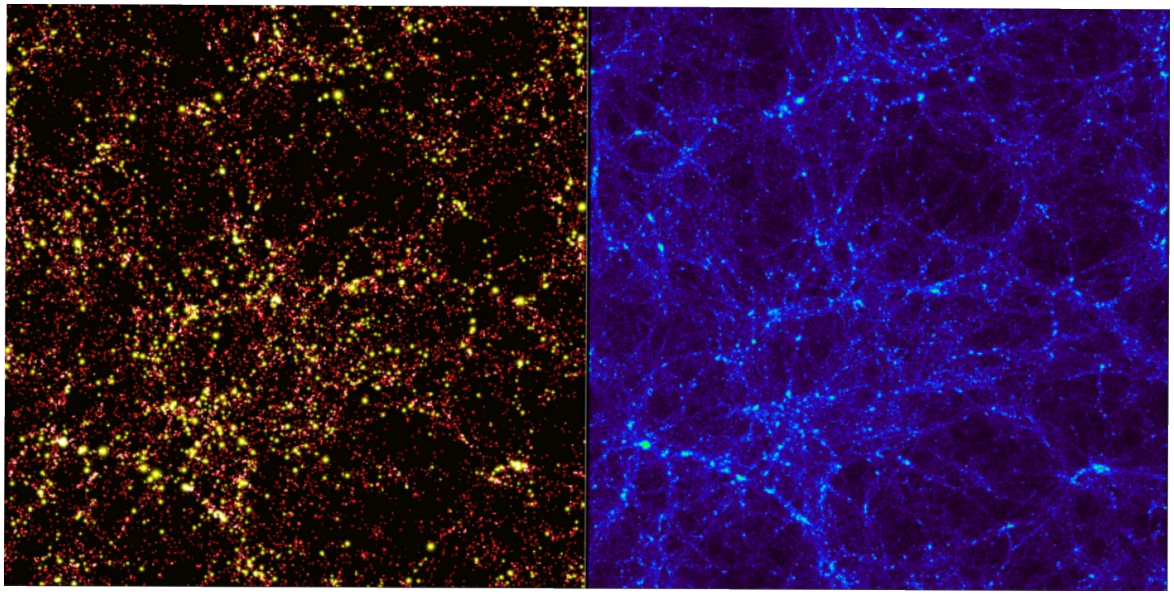
LSS @ DESC



David Alonso
STFC Ernest Ruterford Fellow
University of Oxford

Galaxy clustering

Anze Slosar - BNL

Humna Awan - Rutgers

Eric Gawiser - Rutgers

Javi Sanchez - UCI

Andrina Nicola - Princeton

Adam Broussard - Rutgers

...

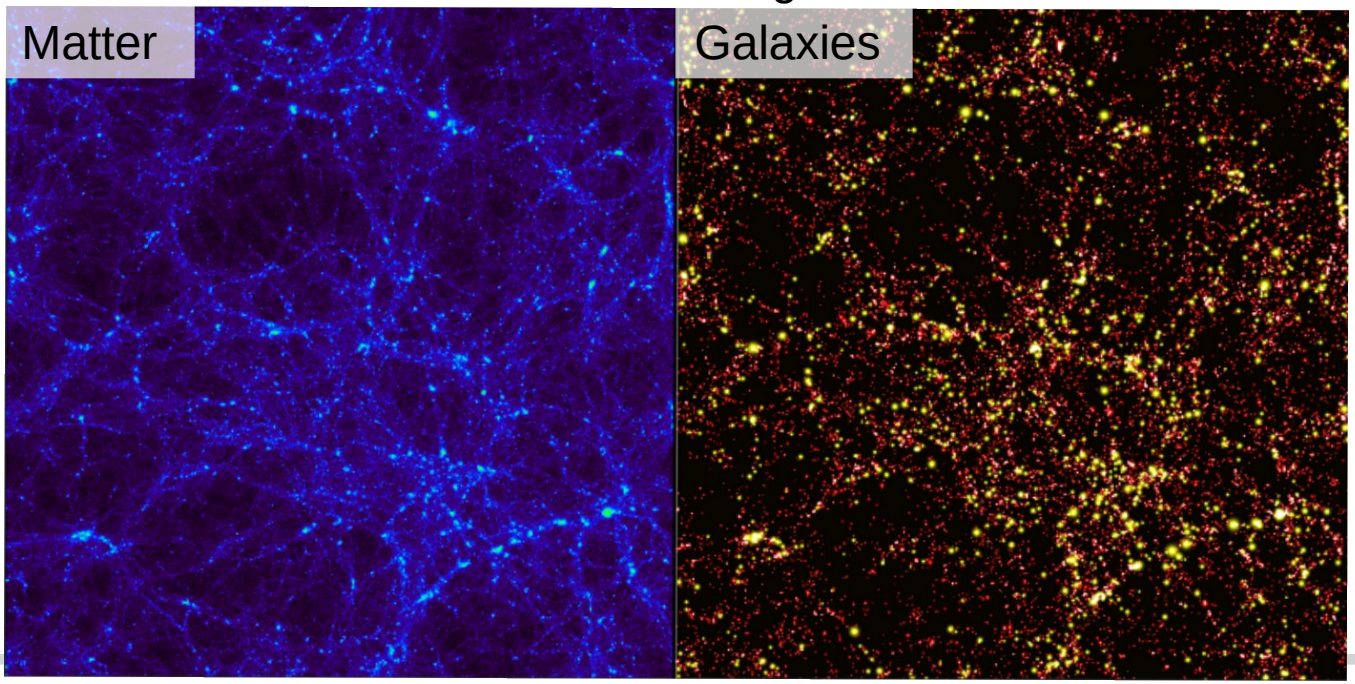
Galaxy clustering

- DE affects cosmic density field
- Galaxy distribution ↔ matter density
- Main systematic → galaxy-matter connection
- In general:

$$\delta_g(x) = f[\delta_M(y)] + \varepsilon(x)$$

- On large scales:

$$\delta_g \sim b_g \, \delta_M$$



Power spectra

What's a power spectrum?

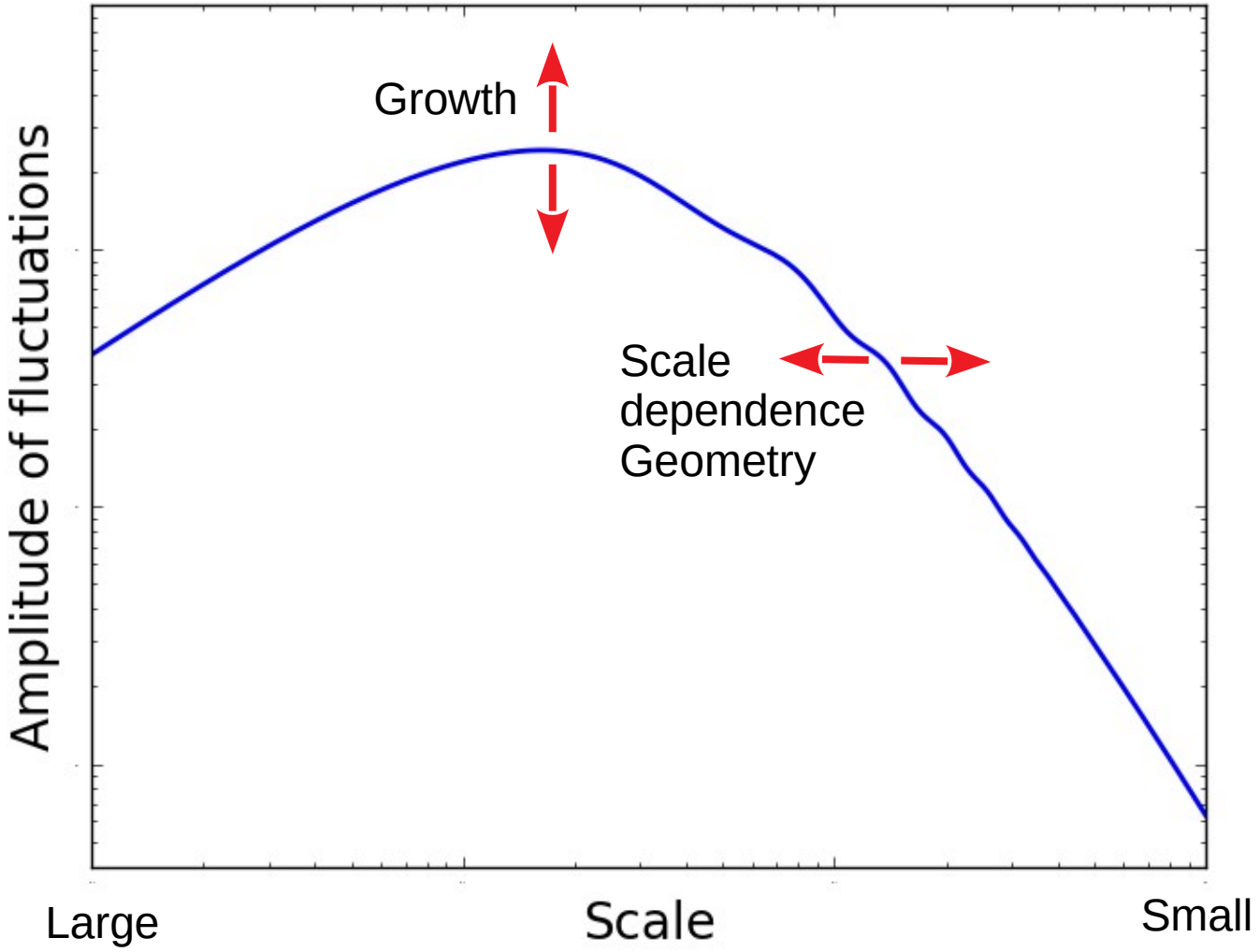
a) It's the variance of a set of equivalent Fourier modes.
i.e. it tells you "how much power is there on these scales?"

$$f(\mathbf{x}) \rightarrow f(\mathbf{k}) = FT[f(\mathbf{x})] \rightarrow C(k) = \langle |f(\mathbf{k})|^2 \rangle$$

b) It's the Fourier-transform of a correlation function:

$$\xi(r) = \langle f(\mathbf{x})\, f(\mathbf{x}+\mathbf{r}) \rangle \rightarrow C(k) = FT(\xi(r))$$

Power spectra



Power spectra

What's a power spectrum?

a) It's the variance of a set of equivalent Fourier modes.
i.e. it tells you "how much power is there on these scales?"

$$f(\mathbf{x}) \rightarrow f(\mathbf{k}) = FT[f(\mathbf{x})] \rightarrow C(k) = \langle |f(\mathbf{k})|^2 \rangle$$

b) It's the Fourier-transform of a correlation function:

$$\xi(r) = \langle f(\mathbf{x})\, f(\mathbf{x}+\mathbf{r}) \rangle \rightarrow C(k) = FT(\xi(r))$$

Why compute power spectra?

a) Cleanly separates theoretically linear and non-linear modes.

b) Different modes are naturally less correlated due to statistical Isotropy. → Simpler covariance, simpler interpretation of error bars.

c) With a smart implementation, it's faster than a 2PCF.

d) Methods developed to deal with systematic contamination optimally

Power spectra

What's a power spectrum?

a) It's the variance of a set of equivalent Fourier modes.
i.e. it tells you "how much power is there on these scales?"

$$f(\mathbf{x}) \rightarrow f(\mathbf{k}) = FT[f(\mathbf{x})] \rightarrow C(k) = < |f(\mathbf{k})|^2 >$$

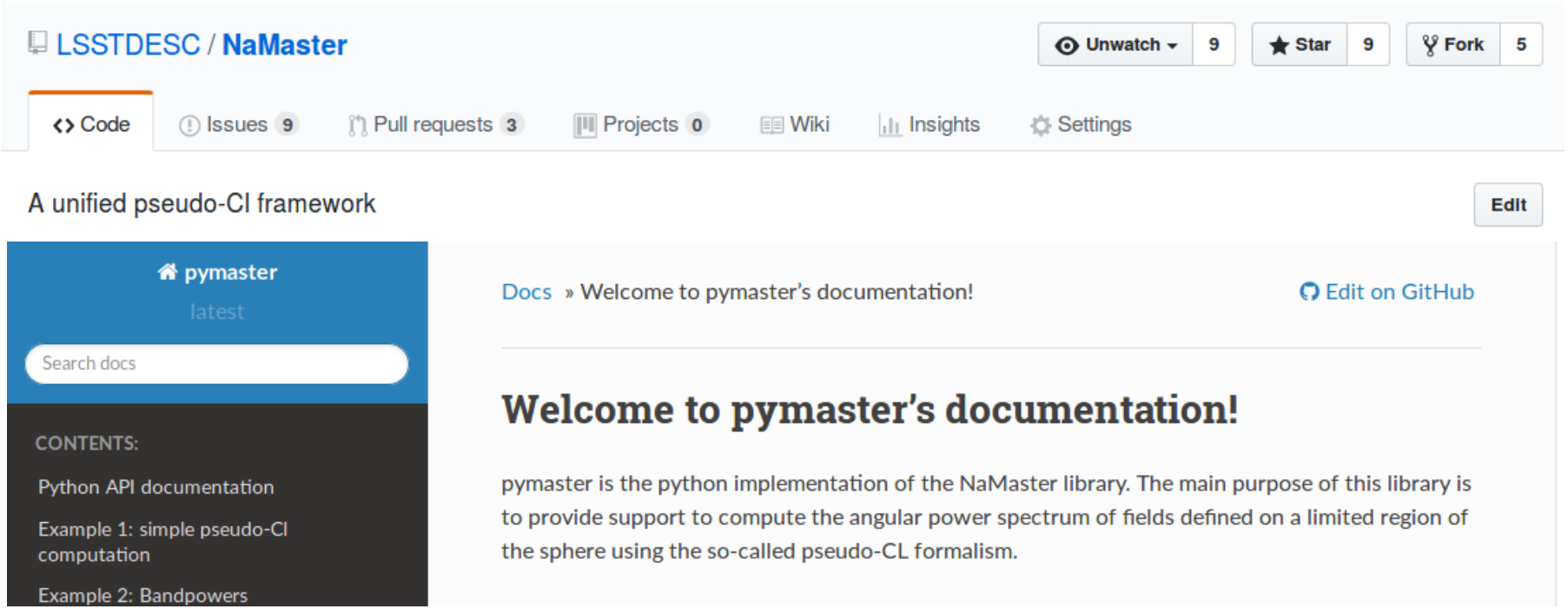
b) It's the Fourier-transform of a correlation function:

$$\xi(r) = < f(\mathbf{x})\, f(\mathbf{x}+\mathbf{r}) > \rightarrow C(k) = FT(\xi(r))$$

Why not compute power spectra?

a) *"Argh, my survey mask is very complicated!"*
(Masks multiply your mask. Due to the convolution theorem, the FT of the observed map is a convolution of FTs. You need to invert the corresponding mode-coupling matrix).
This is a solved problem! (and has been for a while).

Power spectra



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

Power spectra

Example 5: Using workspaces

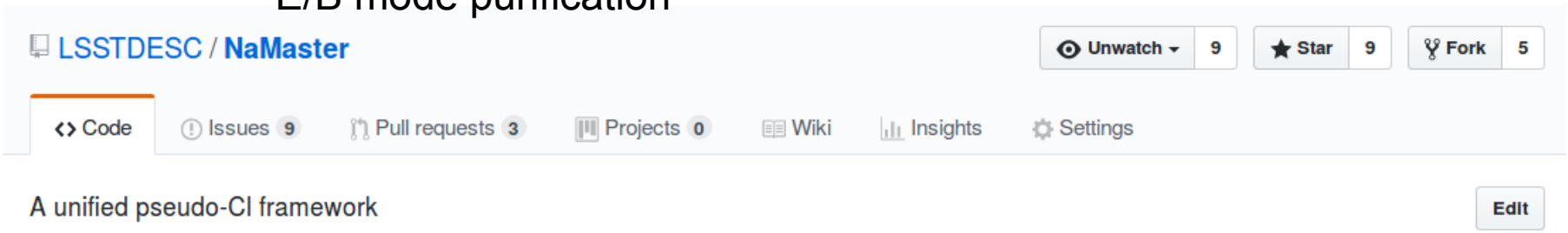
This sample script showcases the use of the NmtWorkspace class to speed up the computation of multiple power spectra with the same mask. This is the most general example in this suite, showing also the correct way to compare the results of the MASTER estimator with the theory power spectrum.

```
import numpy as np
import healpy as hp
import matplotlib.pyplot as plt
import pymaster as nmt

#This script showcases the use of NmtWorkspace objects to speed up the
#computation of power spectra for many pairs of fields with the same masks.

#HEALPix map resolution
nside=256

#We start by creating some synthetic masks and maps with contaminants.
#Here we will focus on the cross-correlation of a spin-2 and a spin-1 field.
#a) Read and apodize mask
mask=nmt.mask_apodization(hp.read_map("mask.fits",verbose=False),1.,apotype="Smooth")
#b) Read maps
mp_t,mp_q,mp_u=hp.read_map("maps.fits",field=[0,1,2],verbose=False)
#c) Read contaminants maps
tm_t,tm_q,tm_u=hp.read_map("temp.fits",field=[0,1,2],verbose=False)
#d) Create contaminated fields
#   Spin-0
f0=nmt.NmtField(mask,[mp_t+tm_t],templates=[[tm_t]])
#   Spin-2
f2=nmt.NmtField(mask,[mp_q+tm_q,mp_u+tm_u],templates=[[tm_q,tm_u]])
#e) Create binning scheme. We will use 20 multipoles per bandpower.
b=nmt.NmtBin(nside,nlb=20)
```

Power spectra

What **features** does it implement?

- Calculate PCL power spectra (including coupling matrix, etc.)
- Capable of doing both:
  - Full spherical case using spherical transforms
  - Flat-sky patches using 2D FFT
- Capable of doing both:
  - Spin-0 fields (density, CMB temperature)
  - Spin-2 fields (shear, CMB polarization)
  - Cross-correlations
- Bells and whistles:
  - Mode deprojection
  - E/B mode purification

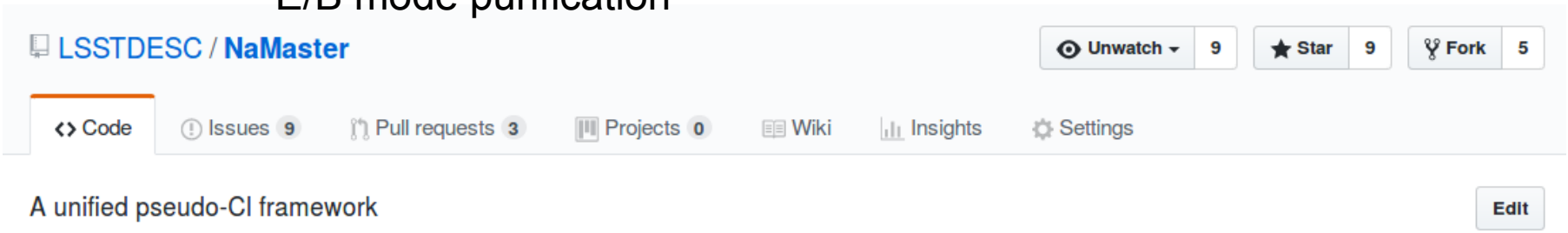


Code: https://github.com/LSSTDESC/NaMaster
Docs: https://namaster.readthedocs.io/en/latest/index.html

Power spectra

What **features** does it implement?

- Calculate PCL power spectra (including coupling matrix, etc.)
- Capable of doing both:
  - Full spherical case using spherical transforms
  - Flat-sky patches using 2D FFT
- Capable of doing both:
  - Spin-0 fields (density, CMB temperature)
  - Spin-2 fields (shear, CMB polarization)
  - Cross-correlations
- Bells and whistles:
  - Mode deprojection
  - E/B mode purification

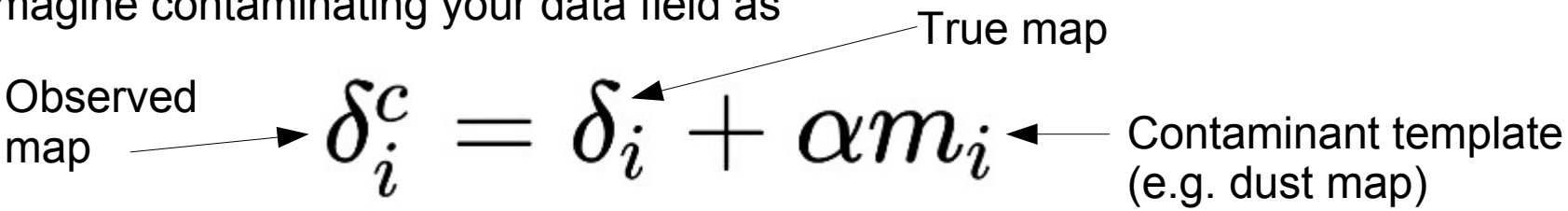


Code: https://github.com/LSSTDESC/NaMaster
Docs: https://namaster.readthedocs.io/en/latest/index.html

Contaminant 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

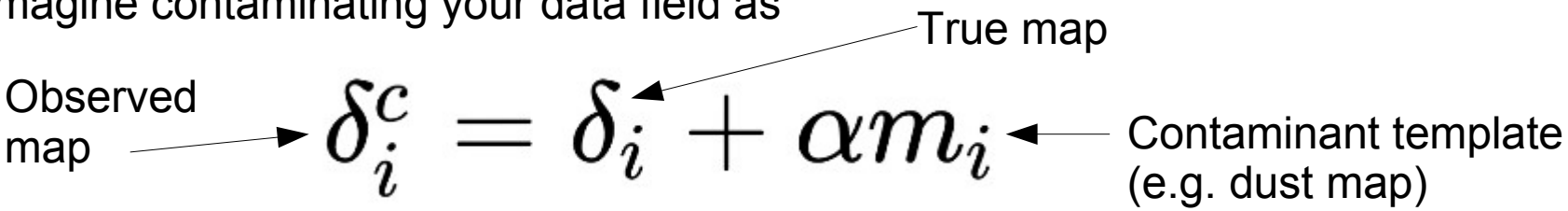


A proper analysis would marginalize over α.

Contaminant 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

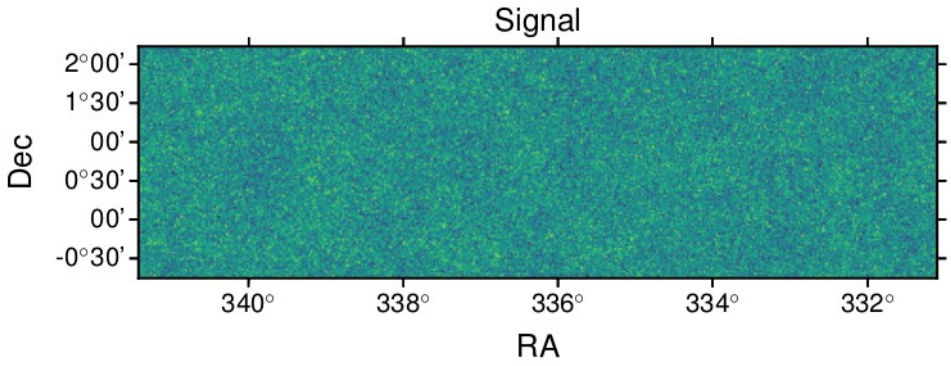


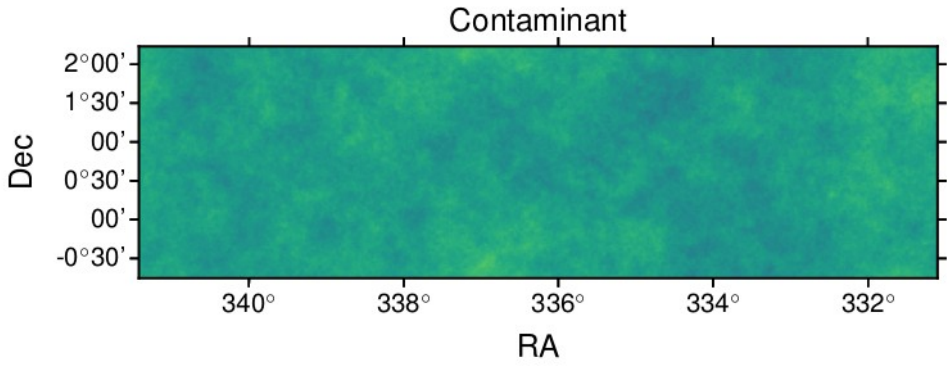
  A proper analysis would marginalize over α.

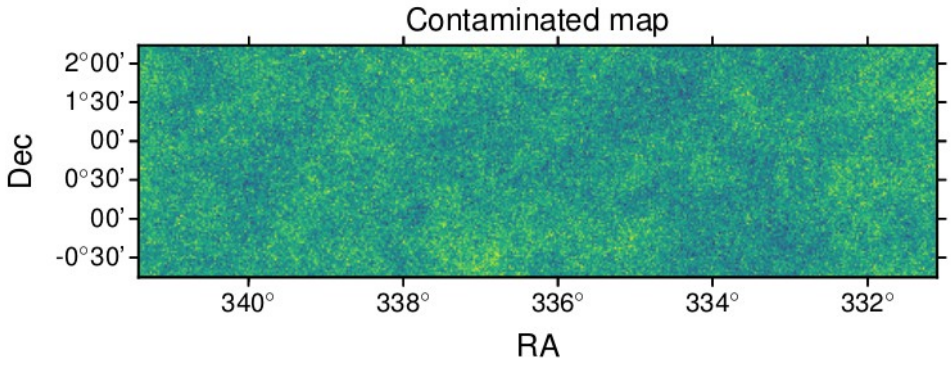
If you do the maths, in PCL this amounts to:

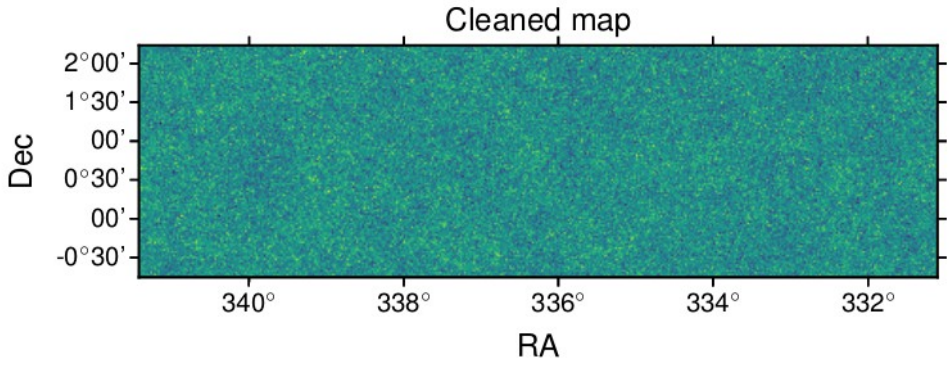
- Finding the best fit value of α.
- Subtracting a contaminant map from the data using this α
- Calculate the PCL estimates and **correct for the bias** this subtraction has produced
- Multiply by the inverse of the mode-coupling matrix

Contaminant deprojection









Power spectra

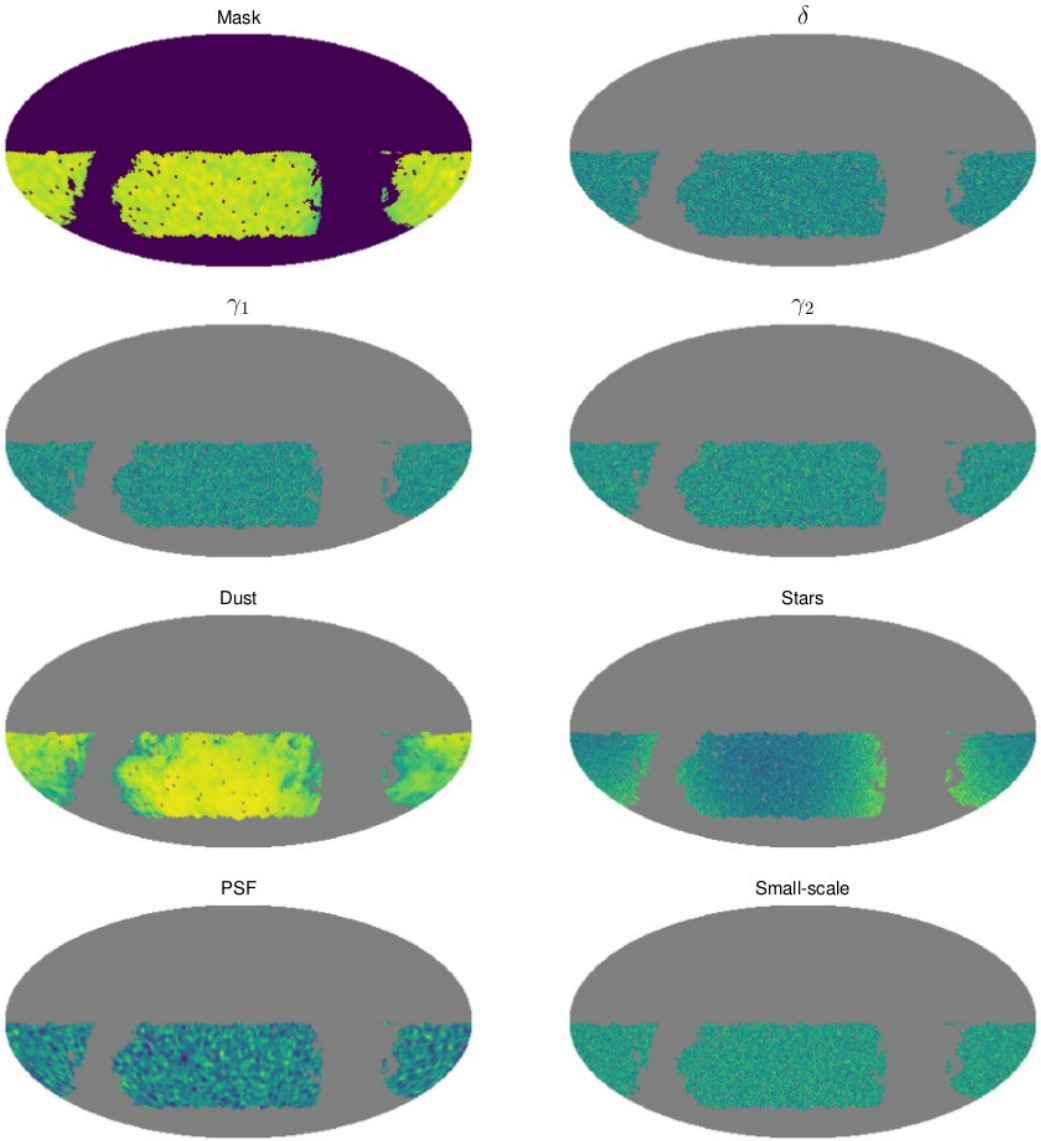
2 validation suites:

- **LSS:** galaxy clustering and lensing with a large set of contaminants.
- **CMB:** B-mode and lensing experiments with foreground contamination.

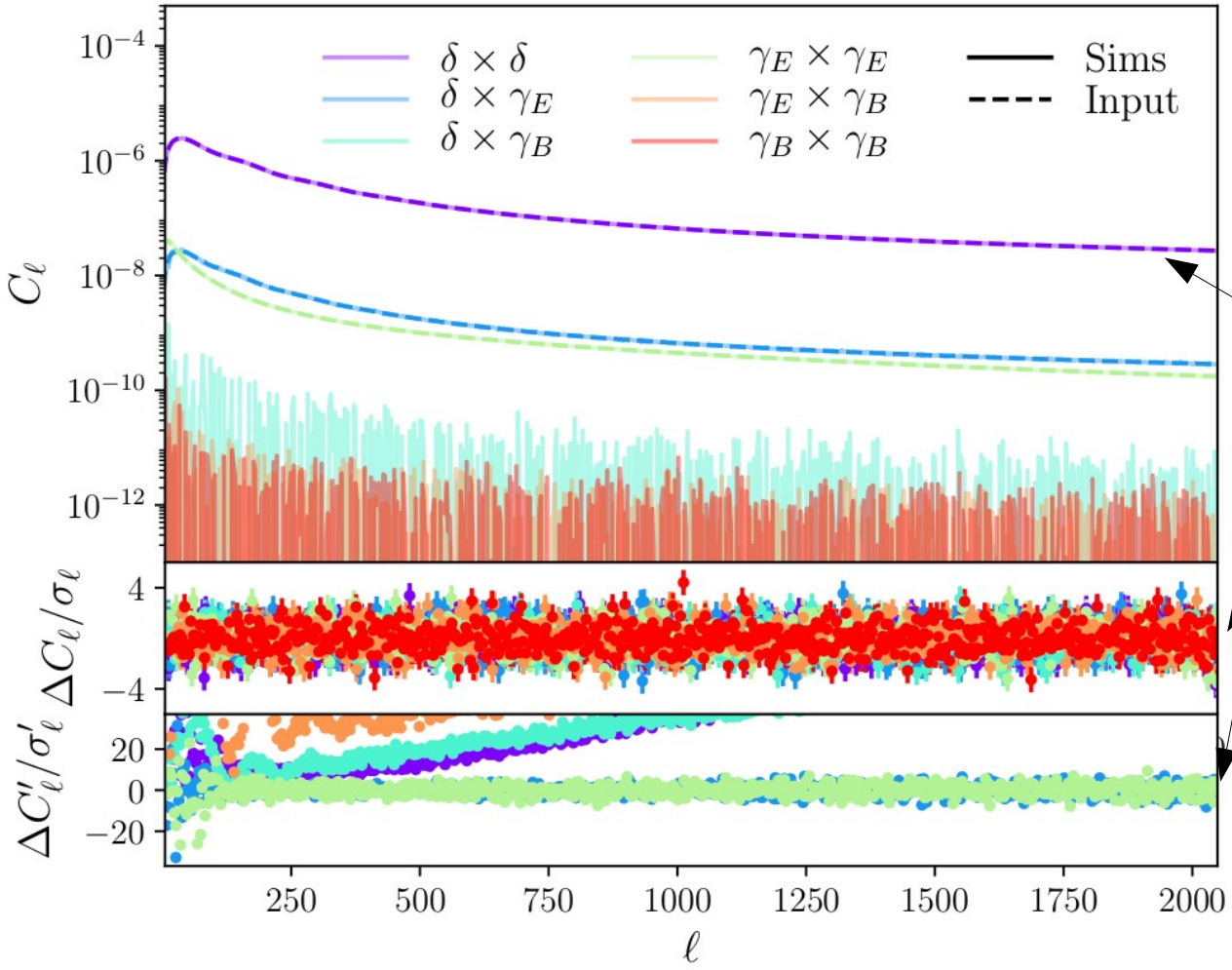
1000 Gaussian simulations

- w./w.o. contaminant deprojection
- w./w.o. E/B purification.
- curved and flat skies.

DA et al. (DESC) 1809.09603



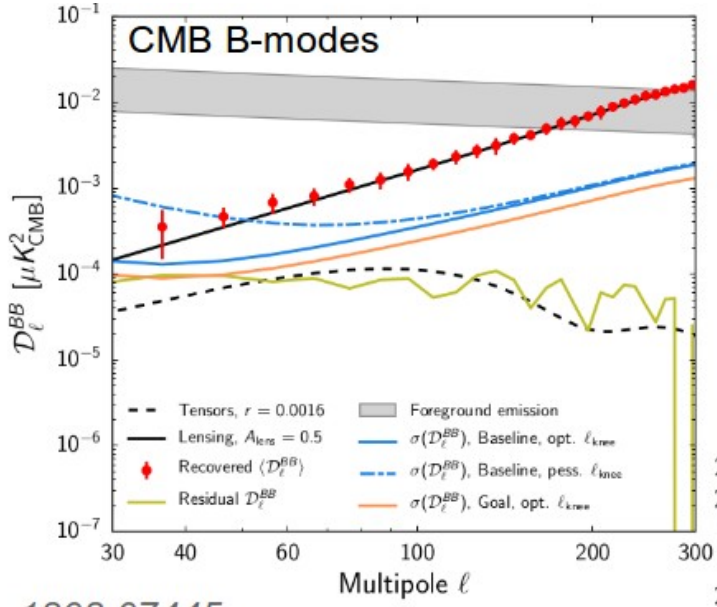
Power spectra



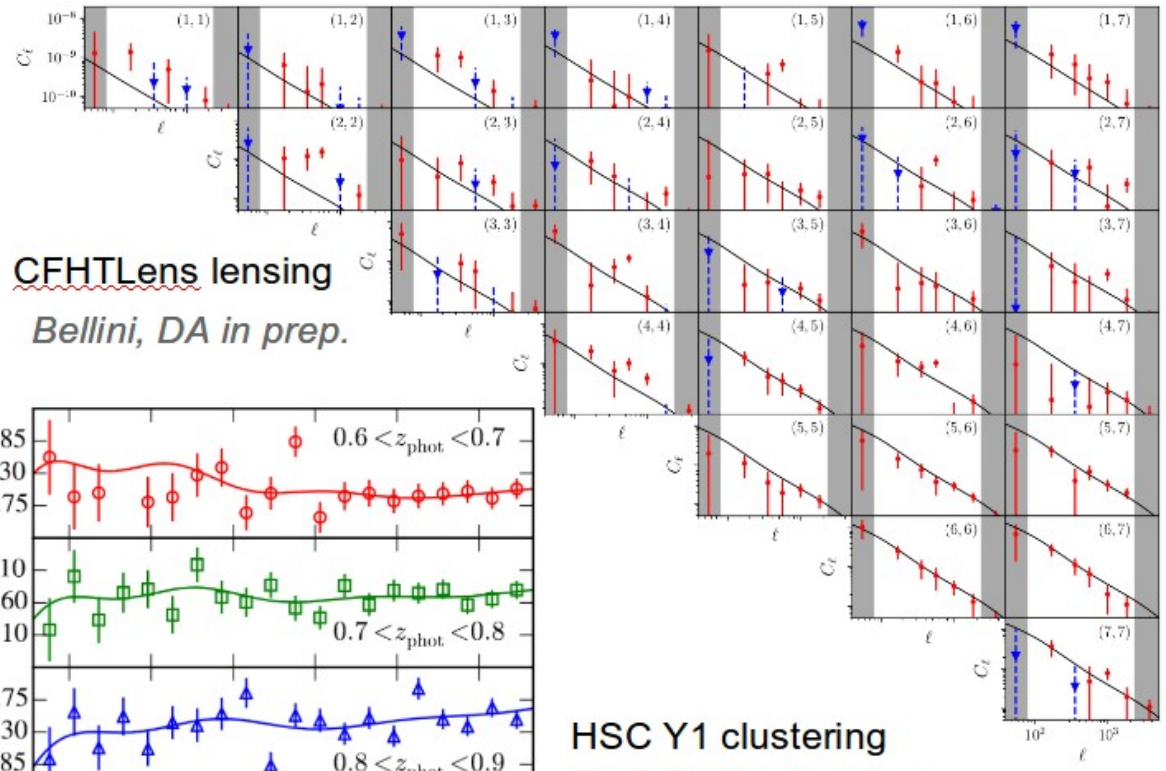
Input power spectra recovered.

Deprojection is important!

Residuals are as expected.



1808.07445



CFHTLens lensing

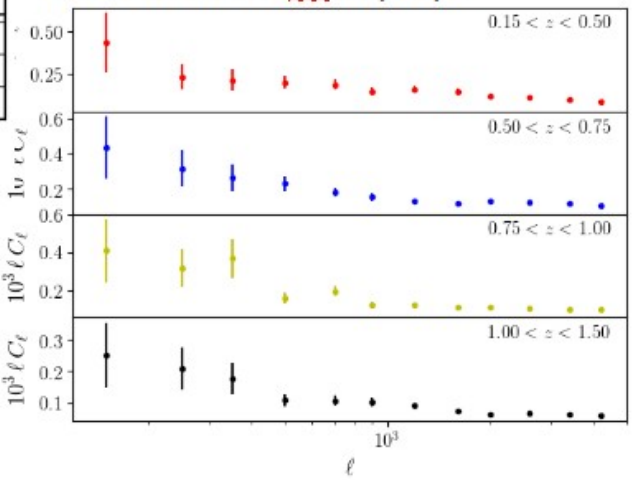
Bellini, DA in prep.

DES Y1 clustering

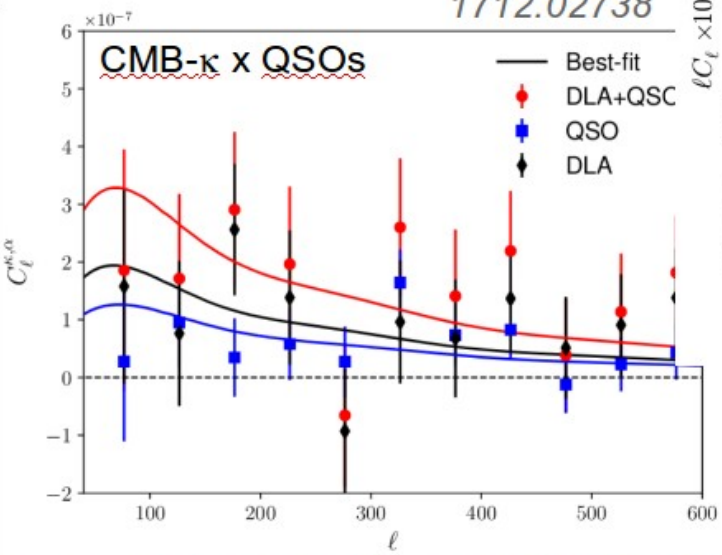
1807.10163

HSC Y1 clustering

DESC LSS et al. in prep.

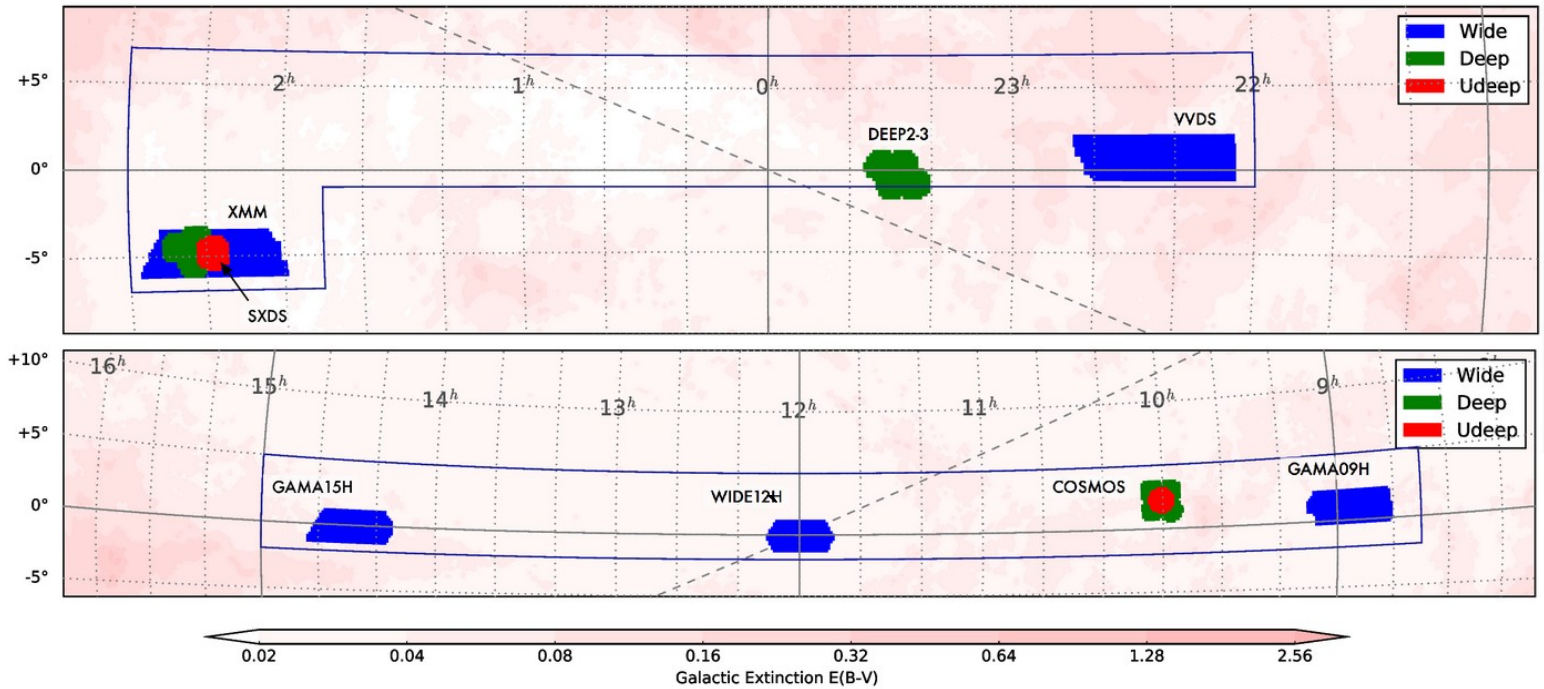


1712.02738



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

HSC analysis



Idea: use public HSC data to perform galaxy clustering analysis.

Motivation:

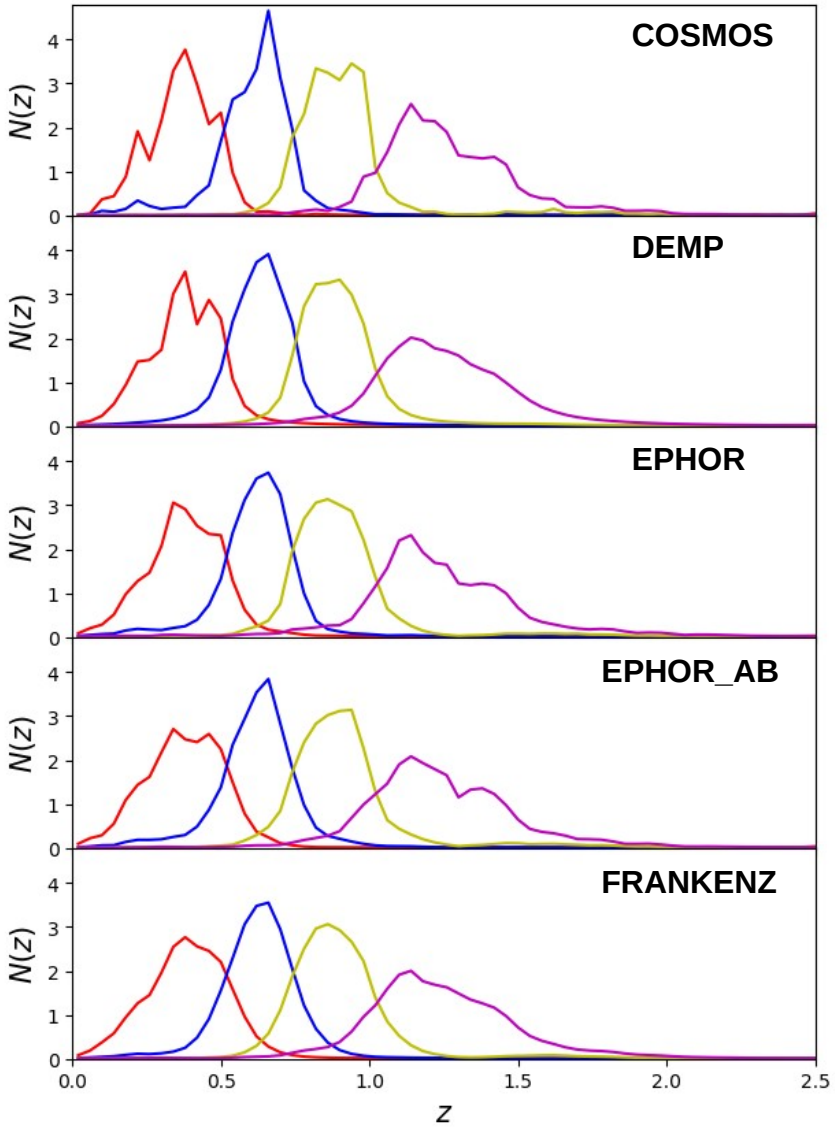
- Same DM pipeline and data format as LSST.
- Learn from these data and use them to improve LSST DM.
- Unused state-of-the-art data!
- New regime of depth and systematics
- Prove that C_ℓ pipelines can deal with this.

HSC analysis

The analysis pipeline uses 100% DESC analysis tools:

- **Ceci** for pipeline prototyping and running.
- DESC-made **tools** for database querying, sample selection, survey geometry and map making.
- Power spectra and Gaussian covariance matrix using **NaMaster**.
- DESC-made **standard** for storage of two-point function data.
- Theory prediction and non-Gaussian covariance with **CCL**.
- Likelihood with **LSSLike** (prototype of TJPCosmo/**firecrown**).
- Pipeline is developed on **github**, and is able to reproduce all current results automatically (starting from zero, including database queries and data download).
- Work from several members of WG (O(10)).
- Our experience has been valuable to spur discussion within LSST DM and DESC about a few issues (survey geometry, mapping, 2-point methods, covariances etc.)

HSC analysis



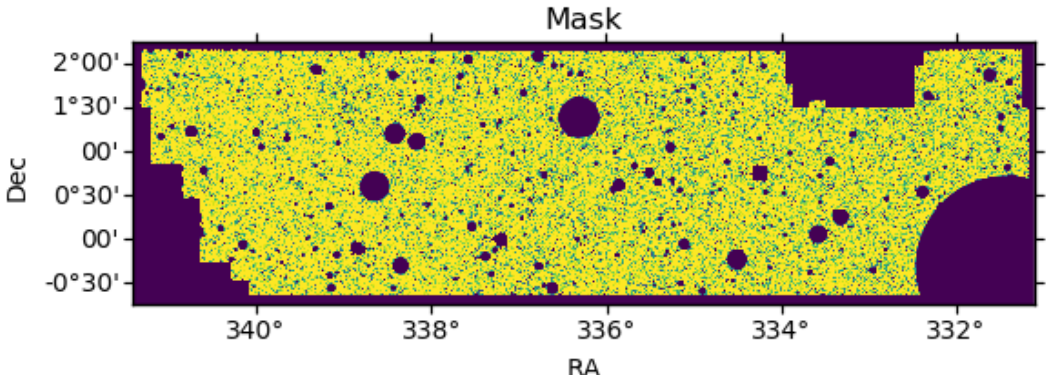
Sample selection:

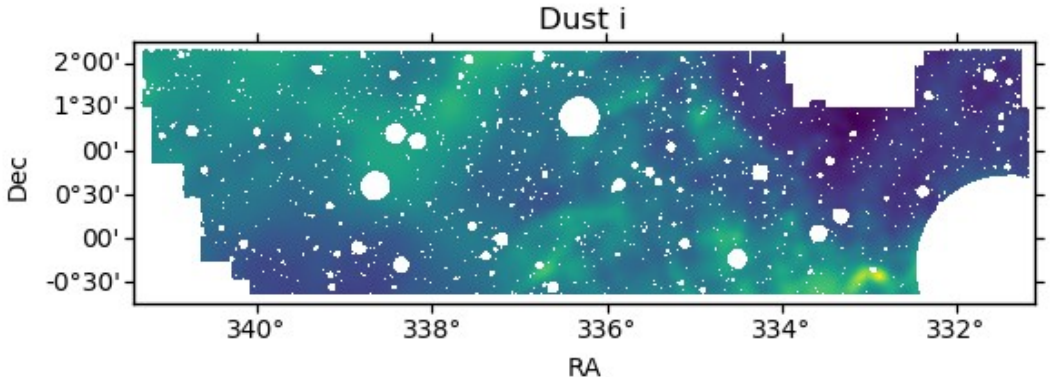
- Standard flags used in the HSC shear sample.
- Strict photometry cuts and blendedness cuts.
- Magnitude limit i<24.5 (well above 10σ depth).
- 4 redshift bins saturate the information content.
- Redshift distributions consistent across different codes/methods.

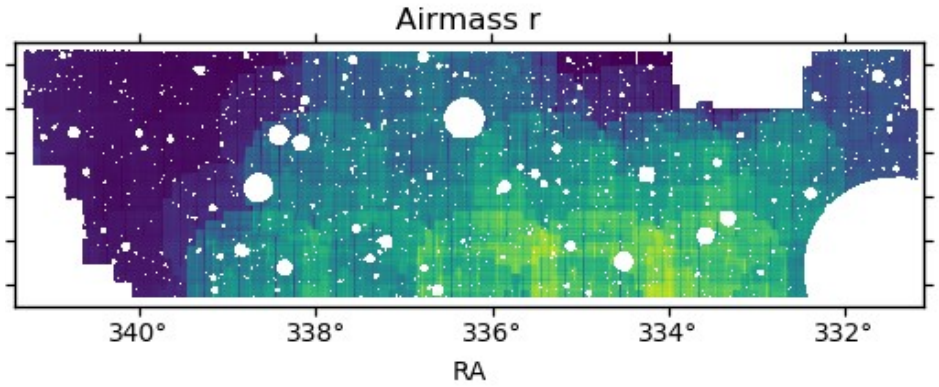
HSC analysis

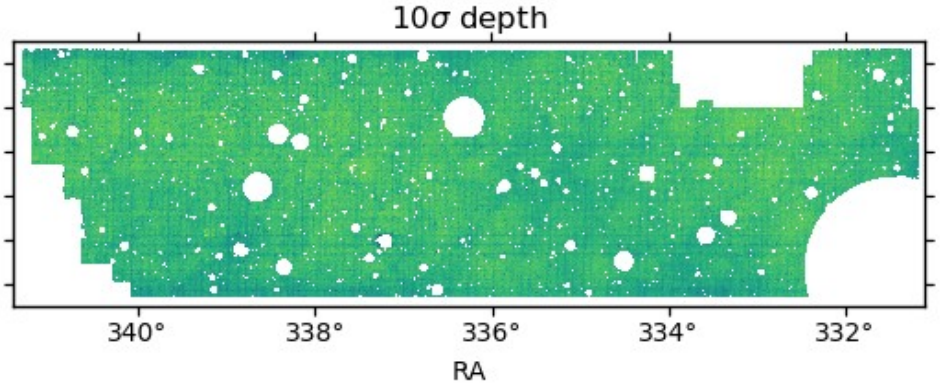
Systematics mapping:

- Masks reconstructed from catalog information. Not ideal, DM aware of this.
- Galactic systematics from external maps or catalogs (dust, stars).
- Observing condition maps from exposure metadata.
- Depth maps from catalog and exposure metadata.
- All relevant systematics are deprojected from our data.

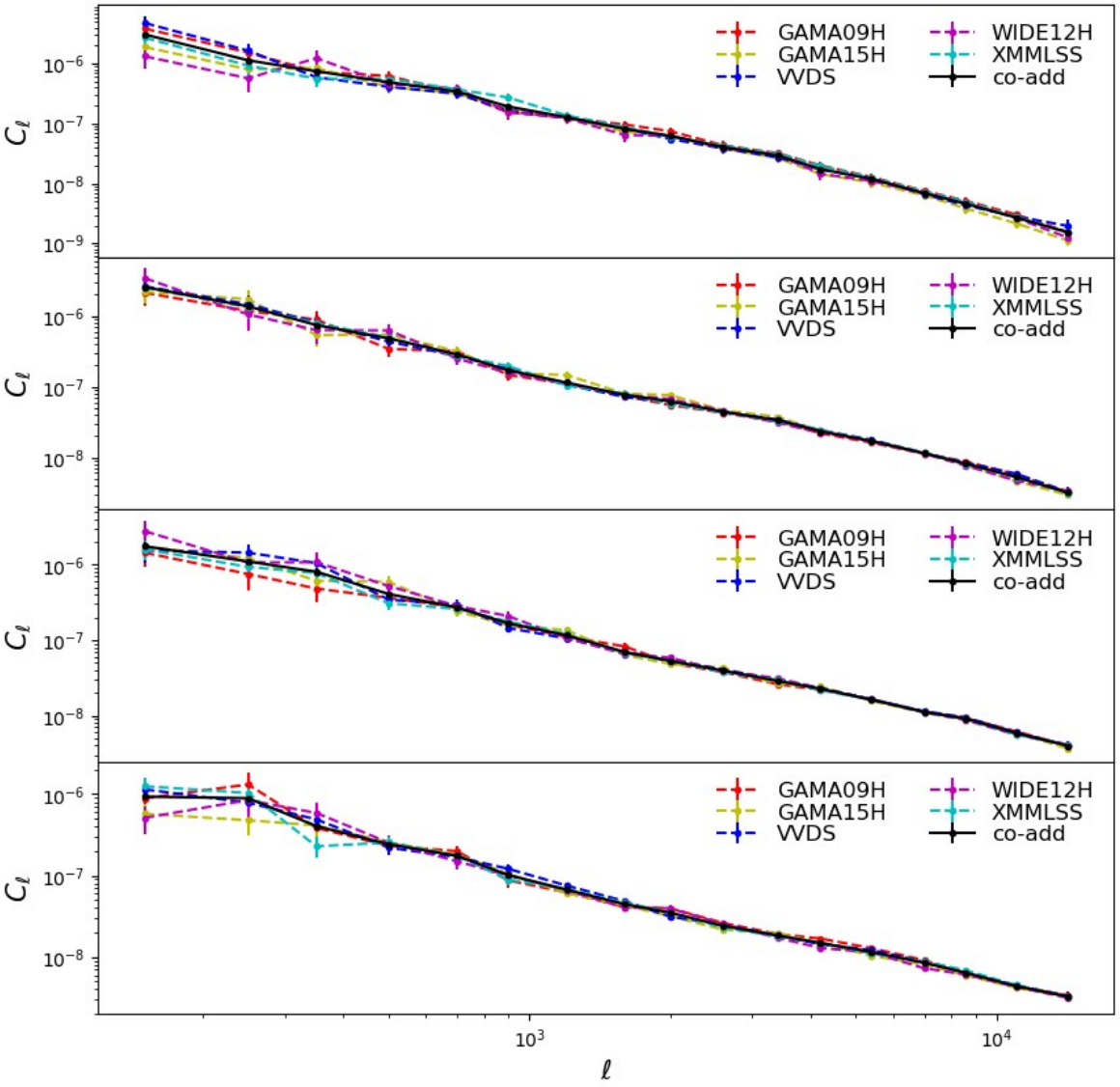








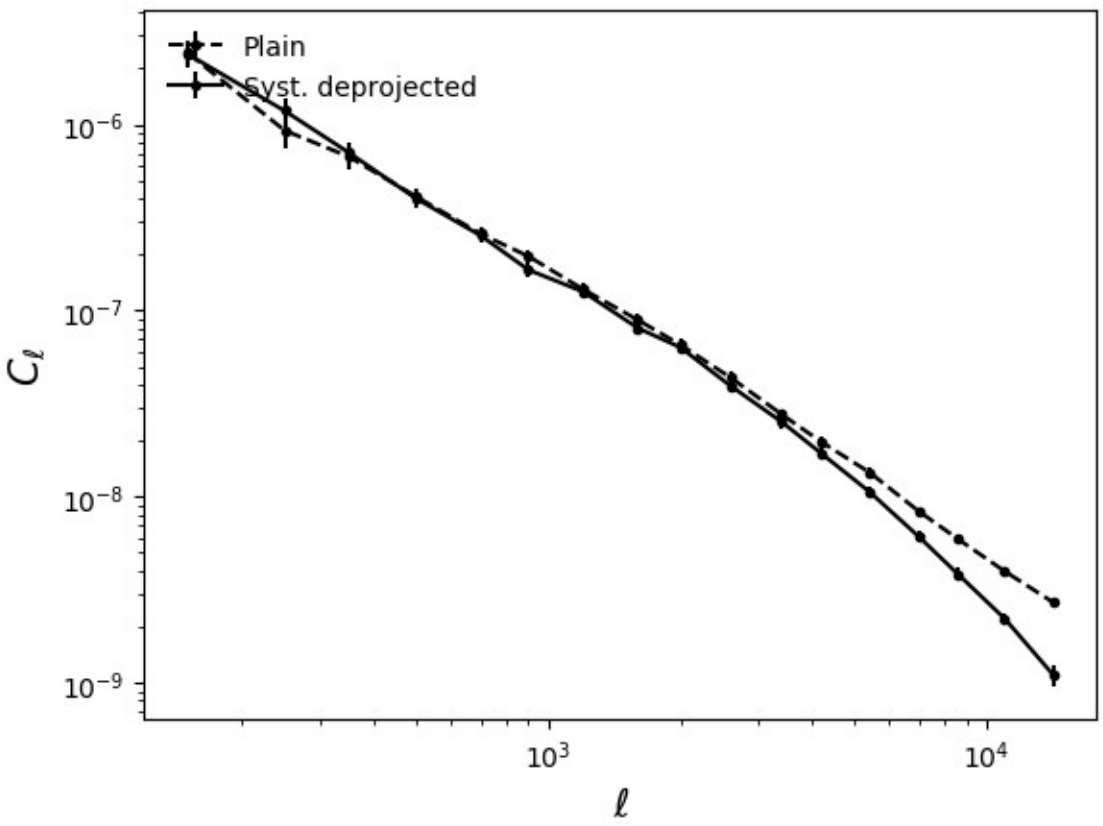
HSC analysis



Power spectra

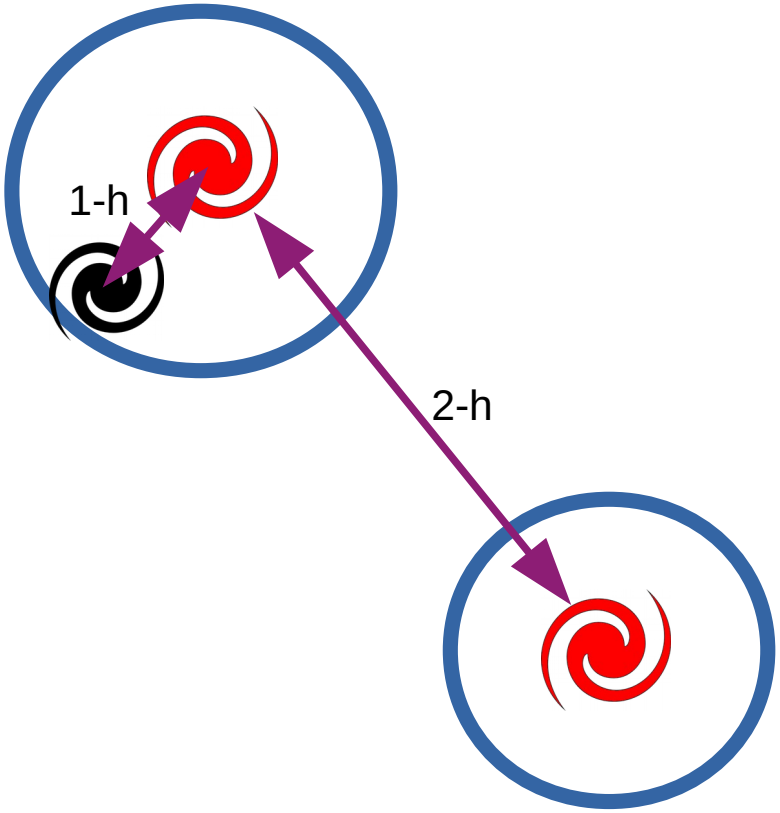
- Computed in 5 different fields and coadded.
- 43 different systematic templates deprojected.
- Power spectra highly compatible across fields.

HSC analysis



Contaminant deprojection can be relevant on both large and small scales.

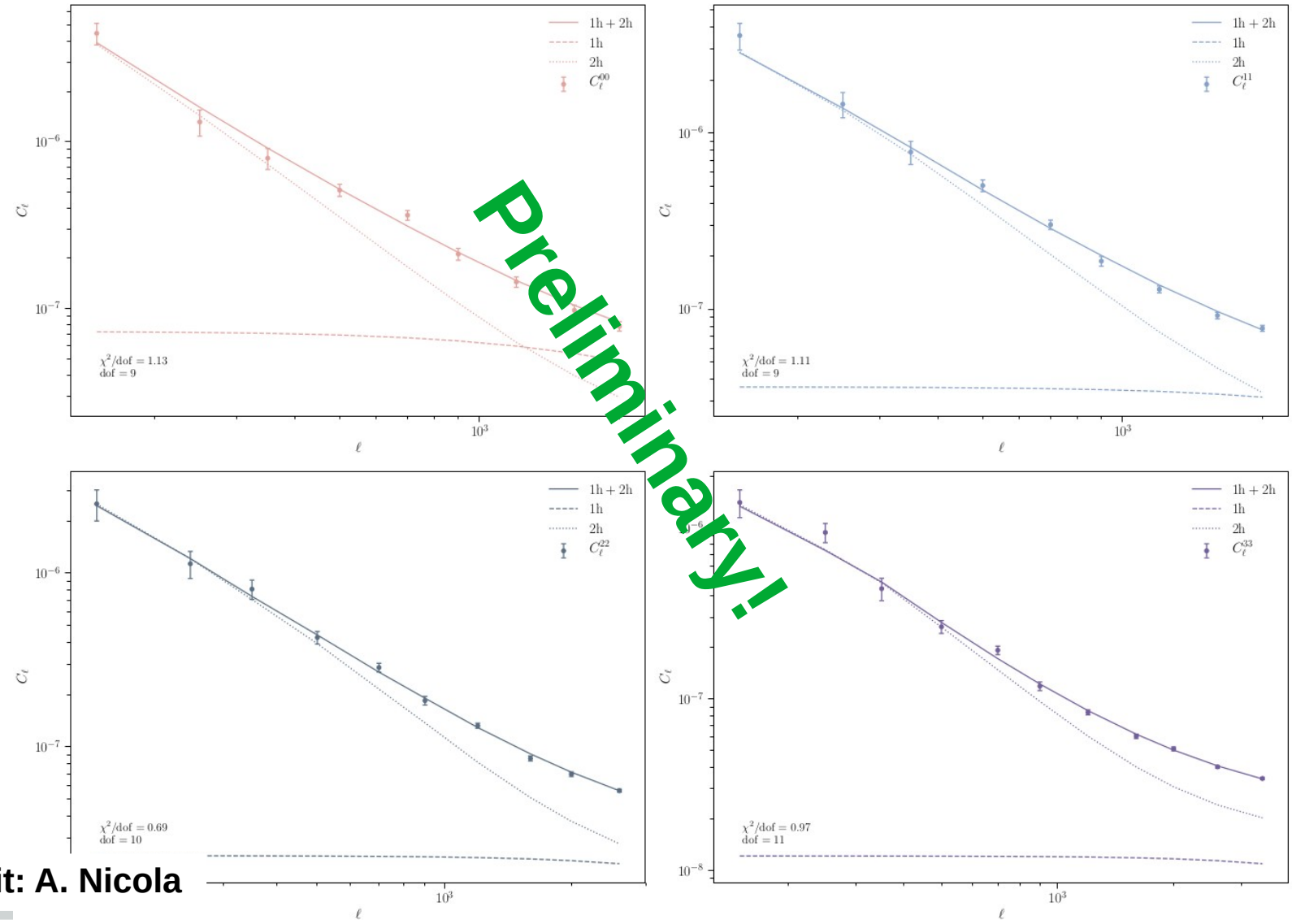
HSC analysis



Theory model

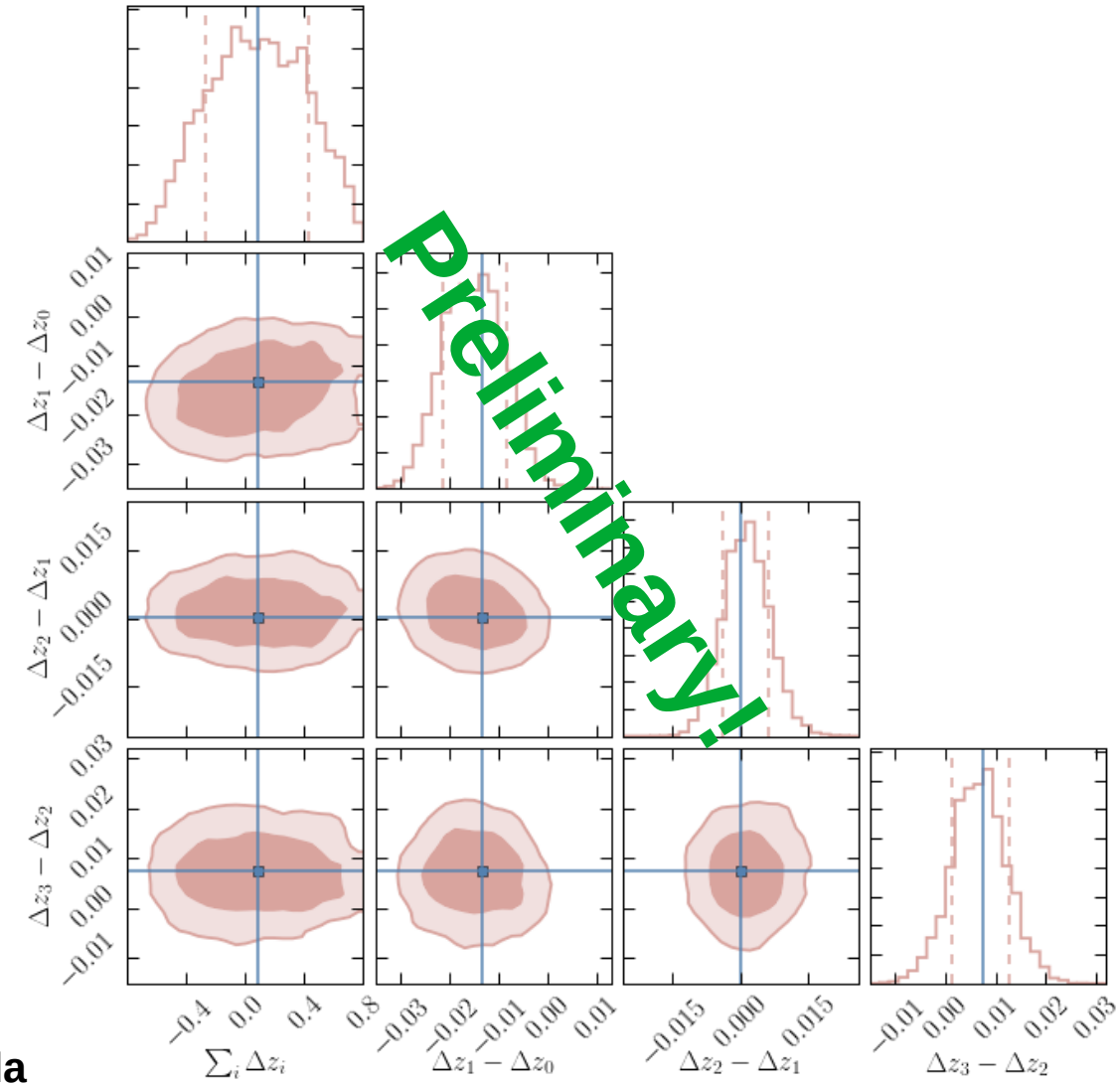
- Largest contiguous area is ~20 deg^2
- Not much info on large scales. Need small-scale model.
- Currently using HOD. Model galaxy-halo connection.
- Good fit down to k~1 Mpc^{-1}

HSC analysis



Credit: A. Nicola

HSC analysis



Credit: A. Nicola

Summary

- LSS constrains dark energy through the statistics of the galaxy distribution. Good probe of local (i.e. not integrated) clustering pattern.
- LSS WG has contributed to a number of DESC projects.
- Design of fast/easy-to-use power spectrum methods (not just for LSS!).
- Analysis of galaxy clustering in HSC data
  - LSST-like data.
  - 100% DESC analysis pipeline.
  - Nice preliminary results: constraints on clustering properties and photo-z systematics.
- More work underway (DC2 projects, TXPipe …). Join us!