Additional probes and alternative techniques for galaxy clustering

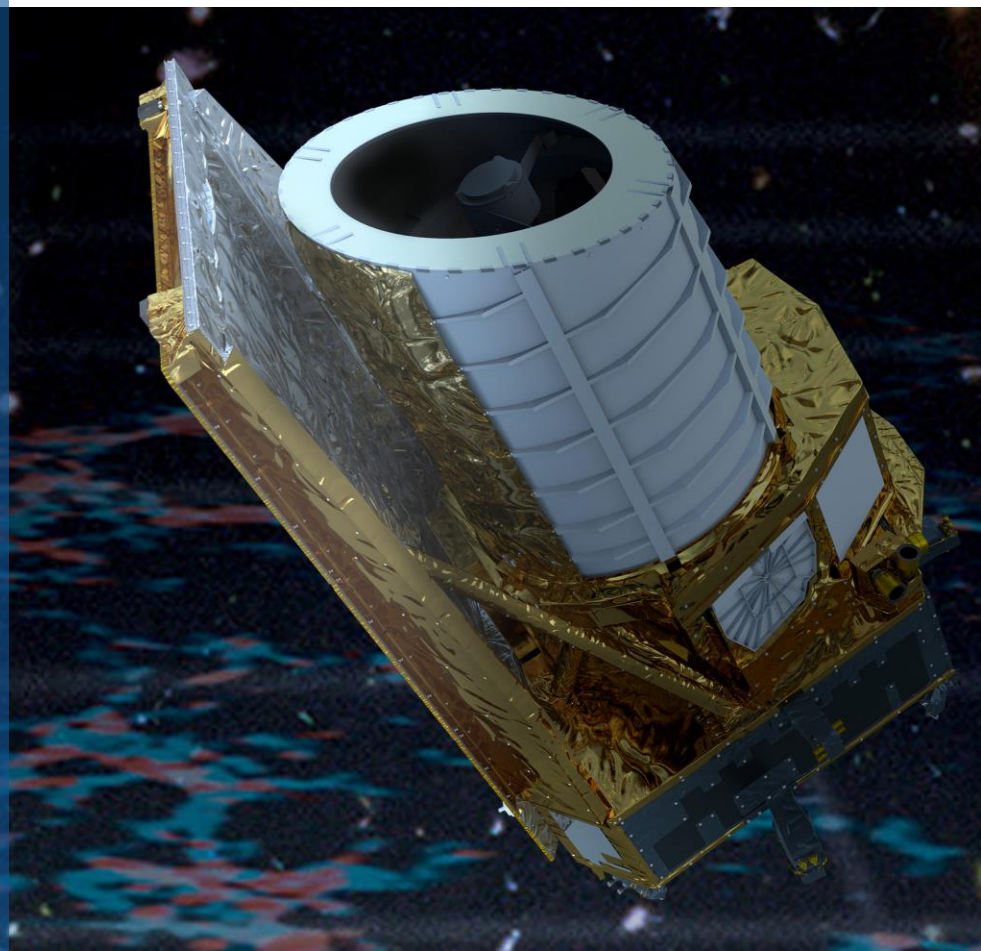
(Galaxy Clustering: Additional Probes work package)

Euclid Galaxy Clustering meeting, Marseille 2024

Florent Leclercq
Institut d'Astrophysique de Paris
CNRS & Sorbonne Université

Cora Uhlemann
Newcastle University

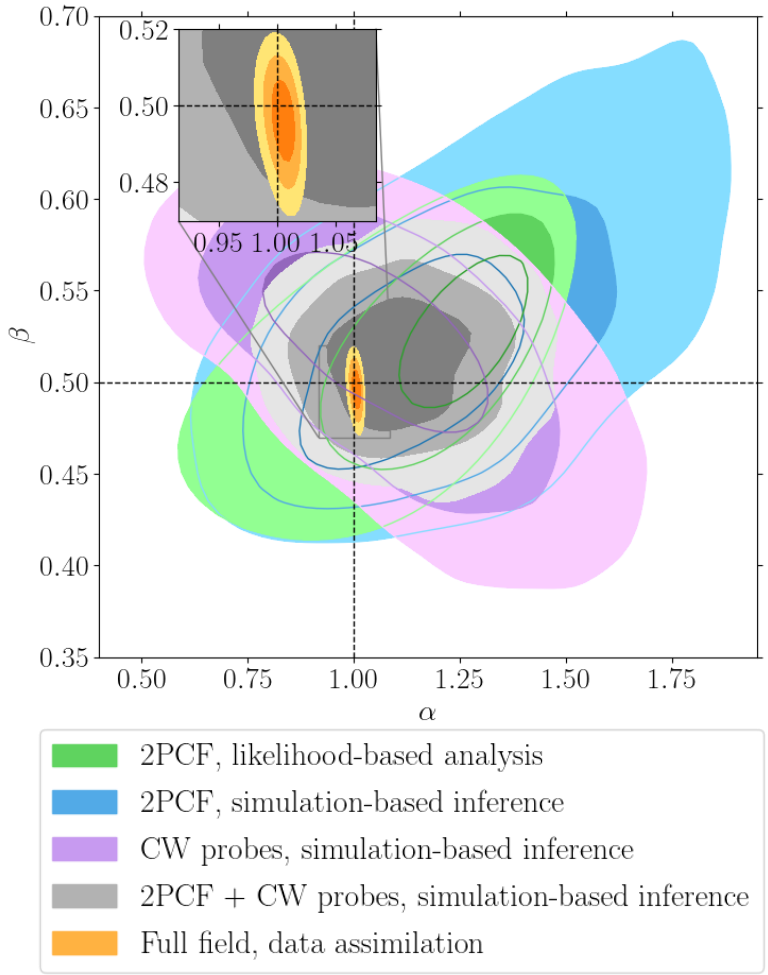
31 January 2024



Going beyond two-point correlations for galaxy clustering: the "implicit" and "explicit" approaches

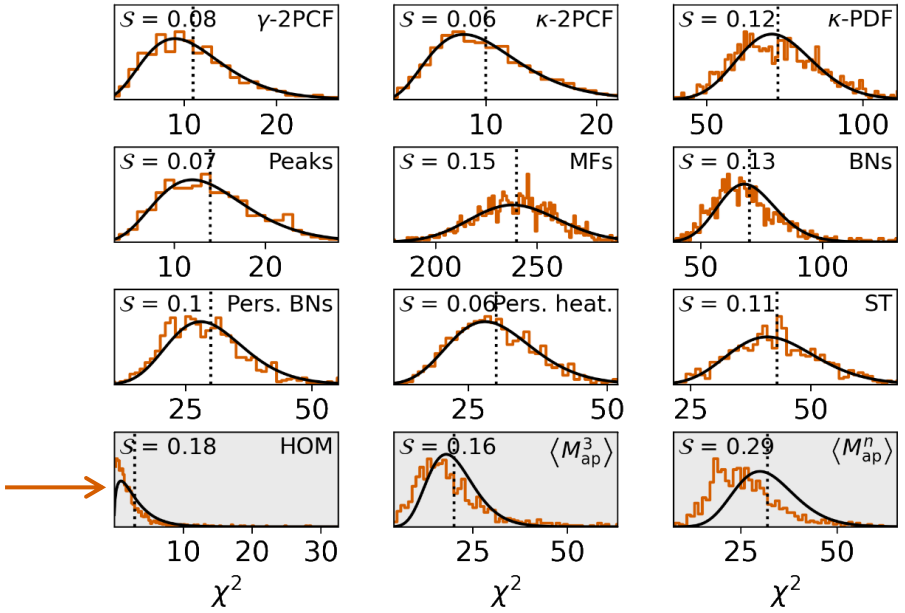
Note:
- likelihood-free inference (LFI) ≈ simulation-based inference (SBI) ≈ implicit likelihood inference (ILI)
- likelihood-based approach = explicit likelihood inference

- A question of accuracy: first, avoid biases.



Leclercq & Heavens, 2103.04158

- Some weak lensing additional probes also have a non-Gaussian distribution.



- A question of precision: can numerical forward models be used to push further than $k \gtrsim 0.15\ h/\mathrm{Mpc}$? The full field contains much more information.

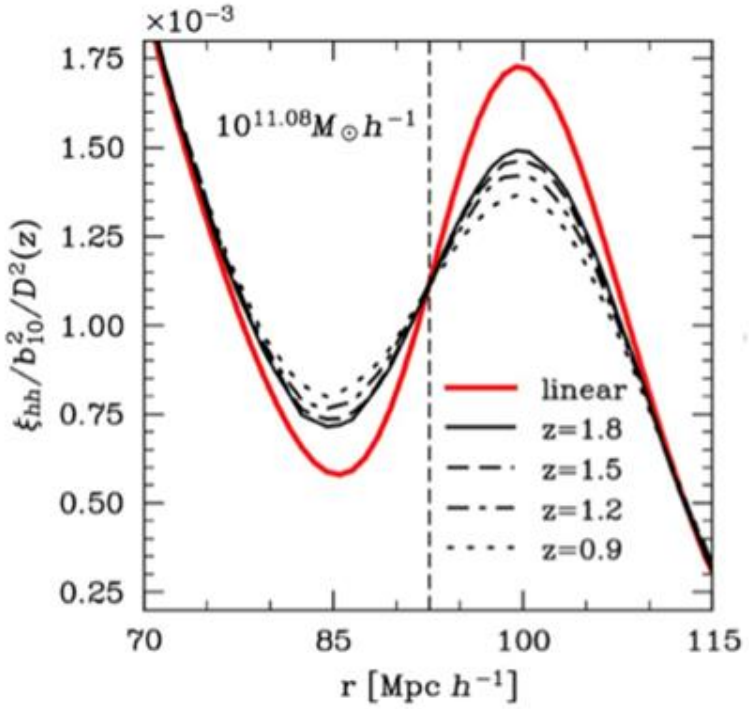
Euclid HOWLS-KP paper 1, Ajani *et al.*, 2301.12890

WP Additional GC Probes

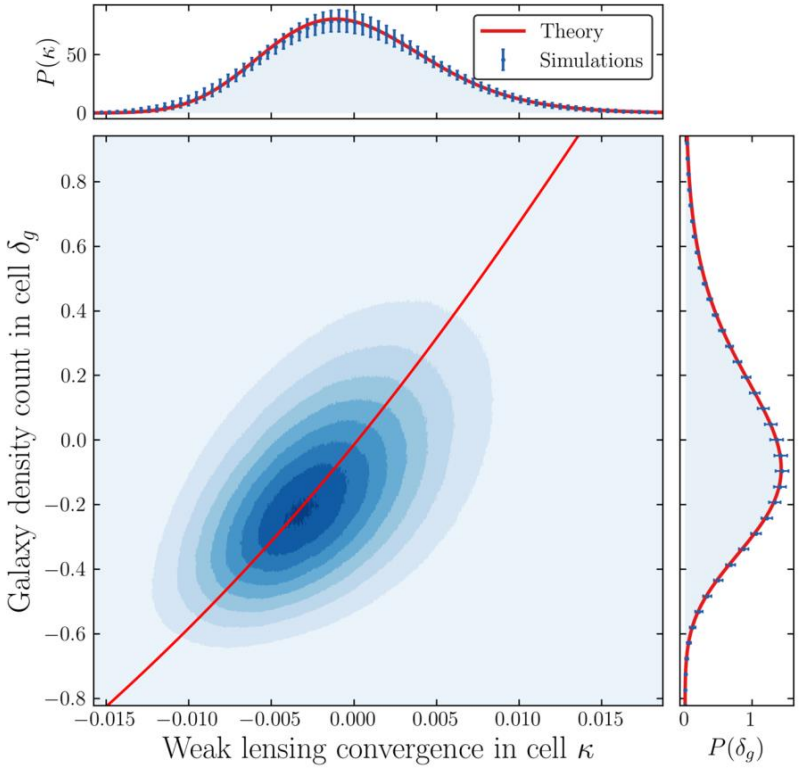
Additional summary statistics & Cross-correlation probes

BAO Linear point standard ruler

between peak & dip $s_{\rm LP} = \frac{s_{\rm p} + s_{\rm d}}{2}$



one-point statistics (incl. photo-z GC & WL)



\+ **more**: angular redshift fluctuations, marked correlations, velocity field, SKAx, …

WP Additional GC Probes

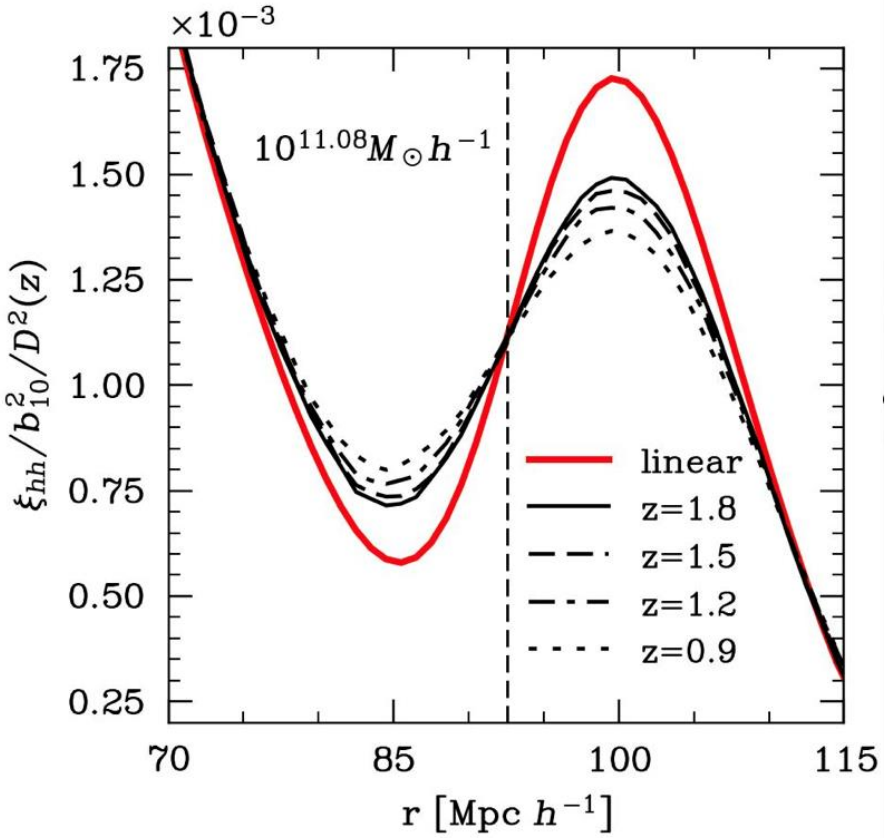
BAO linear point standard ruler

Euclid Flagship I - halo snapshots

- 2pcf measurements
- Development of algorithms for large data-sets
- redshifts: z = 0.9, 1.2, 1.5, 1.8
- Different halo mass cuts
- 2pcf covariance: Gaussian recipe

Flagship halo two-point correlation function

[N-body 2pcf estimated points (connected by straight lines)]

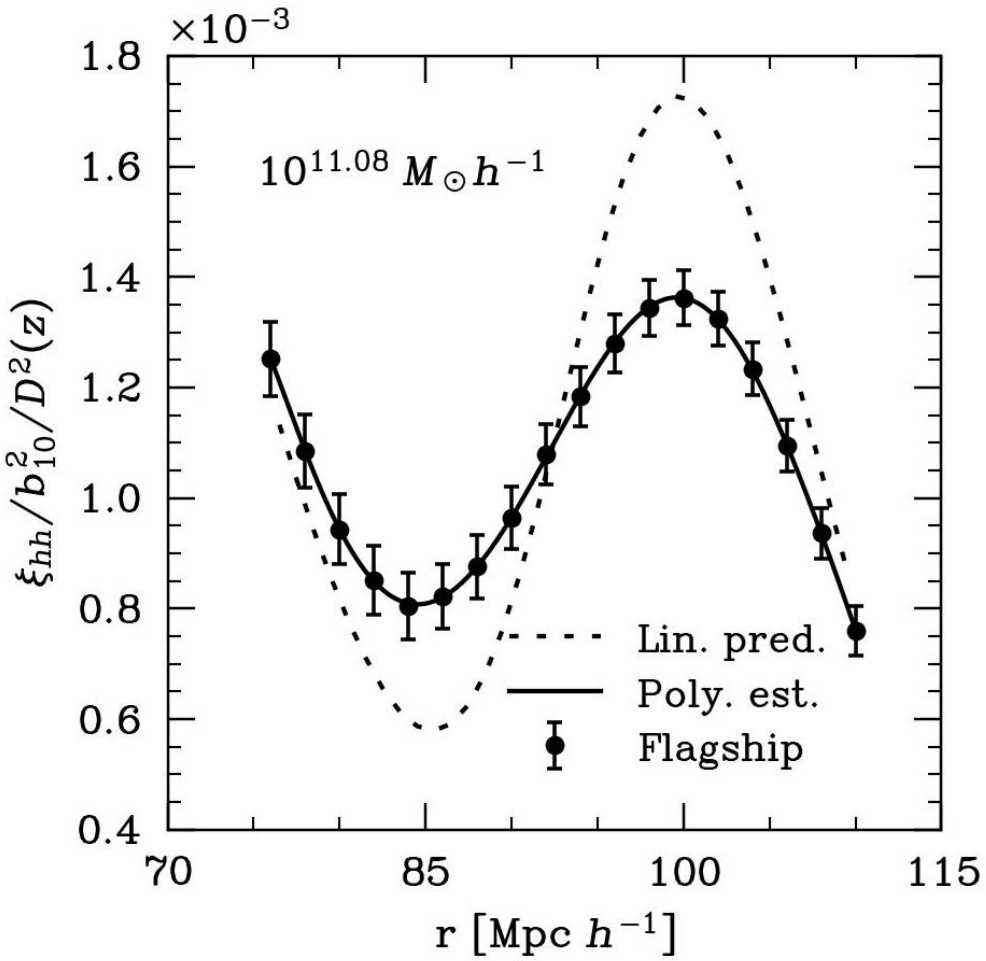


WP Additional GC Probes

BAO linear point standard ruler

Linear point estimation in real space

- Model-independent polynomial estimator
- Validation procedure for unbiased & optimized estimator for each redshift & mass cut
- Uncertainties properly propagated



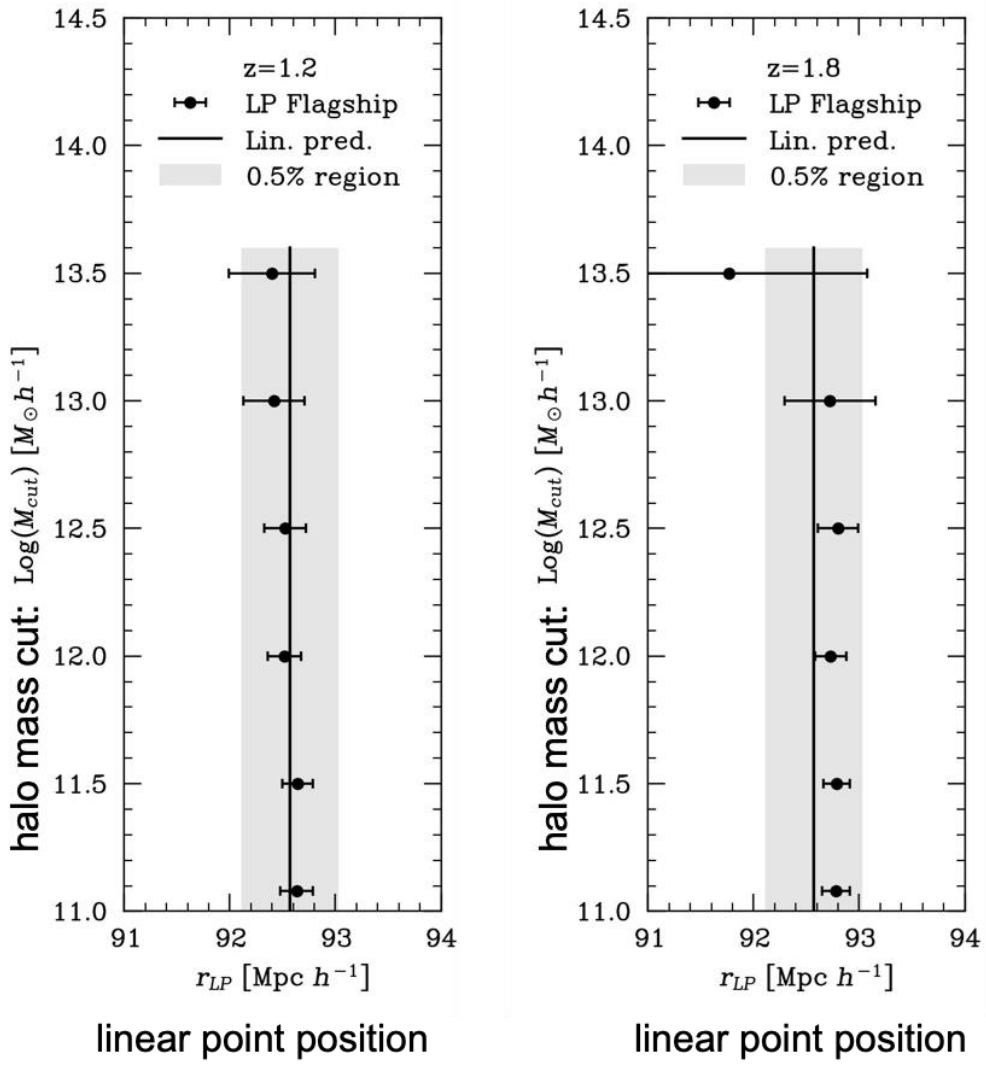
WP Additional GC Probes

BAO linear point standard ruler

Linear point estimation in real space

- Model-independent polynomial estimator
- Validation procedure for unbiased & optimized estimator for each redshift & mass cut
- Uncertainties properly propagated
- Good agreement with linear theory prediction

extraction of **z-space catalogs** just started



Stefano Anselmi, Filippo Oppizzi, Alessandro Renzi, Pier-Stefano Corasaniti, Giovanni Verza, Santiago Casas, …

WP Additional GC Probes

One-point PDF of spec-z galaxy count

Measurement

(Galaxy) counts in cells histogram

Simulation

Halo counts in cells in Flagship I

Spheres radius >10-20 Mpc/h

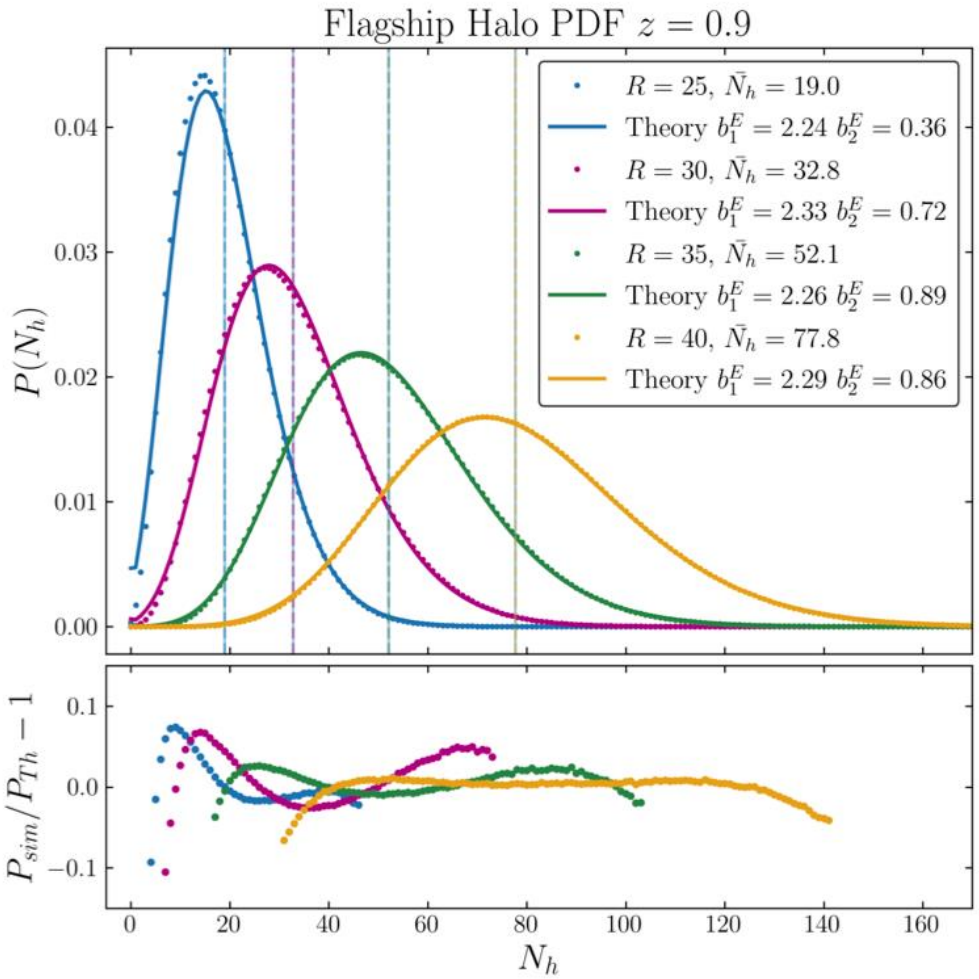
Theory

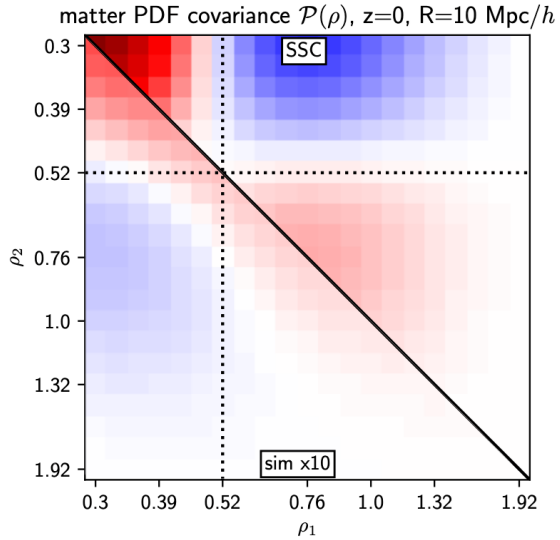
Predicted matter counts in cells

+ Quadratic bias

+ Poisson shot noise

Covariance via 2-point PDF





Lina Castiblanco Tolosa, *Cora Uhlemann*, Alexandre Barthelemy, Sandrine Codis, Beth Gould

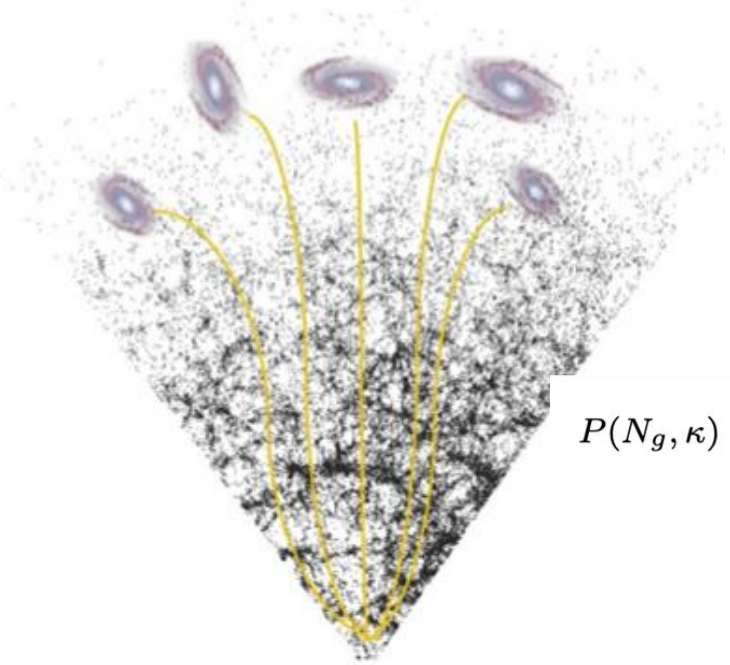
WP Additional GC Probes

One-point PDF of photo-z galaxy count & weak lensing

Joint one-point statistics

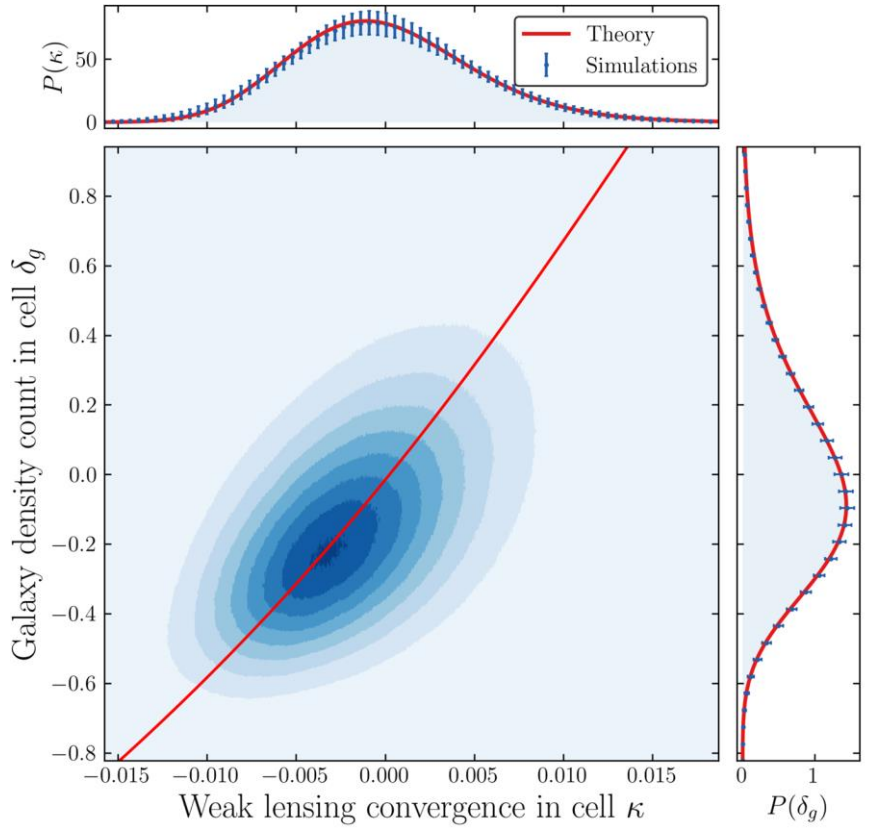
photo-z GC counts x WL convergence

circular apertures of ~10arcmin



$$P(N_g, \kappa) = \int d\delta_m \, p(N_g|\delta_m) P(\delta_m, \kappa)$$

Galaxy bias, Shot Noise — Cosmology



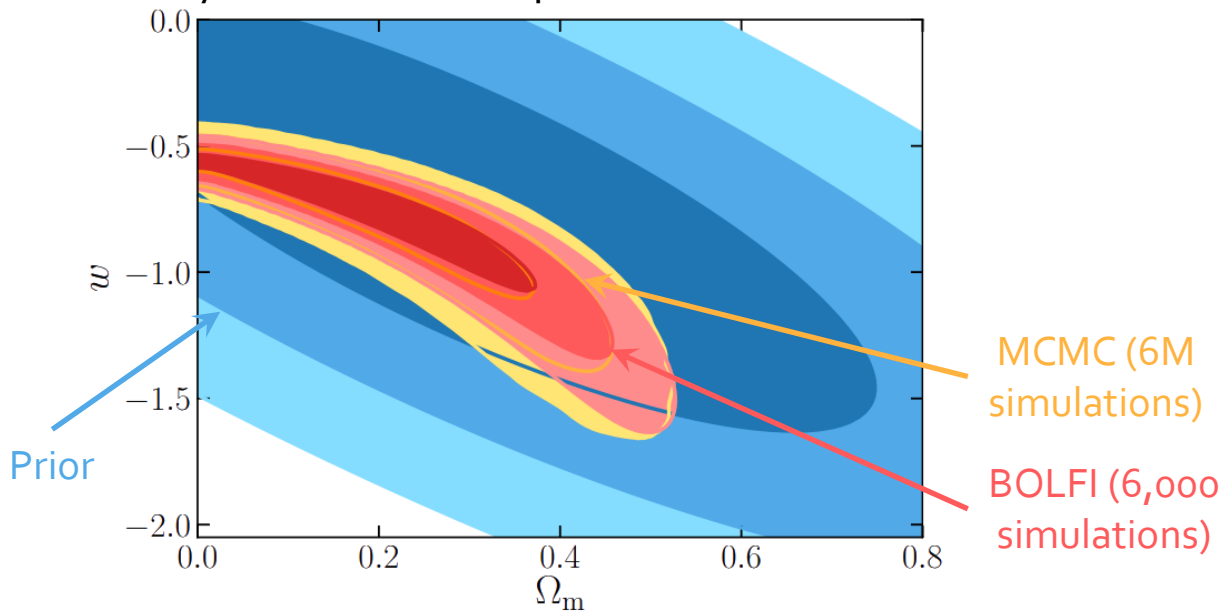
Sims: SLICS, Theory: CosMomentum

Lina Castiblanco Tolosa, *Cora Uhlemann*, Joachim Harnois-Déraps

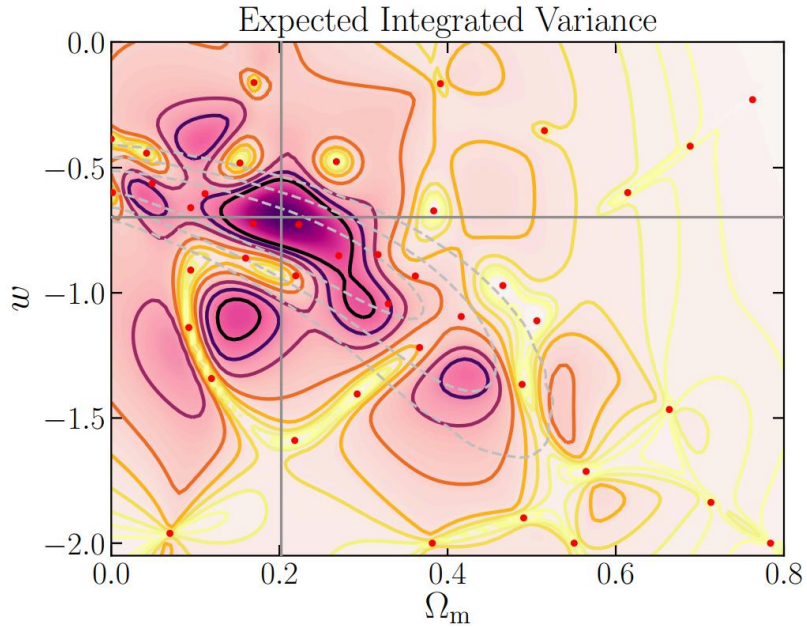
Dealing with expensive simulators in implicit likelihood inference (ILI) problems: The BOLFI algorithm (*Bayesian Optimisation for Likelihood-Free Inference*)

- The simulator will typically be extremely expensive (*N*-body simulation, halo finding, complex observational effects). We can typically afford O(10,000) evaluations.
- Emulation of the data model is not the only option.

Re-analysis of the JLA supernovae data:



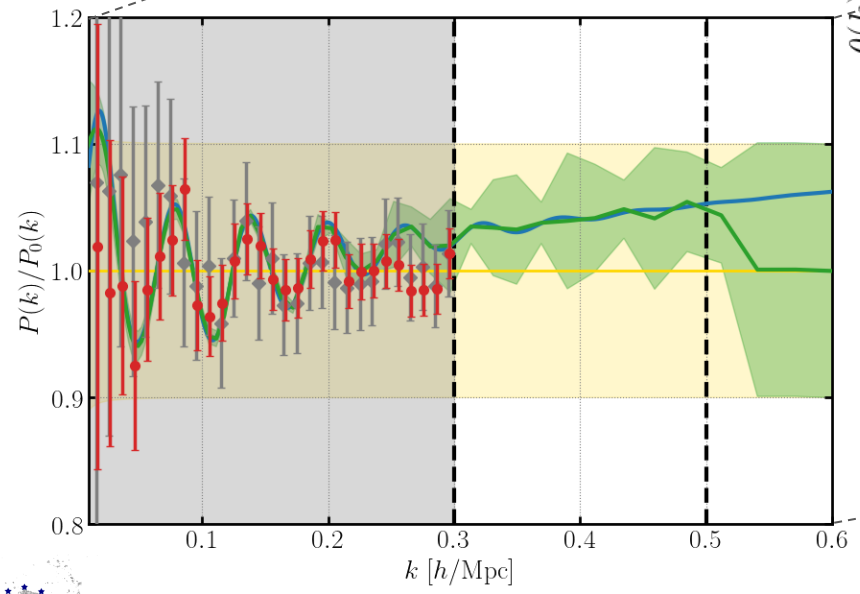
- BOLFI (*Bayesian Optimisation for Likelihood-Free Inference*) uses an acquisition function to place expensive simulations in the parameter space.
- The optimal acquisition function for implicit inference can be derived: the Expected Integrated Variance.

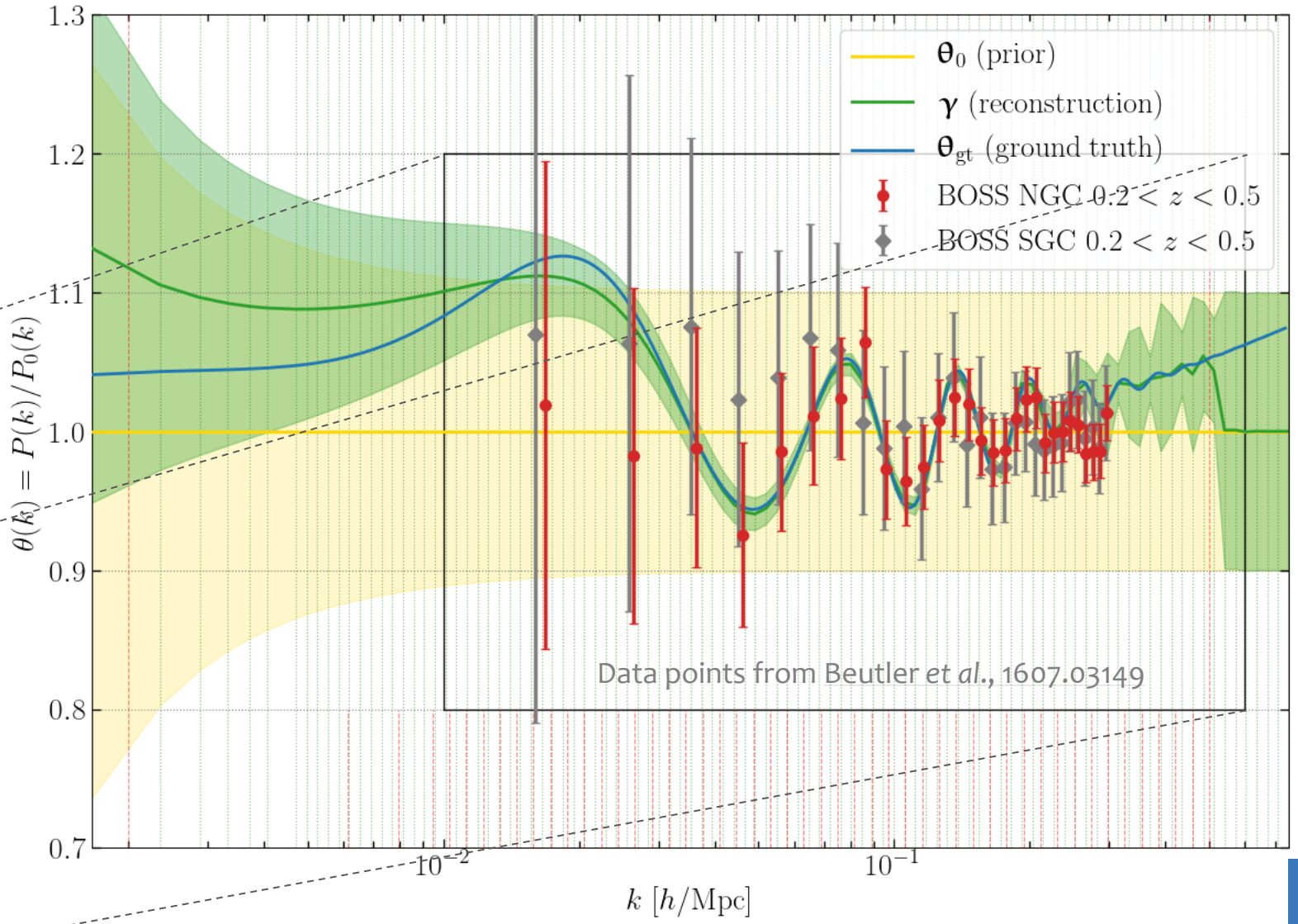


Leclercq, 1805.07152

SELFI (*Simulator Expansion for Likelihood-Free Inference*): ILI of the initial power spectrum
Euclid forecast vs BOSS data

- Numerical data models allow using the galaxy power spectrum as summary statistics up to at least $k \gtrsim 0.5\ h/\mathrm{Mpc}$ safely
- $N_{\mathrm{modes}} \propto k^3$: 5 times more modes are used in the analysis.

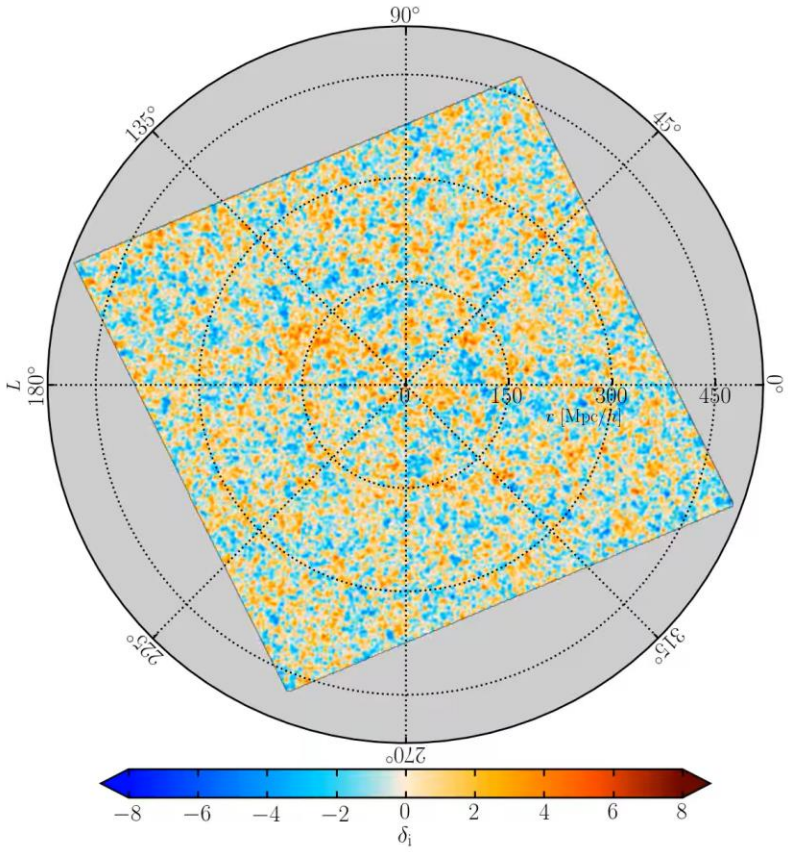




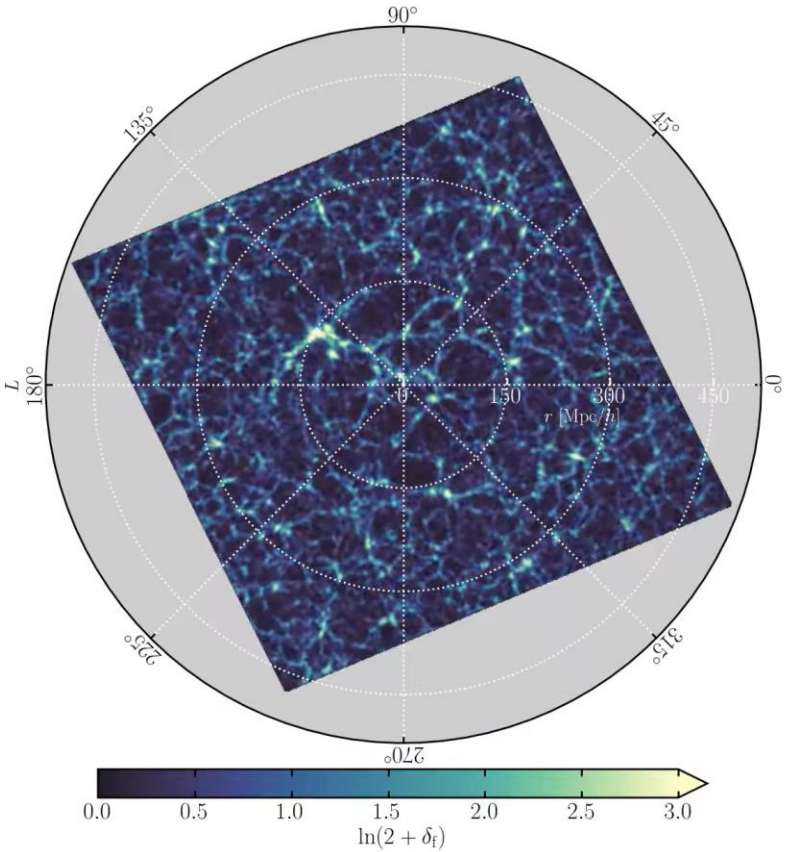
Leclercq *et al.*, 1902.10149, Leclercq, 2209.11057, Hoellinger & Leclercq, in prep.

Inference with an explicit field-level likelihood: The BORG algorithm (*Bayesian Origin Reconstruction from Galaxies*)

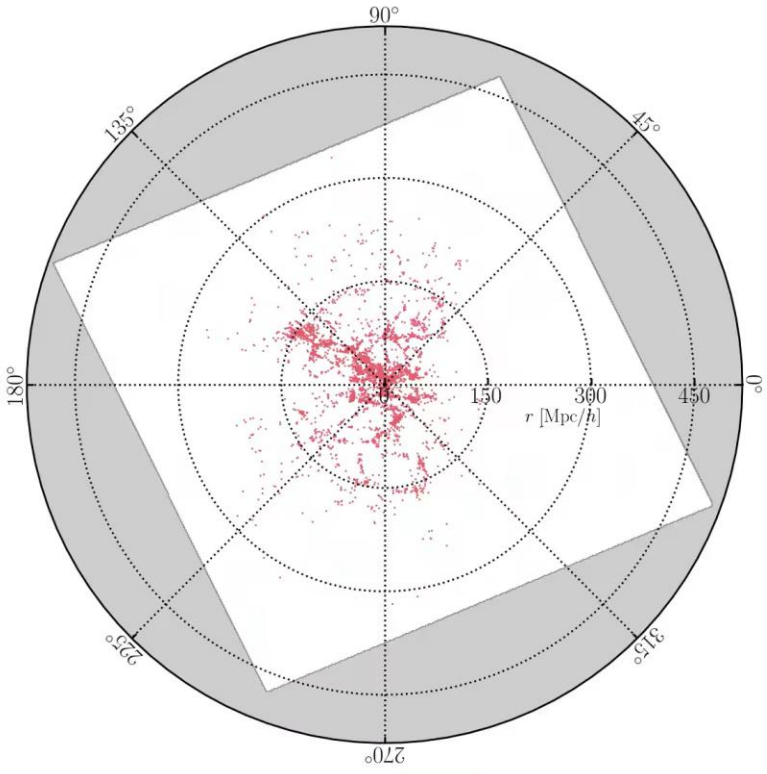
Initial conditions



Final conditions



Observations

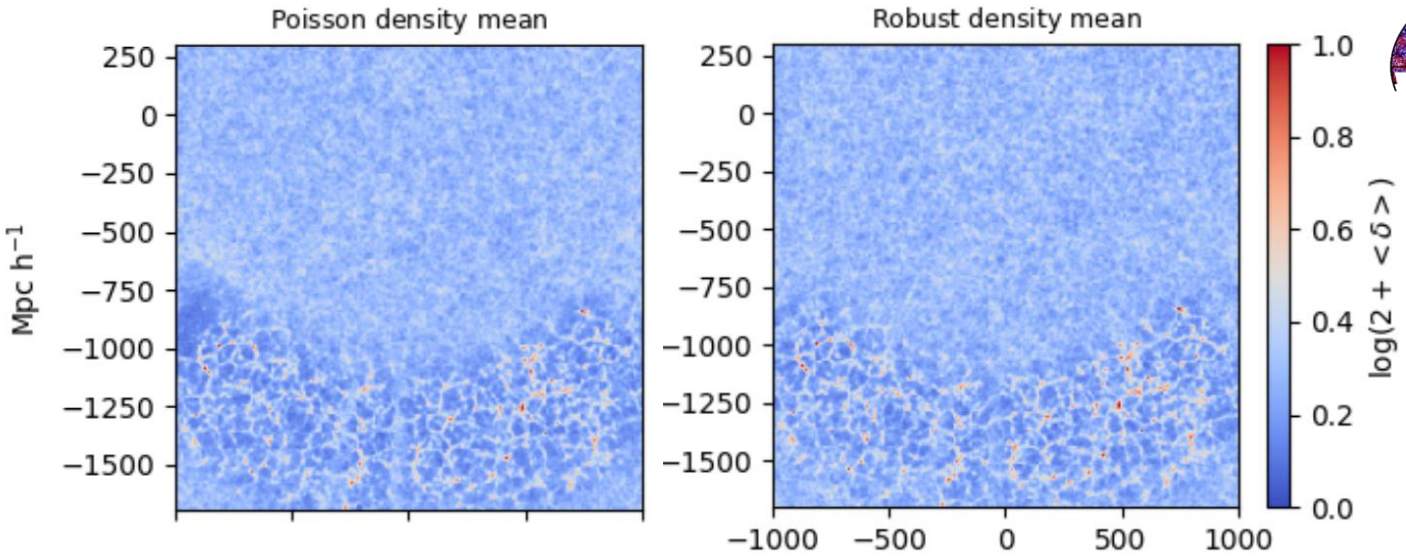


67,224 galaxies, ≈ 17 million parameters, 5 TB of primary data products, 10,000 samples, ≈ 500,000 forward and adjoint gradient data model evaluations, 1.5 million CPU-hours

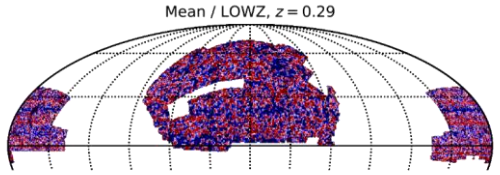
Jasche & Wandelt, 1203.3639; Jasche, Leclercq & Wandelt, 1409.6308; Jasche & Lavaux, 1806.11117; Lavaux, Jasche & Leclercq, 1909.06396

Model misspecification and unknown systematics with an explicit field-level likelihood

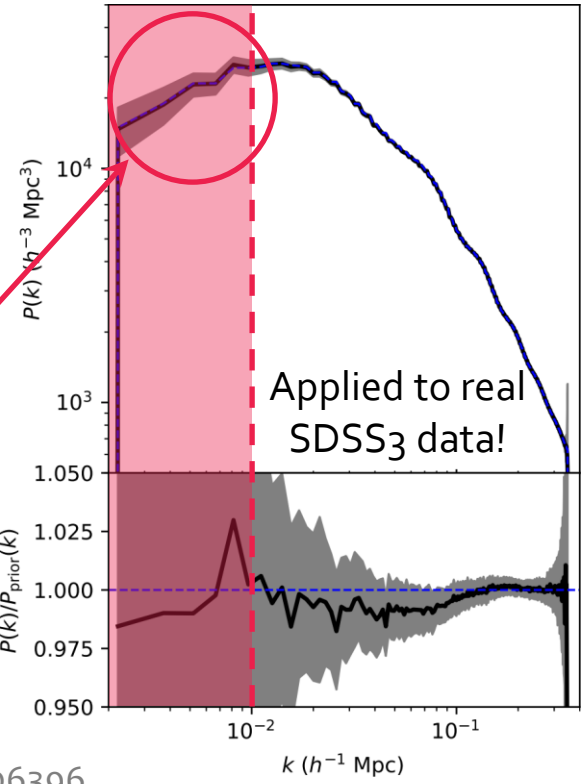
- Systematic effects are an issue of model misspecification: when the model differs from the actual data-generating process, posteriors tend to be biased and/or overly concentrated.
- In cosmology, we are sometimes unable to formulate ***any*** model that fits the data in some regimes.
- Machine-aided report of unknown systematic effects is possible with an explicit field-level likelihood (BORG):



Porqueres *et al.*, 1812.05113



A map of an unknown foreground



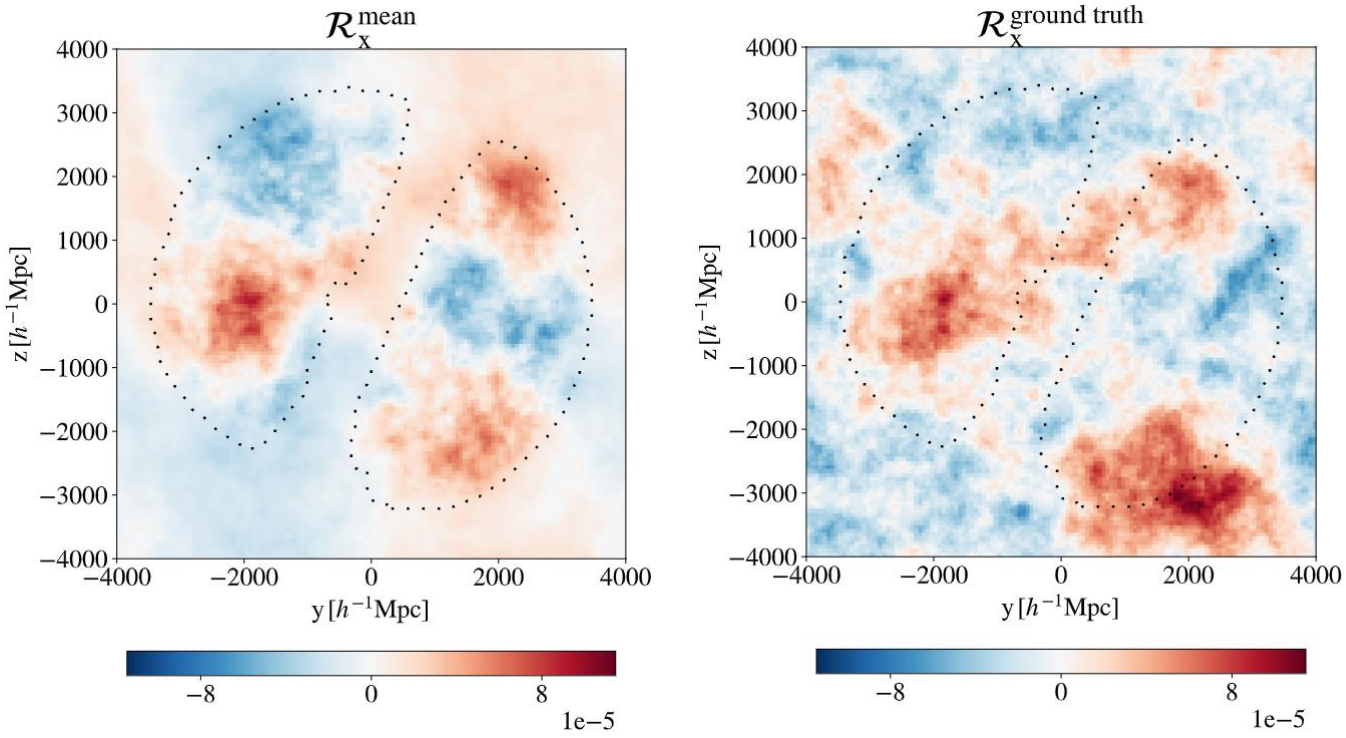
No apparent contamination, even well beyond the turn-over

Lavaux, Jasche & Leclercq, 1909.06396

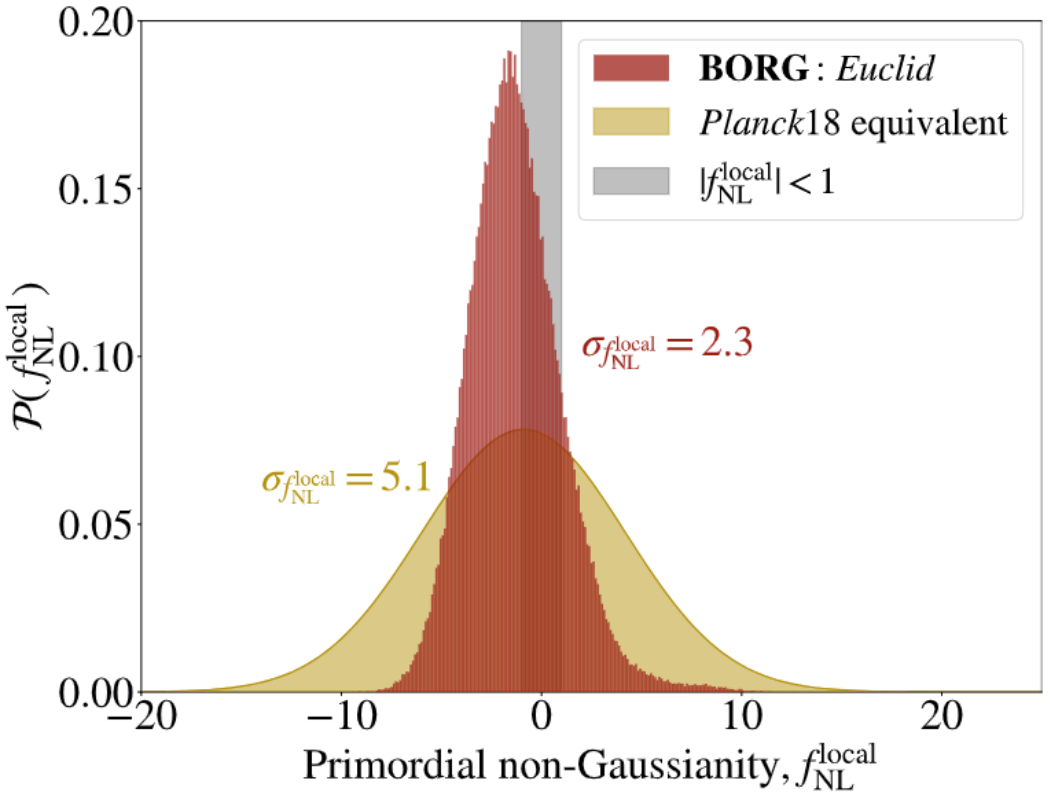
Field-based primordial physics: joint inference of primordial non-Gaussianity and initial conditions

- The physical model can be extended (as long as it is differentiable), e.g. with primordial physics:

Reconstruction of adiabatic curvature fluctuations



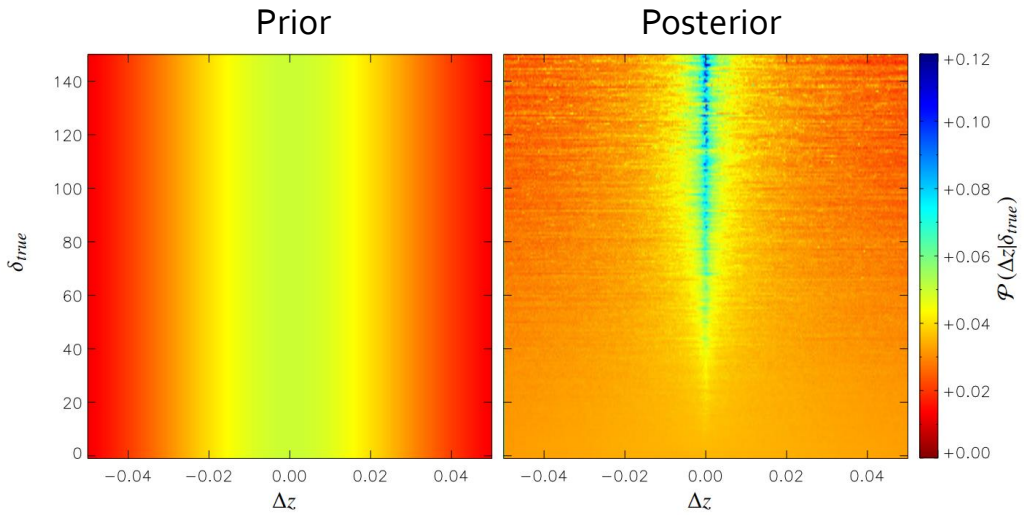
Posterior on $f_{\text{NL}}^{\text{local}}$

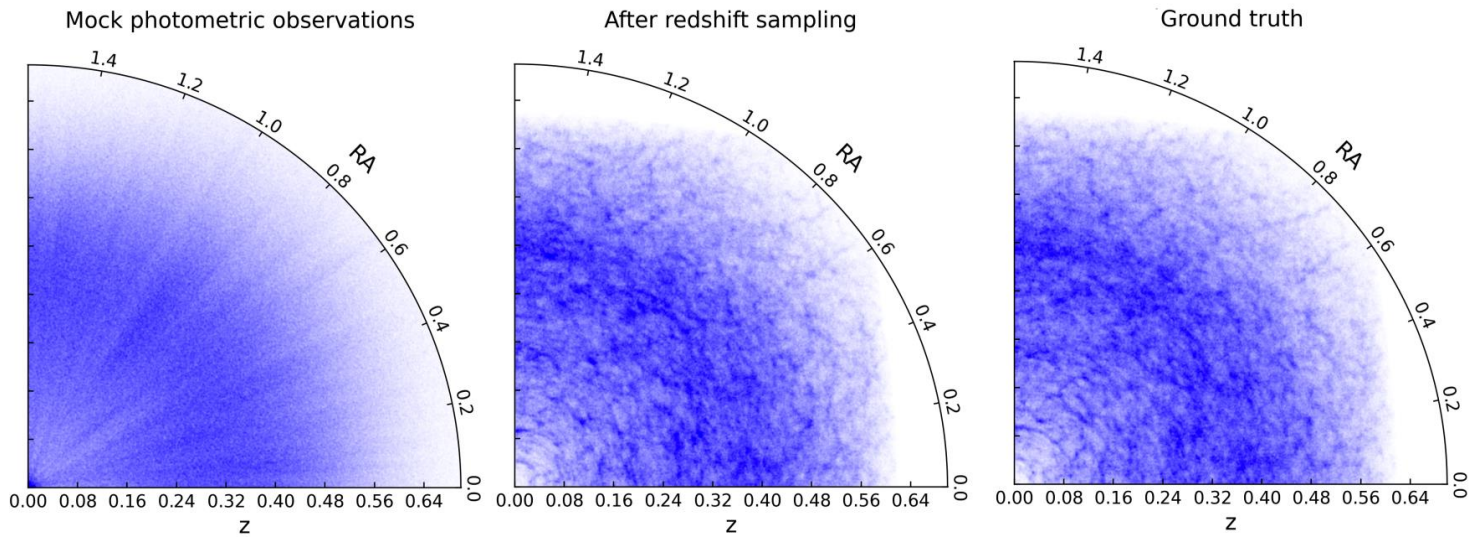


Andrews *et al.*, 2203.08838; Andrews *et al.*, in prep. (Euclid TWG WP4)

Field-based observational uncertainties: joint inference of photometric redshifts and density fields

- Sampling redshifts conditional on the density field sharpens the redshift pdfs…
- and propagates photometric galaxy clustering information to the density field reconstruction.

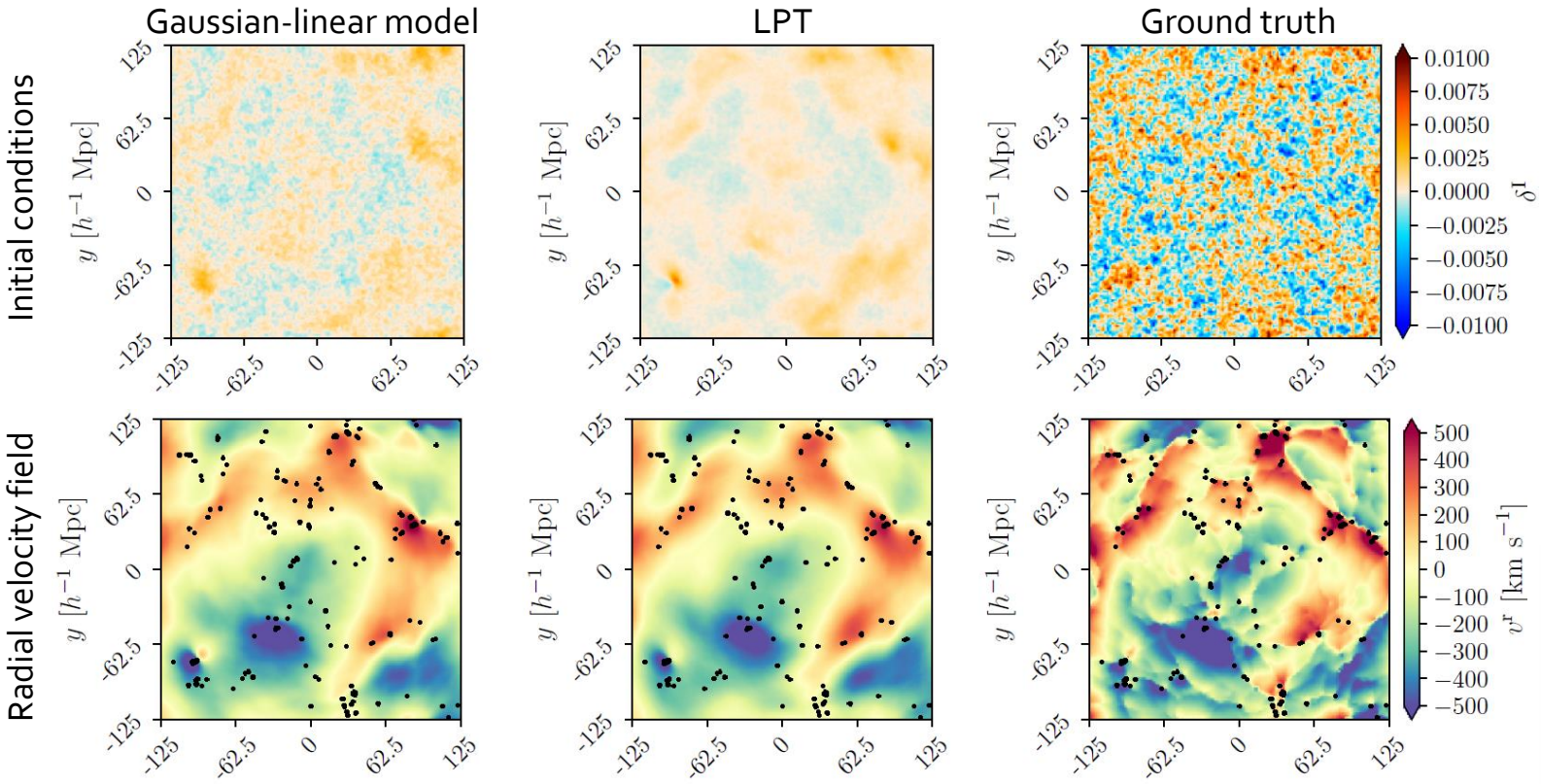




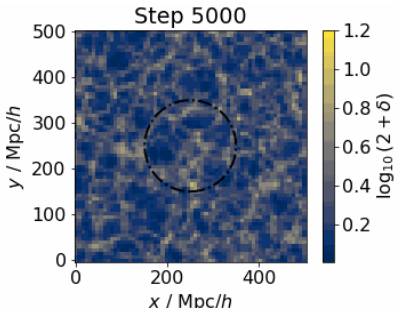
Jasche & Wandelt, 1106.2757; Tsaprazi, Jasche, Lavaux & Leclercq, 2301.03581

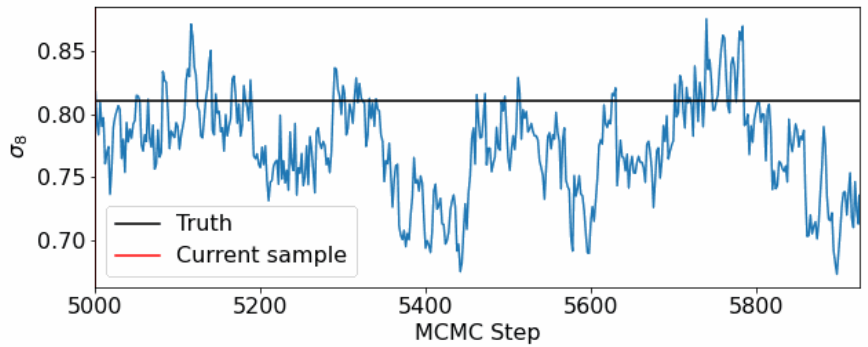
Field-level multi-tracer approach: joint initial conditions and velocity field reconstruction using distance tracers

- A field level approach naturally extends to muti-tracer / multi-wavelength / multi-messenger cosmology.



- New model in development:
  - Homogeneous and inhomogeneous Malmquist bias
  - Non-linear gravity (Lagrangian perturbation theory, ++)
  - Cosmological parameter sampling ($f\sigma_8$)





Lavaux, 1512.04534; Boruah, Lavaux & Hudson, 2111.15535; Prideaux-Ghee, Leclercq, Lavaux, Heavens & Jasche, 2204.00023; Bartlett *et al.*, in prep. See also work from Hoffman, Courtois, Sorce & CLUES team

Take-home messages

- Do not forget additional probes – they carry information needed to get the best science out of Euclid.
- Do not neglect alternative methods – they increase robustness to systematics and answer new questions.
- More details during Splinter 7 on Thursday 1 February at 10:30:
  - Guilhem Lavaux – Constraining primordial non-Gaussianities with field level inference
  - Axel Lapel – Constraining the modified gravity landscape through Bayesian forward modeling of cosmic structures
  - Simon Ding – Fast and differentiable mock catalogues for wide galaxy surveys using physical networks
  - Florent Leclercq – Implicit Likelihood Inference while efficiently checking for survey systematics