Cosmology with photometric redshift surveys: challenges and opportunities

David Alonso
University of Oxford

Geneva, Dec 13th 2019

Me and Geneva

2009 → CERN summer student

Me and Geneva

2009 → CERN summer student

2012 → PhD

Me and Geneva

2009 → CERN summer student

2012 → PhD

2015 → Postdoc

Me and Geneva

2009 → CERN su

2012 → PhD

2015 → Postdoc

2019 → Tenure track

Me and Geneva

2009 →

2012 →

2015 →

2019 → Tenure track

2020s → ????

From data to cosmology

Parameters

A_s n_s
r f_{NL}

Ω_{DE} Ω_{M}
w_a w_0

Initial conditions
Energy components
Background evolution

From data to cosmology

Parameters

A_s n_s
r f_{NL}

Ω_{DE} Ω_{M}
w_a w_0

Initial conditions
Energy components
Background evolution

Observables

$$\Delta(\mathbf{k}, t)$$

$$\langle |\Delta(\mathbf{k}, t)|^2 \rangle$$
$$\Downarrow$$
$$P(k, t)$$

Matter fluctuations
Power spectrum

From data to cosmology

Parameters

A_s n_s
r $f_{\rm NL}$

$\Omega_{\rm DE}$ $\Omega_{\rm M}$
w_a w_0

Initial conditions
Energy components
Background evolution

(Un)observables

$\Delta(\mathbf{k},t)$

$\langle|\Delta(\mathbf{k},t)|^2\rangle$
$\Downarrow$
$P(k,t)$

Matter fluctuations
Power spectrum

Observables

$\Delta^\alpha(\theta,\phi,\lambda)$

α=:
- CMB temperature
- CMB polarisation
- Galaxy density
- Galaxy shapes
- Lyα absorption
- 21cm flux

...

From data to cosmology

Parameters

$A_s \quad n_s$

$r \quad f_{\mathrm{NL}}$

$\Omega_{\mathrm{DE}} \quad \Omega_{\mathrm{M}}$

$w_a \quad w_0$

Initial conditions
Energy components
Background evolution

(Un)observables

$\Delta(\mathbf{k}, t)$

$\langle |\Delta(\mathbf{k}, t)|^2 \rangle$

$\Downarrow$

$P(k, t)$

Matter fluctuations
Power spectrum

Observables

$\Delta^\alpha(\theta, \phi, \lambda)$

α=:

- CMB temperature
- CMB polarisation
- Galaxy density
- Galaxy shapes
- Lyα absorption
- 21cm flux

...

Observations

$\Delta^\alpha_{\mathrm{obs}} \equiv \Delta^\alpha + N^\alpha + S^\alpha$

Object catalogues
Intensity maps
Spectra

From data to cosmology

Parameters	(Un)observables	Observables	Observations

Theoretical uncertainties

Astrophysical uncertainties

Instrumental noise
Inst. systematics
Selection effects

Parameters

$A_s \quad n_s$

$r \quad f_{\rm NL}$

$\Omega_{\rm DE} \quad \Omega_{\rm M}$

$w_a \quad w_0$

Initial conditions
Energy components
Background evolution

(Un)observables

$$\Delta(\mathbf{k}, t)$$

$$\langle|\Delta(\mathbf{k}, t)|^2\rangle$$

$$\Downarrow$$

$$P(k, t)$$

Matter fluctuations
Power spectrum

Observables

$$\Delta^\alpha(\theta, \phi, \lambda)$$

α=:
- CMB temperature
- CMB polarisation
- Galaxy density
- Galaxy shapes
- Lyα absorption
- 21cm flux

...

Observations

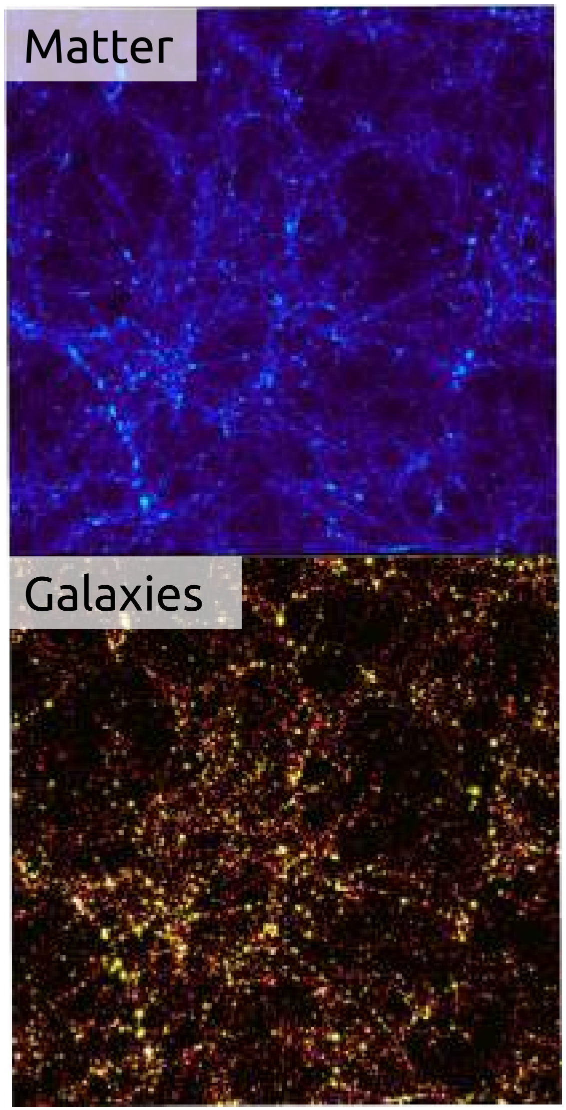
$$\Delta^\alpha_{\rm obs} \equiv \Delta^\alpha + N^\alpha + S^\alpha$$

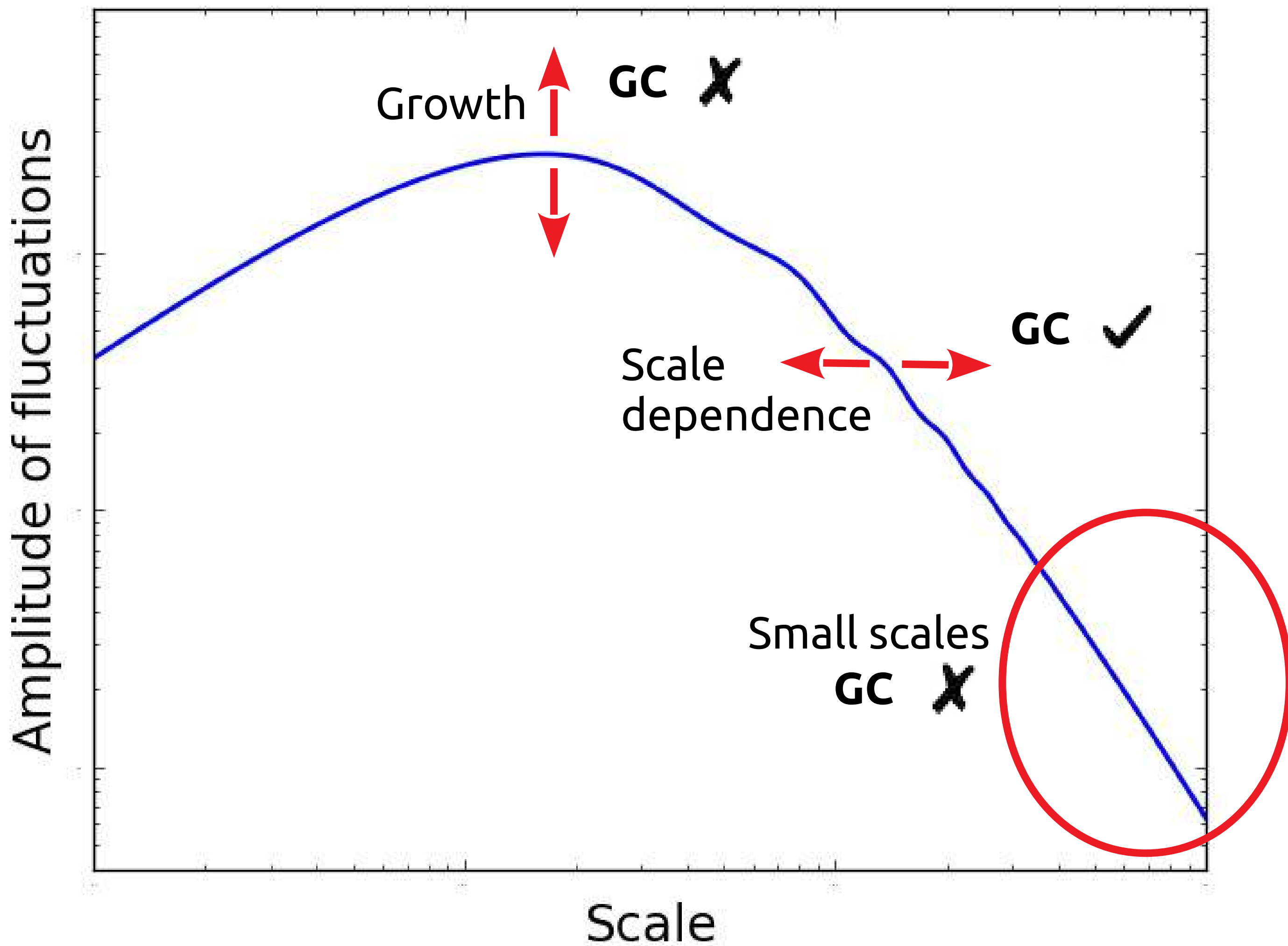
Object catalogues
Intensity maps
Spectra

Photometric surveys

Galaxy clustering:

- $\delta_g = f[\delta_M] \sim b_g \, \delta_M$
- Local
- Spin-0

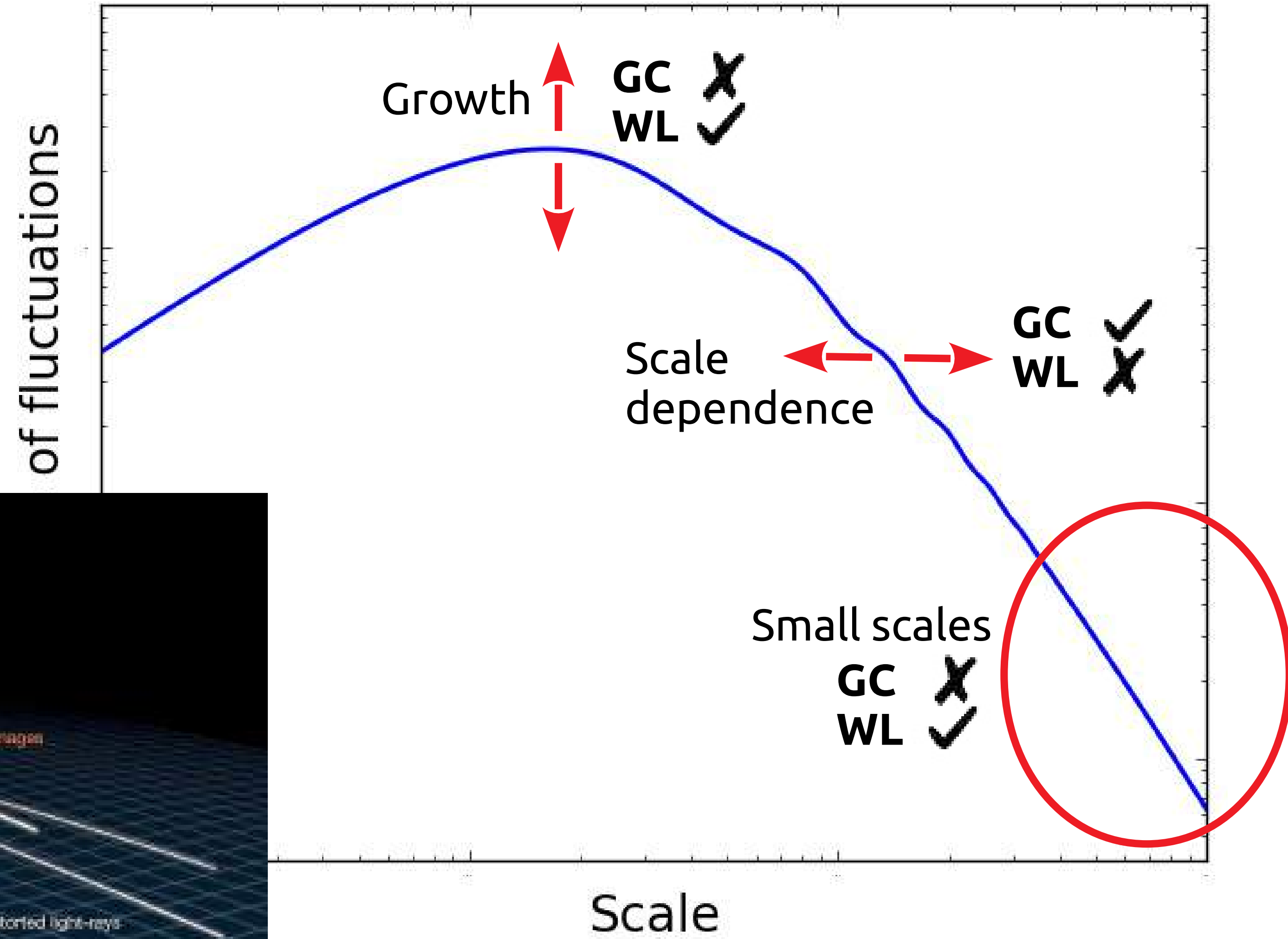


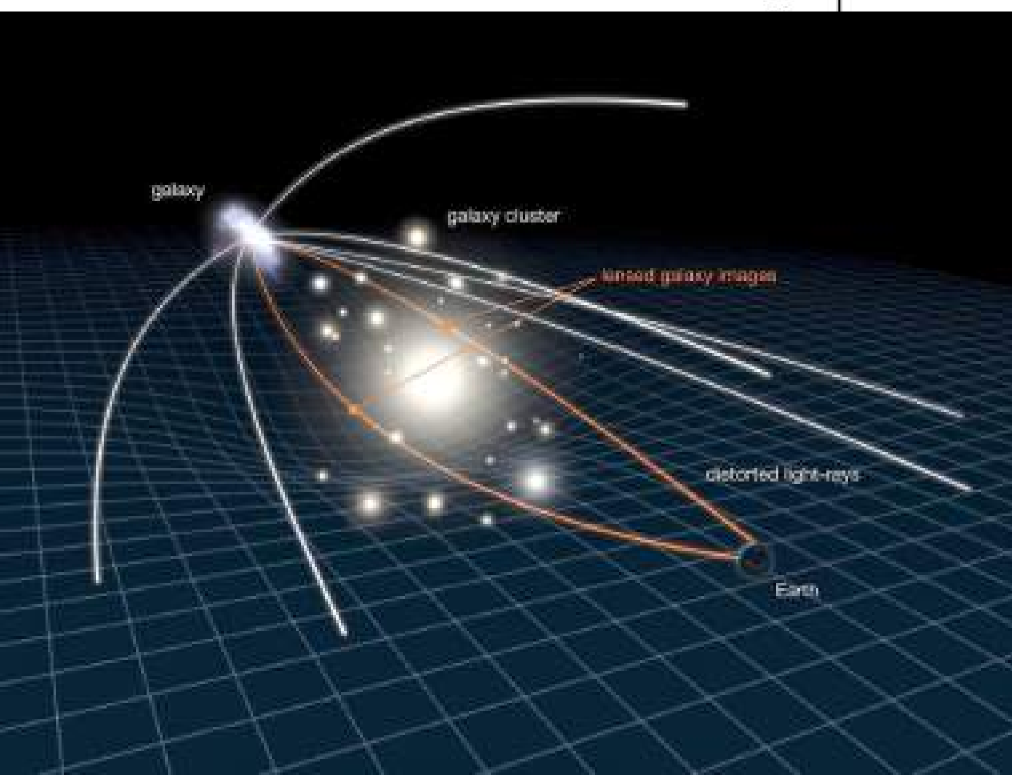


Photometric surveys

Galaxy clustering:
- $\delta_g = f[\delta_M] \sim b_g\, \delta_M$
- Local
- Spin-0

Weak lensing:
- $e_i \sim \gamma_i \sim \delta_M$
- LOS integrated
- Spin-2

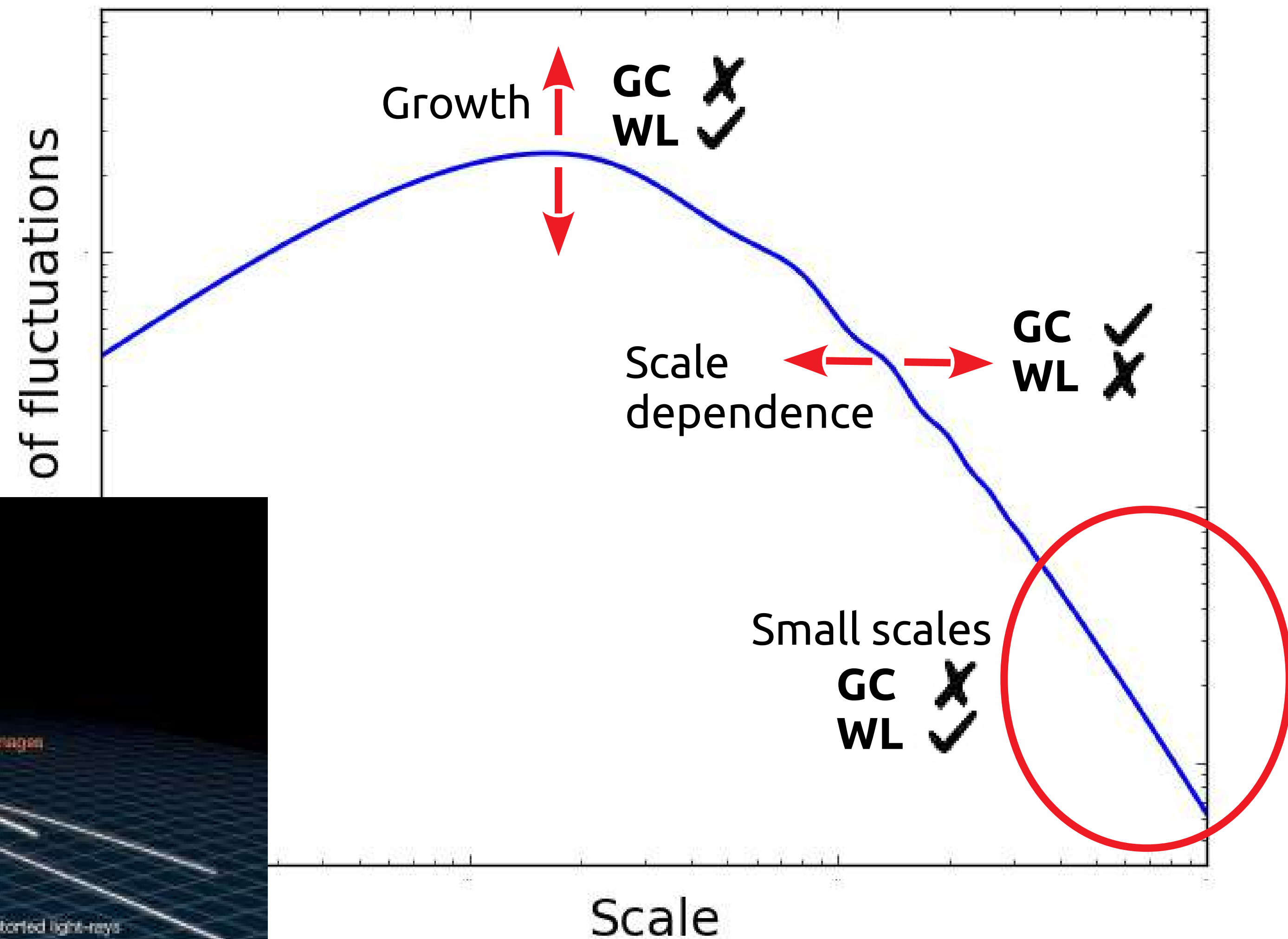


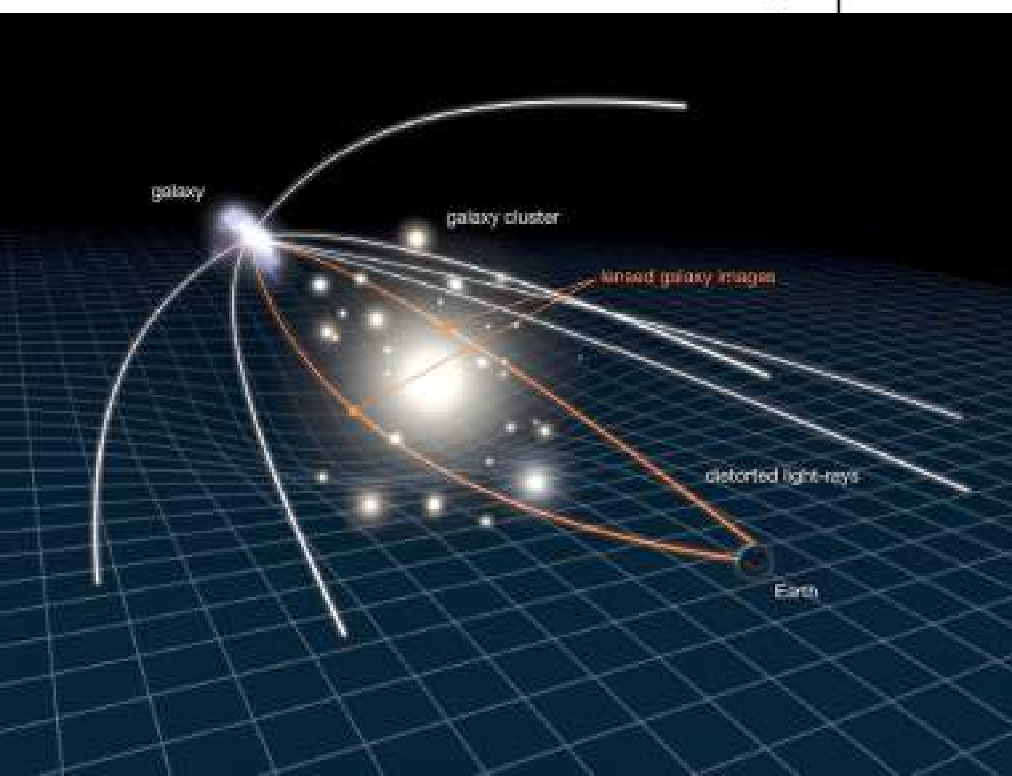


Photometric surveys

Galaxy clustering:
- $\delta_g = f[\delta_M] \sim b_g \, \delta_M$
- Local
- Spin-0

Weak lensing:
- $e_i \sim \gamma_i \sim \delta_M$
- LOS integrated
- Spin-2





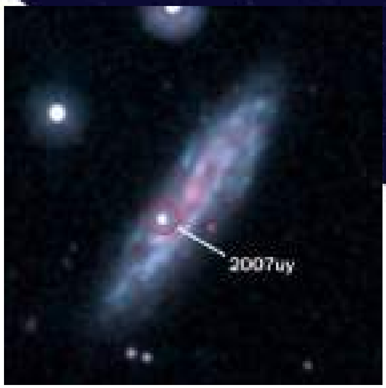
KiDS

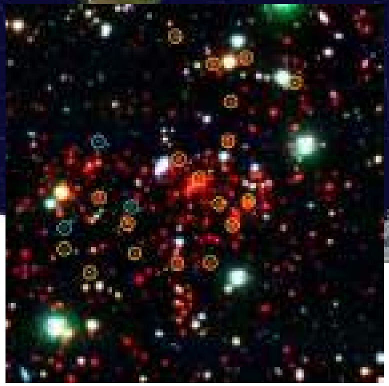
HSC

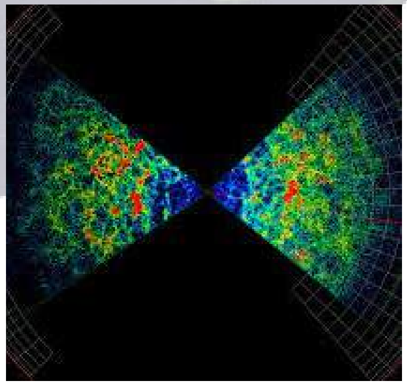
Photometric surveys: the LSST

Outstanding numbers:
- World's largest imager
  8.4 m, 9.6 sq-deg FOV
- Wide: 20K sq-deg
- Deep: r~27
- Fast: ~100 visits per year
- Big data: ~15 TB per day

Dark Energy Science Collaboration:
- Supernovae
- Cluster science
- Strong lensing
- Weak lensing
- Large-scale structure

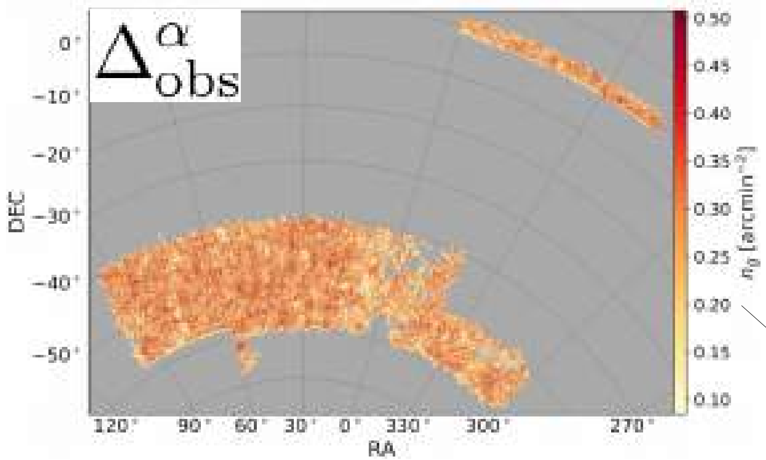






LSST Coll. et al. 0912.0201

Ideal analysis pipeline

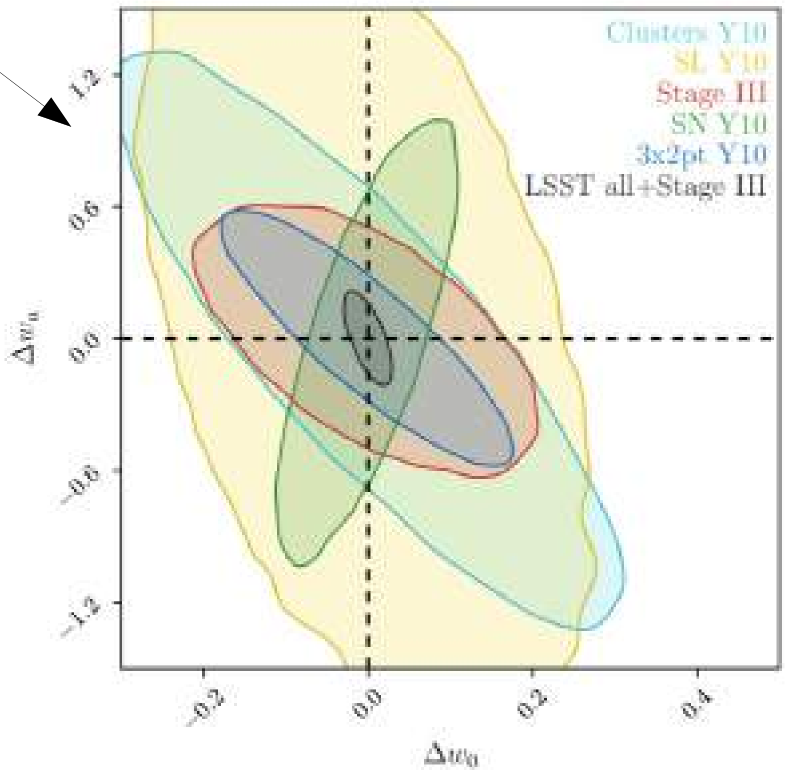


BORG:

Porqueres et al. 1812.05113
Kodi Ramanah et al. 1808.07496
Jasche & Lavaux 1806.11117
Lavaux & Jasche 1509.05040
Jasche & Wandelt 1306.1821

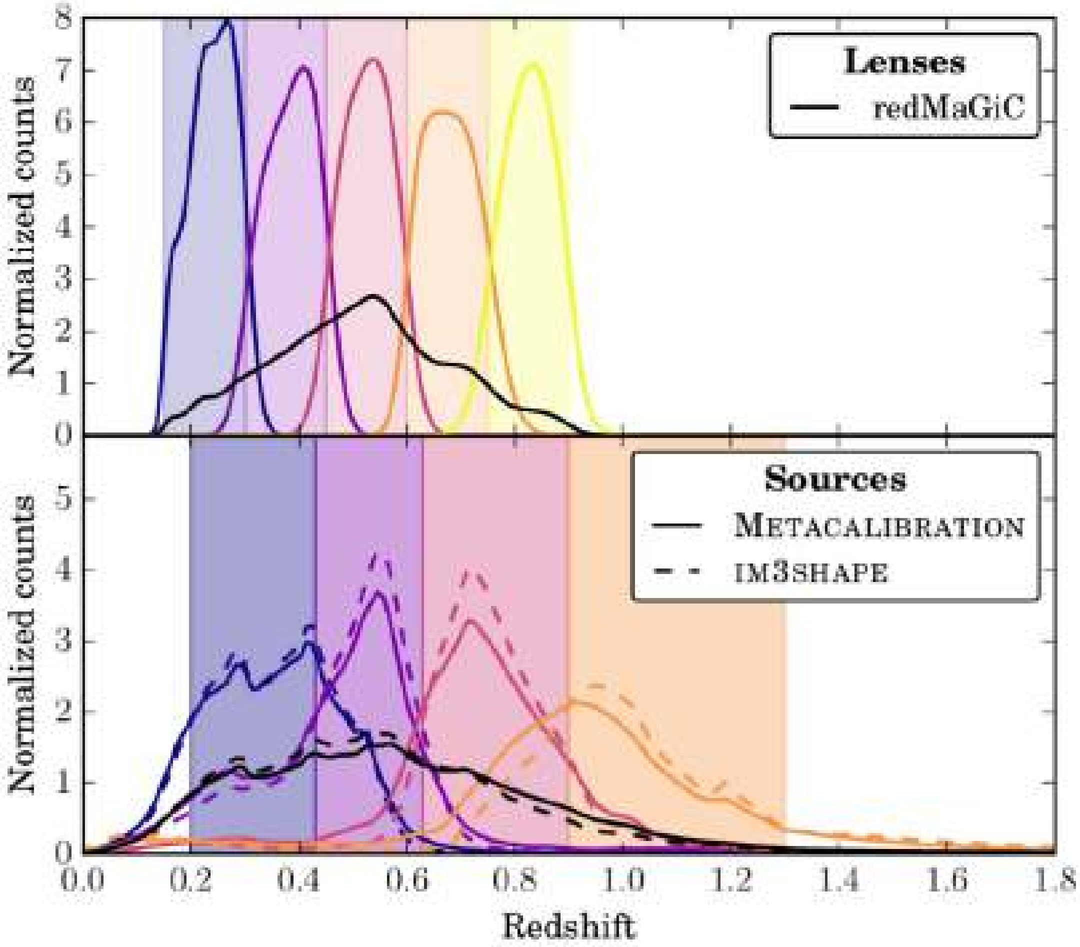
$$p(w|\Delta^{\alpha}_{\rm obs}) = \int \mathcal{D}\Delta^{\alpha}\, \mathcal{D}\Delta\, \mathcal{D}P_k\, p(w|P_k)\, p(P_k|\Delta)\, p(\Delta|\Delta^{\alpha})\, p(\Delta^{\alpha}|\Delta^{\alpha}_{\rm obs})$$

- Cosmological model
- Structure formation model
- Astrophysical model
- Instrument/noise model



2-point tomographic analysis

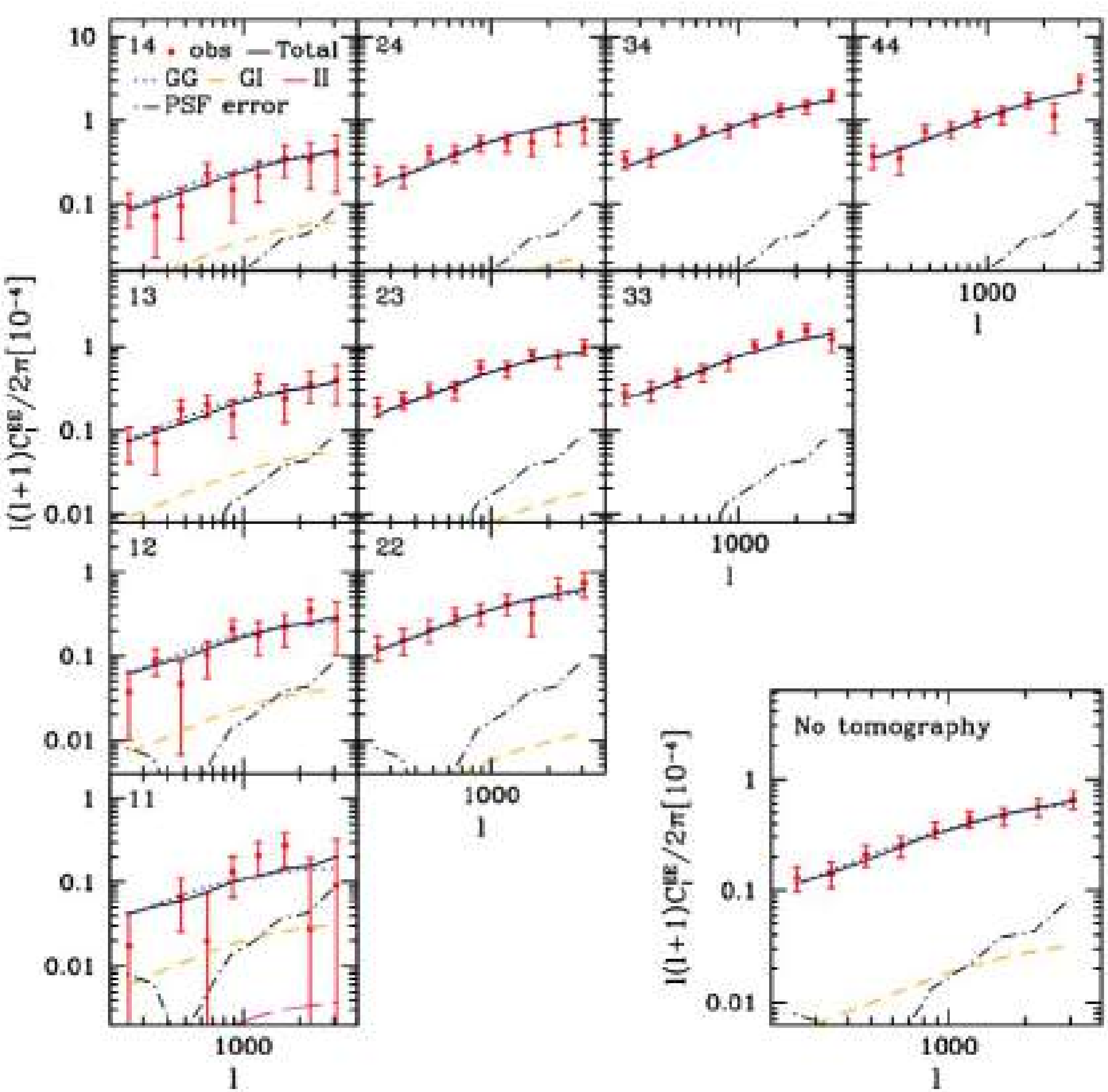
- Photo-zs are complicated.
- Bunch galaxies up into photo-z bins and project onto the sphere.



DES Y1 data
arXiv:1708.01530

2-point tomographic analysis

- Photo-zs are complicated.
- Bunch galaxies up into photo-z bins and project onto the sphere.
- Compute all possible two-point cross-correlations (different bins, different observables).

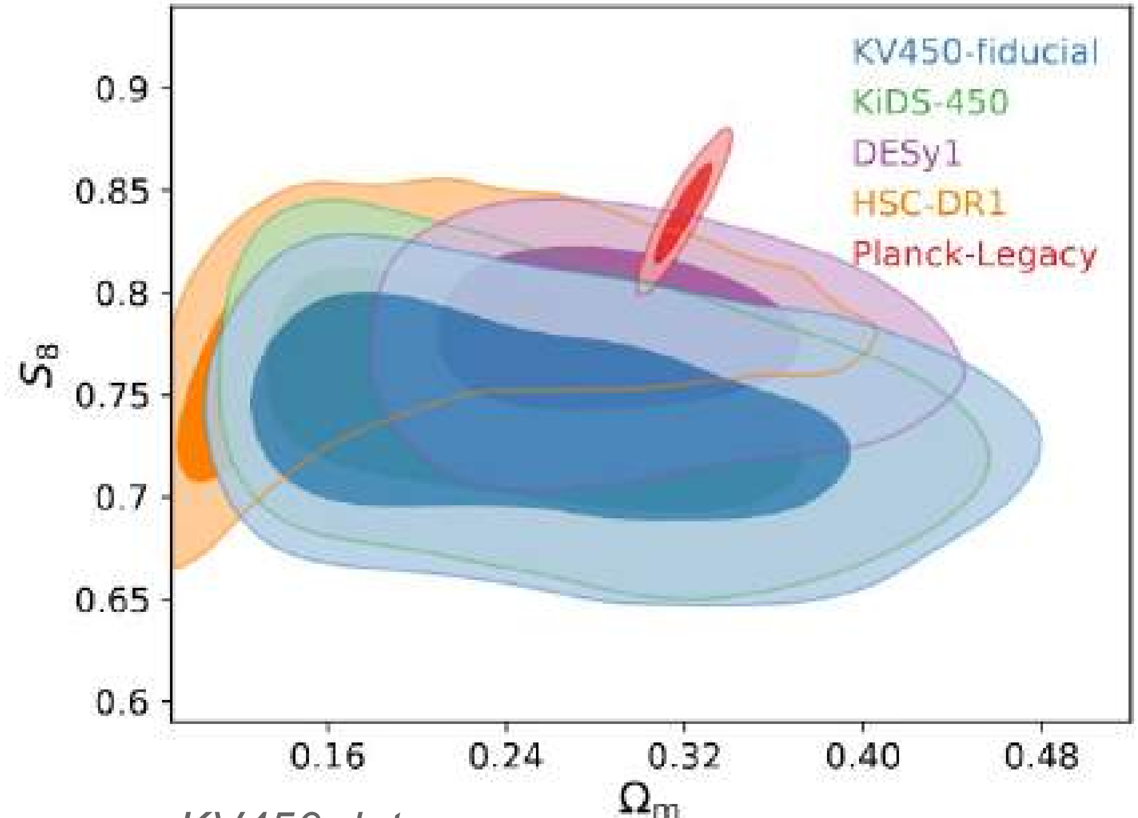


HSC Y1 data
arXiv:1809.09148

2-point tomographic analysis

$$-2 \log P(d|\theta) = (d-t(\theta))^T C^{-1} (d-t(\theta)) + L_0$$

- Photo-zs are complicated.
- Bunch galaxies up into photo-z bins and project onto the sphere.
- Compute all possible two-point cross-correlations (different bins, different observables).
- Constrain parameters using a Gaussian likelihood.



KV450 data
arXiv:1812.06076

2-point tomographic analysis

Covariance matrix

Gaussian likelihood

$$-2 \log P(d|\theta) = (d - t(\theta))^T C^{-1} (d - t(\theta)) + L_0$$

Vector of cross-correlations

Theory prediction

Computing two-point functions

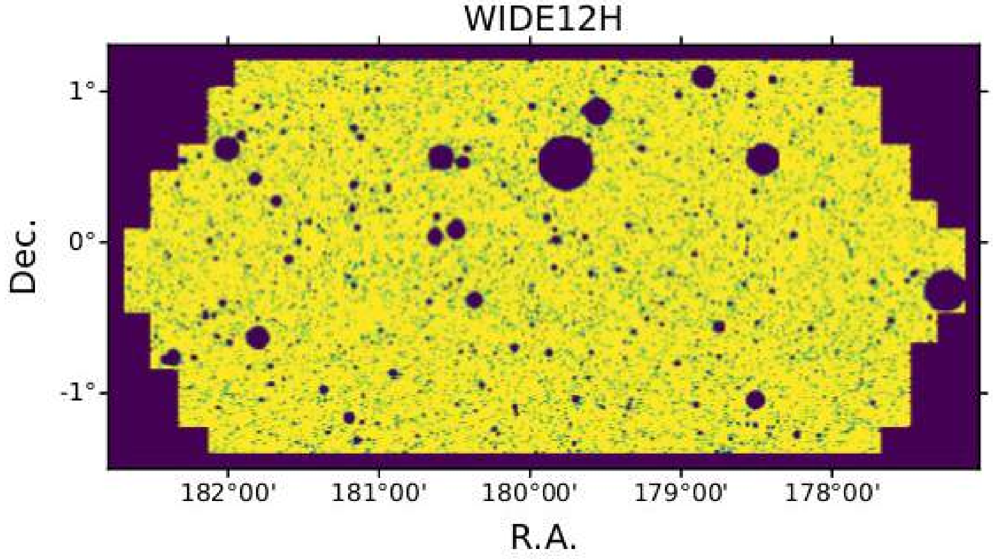
Covariance matrix

Gaussian likelihood

$$-2 \log P(d|\theta) = (d - t(\theta))^T C^{-1} (d - t(\theta)) + L_0$$

Vector of cross-correlations

Theory prediction



A unified pseudo-C_ℓ estimator

DA, F.J. Sanchez, A. Slosar

arXiv:1809.09603

PCL facts

- **Why C_ℓ?** (as opposed to $\xi(\theta)$)
  - k-cuts are easy to interpret. No Hankel transform
  - Covariance is a lot more diagonal
  - Good computational scaling ($\sim N^{3/2}$)

- **PCL vs. QMV**
  - PCL == QMV when the covariance matrix is diagonal
  - PCL is precise enough in many common scenarios
  - QMV ~ N^3, PCL ~ $N^{3/2}$
    (The trick is being able to estimate mode coupling analytically)

Mode-coupling

Inverse-variance-weight data

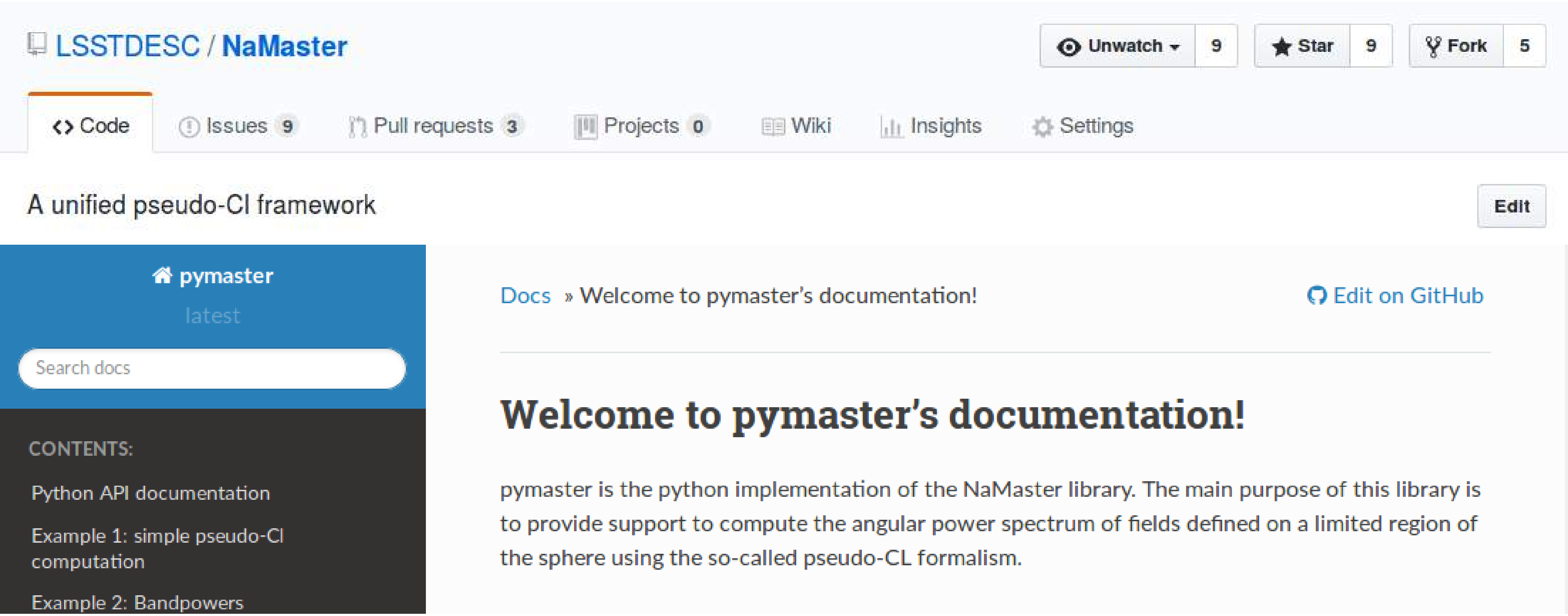
$$F_{ij} B_j = \frac{1}{2} \mathbf{a}^\dagger \tilde{C}^{-1} P_i \tilde{C}^{-1} \mathbf{a} - \frac{1}{2} \mathrm{Tr}\left[\tilde{C}^{-1} P_i \tilde{C}^{-1} N\right]$$

Fourier-transform and square

Noise bias

Tegmark astro-ph/9611174
Efstathiou astro-ph/0307515
Leistedt et al. arXiv:1306.0005

A unified pseudo-C_ℓ code



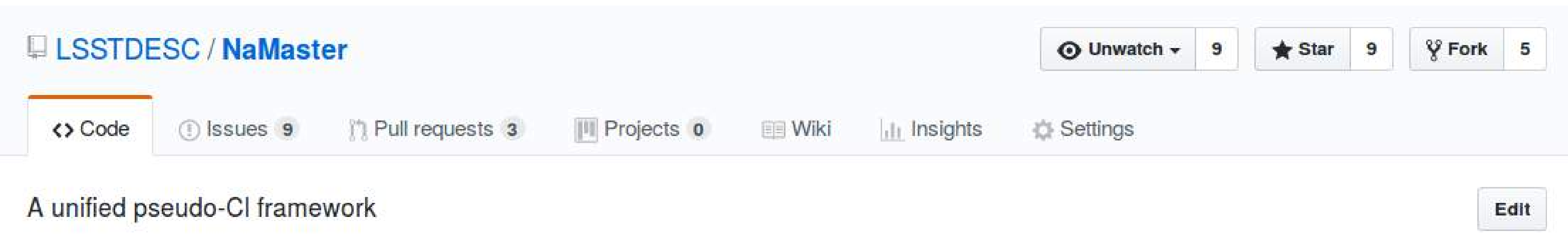
Code: https://github.com/LSSTDESC/NaMaster
Docs: https://namaster.readthedocs.io/en/latest/index.html

A unified pseudo-C_ℓ code

http://www2.iap.fr/users/hivon/software/PolSpice/
https://gitlab.in2p3.fr/tristram/Xpol

What **features** does it implement?

- Calculate PCL power spectra (including coupling matrix, etc.)
- In curved and flat skies
- Spin-0 (density, CMB T) and spin-2 (shear, CMB Q/U) quantities
- Bells and whistles:
  - Mode deprojection
  - E/B mode purification
- Gaussian covariances



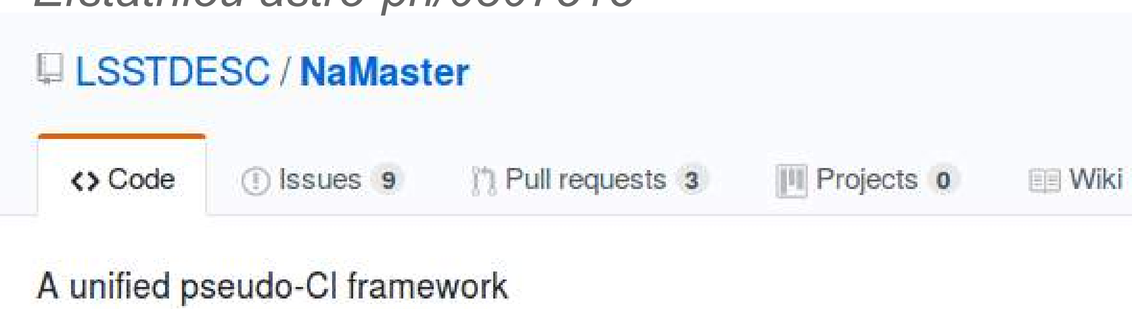
Code: https://github.com/LSSTDESC/NaMaster
Docs: https://namaster.readthedocs.io/en/latest/index.html

A unified pseudo-C_ℓ code

What **features** does it implement?

- Calculate PCL power spectra (including coupling matrix, etc.)
- In curved and flat skies
- Spin-0 (density, CMB T) and spin-2 (shear, CMB Q/U) quantities
- Bells and whistles:
  - Mode deprojection
  - E/B mode purification
- Gaussian covariances → Garcia-Garcia C., DA, Bellini E. arXiv:1906.11765

Efstathiou astro-ph/0307515



Code: https://github.com/LSSTDESC/NaMaster
Docs: https://namaster.readthedocs.io/en/latest/index.html

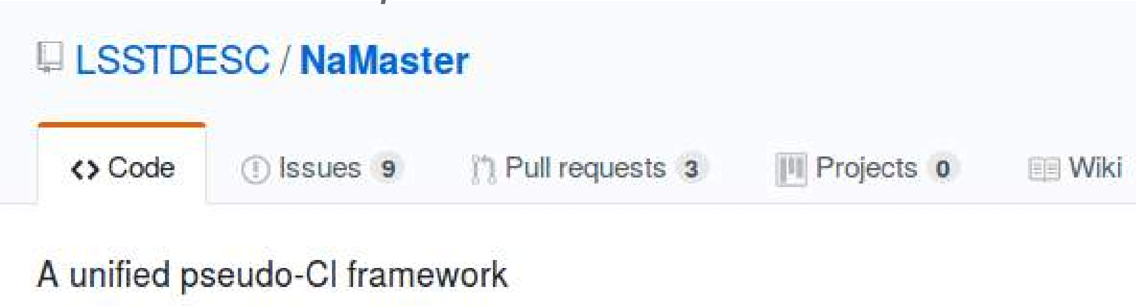
A unified pseudo-C_ℓ code

What **features** does it implement?

- Calculate PCL power spectra (including coupling matrix, etc.)
- In curved and flat skies
- Spin-0 (density, CMB T) and spin-2 (shear, CMB Q/U) quantities
- Bells and whistles:
  - Mode deprojection
  - E/B mode purification
- Gaussian covariances

Garcia-Garcia C., DA, Bellini E.
arXiv:1906.11765

Efstathiou astro-ph/0307515



Code: https://github.com/LSSTDESC/NaMaster
Docs: https://namaster.readthedocs.io/en/latest/index.html

Mode deprojection

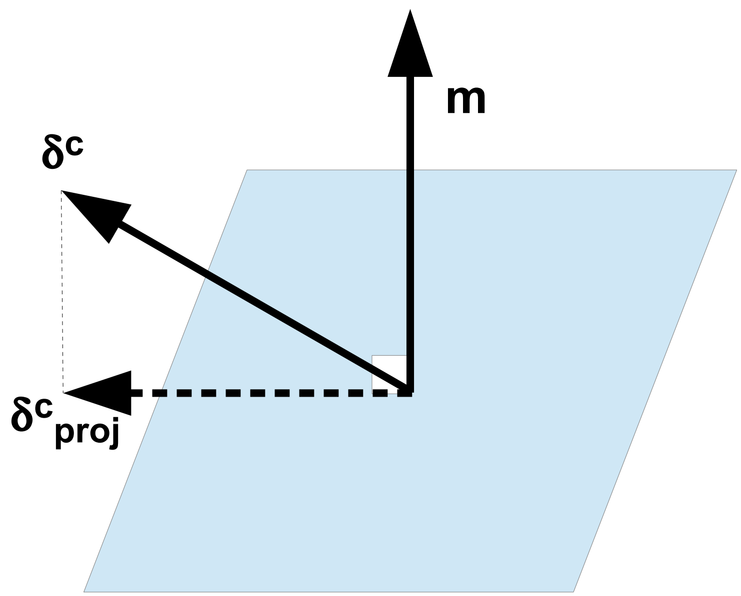
A. Slosar: "The greatest thing since sliced bread"

- **Masking:** if I have a bad pixel, I make sure it doesn't get used.
- **Mode deprojection** is the extension of this idea into an arbitrary linear combination of pixels.
  Imagine contaminating your data field as

True map

Observed map → $$\delta_i^c = \delta_i + \alpha m_i$$ ← Contaminant template (e.g. dust map)

A proper analysis would marginalize over α.



Leistedt et al. 1306.0005
Elsner et al. 1609.03577

Mode deprojection

A. Slosar: "The greatest thing since sliced bread"

- **Masking:** if I have a bad pixel, I make sure it doesn't get used.
- **Mode deprojection** is the extension of this idea into an arbitrary linear combination of pixels.
  Imagine contaminating your data field as

$$\delta_i^c = \delta_i + \alpha m_i$$

Observed map → δ_i^c; True map → δ_i; Contaminant template (e.g. dust map) → m_i

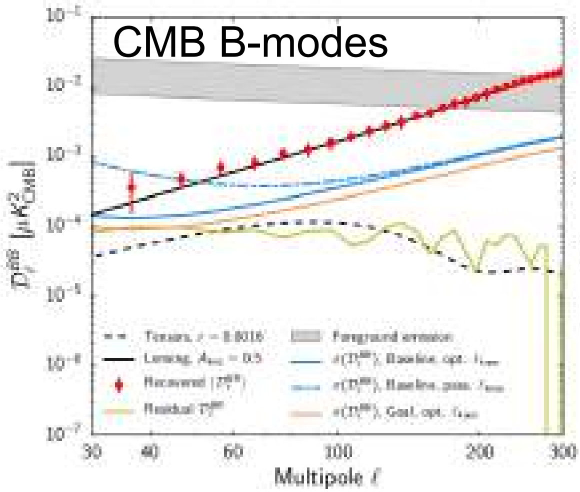
A proper analysis would marginalize over α.

If you do the maths, in PCL this amounts to:

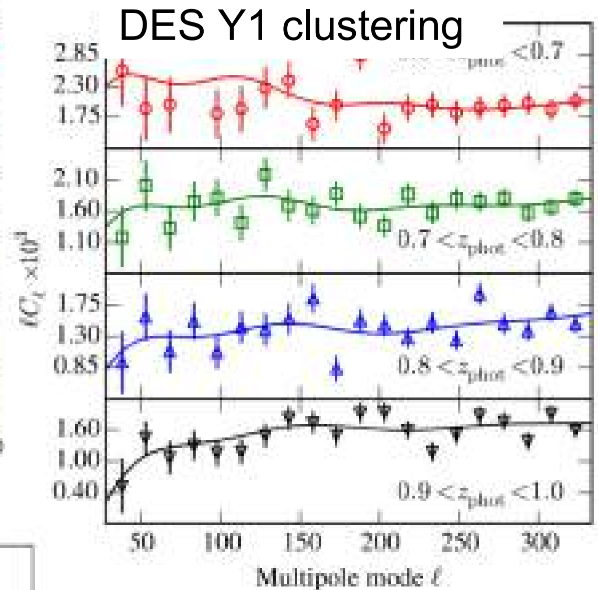
- Finding the best fit value of α.
- Subtracting a contaminant map from the data using this α
- Calculate the PCL and **correct for the bias** this subtraction has produced
- Multiply by the inverse of the mode-coupling matrix

Leistedt et al. 1306.0005
Elsner et al. 1609.03577

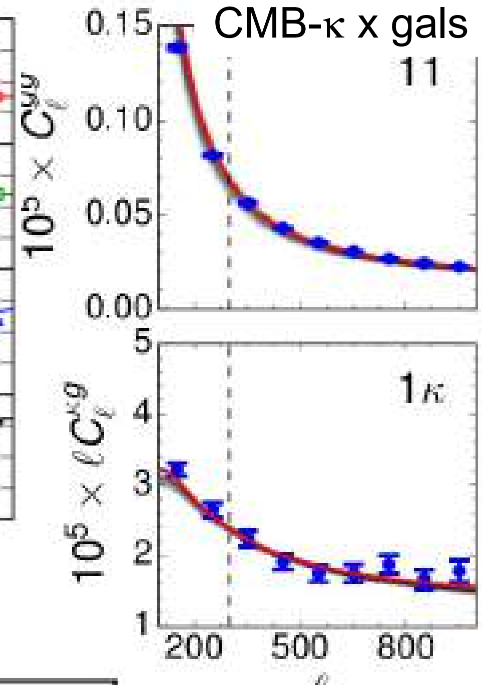
NaMaster



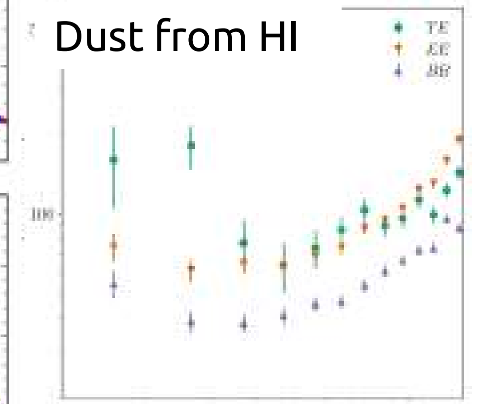
SO et al. 1808.07445



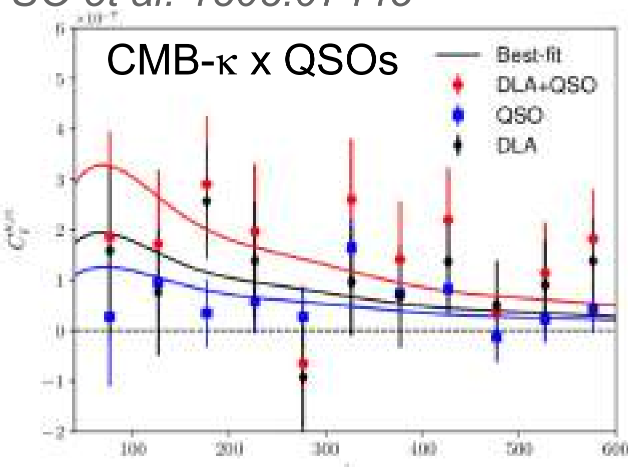
DES et al. 1807.10163



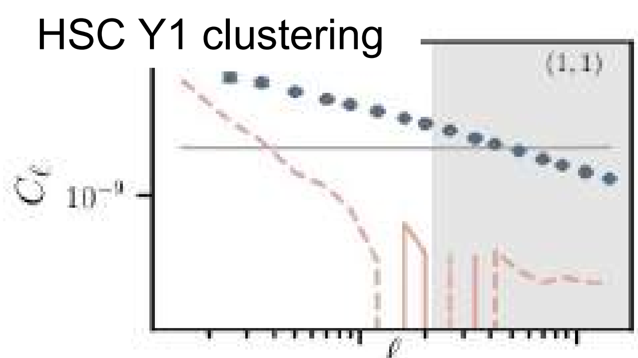
Krolewski et al. 1909.07412



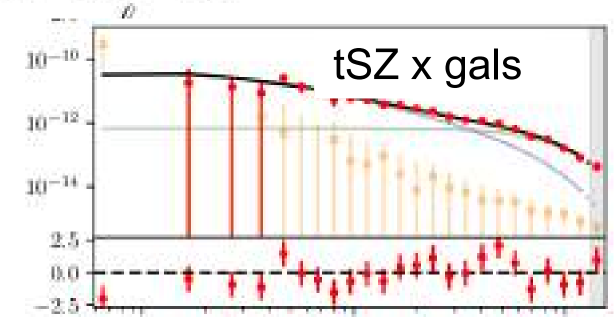
Hensley & Clark 1909.11673



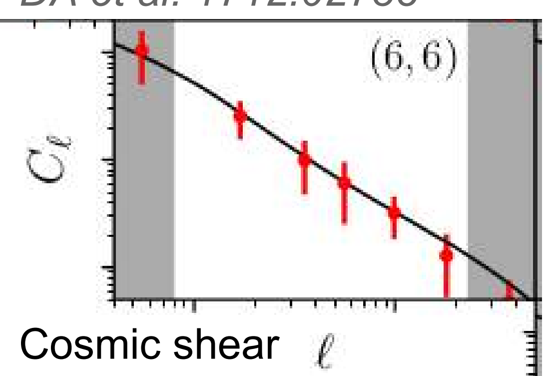
DA et al. 1712.02738



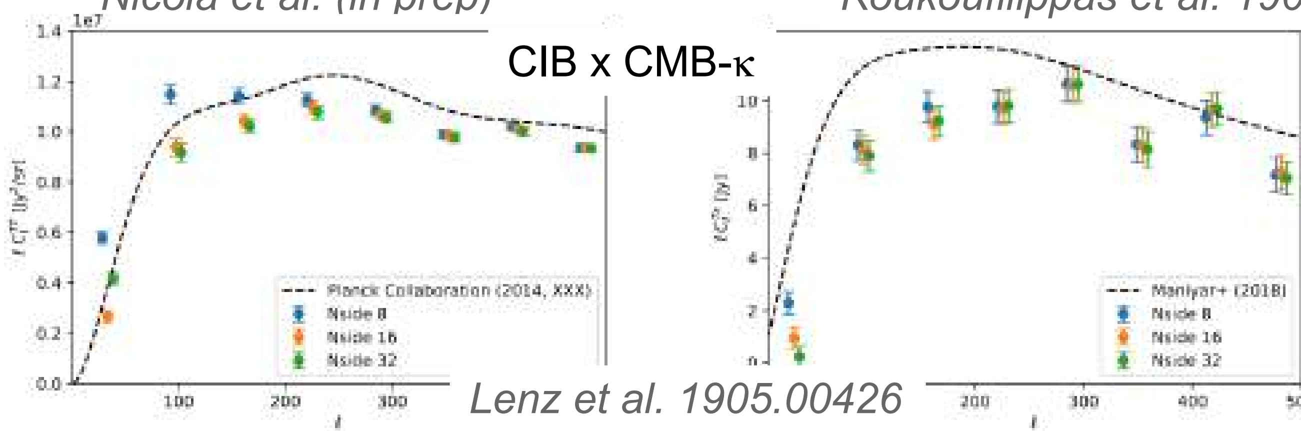
Nicola et al. (in prep)



Koukoufilippas et al. 1909.09102



Bellini et al. 1903.04957



Lenz et al. 1905.00426

•••

Example: tomographic analysis

Covariance matrix

Gaussian likelihood

$$-2 \log P(d|\theta) = (d - t(\theta))^T C^{-1} (d - t(\theta)) + L_0$$

Vector of cross-correlations

Theory prediction

Accuracy of $t(\theta)$ >> statistical power.
LSST's statistical power will be awesome.

Requirements for LSST:

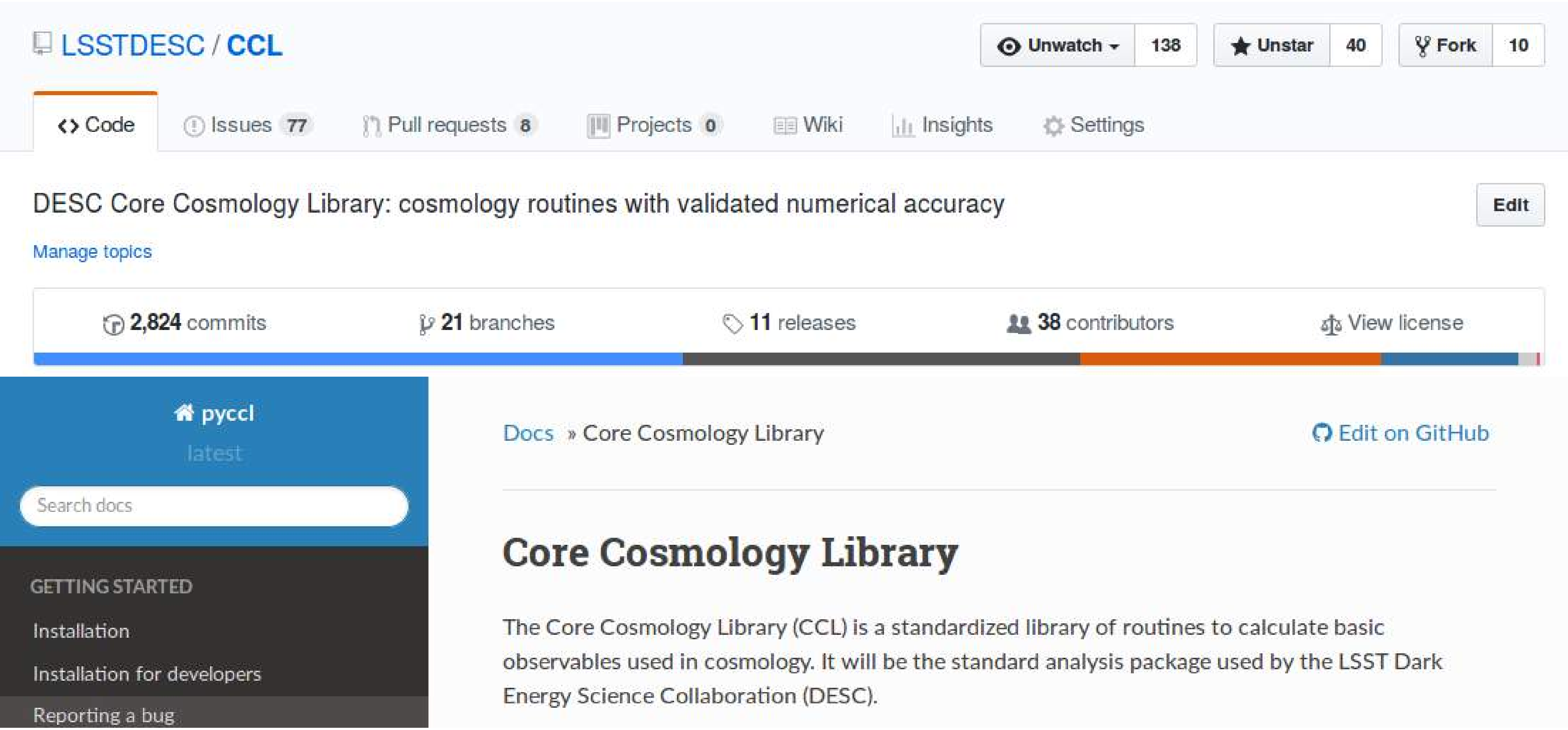
- Accuracy (errors well below statistical uncertainties)
- Robustness (thorough code validation and comparison)
- Flexibility (many observables, many cosmological models, ability to vary models and absorb systematics)
- Numerical performance (reasonable MCMC-ing time)

Robust theory predictions

Core Cosmology Library: precision cosmological predictions for LSST

Chisari E., DA, E. Krause +27, arXiv:1812.05995

The Core Cosmology Library



Code: https://github.com/LSSTDESC/CCL
Docs: https://ccl.readthedocs.io/en/latest/
Latest release: https://github.com/LSSTDESC/CCL/releases/tag/v2.0.1

Code validation

Strict code validation requirements

- All calculations are performed with at least one different independent code.
- Agreement must be found within well-motivated/crazy stringent requirements.
- Alternative calculations are kept as benchmarks.
- CCL is automatically compared against benchmarks whenever a new addition is made to the code.
- Unit tested (~95%).

Code validation

Currently implemented:

- Background quantities and linear growth.
- Matter power spectrum
  - Links to CAMB, CLASS, CosmicEmu, fast approximations
- Halo model:
  - Mass function
  - Bias
  - Concentrations
  - Profiles
  - Halo model power spectra
  - Easily generalisable
- Angular power spectra
  - Galaxy clustering, cosmic shear, CMB lensing
  - Easily generalisable
- Angular correlations functions
- 3D correlation functions

Used in a number of real-life analyses:

Cosmic shear: arXiv:1903.04957
Intrinsic alignments: arXiv:1911.01582,1901.09925
Cross-correlations: arXiv:1712.02738,1909.09102

Code: https://github.com/LSSTDESC/CCL
Docs: https://ccl.readthedocs.io/en/latest/
Latest release: https://github.com/LSSTDESC/CCL/releases/tag/v2.0.1

Covariances and data compression

Gaussian likelihood

$$-2 \log P(d|\theta) = (d - t(\theta))^T C^{-1} (d - t(\theta)) + L_0$$

Covariance matrix

Vector of cross-correlations

Theory prediction

Covariance matrices and data compression

A tomographic two-point function analysis already compresses the initial data vector significantly:

Catalogue with ~billions of objects and >5 quantities per object

↓

A number of cross-correlations between sub-samples of these

What is the actual number of cross correlations?

Covariance matrices and data compression

A tomographic two-point function analysis already compresses the initial data vector significantly:

Catalogue with ~billions of objects and >5 quantities per object

↓

A number of cross-correlations between sub-samples of these

What is the actual number of cross correlations?

Let's take an ideal LSST as an example:
10 redshift bins for lensing. 10 bins for clustering. 15 angular bins.

$$N_d = N_\theta \, N_{bin} \, (N_{bin}+1) \, / \, 2 = 3150$$

Compression factor: $\sim 3 \times 10^6$ → pretty good!

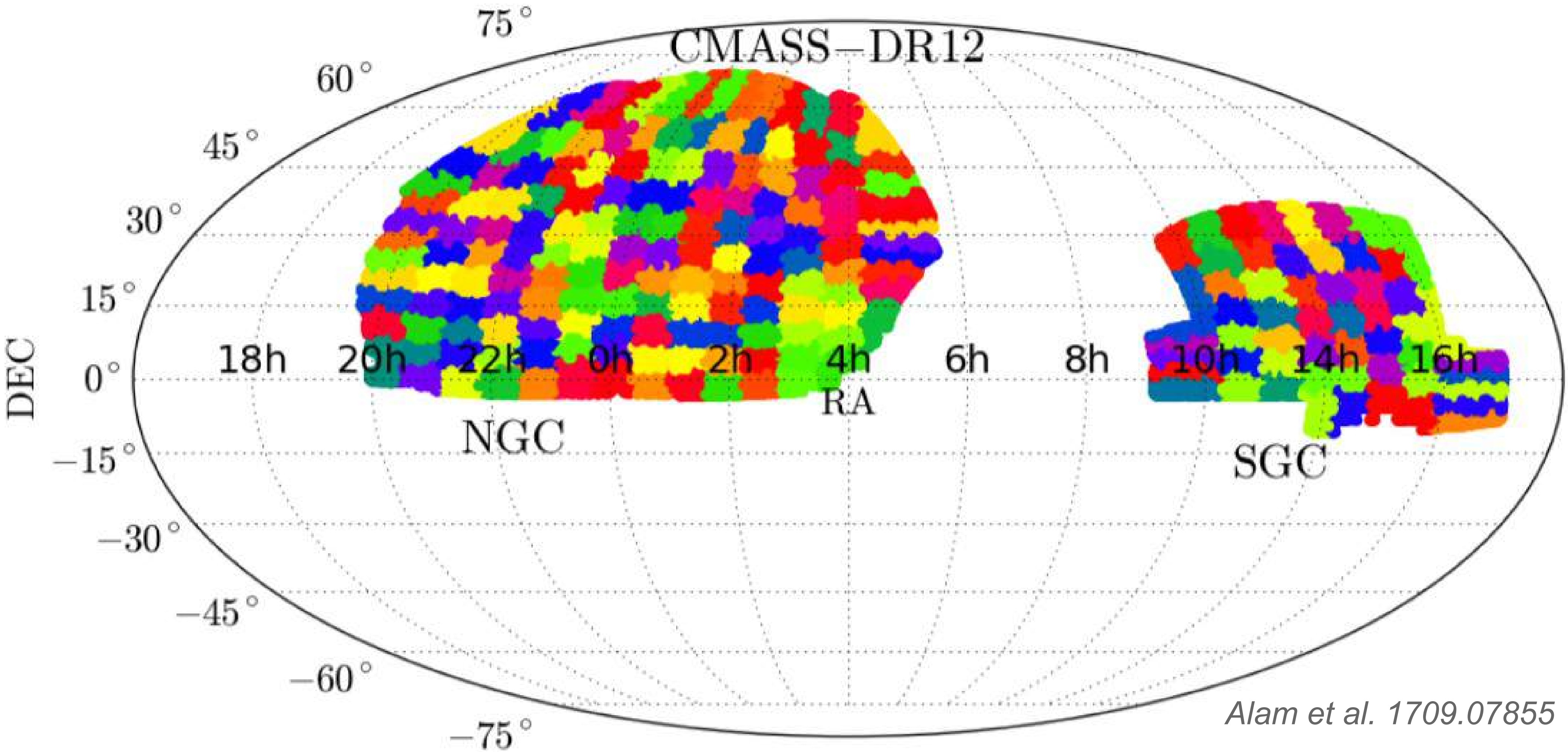
Achieved by:
- Selecting only the most informative summary statistic.
- Averaging over equivalent modes (e.g. using statistical isotropy).

However, now we need to compute the data **covariance matrix**.

Computing the covariance matrix

Different methods:

- Jackknife/bootstrap: use sub-samples of your own data.

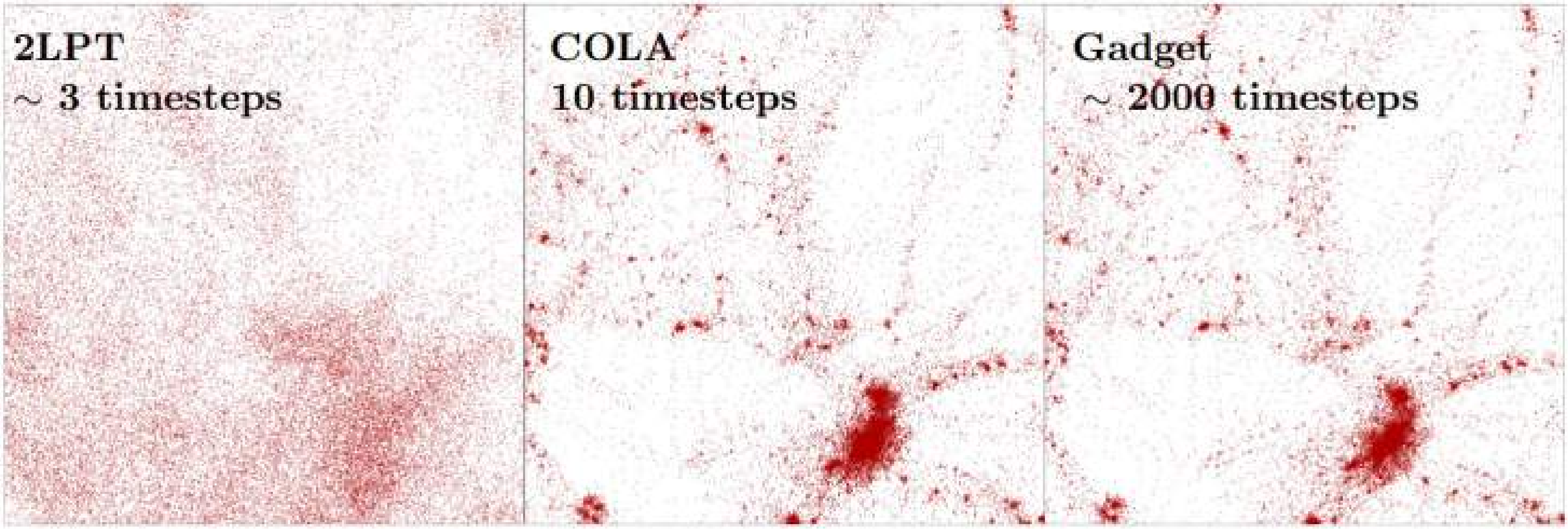


Alam et al. 1709.07855

Computing the covariance matrix

Different methods:

- Jackknife/bootstrap: use sub-samples of your own data.
- Mock catalogues: based on N-body sims or fast methods (Gaussian, FLASK, 2LPT, COLA, PINOCHIO, PTHALOS, QuickPM …)



Tassev et al. 1301.0322

For both of these, rule of thumb is $N_{samples}$ > 10 x (size of data vector).
Then, $O(3 \times 10^4)$ mocks/JKs are needed (covering the same volume as LSST).

Computing the covariance matrix

Different methods:

- Jackknife/bootstrap: use sub-samples of your own data.
- Mock catalogues: based on N-body sims or fast methods (Gaussian, lognormal, 2LPT, COLA, PINOCHIO, PTHALOS, QuickPM …)
- Analytical covariance matrix:
  Gaussian disconnected part:

Krause & Eifler 1601.05779

$$\mathrm{Cov}^{\mathrm{G}}\left(C^{ij}_{AB}(l_1), C^{kl}_{CD}(l_2)\right) = \frac{4\pi\delta_{l_1 l_2}}{\Omega_{\mathrm{s}}(2l_1+1)\Delta l_1}\left[\left(C^{ik}_{AC}(l_1)+\delta_{ik}\delta_{AC}N^i_A\right)\left(C^{jl}_{BD}(l_2)+\delta_{jl}\delta_{BD}N^j_B\right)+\left(C^{il}_{AD}(l_1)+\delta_{il}\delta_{AD}N^i_A\right)\left(C^{jk}_{BC}(l_2)+\delta_{jk}\delta_{BC}N^j_B\right)\right]$$

SSC

$$\mathrm{Cov}^{\mathrm{SSC}}\left(C^{ij}_{AB}(l_1), C^{kl}_{CD}(l_2)\right) = \int \mathrm{d}\chi\, \frac{q^i_A(\chi)q^j_B(\chi)q^k_C(\chi)q^l_D(\chi)}{\chi^4}\,\frac{\partial P_{AB}(l_1/\chi, z(\chi))}{\partial\delta_b}\,\frac{\partial P_{CD}(l_2/\chi, z(\chi))}{\partial\delta_b}\sigma_b(\Omega_{\mathrm{s}}; z(\chi))$$

Relevant connected parts

$$\mathrm{Cov}^{\mathrm{NG,0}}\left(C^{ij}_{AB}(l_1), C^{kl}_{CD}(l_2)\right) = \frac{1}{\Omega_{\mathrm{s}}}\int_{|\mathbf{l}|\in l_1}\frac{d^2\mathbf{l}}{A(l_1)}\int_{|\mathbf{l}'|\in l_2}\frac{d^2\mathbf{l}'}{A(l_2)}\int \mathrm{d}\chi\,\frac{q^i_A(\chi)q^j_B(\chi)q^k_C(\chi)q^l_D(\chi)}{\chi^6}\,T^{ijkl}_{ABCD}(\mathbf{l}/\chi, -\mathbf{l}/\chi, \mathbf{l}'/\chi, -\mathbf{l}'/\chi; z(\chi))$$

+ double Hankel transform if you work in real space
+ probably worry about survey geometry (mode coupling)

$$\left\langle \Delta\tilde{C}^{ab}_{\ell}\Delta\tilde{C}^{cd}_{\ell'}\right\rangle = \sum_{mm'}\sum_{\mathbf{l}_1\mathbf{l}_2}\left(C^{ac}_{\ell_1}C^{bd}_{\ell_2}W^a_{\mathbf{l}\mathbf{l}_1}W^b_{\mathbf{l}\mathbf{l}_2}W^c_{\mathbf{l}\mathbf{l}_1}W^d_{\mathbf{l}\mathbf{l}_2} + C^{ad}_{\ell_1}C^{bc}_{\ell_2}W^a_{\mathbf{l}\mathbf{l}_1}W^b_{\mathbf{l}\mathbf{l}_2}W^c_{\mathbf{l}\mathbf{l}_2}W^d_{\mathbf{l}\mathbf{l}_1}\right)$$

Computation scales very bad: $O(N_\theta^2\, N_{bin}^4)$

Computing the covariance matrix

Different methods:

- Jackknife/bootstrap: use sub-samples of your own data.
- Mock catalogues: based on N-body sims or fast methods (Gaussian, lognormal, 2LPT, COLA, PINOCHIO, PTHALOS, QuickPM …)
- Analytical covariance matrix:
  Gaussian disconnected part:

Krause & Eifler 1601.05779

$$\mathrm{Cov}^{\mathrm{G}}\left(C^{ij}_{AB}(l_1), C^{kl}_{CD}(l_2)\right) = \frac{4\pi\delta_{l_1 l_2}}{\Omega_s(2l_1+1)\Delta l_1}\left[\left(C^{ik}_{AC}(l_1)+\delta_{ik}\delta_{AC}N^i_A\right)\left(C^{jl}_{BD}(l_2)+\delta_{jl}\delta_{BD}N^j_B\right)+\left(C^{il}_{AD}(l_1)+\delta_{il}\delta_{AD}N^i_A\right)\left(C^{jk}_{BC}(l_2)+\delta_{jk}\delta_{BC}N^j_B\right)\right]$$

SSC

$$\mathrm{Cov}^{\mathrm{SSC}}\left(C^{ij}_{AB}(l_1), C^{kl}_{CD}(l_2)\right) = \int d\chi\, \frac{q^i_A(\chi)q^j_B(\chi)q^k_C(\chi)q^l_D(\chi)}{\chi^4}\, \frac{\partial P_{AB}(l_1/\chi, z(\chi))}{\partial \delta_b}\frac{\partial P_{CD}(l_2/\chi, z(\chi))}{\partial \delta_b}\sigma_b(\Omega_s; z(\chi))$$

Relevant connected parts

$$\mathrm{Cov}^{\mathrm{NG},0}\left(C^{ij}_{AB}(l_1), C^{kl}_{CD}(l_2)\right) = \frac{1}{\Omega_s}\int_{|\mathbf{l}|\in l_1}\frac{d^2\mathbf{l}}{A(l_1)}\int_{|\mathbf{l}'|\in l_2}\frac{d^2\mathbf{l}'}{A(l_2)}\int d\chi\,\frac{q^i_A(\chi)q^j_B(\chi)q^k_C(\chi)q^l_D(\chi)}{\chi^6}\,T^{ijkl}_{ABCD}(\mathbf{l}/\chi, -\mathbf{l}/\chi, \mathbf{l}'/\chi, -\mathbf{l}'/\chi; z(\chi))$$

+ double Hankel transform if you work in real space
+ probably worry about survey geometry (mode coupling)

Garcia-Garcia et al.
arXiv:1906.11765

$$\left\langle \Delta\tilde{C}^{ab}_{\ell}\Delta\tilde{C}^{cd}_{\ell'}\right\rangle = \sum_{mm'}\sum_{\mathbf{l}_1\mathbf{l}_2}\left(C^{ac}_{\ell_1}C^{bd}_{\ell_2}W^a_{\mathbf{l}\mathbf{l}_1}W^b_{\mathbf{l}\mathbf{l}_2}W^c_{\mathbf{l}'\mathbf{l}_1}W^d_{\mathbf{l}'\mathbf{l}_2}+C^{ad}_{\ell_1}C^{bc}_{\ell_2}W^a_{\mathbf{l}\mathbf{l}_1}W^b_{\mathbf{l}\mathbf{l}_2}W^c_{\mathbf{l}'\mathbf{l}_2}W^d_{\mathbf{l}'\mathbf{l}_1}\right)$$

Computation scales very bad: $O(N_\theta^2\, N_{bin}^4)$

Computing the covariance matrix

Different methods:

- Jackknife/bootstrap: use sub-samples of your own data.
- Mock catalogues: based on N-body sims or fast methods (Gaussian, lognormal, 2LPT, COLA, PINOCHIO, PTHALOS, QuickPM …)
- Analytical covariance matrix:

All of these cases would benefit massively from reducing the size of the data vector.

Can we compress further?

Covariance matrices and data compression

Science-driven 3D data compression

DA, arXiv:1707.08950

Sheer shear:

weak lensing with one mode

E. Bellini, DA et al.
arXiv:1903.04957

The Karhunen-Loeve transform

Idea: find the linear combinations of your data that contain most of the information about a given parameter θ.

$$y_p \equiv \mathbf{e}_p^\dagger \mathbf{x}$$

Data: **x** → maps/$a_{\ell m}$s of a given set of tomographic observables (e.g. galaxy overdensity or shear in a set of redshift bins).

The linear coefficients **e** can be found as the eigenvectors of a generalized eigenvalue equation:

$$\partial_\theta \mathsf{C}\,\mathbf{e}_p = \lambda_p \mathsf{C}\,\mathbf{e}_p$$

Covariance of **x**

One generic parameter we could optimize for is the overall S/N amplitude. Maximizing this should provide us with most of the information about any parameter in most cases.

In this case, the eigenvalue equation reads:

$$(\mathsf{S} + \mathsf{N})\mathbf{e}_p = \lambda_p \mathsf{N}\mathbf{e}_p$$

Signal covariance

Noise covariance

Resulting modes y_p are uncorrelated and contain the maximum amount of information (info(y_0) > info(y_1) > …).

The Karhunen-Loeve transform

Example: galaxy clustering with spectroscopic redshifts.

x → galaxy overdensity in an infinitesimal redshift bin.

C → all possible cross-power spectra between bins (noise + signal)
N → flat, diagonal shot-noise power spectrum

The solution to the generalized eigenvalue equation (KL modes) is

$$e_{k,\ell}(z) \propto j_\ell\left(k\chi(z)\right)$$

i.e. KL transform in this case is the harmonic-Bessel transform.

The covariance of the resulting KL modes is

$$\lambda_{k,\ell} \propto P(k)$$

i.e. in this case the KL transform tells you to just compute the Fourier transform and estimate the 3D power spectrum (as expected!).

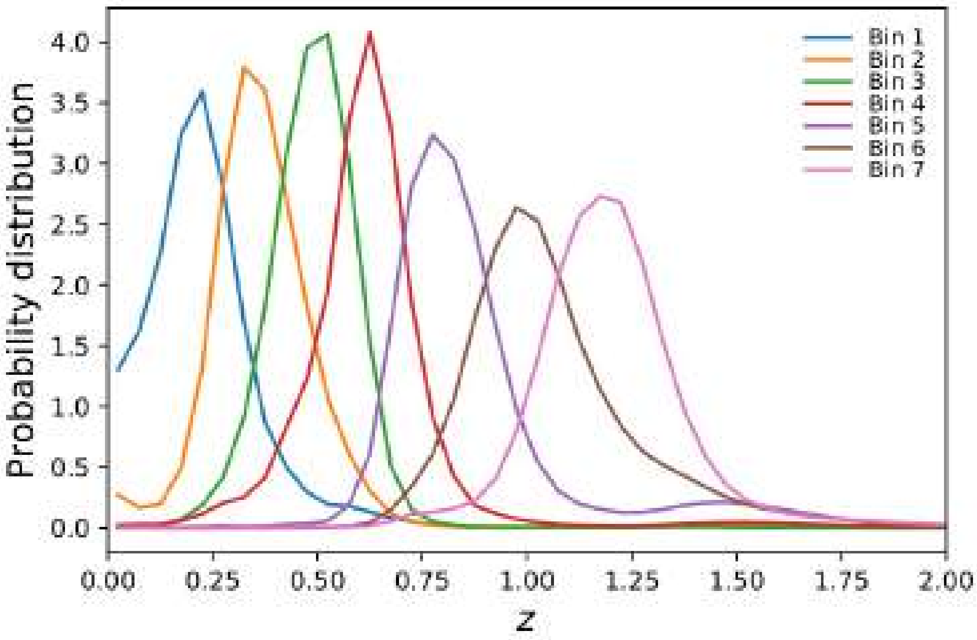
The KL eigenmodes are the generalization of a P(k) analysis to other types of data.

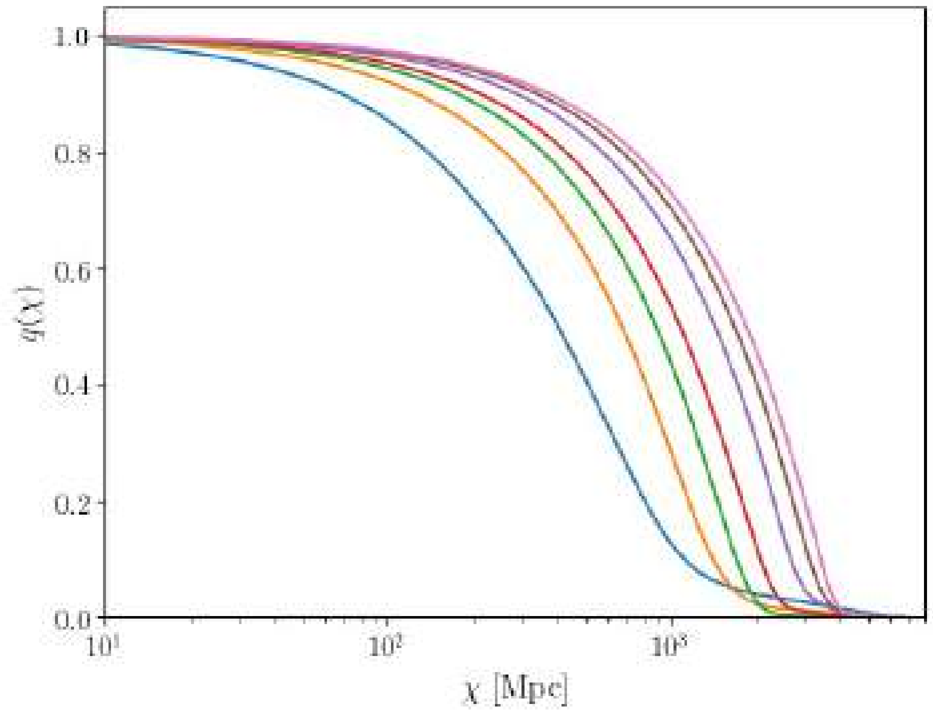
Data compression

Example: cosmic shear

$$C_{AB}^{ij}(l) = \int d\chi \frac{q_A^i(\chi) q_B^j(\chi)}{\chi^2} P_{AB}(l/\chi, z(\chi))$$

$$q_\kappa^i(\chi) = \frac{3H_0^2\Omega_m}{2c^2} \frac{\chi}{a(\chi)} \int_\chi^{\chi_h} d\chi' \frac{n_{\rm source}^i(z(\chi'))dz/d\chi'}{\bar{n}_{\rm source}^i} \frac{\chi' - \chi}{\chi'}$$



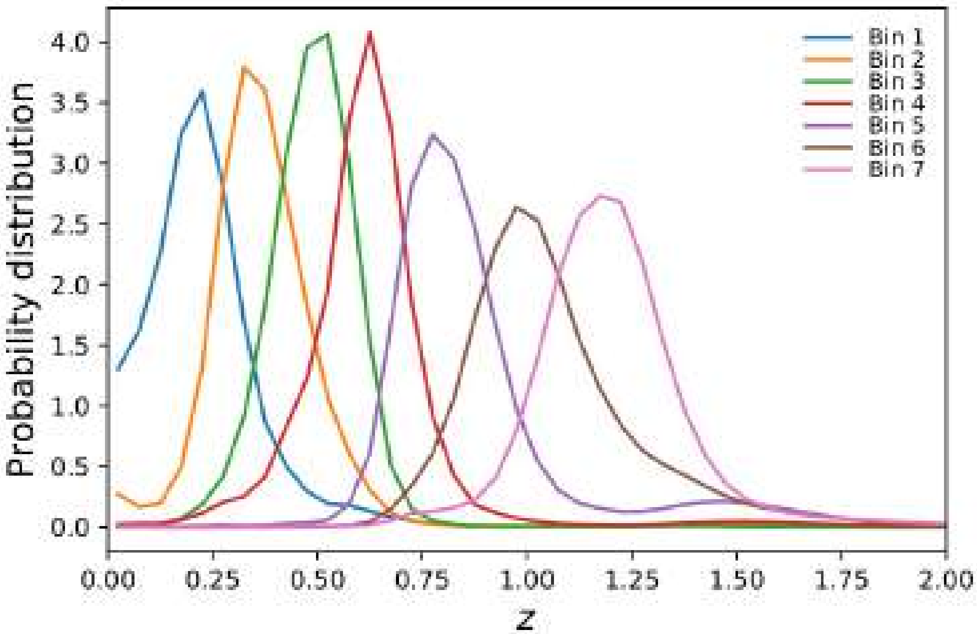


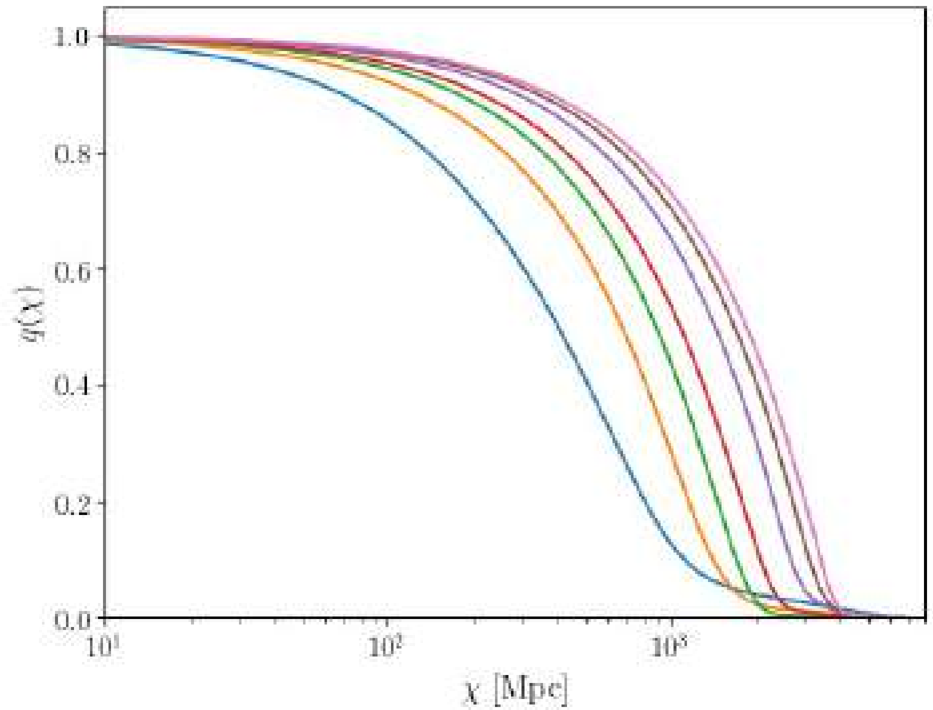
Data compression

Example: cosmic shear

$$C^{ij}_{AB}(l) = \int d\chi \frac{q^i_A(\chi) q^j_B(\chi)}{\chi^2} P_{AB}(l/\chi, z(\chi))$$

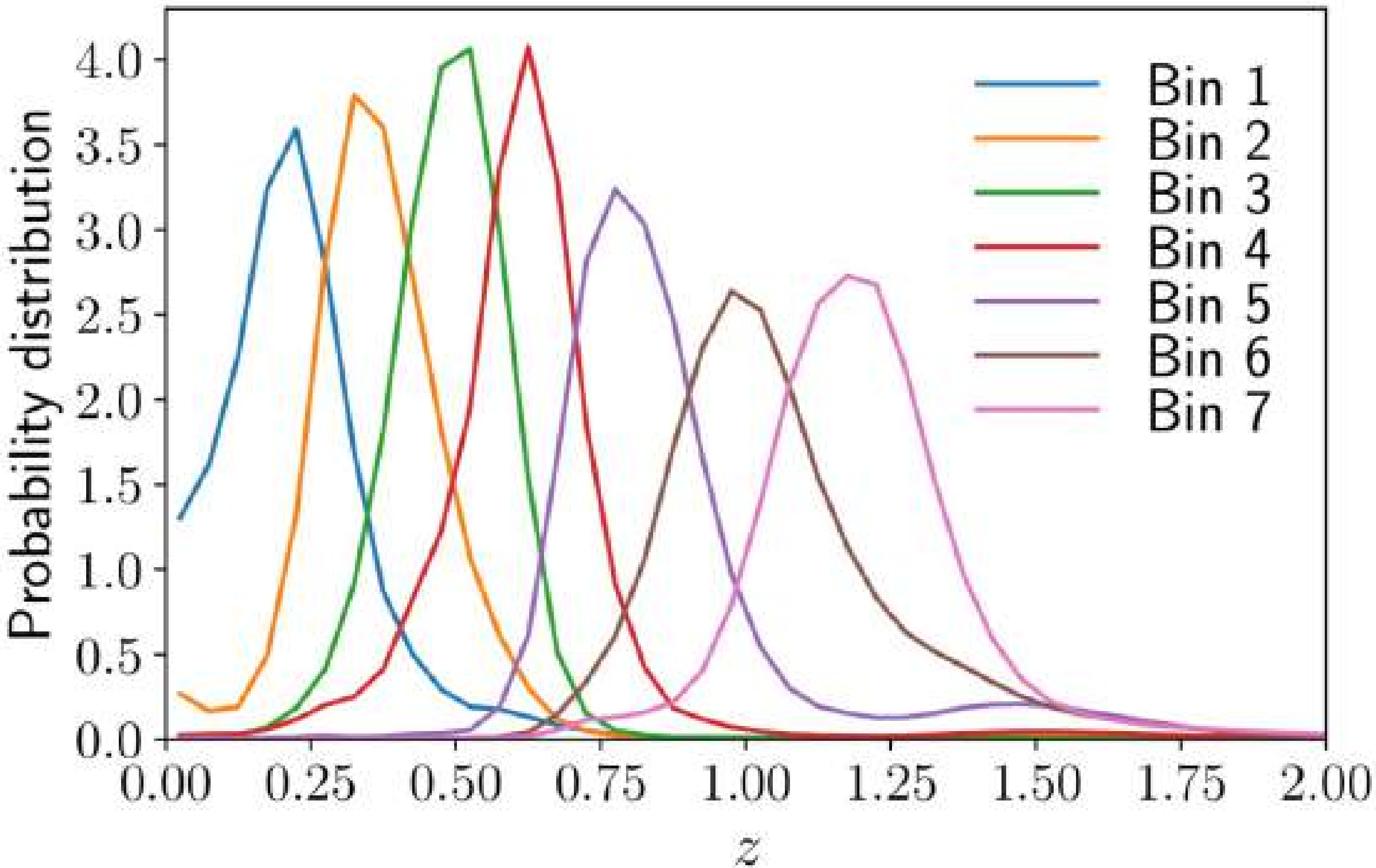
$$q^i_\kappa(\chi) = \frac{3H_0^2\Omega_m}{2c^2} \frac{\chi}{a(\chi)} \int_\chi^{\chi_h} \mathrm{d}\chi' \frac{n^i_{\rm source}(z(\chi'))dz/d\chi'}{\bar{n}^i_{\rm source}} \frac{\chi' - \chi}{\chi'}$$





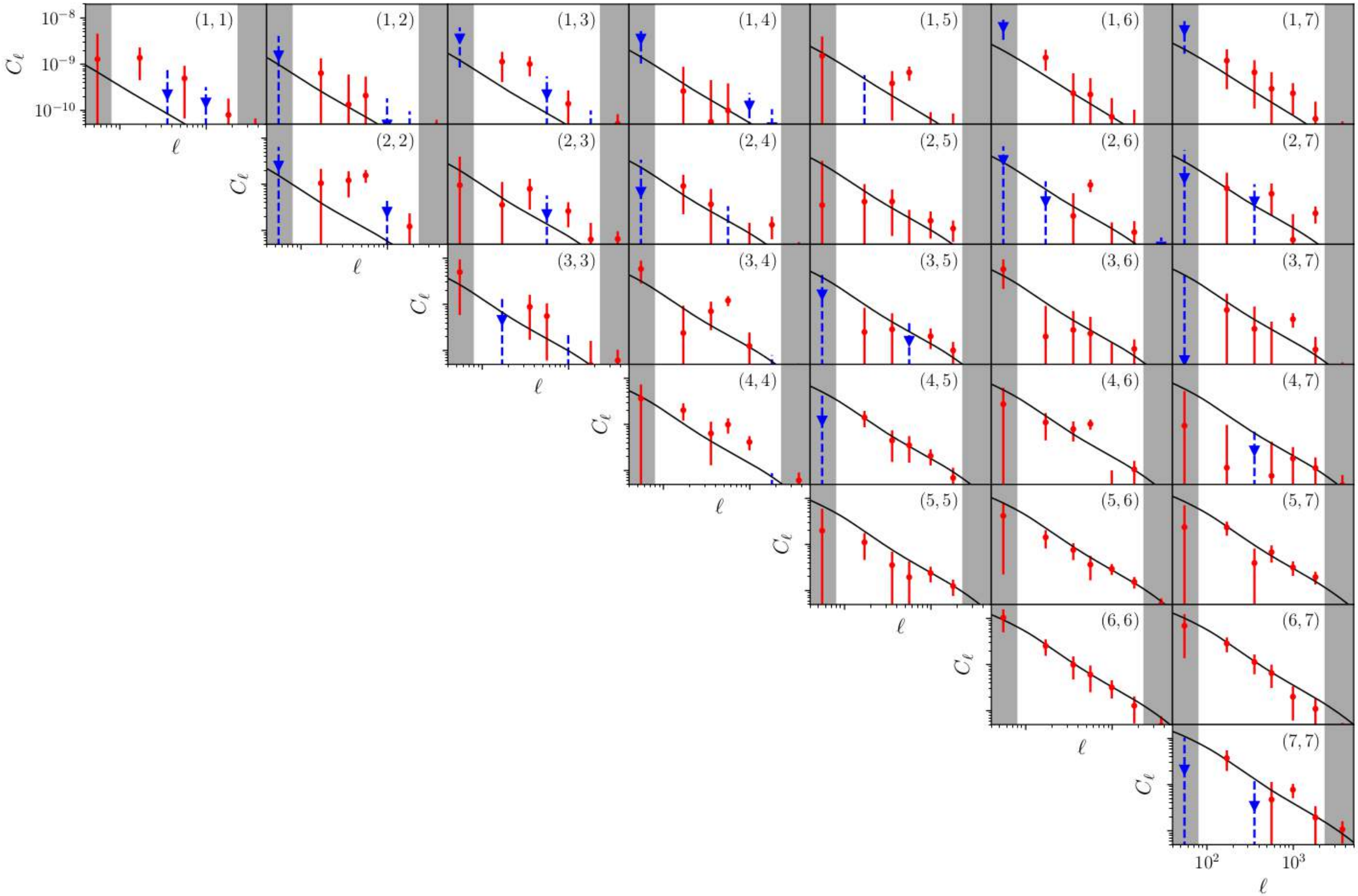
Different bins are very correlated.
Correlation → you have fewer *d.o.f.s* than you think.
You **can** compress further!

The KL transform: application to CFHTLens



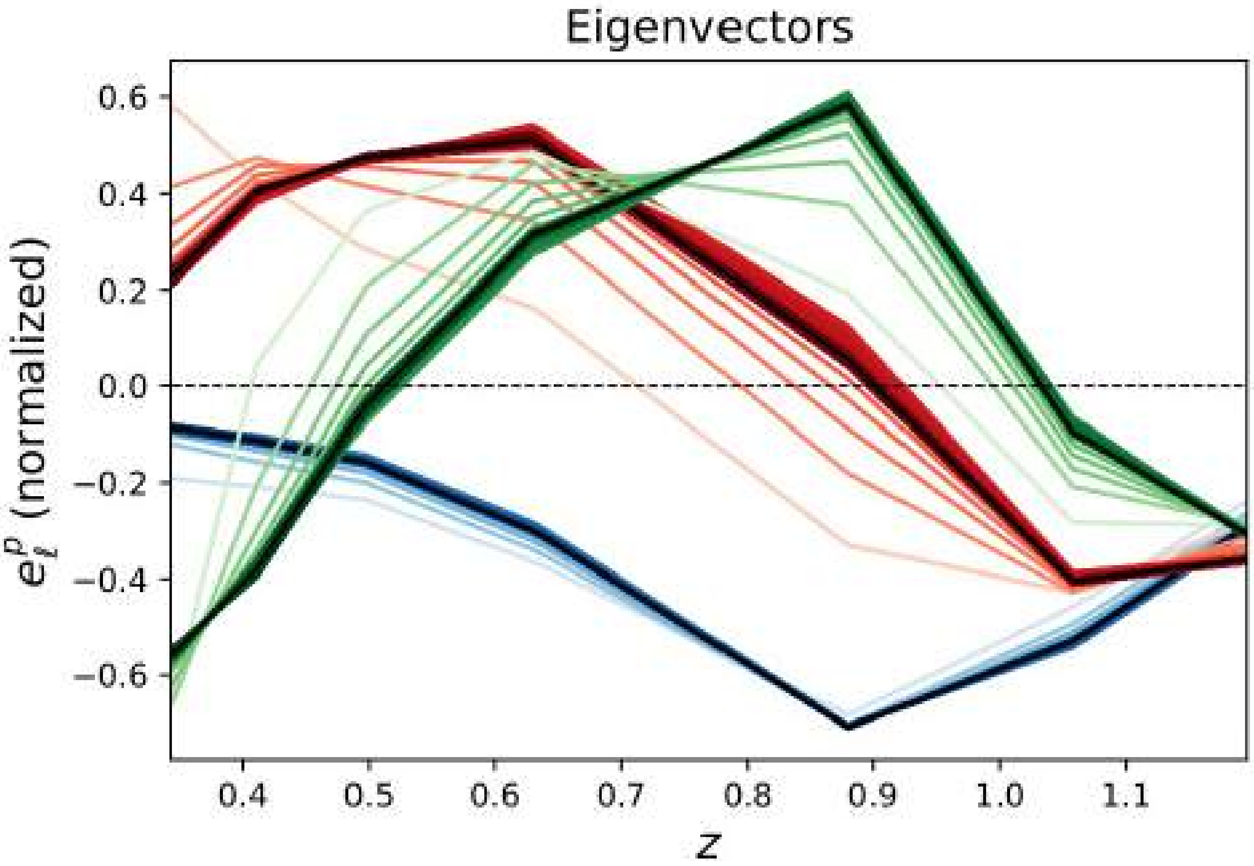
- Latest analysis (Joudaki et al. 2016) uses 7 tomographic bins in real space.

The KL transform: application to CFHTLens



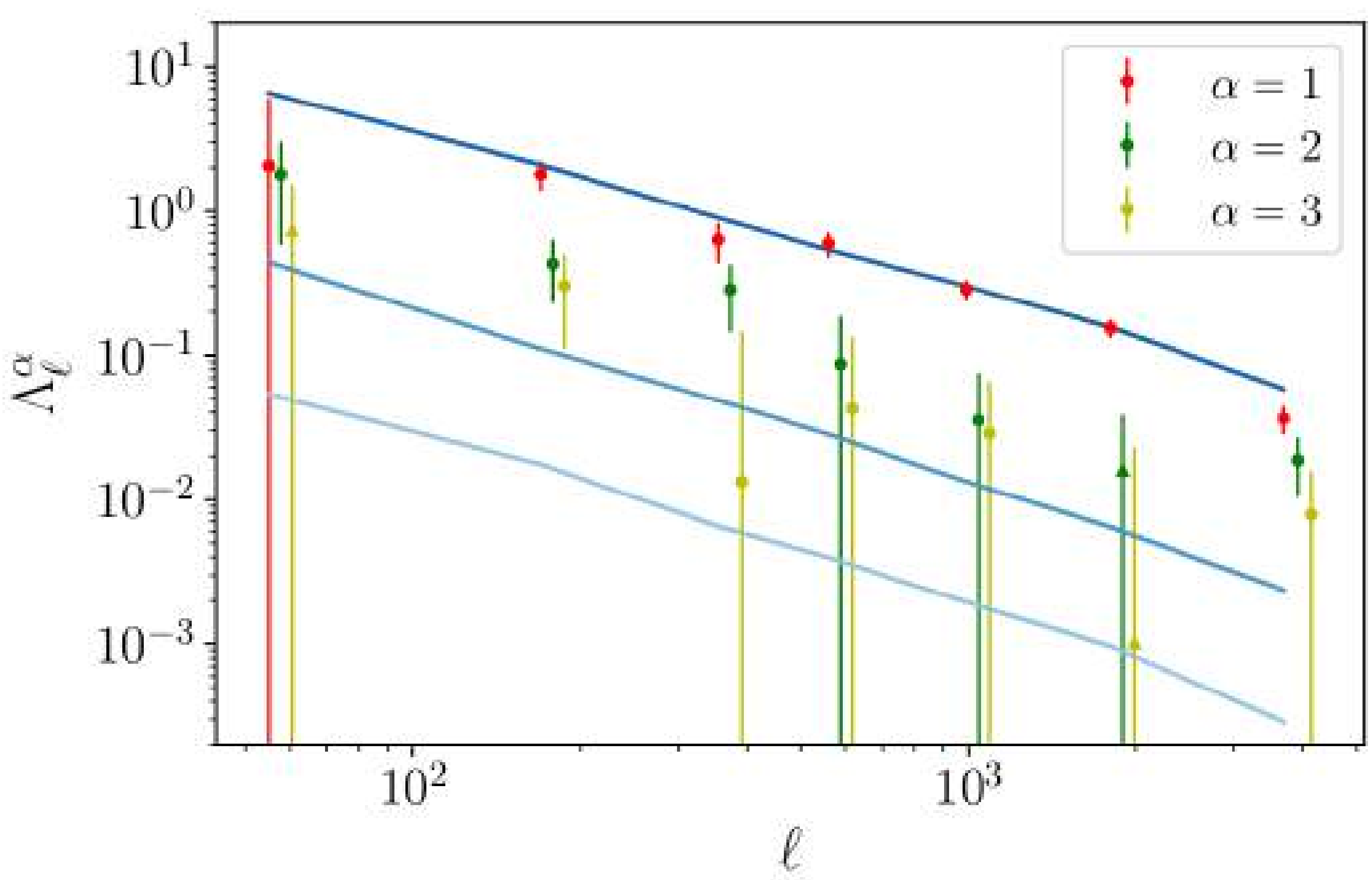
- Latest analysis (Joudaki et al. 2016) uses 7 tomographic bins in real space.
- Size of data vector: 2 x 5 x (7x8)/2 = 280 elements
  Power spectra estimated with NaMaster.

The KL transform: application to CFHTLens



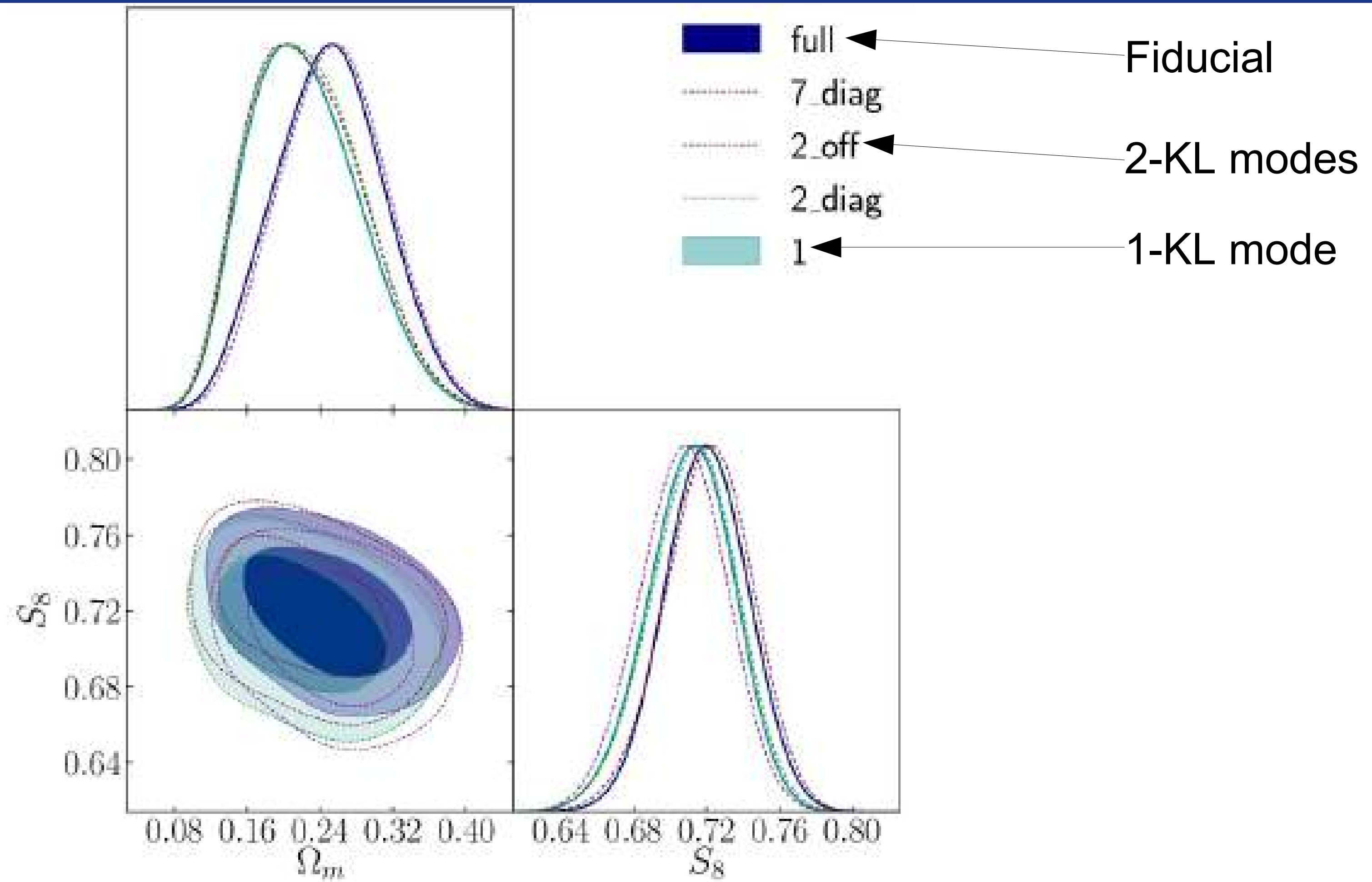
- Latest analysis (Joudaki et al. 2016) uses 7 tomographic bins in real space.
- Size of data vector: 2 x 5 x (7x8)/2 = 280 elements.
- Eigenvectors close to scale-independent.
  Think of them as redshift-dependent galaxy weights.

The KL transform: application to CFHTLens



- Latest analysis (Joudaki et al. 2016) uses 7 tomographic bins in real space.
- Size of data vector: 2 x 5 x (7x8)/2 = 280 elements.
- Eigenvectors close to scale-independent.
  Think of them as redshift-dependent galaxy weights.
- Majority of the signal concentrated in 1st KL mode.

The KL transform: application to CFHTLens



- Latest analysis (Joudaki et al. 2016) uses 7 tomographic bins in real space.
- Size of data vector: 2 x 5 x (7x8)/2 = 280 elements.
- Eigenvectors close to scale-independent.
  Think of them as redshift-dependent galaxy weights.
- The first 1-2 modes are able to recover the full constraining power.
  Compression factor ~19-30!

Data compression

Other uses of the KL transform:

- Large-scale effects: optimize fNL constraints.
- Systematics: remove modes that are most sensitive to e.g. intrinsic alignments, magnification …
  (basically put everything you don't like in the noise component)
- Foreground removal in 21cm experiments

Extreme data compression:

- *Alsing & Wandelt 1712.00012, Alsing et al. 1801.01497.*
- One summary statistic per free parameter.
- Can be made robust to systematics.
- Potentially more sensitive to modeling errors. Missing systematics may be more difficult to detect (KL at least gives you maps to inspect).

Parameter-dependent covariances

Gaussian likelihood

$$-2 \log P(d|\theta) = (d-t(\theta))^T C^{-1}(\theta) (d-t(\theta)) + L_0 \;?$$

- Do we have to take into account the parameter dependence of the covariance matrix?
- I.e. do we need to compute a new covariance at every point in an MCMC chain?

The effect on cosmological parameter estimation of a parameter-dependent covariance matrix

Kodwani D., DA, P. Ferreira

arXiv:1811.11584

Parameter-dependent covariances

$$-2 \log P(d|\theta) = (d-t(\theta))^T C^{-1}(\theta) (d-t(\theta)) + L_0 \ ?$$

- Do we have to take into account the parameter dependence of the covariance matrix?
- I.e. do we need to compute a new covariance at every point in an MCMC chain?
- Carron 2016: for Gaussian fields it's not only unnecessary, it's incorrect.
- The galaxy overdensity and cosmic shear aren't Gaussian, so do we need to worry about this at all?

The math

The information content of the covariance matrix can be quantified approximating the likelihood as Gaussian around the maximum (i.e. a la Fisher).

- Effect on parameter uncertainties:

$$\mathcal{F}_{\mu\nu} = \partial_\mu \mathbf{t}^T \Sigma^{-1} \partial_\nu \mathbf{t} + \frac{1}{2} \mathrm{Tr} \left(\Sigma^{-1} \partial_\mu \Sigma \Sigma^{-1} \partial_\nu \Sigma \right)$$

- Effect on parameter bias:

$$\Delta\theta_\mu = -\frac{1}{2} \mathcal{F}^{-1}_{\mu\nu} \mathcal{F}^{-1}_{\rho\tau} \partial_\rho \mathbf{t}^T \Sigma^{-1} \partial_\nu \Sigma \Sigma^{-1} \partial_\tau \mathbf{t}$$

The math

The information content of the covariance matrix can be quantified approximating the likelihood as Gaussian around the maximum (i.e. a la Fisher).

- Effect on parameter uncertainties:

$$\mathcal{F}_{\mu\nu} = \partial_\mu \mathbf{t}^T \Sigma^{-1} \partial_\nu \mathbf{t} + \frac{1}{2} \mathrm{Tr}\left(\Sigma^{-1} \partial_\mu \Sigma \Sigma^{-1} \partial_\nu \Sigma\right)$$

- Effect on parameter bias:

$$\Delta\theta_\mu = -\frac{1}{2} \mathcal{F}^{-1}_{\mu\nu} \mathcal{F}^{-1}_{\rho\tau} \partial_\rho \mathbf{t}^T \Sigma^{-1} \partial_\nu \Sigma \Sigma^{-1} \partial_\tau \mathbf{t}$$

Let's examine the dependence on f_{sky}.

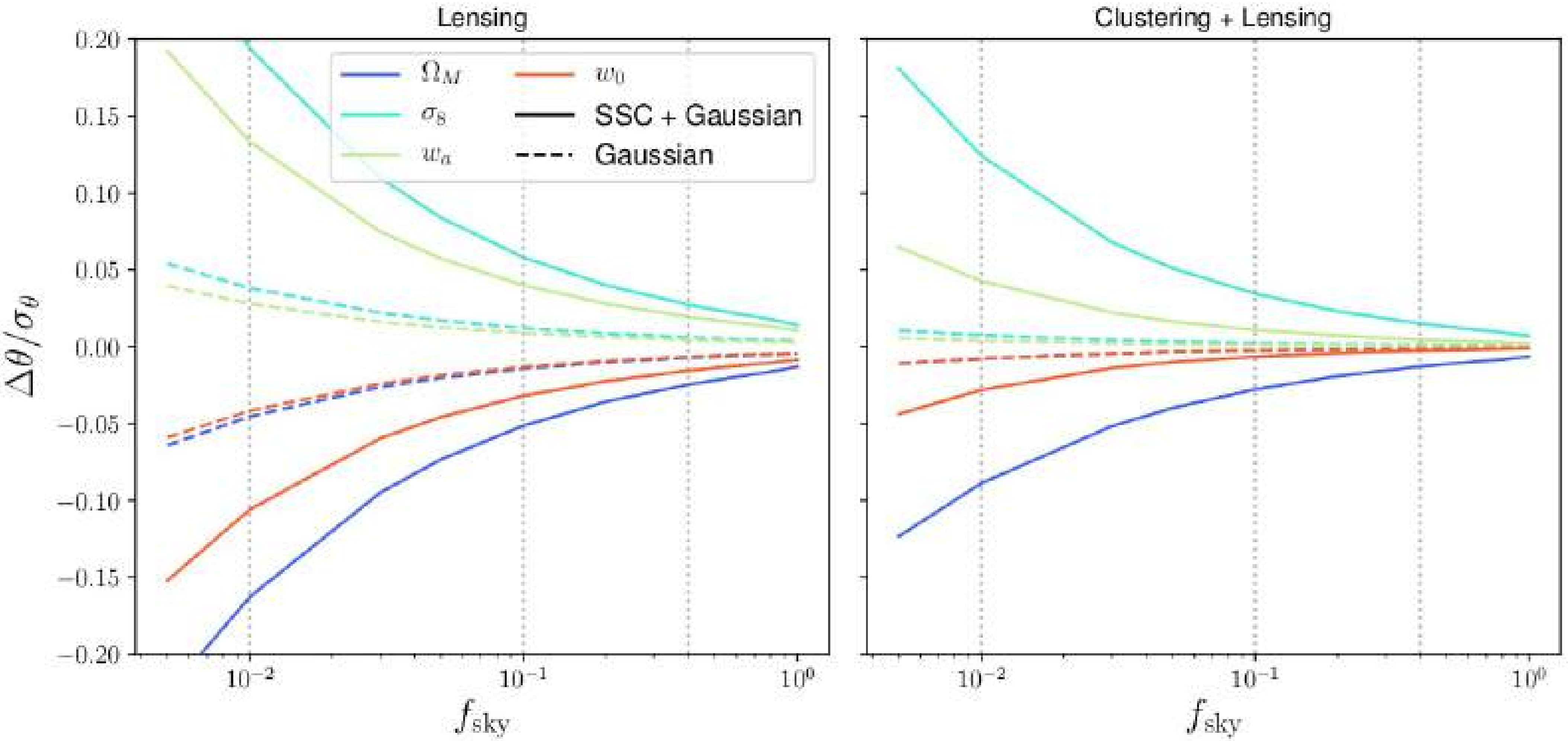
Roughly: $\Sigma \propto f_{\rm sky}^{-1}$

Then: $\Delta\theta \propto f_{\rm sky}^{-1}$, $\delta\sigma(\theta) \propto f_{\rm sky}^{-3/2}$

In general, the effects of a parameter-dependent covariance shrink with the number of modes in the analysis (same also with ℓ_{max}).

Parameter-dependent covariances

Results: parameter uncertainties



The parameter dependence of the covariance is irrelevant in all cases.

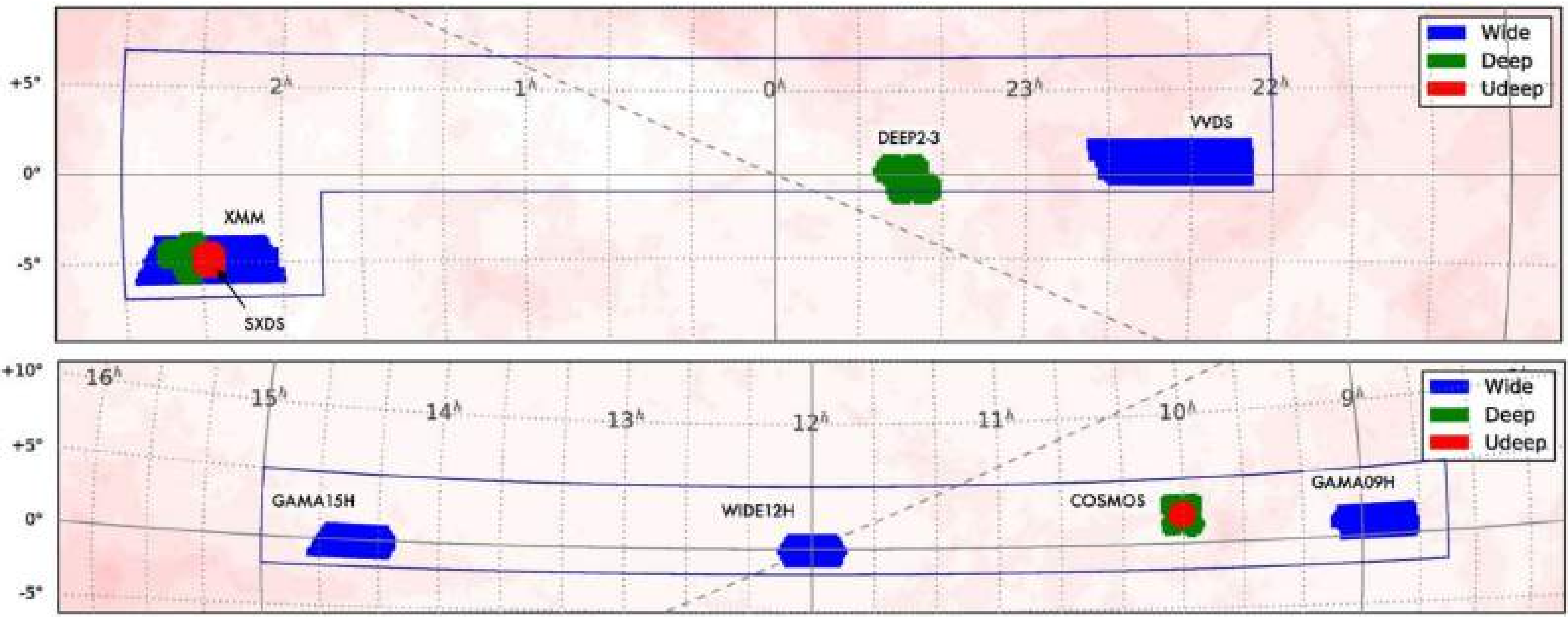
Fully worked-out example

Tomographic galaxy clustering with the Subaru Hyper Suprime-Cam first-year public data release

Nicola A., DA, Slosar A., F.J. Sanchez et al. (LSST DESC)

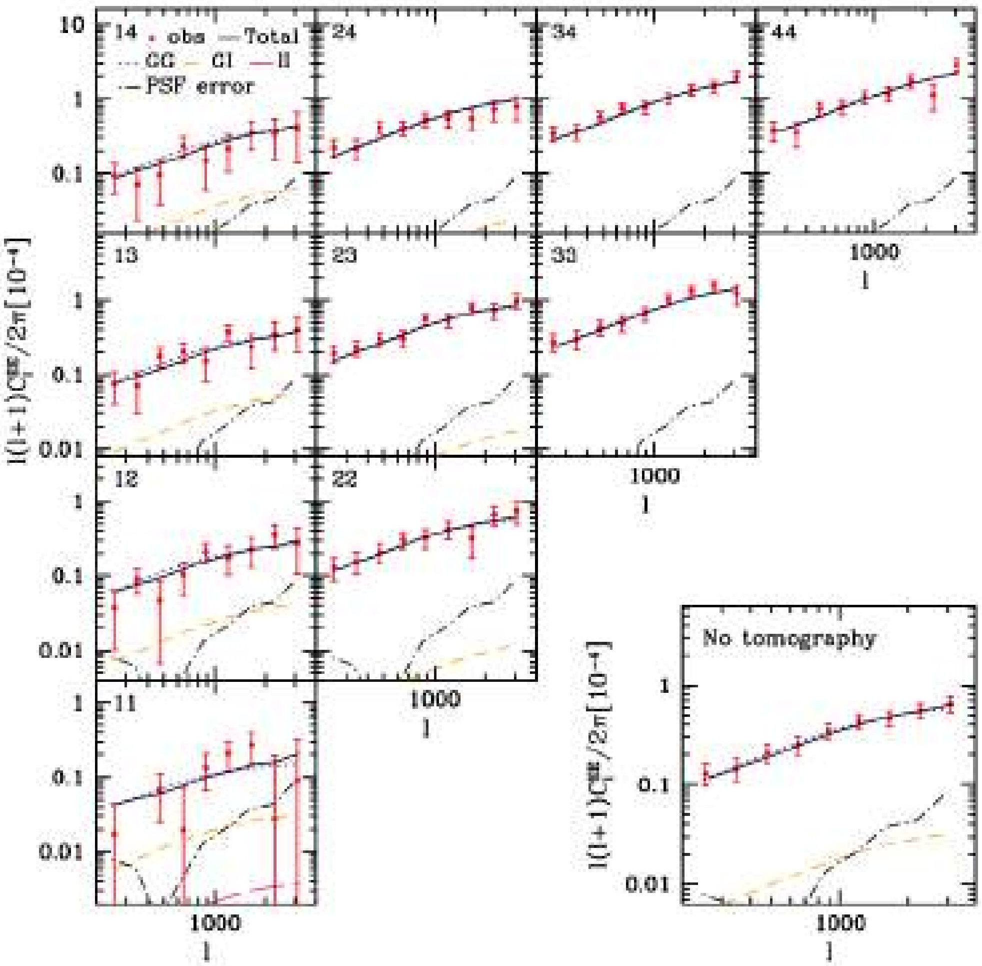
(arXiv coming soon!)

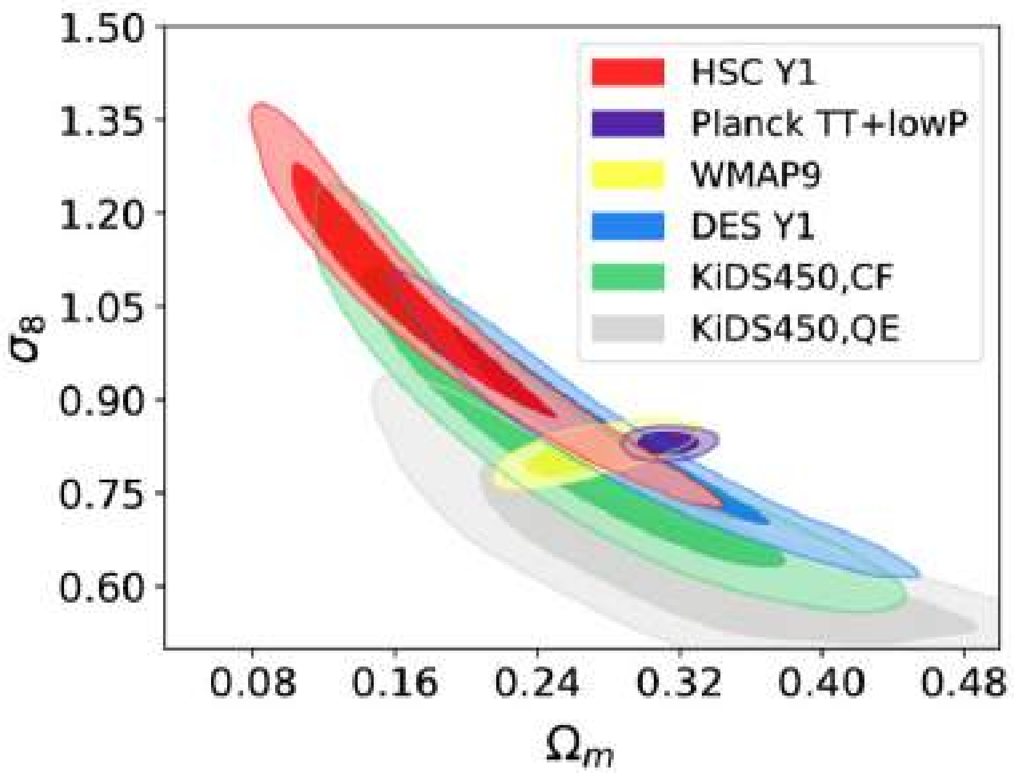
The HSC survey



- HSC end goal: 5-year survey covering 1400 sq-deg
- Deep (r_{lim} ~ 26), very good seeing (0.6'')
- 5 bands: grizy
- 1st-year data release: full depth on ~150 sq-deg in 6 fields (+ a few deeper fields)
- Precursor to LSST. Common DM pipeline, similar depth.

The HSC survey





Hikage et al. arXiv:1809.09148

Cosmological constraints from shear

Why is DESC interested?

- Same DM pipeline, similar depth: ideal testbed for our pipelines.
- Test viability of Fourier-space clustering analysis in the presence of sky systematics.
- Photometric clustering has focused on small samples with good photo-zs (e.g. LRGs, redMaGiC)
  - This will mean losing almost all of our galaxies in LSST. Can we do better?
  - How much more stringent are photo-z calibration requirements for galaxy clustering?
- No-one has looked at photometric clustering in HSC.

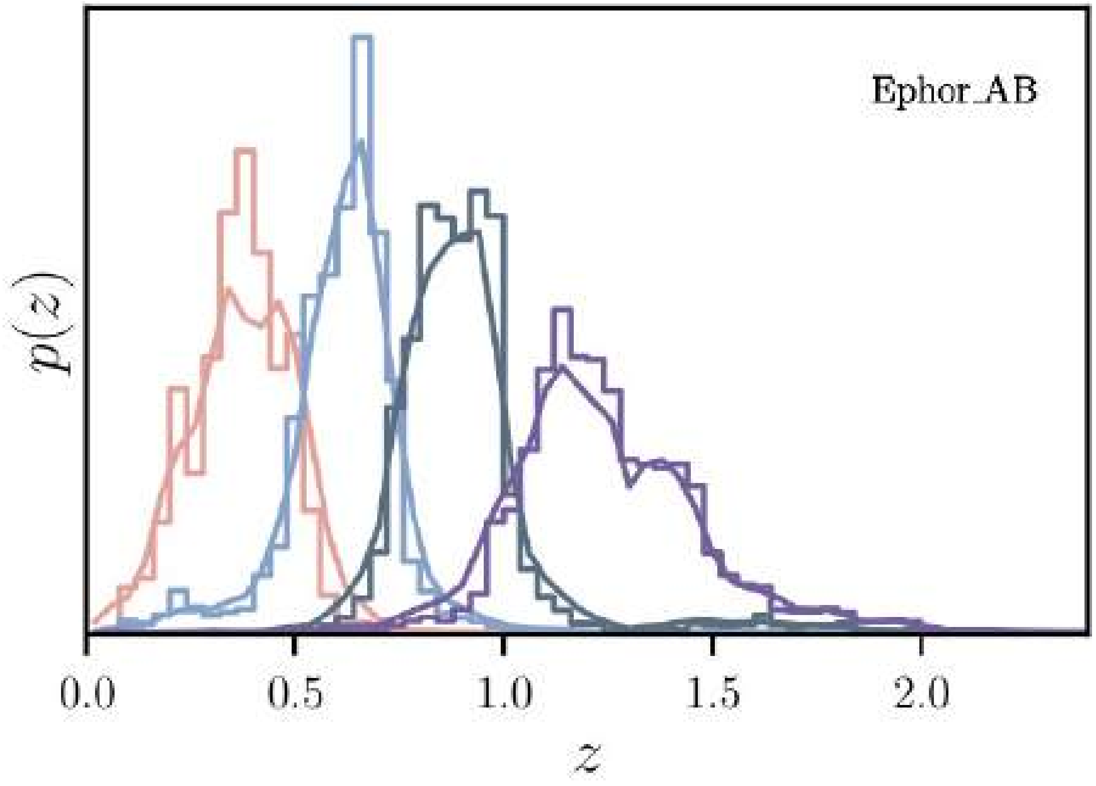
LSST DESC
Dark Energy Science Collaboration

Sample selection

Cut	Comment
detect_is_primary=True	Basic quality cuts, see [35, 43]
icmodel_flags_badcentroid=False	
icentroid_sdss_flags=False	
iflags_pixel_edge=False	
iflags_pixel_interpolated_center=False	
iflags_pixel_saturated_center=False	
iflags_pixel_cr_center=False	
iflags_pixel_bad=False	
iflags_pixel_suspect_center=False	
iflags_pixel_clipped_any=False	
meas.ideblend_skipped=False	
iblendedness_abs_flux$< 10^{-0.375}$	
[g,r,z,y]centroid_sdss_flags=False	Strict photometry cuts
[g,r,i,z,y]cmodel_flux_flags=False	
[g,r,i,z,y]flux_psf_flags=False	
[g,r,z,y]flags_pixel_edge=False	
[g,r,z,y]flags_pixel_interpolated_center=False	
[g,r,z,y]flags_pixel_saturated_center=False	
[g,r,z,y]flags_pixel_cr_center=False	
[g,r,z,y]flags_pixel_bad=False	
icmodel_mag−a_i < 24.5	Magnitude limit
icmodel_flux$>$ 10 icmodel_flux_err	10σ detections
[g,r,y,z]cmodel_flux$>$ 5 [g,r,y,z]cmodel_flux_err	5σ detection (required only in 2 other bands)
iclassification_extendedness=1	Star-galaxy separator

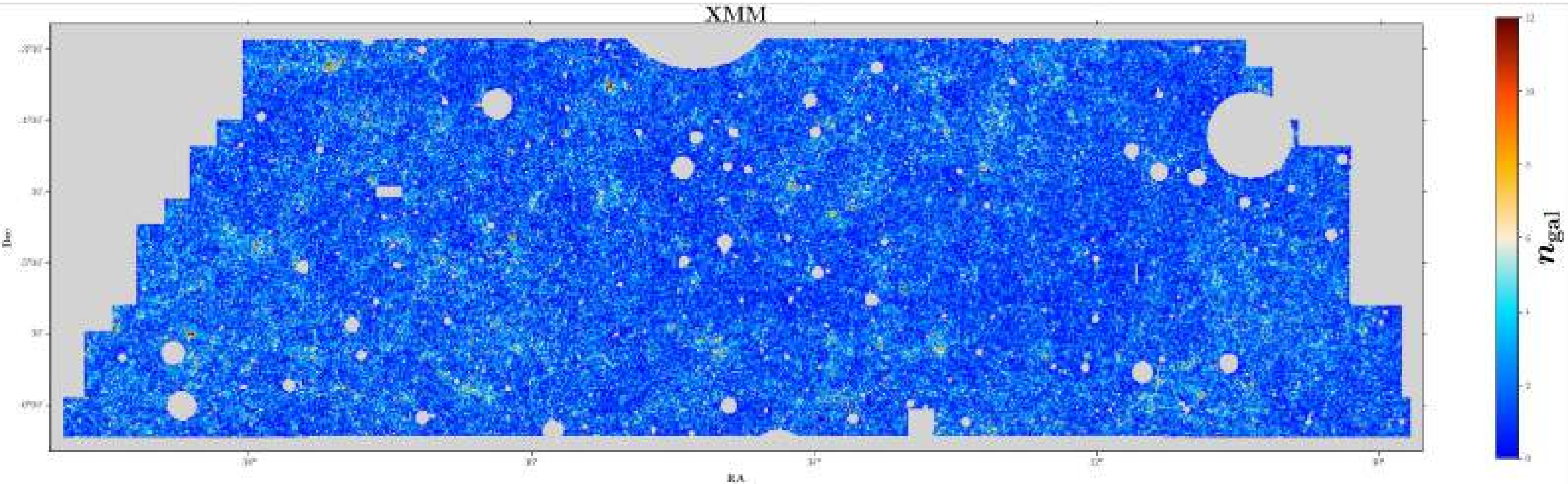
- Similar to HSC shear sample (no shape cuts)
- Bright magnitude cut (5σ limit is ~26)
- This improves sample homogeneity

Sample selection



- Similar to HSC shear sample (no shape cuts)
- Bright magnitude cut (5σ limit is ~26)
- This improves sample homogeneity
- Split into 4 redshift bins
- Photo-z posteriors from several codes

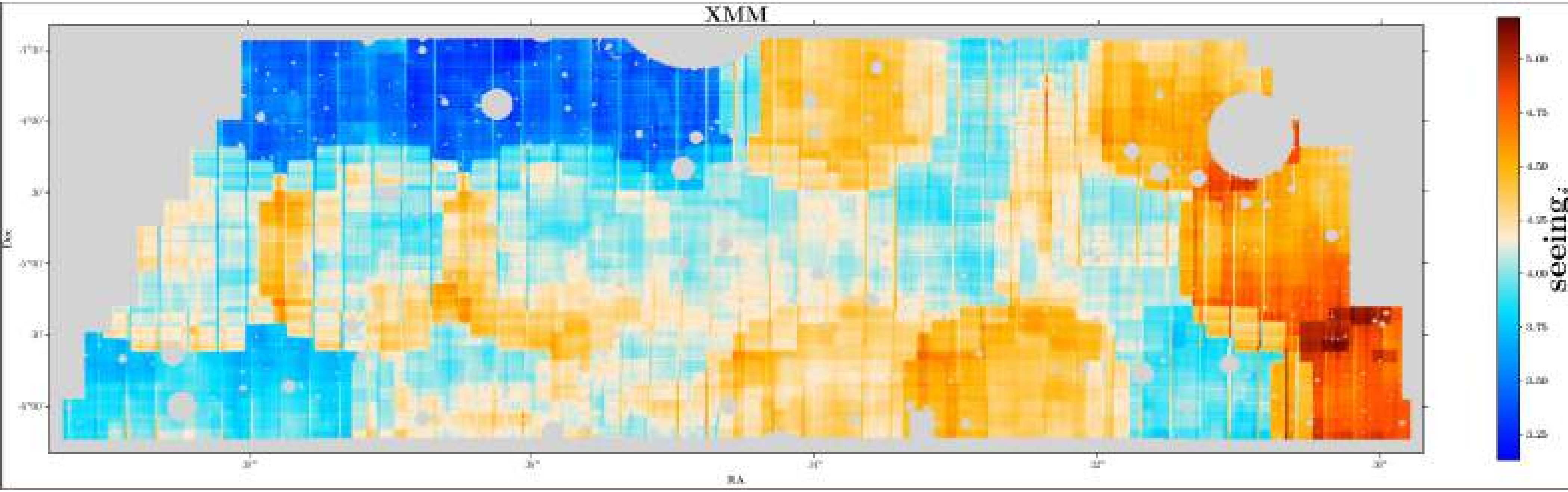
Maps and masks



- We pixelise each field using square **pixels 0.6’** in size.
- **Plate-Carrée** projection.
- Small fields (<20 sq-deg) → **flat-sky** approximation.
- Mask: footprint + bright-object mask + depth mask.
- Overdensity maps for each field and redshift bin.

$$\delta_p = \frac{n_p}{\bar{n} w_p} - 1 \qquad \bar{n} = \frac{\langle n_p w_p \rangle}{\langle w_p \rangle}$$

Contaminant maps



- We generate sky maps for all quantities that could potentially cause systematic fluctuations in δ_g.
- Observing conditions are mapped in all bands.
- 47 maps in total (per field).
- We deproject all of these in all power spectra.

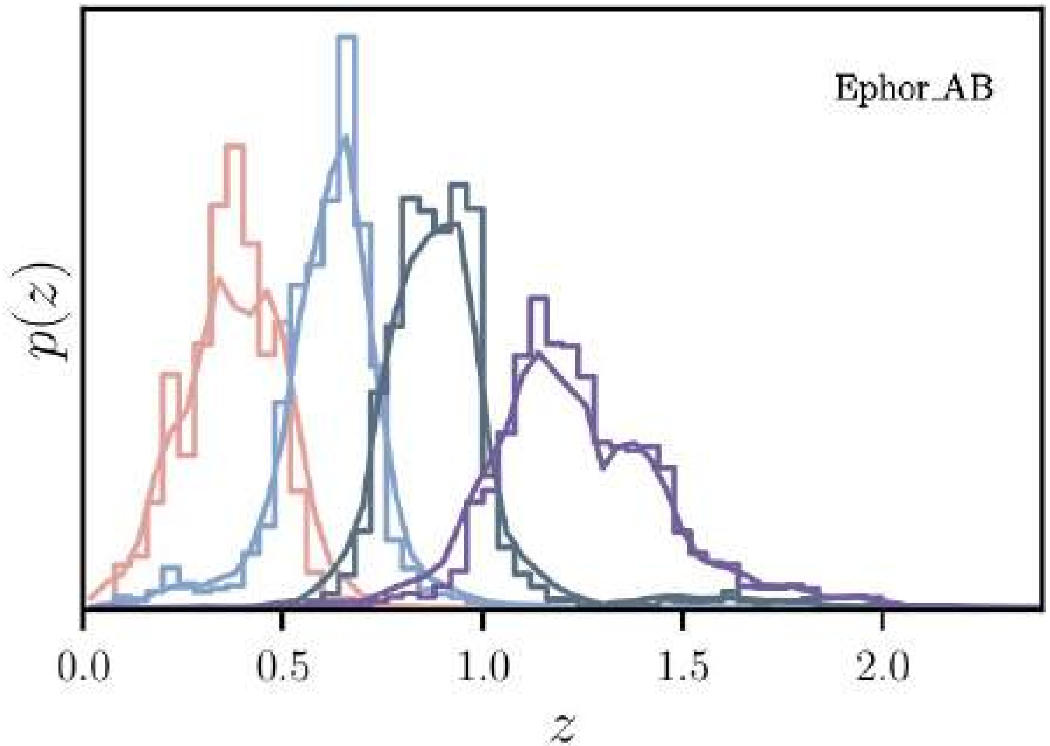
Stars	Airmass	PSF ellipticity	Sky-sigma
Dust	CCD temperature	Exposure time	# visits
10σ depth	Seeing FWHM	Sky level	

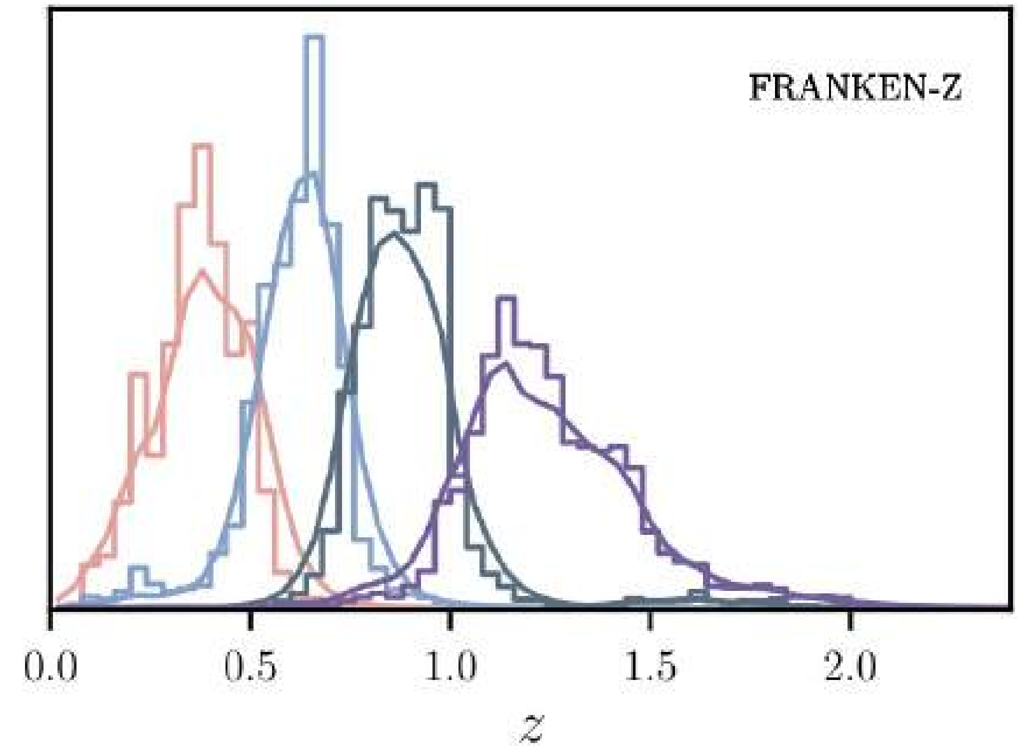
Redshift distributions

- The redshift distributions are a central part of the theory prediction

$$C_\ell^{ab} = \int d\chi \, \frac{H^2}{\chi^2} \, p_a(z) \, p_b(z) \, P_{ab}\left(k = \ell + \frac{1}{2}, z\right)$$

- Our fiducial distributions are determined from the COSMOS 30-band photometric catalog.
- COSMOS objects are reweighted in color space to match our sample.
- We obtain alternative estimates of the p(z)s by stacking the pdfs of all objects in each bin for 4 different photo-z codes.

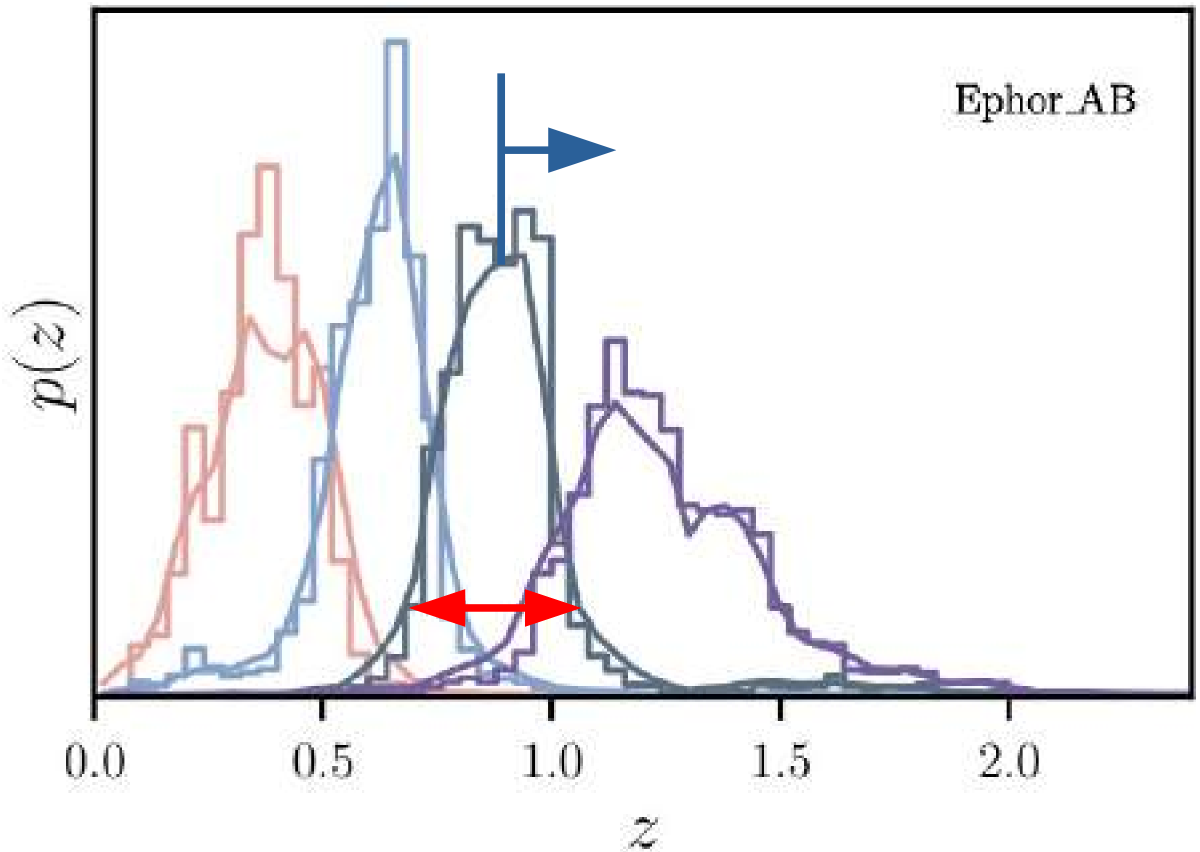




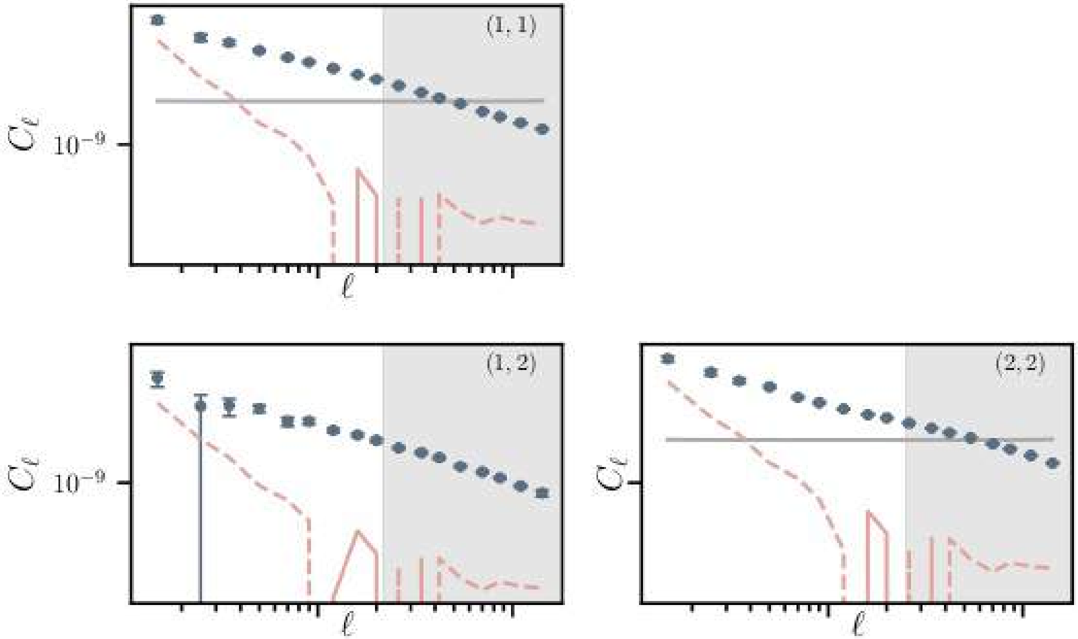
Redshift distributions

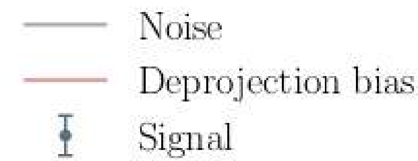
- None of these estimates are exact: we need to marginalise over residual uncertainties.
- N(z) uncertainties have traditionally been summarized by a single **“shift” parameter** in shear analyses.
- Clustering is potentially more sensitive to this, so we extend this by adding a **“width” parameter**.
- We vary these within broad priors.

$$p_i(z) = \hat{p}_i(z_c + (1 + z_{w,i})(z - z_c) + \Delta z_i)$$

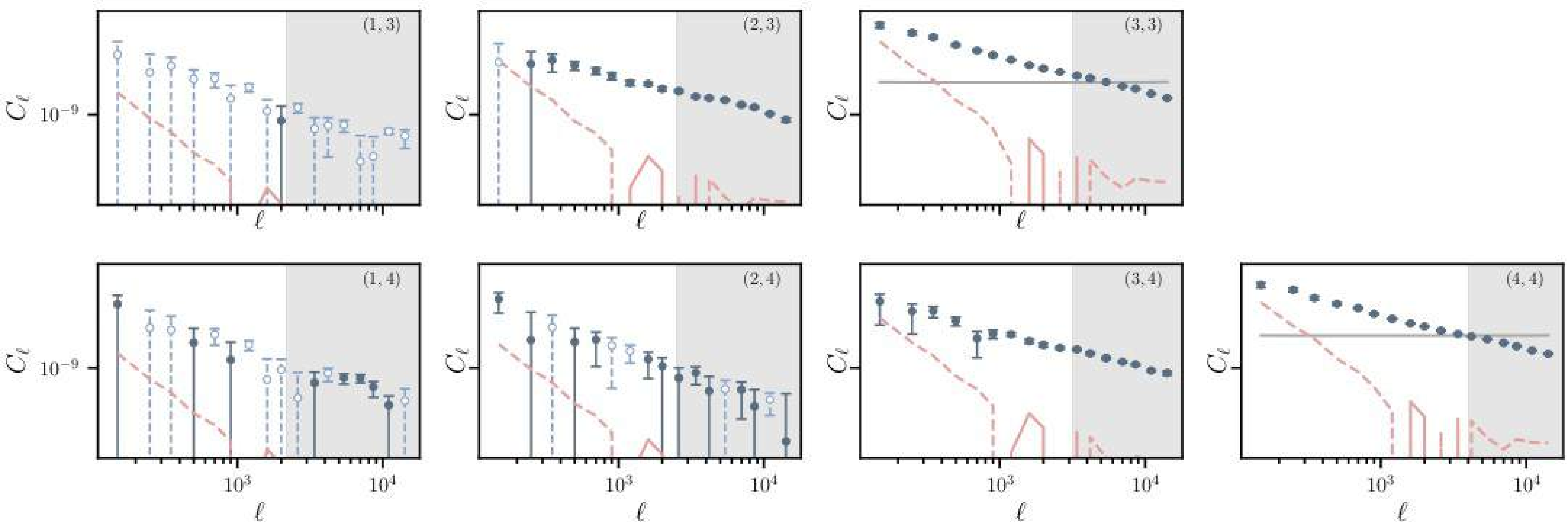


Power spectra





- Power spectra estimated with NaMaster
- 47 templates deprojected.
- Analytic shot-noise correction.



Covariance matrices

We use an **analytical covariance** matrix that includes:

- Mode coupling in the Gaussian part due to **survey geometry** (computed with NaMaster).

$$\mathrm{Cov}\left(\tilde{C}_\ell^{ab}, \tilde{C}_{\ell'}^{cd}\right) = (2\ell'+1)^{-1}\left[C_{(\ell}^{ac}C_{\ell')}^{bd}M_{\ell\ell'}(w_a w_c, w_b w_d) + C_{(\ell}^{ad}C_{\ell')}^{bc}M_{\ell\ell'}(w_a w_d, w_b w_c)\right]$$

- Mode-coupling due to **non-linear** growth using a perturbation theory + halo model approach.

$$\mathrm{Cov}_{\mathrm{NG}}(C_\ell^{ab}, C_{\ell'}^{cd}) = \frac{1}{4\pi f_{\mathrm{sky}}}\int_{|\boldsymbol{\ell}|\in\ell_1}\int_{|\boldsymbol{\ell}'|\in\ell_2}\int \frac{\mathrm{d}^2\boldsymbol{\ell}}{A(\ell_1)}\frac{\mathrm{d}^2\boldsymbol{\ell}'}{A(\ell_2)}\,\mathrm{d}\chi\,\frac{q^a(\chi)q^b(\chi)q^c(\chi)q^d(\chi)}{\chi^6}\times T^{abcd}(\boldsymbol{\ell}/\chi, -\boldsymbol{\ell}/\chi, \boldsymbol{\ell}'/\chi, -\boldsymbol{\ell}'/\chi).$$

$$T^{abcd} = T^{abcd,1h} + (T_{22}^{abcd,2h} + T_{13}^{abcd,2h}) + T^{abcd,3h} + T^{abcd,4h}$$

- Mode-coupling due to **super-survey** modes.

$$\mathrm{Cov}_{\mathrm{SSC}}(C_\ell^{ab}, C_{\ell'}^{cd}) = \int \mathrm{d}\chi\,\frac{q^a(\chi)q^b(\chi)q^c(\chi)q^d(\chi)}{\chi^4}\times \frac{\partial P_{ab}(\ell/\chi, z(\chi))}{\partial\delta_{\mathrm{LS}}}\frac{\partial P_{cd}(\ell'/\chi, z(\chi))}{\partial\delta_{\mathrm{LS}}}\sigma_b^2(z(\chi))$$

- Covariances are estimated in each field and then coadded.

Covariance matrices

Covariances are model dependent!

We use a four-step process:

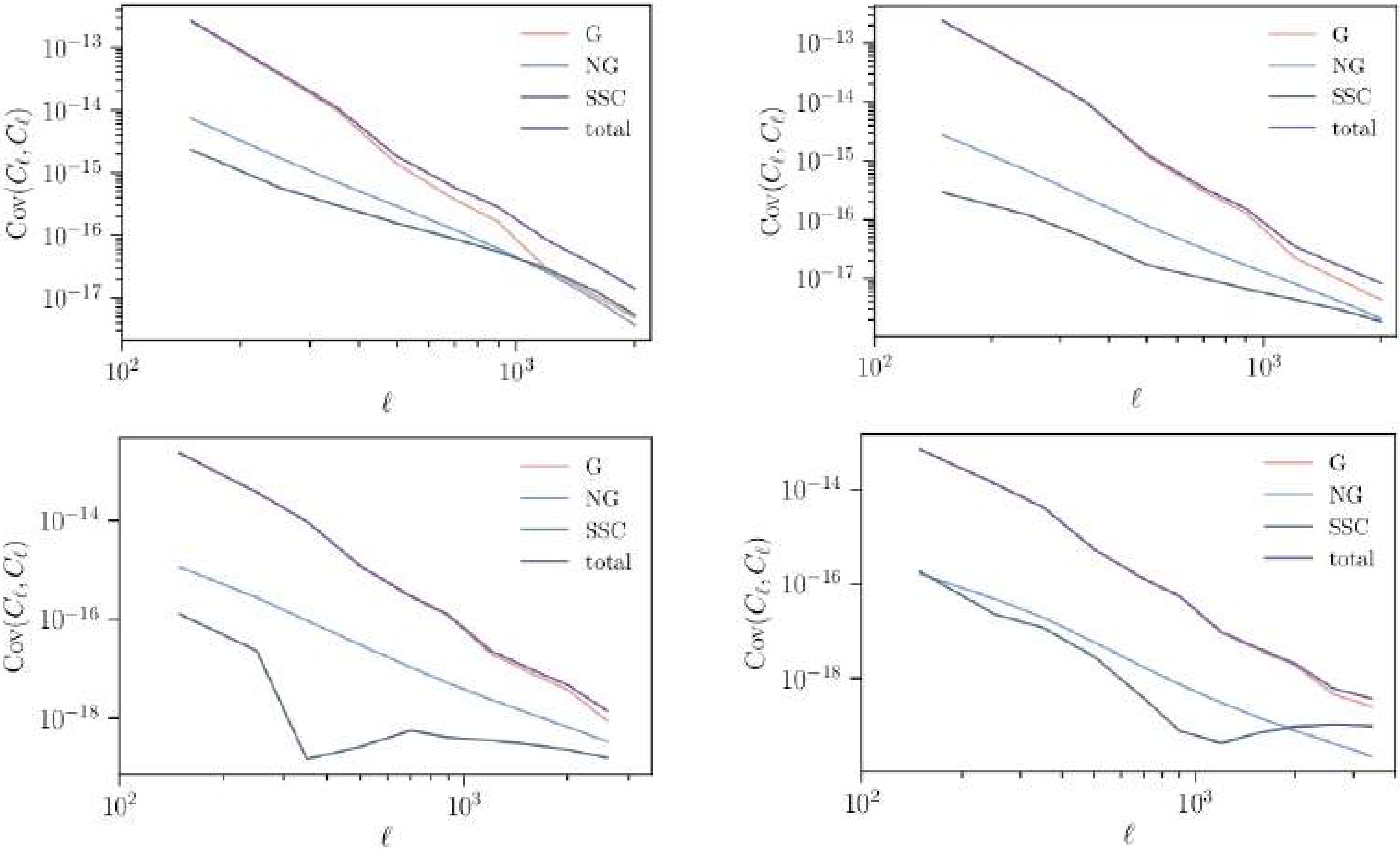
1. Gaussian covariance from measured power spectra.

2. Obtain best-fit parameters and compute corresponding covariance.

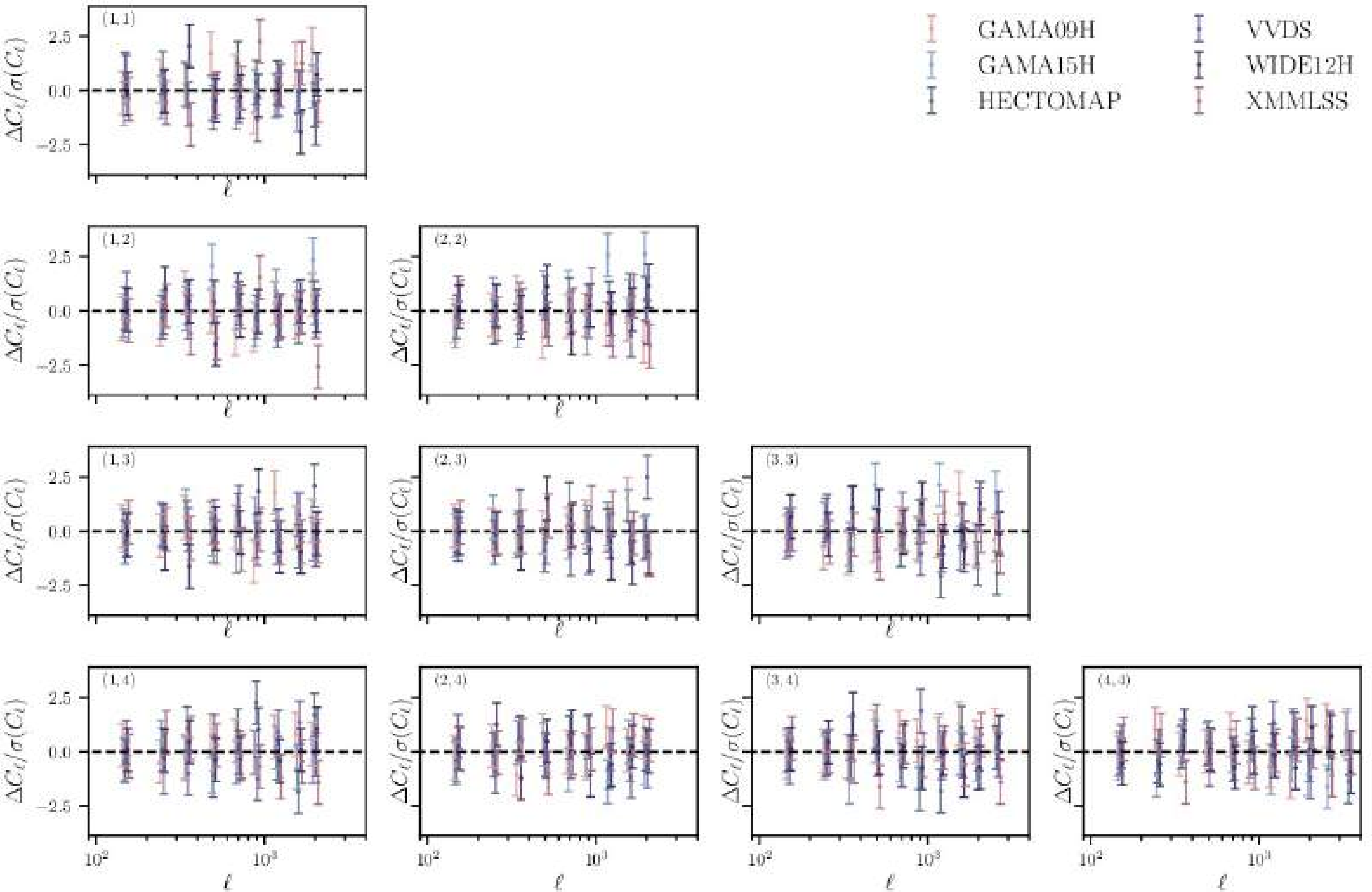
3. Run chains with this covariance.

4. If new best-fit is too far from the previous one, GOTO 2.

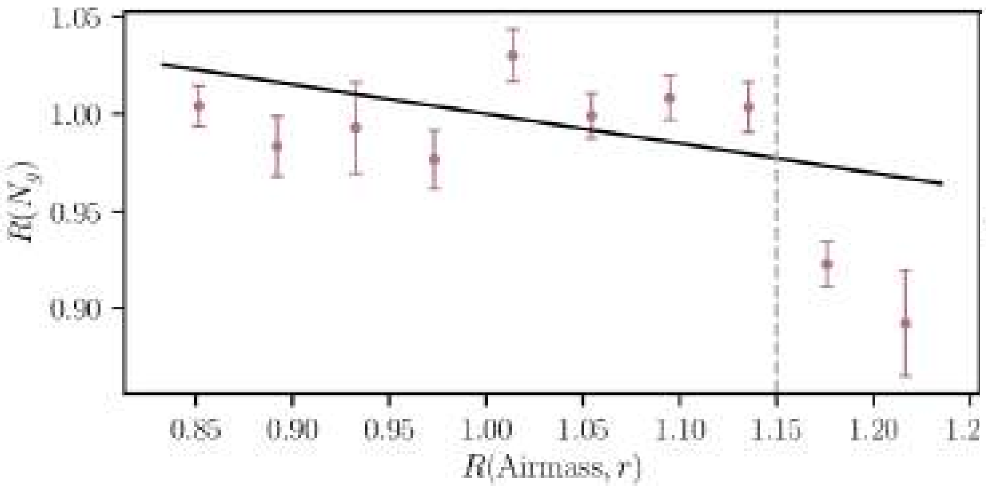
Covariance matrices

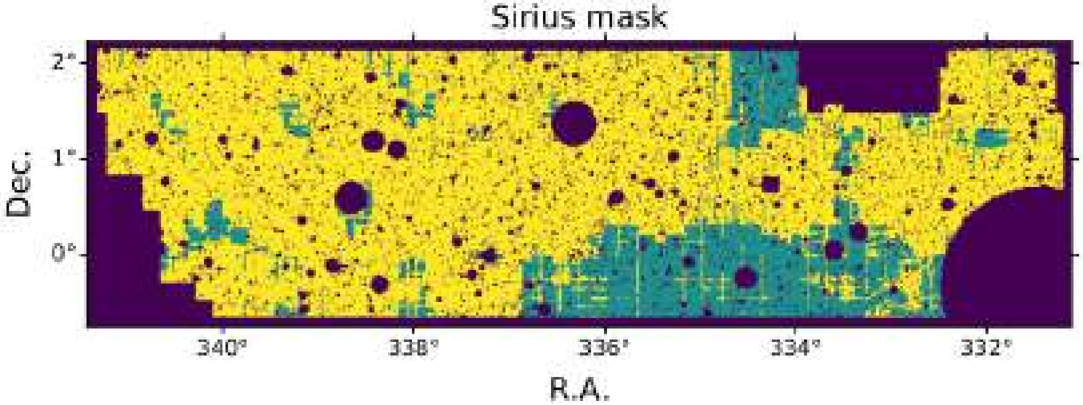


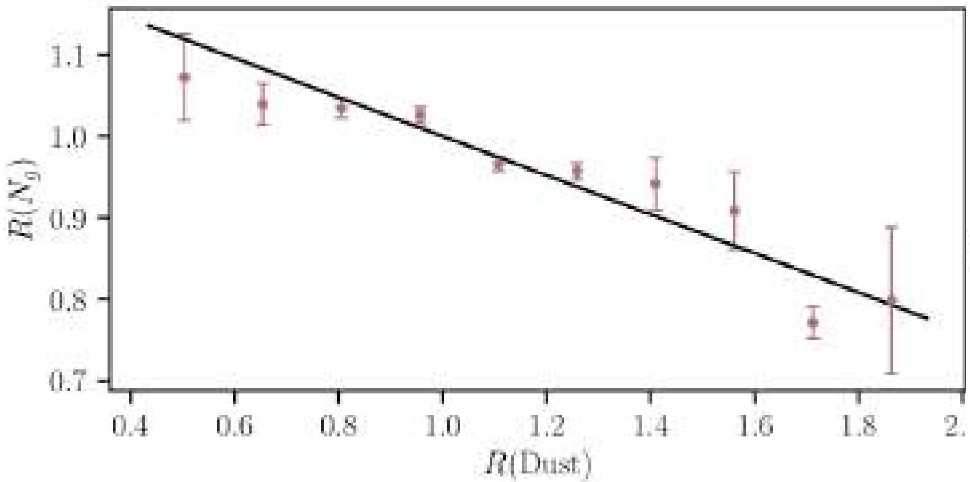
Robustness tests: consistency across fields

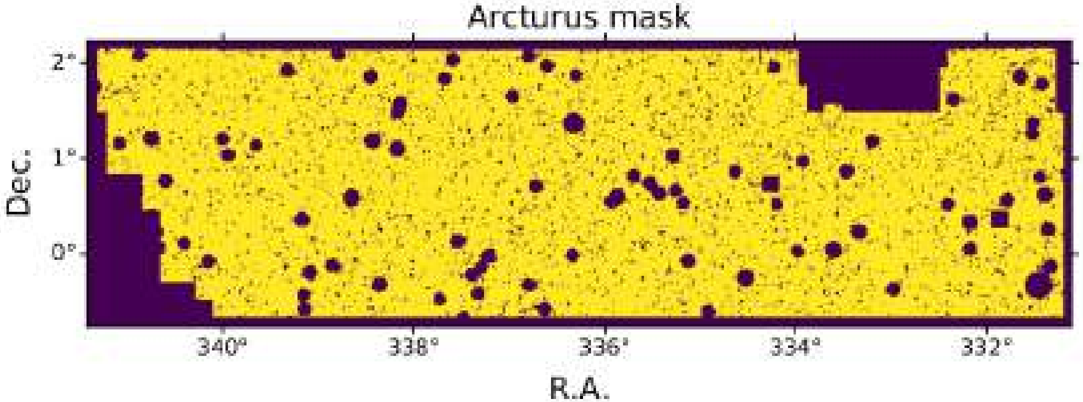


Robustness tests: masks and contaminants

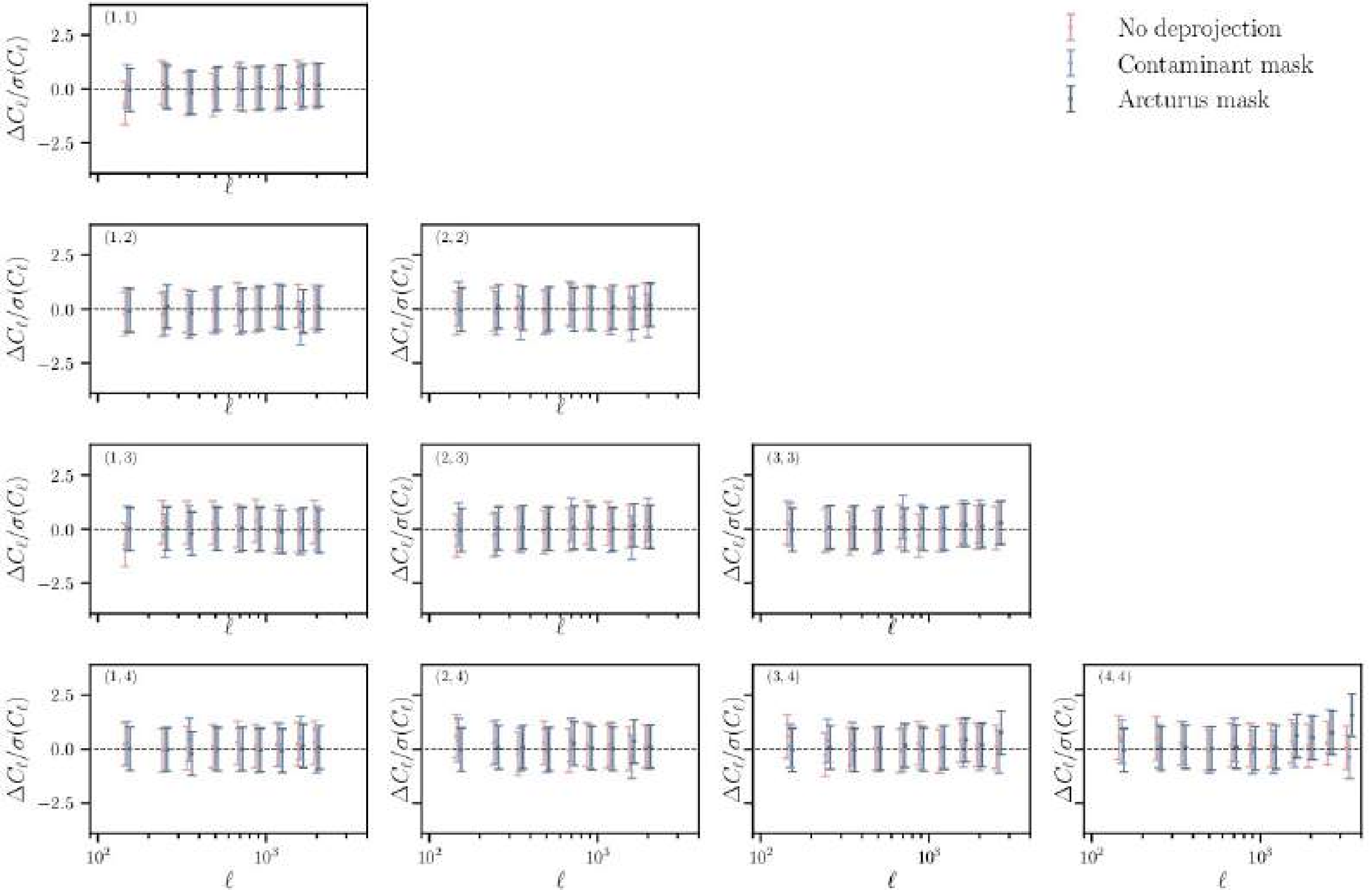






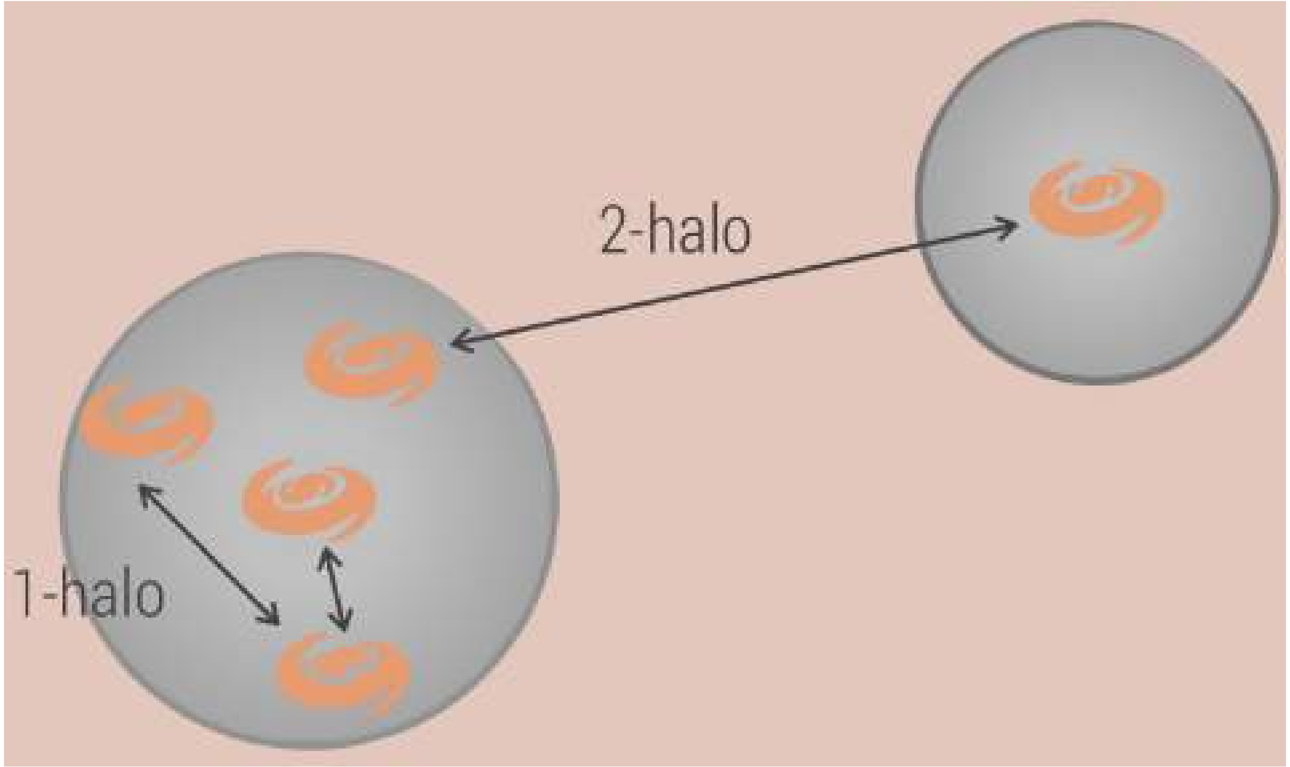


Robustness tests



HOD parameterisation

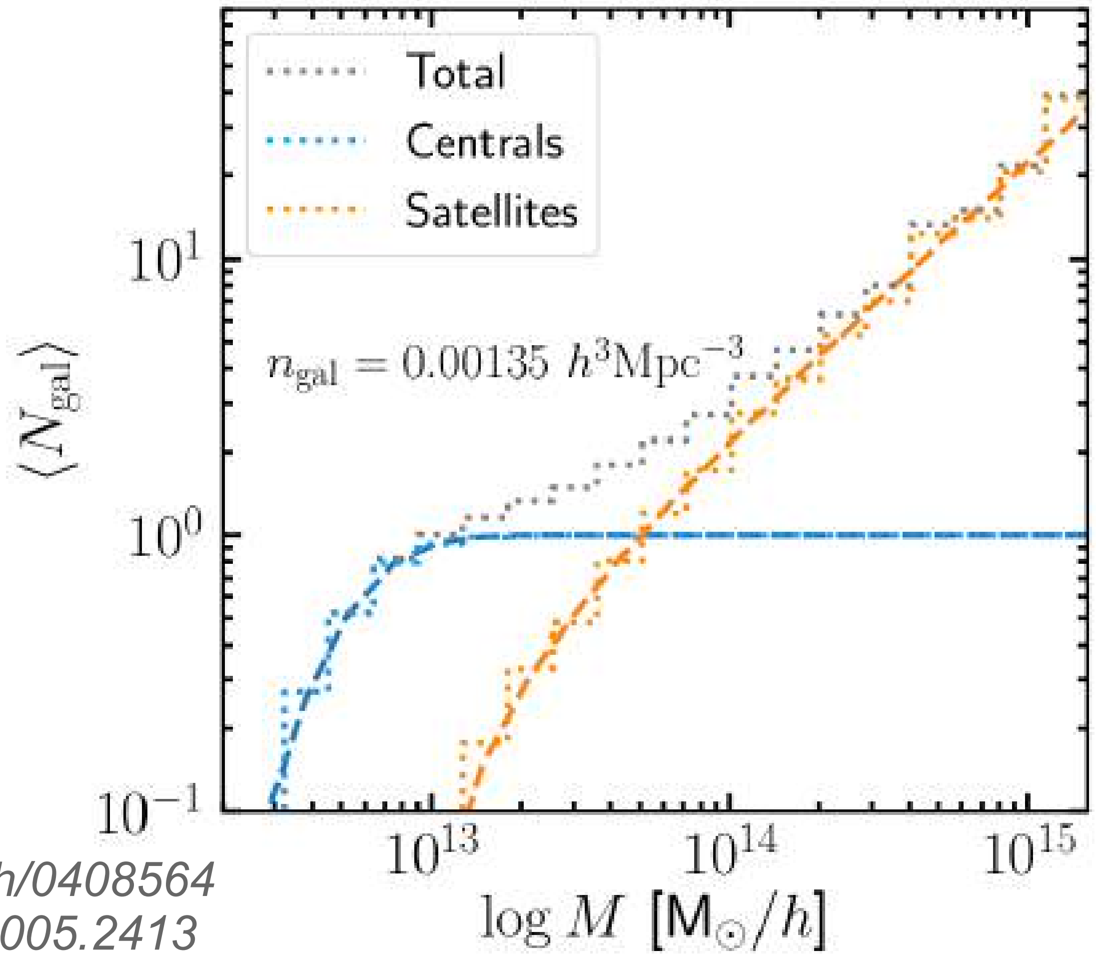
$$P_{gg}(z,k) = P_{gg,1\mathrm{h}}(z,k) + P_{gg,2\mathrm{h}}(z,k)$$



$$P_{gg,1\mathrm{h}}(k) = \frac{1}{\bar{n}_g^2} \int \mathrm{d}M \, \frac{\mathrm{d}n}{\mathrm{d}M} \bar{N}_c \, \left[\bar{N}_s^2 u_s^2(k) + 2\bar{N}_s u_s^2(k)\right]$$

$$P_{gg,2\mathrm{h}}(k) = \left(\frac{1}{\bar{n}_g} \int \mathrm{d}M \, \frac{\mathrm{d}n}{\mathrm{d}M} \, b_h(M) \, \bar{N}_c \, \left[1 + \bar{N}_s u_s(k)\right]\right)^2 P_{\mathrm{lin}}(k)$$

HOD parameterisation



Hadzhyiska et al. arXiv:1911.02610

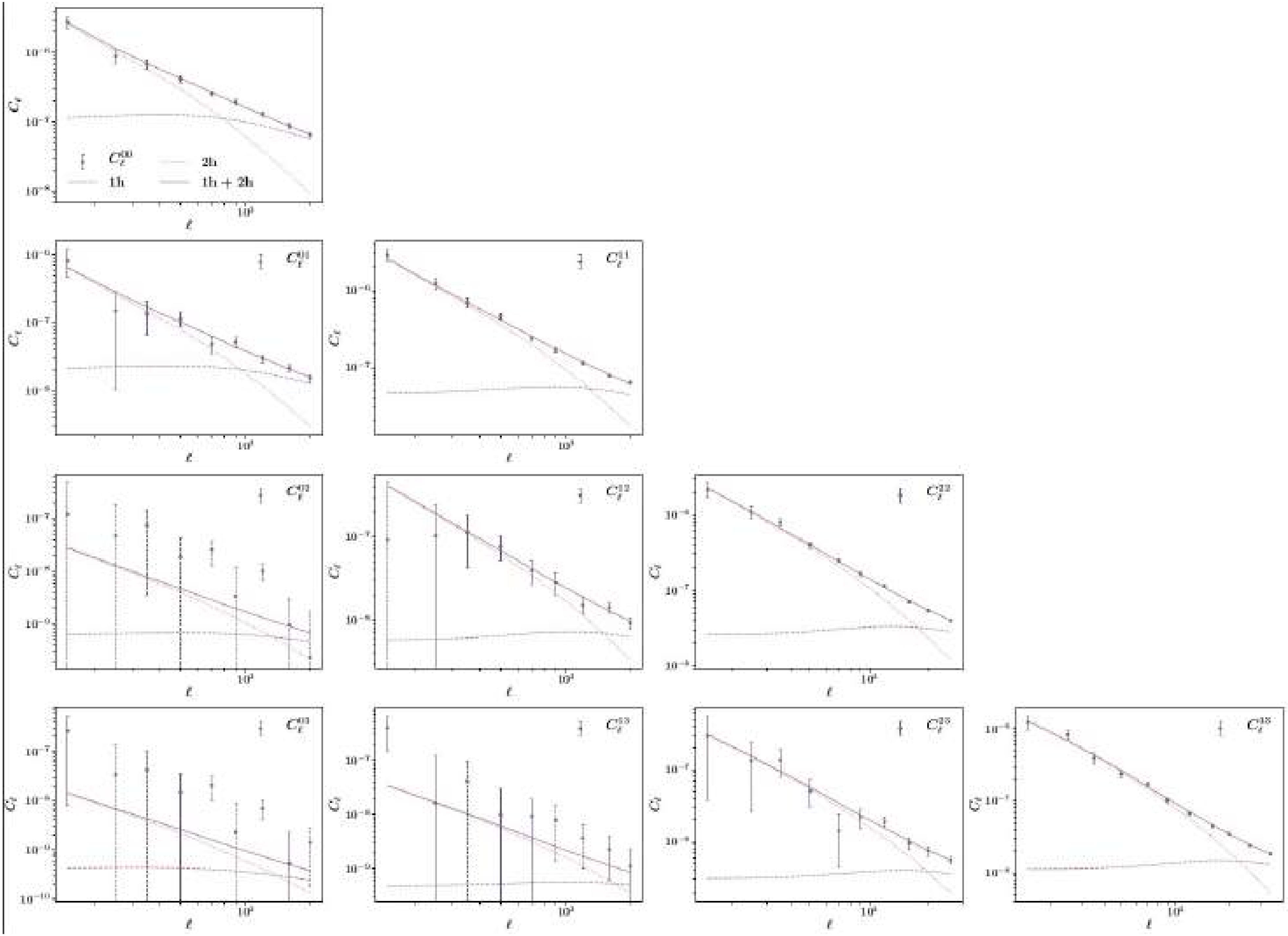
Zheng et al. astro-ph/0408564
Zehavi et al. arXiv:1005.2413

$$\bar{N}_c(M) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\log(M/M_{\rm min})}{\sigma_{\log M}}\right)\right]$$

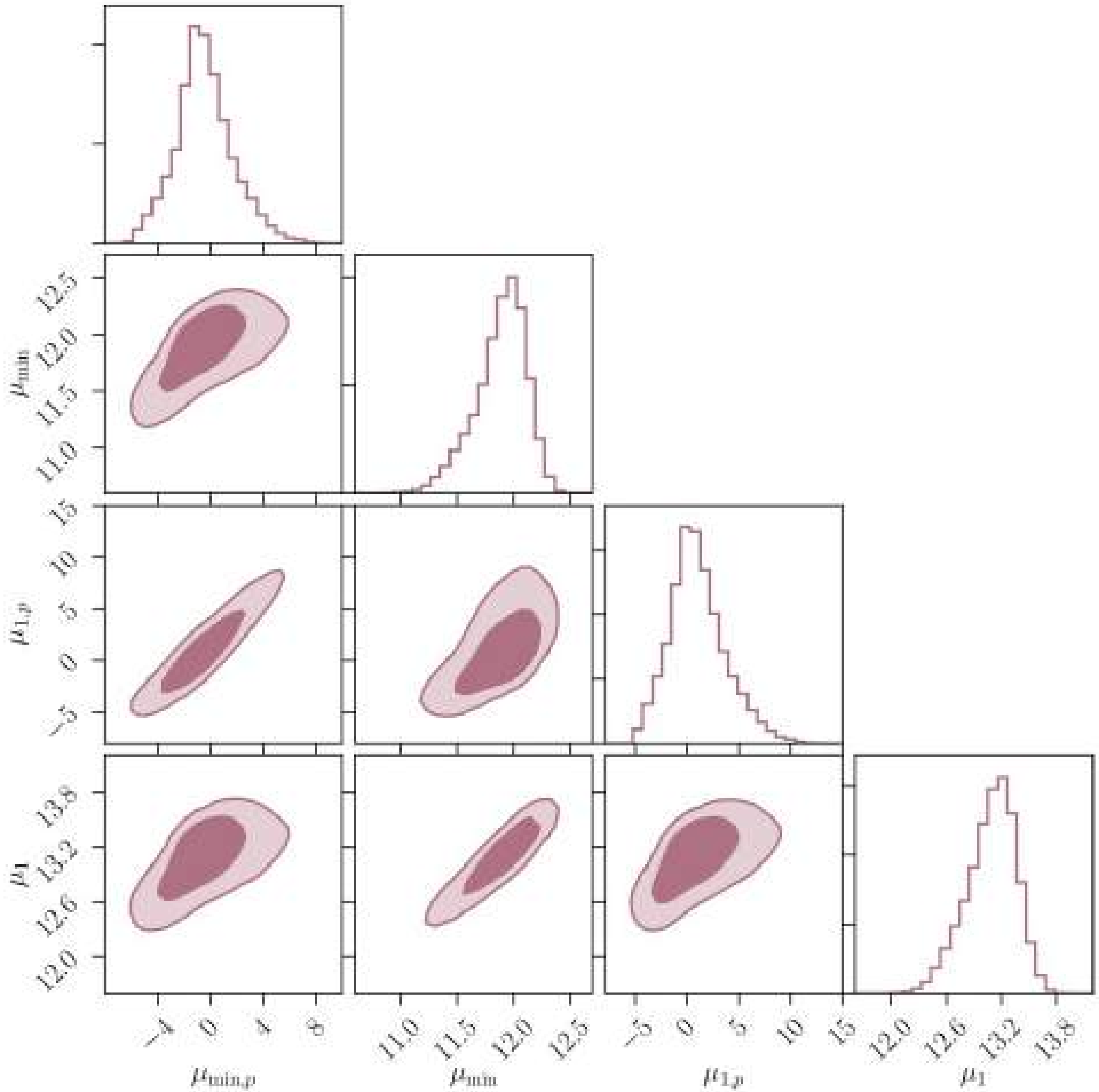
$$\bar{N}_s(M) = \Theta(M - M_0)\left(\frac{M - M_0}{M_1'}\right)^\alpha,$$

$$\log M_i(z) = \mu_i + \mu_{i,p}\left(\frac{1}{1+z} - \frac{1}{1+z_p}\right)$$

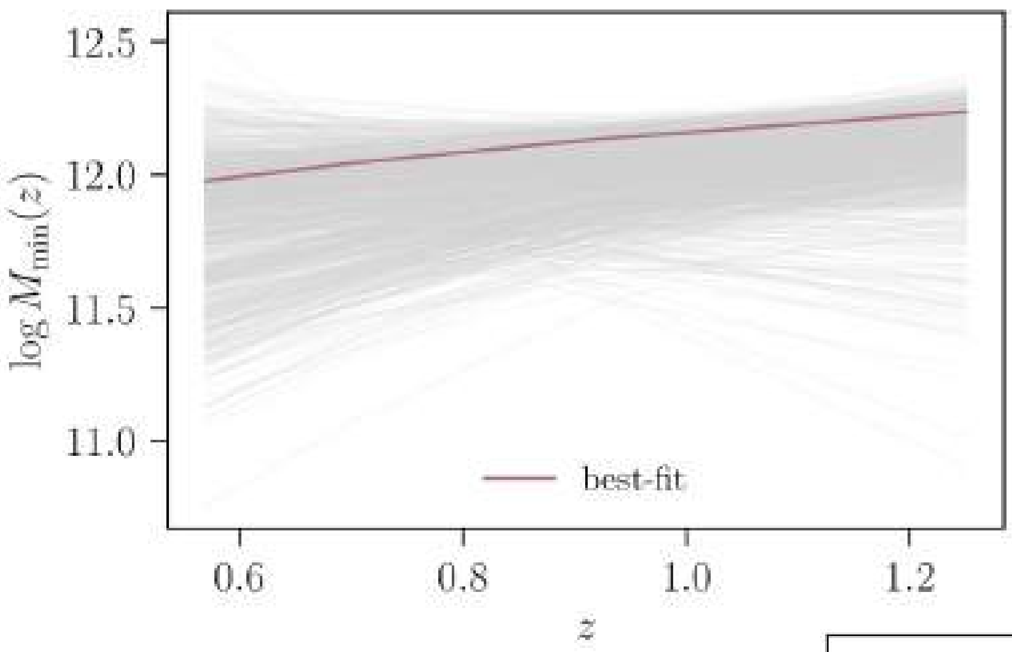
Results: HOD constraints

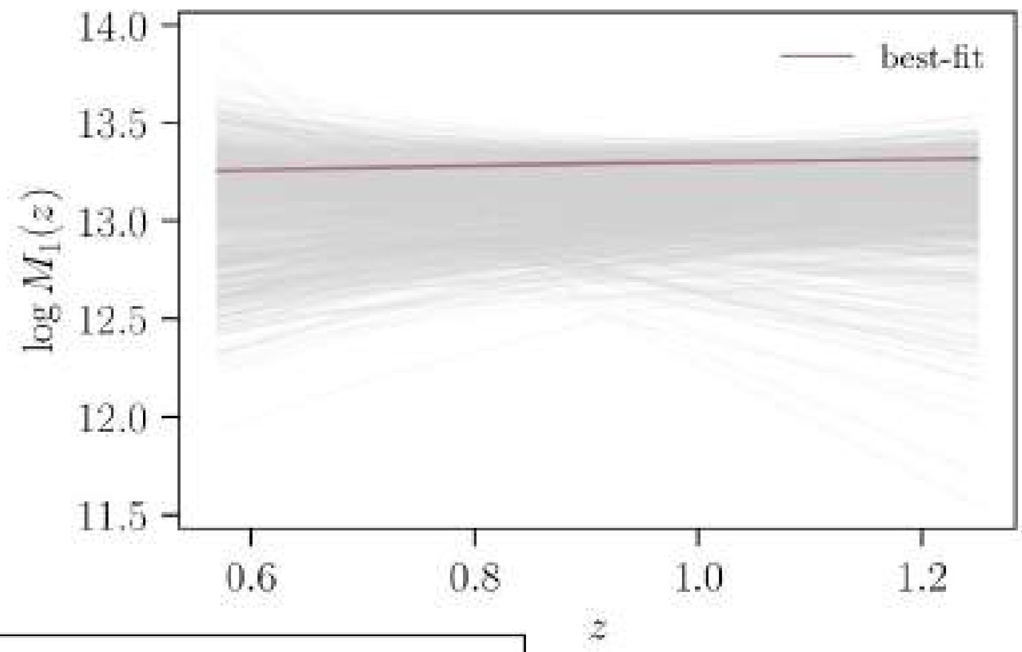


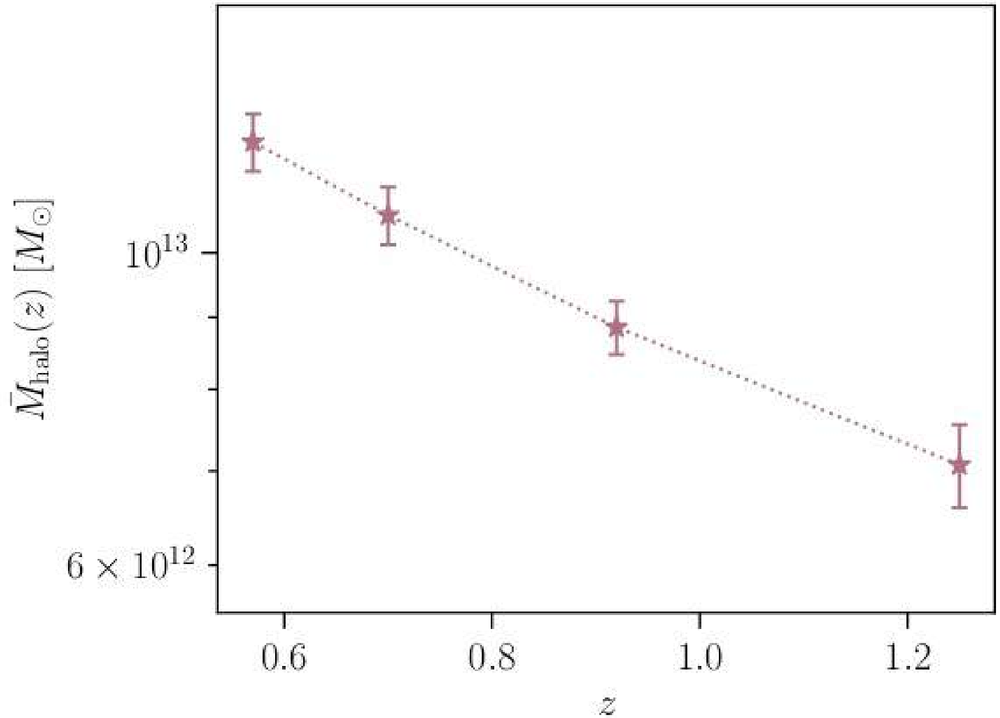
Results: HOD constraints



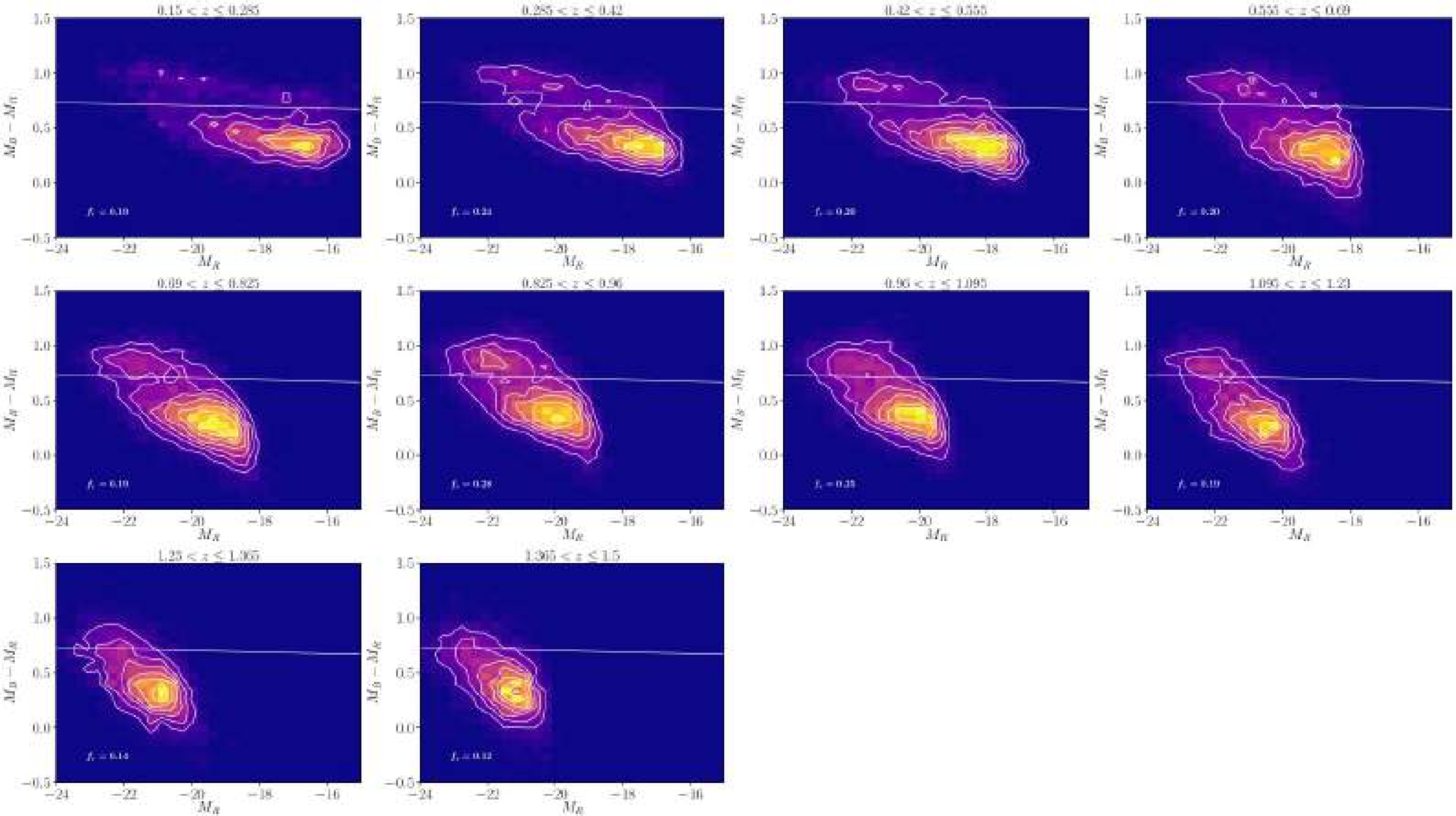
Results: HOD constraints



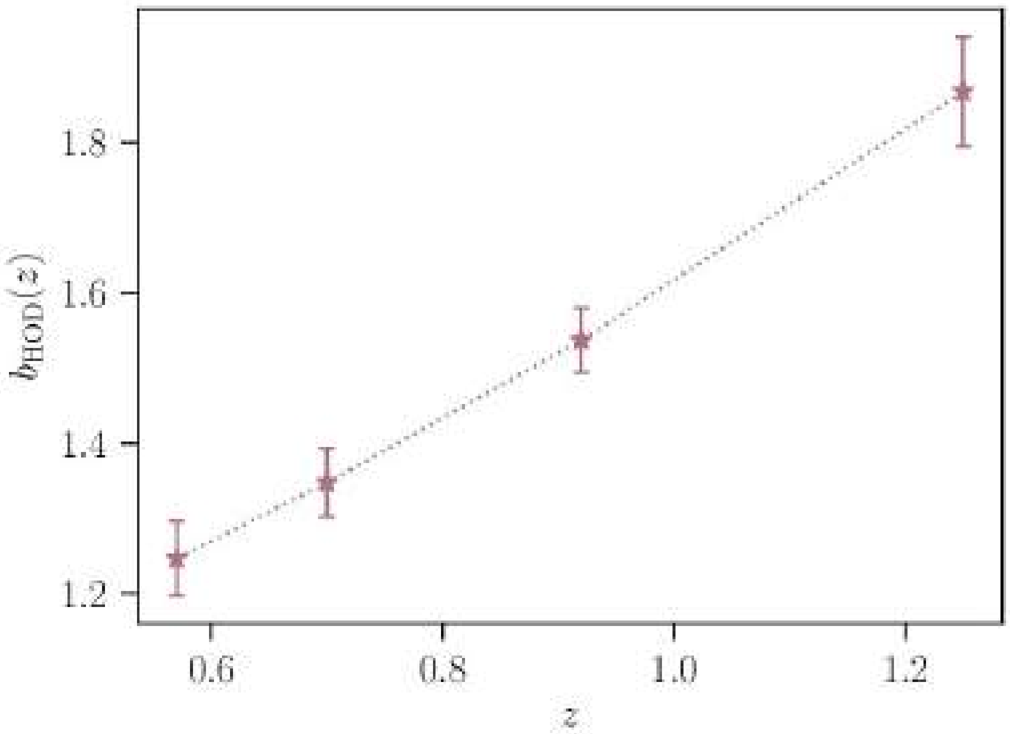


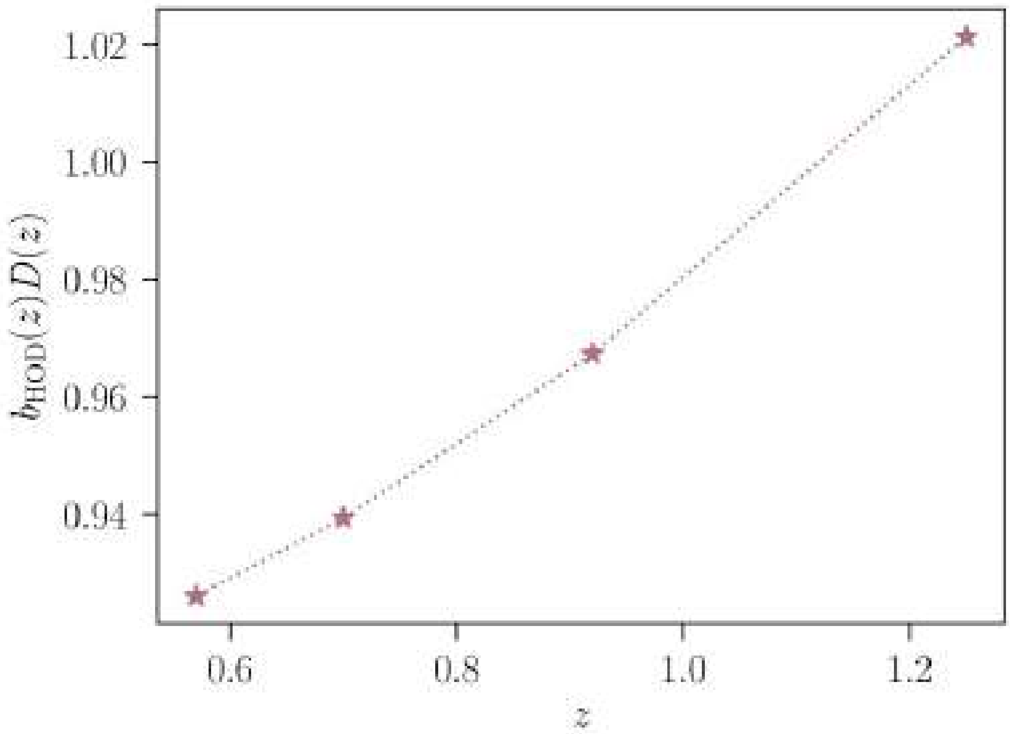


Results: HOD constraints

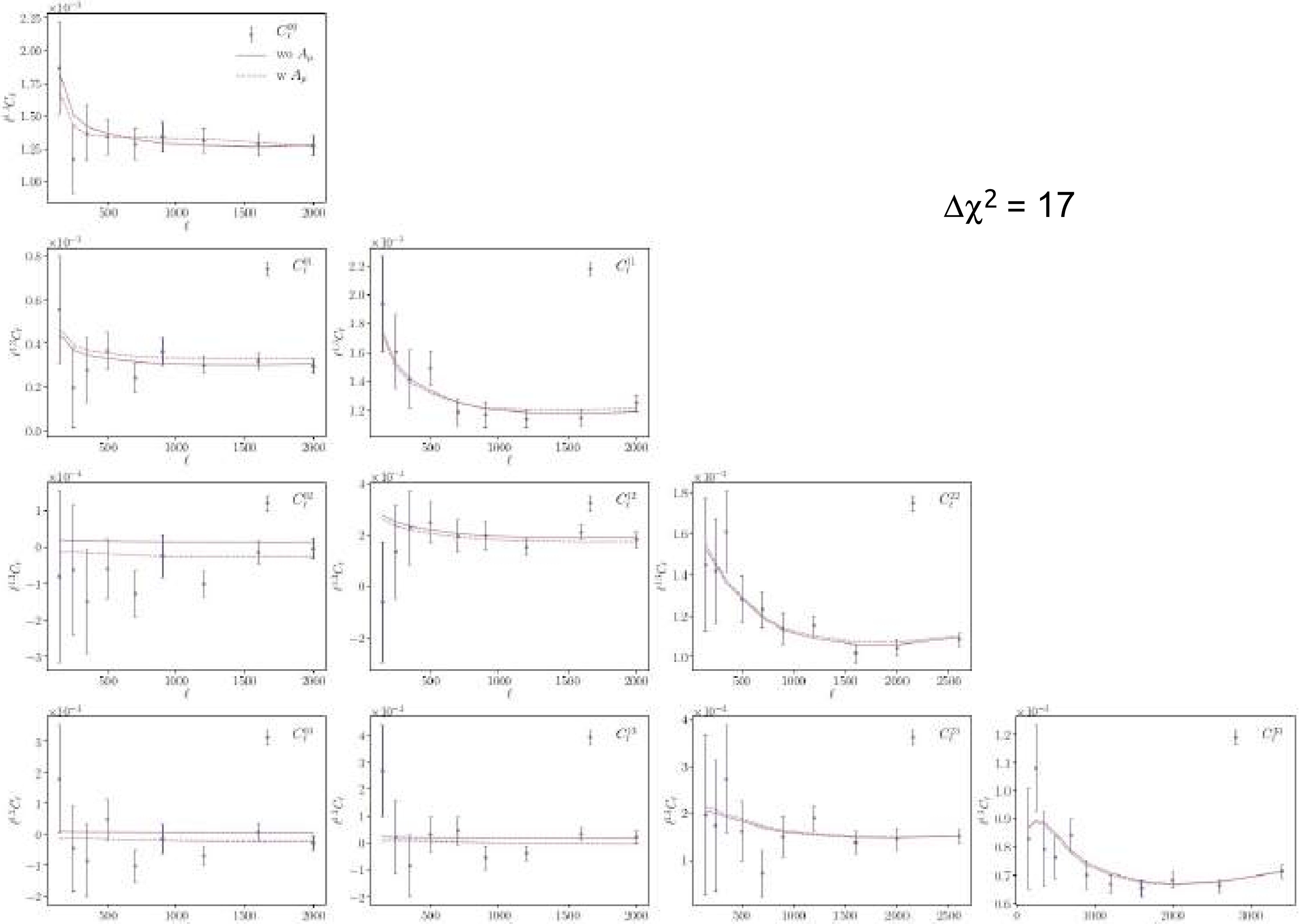


Results: HOD constraints



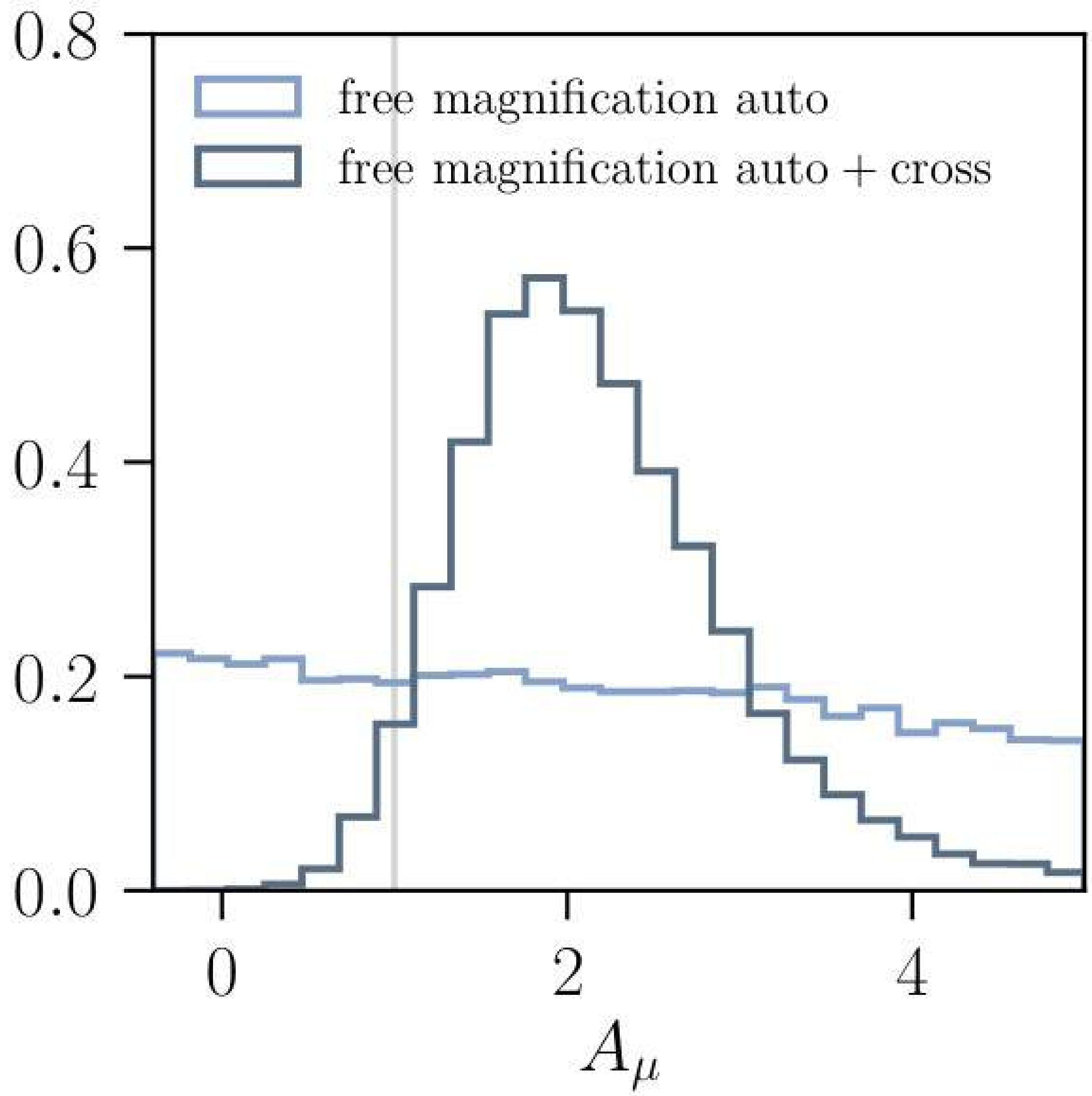


Results: Magnification



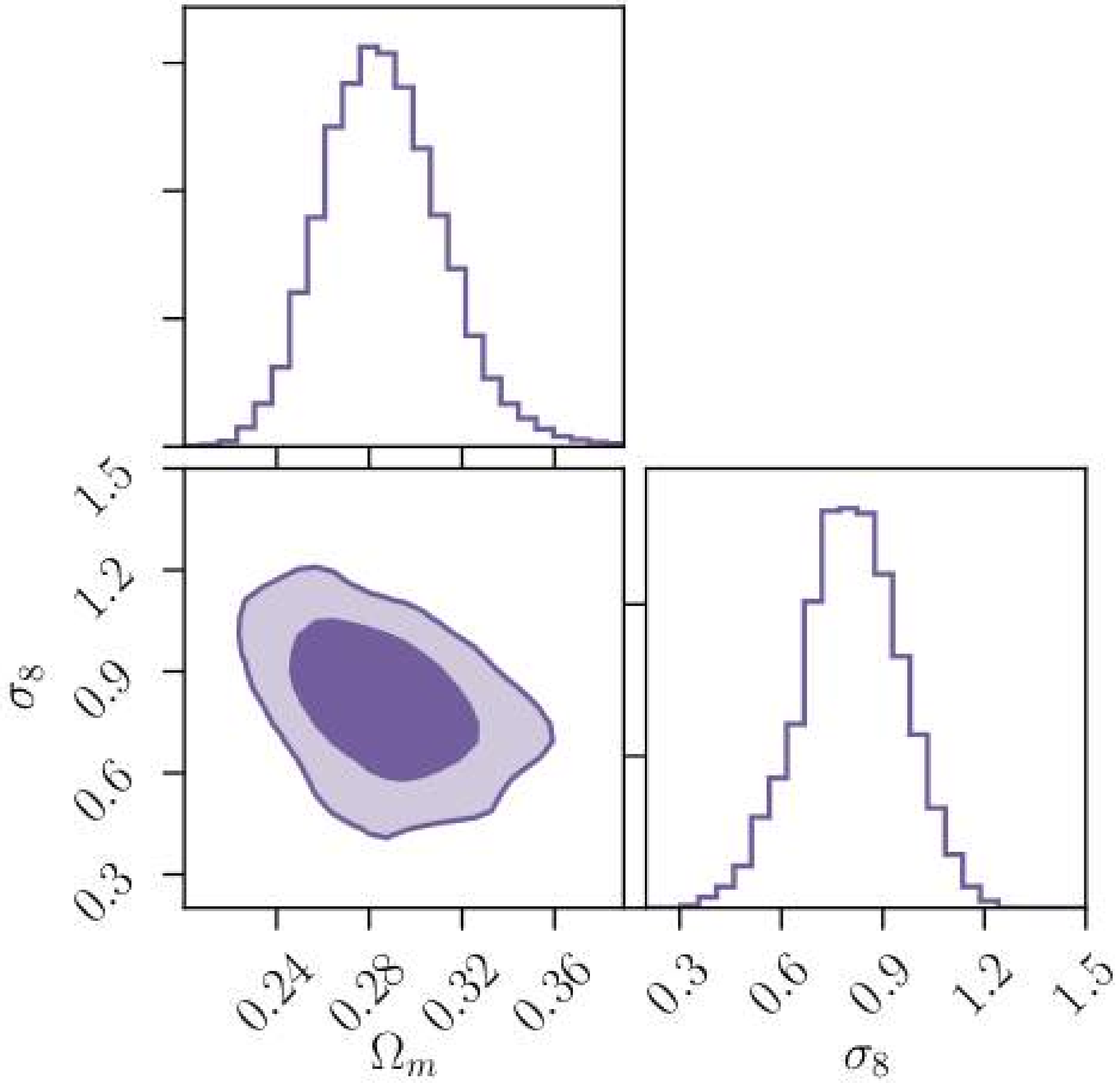
$\Delta\chi^2 = 17$

Results: Magnification

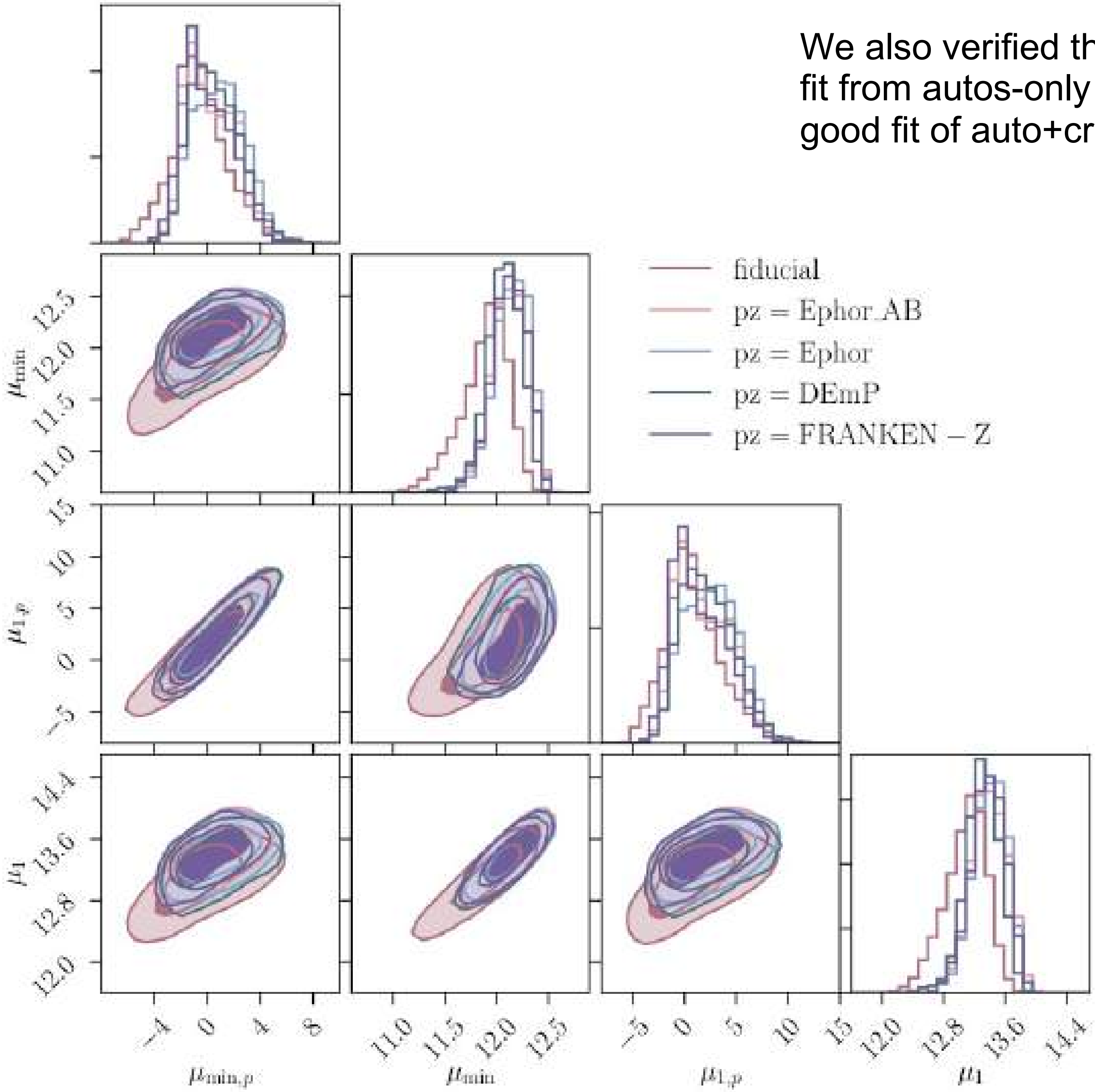


~3σ detection

Results: Cosmology

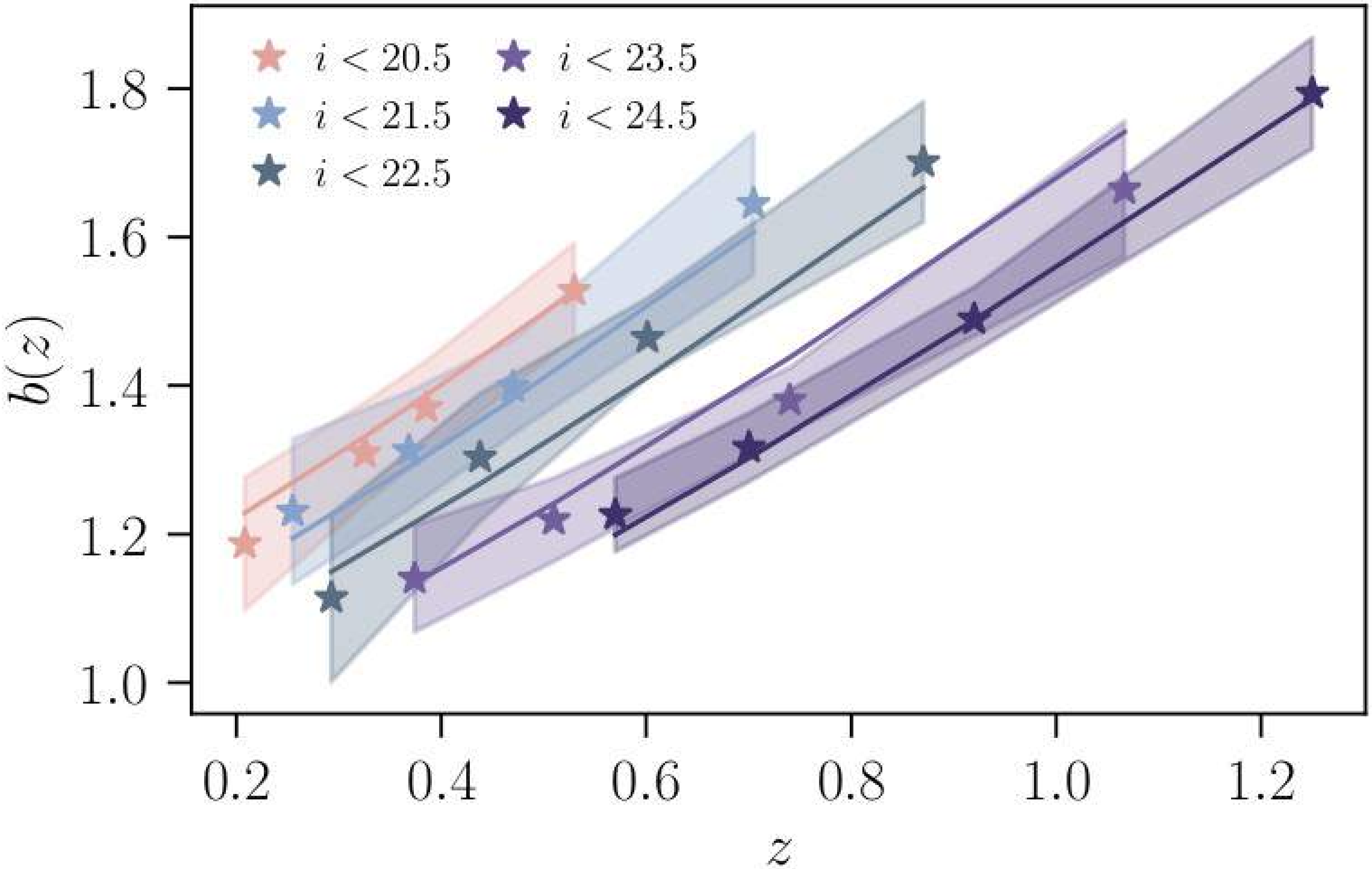


Results: photo-z systematics



We also verified that the best-fit from autos-only is also a good fit of auto+cross.

Results: galaxy bias



$$b(z, m_{\mathrm{lim}}) = \bar{b}(m_{\mathrm{lim}}) D(z)^{\alpha}$$

$$\alpha = -1.30 \pm 0.19$$

$$\bar{b}(m_{\mathrm{lim}}) = b_1(m_{\mathrm{lim}} - 24) + b_0$$

$$b_0 = 0.8346 \pm 0.161$$

$$b_1 = -0.0624 \pm 0.0070$$

Summary

- **Tools for cosmological analyses:**
  - Power spectra and Gaussian covariances: NaMaster *(arXiv:1809.09603)*.
  - Accurate theory predictions: CCL *(arXiv:1812.05995)*.

- **Data compression for weak lensing** *(arXiv:1903.04957)*:
  - KL transform == P(k) analysis for weak lensing.
  - Information contained in a small number of modes N.
  - N=1-2 for current data.

- **Parameter-dependent covariances** *(arXiv:1811.11584)*:
  - Don't worry about this.

- **Galaxy clustering in HSC DR1** *(coming soon)*:
  - Fourier-space analysis with comprehensive systematic deprojection.
  - Mild redshift dependence of HOD due to red galaxy dropout.
  - Simple prescription for $b(z, m_{lim})$
  - Additional sensitivity to photo-z uncertainties (challenge and opportunity)
  - 3σ detection of cosmic magnification
  - Consistent cosmological constraints.

Merci beaucoup!