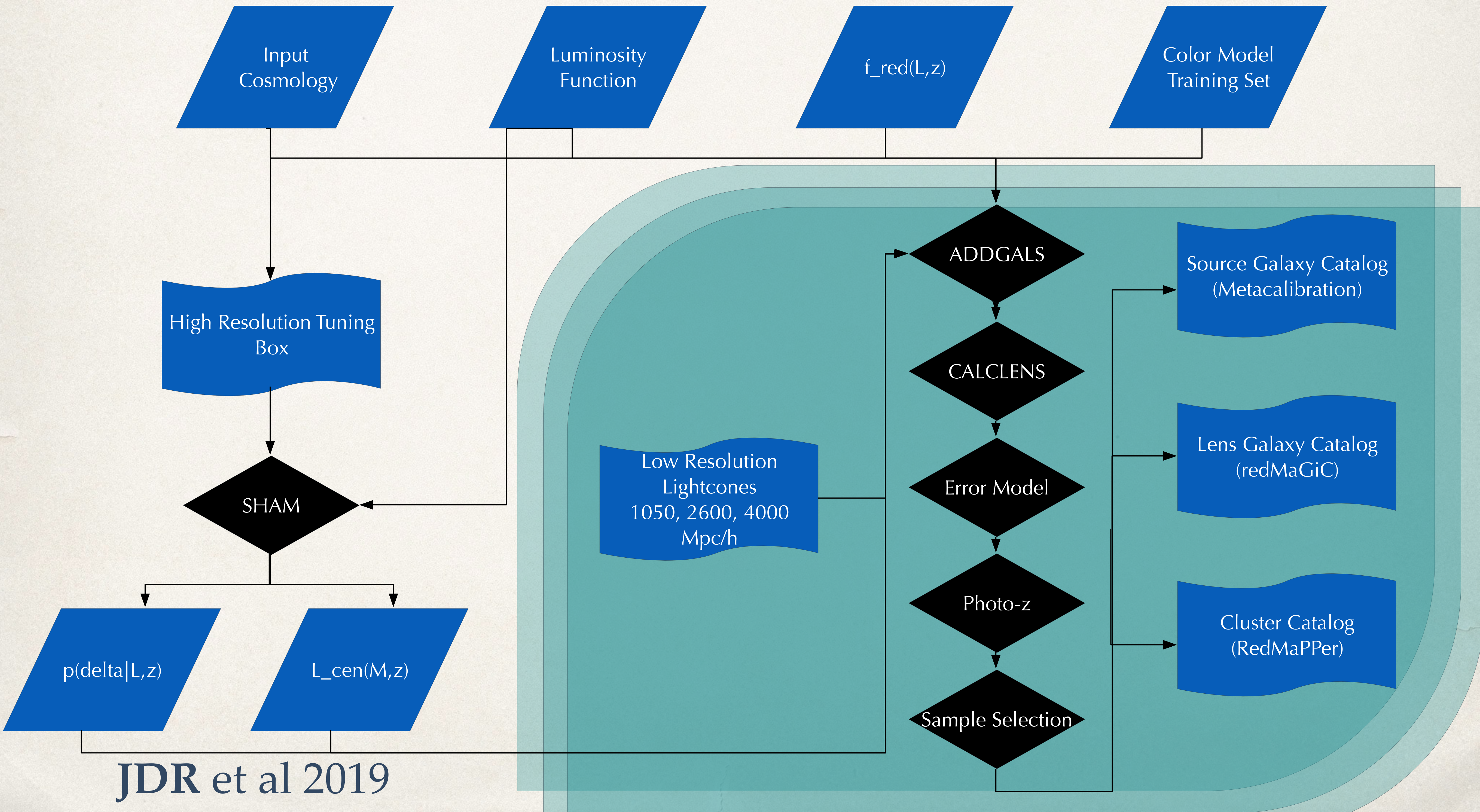
The Dark Energy Survey

- Imaging survey of the southern sky
  - 4m Blanco Telescope on Cerro Tololo, Chile
  - ~5000 sq. degrees
  - 5 bands: *grizy*
- Done taking 6 years of data, results published for first year (Y1) and working on analyzing first 3 years (Y3)



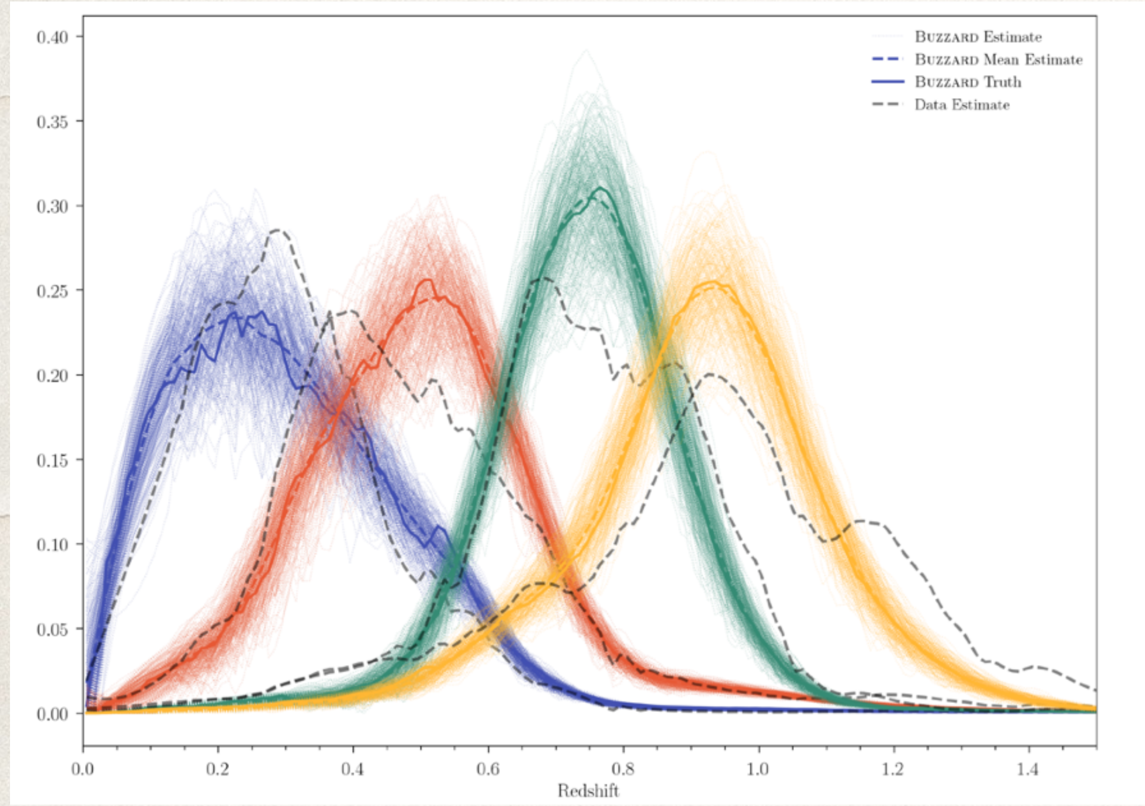


The Buzzard Flock

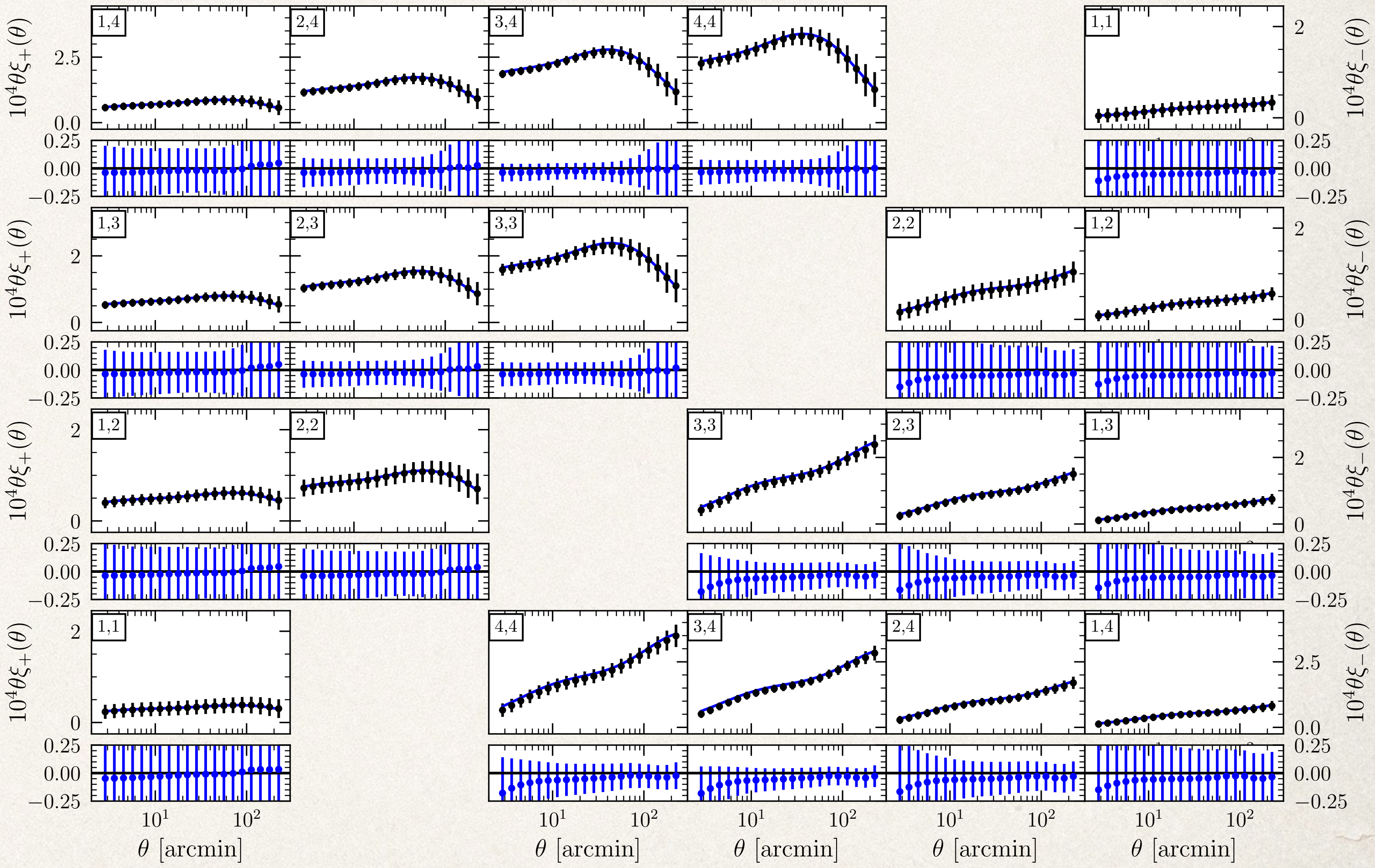


JDR et al 2019

Highlight: Validating the 3x2pt Pipeline



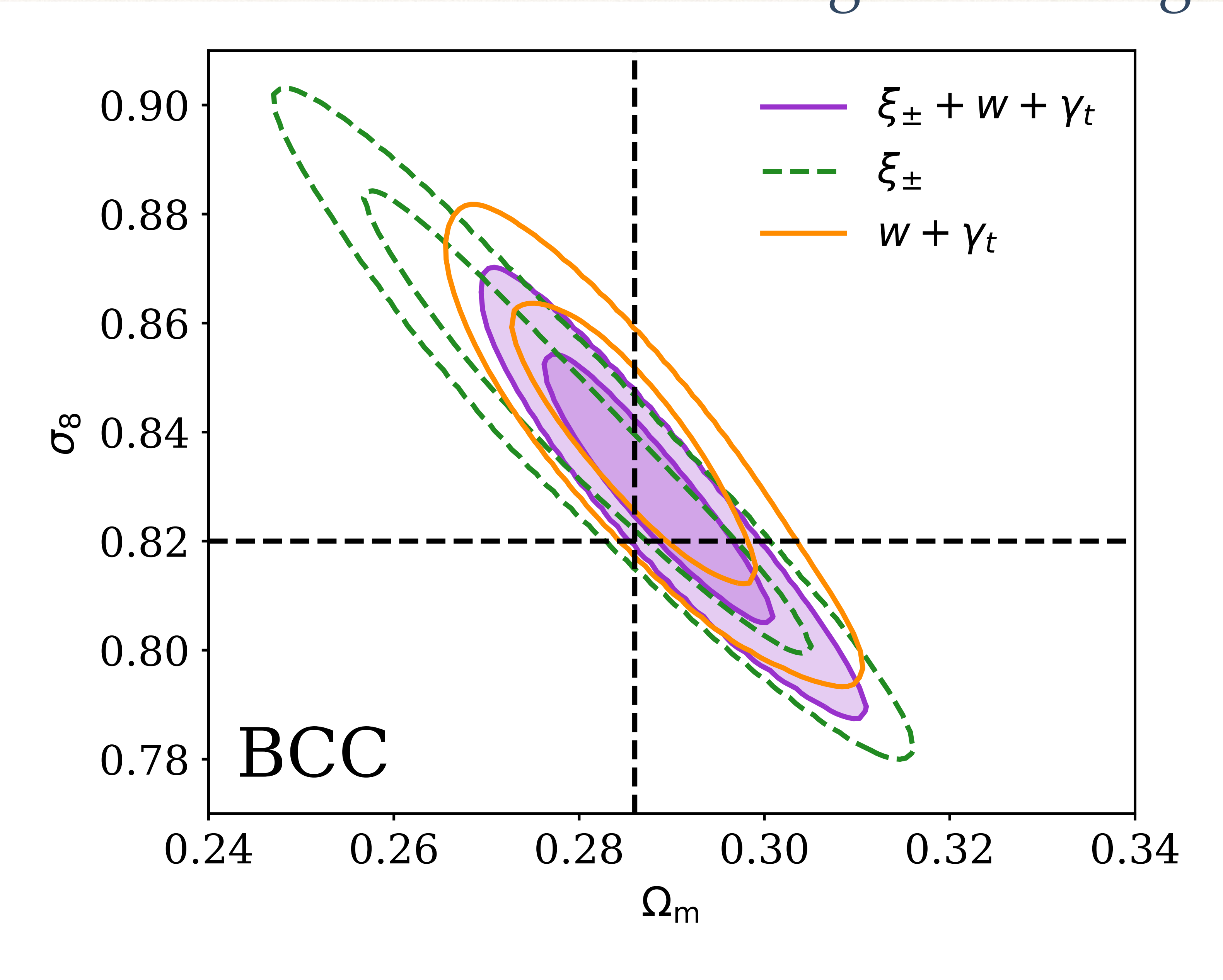
Myles, Alarcon et al. 2020
(incl. **JDR**)



JDR et al 2021 (under DES review)

Highlight: Validating the 3x2pt Pipeline

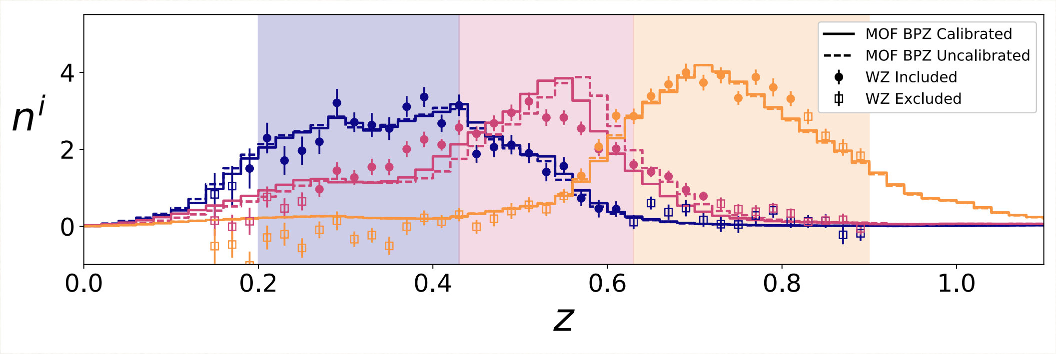
Constrained biases on inference to <1 sigma with high confidence



MacCrann, **JDR** et al. 2018

Buzzard sims used in a 11/14 of "DES Y1 Results"

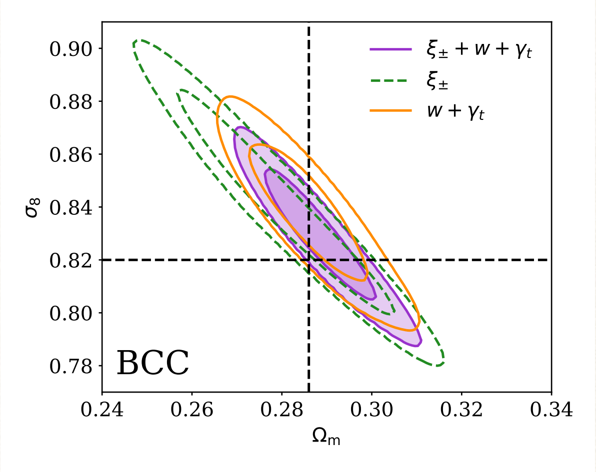
Redshift Estimation



*Gatti, Vielzeuf et al. (incl. **JDR**) 2018*

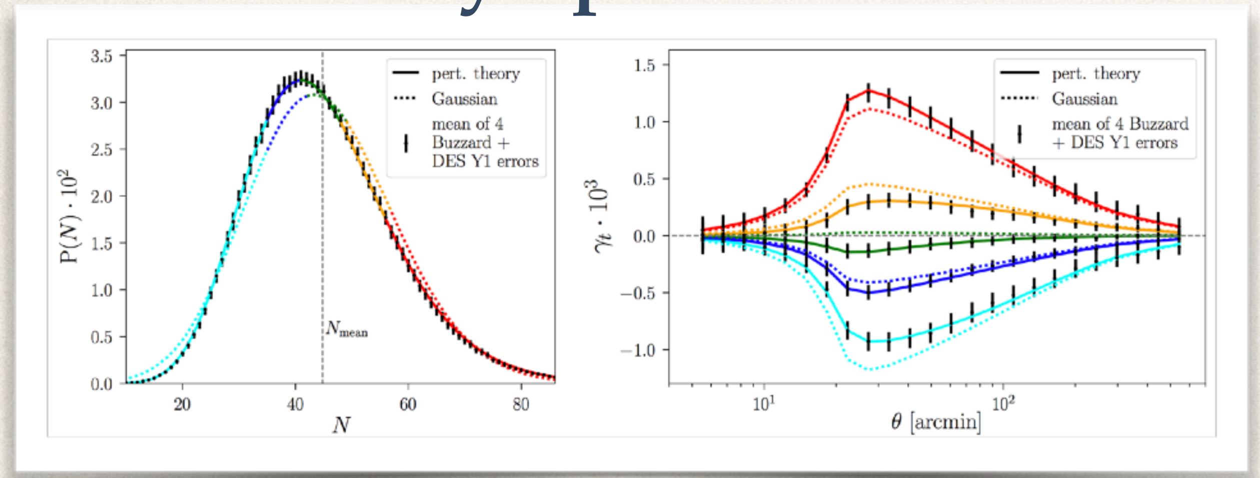
*Hoyle et al. (incl. **JDR**) 2018*

3x2pt Parameter Inference



*MacCrann, **JDR** et al. 2018*

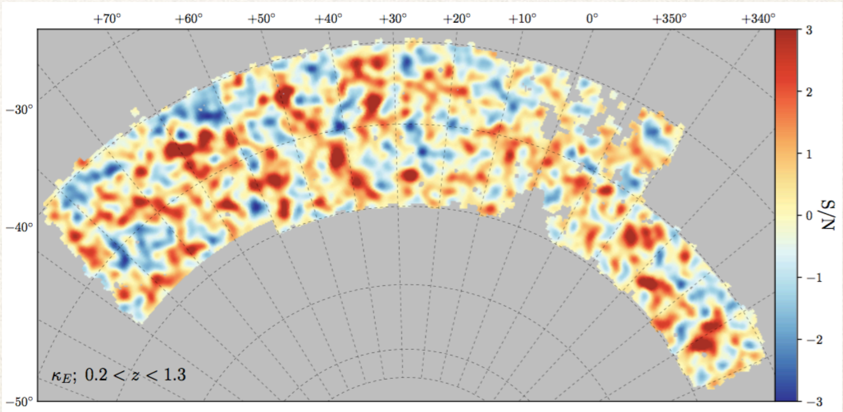
Density Split Statistics



*Gruen, Friedrich, Krause, **JDR** et al.*

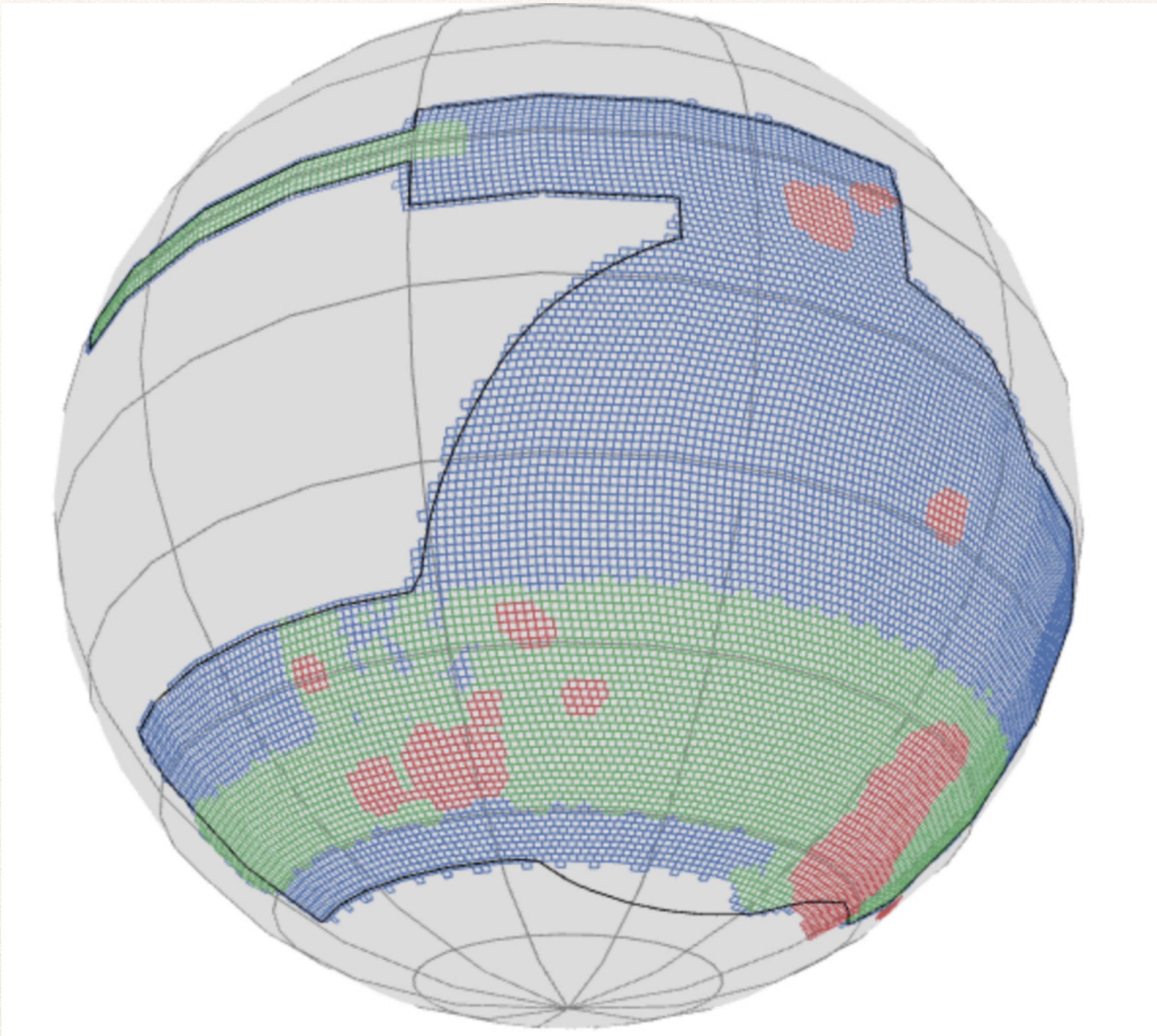
*Friedrich, Gruen, **JDR**, Krause et al.*

Mass Mapping



*Chang et al. (incl. **JDR**) 2018*

DES Y3:
Unblinded and writing papers



DESI is next!

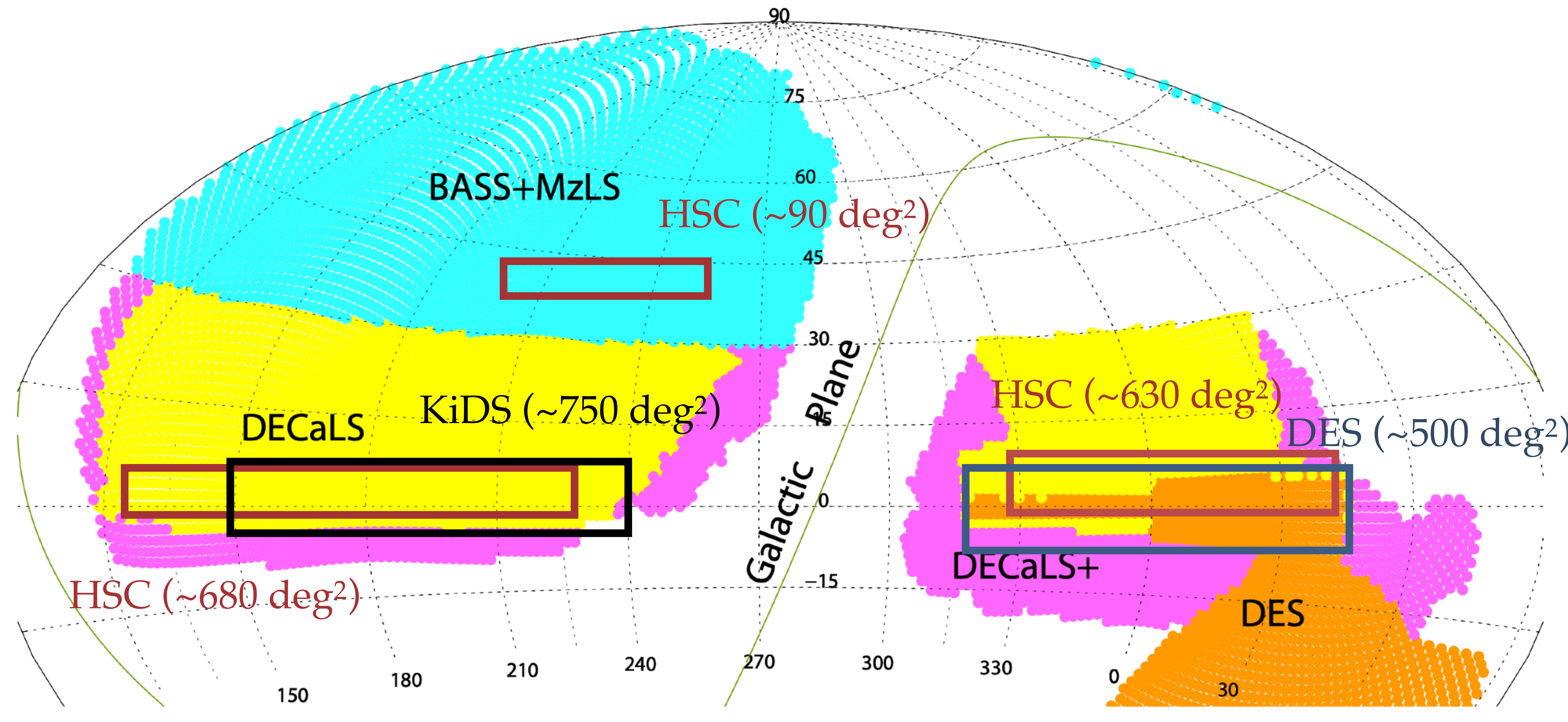


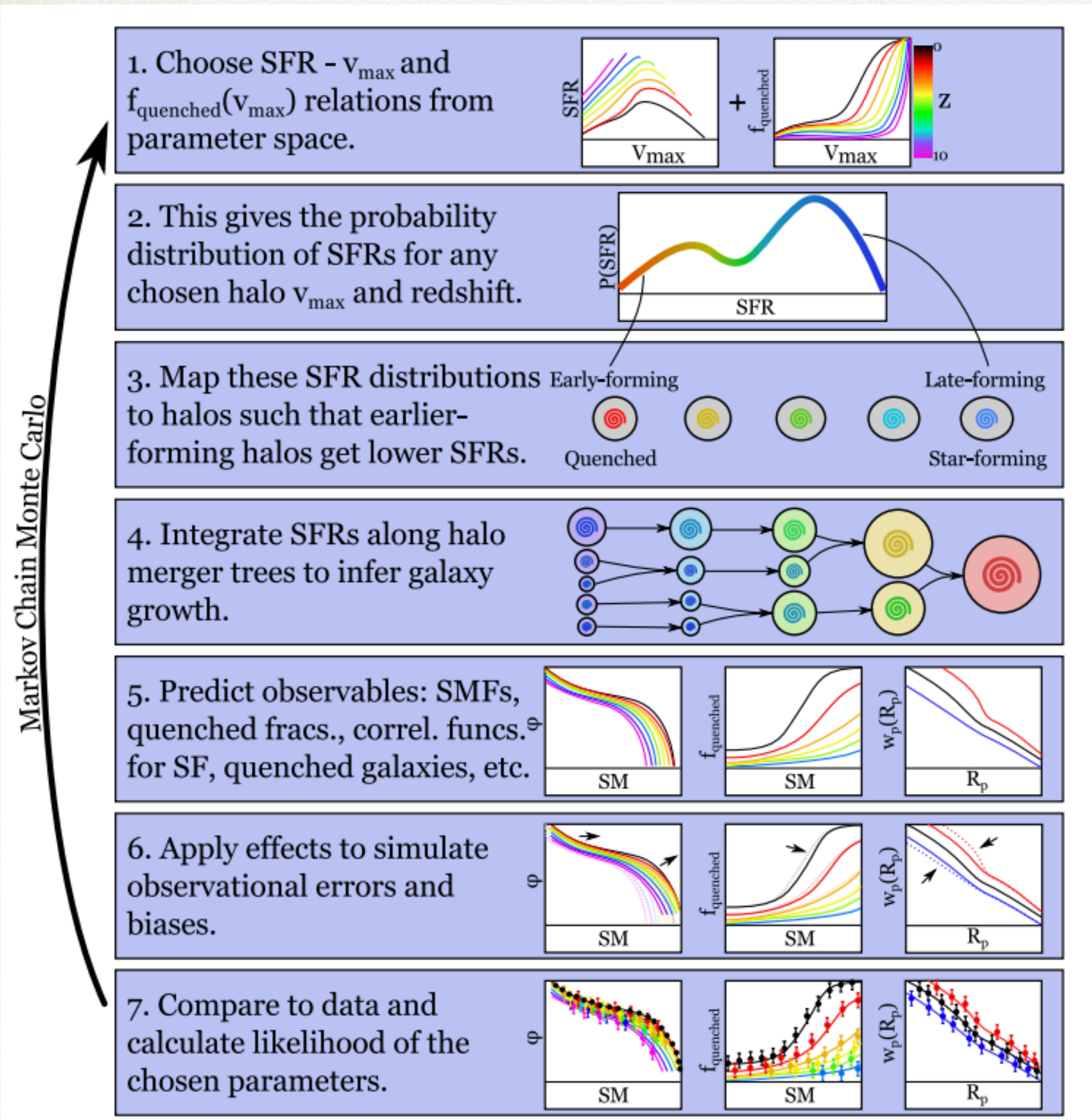
Figure from 1611.00036 w/ additions by C. Blake

Synthetic skies of the future

What do we need to test combined BAO/RSD/WL/CMB lensing analyses (e.g. DESI x LSST x SO/S4)

Real-data complexity	Included in Buzzard
Non-linear matter distribution & galaxy bias	✓
Velocity bias	✓
Galaxy lensing	✓
Photometric redshift biases	✓
Spatially varying photometric noise	✓
Realistic target selection/fiber assignment/redshift failure	(✓)
CMB lensing	(✓)
Baryonic effects on matter distribution	✗
Correlated CMB foregrounds	✗

Synthetic skies of the future



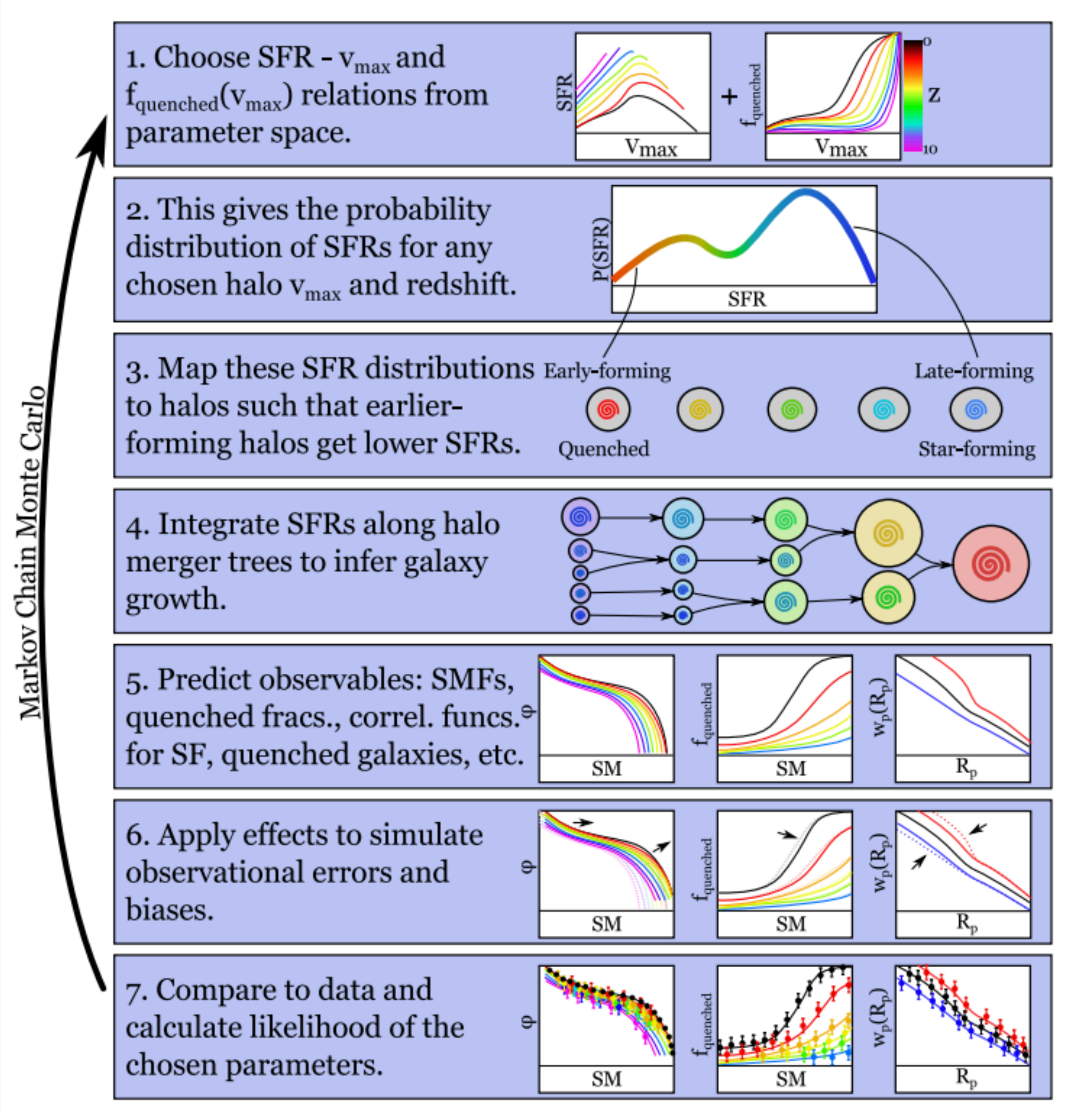
Behroozi et al. 2018

Similar simulation techniques are capable of including more relevant correlations

But…

they are extremely expensive to fit to data, slowing down model development.

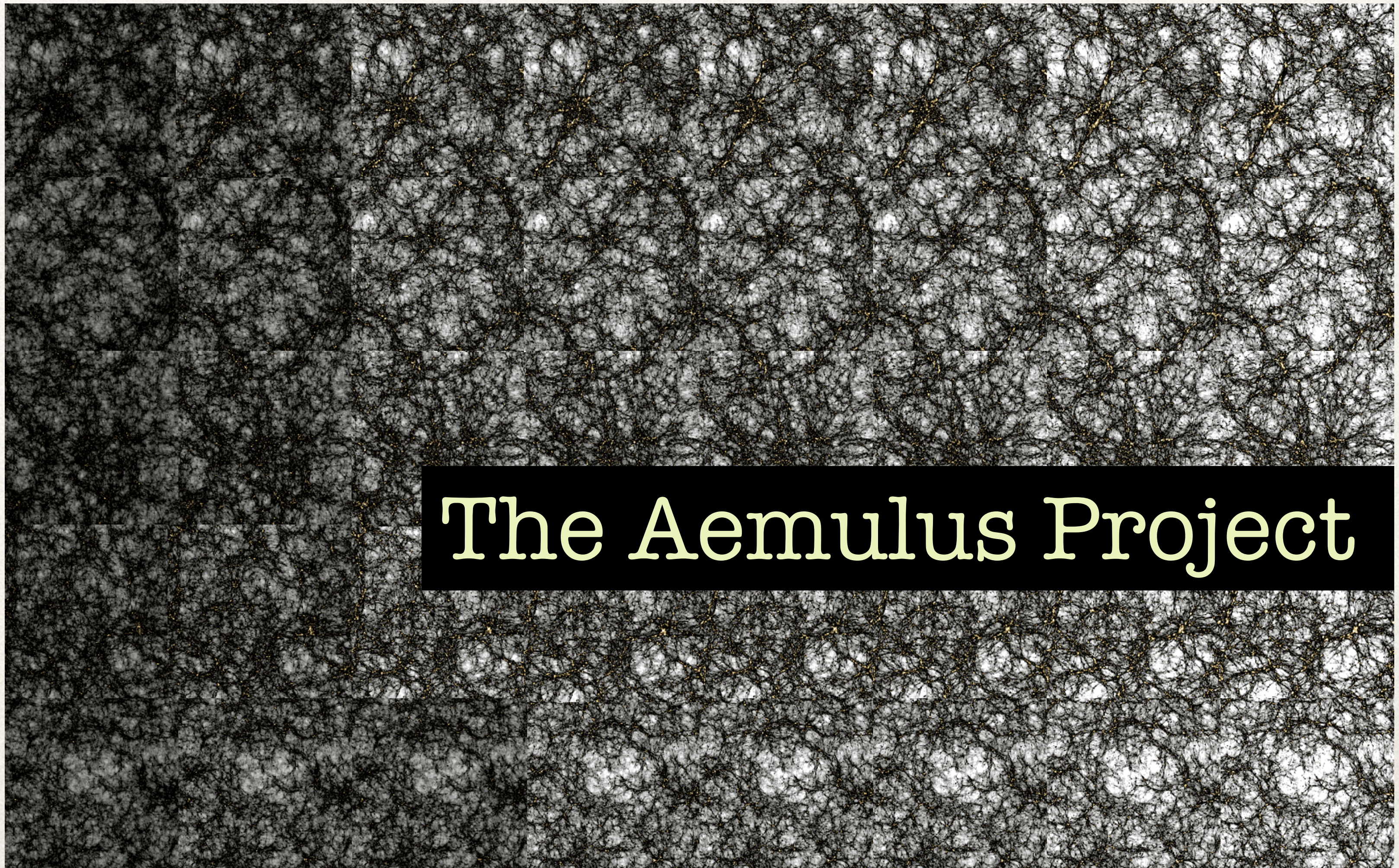
ML/HPC is driving progress!



Behroozi et al. 2018

Beginning to apply automatic differentiation/GPU acceleration techniques developed for ML applications —> Allows for better sampling methods (HMC) that will **drastically speed up model development cycles**, and allow for the incorporation of more realism.

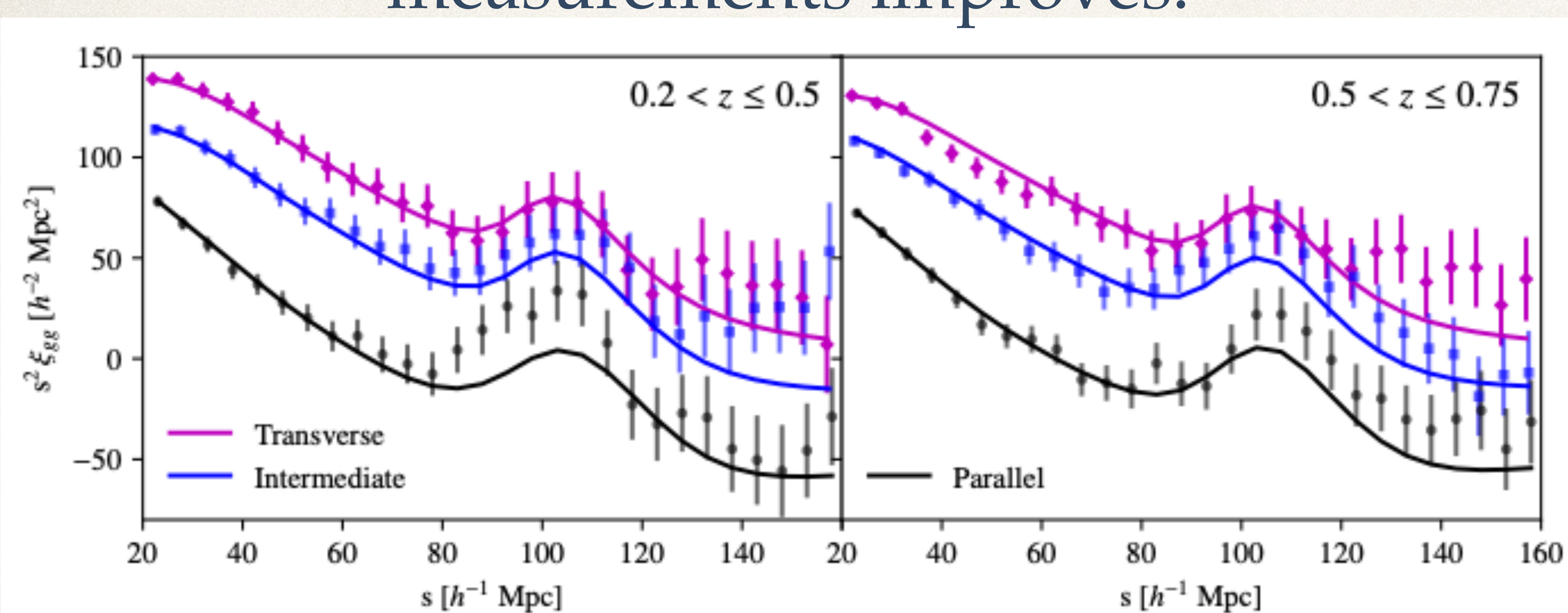
Part II: Accurate and precise models for lensing and clustering cross-correlations

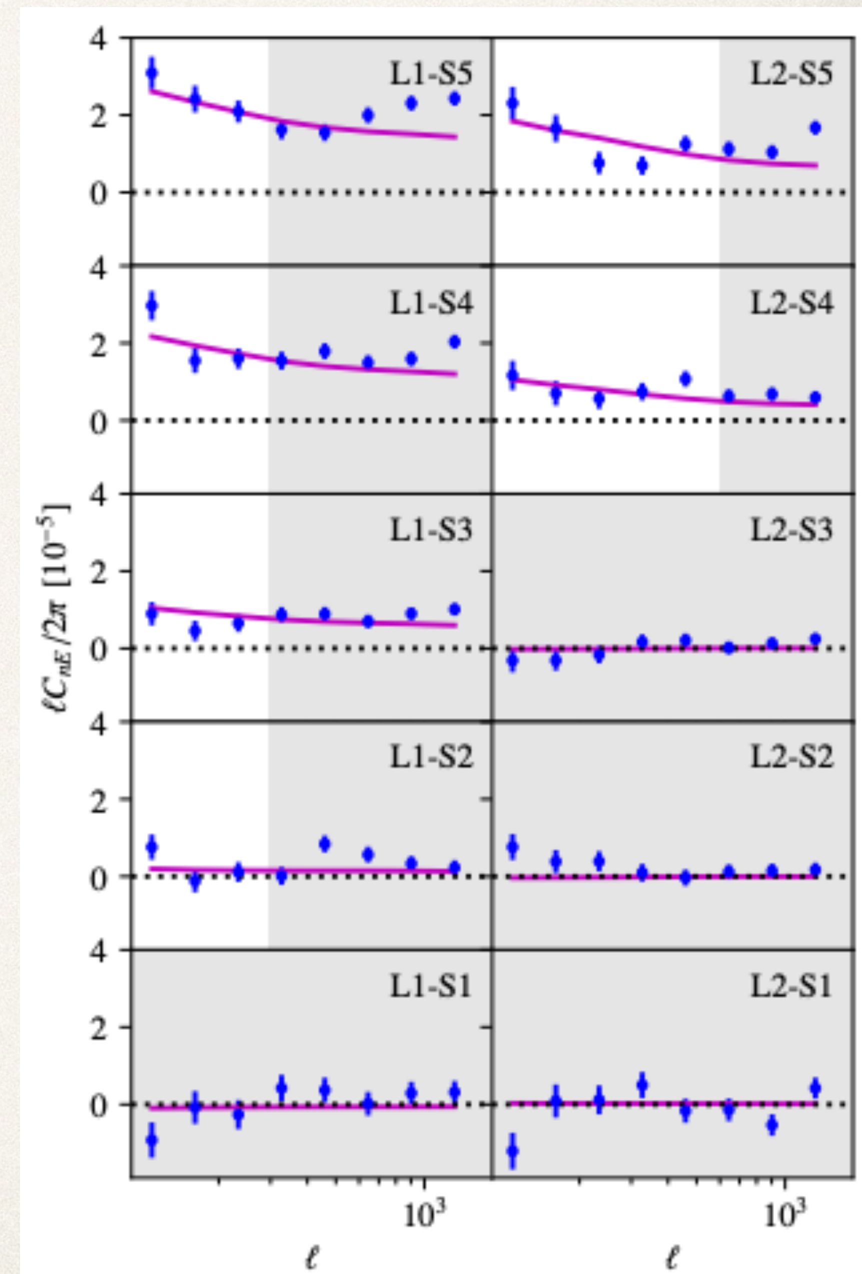


Limitations of current analyses

One of the main ingredients required to constrain cosmological parameters from galaxy clustering or lensing observations is a model for the galaxy power spectrum.

Current models require drastic scale cuts, and will only get worse as precision of measurements improves.

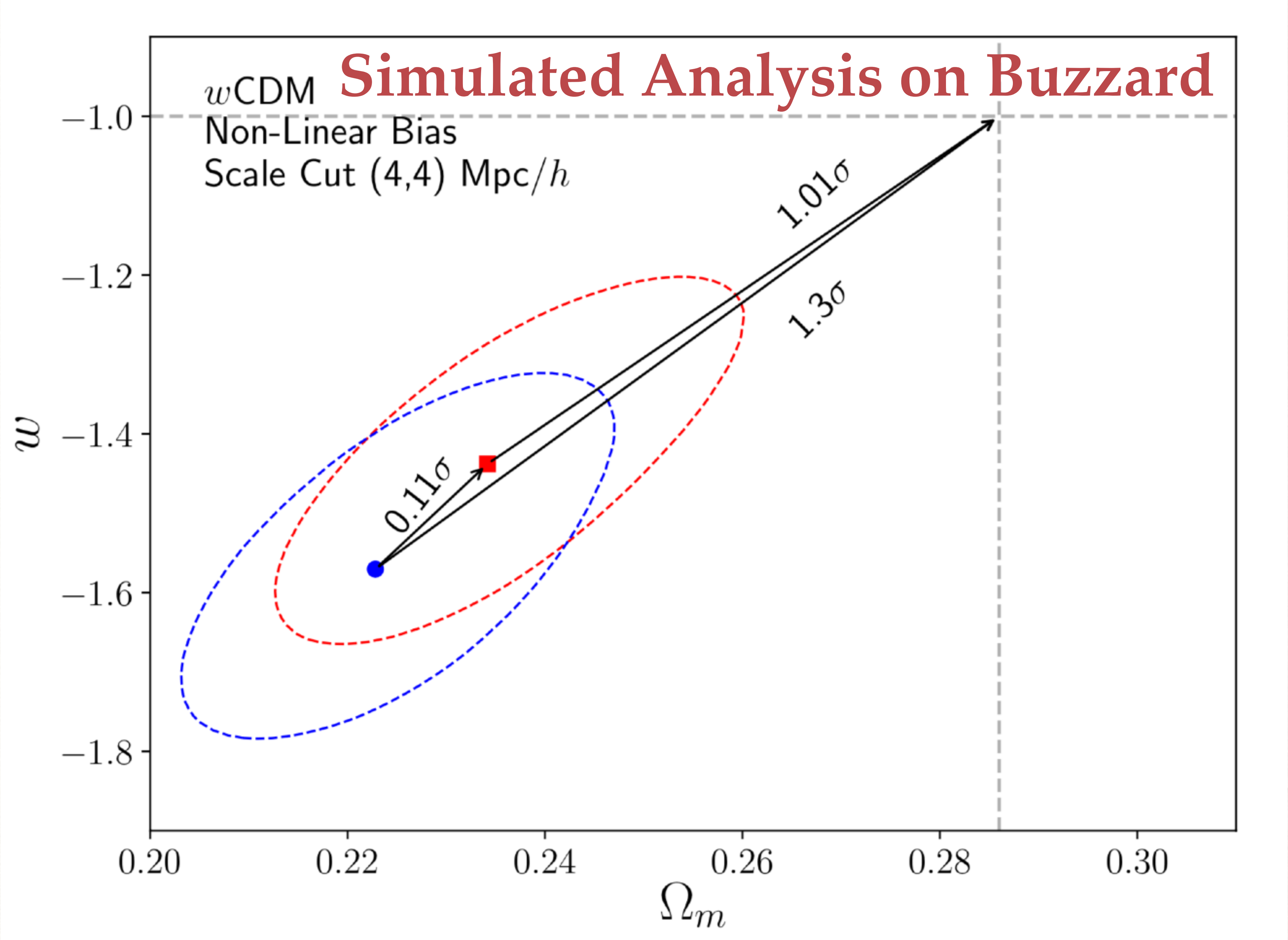




Heymans, Troster et al. 2020

PT in current analyses

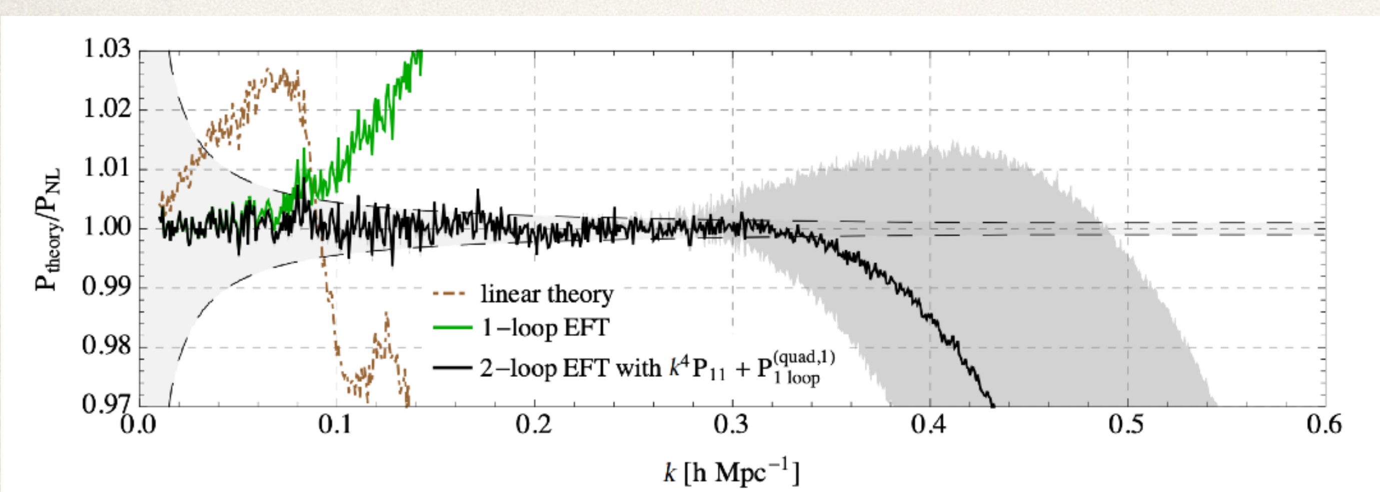
- Perturbation theory models designed for redshift space clustering are **not straight-forward** to apply to x-corr.
- Need to worry about projection effects.



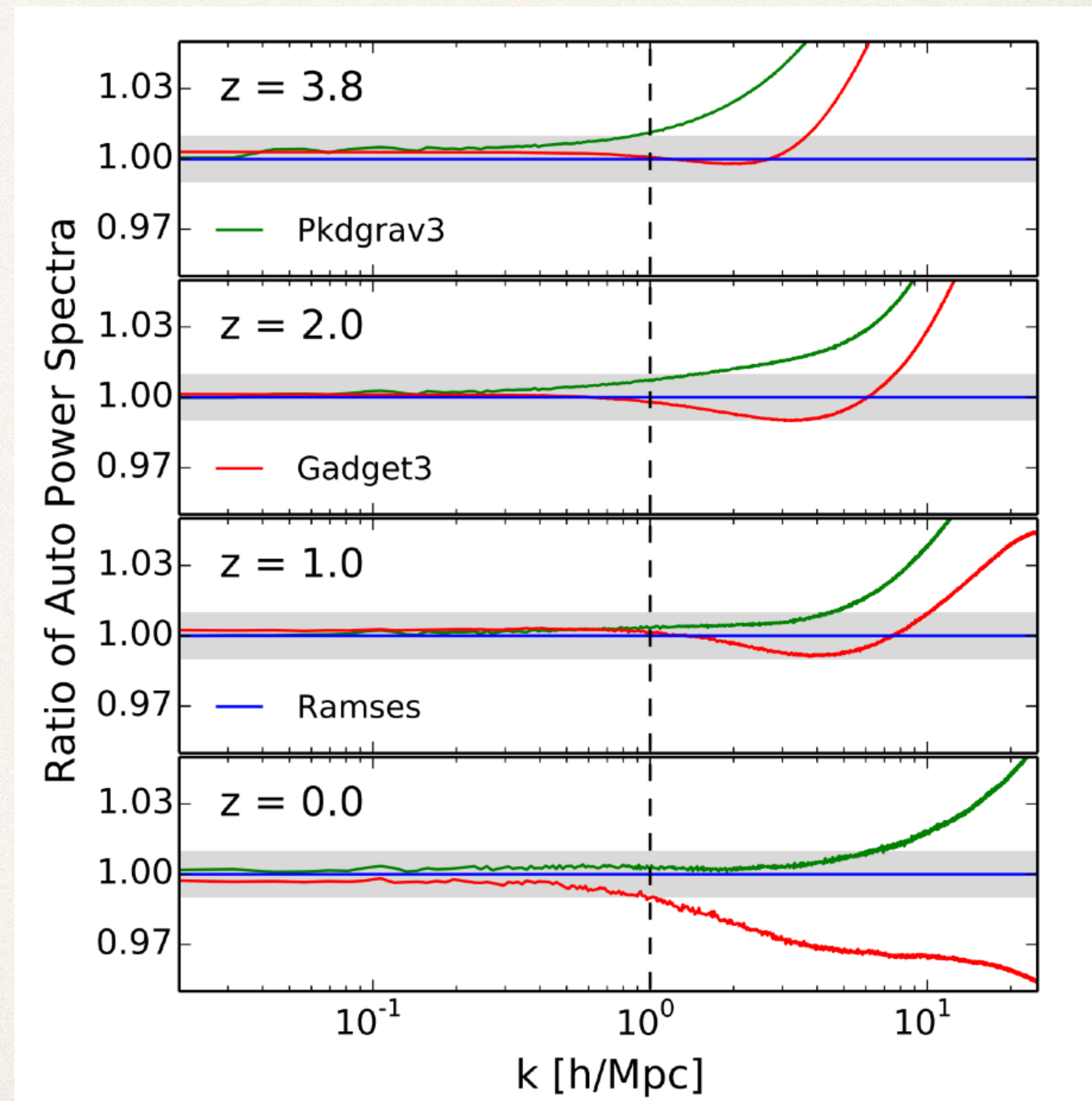
Pandey, JDR et al. 2021 (under DES review)

Simulation or Perturbation theory?

- Perturbation theory is general, but pushing to higher order yields diminishing returns
- Simulations of dark matter are converged to k~1.



Foreman et al 2015



Schneider et al 2016

Simulation or Perturbation theory?

Zhai et al. (incl. JDR) 2018

	Parameter	Meaning	Range
Cosmology	Ω_m	The matter energy density	[0.255, 0.353]
	Ω_b	The baryon energy density	[0.039, 0.062]
	σ_8	The amplitude of matter fluctuations on 8 h^{-1}Mpc scales.	[0.575, 0.964]
	h	The dimensionless Hubble constant	[0.612, 0.748]
	n_s	The spectral index of the primordial power spectrum	[0.928, 0.997]
	$w^{\dagger}$	The dark energy equation of state	[-1.40, -0.57]
	$N_{\rm eff}{}^{\dagger}$	The number of relativistic species	[2.62, 4.28]
	$\gamma_f{}^{\dagger}$	The amplitude of halo velocity field relative to wCDM+GR	[0.5, 1.5]
HOD	$\log M_{\rm sat}$	The typical mass scale for halos to host one satellite	[13.8, 14.5]
	α	The power-law index for the mass dependence of the number of satellites	[0.2, 1.8]
	$\log M_{\rm cut}$	The mass cut-off scale for the satellite occupatioin function	[10.0, 13.7]
	$\sigma_{\log M}$	The scatter of halo mass at fixed galaxy luminosity	[0.05, 0.6]
	$\eta_{\rm con}{}^{\dagger}$	The concentration of satellites relative the dark matter halo	[0.2, 2.0]
	$\eta_{\rm vc}{}^{\dagger}$	The velocity bias for central galaxies	[0.0, 0.7]
	$\eta_{\rm vs}{}^{\dagger}$	The velocity bias for satellite galaxies	[0.2, 2.0]

How to know when to stop?

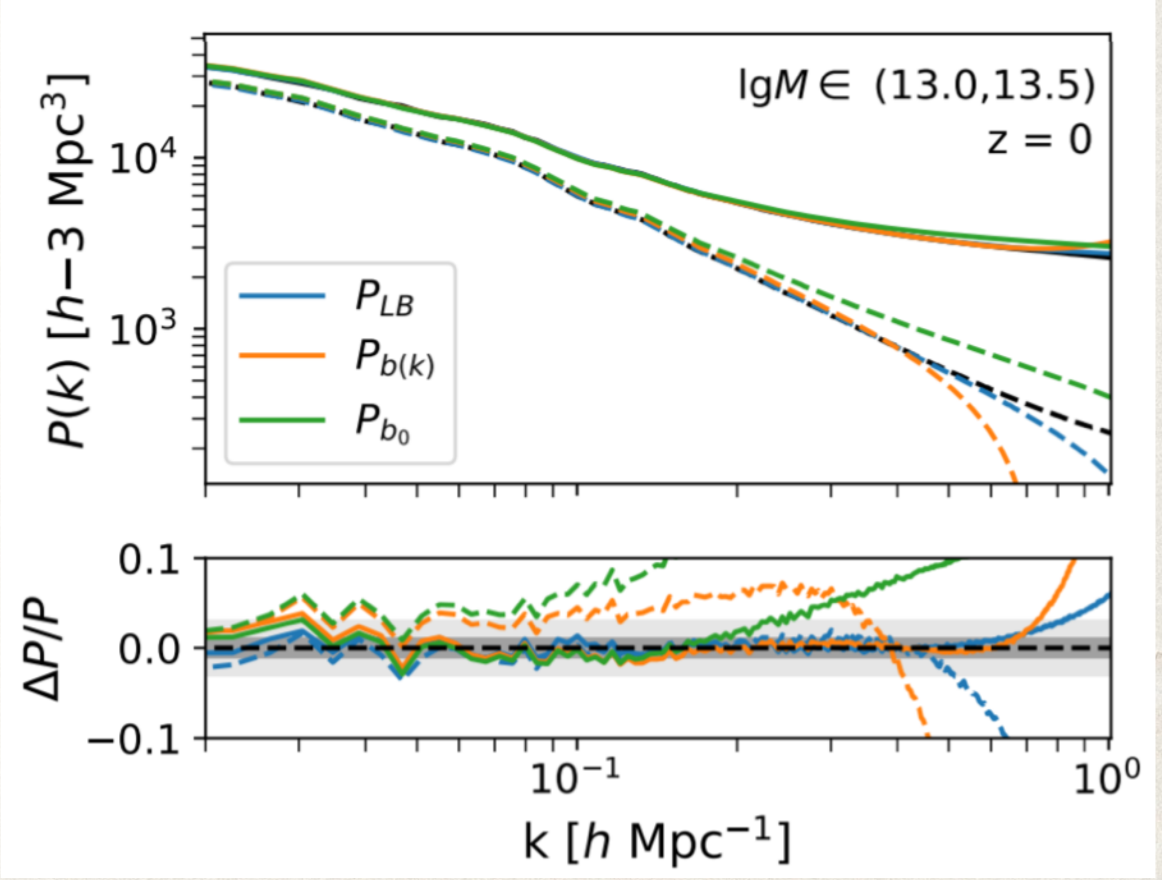
Simulation **and** Perturbation theory!

Rather than halo based models, use symmetries based expansions popular in perturbation theory models. Only keep fields allowed by rotation/ Galilean invariance and the equivalence principle

$$\delta_g(\mathbf{q}) = F[\delta_L(\mathbf{q}), \delta_L^2(\mathbf{q}), s^2(\mathbf{q}), \nabla^2\delta_L(\mathbf{q})]$$

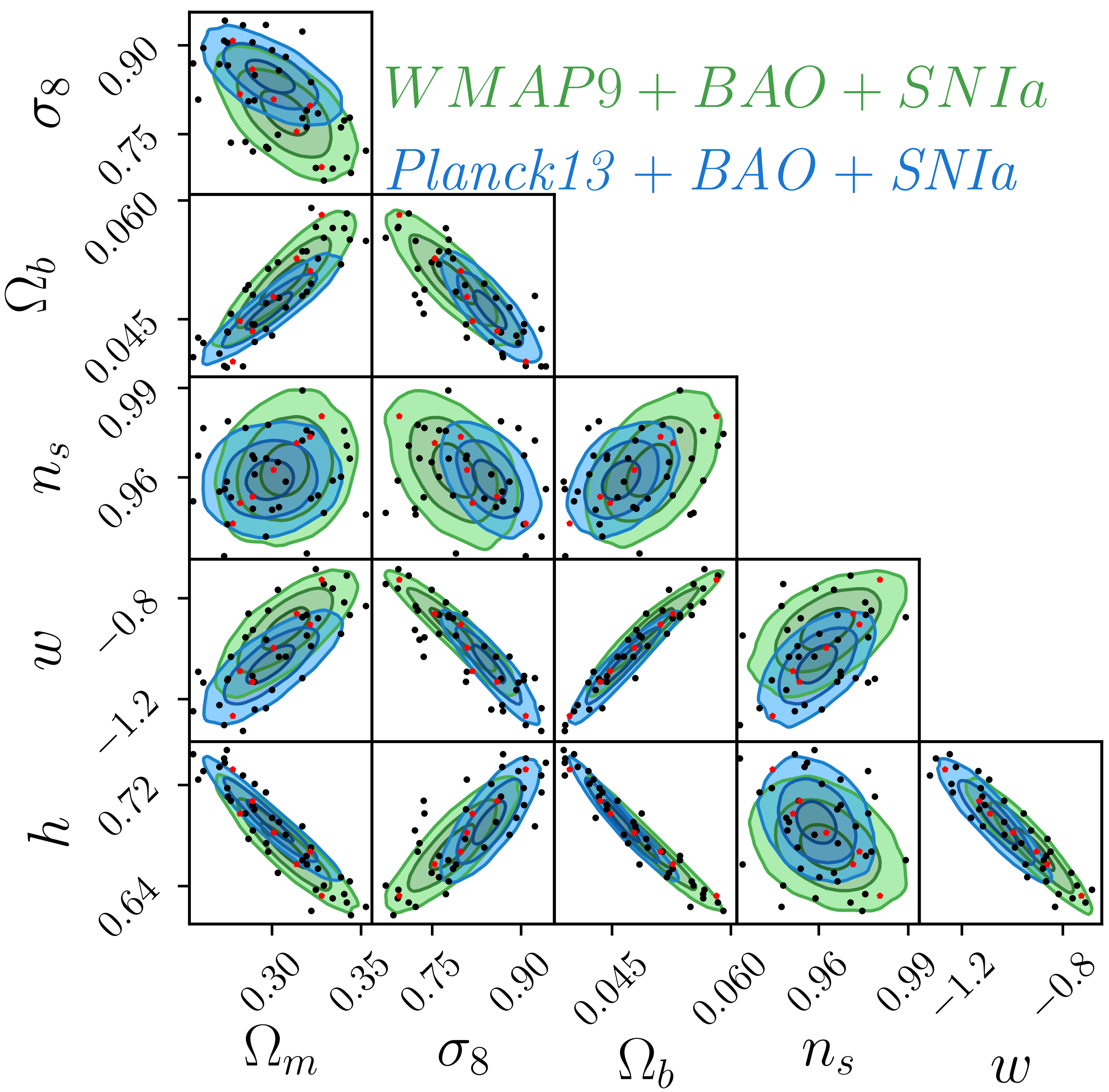
Proof of concept showed that this idea extends the reach of PT by a factor of ~2 in scale.

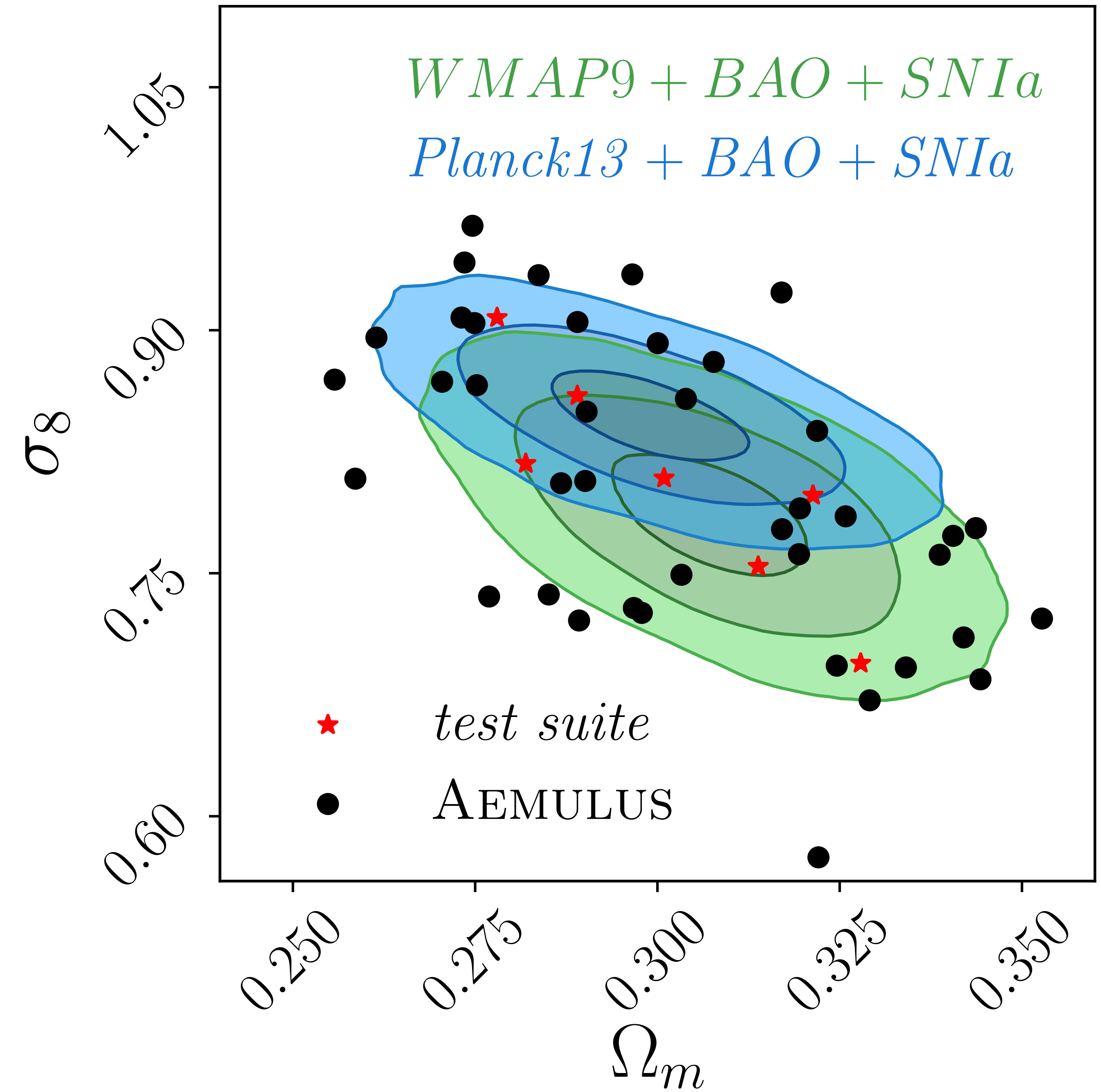
In order to apply to data we need a model for the cosmology dependence of these spectra.



Modi, Chen & White 2019

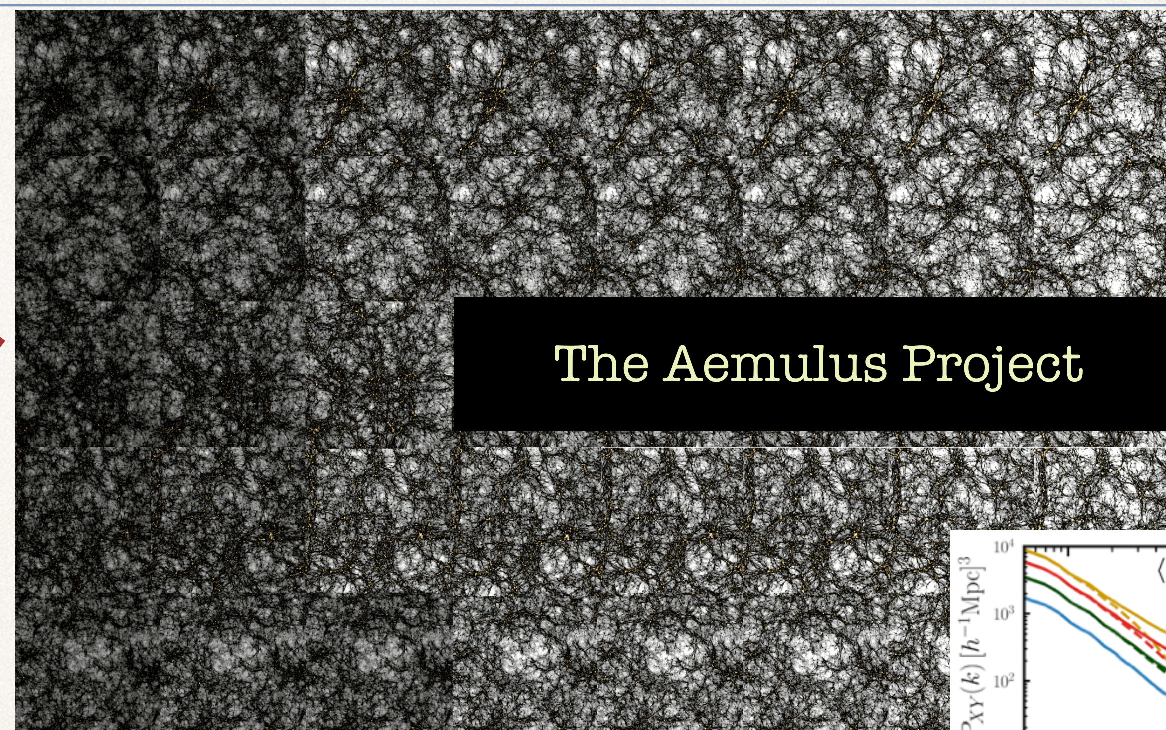
Sampling Cosmological Parameter Space



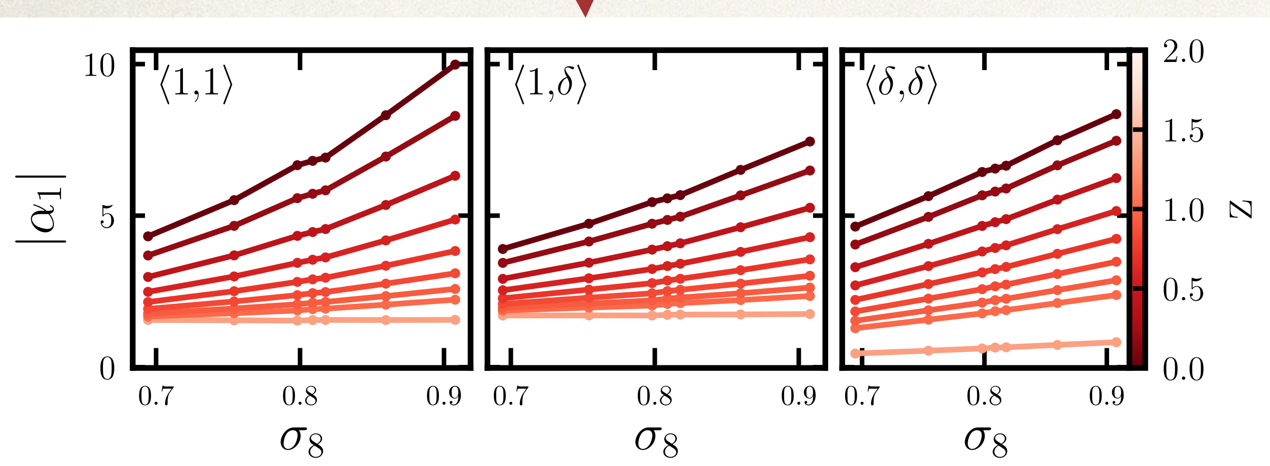


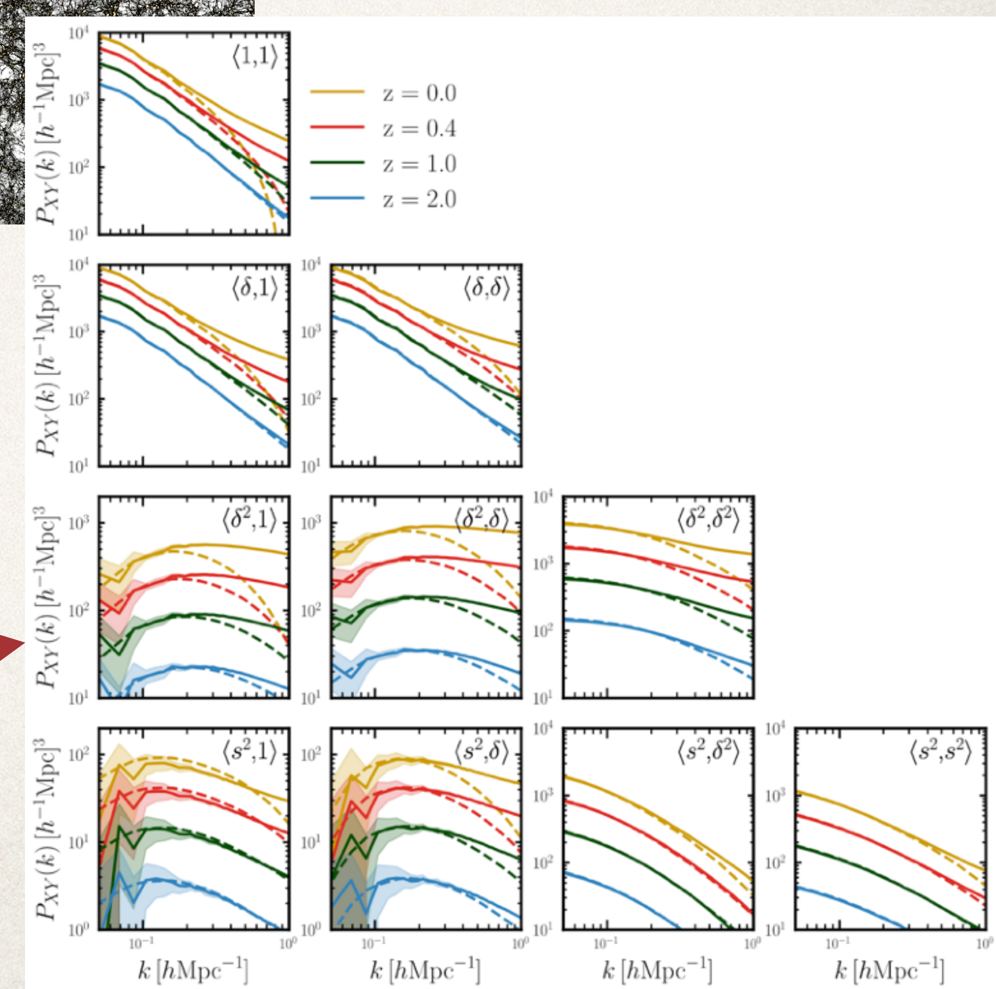
JDR et al. 2018

Emulating Component Spectra

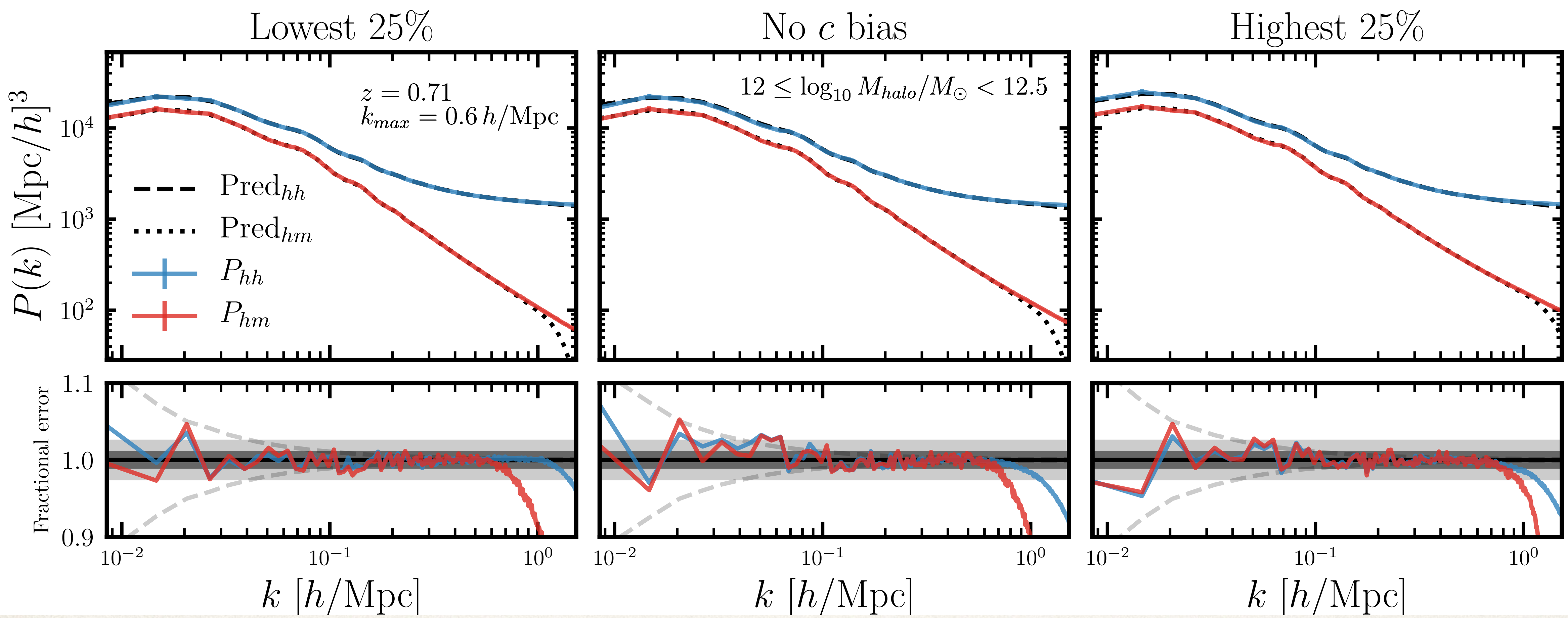


Kokron, JDR, Chen, White & Wechsler
arxiv:2101.11014





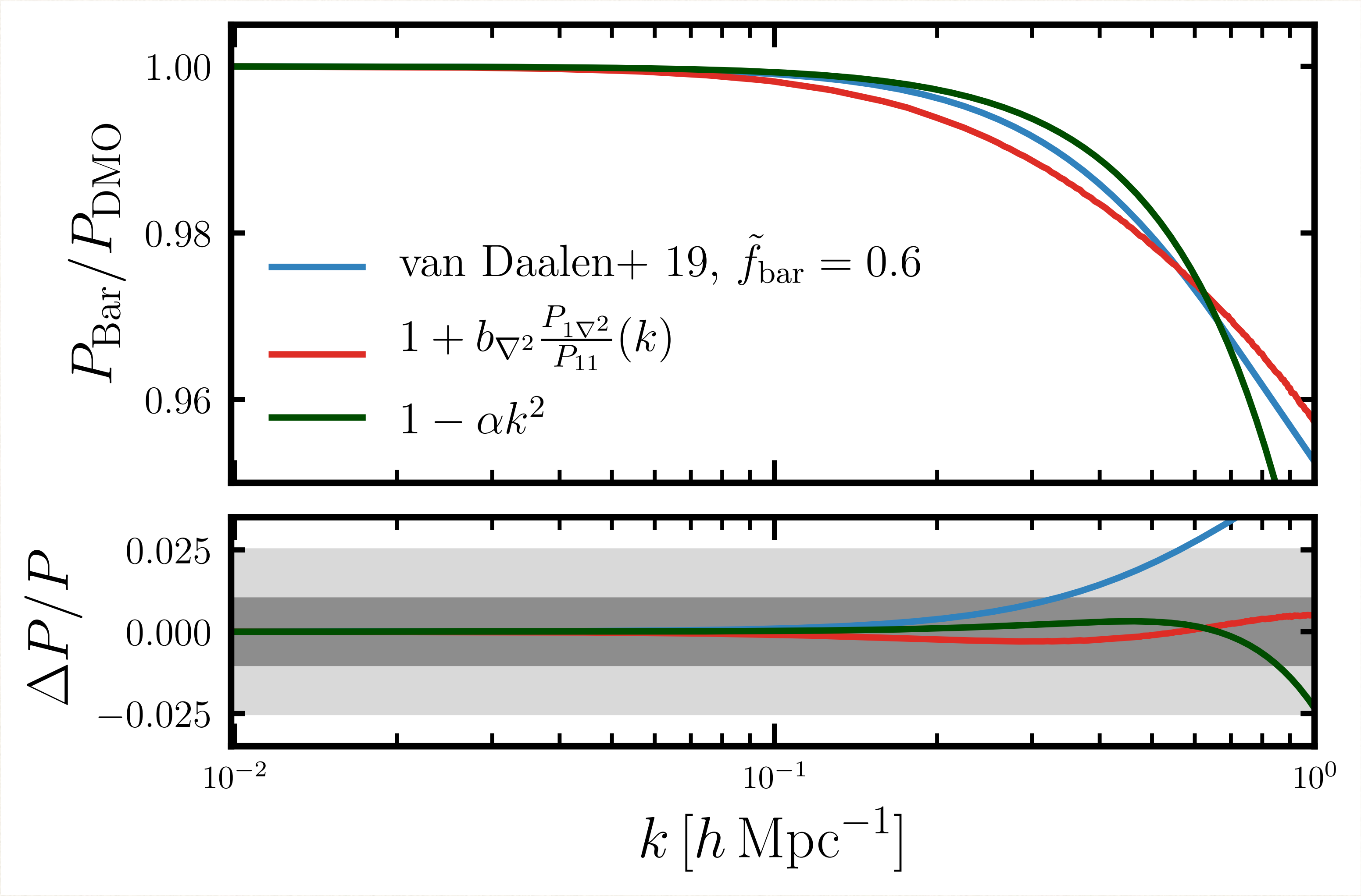
Ability to fit complex samples



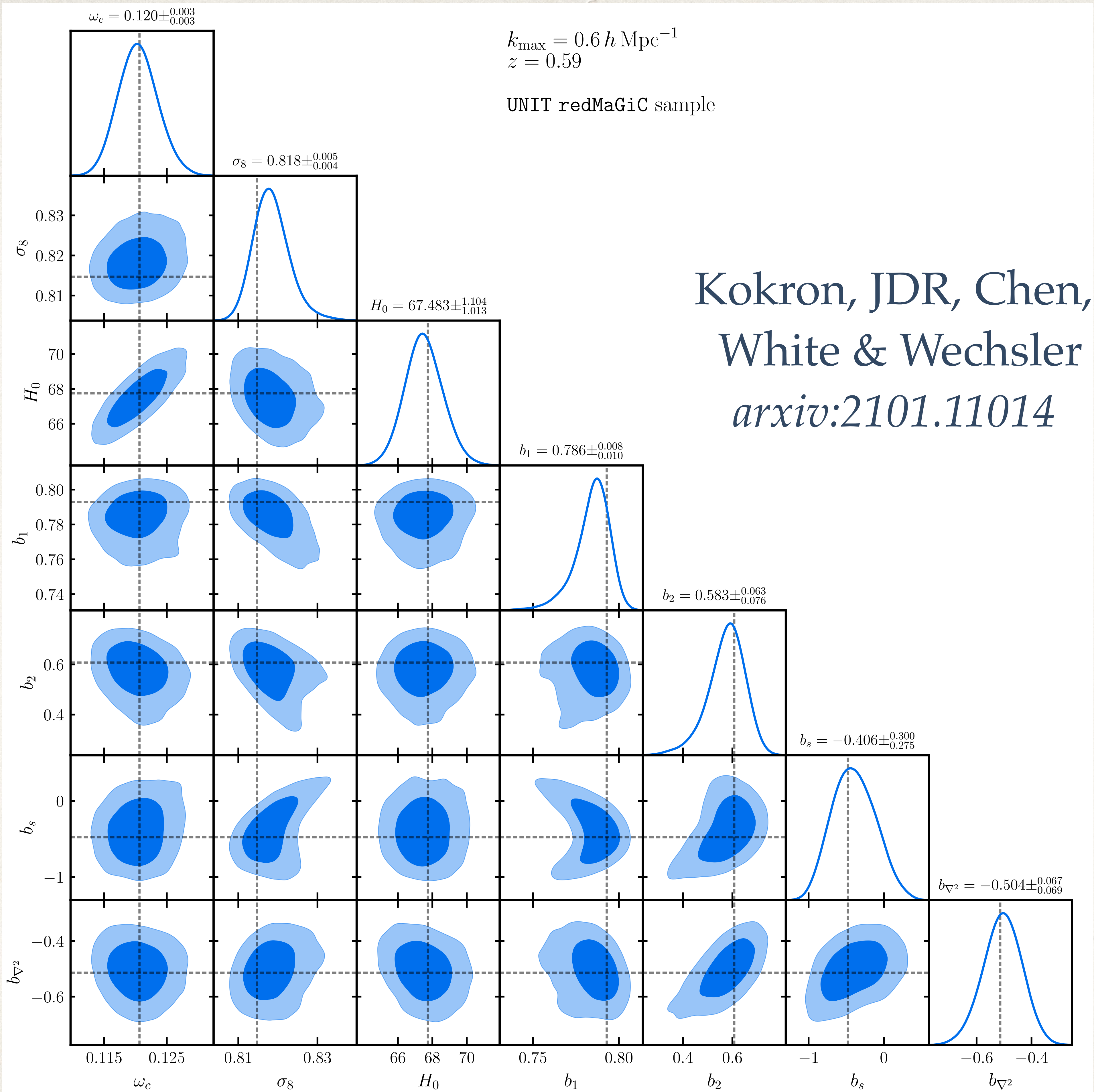
This model can handle modest amounts of assembly bias.

Fits to kmax~0.6 lead to agreement better than 1% out to k~1

“Built in” accounting of baryonic effects



The del^2 term in the bias expansion is a good approximation to the effect of baryons on the matter distribution at intermediate scales.



Kokron, JDR, Chen, White & Wechsler

arxiv:2101.11014

Future Directions for Hybrid Simulation/PT models

- Apply to joint (CMB and galaxy) lensing and RSD analyses in current data
- Extensible to higher order statistics / field level analyses —> better initial density reconstruction for BAO?

Emulators in DESI

As DESI emulator sub-working group chair, I am helping leading the effort to prepare for cross-correlation analyses:

- Emulator mock challenge
  - Focused on modeling systematics for combined probe analyses.
  - Will provide insight into strengths and weaknesses different emulation techniques
- Analysis pipeline development
  - Need to integrating emulators into frameworks that exist for joint-probe analyses (cosmosis/cobaya etc)
  - Think about optimal sampling methods to expedite analysis

Summary

Cross correlation science in DESI will lead to exciting insight about dark energy!

- We have designed an algorithm that allows us to produce realistic suites of galaxy catalogs.
  - Used to test DES Y1/Y3 and DESI cosmology analyses
  - ML inspired techniques will continue to improve realism
- Combinations of perturbation theory and simulation lead to a hybrid model that is the best of both worlds
  - Emulator for lensing and clustering x-corr ready to be applied to current data
  - Plethora of future directions to model higher order stats.

Thanks!