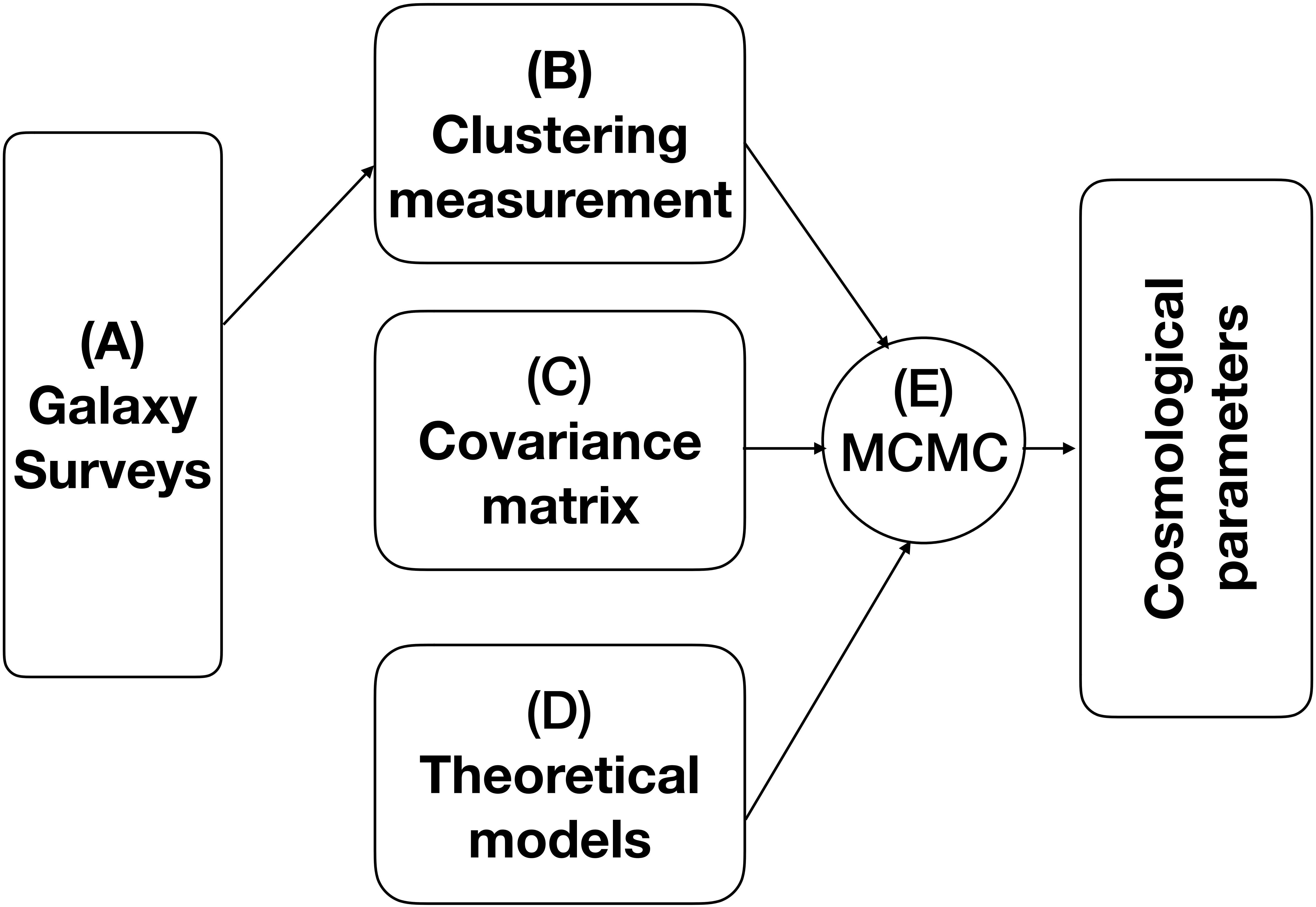
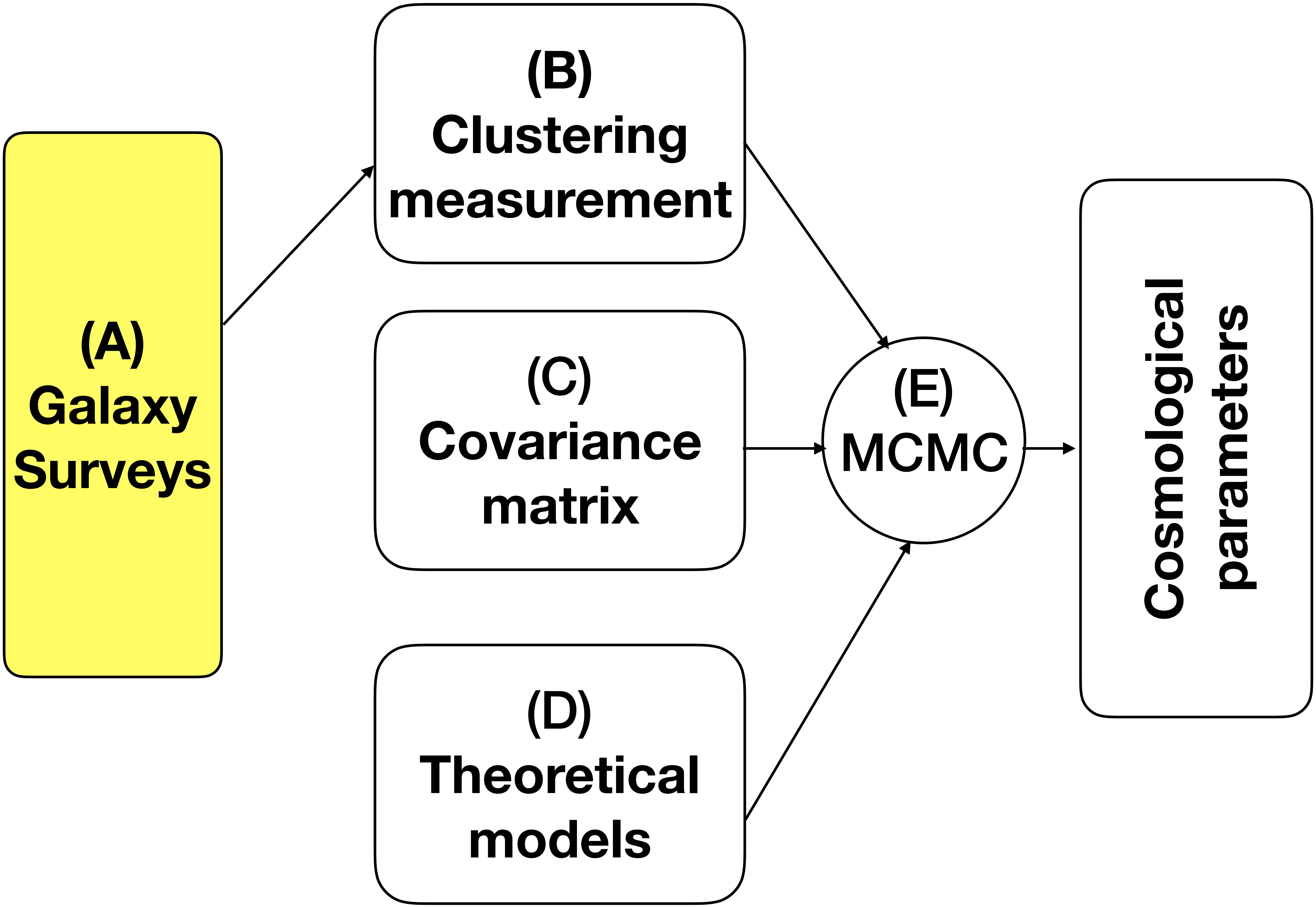
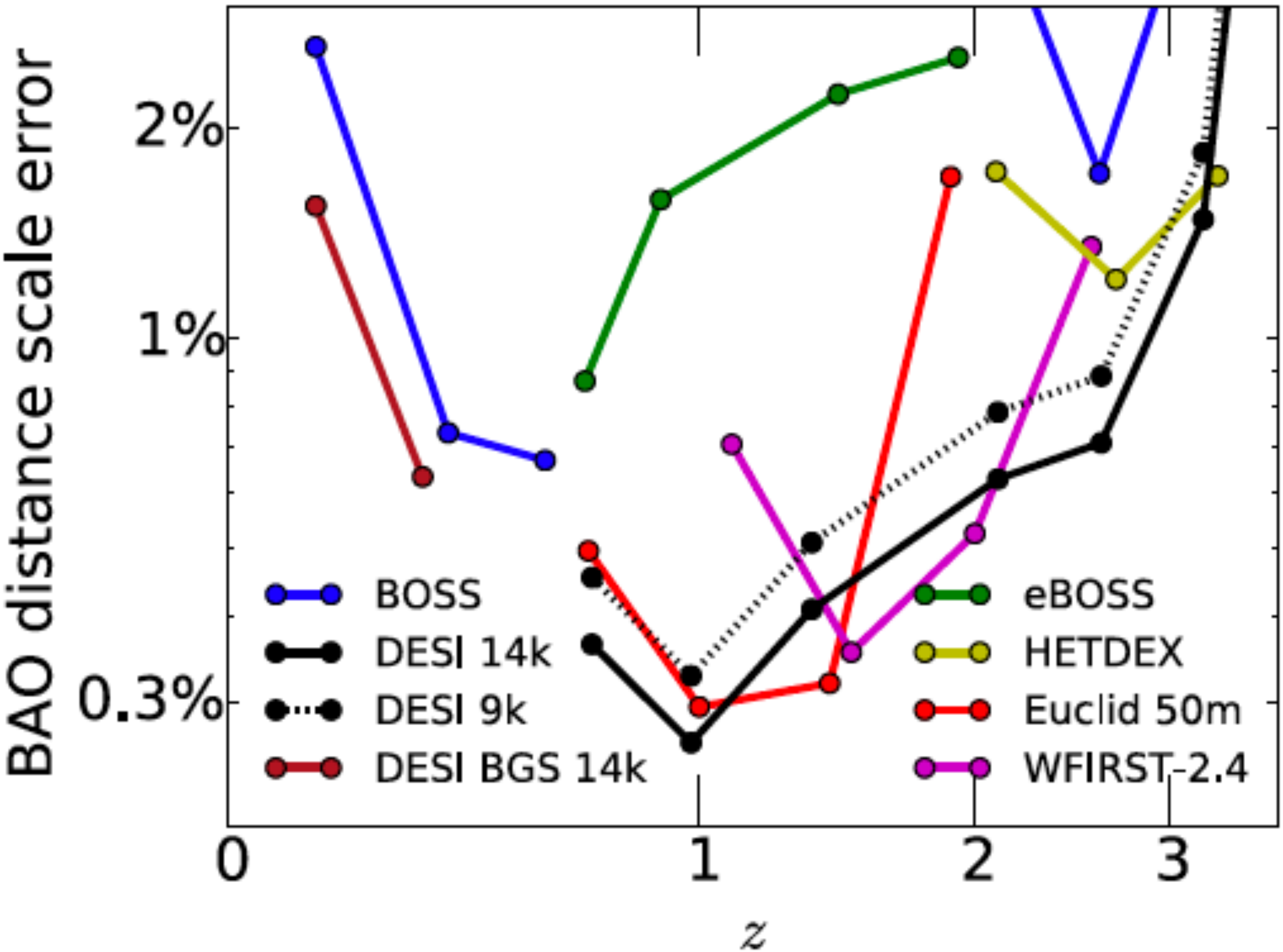
Accurate and precise cosmological constraints from the large-scale structure of galaxy clustering

Chia-Hsun Chuang
Research Scientist
KIPAC (Stanford/SLAC)

KIPAC KAVLI INSTITUTE FOR PARTICLE ASTROPHYSICS AND COSMOLOGY


(A)
Galaxy Surveys
(B)
Clustering measurement
(C)
Covariance matrix
(D)
Theoretical models
(E)
MCMC
Cosmological parameters


(A)
Galaxy
Surveys
(B)
Clustering
measurement
(C)
Covariance
matrix
(D)
Theoretical
models
(E)
MCMC
Cosmological
parameters


BAO distance scale error
2%
1%
0.3%
0
1
2
3
z
BOSS
DESI 14k
DESI 9k
DESI BGS 14k
eBOSS
HETDEX
Euclid 50m
WFIRST-2.4

DESI (Dark Energy Spectroscopic Instrument)

- 4-meter Mayall Telescope at Kitt Peak National Observatory. (Kitt Peak National Observatory (KNPO) is located 56 miles southwest of Tucson, Arizona)

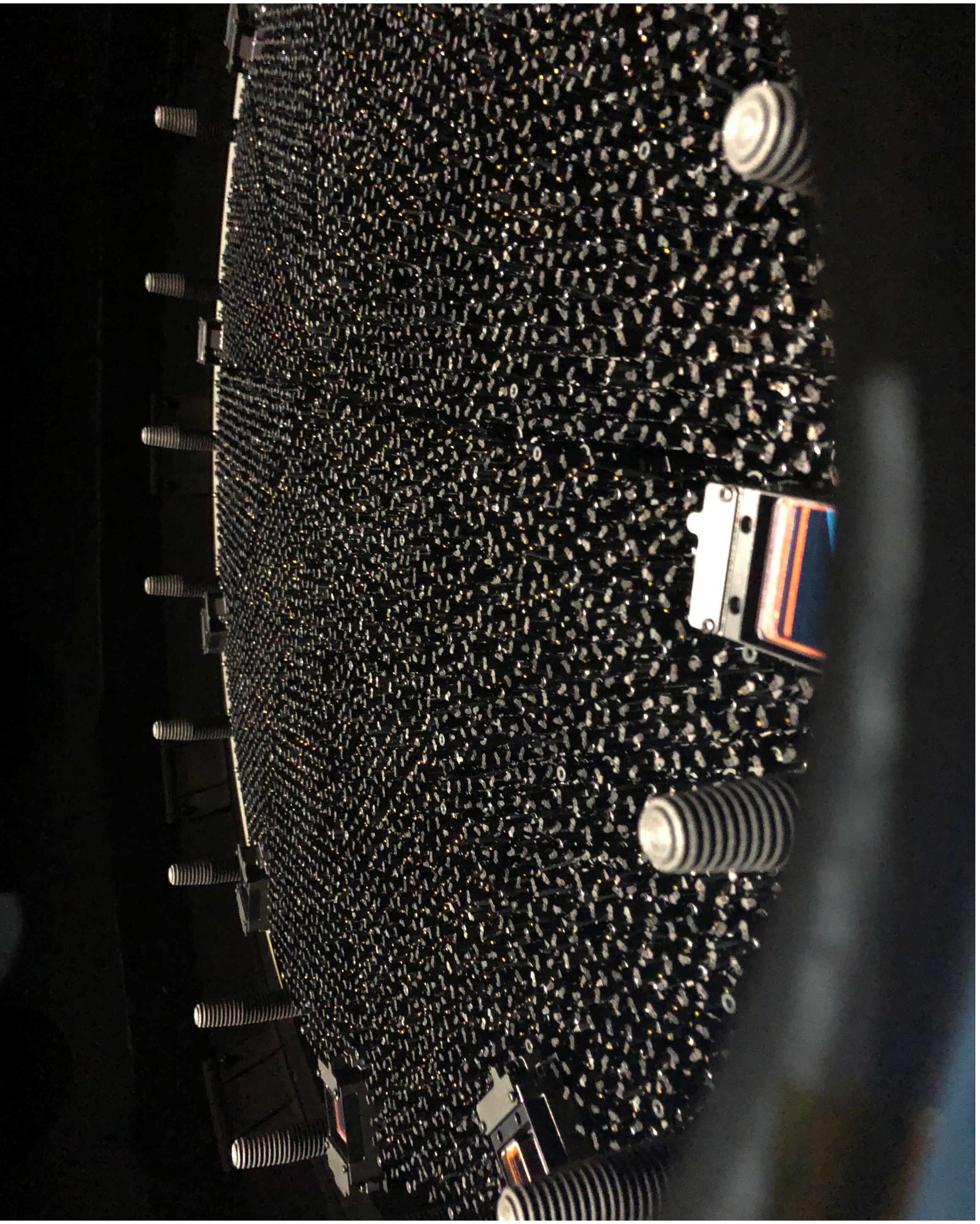
- 5000 robotic fiber positioners

- DESI will measure the spectra of more than 30 million galaxies and quasars covering 14,000 square degrees

- The spectrographs will cover a spectral range of 360 nanometers (nm) to 980 nm with a resolution of 2,000 to 5,000, enabling DESI to probe redshifts up to 1.7 for emission line galaxies and 3.5 for the Lyman-α spectra from quasars

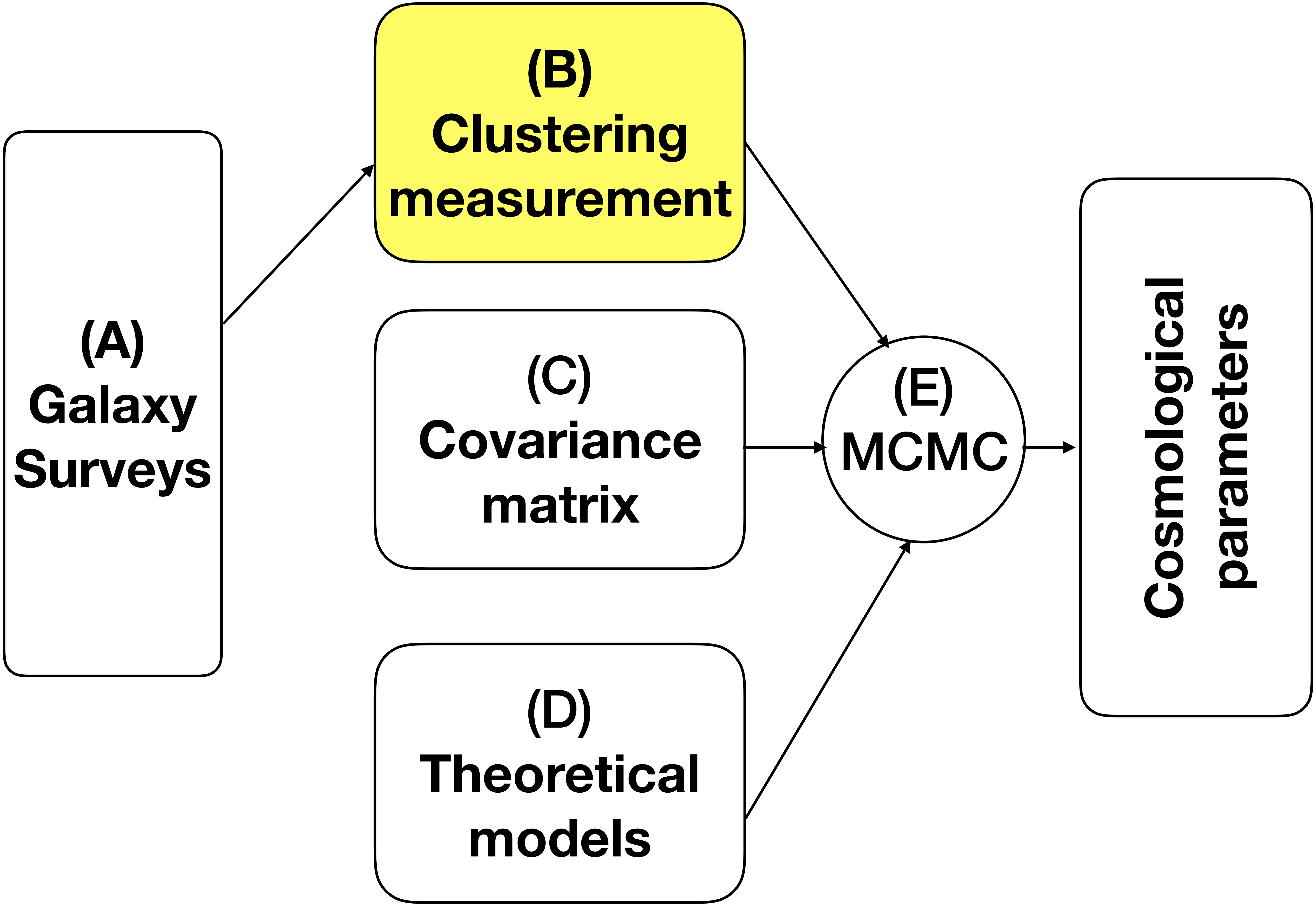
- We have the first light a few weeks ago!!!

Positioners on the focal plate!

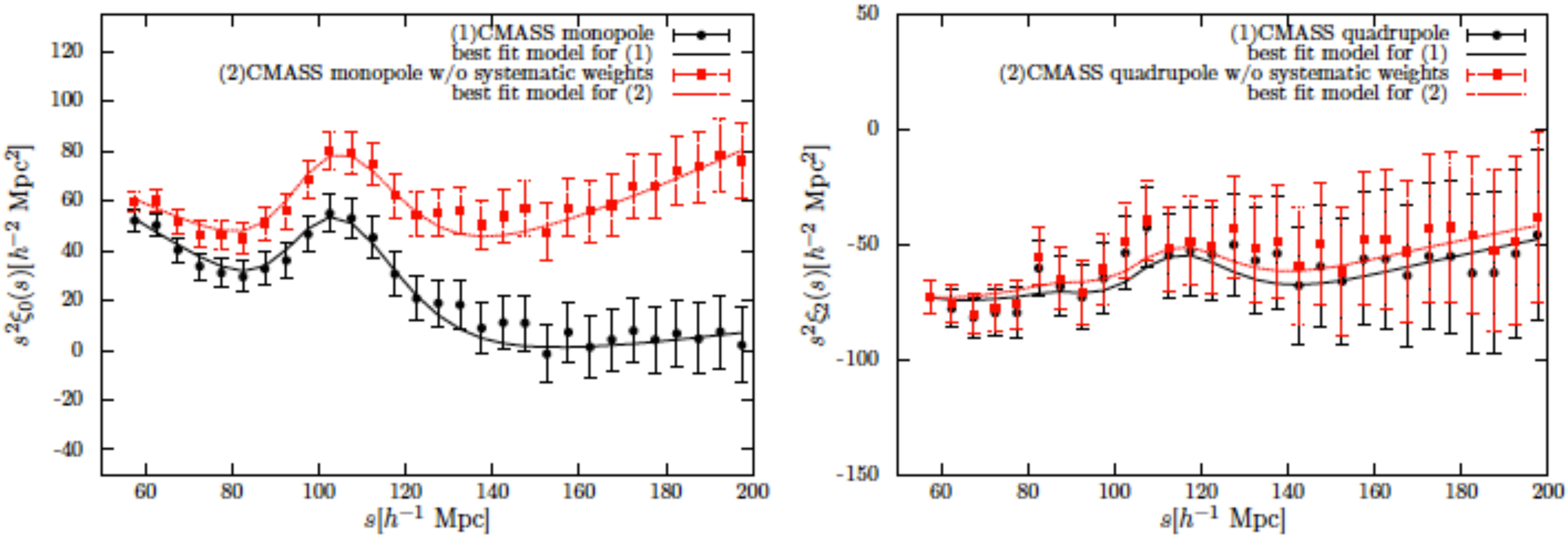


5000 positioners at the focal plate

The other side of the focal plate

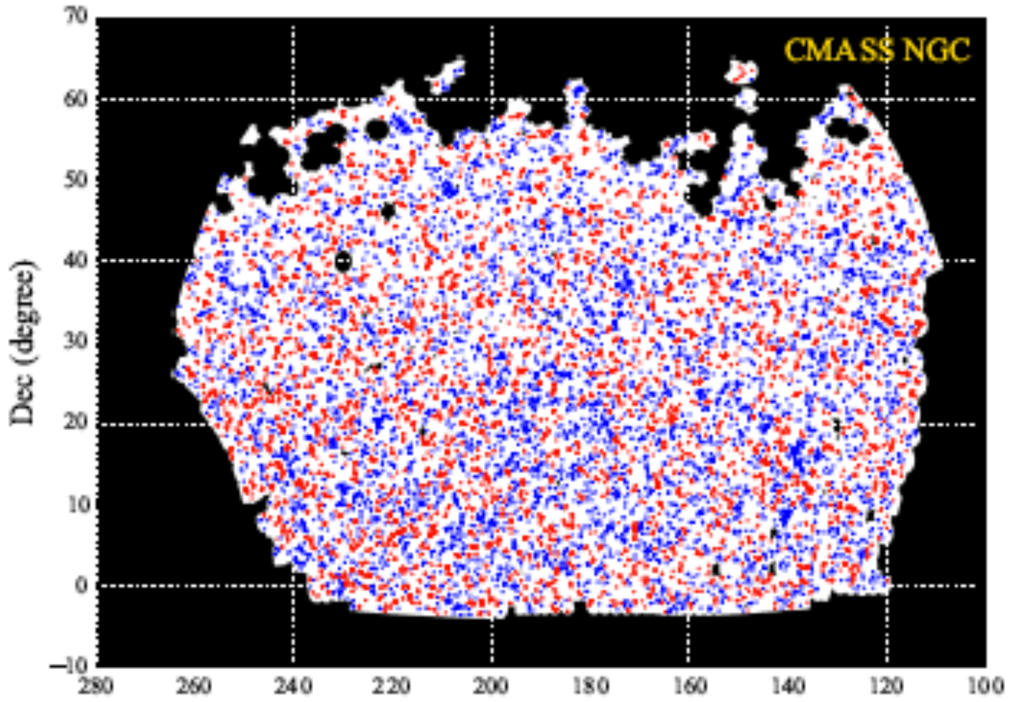

(A)
Galaxy
Surveys
(B)
Clustering
measurement
(C)
Covariance
matrix
(D)
Theoretical
models
(E)
MCMC
Cosmological
parameters

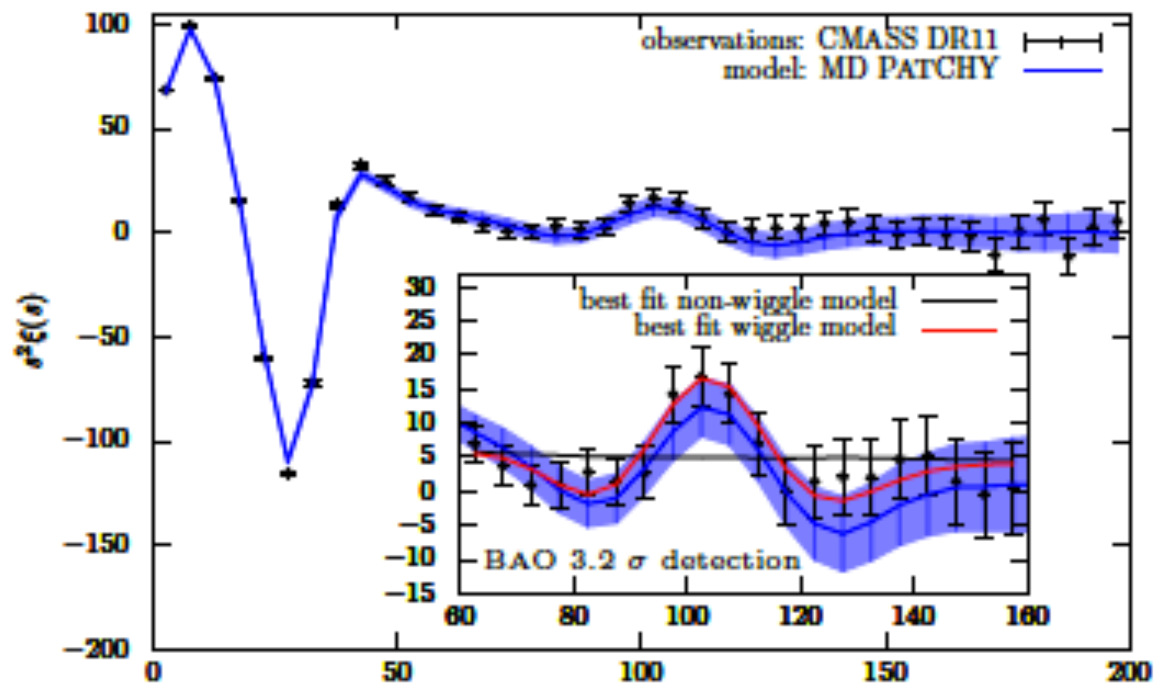
Galaxy clustering from BOSS (Chuang et al. 2017)



Correct the systematics from star density

Void clustering from BOSS
(Kitaura, Chuang et al. 2016; Zhao, Chuang, et al. 2018)





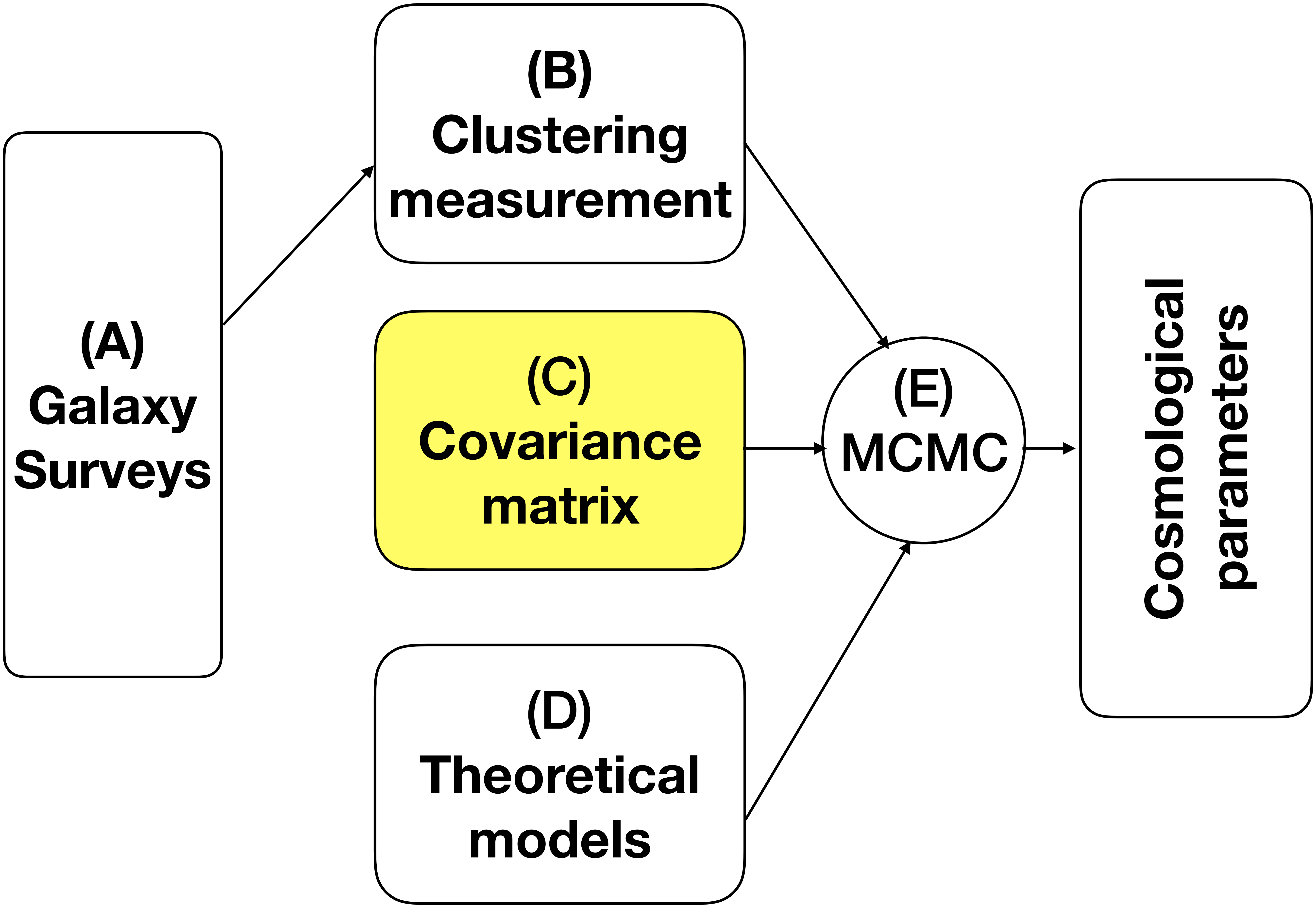
α	0.2 < z < 0.5	0.5 < z < 0.75
galaxy	0.9981 ± 0.0132	0.9996 ± 0.0123
void	0.9962 ± 0.0202	1.0177 ± 0.0575
combine (w = –0.07)	0.9981 ± 0.0114	0.9998 ± 0.0110
Improvement on σ_α	13.7%	11.1%
galaxy (Vargas-Magana et al.)	0.9986 ± 0.0136	1.0007 ± 0.0121

The gain is equivalent to increasing the survey volume by 20% on top of BAO reconstruction technique

Can we improve growth rate measurement by including voids as well?

- Very challenging to get unbiased measurements!

Voids are defined based on the galaxy sample. The selection has suffered the redshift distortion effect in the galaxy sample. (see Chuang, Kitaura et al. 2017)


(A)
Galaxy
Surveys
(B)
Clustering
measurement
(C)
Covariance
matrix
(D)
Theoretical
models
(E)
MCMC
Cosmological
parameters

How to estimate the covariance matrix for galaxy clustering?

1. Based on a large number of mocks
2. Internal estimation based the data sample itself (e.g., jackknife method and bootstrap method)
3. Analytic estimation

How to construct a large number of mocks?

1. Generate approximative dark matter density fields by
   - running low mass/time resolution Nbody simulations, OR
   - constructing fields based on perturbation theory. Cheaper
2. Populate haloes using
   - halo finders, OR
   - some bias models Cheaper

EZmock (Effective Zel'dovich approximation mock; Chuang, Kitaura et al. 2015, arXiv:1409.1124)

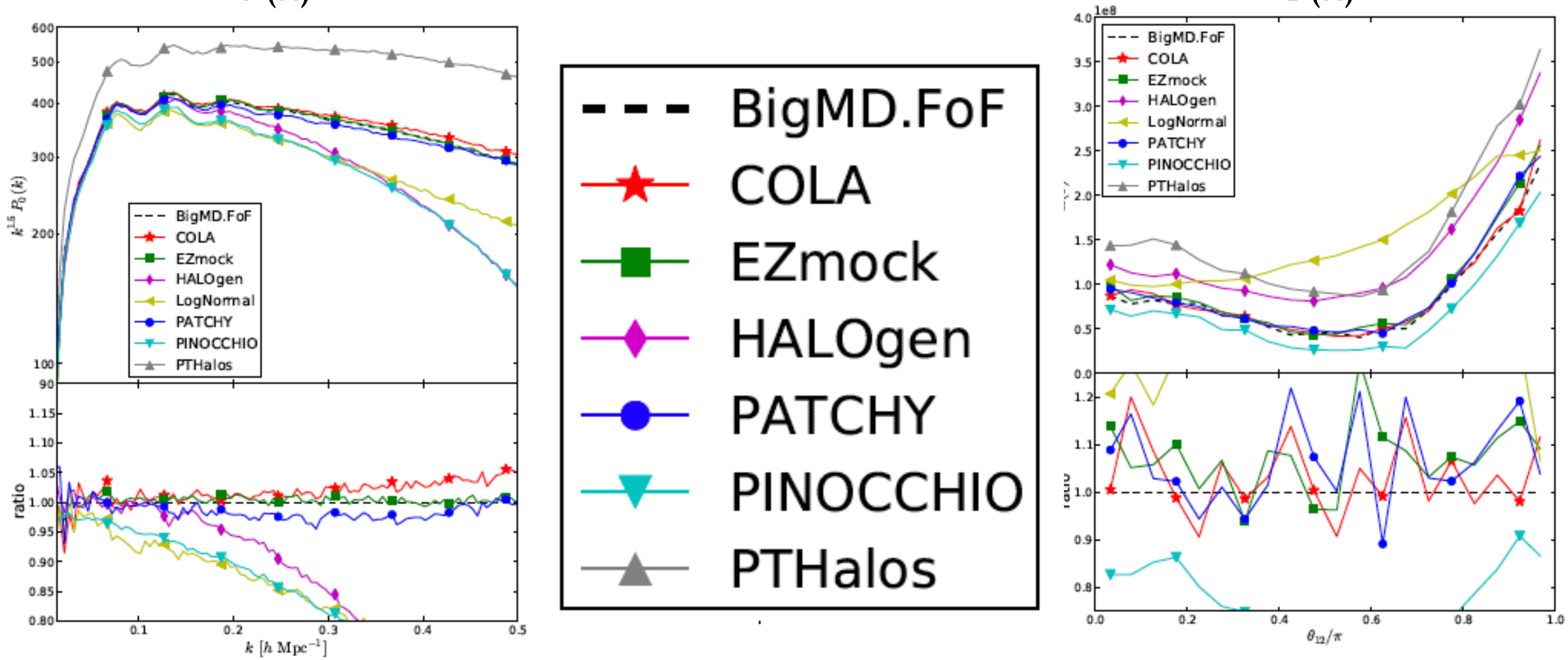
Cheapest!

- Construct the Zel'dovich approximation (ZA) dark matter density field
- Measure the parameters of the effective bias model which maps the ZA density field to a reference galaxy catalogue
- Replace the random Gaussian density field to construct thousands of mocks (> 100,000 mocks with the same computation time of one N-body simulation)

Mock catalogue comparison project

(Chuang et al. 2015, arXiv:1412.7729)

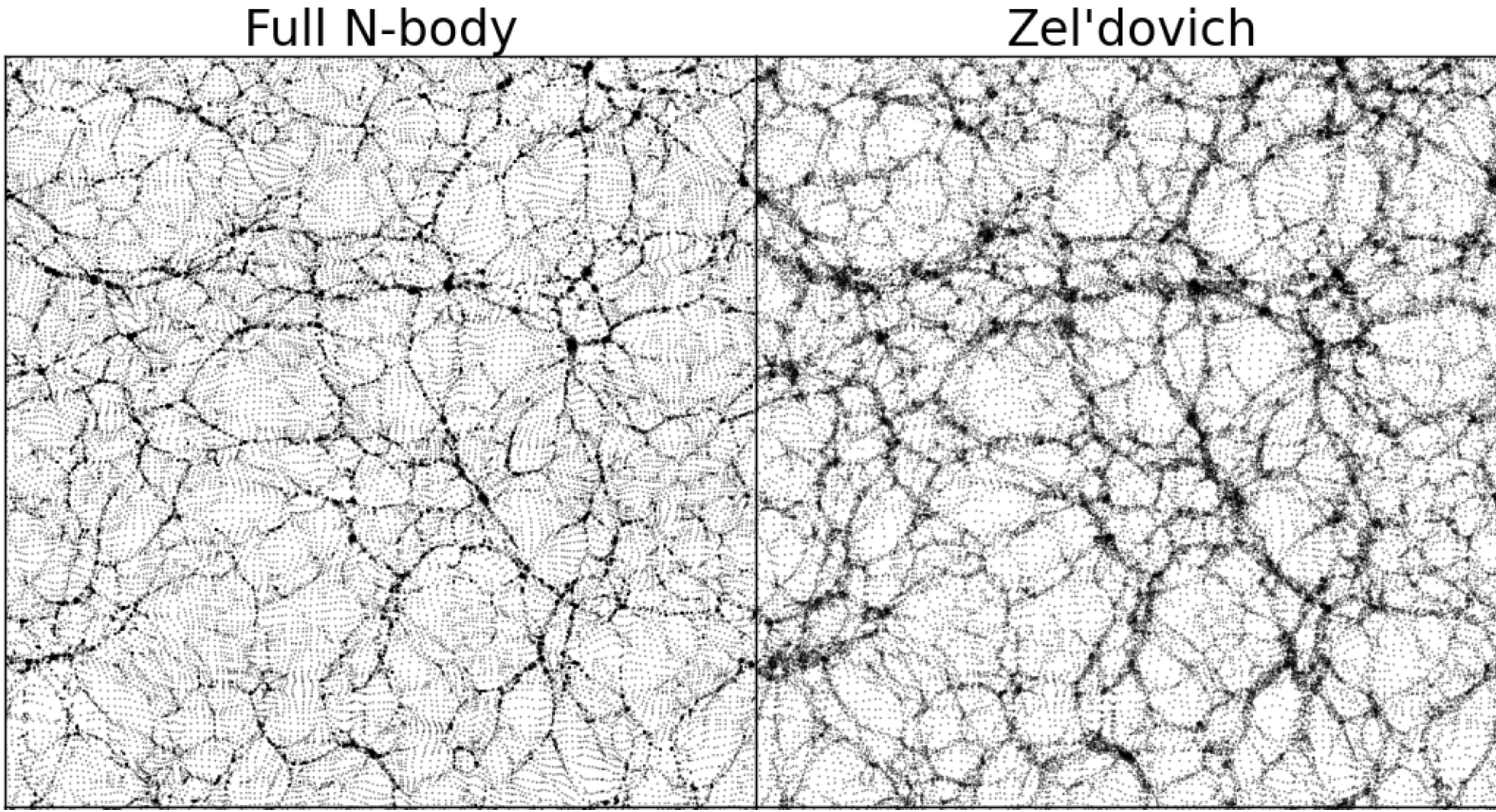
P(k) B(k)



Only COLA, EZmock, and PATCHY reach percentage accuracy at small scales for 2- and 3-point clustering statistics.

Zel'dovich approximation introduces cosmic web (higher order statistics)

Thousands of steps | **One step**



Plot taken from Neyrinck 2013

DESI covariance matrix mock challenge

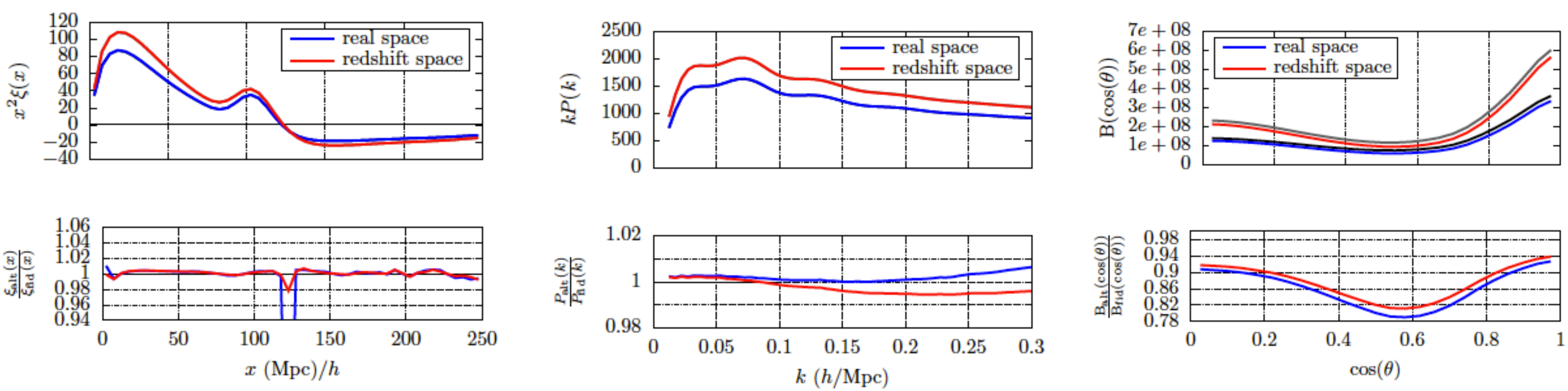
- We have started the DESI mock challenge stage 1, 2, 3: computing PK/CF for (1) box (2) cut-sky (3) lightcone DESI-like galaxy catalogs.
- We would like to invite people to construct the covariance matrices corresponding to these measurements.
- More discussion will happen at the DESI OSU meeting.

Robustness of the covariance matrix of galaxy clustering (Baumgarten & Chuang 2018)

- We test how the covariance matrix depends on the fiducial cosmology used by generating the mock catalogues.
- We test how the covariance matrix depends on different biased samples.
- To have the perfect control of the other factors, we use EZmocks (Chuang, Kitaura, et al. 2015) of which the 2-point and 3-point can be tuned to fit a reference data. Each set has 3000 EZmock boxes.

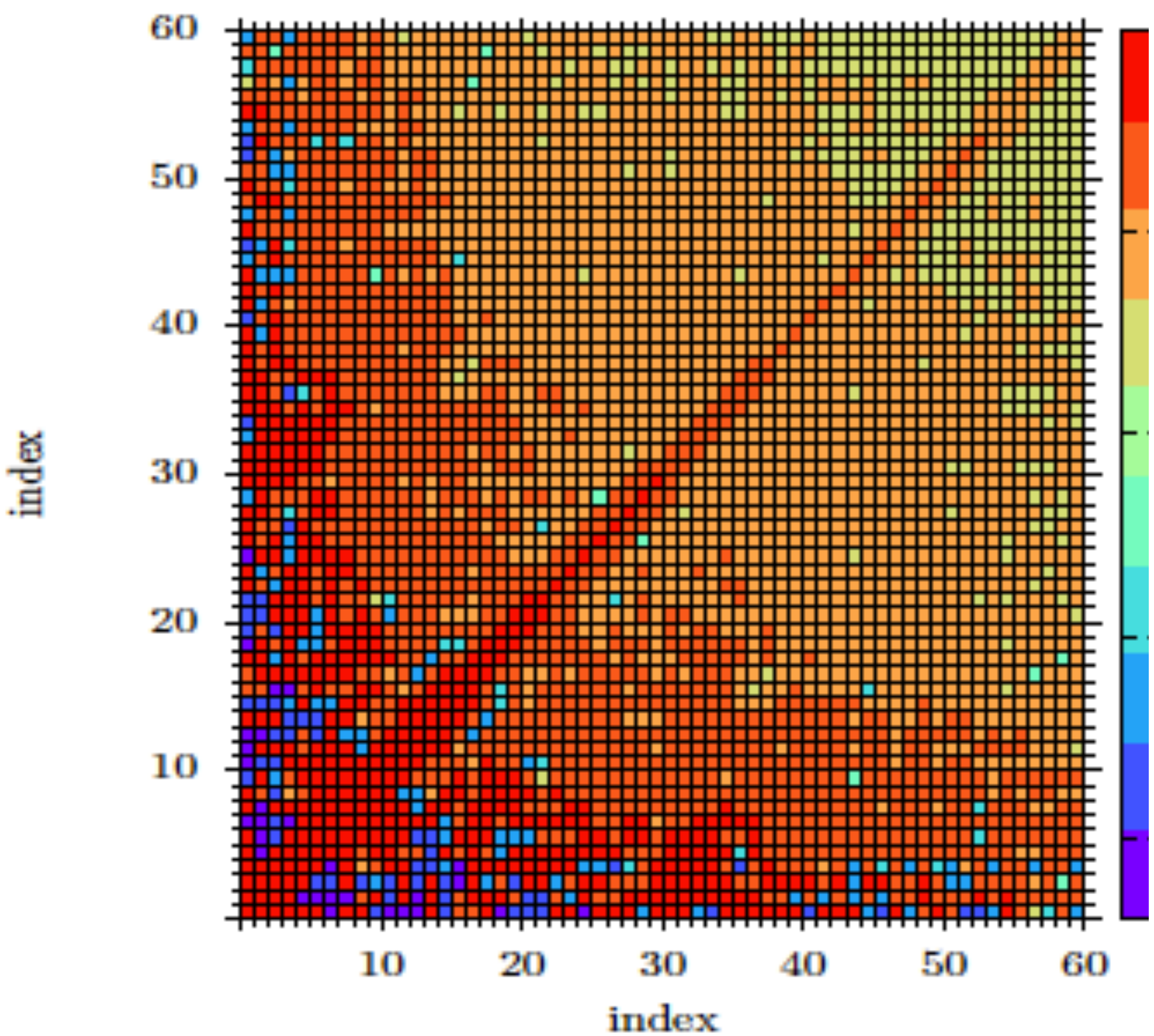
Mocks of different biased sample

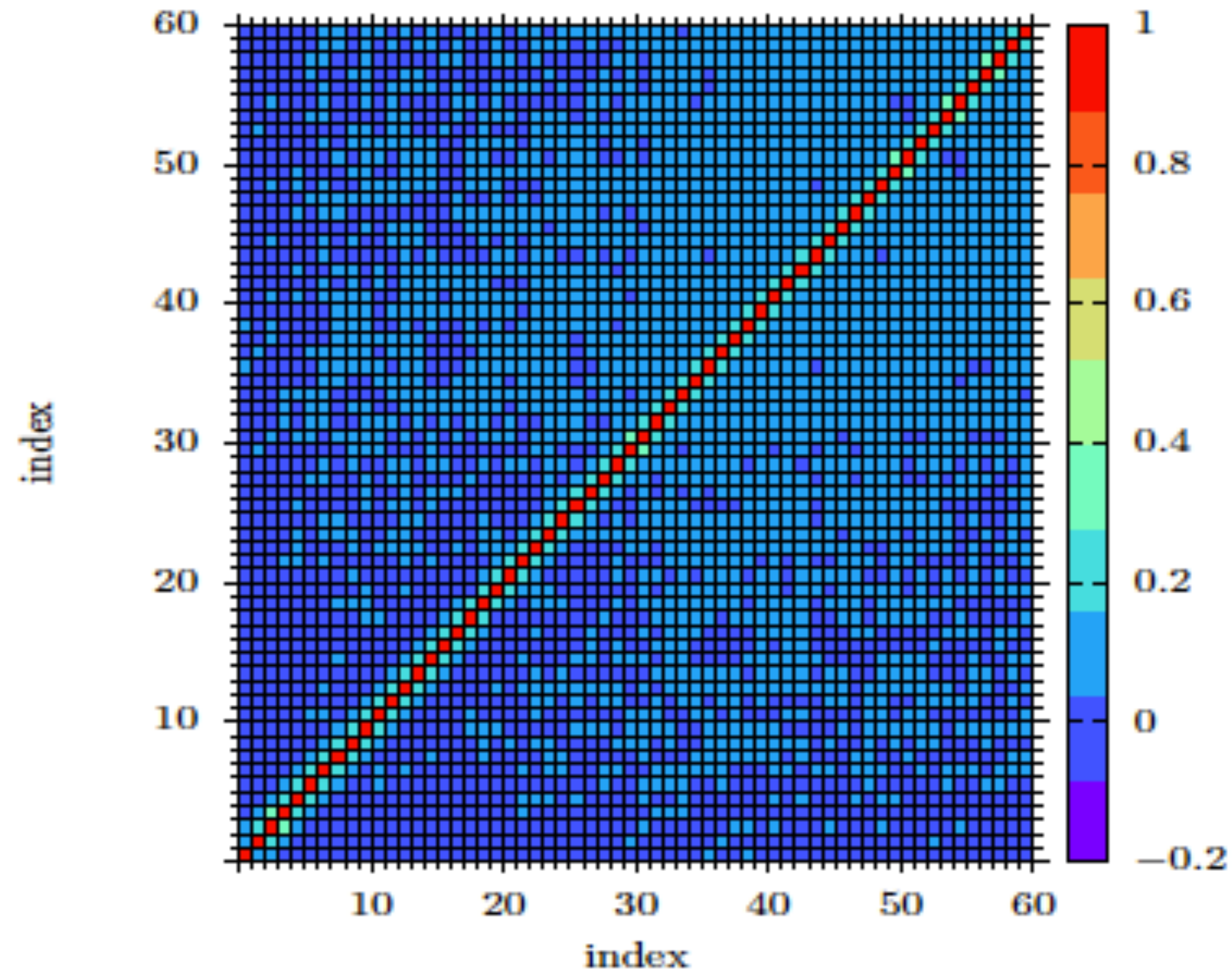
- We expect that the covariance matrix of 2-point clustering measurement is sensitive to the 2-point clustering.
- What we are interested is the impact of 3-point statistics.



Baumgarten & Chuang 2018

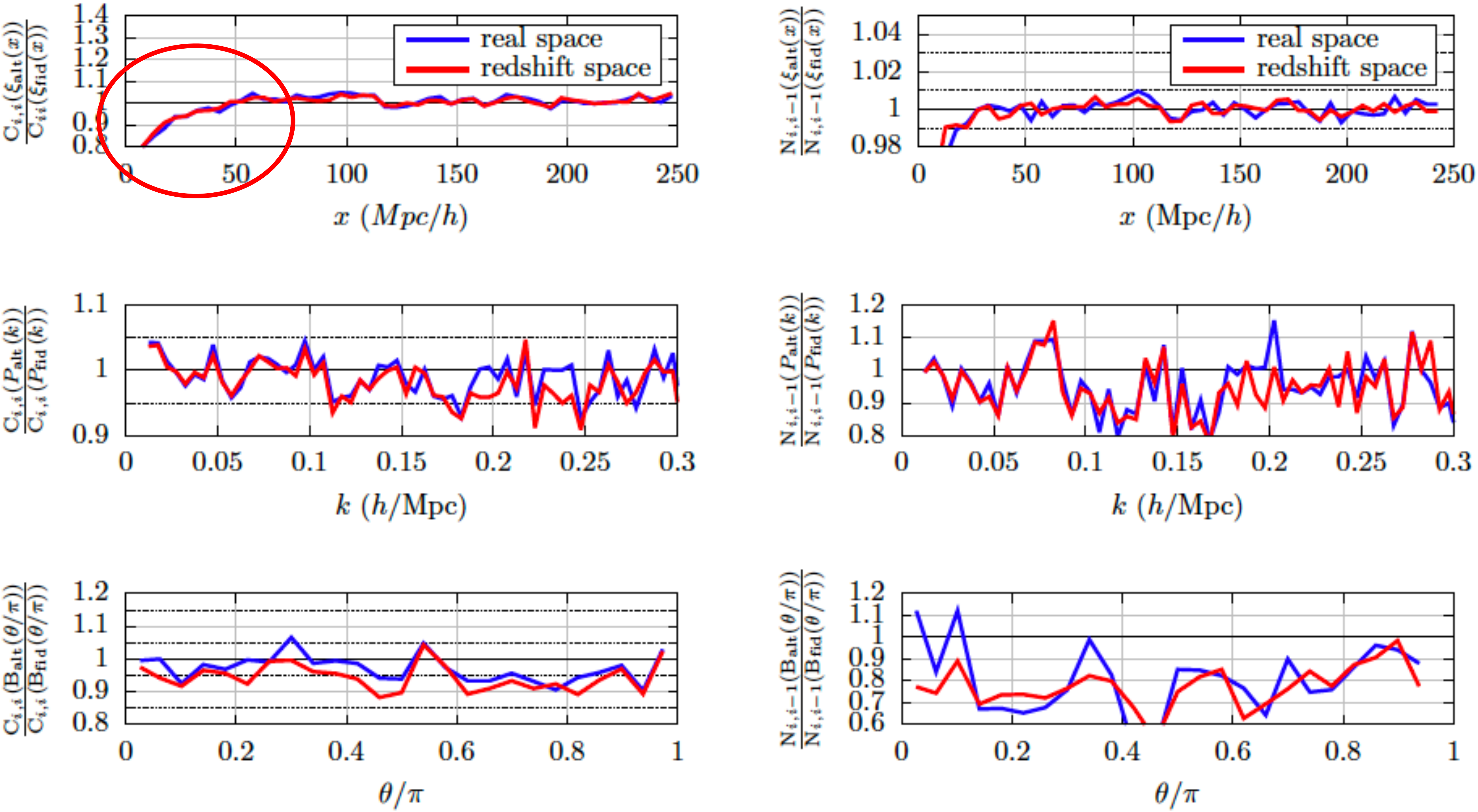
Covariance matrix & Normalized covariance matrix





Baumgarten & Chuang 2018

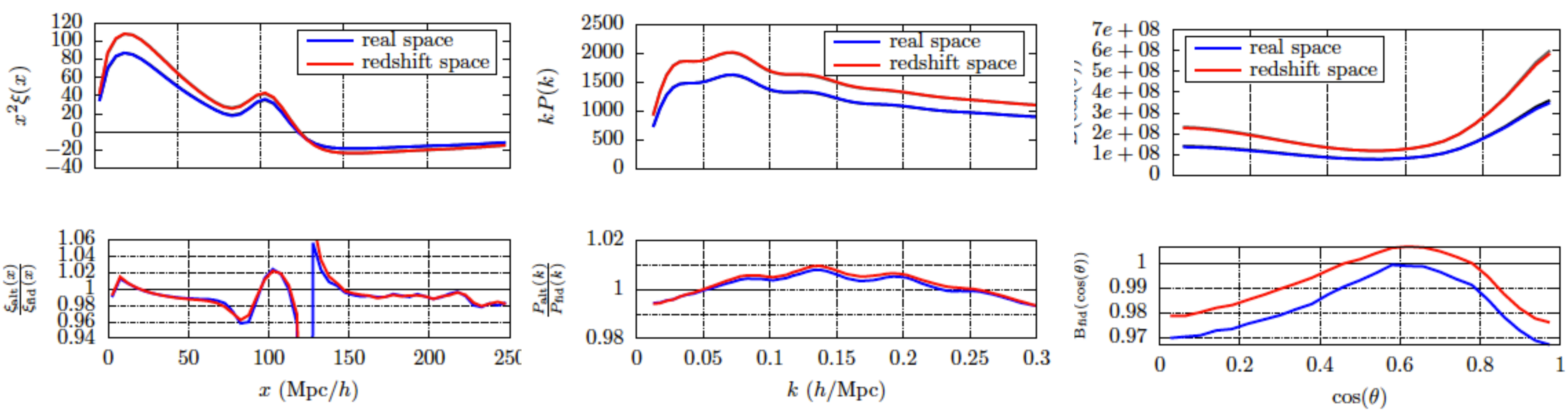
When the 3 point clustering is biased, we observe bias in the covariance matrix at scale r< 40 Mpc/h



Baumgarten & Chuang 2018

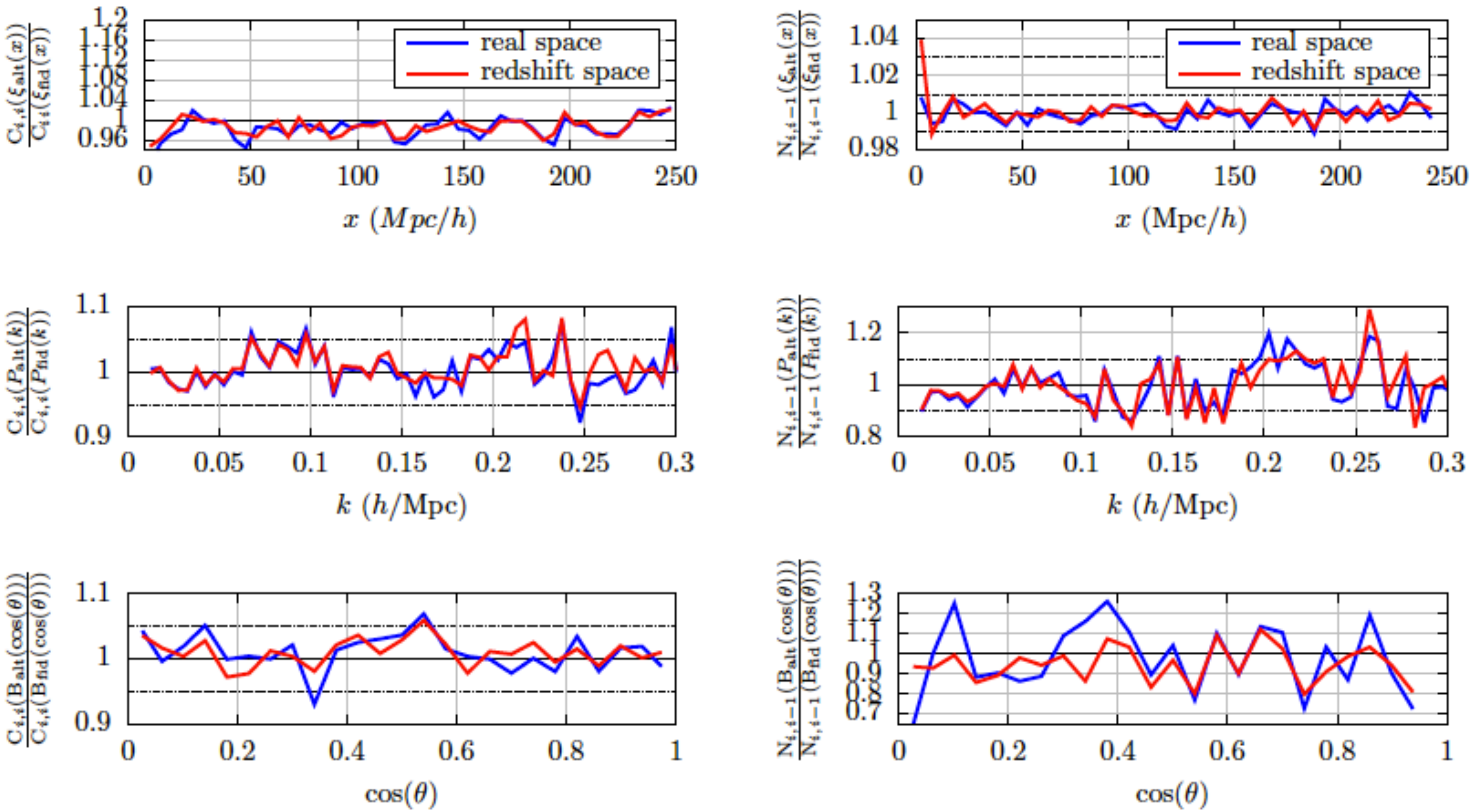
Mocks with different fiducial cosmologies

- We vary σ_8 since it has largest uncertainty based on CMB measurements.



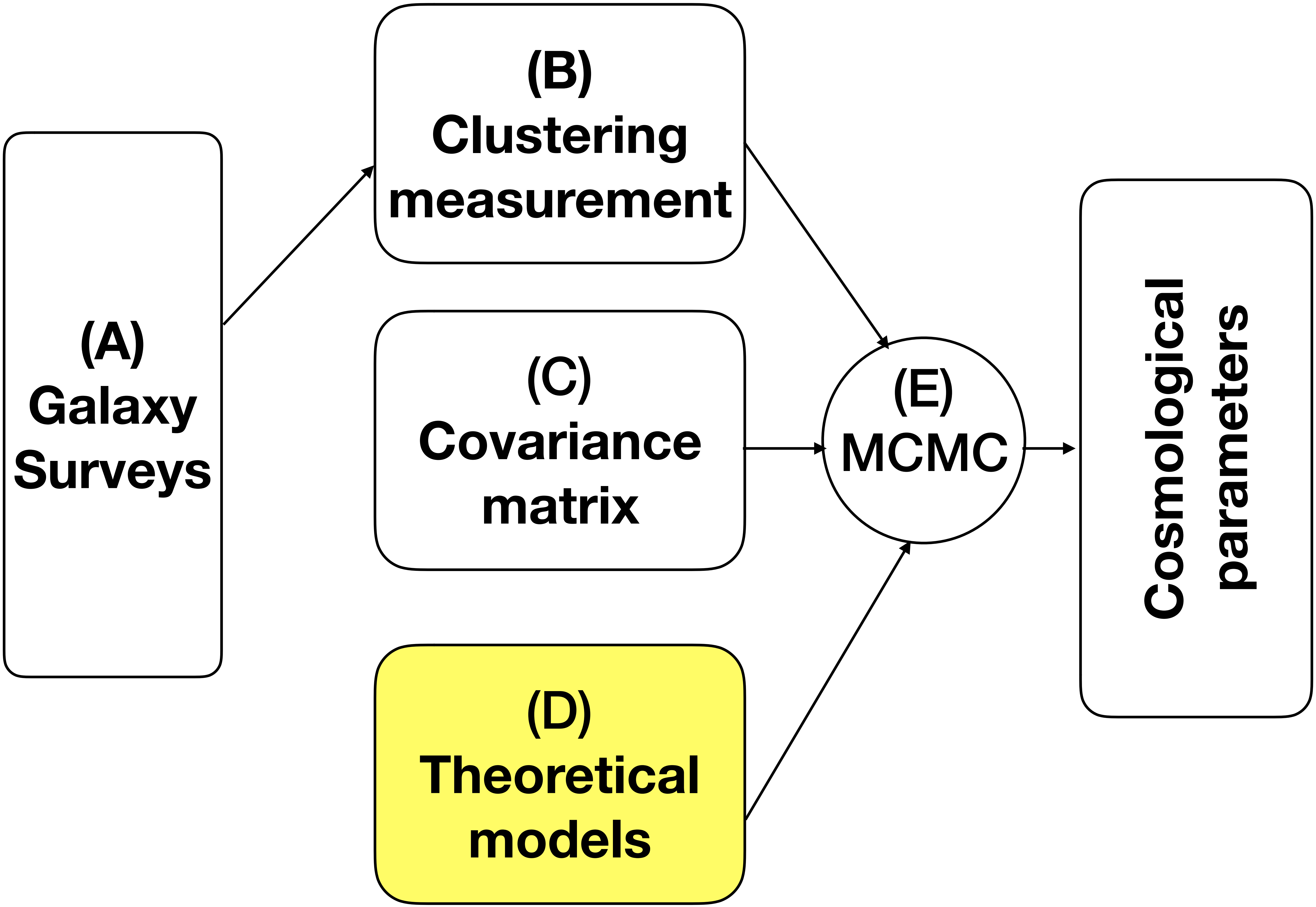
Baumgarten & Chuang 2018

No obvious bias once we calibrate the 2 and 3 point clustering statistics.



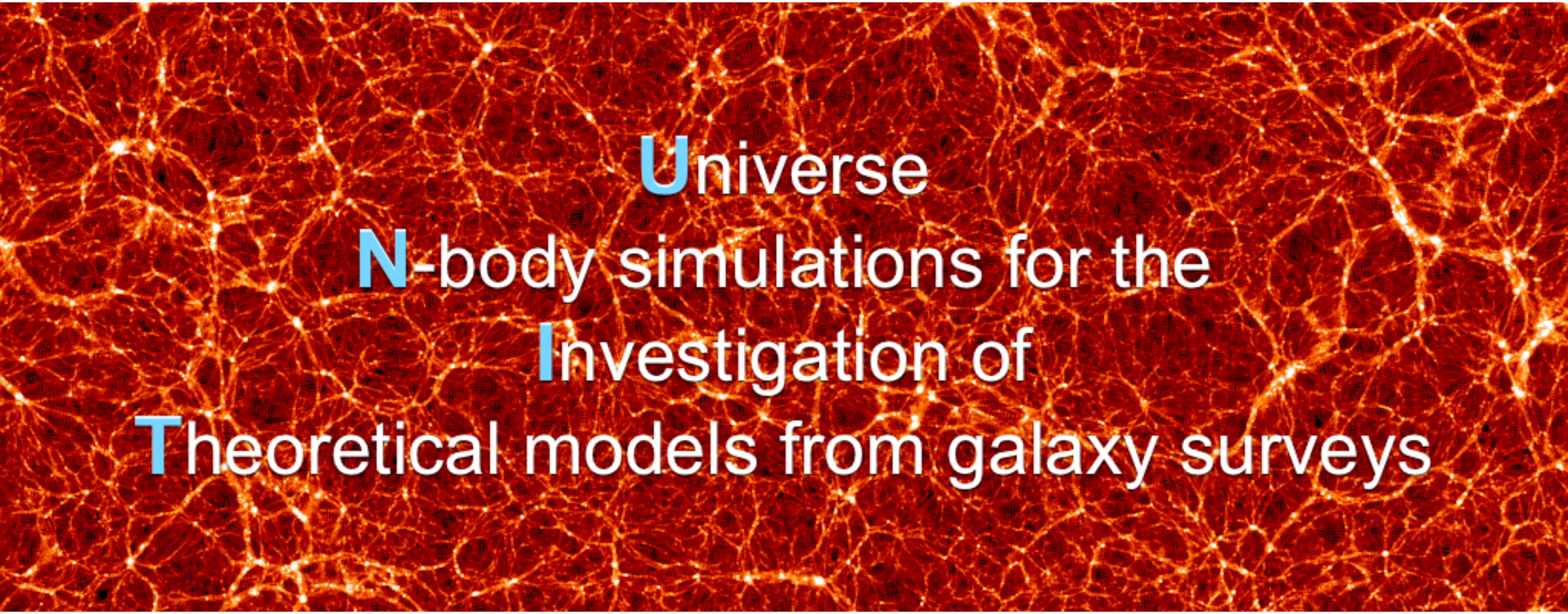
Baumgarten & Chuang 2018

- An accurate estimation of galaxy bias, or an accurate cosmological parameter set is NOT compulsory to make precision cosmological analysis from galaxy clustering, as long as the 2 and 3-point statistics are accurately fitting observations, since then systematic deviations in the galaxy bias and the cosmological parameter set compensate each other yielding unbiased covariance matrices.


(A)
Galaxy
Surveys
(B)
Clustering
measurement
(C)
Covariance
matrix
(D)
Theoretical
models
(E)
MCMC
Cosmological
parameters

UNIT *project*

(Chuang et al. 2019; arXiv:1811.02111; www.unitsims.org)



- We have delivered the simulations which could be used to validate the theoretical models of large scale structure for future surveys, such as DESI, Euclid, etc.

UNIT team

Chia-Hsun Chuang (KIPAC, Stanford/SLAC, USA)
Gustavo Yepes (UAM, Spain)
Fracisco-Shu Kitaura (IAC/ULL, Spain)
Risa Wechsler (KIPAC, Stanford/SLAC, USA)
Yu Feng (BCCP, USA)
Shadab Alam (IfA, UK)
Arka Banerjee (KIPAC, Stanford/SLAC, USA)
Robert Benton Metcalf (DIFA/INAF, Italy)
Alexander Knebe (UAM, Spain)
Carlo Giocoli (DIFA/INAF/INFN, Italy)
Marcos Pellejero-Ibanez (IAC/ULL, Spain)
Sergio Rodriguez-Torres (UAM, Spain)
Joseph DeRose (Stanford, USA)
Chun-Hao To (Stanford, USA)
Cheng Zhao (NAOC, China)

Chuang et al. 2019 (UNIT project)

simulation code	amplitude	phases	box side length	number of particles	particle M $[h^{-1}\,M_\odot]$	force resolution	number of boxes
Gadget G	fixed	regular	$1\ h^{-1}$Gpc	4096^3	1.2×10^9	$6\ h^{-1}$kpc	2
Gadget $\overline{\mathrm{G}}$	fixed	inverse-phase of G	$1\ h^{-1}$Gpc	4096^3	1.2×10^9	$6\ h^{-1}$kpc	2
FastPM A	non-fixed	regular-reference LR	$1\ h^{-1}$Gpc	1024^3	7.68×10^{10}	$1.46\ h^{-1}$Mpc	100
FastPM $\overline{\mathrm{A}}$	non-fixed	inverse-phase of A	$1\ h^{-1}$Gpc	1024^3	7.68×10^{10}	$1.46\ h^{-1}$Mpc	100
FastPM B	fixed	regular	$1\ h^{-1}$Gpc	1024^3	7.68×10^{10}	$1.46\ h^{-1}$Mpc	100
FastPM $\overline{\mathrm{B}}$	fixed	inverse-phase of B	$1\ h^{-1}$Gpc	1024^3	7.68×10^{10}	$1.46\ h^{-1}$Mpc	100
FastPM C	non-fixed	regular-reference HR	$250\ h^{-1}$Mpc	1024^3	1.2×10^9	$0.36\ h^{-1}$Mpc	100
FastPM $\overline{\mathrm{C}}$	non-fixed	inverse-phase of C	$250\ h^{-1}$Mpc	1024^3	1.2×10^9	$0.36\ h^{-1}$Mpc	100
FastPM D	fixed	regular	$250\ h^{-1}$Mpc	1024^3	1.2×10^9	$0.36\ h^{-1}$Mpc	100
FastPM $\overline{\mathrm{D}}$	fixed	inverse-phase of D	$250\ h^{-1}$Mpc	1024^3	1.2×10^9	$0.36\ h^{-1}$Mpc	100

Table 1. Overview of the set of simulations performed for this study and their corresponding parameter settings, including 800 FastPM and 2 pairs of Gadget simulations. LR and HR refer to low and high resolution, respectively.

A efficient method to get the ensemble mean of galaxy clustering from simulations

- Angule and Pontzen 2016 proposed a method to run a couple of simulations with special initial conditions.

  A. Amplitude of each k mode is fixed to the input linear power spectrum

  B. Run a paired simulation with inverse phases

- By averaging the measurements from these simulations have sample variance removed without bias introduced.

- Villaescusa-Navarro et al. (2018) included baryon and find no bias down to very small scales (k~10 h/Mpc) either.

Access **UNIT** simulations

Public available at www.unitsims.org

Cosmological parameter: $\Omega m = 0.3089$, $h = 0.6774$, $ns = 0.9667$ and $\sigma 8 = 0.8147$

Halo finder code: Rockstar

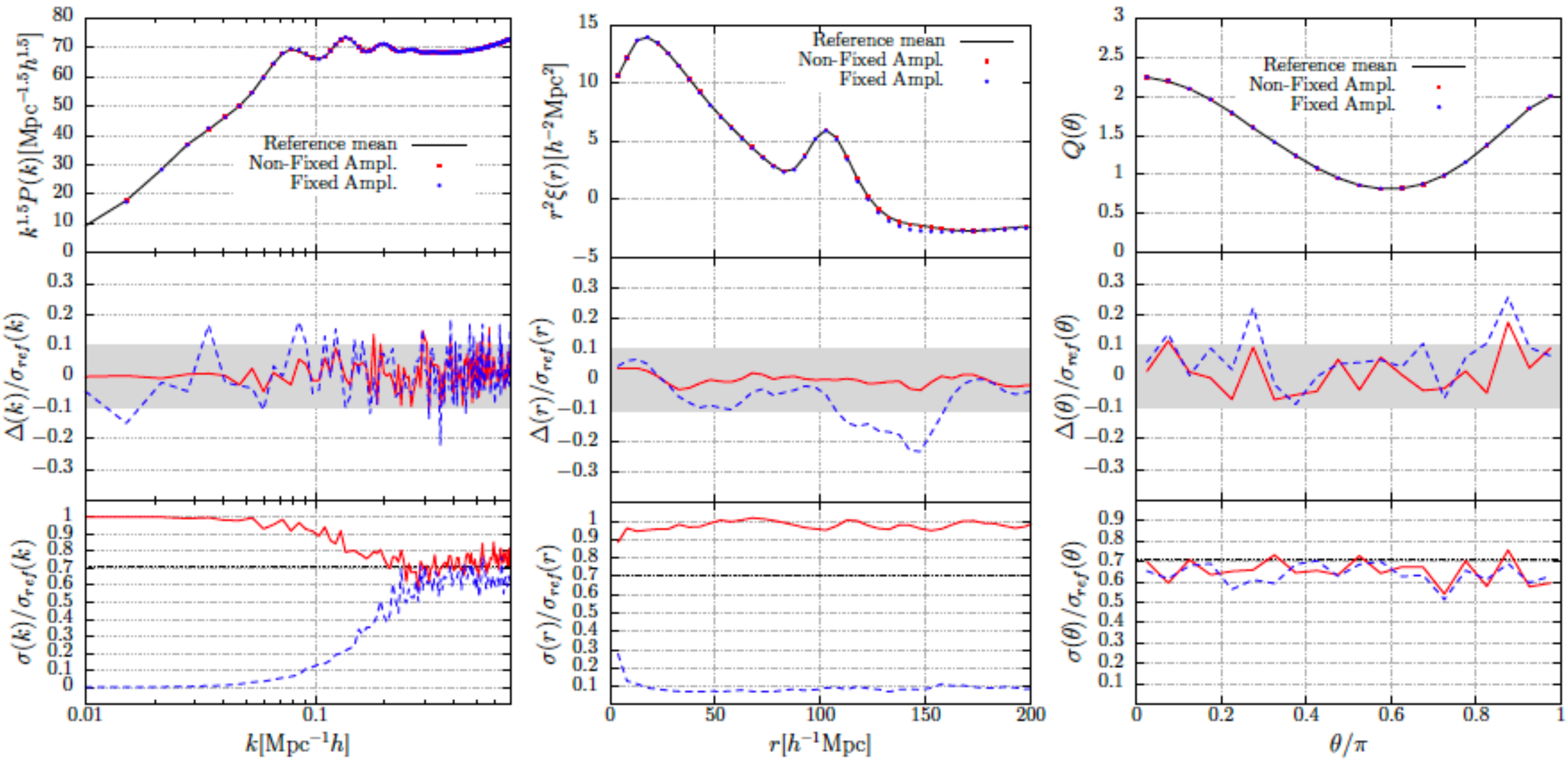
Halo catalogues available for 128 redshifts (equal ln(a))

Density fields and 0.5% particles are available for all the 128 redshifts as well.

A few snapshots are available upon request

800 FastPM halo catalogues with different ICs (fixed and non-fixed amplitude)

Dark matter PK, CF, BK

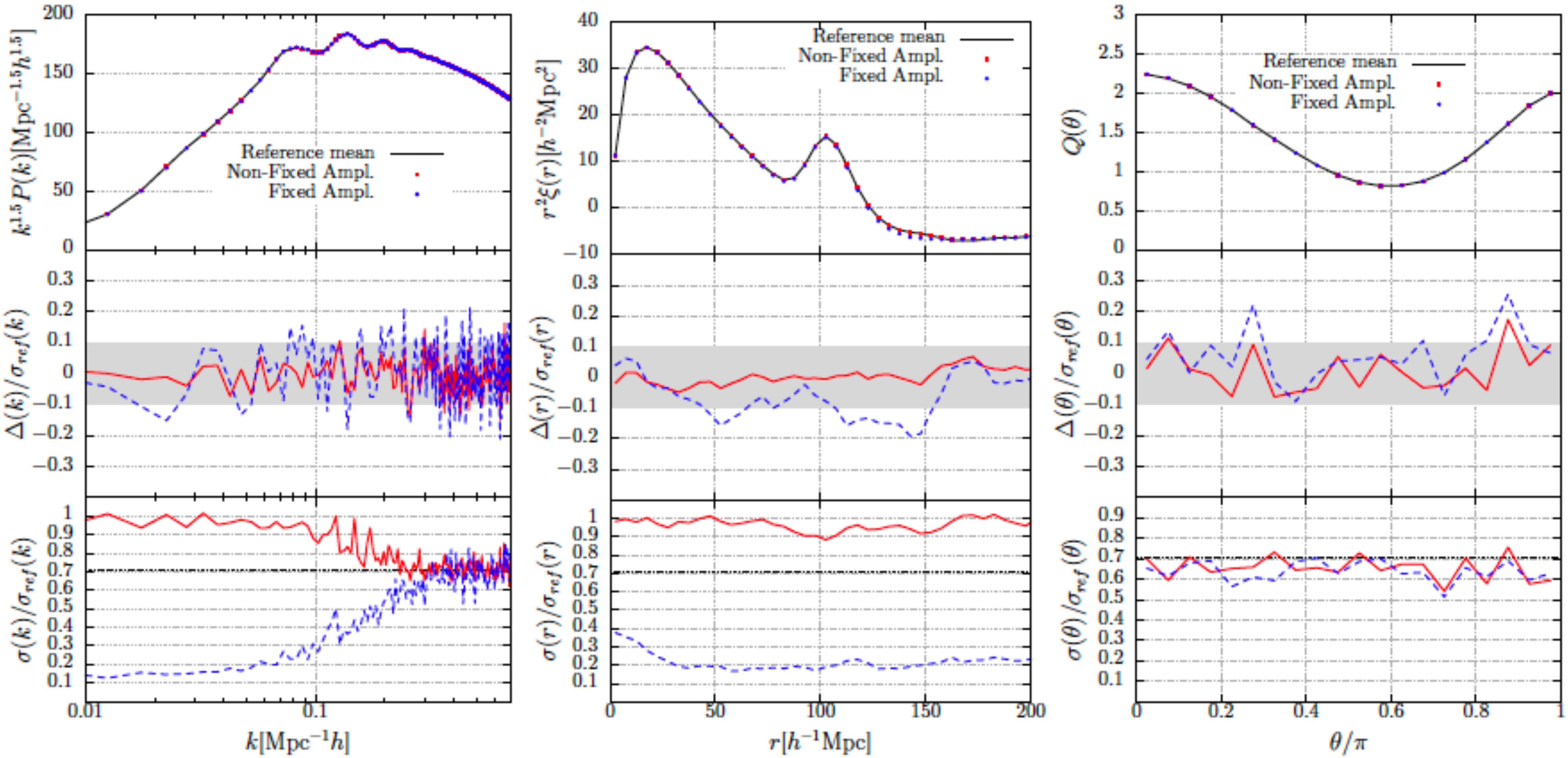


(a) PK-particles-FastPM (b) CF-particles-FastPM (c) BK-particles-FastPM

We also checked other statistics…

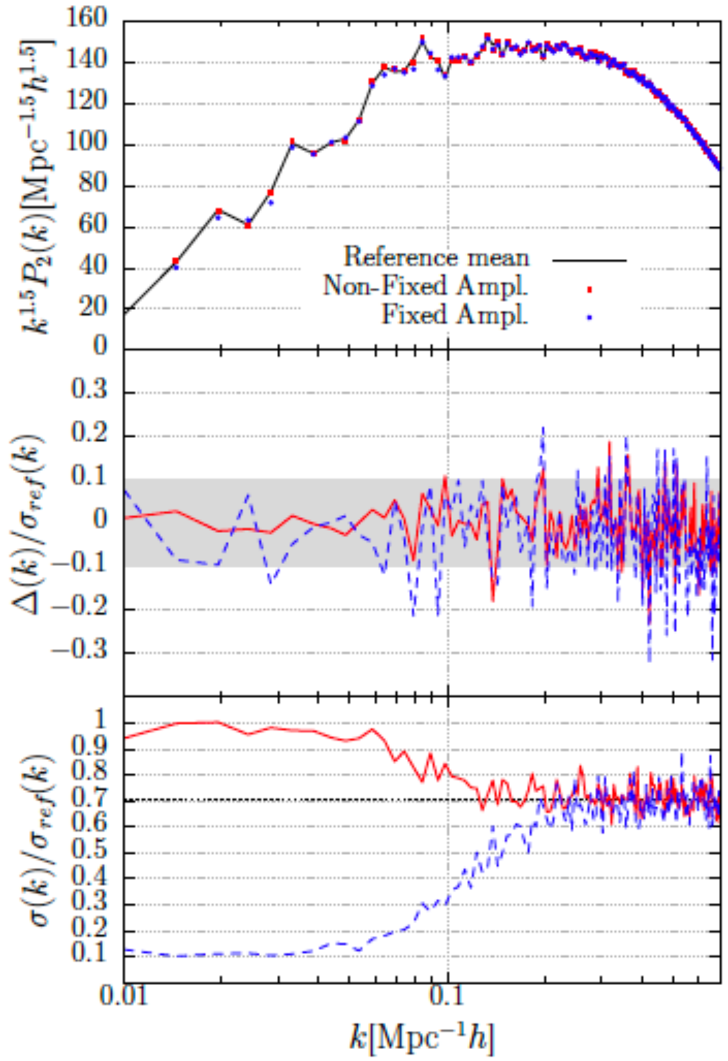
- Halo 2-point and 3-point clustering
- Halo clustering in redshift space
- Void-void clustering and void-galaxy clustering

Halo PK, CF, BK

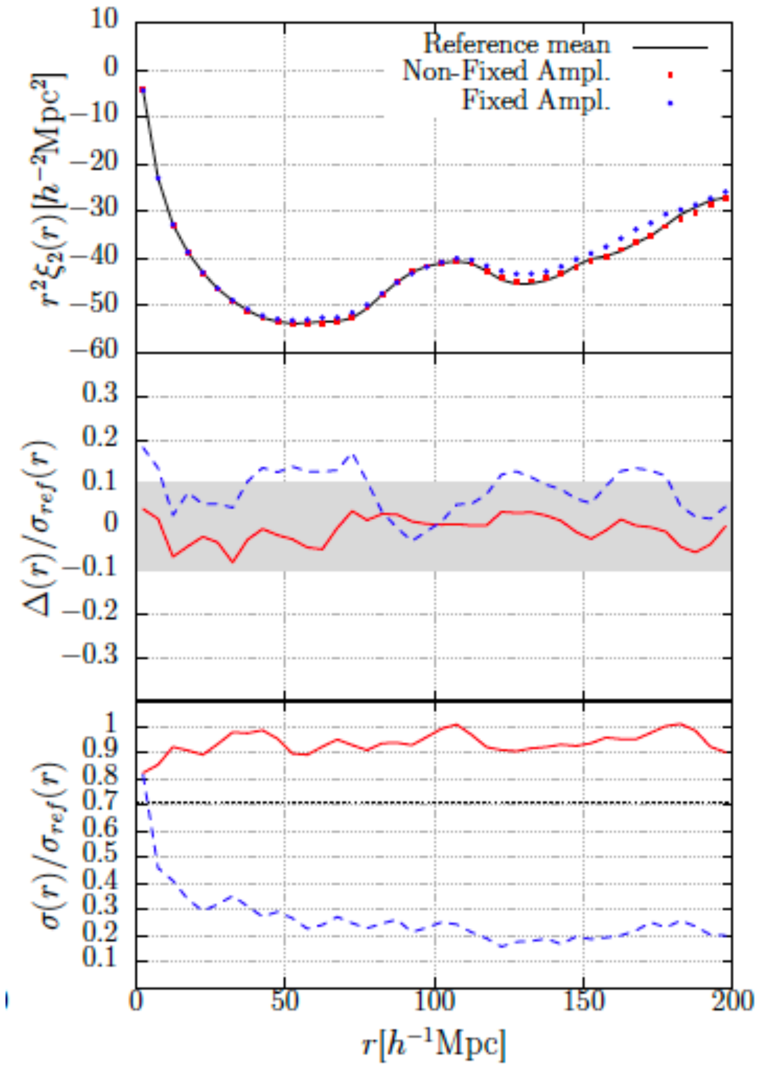


(a) PK-halos-FastPM (b) CF-halos-FastPM (c) BK-halos-FastPM

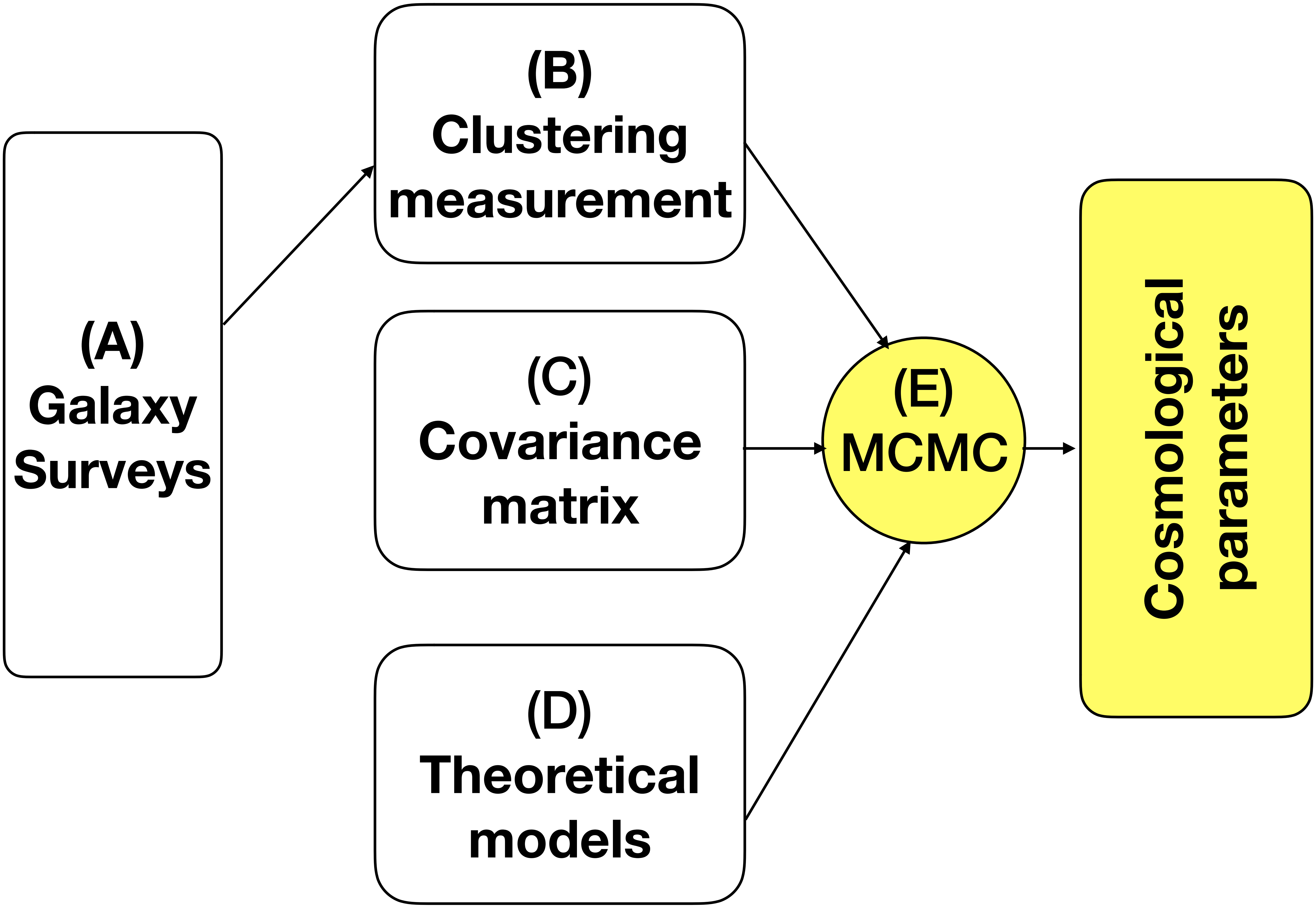
Halo quadrupoles (PK, CF)



(b) PK-quad-halos-FastPM



(d) CF-quad-halos-FastPM


(A) Galaxy Surveys
(B) Clustering measurement
(C) Covariance matrix
(D) Theoretical models
(E) MCMC
Cosmological parameters

Priors for Markov Chain Monte Carlo (MCMC) analysis

- For galaxy clustering, we have at least 8 parameters for fitting:
  - $H(z)$, $D_A(z)$, $\Omega_m h^2$, β, and $b\ \sigma_8$ are well constrained
  - $\Omega_b h^2$, n_s, and f (or bias) are NOT well constrained
- How to handle those parameters not well constrained by galaxy clustering? We need priors

Informative Priors

- **Strong priors:**

1. Fix $\Omega_b h^2$, n_s, and $\Omega_m h^2$ to the best fit values from CMB
2. Use 1σ Gaussian priors of $\Omega_b h^2$, n_s, and $\Omega_m h^2$ from CMB

Concerns raise when combining the CMB data later

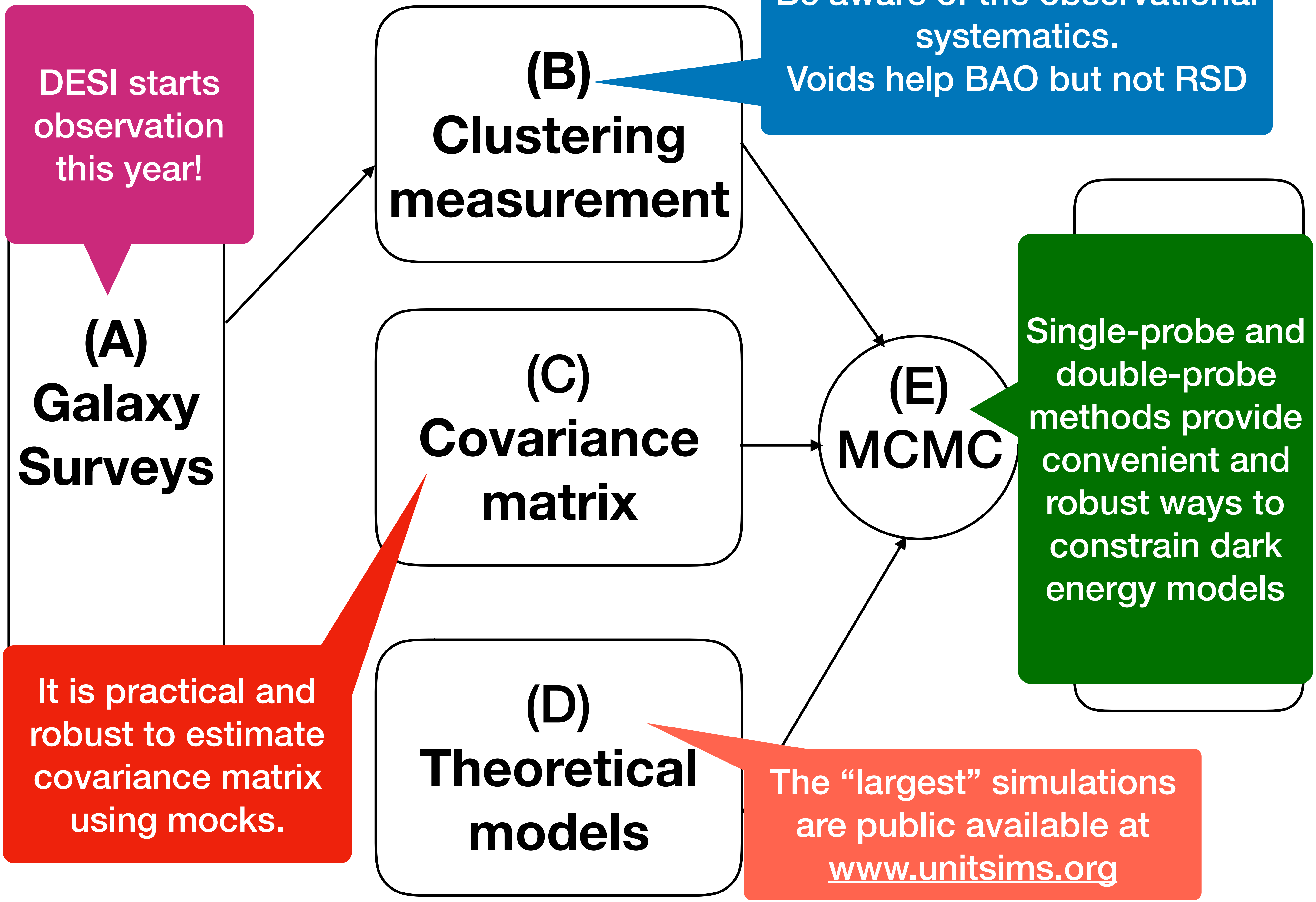
- **Weak priors:**

➢ 10 σ flat priors on $\Omega_b h^2$ and n_s measured from CMB -> Single-probe (Chuang, Pellejero-Ibanez, et al. 2017)

- **No priors:**

➢ Use joint data set of CMB and galaxy clustering -> Double-probe (Pellejero-Ibanez, Chuang, et al. 2017)

Summary


DESI starts observation this year!
(A) Galaxy Surveys
(B) Clustering measurement
Be aware of the observational systematics.
Voids help BAO but not RSD
(C) Covariance matrix
It is practical and robust to estimate covariance matrix using mocks.
(D) Theoretical models
The “largest” simulations are public available at www.unitsims.org
(E) MCMC
Single-probe and double-probe methods provide convenient and robust ways to constrain dark energy models

Summary

- Carefully handling observational systematics is very critical.
- We could gain more cosmological information by adding void clustering. However, it is dangerous to include RSD measurements from voids.
- It is practical to estimate covariance matrix with mocks for future surveys. It is critical for mocks to reproduce the 2nd and higher order clustering. It is not critical to use the wrong fiducial cosmology.
- We have prepared a set of simulations with effectively >7 times of DESI survey volume to test LSS models.
- We developed/improved the single probe and double probe methodologies and applied to the BOSS final galaxy sample. They are self-consistent and convenient ways to study dark energy models.