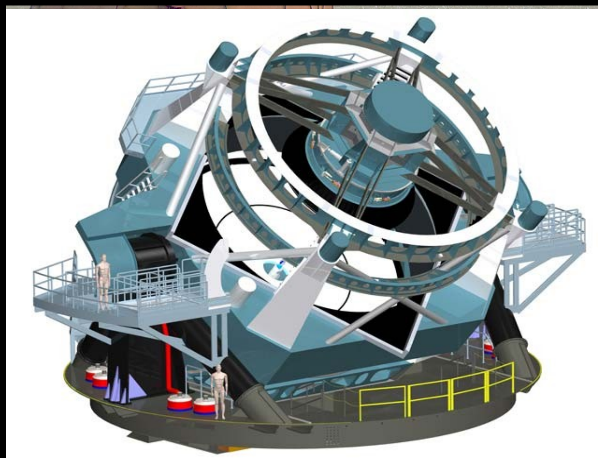
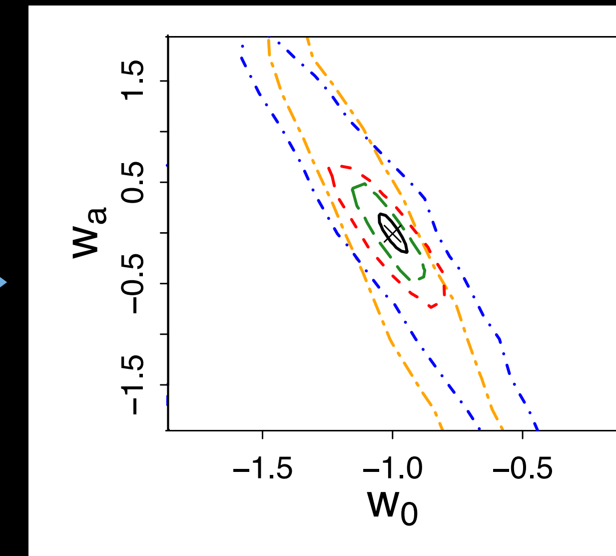




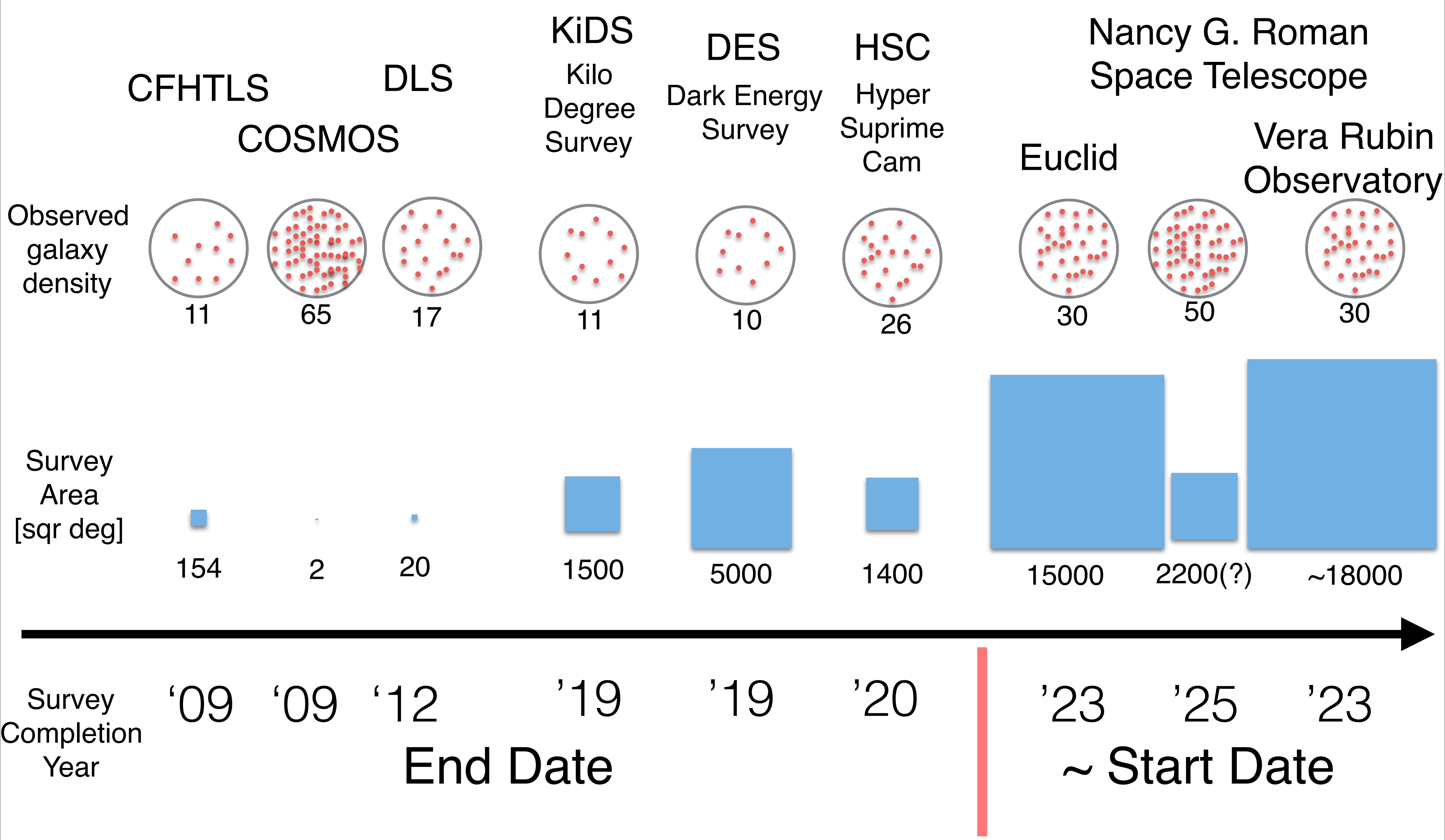
Large-Scale Structure Cosmology in the Systematics-Limited Regime

with Xiao Fang, Hung-Jin Huang, Vivian Miranda, Shivam Pandey, Chun-Hao To
and the Dark Energy Survey Collaboration, Vera Rubin Observatory Dark Energy Science Collaboration, and Roman Space Telescope Science Investigation Teams

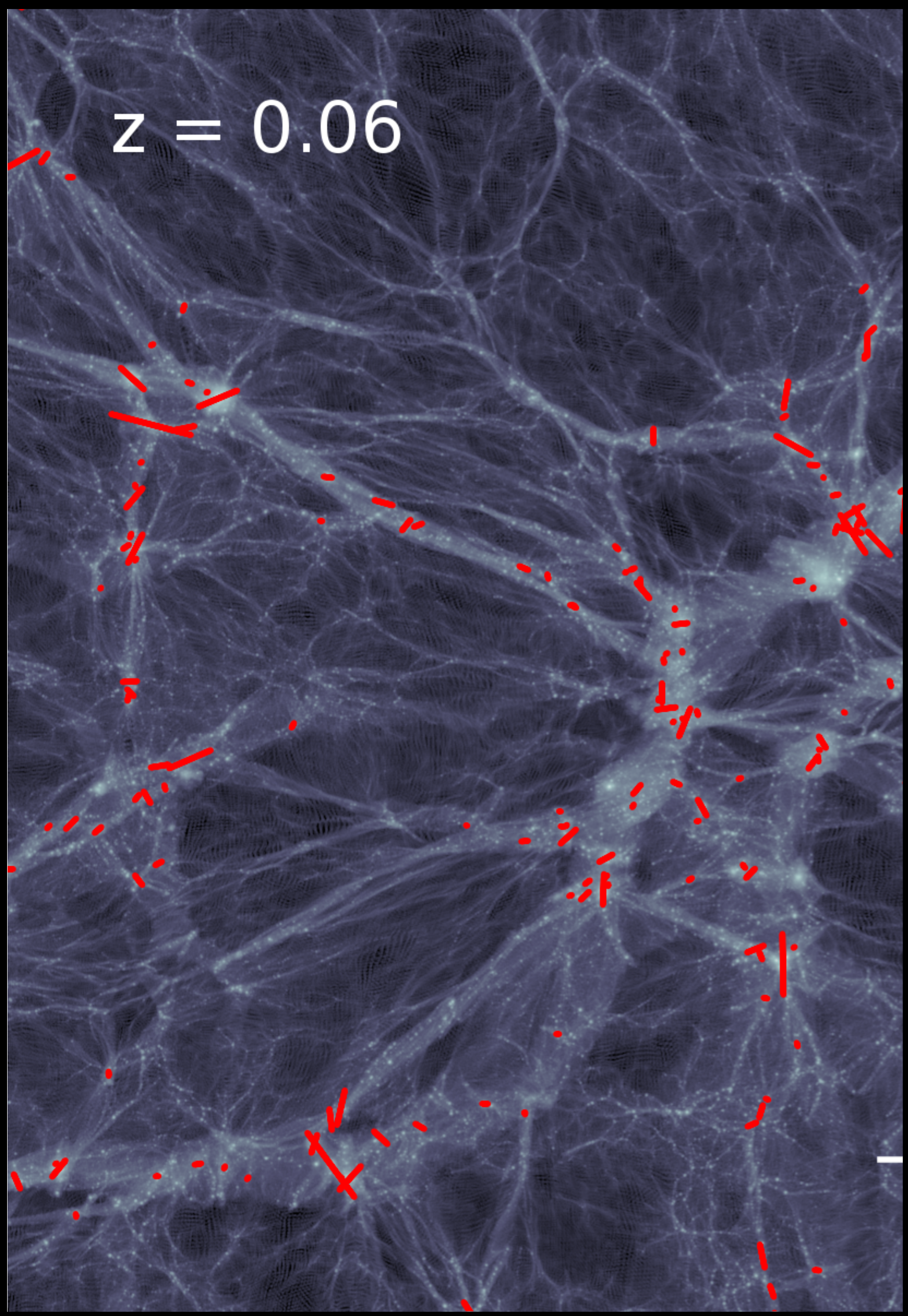
Elisabeth Krause
University of Arizona

Cambridge-LMU workshop, 7.1.2021

Photometric LSS Surveys



Galaxies as (Idealized) Tracers



Khandai+2014, Tenneti+2014

Observable: positions/galaxy density

$$\delta_{\rm g} = b_1\delta + b_s\delta^2 + b_s s^2 + \cdots$$

(e.g, McDonald & Roy 2009, Desjaques, Jeong & Schmidt 2018)

Observable: shapes

$$\gamma^{\rm obs} = \gamma^{\rm G} + \gamma^{\rm I}$$ (weak lensing + intrinsic shape)

intrinsic shape from collapse in tidal field

$$\gamma^{\rm I}_{ij} = C_1 s_{ij} + C_2 s_{ik}s_{kj} + C_\delta \delta s_{ij} + C_t t_{ij} + \cdots$$

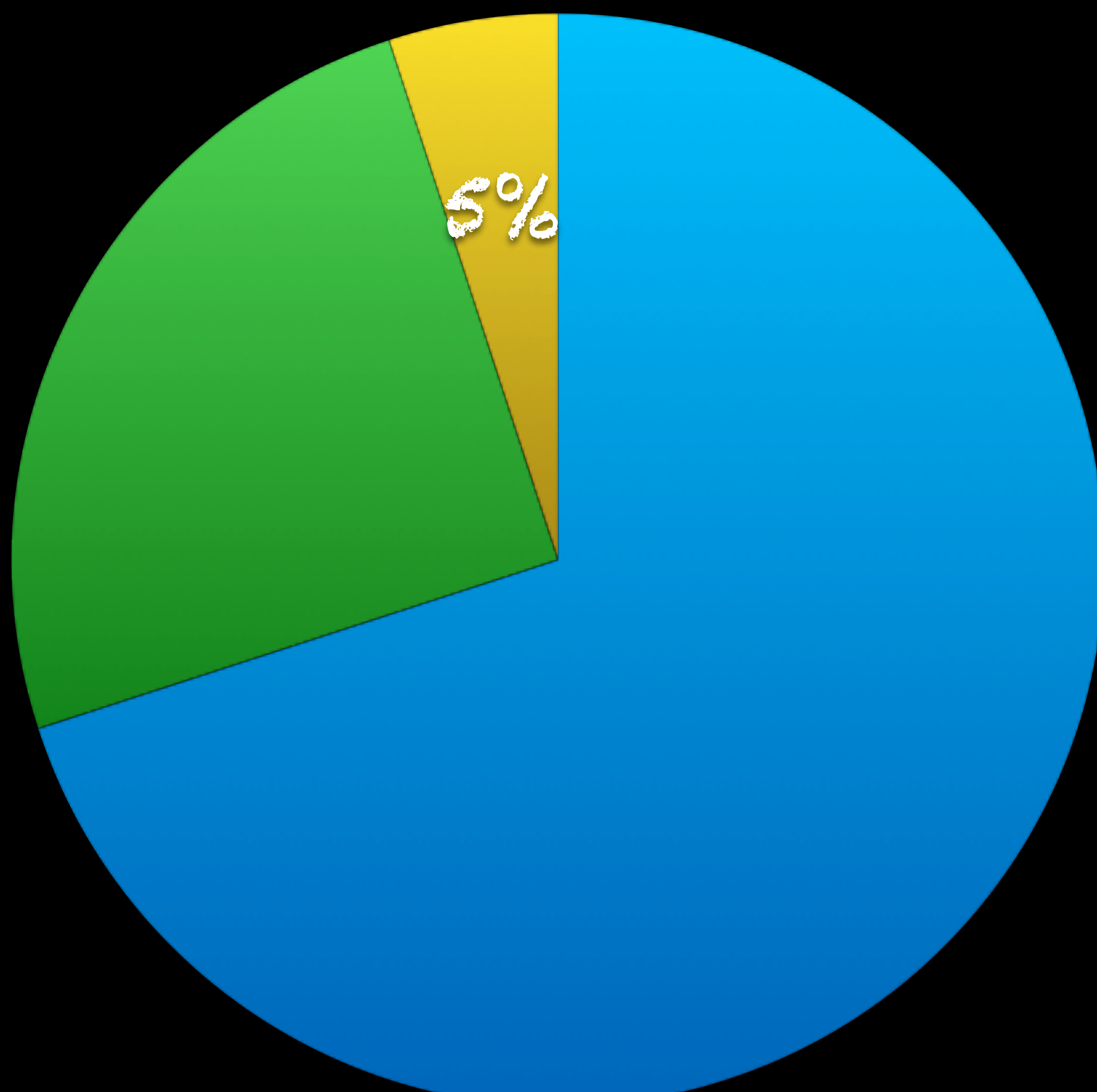
(e.g, Blazek+ 2015, Schmidt+ 2015, Vlah+ 2020ab)

Predict (large-scale) scale dependence for specific galaxy type (expansion coeffs)

Need astrophysics to understand time dependence!

Preview: Cosmology Analyses, ca. 2025

Cosmology Parameters



95% Systematics Parameters

- *known unknowns*
- *unknown unknowns*

From Cosmology to Observations

Parameters	**(Unobservables)**	**Observables**	**Observations**
$\boldsymbol{\lambda}_{\rm cosmo}$ $\boldsymbol{\lambda}_{\rm th}$		$\boldsymbol{\lambda}_{\rm astro}$	$\boldsymbol{\lambda}_{\rm obs}$
initial conditions	<u>3D matter fluctuations</u>	<u>(projected) tracers</u>	<u>maps, catalogs</u>
energy components	matter power spectrum	tracer power spectra	tracer power spectra
background evol.	halo mass function	cluster counts	cluster counts
	…	….	….
			as measured from data
this talk:		$\{C_{ab}(\ell), N\}$	$\{\hat{C}_{ab}(\ell), \hat{N}\}$

focus on astrophysical systematics

<u>see DES-Y3 early papers, Niall's talk (Friday) for observational systematics examples!</u>

From Observations to Cosmology

$$p(\boldsymbol{\lambda}_{\rm cosmo}|\{\hat{C}(\ell),\hat{N}\}) = p(\boldsymbol{\lambda}_{\rm cosmo}) \int d\,\boldsymbol{\lambda}_{\rm th+astro+obs}\, p(\boldsymbol{\lambda}_{\rm th+astro+obs})\, p\left(P_{\rm m}, n(M)|\boldsymbol{\lambda}_{\rm cosmo+th}\right)$$
$$p(\{C(\ell),N\}|P_{\rm m}, n(M), \boldsymbol{\lambda}_{\rm astro}) p(\{\hat{C}(\ell),\hat{N}\}|\{C(\ell),N\},\boldsymbol{\lambda}_{\rm obs})$$

From Observations to Cosmology

Science Case

parameters of interest
which science?

large data vector
which probes + scales?

"systematic effects"
may outnumber cosmo params
parameterize + prioritize!

systematics prior
large prior volume
validate (external data, simulations)

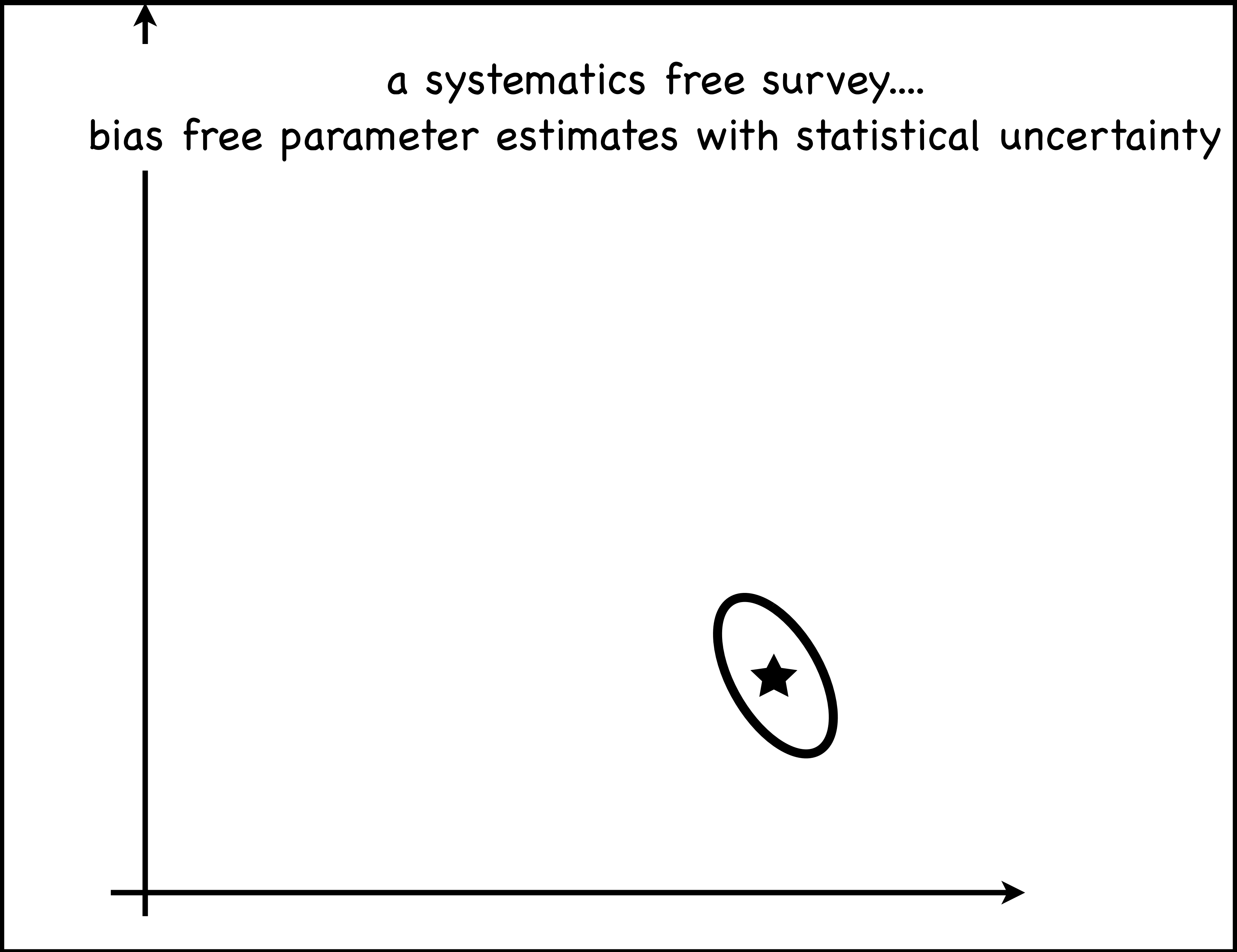
$$p(\boldsymbol{\lambda}_{\rm cosmo}|\{\hat{C}(\ell),\hat{N}\}) = p(\boldsymbol{\lambda}_{\rm cosmo}) \int d\boldsymbol{\lambda}_{\rm th+astro+obs}\, p(\boldsymbol{\lambda}_{\rm th+astro+obs})\, p(P_{\rm m}, n(M)|\boldsymbol{\lambda}_{\rm cosmo+th})$$
$$p(\{C(\ell),N\}|P_{\rm m}, n(M), \boldsymbol{\lambda}_{\rm astro})\, p(\{\hat{C}(\ell),\hat{N}\}|\{C(\ell),N\}, \boldsymbol{\lambda}_{\rm obs})$$

Cosmology Priors

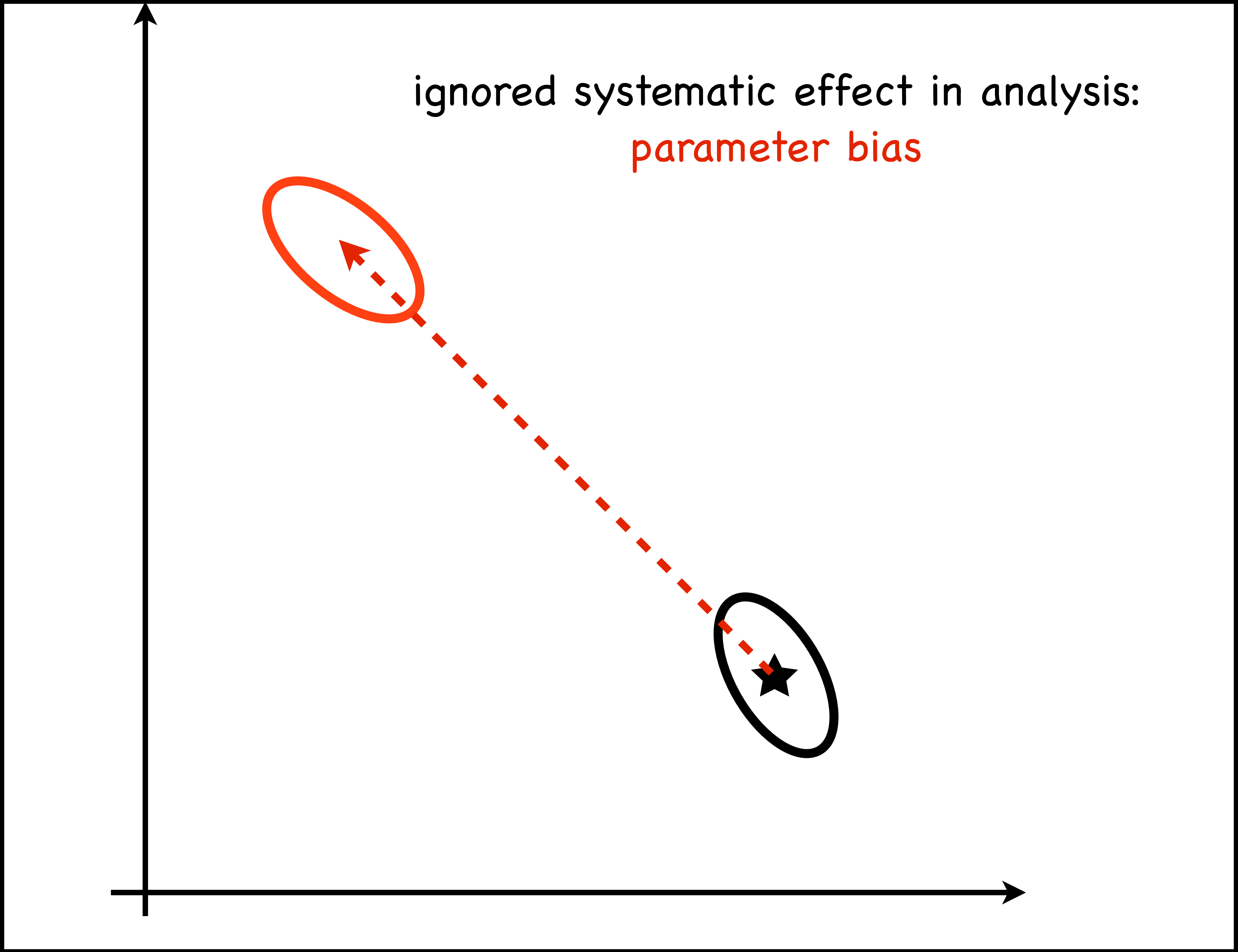
Likelihood
for observables + systematics
requires (data, sys) covariances

Model Data Vector
consistent modeling of all observables
including all (cosmo + nuisance) parameters

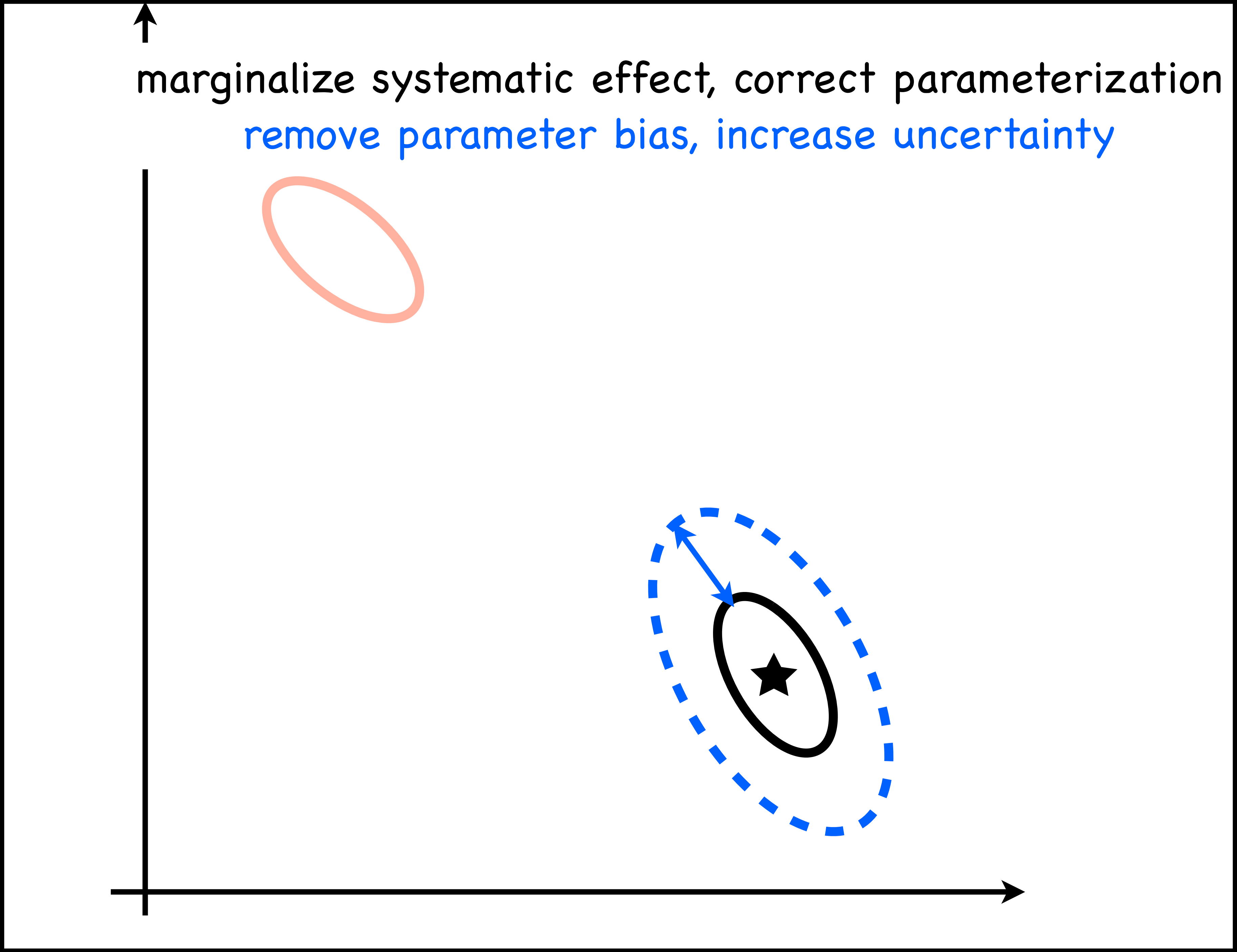
The Trouble with Systematics



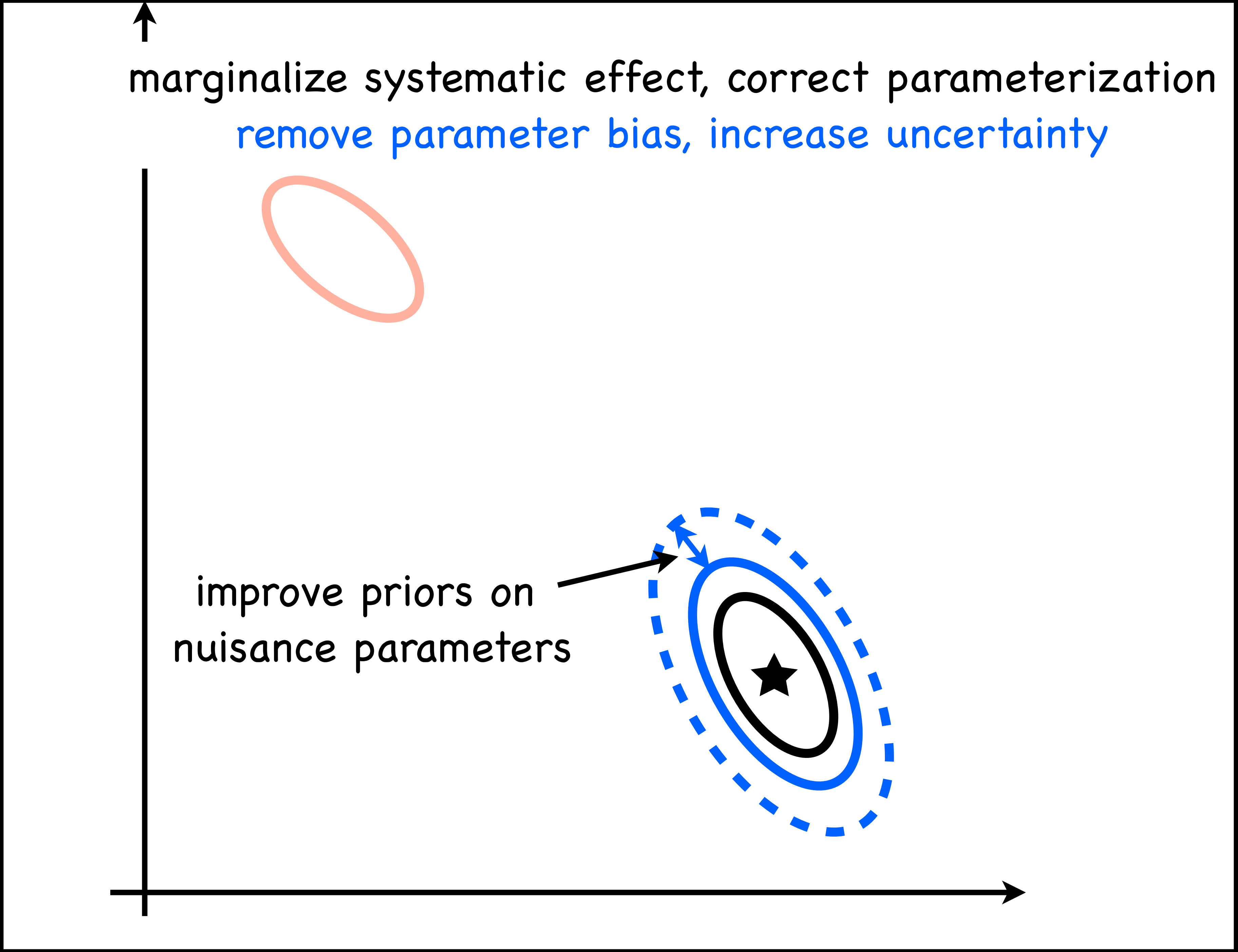
The Trouble with Systematics



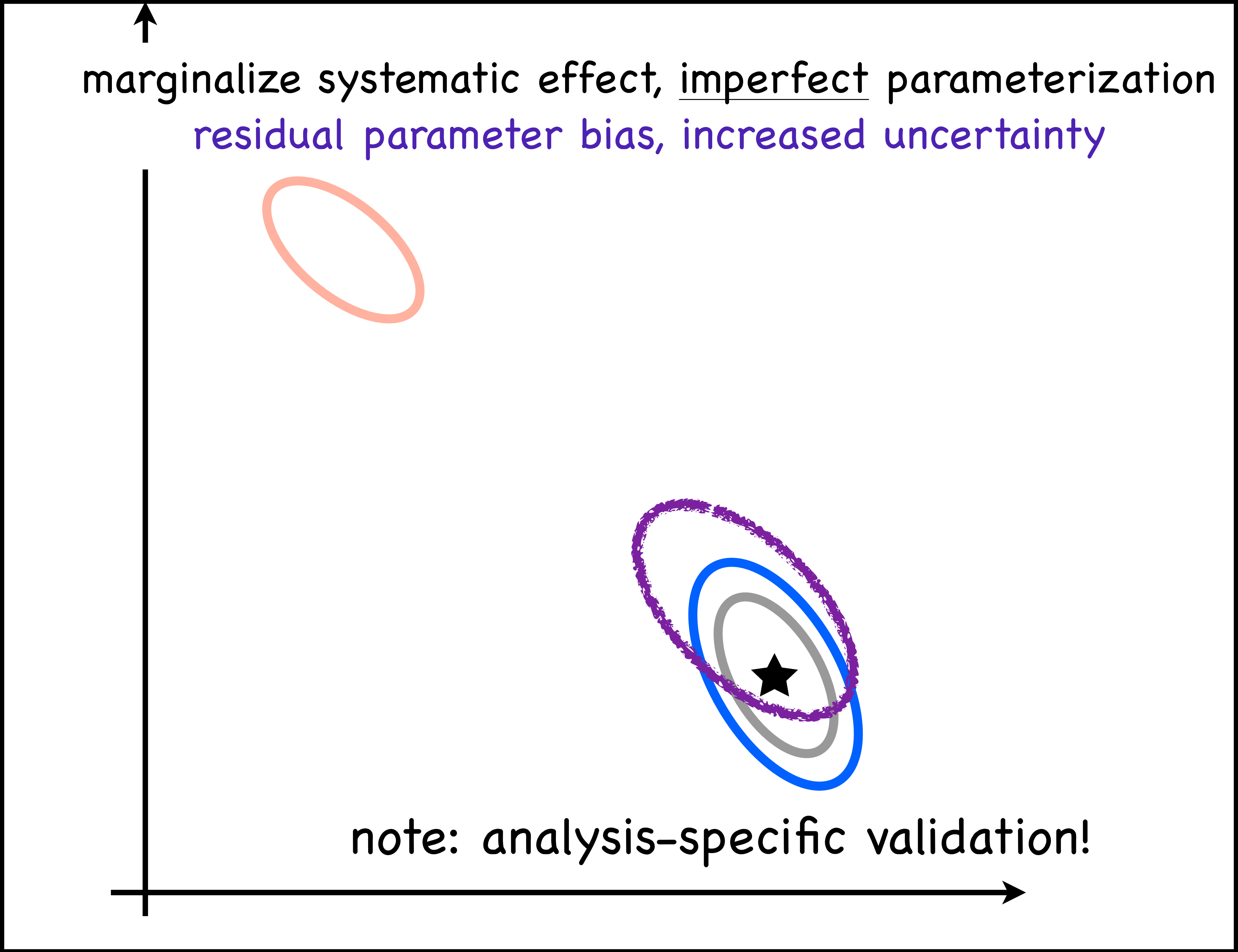
The Trouble with Systematics



The Trouble with Systematics



The Trouble with Systematics

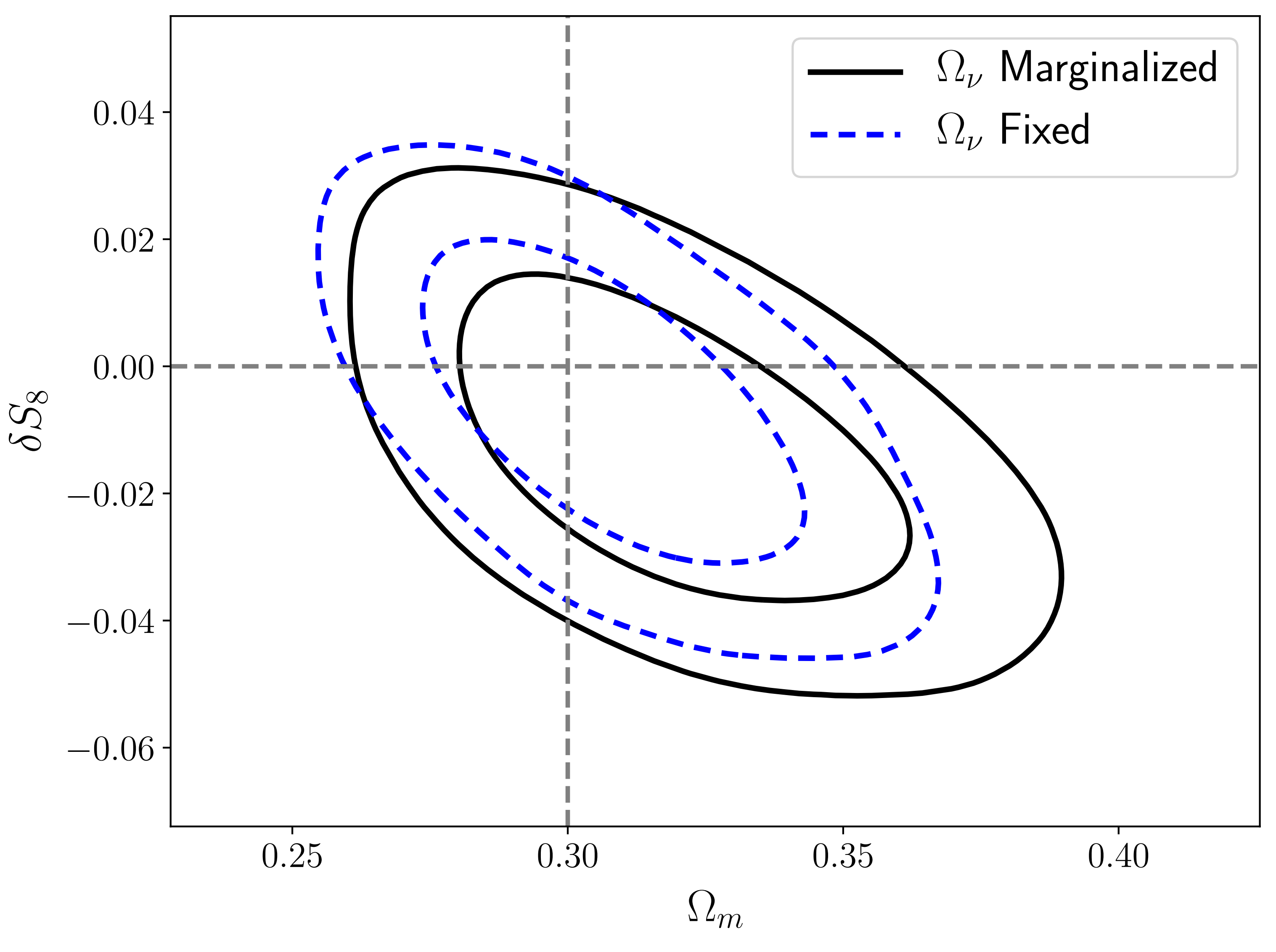


Combined Probes Systematics

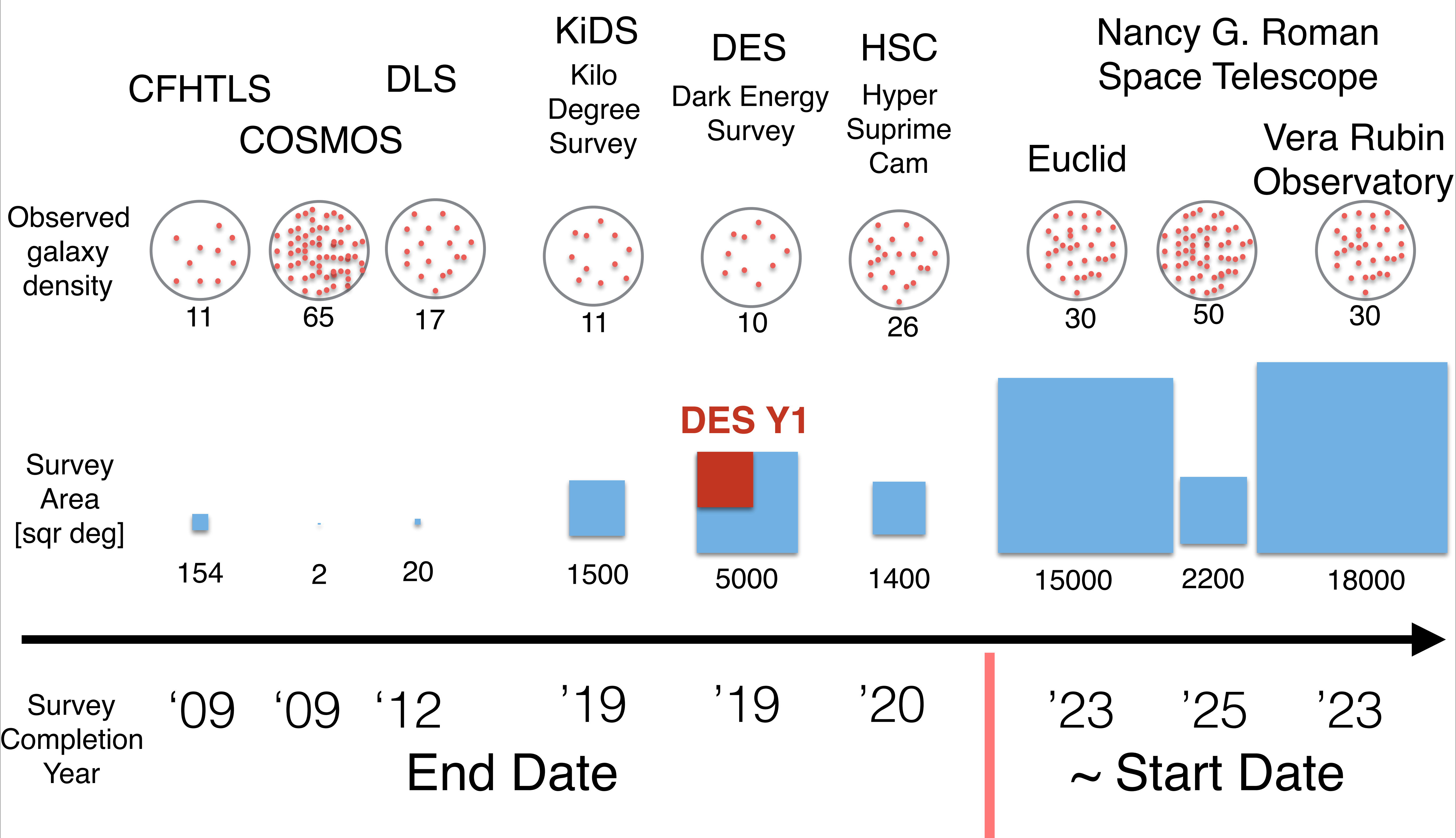
- "Precision cosmology": excellent statistics - systematics limited
  - (and person-power limited!)
- Easy to come up with large list of systematics + nuisance parameters
  - galaxies: LF, bias (e.g., 5 HOD parameters + b_2 per z-bin,type)
  - cluster mass-observable relation: mean relation + scatter parameters
  - shear calibration, photo-z uncertainties, intrinsic alignments,...
  - Σ(poll among DES working groups) ~ 500-1000 parameters [2013 estimate]
- Self-calibration + marginalization?
  - costly (computationally, constraining power)

The Trouble with Systematics

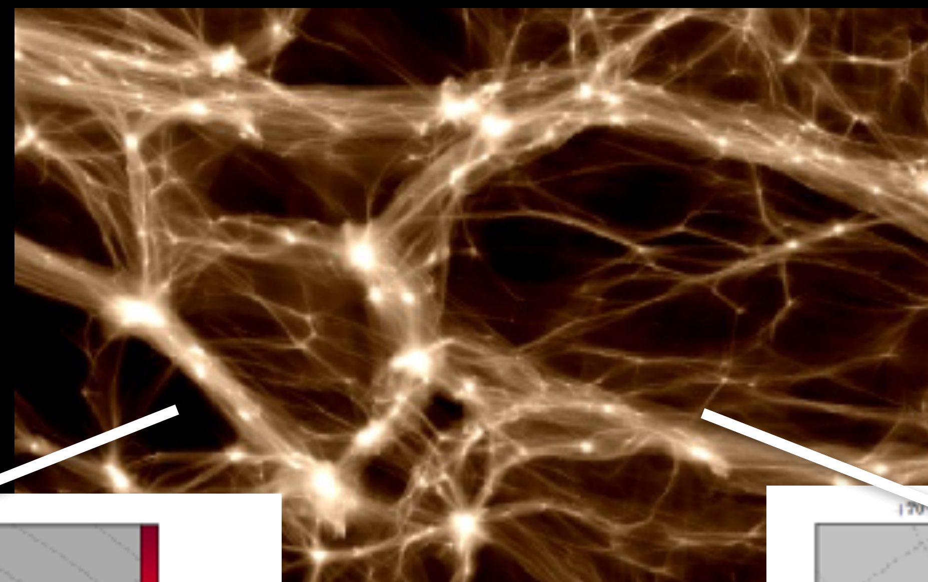
idealized DES-Y3 analysis, input = analysis model
parameter correlations may bias marginalized posteriors



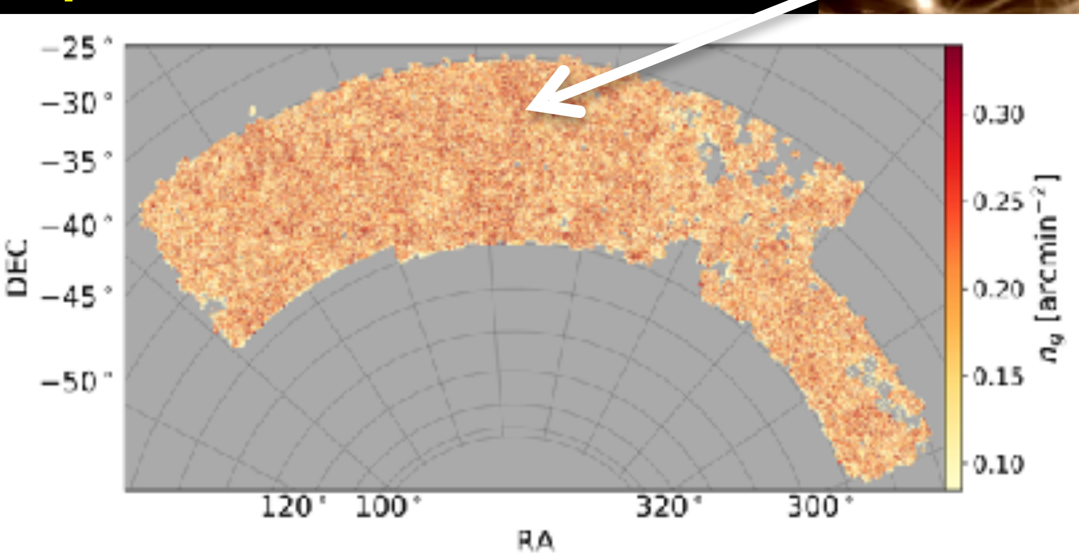
Photometric LSS Surveys



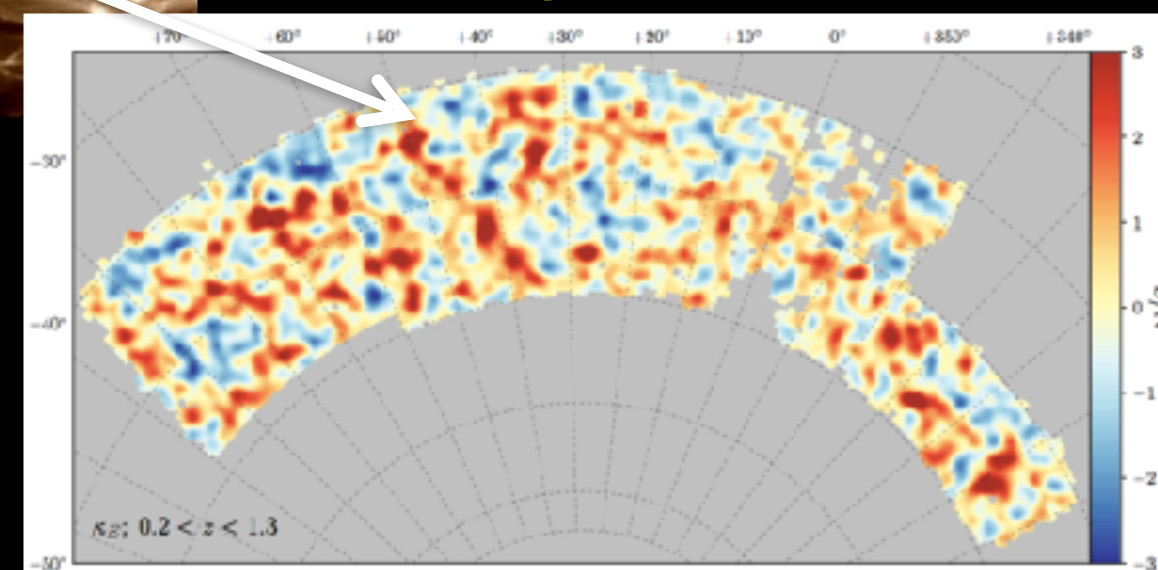
DES Year 1 WL x LSS Analysis



660K redMaGiC galaxies
split in 5 redshift bins



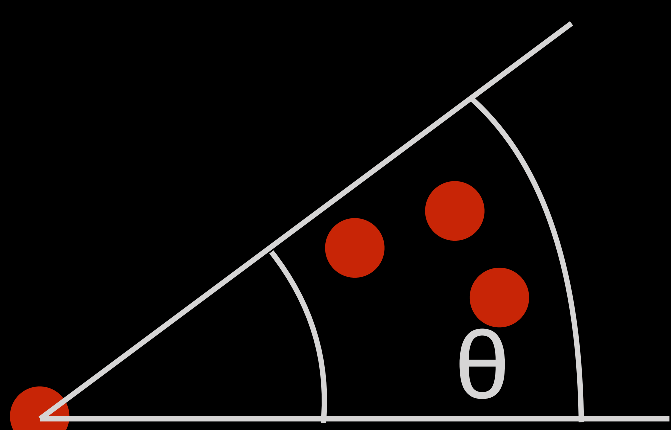
26M source galaxies
split in 4 redshift bins

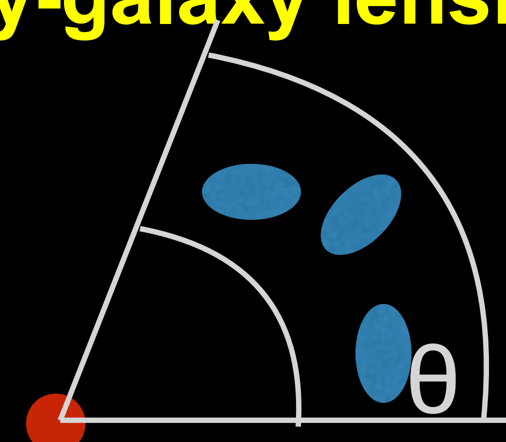


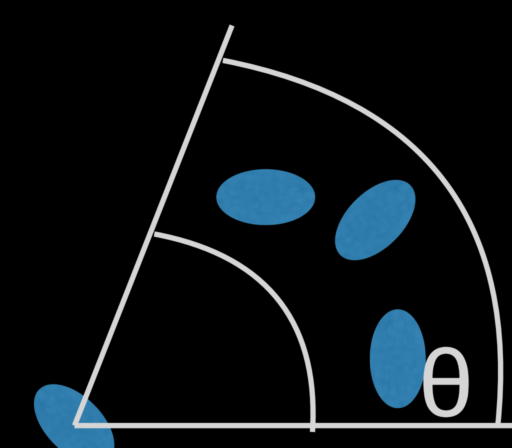
galaxies x galaxies:
angular clustering

galaxies x lensing:
galaxy-galaxy lensing

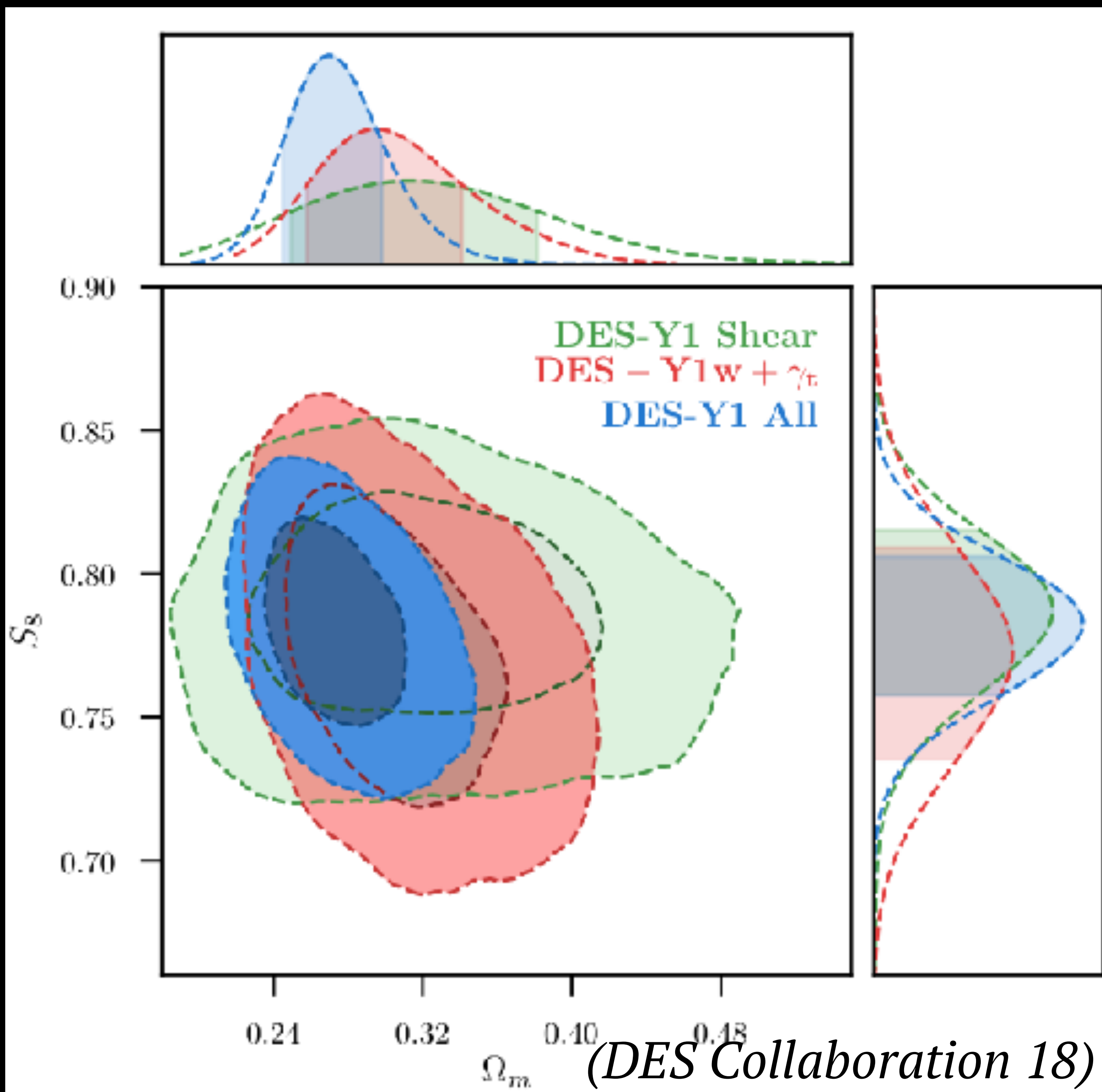
lensing x lensing:
cosmic shear







DES Y1 Results:
LCDM Multi-Probe Constraints



(DES Collaboration 18)

Amplitude of Structure Growth

Matter Density

- marginalized 4 cosmology parameters, 10 clustering nuisance parameters, and 10 lensing nuisance parameters
- consistent cosmology constraints from weak lensing and clustering in configuration space
- Central values differ by >1σ from Planck, in the same direction as other lensing analyses (CFHTLS, KiDS-1000, HSC)

DES-Y1 Systematics Modeling + Mitigation

baseline systematics marginalization (20 parameters)

- linear bias of lens galaxies, per lens z-bin
- lens galaxy photo-zs, per lens z-bin
- source galaxy photo-zs, per source z-bin
- multiplicative shear calibration, per source z-bin
- intrinsic alignments, power-law/free amplitude per per source z-bin

-> this list is known to be incomplete

how much will **known, unaccounted-for** systematics bias Y1?

-> remove contaminated data points (*i.e., throw out large fraction of S/N*)

-> choice of parameterizations ≠ universal truth

are these **parameterizations sufficiently flexible** for Y1?

1706.09359

Systematics Mitigation
incomplete model - scale cuts

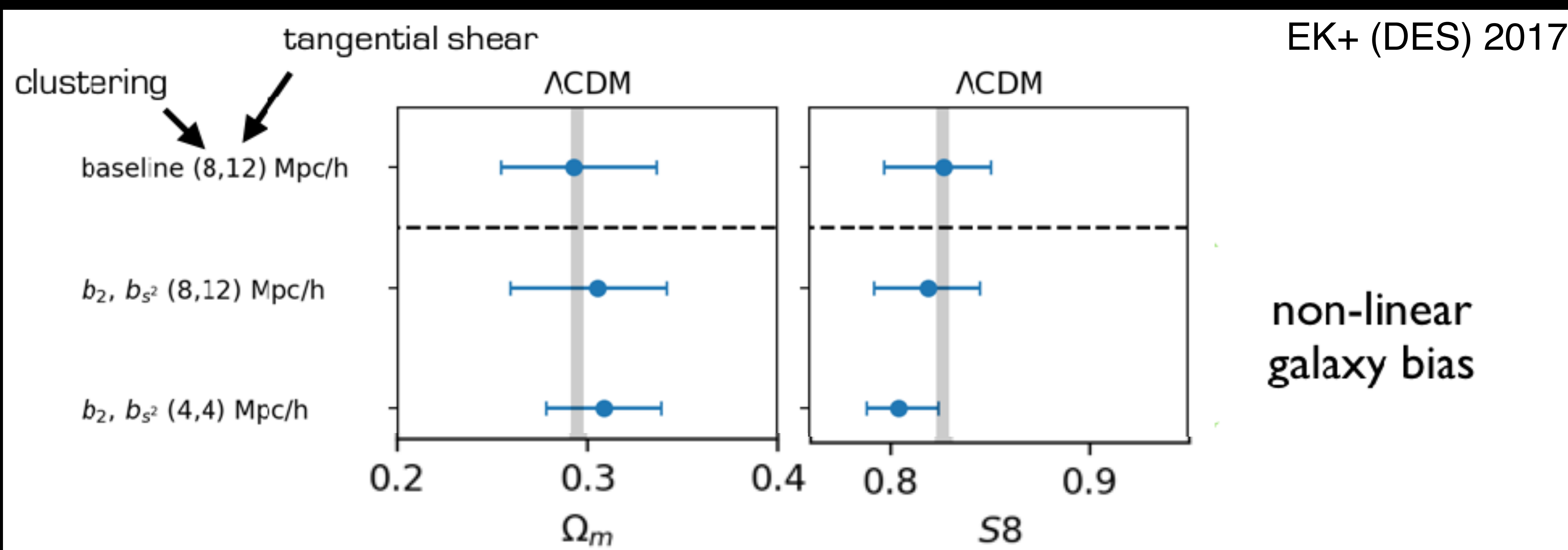
-> this list is known to be incomplete

how much will known, unaccounted-for systematics bias Y1 results?

Example: generate input 'data' incl. 2nd order galaxy bias

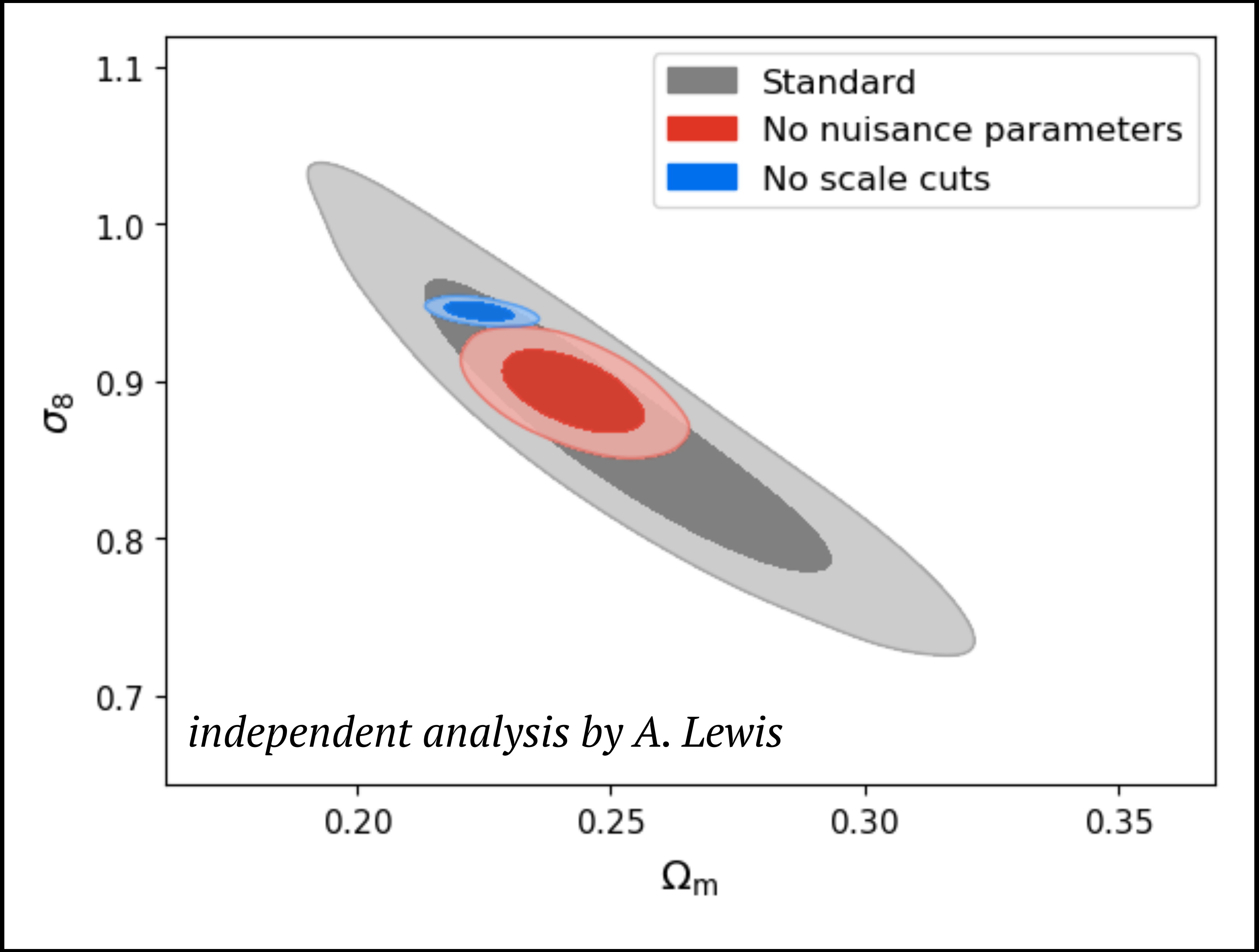
enhances clustering signal on small physical scales

determine scale cuts to minimize parameter biases



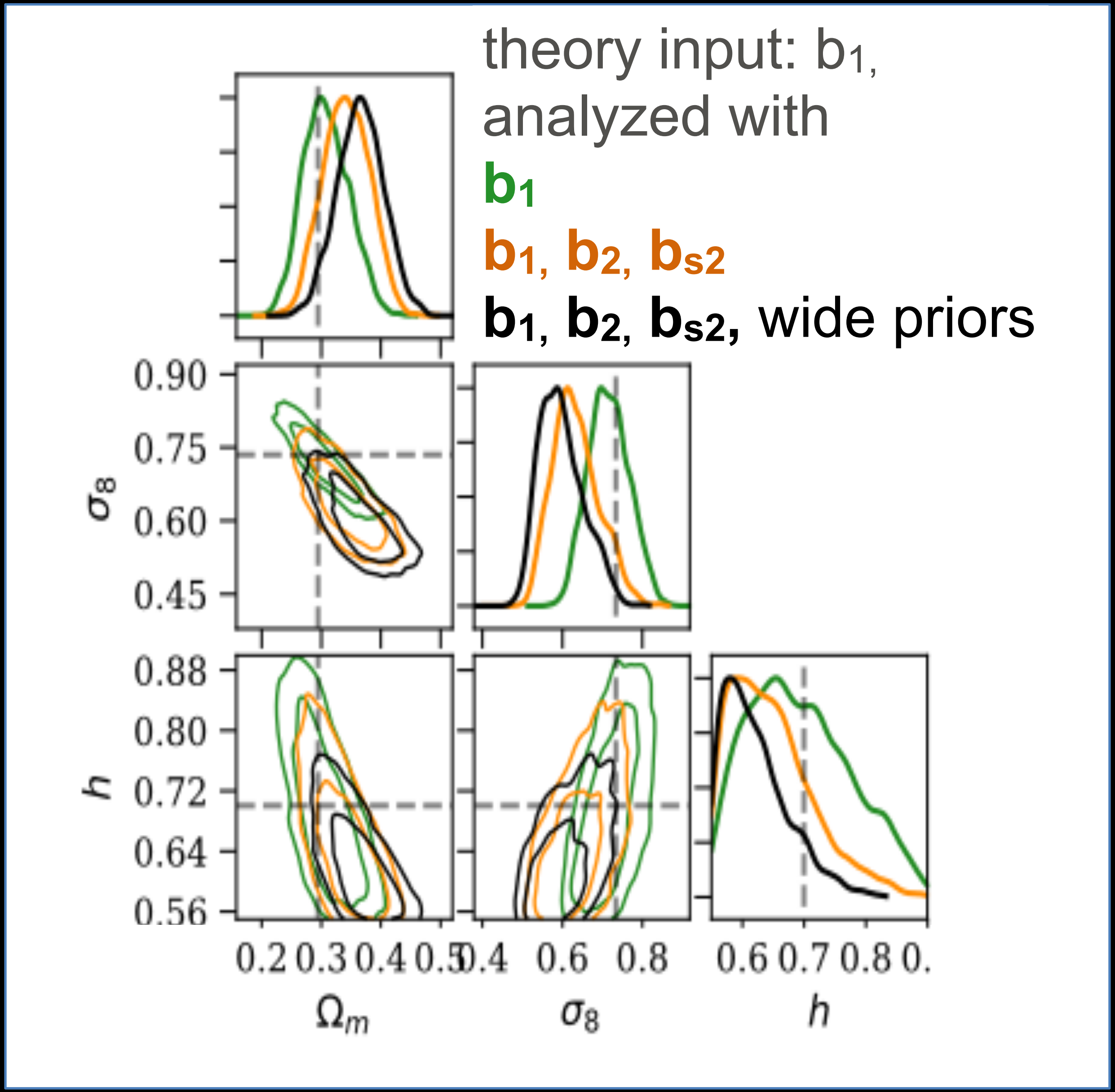
DESY1 Systematics Mitigation

Does it hurt?



DES-Y1 Systematics Modeling + Mitigation why such simple models?

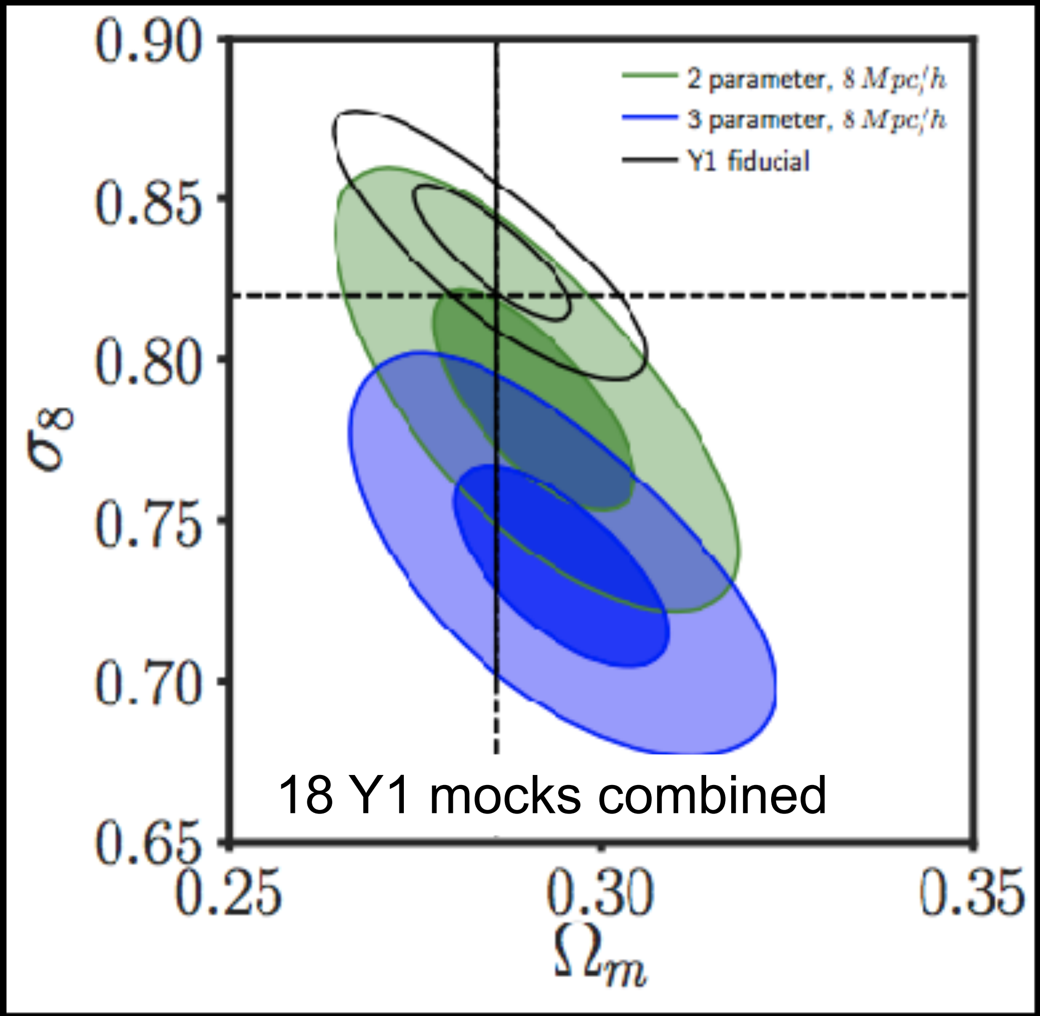
- More accurate (+more complex) systematics models have been around for decades… why not use them?
- Sampling over poorly constrained model parameters may bias inferred cosmology (if model parameters are degenerate with cosmology)
- Model evaluation time is important when running hundreds of chains
- (save most accurate model for validation)

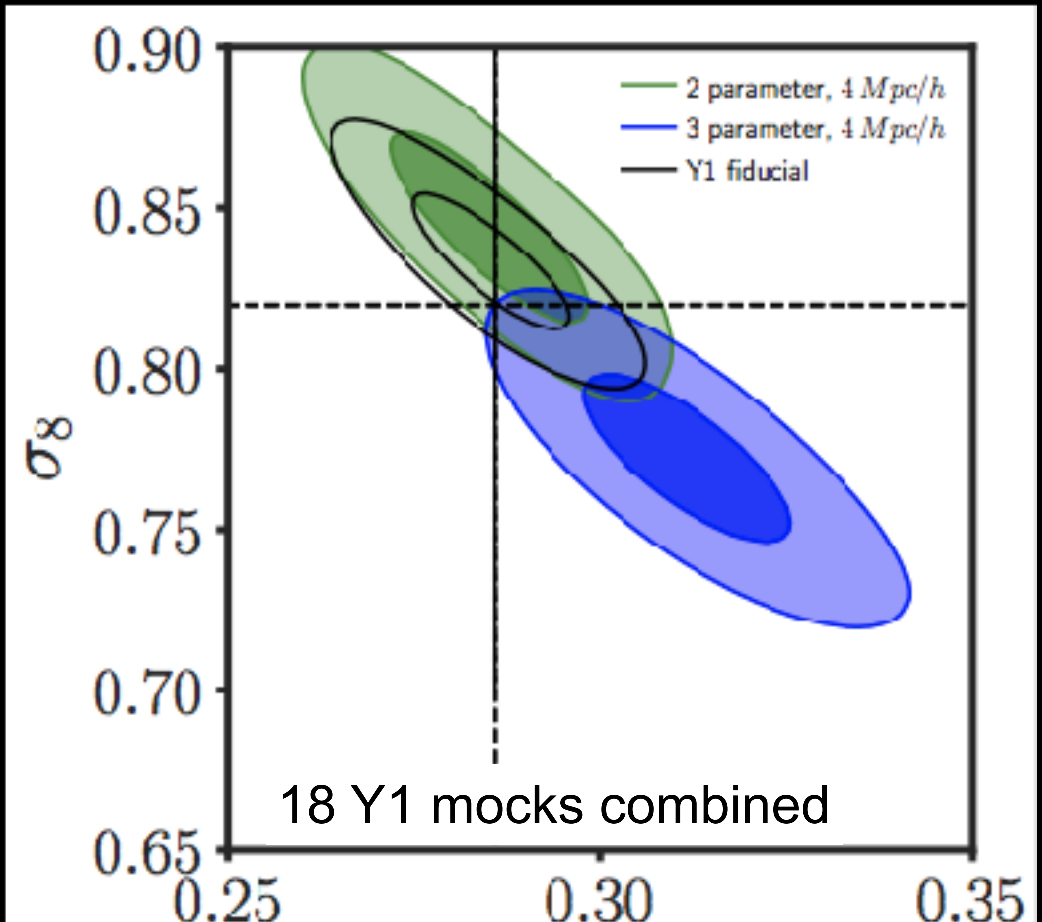


Constraining power influences allowed model complexity

Lesson: simulate analyses early and often!

DES-Y1 Systematics Modeling + Mitigation: non-linear bias, simulation input





work by Shivam Pandey & Niall MacCrann

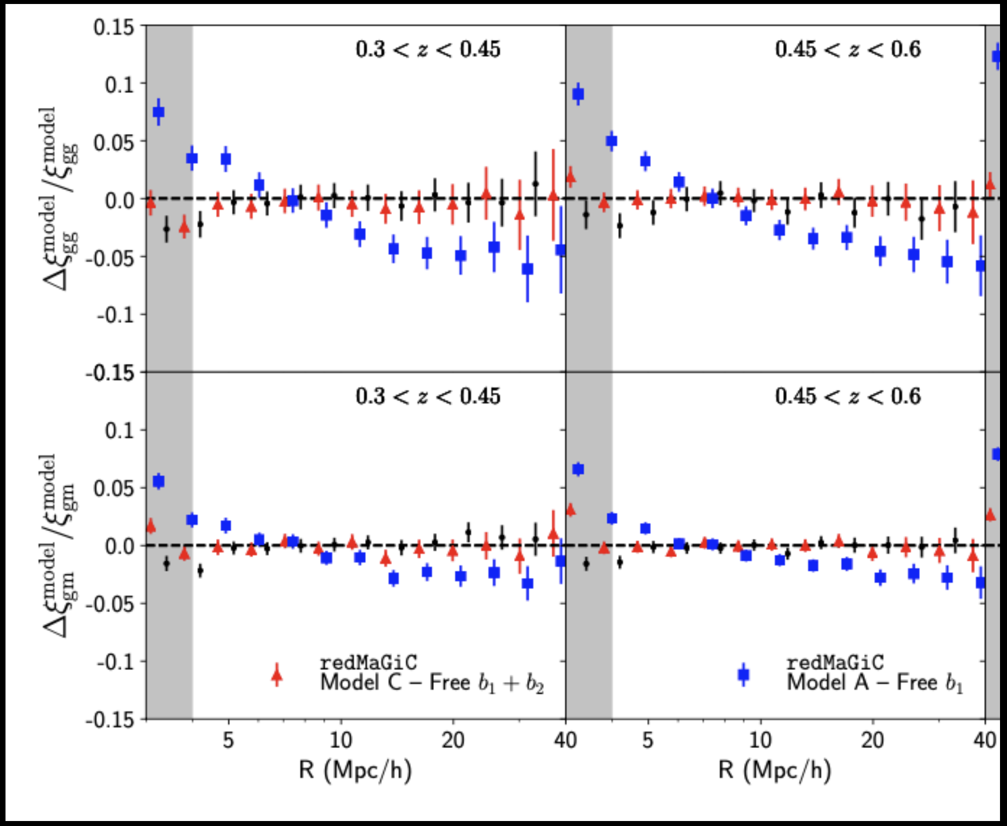
Y1 fiducial: $\mathbf{b_1}$

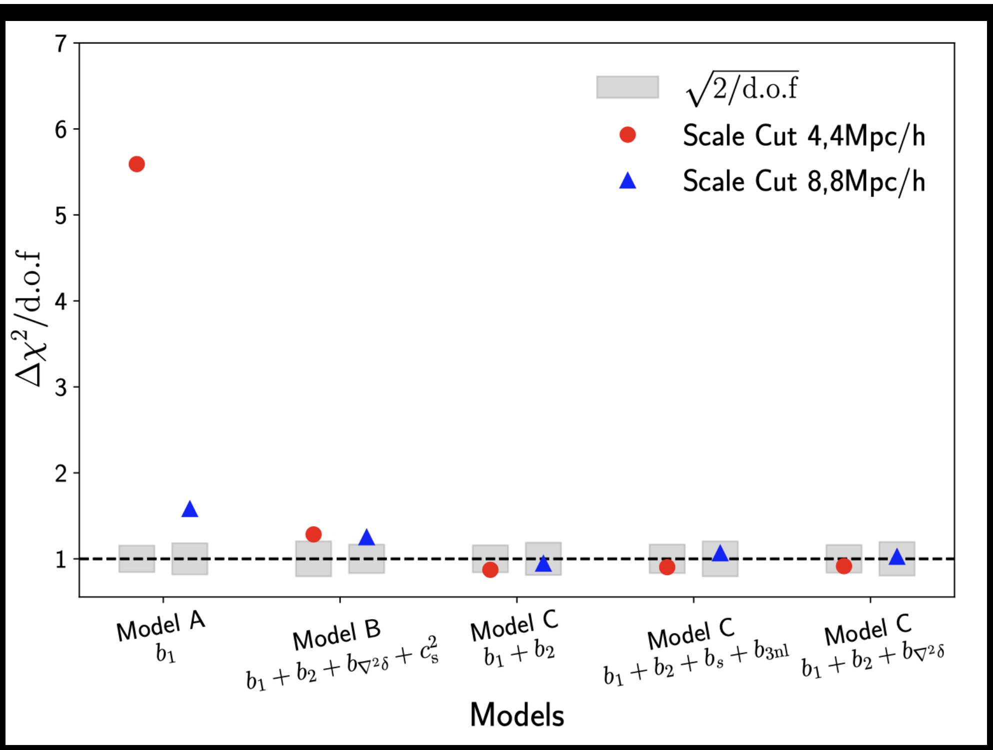
2 parameters: $\mathbf{b_1 + b_k\, k^2} + b_2(b_1) + b_{s2}(b_1) + b_{3nl}(b_1)$

3 parameters: $\mathbf{b_1 + b_k\, k^2 + b_2} + b_{s2}(b_1) + b_{3nl}(b_1)$

NOTE: good fit to sims with 3 bias parameters *at fixed cosmology*

Systematics Opportunities and Challenges: Non-Linear Bias Modeling



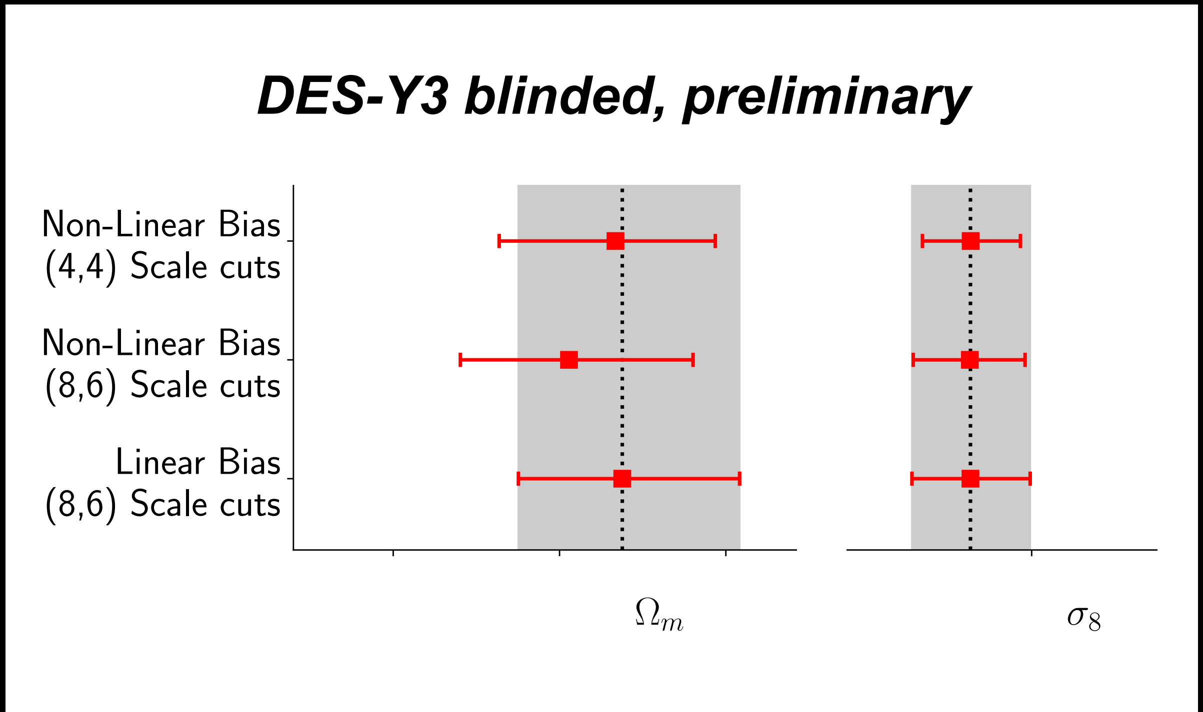


Pandey+ 2020: minimal bias model for DES-Y3 analysis

- analyze galaxy-galaxy and galaxy-matter correlation function in of redMaGiC galaxies in DES mocks
- fit galaxy bias models with varying complexity/number of parameters
- neglecting higher-derivative bias, and setting bs, b3nl to coevolution values sufficent for DES-Y3 analysis

✓ 2 parameter model, reduced prior volume

Systematics Opportunities and Challenges: Non-Linear Bias Modeling



Pandey+ in prep: DES-Y3 clustering + g-g lensing analysis

- increased statistical power and reduced model complexity enable analysis with non-linear bias modeling
- linear bias x non-linear matter power spectrum sufficient for > 8 Mpc/h
- limited increase in constraining power when including smaller scales + non-linear bias model

Systematics Opportunities and Challenges: Baryonic Effects in WL Analyses

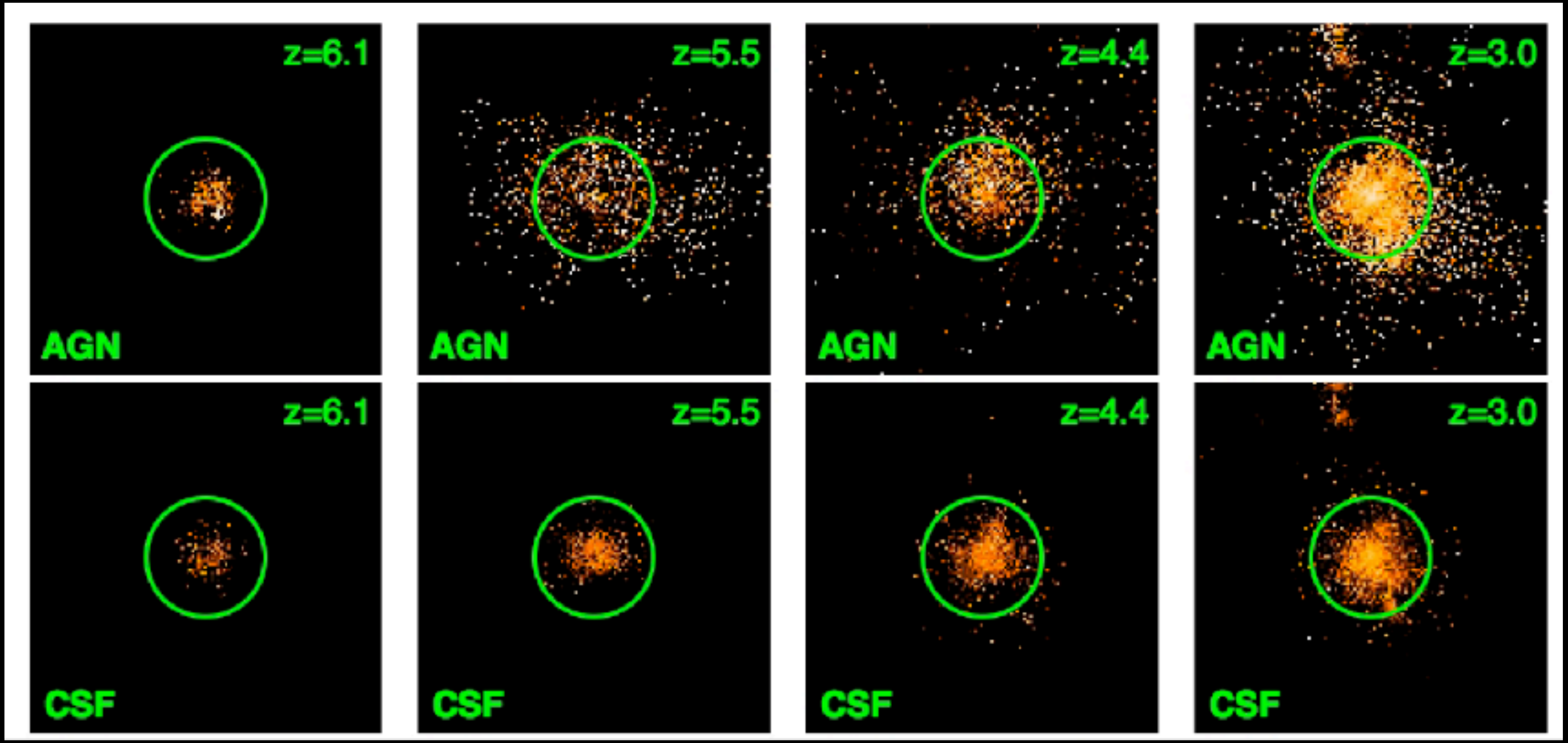
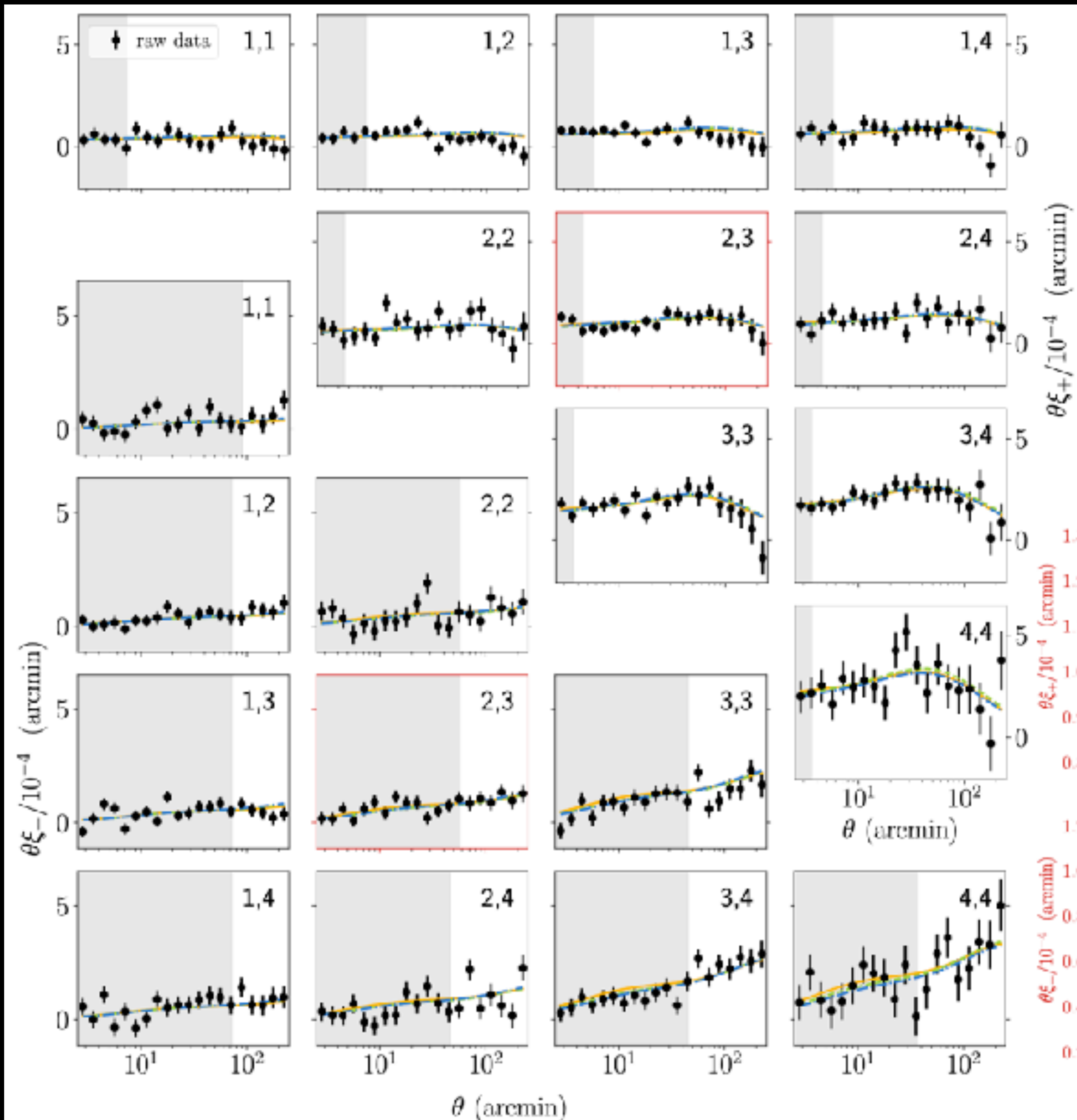


illustration from OWLS collaboration

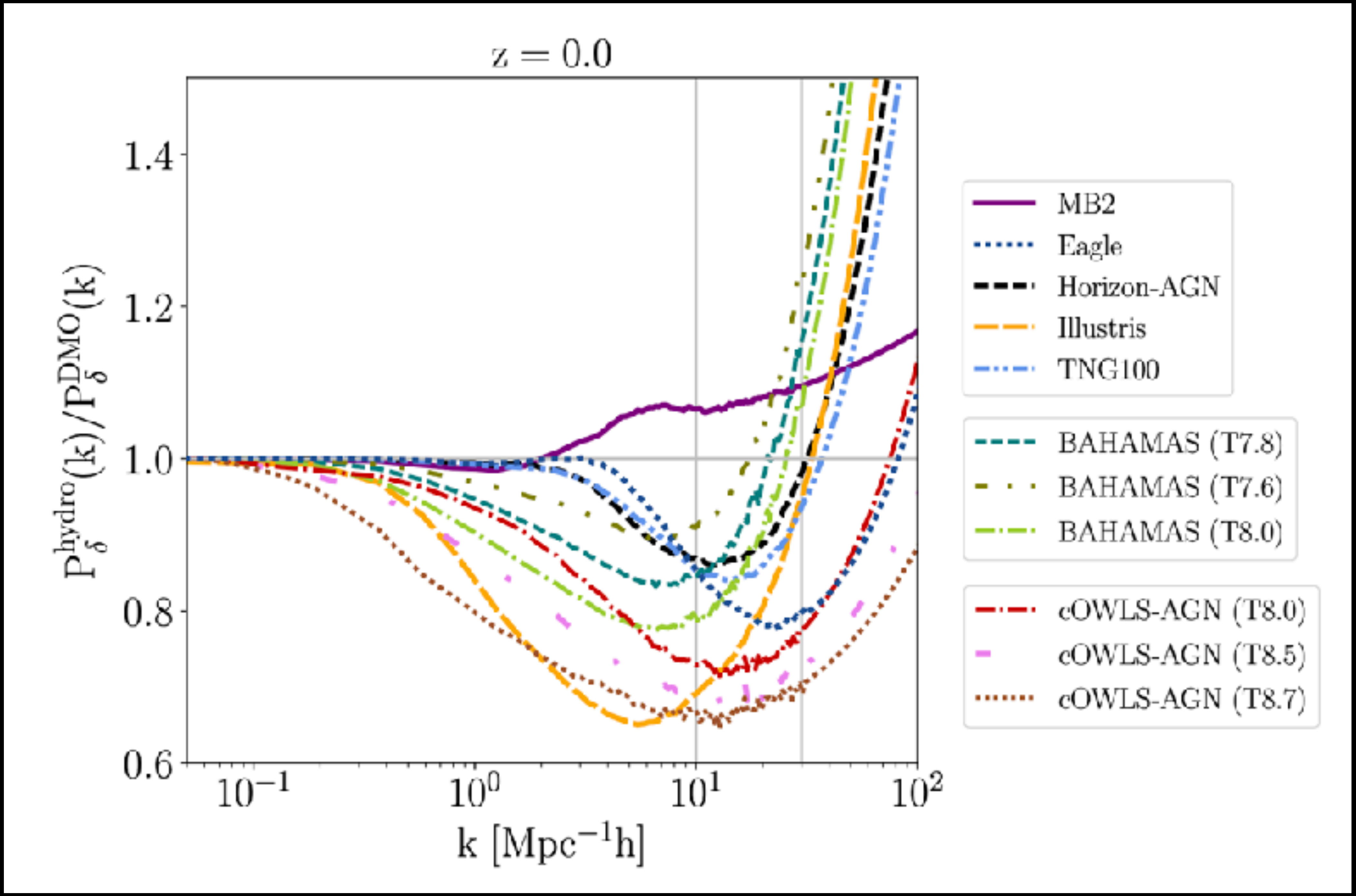
DESY1 WL Correlation functions



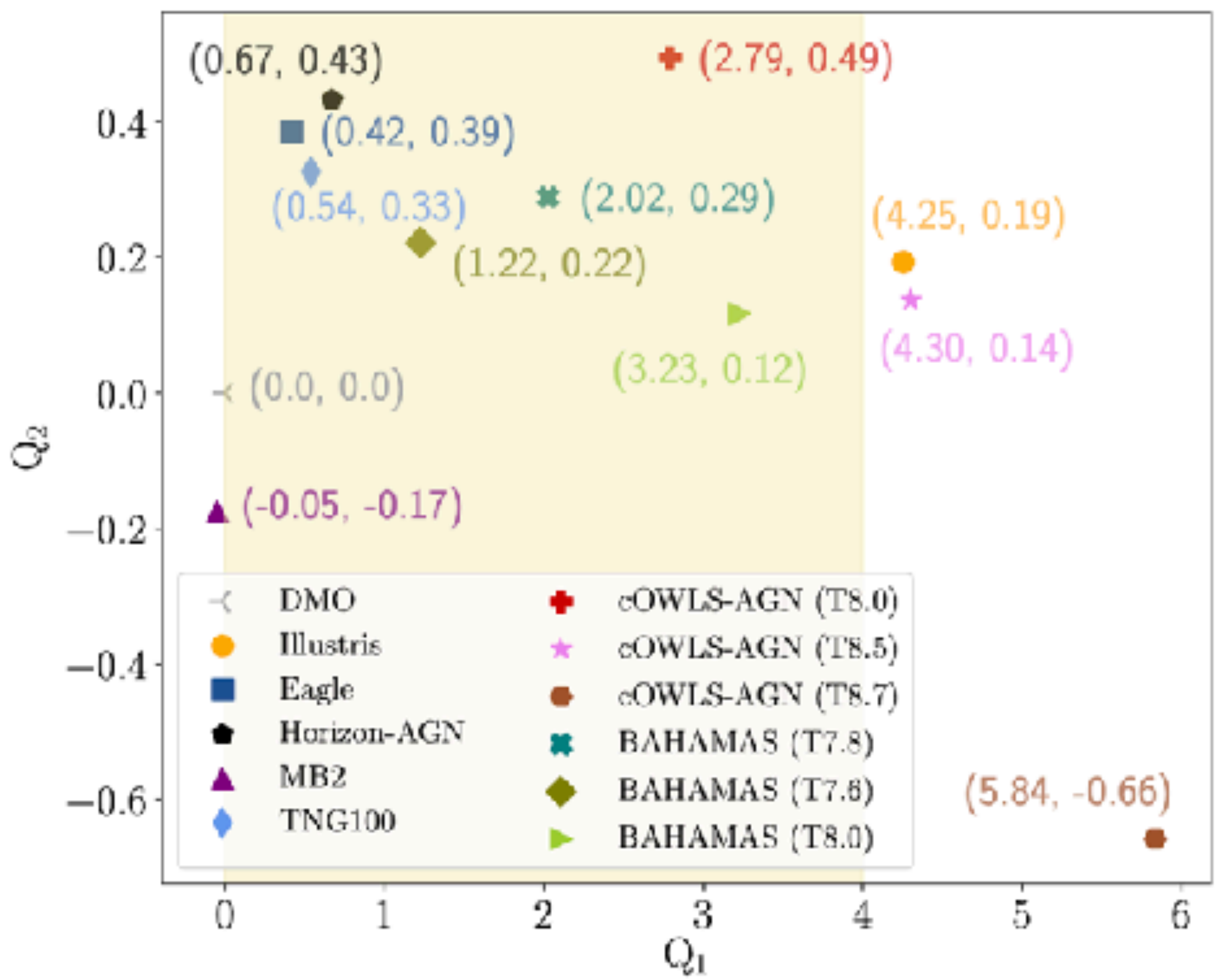
Small scale correlation function data points were being cut out mostly (almost entirely) because of baryonic effects

Huang+2020: reanalyze DESY1 including all WL measurements down to 2.5’

Baryonic Effects in WL Analyses



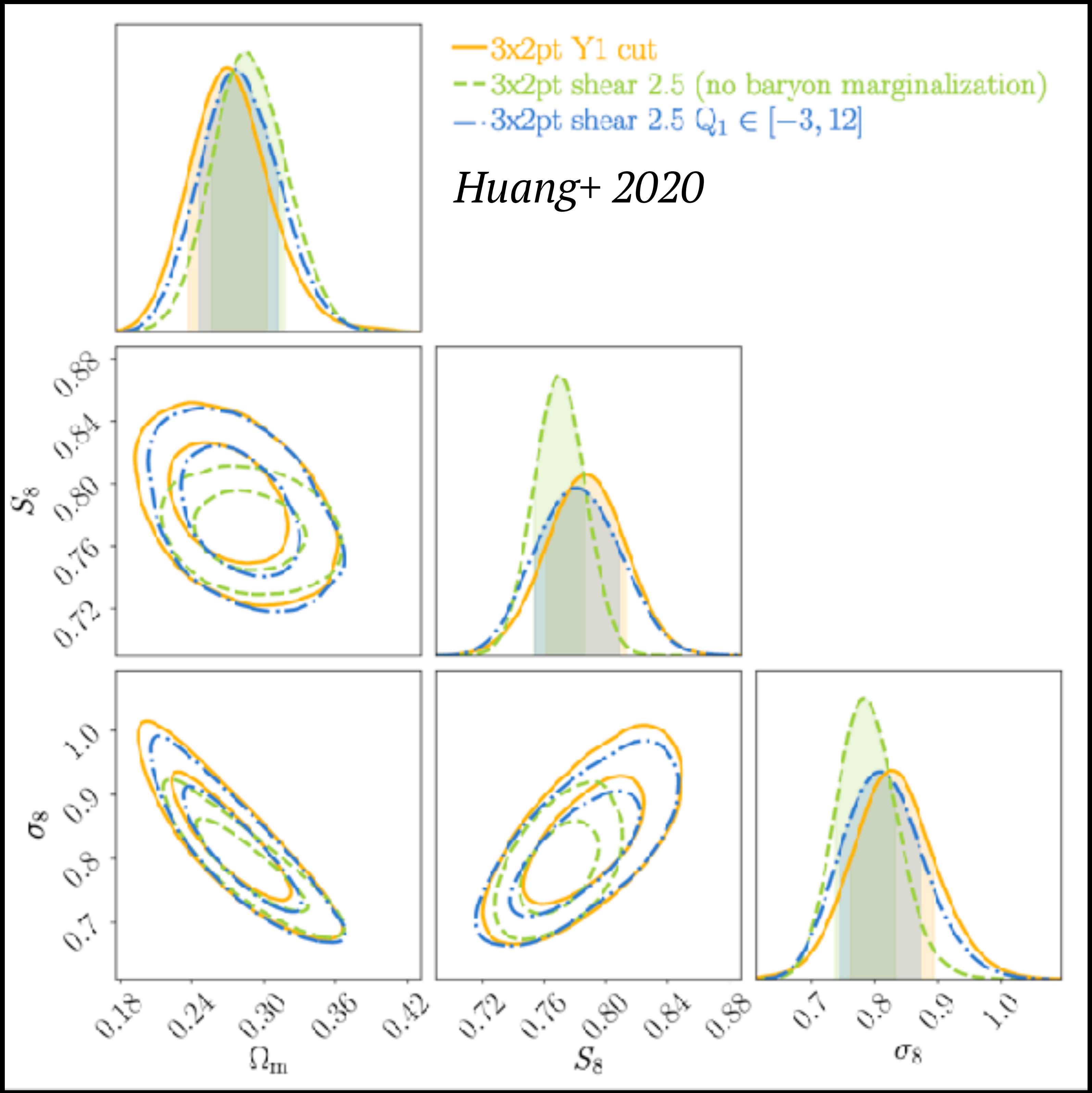
Baryonic Effects in WL Analyses PCA Baryon Impact Model



Huang+ 2020

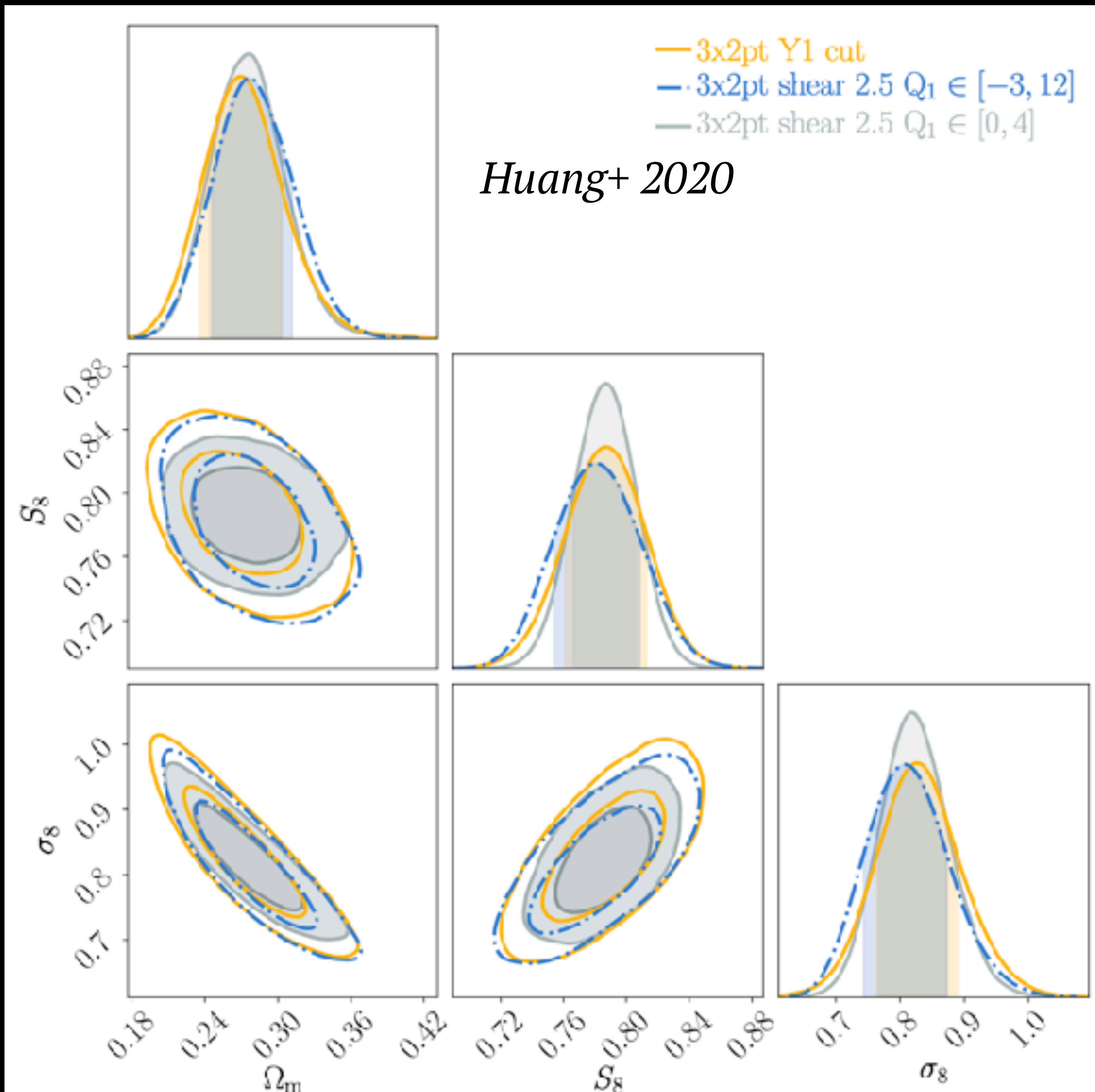
- PC amplitudes Q_i from hydro power spectra
- use Q_i as continuous baryon parameters
- marginalizing over Q_1 is sufficient given DES-Y1 constraining power
- Q_1 non-informative prior: [-3,12]
- Q_1 informative prior: [0,4]

Baryonic Effects in WL Analyses Cosmology Constraints



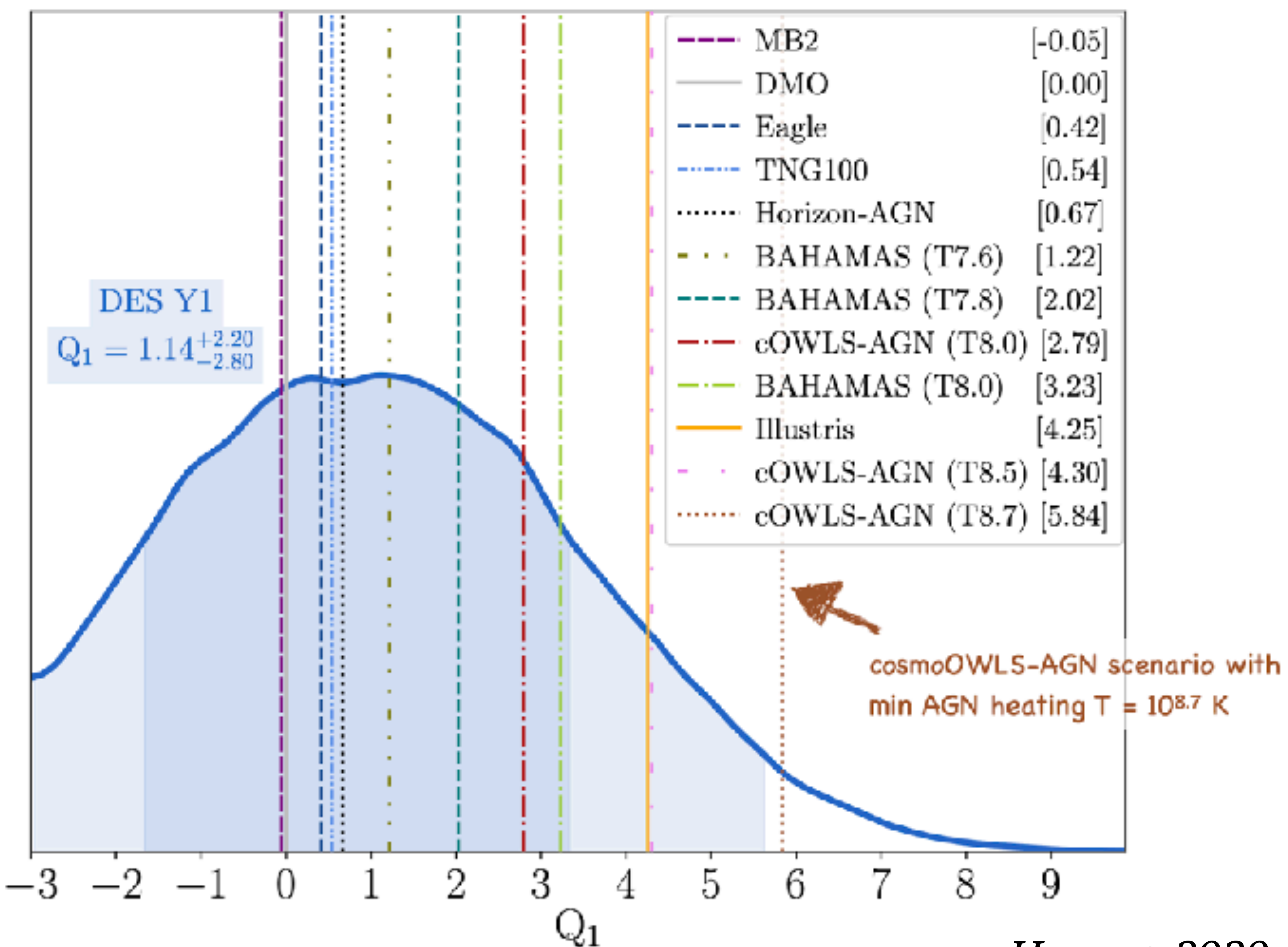
- Green: DES-Y1 including all scales, and baryons are not included in the modeling (don't do that…)
- Orange: DES-Y1 baseline (conservative scale cuts)
- Blue: DES-Y1 including all scales, and baryonic effects are modeled using **PCA with non-informative prior**

Baryonic Effects in WL Analyses Cosmology Constraints



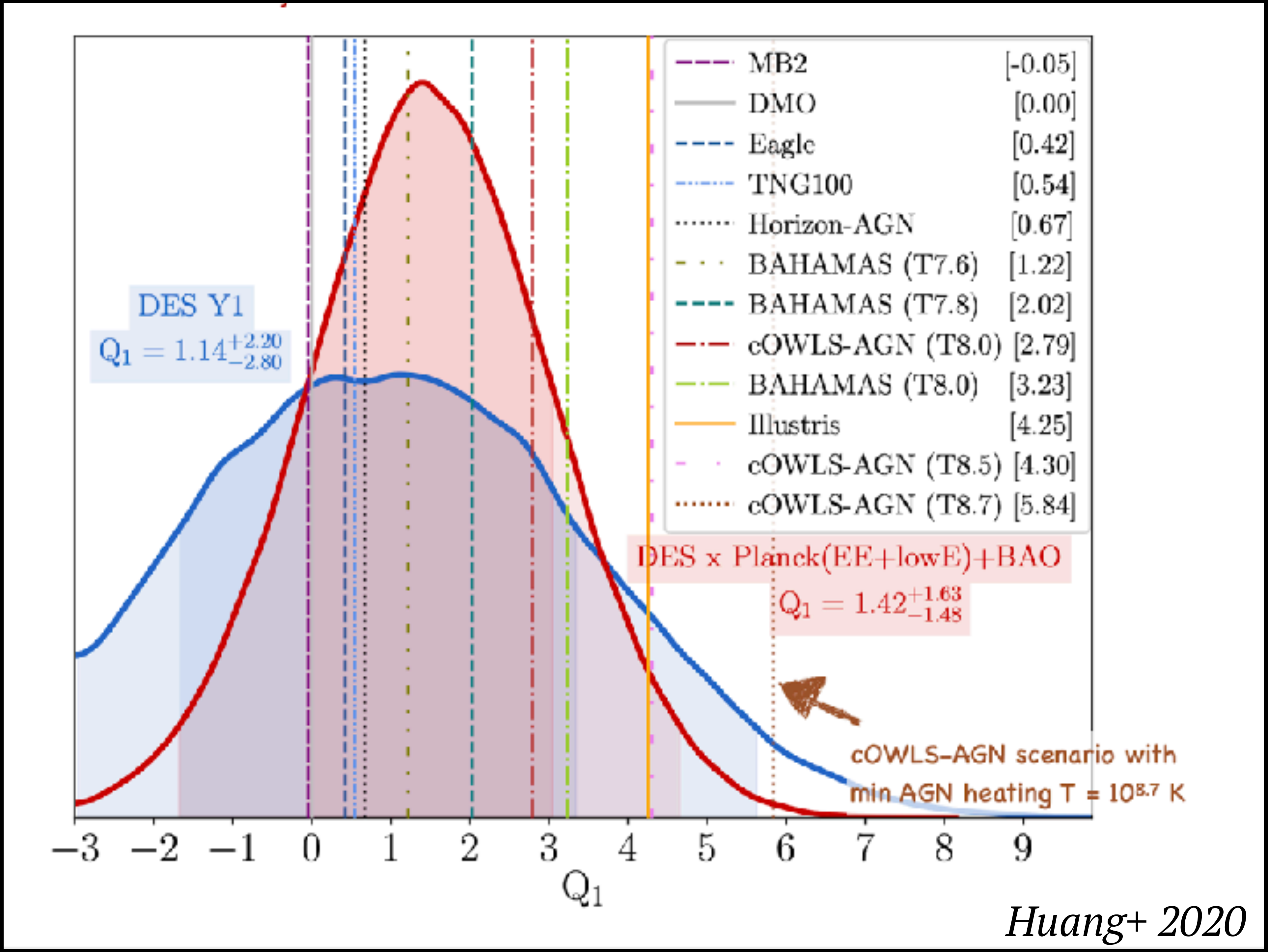
- Orange: DES-Y1 baseline (conservative scale cuts)
- Blue: DES-Y1 including all scales, and baryonic effects are modeled using **PCA with non-informative prior**
- Grey: DES-Y1 including all scales, and baryonic effects are modeled using **PCA with informative prior**

Baryonic Effects in WL Analyses
Feedback Constraints

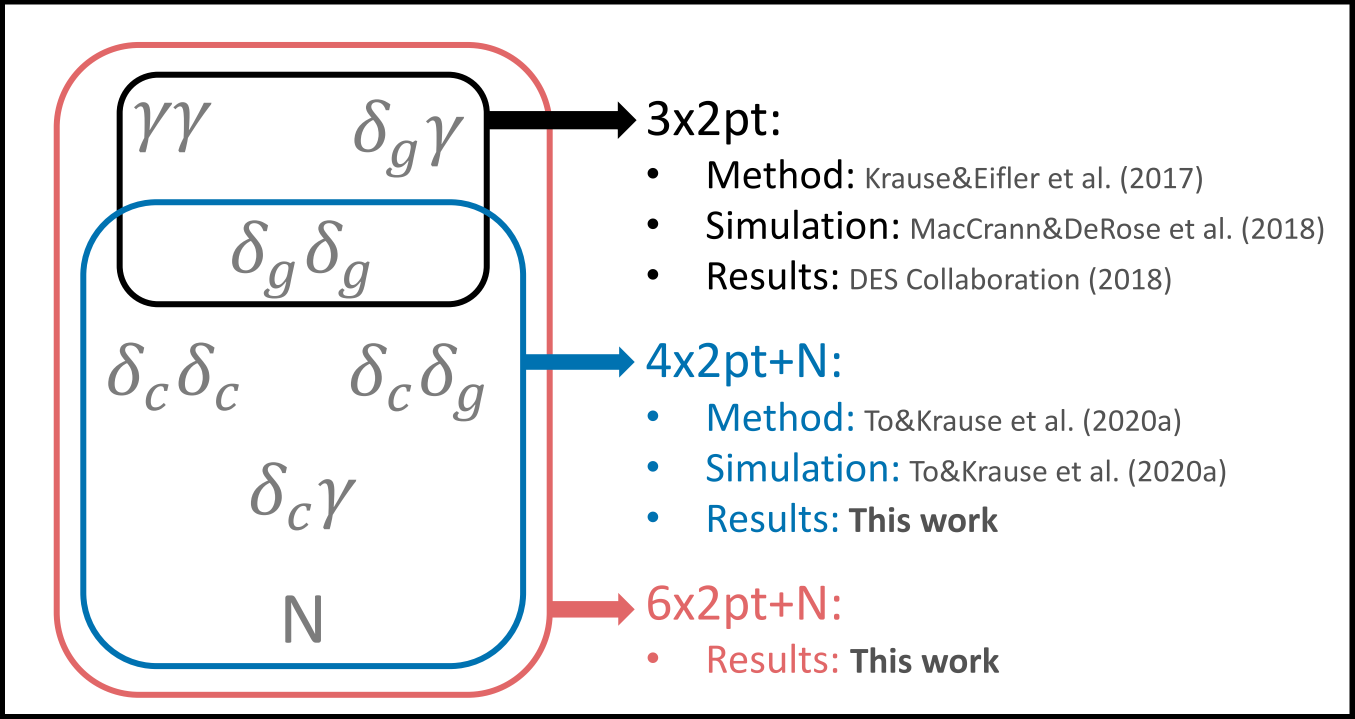


Huang+ 2020

Baryonic Effects in WL Analyses
Feedback Constraints



Systematics Opportunities and Challenges: Cluster Counts x 2PCFs



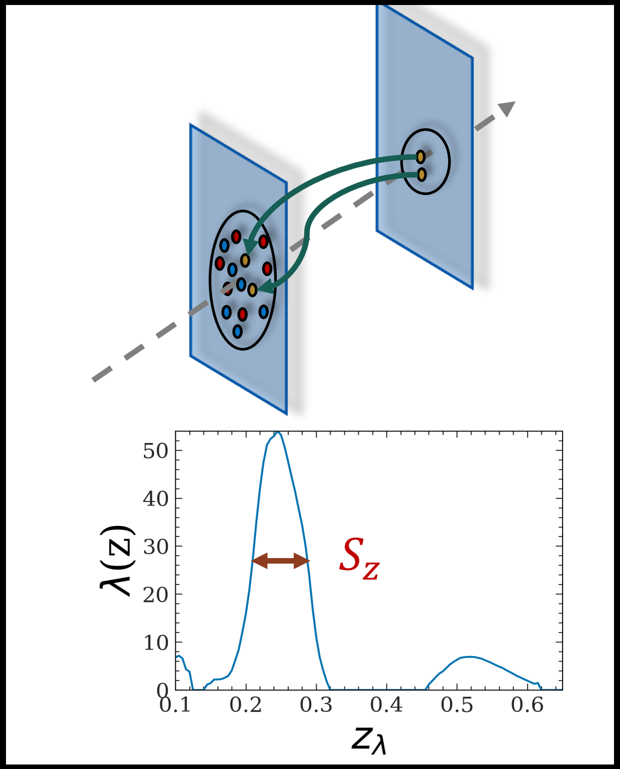
Analysis in comparison		Pros of this analysis
DES Y1 cluster analysis [DES collaboration 2020]	This analysis	
• Small scale	• Large scale, 2-halo regime	Safe from many systematics (e.g. baryonic effects, mis-centering)
• Two step analysis: Weak lensing → mass + N → Cosmology	• One step analysis: Data vector → Cosmology	Easy to be combined with other cosmology probes (e.g. 3x2pt)

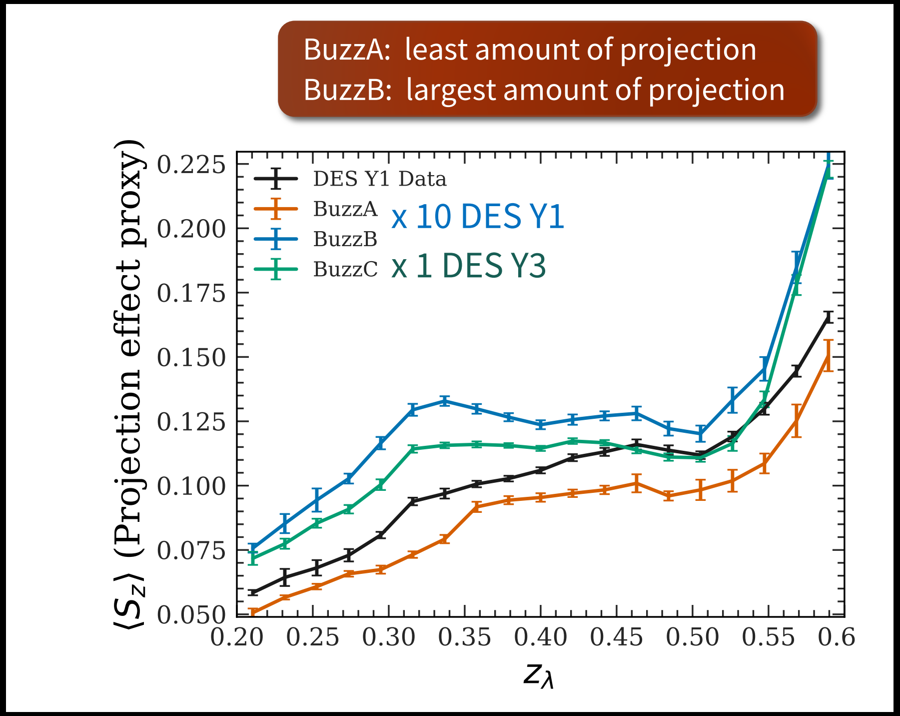
To&Krause+ 2020a,b: Cluster cosmology constraints from abundances and large-scale 2PCF

- joint likelihood analysis validated on DES-like mock catalogs (Buzzard, DeRose+2020)
- MOR calibrated from large-scale clustering, account for selection bias

✓ cosmology constraints consistent with other DES probes
(but not with main DES-Y1 cluster analysis)

Systematics Opportunities and Challenges: Cluster Counts x 2PCFs





End-to-end validation on simulations (To,Krause+2020a; DeRose+2019, Risa's talk):

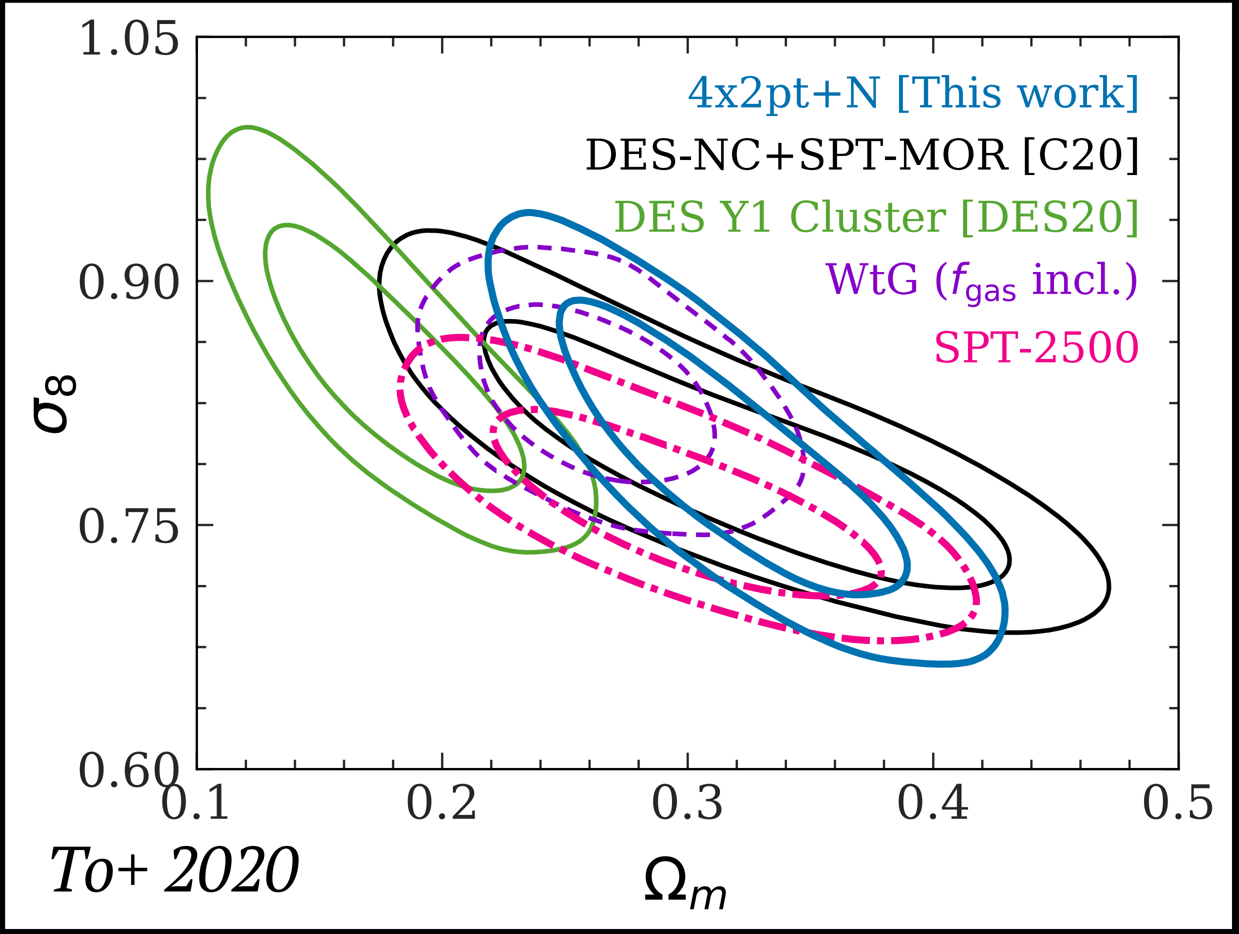
projection effects, orientation bias key systematics for cluster selection

validate analysis on custom Buzzard simulations spanning range of possible projection effects

recover input cosmology across all realizations

Systematics Opportunities and Challenges: Cluster Counts x 2PCFs

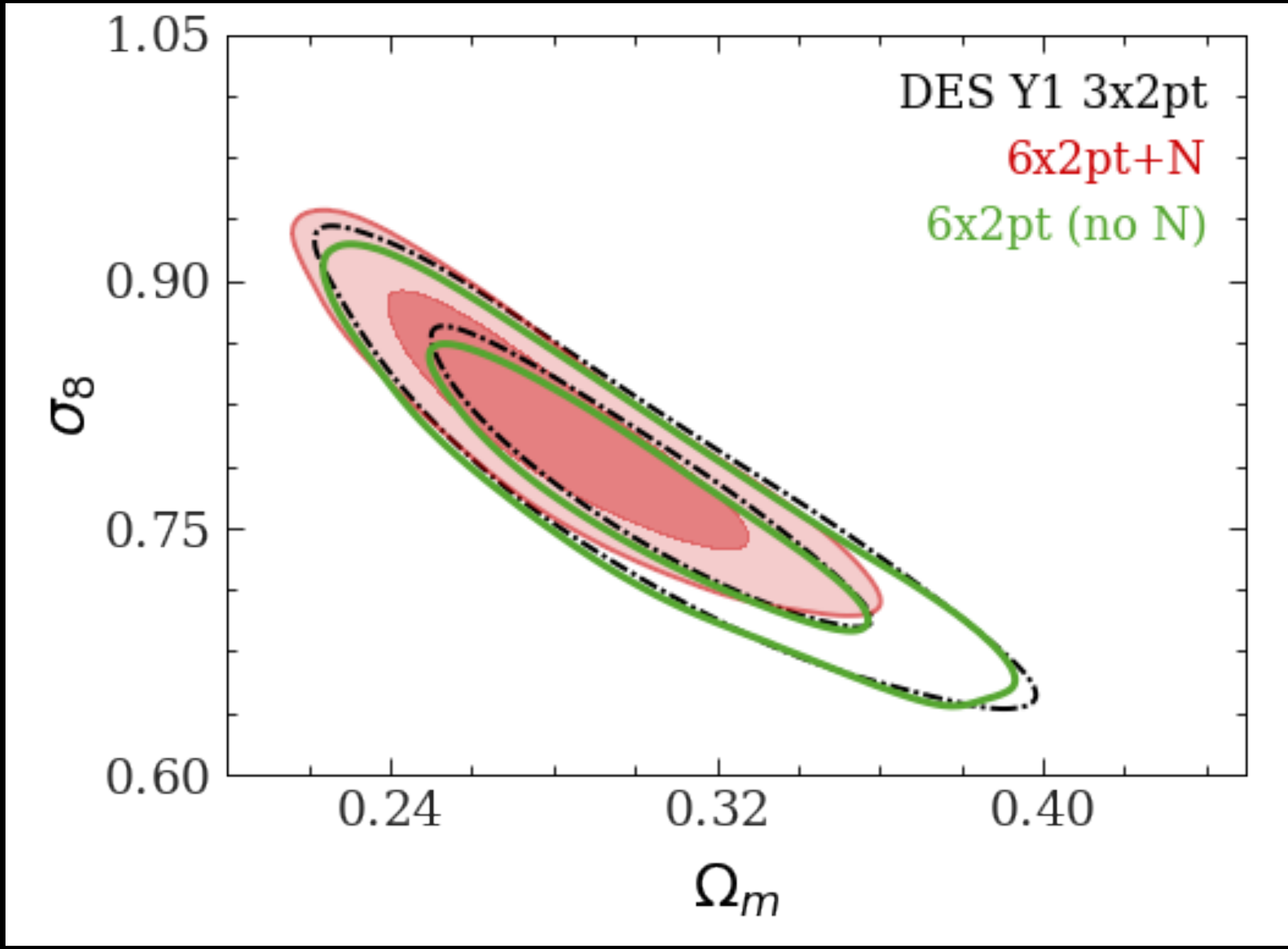
Name		
4x2pt+N	DES	4 two-point correlation (large scale only)
DES-NC +SPT-MOR	DES	SPT multiwavelength data (high mass, small scale only)
DES Y1	DES	Weak lensing (large + small scale)
WtG	X-ray	Weak lensing (high mass)
SPT-2500	SPT	SPT multiwavelength data (high mass, small scale only)

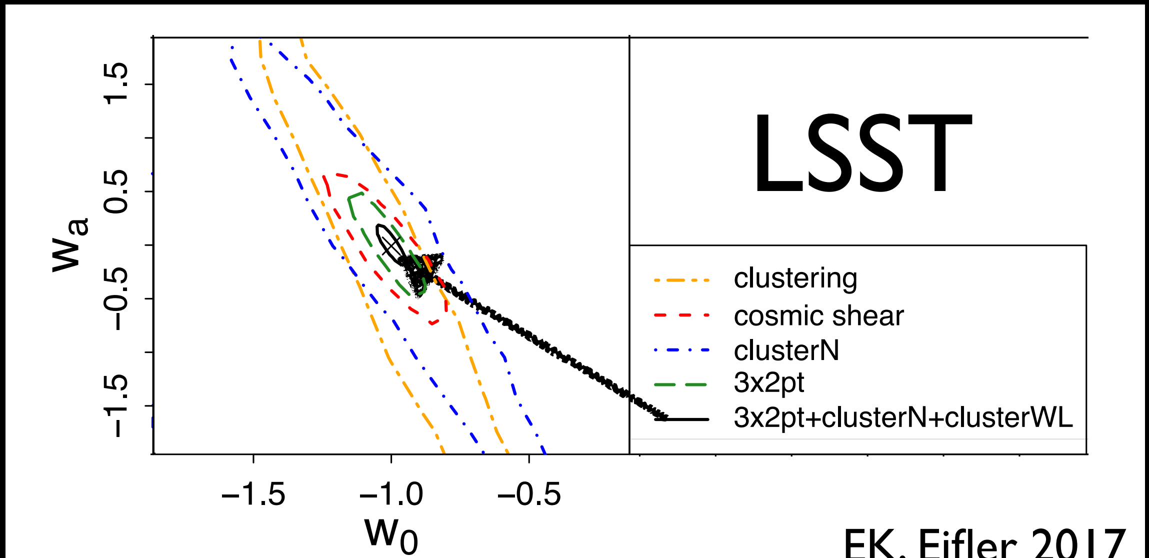


Compare 2 and 3: DES-Y1 problem associated with low-mass clusters
Compare 1 and 3: DES-Y1 problem associated with small-scale lensing

Systematics Opportunities and Challenges: Cluster Counts x 2PCFs

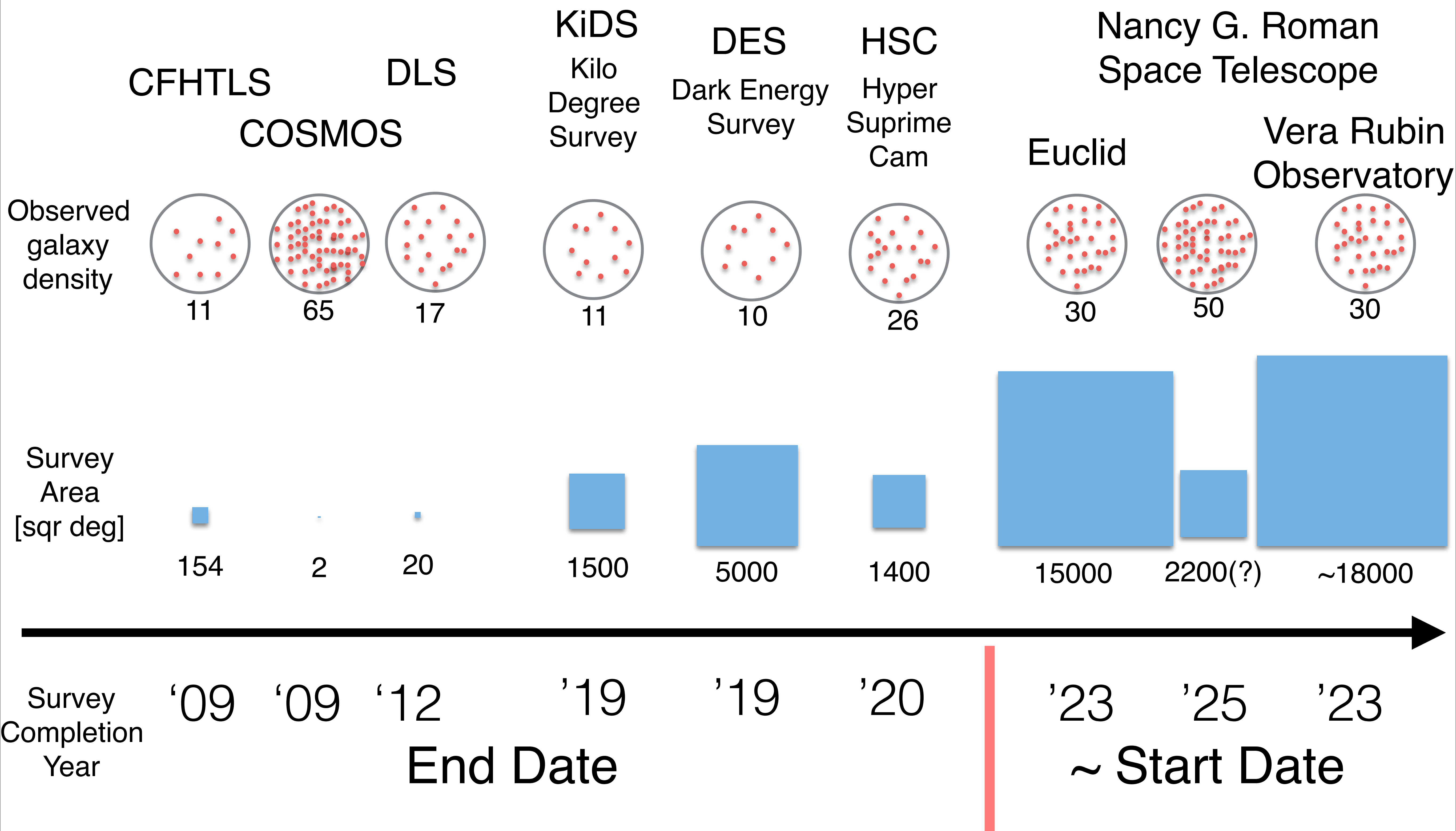
Additional constraining power from number counts, not additional 2PCFs
Much constraining power to be gained iff accurate MOR calibration



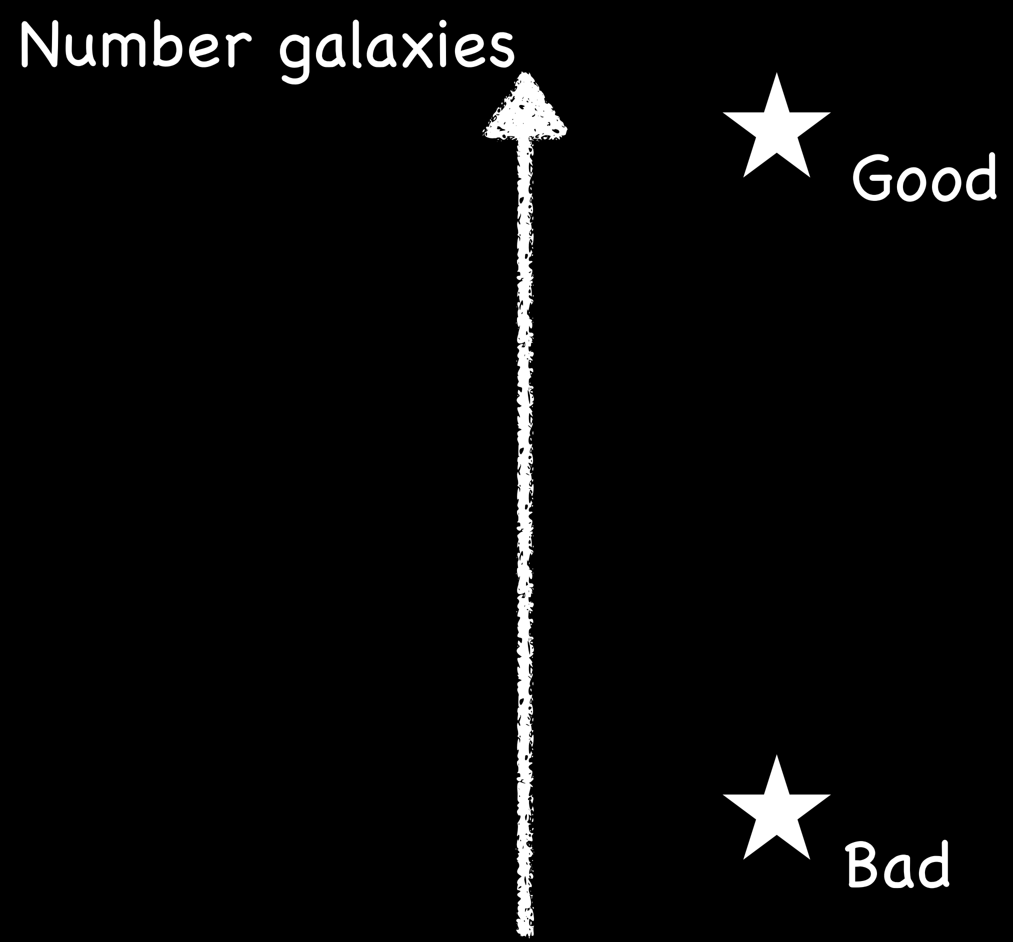


EK, Eifler 2017

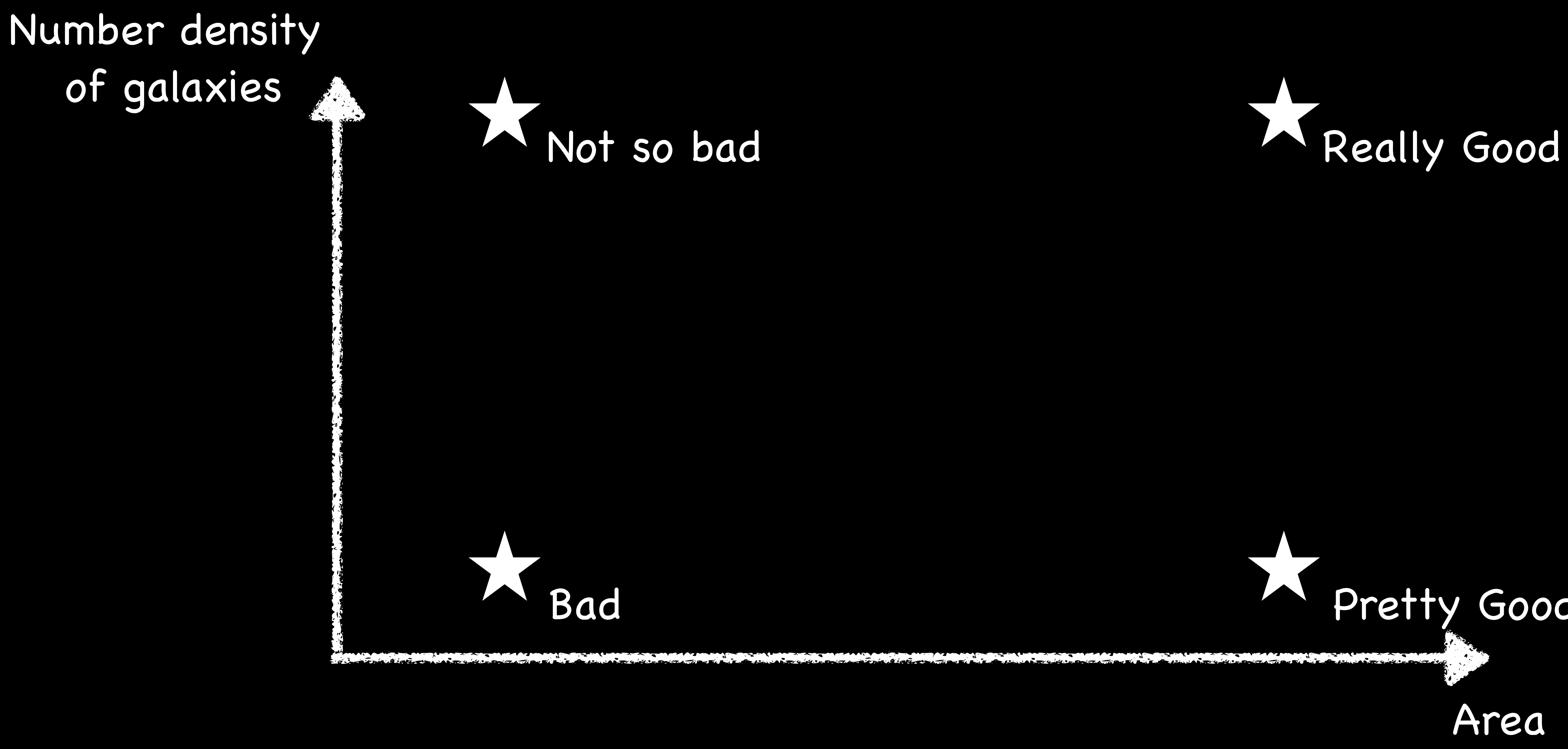
Photometric Dark Energy Surveys



Survey Optimization I



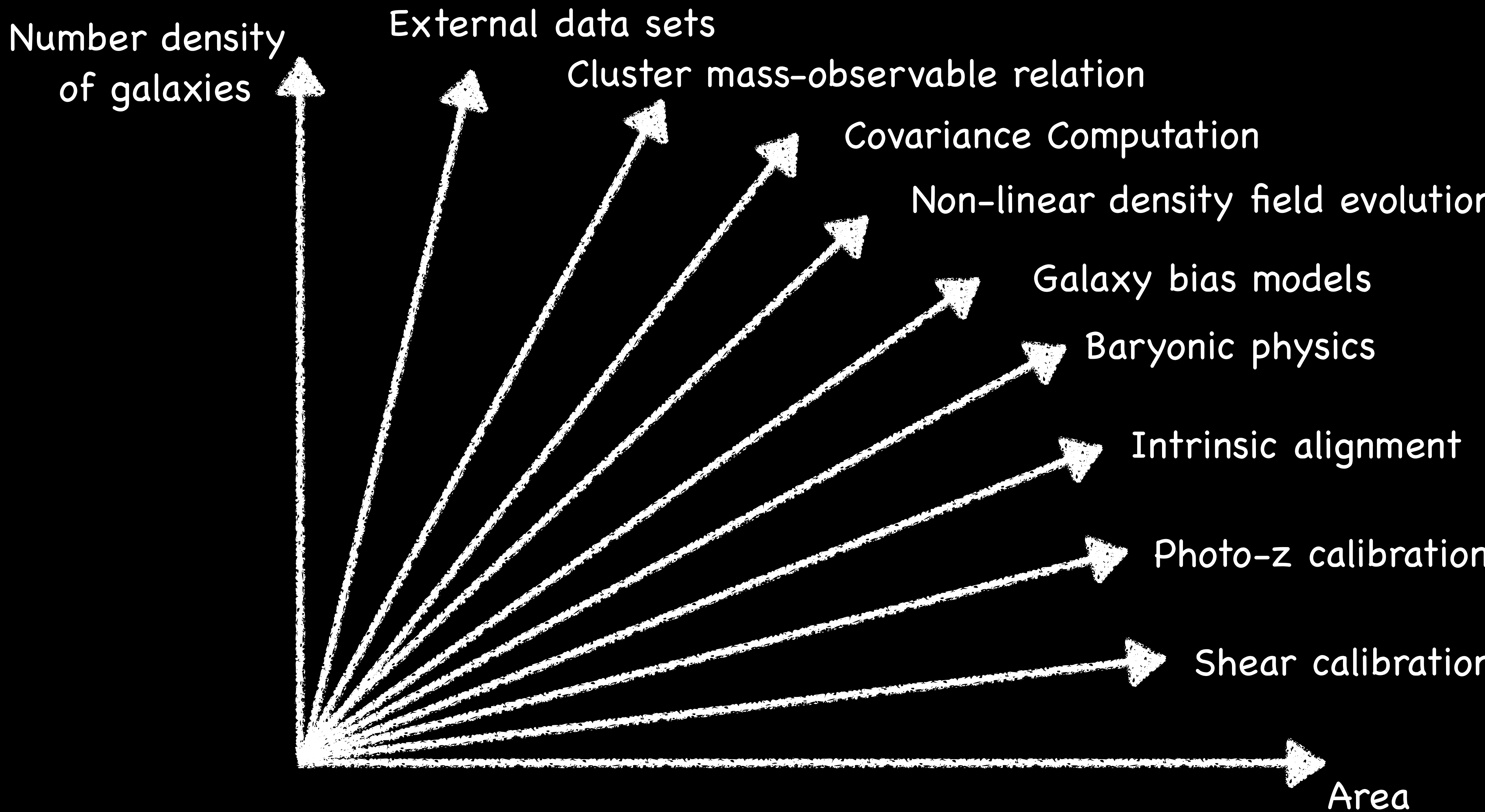
Survey Optimization II



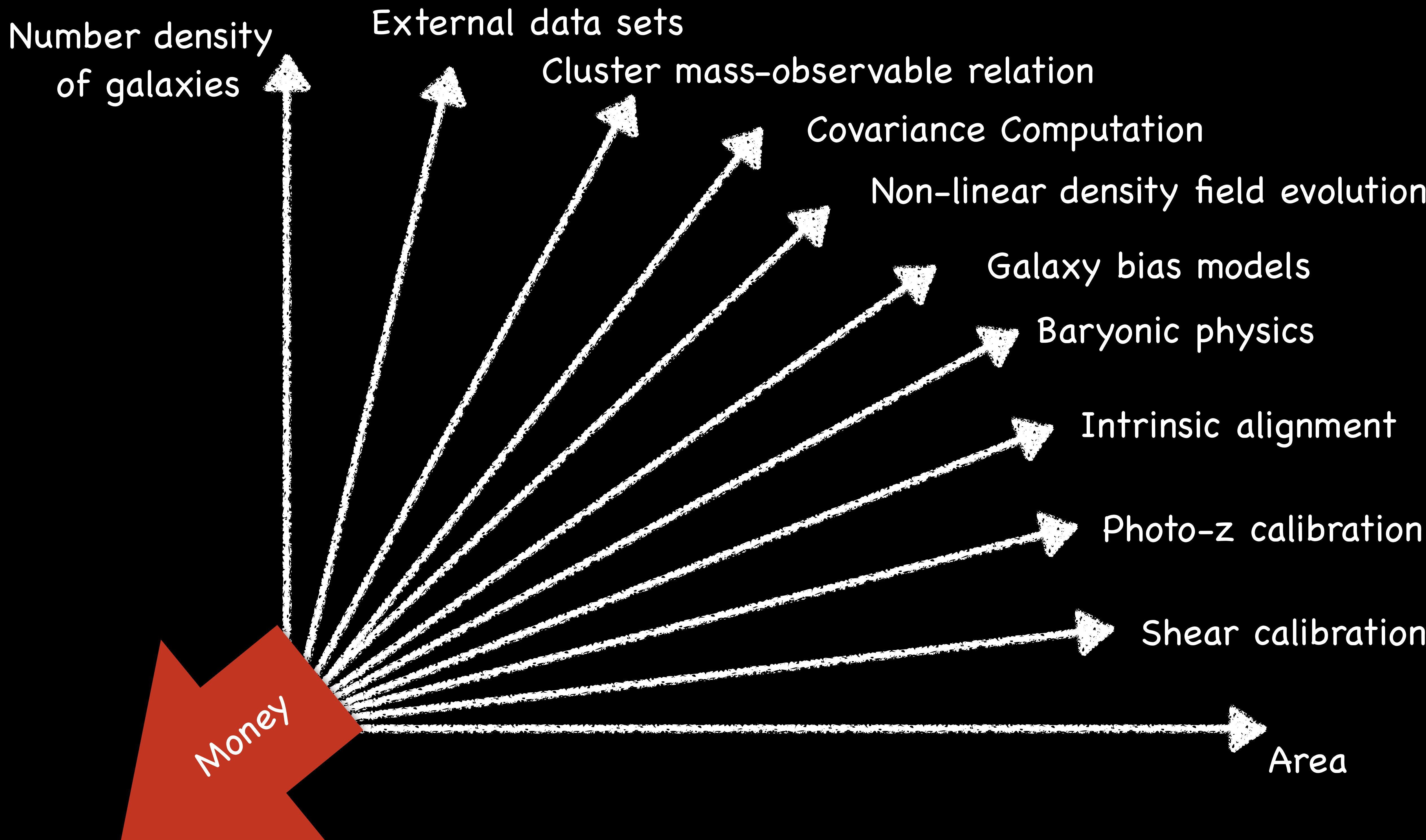
Statistical error bars only (simplified):

- Area is more important than depth
- Even more true since non-gaussian Covariances became fashionable

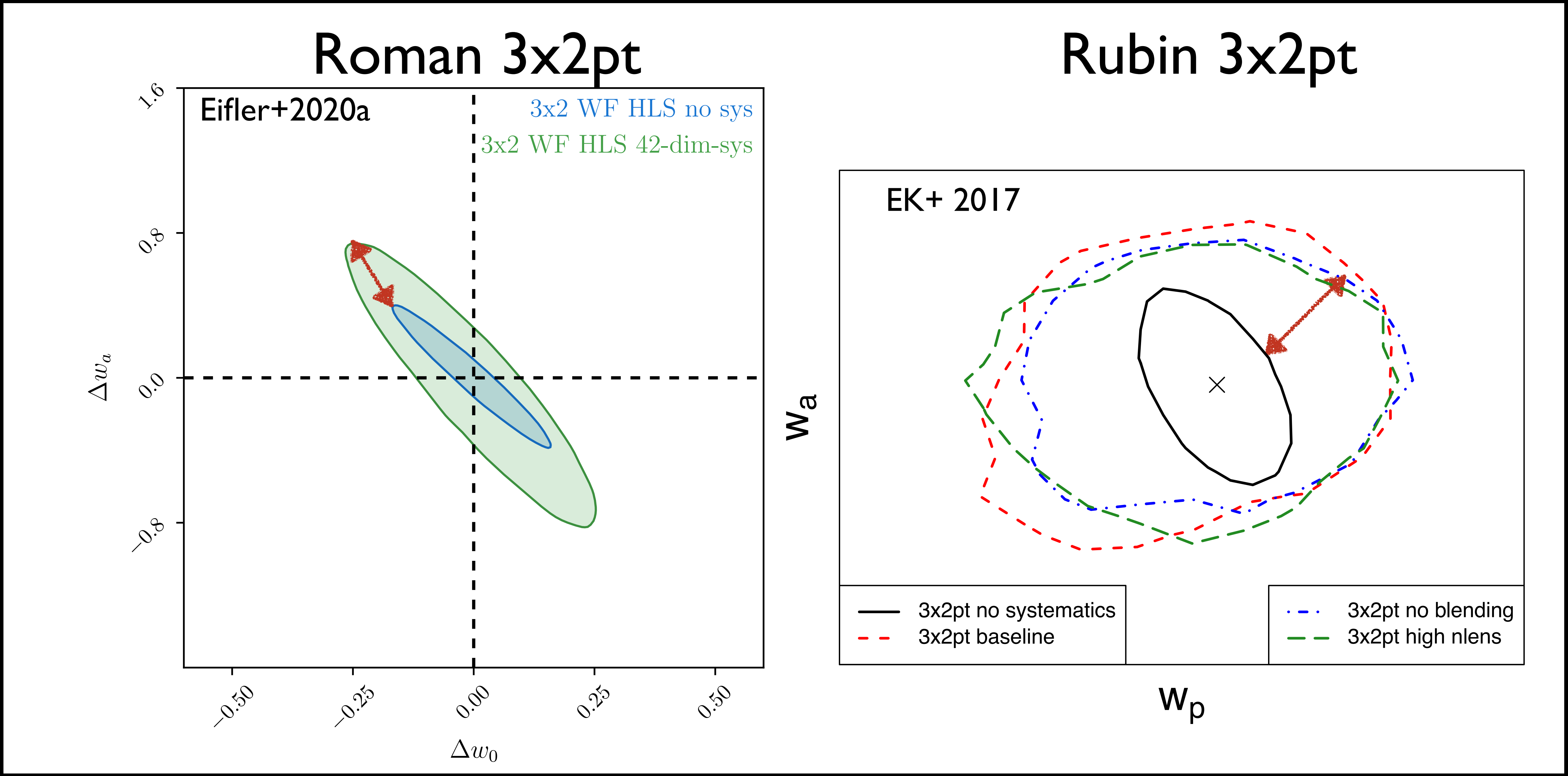
Survey Optimization III



Survey Optimization III



Stage-IV 3x2pt forecasts
(*details matter*)

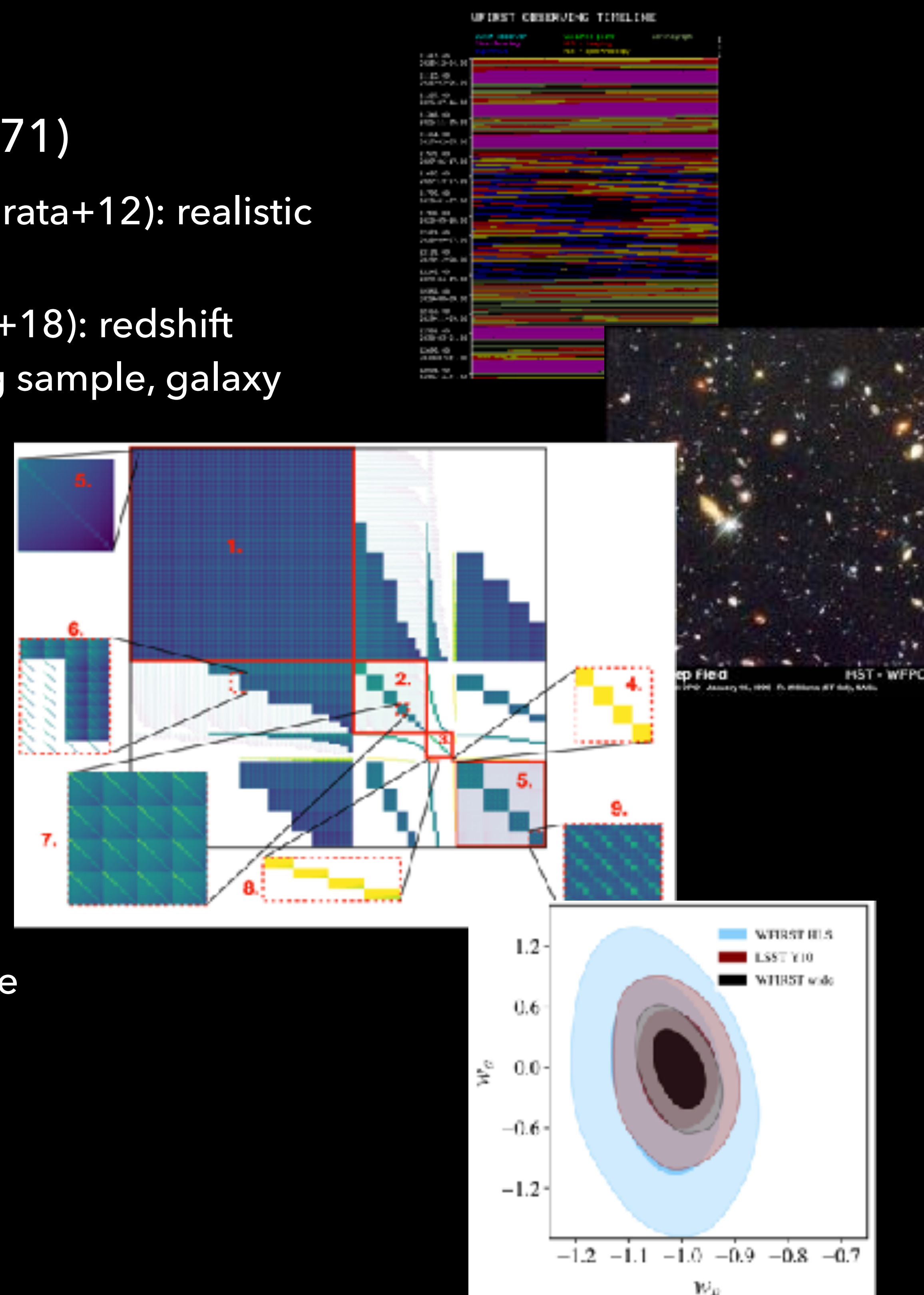


marginalized over {linear galaxy bias, lens photo-z, source photo-z} per tomography bin

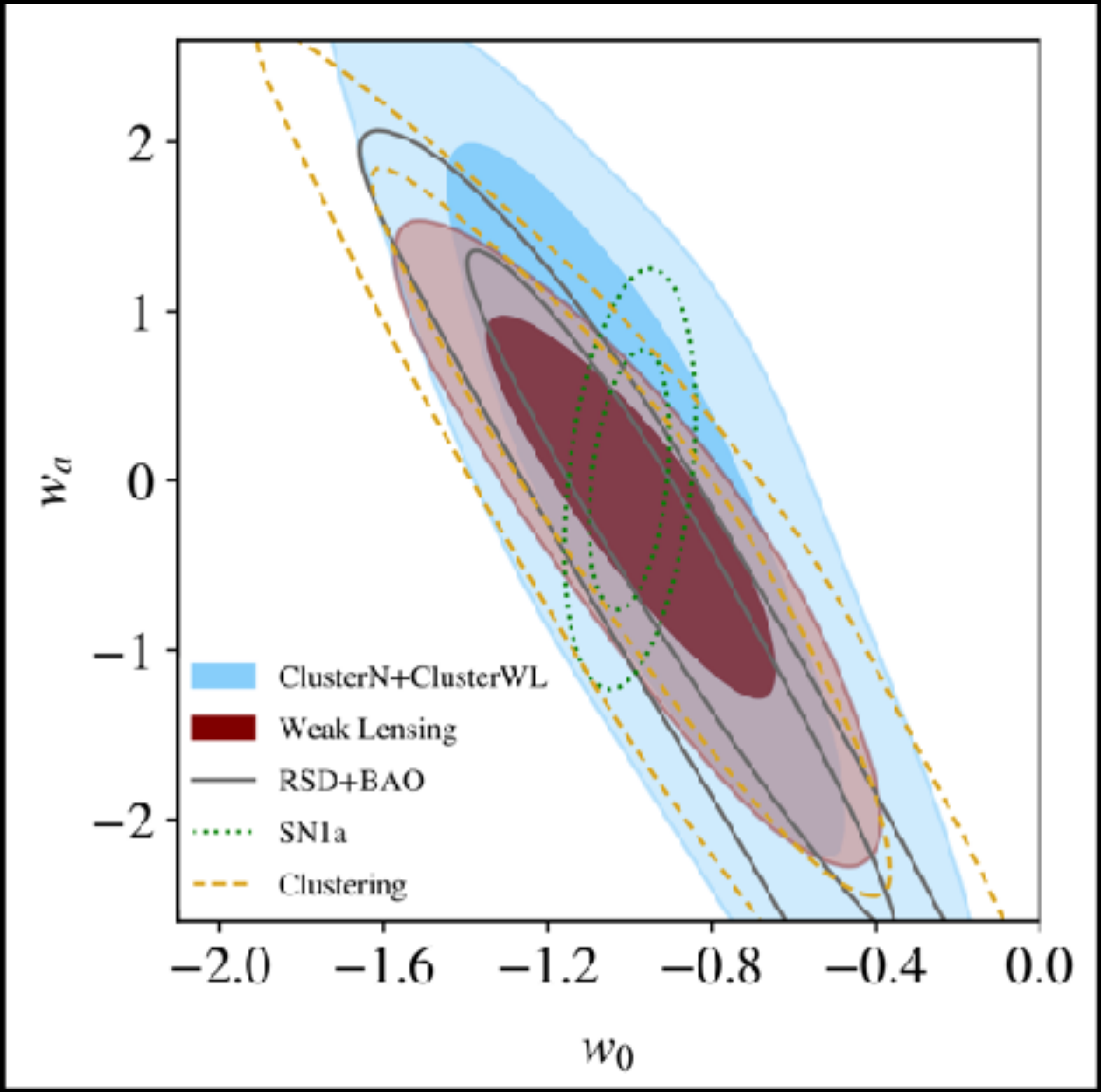
Roman Space Telescope Forecasting

Forecast Machinery (Eifler+2004.05271)

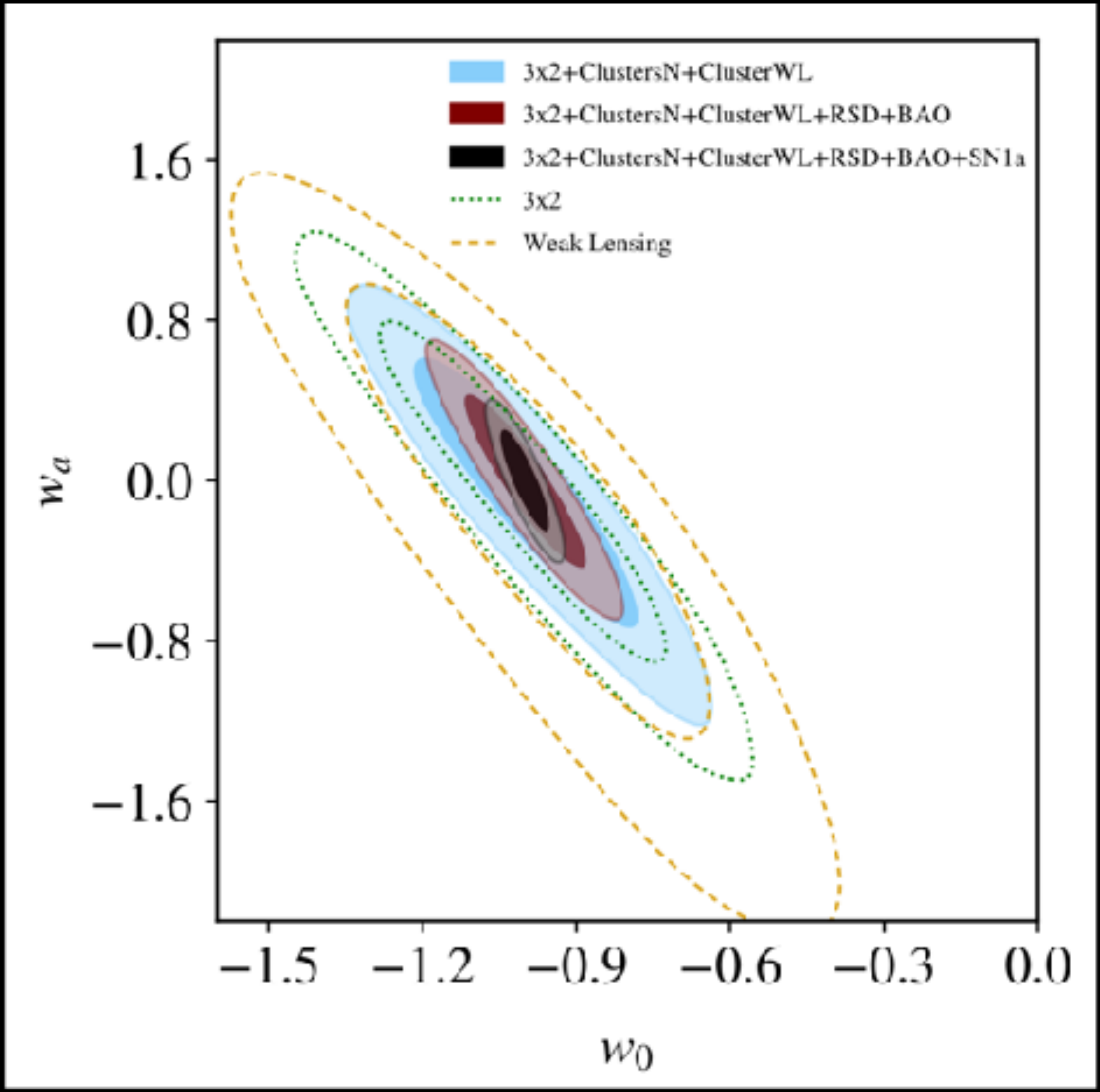
- WFIRST Exposure Time Calculator (Hirata+12): realistic survey area + depth
- CANDELS WFIRST catalog (Hemmati+18): redshift distribution for lensing and clustering sample, galaxy clusters
- Combine
    - Cosmic shear
    - Galaxy-Galaxy Lensing
    - Galaxy Clustering (photo)
    - Cluster Number Counts
    - Cluster Weak Lensing
    - Galaxy Clustering (Spectro)
    - SN1a (Hounsell+2018)
- Non-Gaussian Multi-Probe Covariance
- 80+ systematic parameters
- full simulated likelihood analyses



Roman Forecasts: Reference Survey

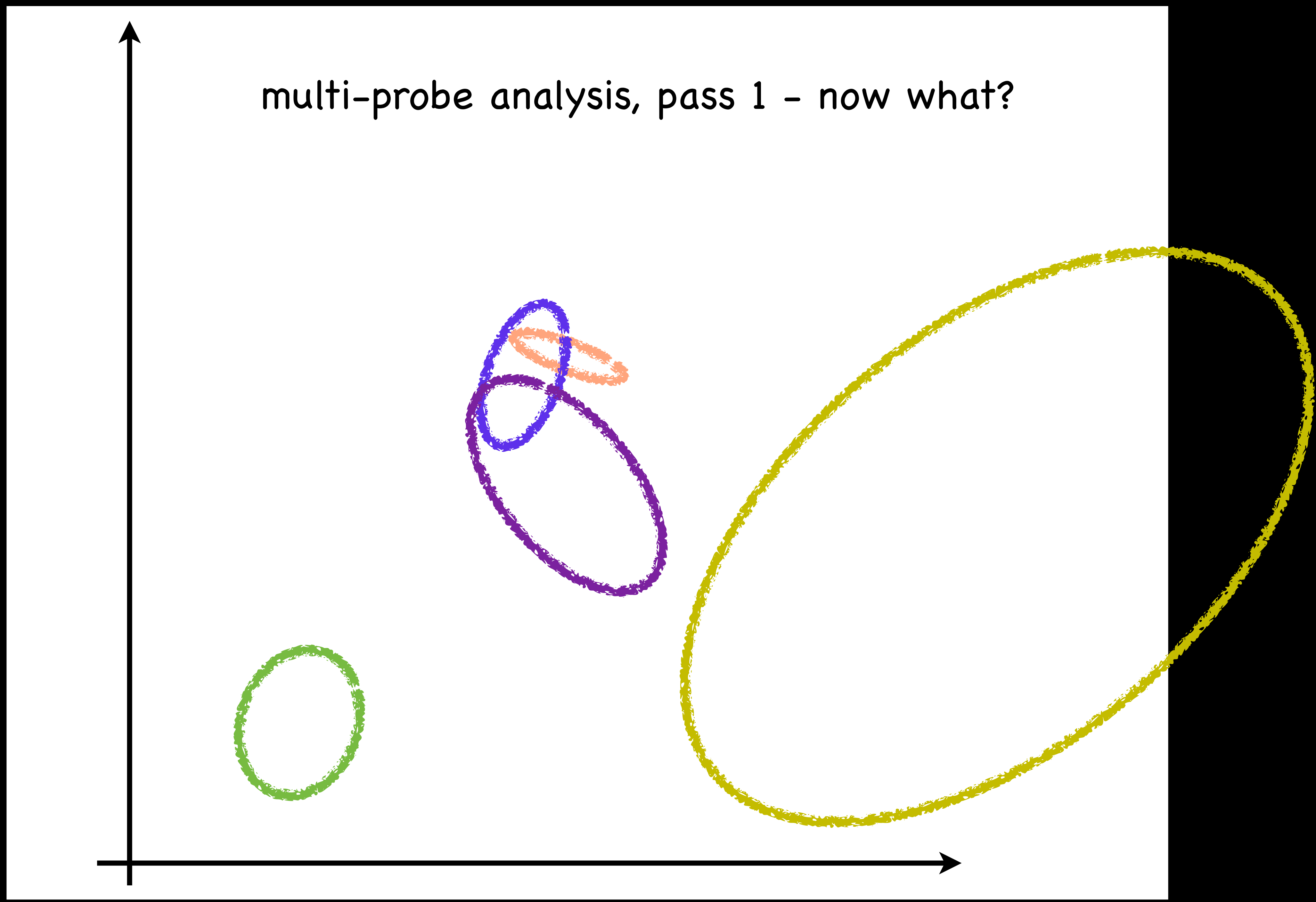


individual probes



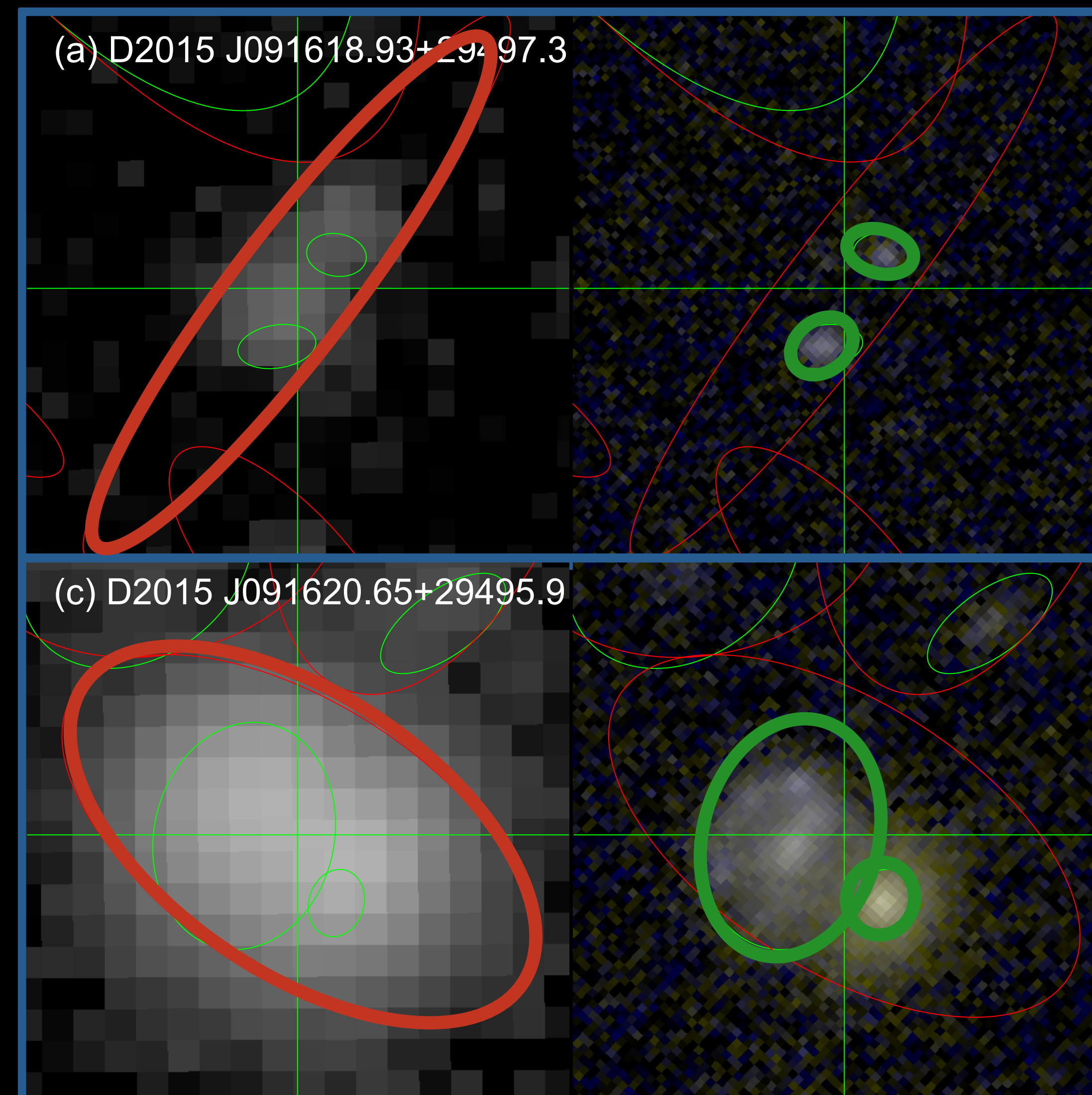
combined probes

Unknown Systematics? vs. New Physics?



Unknown Systematics? vs. New Physics?

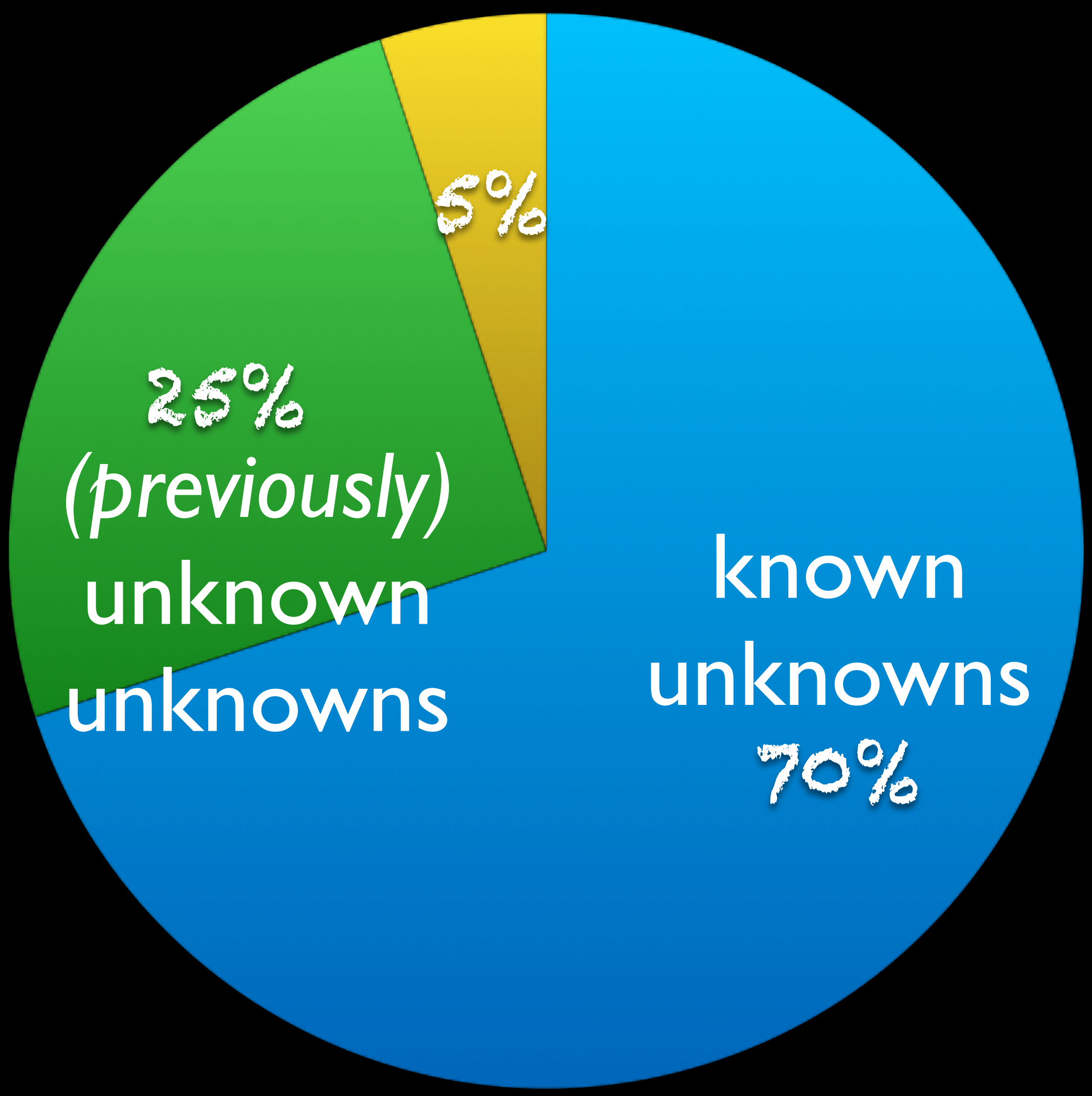
- scale dependence?
- dependence on galaxy/cluster selection?
- calibrate with more accurate measurements
  - spectroscopic redshifts
  - low-scatter cluster mass proxies
  - galaxy shapes from space-based imaging [potentially expensive]
- correlate with other surveys
  - compare to predicted cross-correlations
  - constrain uncorrelated systematics



Subaru HST-ACS
ground vs. space-based shape measurements
Dawson+ 2016

Cosmology Analysis Parameters

Cosmology Parameters



Systematics Parameters

- observational systematics
  - survey specific
- astrophysical systematics
  - observable + survey specific

Conclusions

- We're entering the decade of very large galaxy surveys
  - BOSS, KiDS,DES, HSC, PFS -> DESI, LSST, Euclid,WFIRST,…
  - + radio surveys: impressive forecasts, complementary systematics
- (Most) cosmological constraints will be systematics limited
  - require accurate systematics *parameterizations+priors*
- Need different probes and analysis methods to enable accurate cosmology
  - identify and understand systematics effects
  - maximize constraining power
- Precision cosmology requires collaboration across surveys + wavelengths, plan for analysis frameworks to combine data from all surveys

Always working on that next code…

CoCoA (Cobaya-CosmoLike Architecture)

Lead developer of the framework: Vivian Miranda, https://github.com/CosmoLike/cocoa

- Idea: Combine the python Cobaya framework with CosmoLike
- Access to samplers+external likelihood of Cobaya
- Automated updates of camb/class
- Large scale structure modeling of observables, systematics and multi-probe covariances from CosmoLike
- Each project gets its own likelihood module

- Easy installation and updating on usual suspects HPC systems (Pleiades, NERSC)
- Docker version available

Always working on that next code…

CosmoCov

Lead developer: Xiao Fang, https://github.com/CosmoLike/CosmoCov

Idea: Reliable analytic covariances for 3x2pt

Features:

- Non-Gaussian connected terms+ SuperSample terms
- Response covariance module
- Flat and curved sky
- Stable and fast Real Space transformation using 2D FFTlog algorithm
- Near future extensions beyond 3x2pt (code exists but not public yet)
  - CMB lensing
  - Clusters
  - tSZ 2pt functions