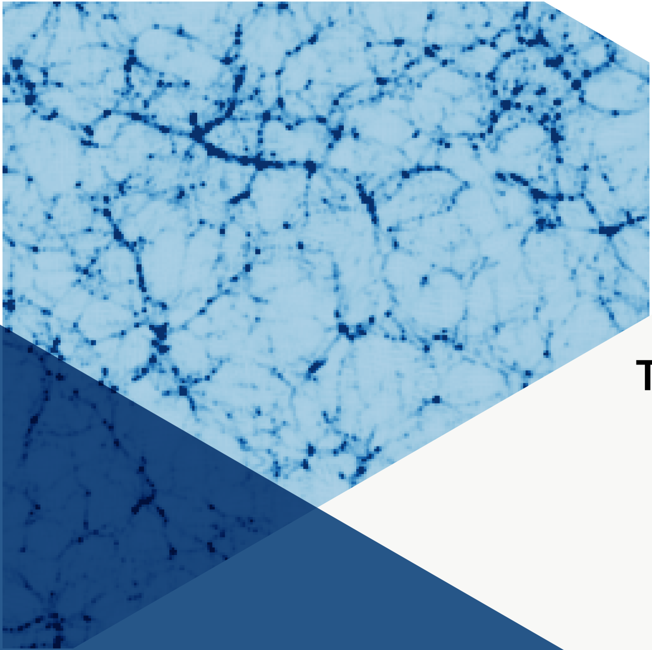


PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

The Initial Shape of the Universe

for the focus week at IFPU

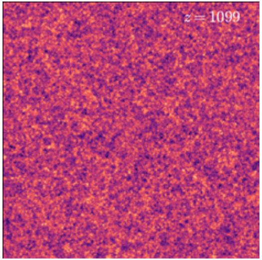
Juan M. Calles H.

July 1, 2022

What?

▶Q: What is the nature of the primordial perturbations that set up the structure formation of the Universe?

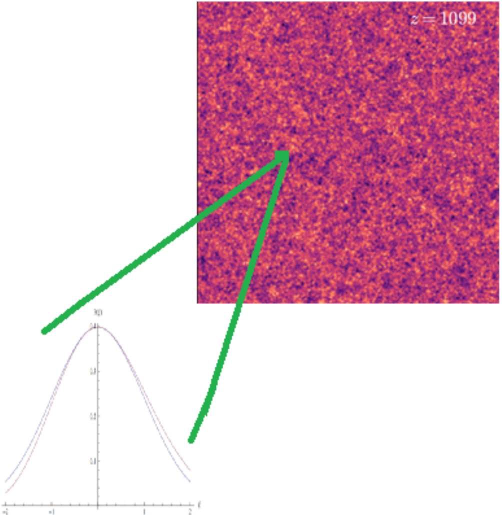
Einstein Toolkit. arXiv: 1807.01711



What?

▶Q: What is the nature of the primordial perturbations that set up the structure formation of the Universe?

Einstein Toolkit. arXiv: 1807.01711



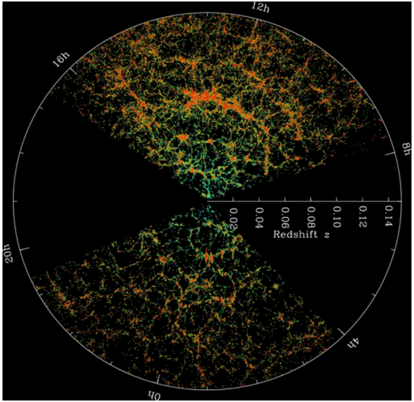
Why?

▶ Which is the primordial mechanics that generates the seeds of density perturbations?

▶ In the inflationary scenario, these perturbations come from quantum fluctuations of scalar fields in an expanding background.

▶ Primordial non-Gaussianity is a key aspect of these initial conditions, and provides an important link between high energy physics and cosmology.

Where?



What can cosmological observation tell us about the initial conditions and the dynamics of the Universe?

▶Q:What is the maximun amount of information we can extract from cosmological observations?

Plan of the talk

We want to **explore the potential of topological data analysis in detecting primordial non-Gaussianity** through observations of the large scale structures of the universe.

▶ **Based on arXiv: 2203.08262 [astro-ph.CO] [1]. Matteo Biagetti, JC, Lina Castiblanco, Alex Cole, and Jorge Noreña.**

▶ **Related to arXiv: 2009.04819 [astro-ph.CO] [2]. Matteo Biagetti, Alex Cole, and Gary Shiu.**

Overview

1. Introduction
2. Persistence Homology Pipeline
3. Future work and Summary

Introduction

Initial conditions from the late Universe

- Dealing with galaxies as bias tracer of the underlying matter distributions.
- Contamination from gravitational evolution leads to non-Gaussianities in the late Universe.
- Redshift Space Distortion.
- Large scale probes: Newtonian vs Relativistic.

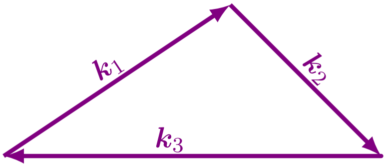
How to measure?

- Typically N-point correlation functions are used to extract information about the statistical distribution of primordial perturbations. Being the lowest order the powerspectrum,

$$P(k) \propto \langle \Phi_{\boldsymbol{k}} \Phi_{\boldsymbol{k'}} \rangle \tag{1}$$

- The powerspectrum only give us information about the free field theory during inflation.
- The next leading statistics is the bispectrum,

$$B(\boldsymbol{k_1}, \boldsymbol{k_2}, \boldsymbol{k_3}) \propto \langle \Phi_{\boldsymbol{k_1}} \Phi_{\boldsymbol{k_2}} \Phi_{\boldsymbol{k_3}} \rangle \tag{2}$$



Non-Gaussian Initial Conditions I

- Truly Gaussian seeds lead to a banishing primordial bispectrum.

- For the powerspectrum of a non-Gaussian Field

$$P(k) \propto P^I(k) + 2\int d\boldsymbol{k}_2 F_2(\boldsymbol{k}_2, -\boldsymbol{k}-\boldsymbol{k}_2) B^I(\boldsymbol{k}, \boldsymbol{k}_2, -\boldsymbol{k}-\boldsymbol{k}_2) + \cdots \quad (3)$$

- This information can be categorized into two features of the primordial bispectrum:
  1. The shape, the relation between the momentas.
  2. The amplitude, usually denoted as f_{NL}.

Non-Gaussian Initial Conditions II

- Different physical scenarios generate distinguishable signal that peaks at different shapes of the bispectrum.
- **Local type.**

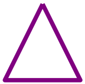
$$B_\phi^{\text{local}}(k_1, k_2, k_3) = \frac{6}{5} f_{\text{NL}}^{\text{loc}} \Big[P_\phi(k_1) P_\phi(k_2) + 2 \text{ perms.} \Big], \tag{4}$$

1. Bispectrum peaks at the squeezed configuration. 
2. Measurements of $f_{\text{NL}}^{\text{loc}} > 1$ would rule out single field slow-roll inflation.

Non-Gaussian Initial Conditions III

- **Equilateral type.**

$$B_\phi^{\rm equi}(k_1,k_2,k_3)=\frac{18}{5}f_{\rm NL}^{\rm equi}\Big[-P_\phi(k_1)P_\phi(k_2)-2\ \text{perms.}-2P_\phi^{2/3}(k_1)P_\phi^{2/3}(k_2)P_\phi^{2/3}(k_3)+ P_\phi^{1/3}(k_1)P_\phi^{2/3}(k_2)P_\phi(k_3)+5\ \text{perms.}\Big]\,. \tag{5}$$

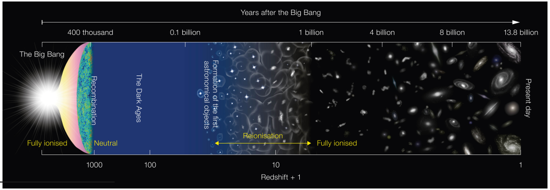


1. Bispectrum peaks at the equilateral configuration.
2. Genereted by alternative inflationary models.

- Given a Bispectrum one can reconstruct the primordial field and implement it in a Cosmological Simulation to analyze its evolutions.

Current constrain

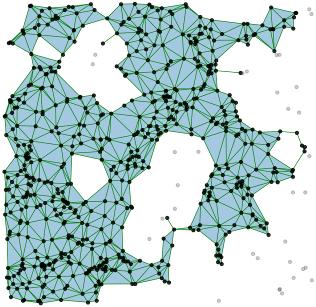
- **Planck collaboration**[3]: for the **local type** $f_{\rm NL}^{\rm loc} = (-0.9 \pm 5.)$, and for the **equilateral type** $f_{\rm NL}^{\rm equi} = (-26 \pm 47)$ at (68% CL, statistical).
- **eBOSS collaboration** [4]: for the **local type** $-51 < f_{\rm NL}^{\rm loc} < 21$ at (95% CL, statistical).
- The next generation of surveys like SPHEREx, LSST and Euclid are expected to approach $f_{\rm NL}^{\rm loc} = \mathcal{O}(1)$ and $f_{\rm NL}^{\rm equi} = \mathcal{O}(50)$ [5][1].



[1] Karagiannis et al, 1801.09280

Why topological data analysis?

- Complementary to the usual information accessed by low-order correlation functions.
- Characterize cosmological information of real-space galaxy maps, via features such as clusters, filament, loops, and cosmic voids.
- Q: Can we predict deviation from Gaussianity based on topological features?



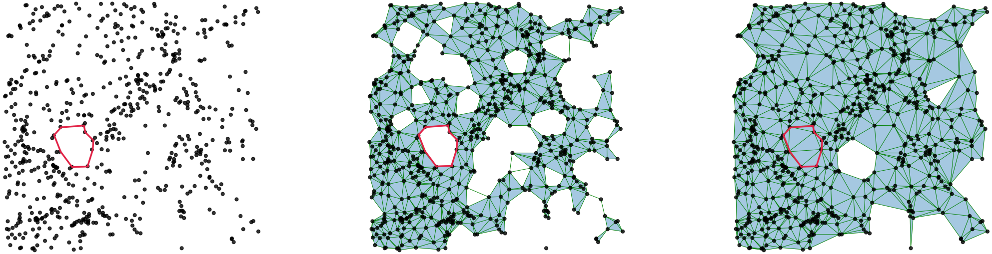
The Exploring the Origin of Structure
EOS DATASET[2]

- This is a suite of Halo catalogs generated using full N-body simulations with Gaussian and non-Gaussian initial conditions at large scale.
- We focus on Local non-Gaussian with $\boldsymbol{f_{\mathrm{NL}}^{\mathrm{loc}} = 10}$ and Equilateral non-Gaussian with $\boldsymbol{f_{\mathrm{NL}}^{\mathrm{equi}} = -30}$ types of initial conditions for our analysis.
- With flat $\boldsymbol{\Lambda}$**CDM** cosmology and cosmological parameter: $\sigma^8 = 0.85$, $h = 0.7$ and $\Omega_m = 0.3$.
- Total volume simulation for each cosmology of 120 (Gpc/h)^3.

[2] More information available at `https://mbiagetti.gitlab.io/cosmos/nbody/eos/`

Intuition of Persistent Homology I

- We start from a collection of point, such as halo positions, and we want to **characterize the structure of the data via its relevant features**, like voids, filaments, and clusters, as a function of scale.

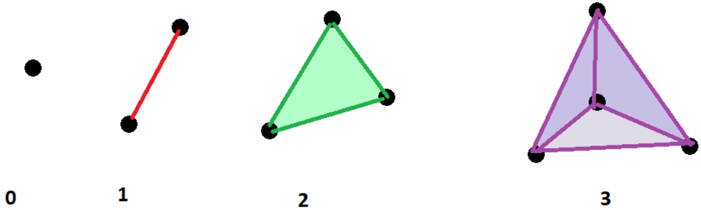


Intuition of Persistent Homology II

- **Track each feature by corse-graning at different increasing scales**, e.g. track the distribution birth and death of the features as a function of scale.
- Understanding the underlying geometric structure of the data. **Separate noise from real features via the persistence** of the feature, e.g. it's "life's spawn."
- Q: **How this distribution reacts to changes in initial conditions w.r.t primordial non-Gaussianity?**

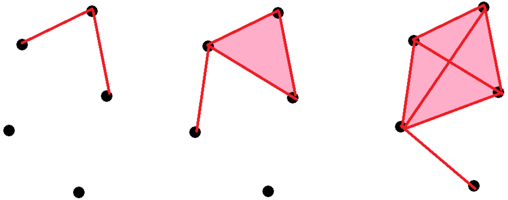
Basics of Persistent Homology I

- Define topological feature such as a connected component, a hole or a cavity on the data.
- Going from points to features: we define simplices as the smallest shape that enclose all the points.



Basics of Persistent Homology II

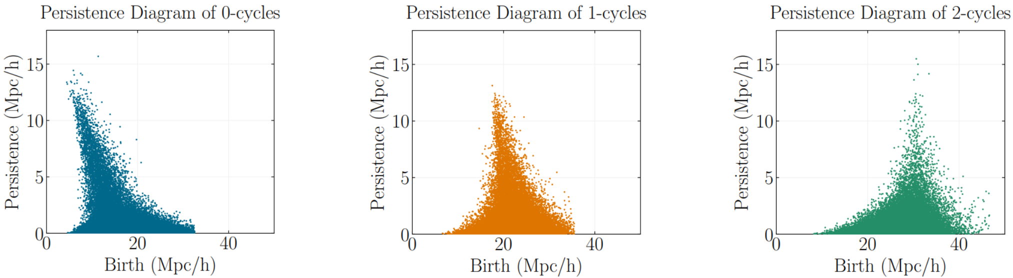
- Connecting the simpleces in the point cloud to form a simplicial complex.
- We include "scale" via Filtration. Which is growing family of simplicial complexes of sub-sequence subcomplex, often roughly describe the structure of the point cloud.



- For each nested sequence we record when each topological feature is created or destroyed in that scale.

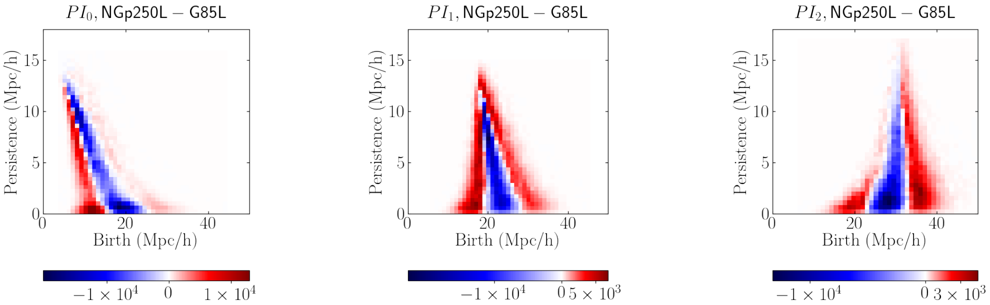
Persistent Diagrams

- We build scatter plots of the creation an annihilation of individuals features contributing to a particular connected component as a function of scale.
- Gain physical intuition and visualizing the density of topological features.
- 0-cycles: clusters of halos, 1-cycles: loops from connecting halos, and 2-cycles: voids.



Persistent Images

- Smoothed histograms of a persistence diagram.
- Identify deviation of primordial local non-Gaussianity between $f_{\mathrm{NL}}^{\mathrm{loc}} = 250$ and $f_{\mathrm{NL}}^{\mathrm{loc}} = 0$ simulations.



Challanges

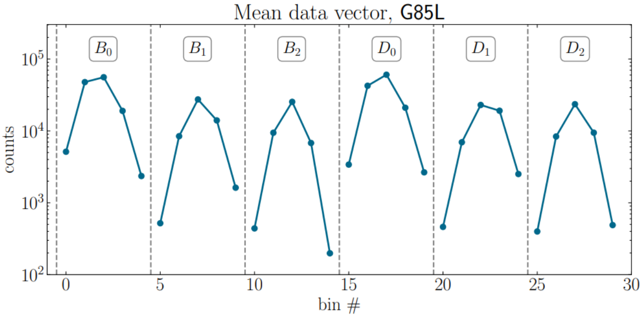
- Build a reliable estimator that quantify the deviation from primordial non-Gaussianity.
- Persistence Diagram has limited constraining power due to the large amount of information.
- Choose a statistical representation to build covariance matrix.
- Computational cost, large simulation volumes.

Persistence Homology Pipeline

Summary Statistics I

- Define a compressed representation of histograms of births and deaths for each of the cycles.
- Total of 120 vectors per cosmology, corresponding to the 15 realizations divided into 8 sub-boxes.
- The data vector bins has to be Gaussian distributed for a reliable Fisher matrix.

Summary Statistics II



- The mean data vector has 30 bins and typical scale size between $\sim 10 - 40$ Mpc$/h$ and at least $\mathcal{O}(200)$ features per bin.

Towards a Fisher Matrix I

- For our mean data vector D we construct the Fisher matrix via,

$$F_{ij} = D_{,i}^T C^{-1} D_{,j} + \frac{1}{2}\mathrm{Tr}\left[(C^{-1}C_{,i})(C^{-1}C_{,j})\right] \tag{6}$$

  we neglect the second term.

- Numerical derivatives are defined with respect to the paramater $\theta_i = f_{\mathrm{NL}}^{\mathrm{loc}}, f_{\mathrm{NL}}^{\mathrm{equi}}$

$$D_{,i} = \frac{D(\theta_i) - D(\theta_i = 0)}{\theta_i} \tag{7}$$

Towards a Fisher Matrix II

- The covariance matrix C

$$C_{kl} = \frac{1}{N_{sim} - 1} \sum_{i=1}^{N_{sim}} (D_k - D_{mean})(D_l - D_{mean})^T \tag{8}$$

where runs $\{k, l\}$over 30 bins and the average is over 120 realizations.

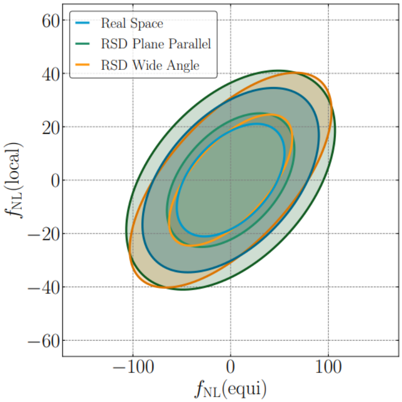
- Marginalized information on model parameter θ_i corresponds to a $1 - \sigma$ constraint of

$$\sigma_i = \sqrt{(F^{-1})_{ii}} \tag{9}$$

Redshift Space Distortion Fisher Matrix

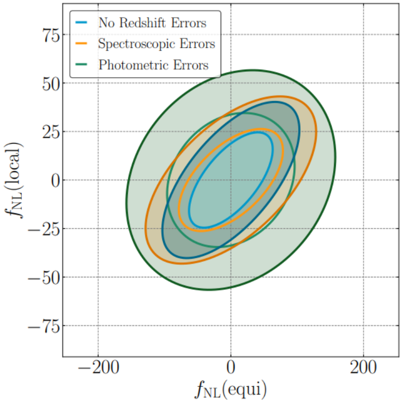
Findings!!

- We compare halos in redshift space in two scenarios:
  - displace halos along $\hat{z}$ of the axes parallel.
  - displace them along the vector between the observer (at the origin of the box) and the halo.
- $\Delta f_{\rm NL}^{\rm loc} \sim 16$ and $\Delta f_{\rm NL}^{\rm equi} \sim 41$ wide angle observer.



... Including redshift error

Findings!!



- We translate redshift uncertainties with velocity uncertainties at fixed redshift in the box.
- Spectroscopic sample:

$$\boldsymbol{s} = \boldsymbol{x} + \frac{v}{\mathcal{H}}\hat{\boldsymbol{n}} + \frac{v_{spec}}{\mathcal{H}}\hat{\boldsymbol{n}} \tag{10}$$

degradation $\sim 20\%$.

- Photometric sample:

$$\boldsymbol{s} = \boldsymbol{x} + \frac{v_{photo}}{\mathcal{H}}\hat{\boldsymbol{n}} \tag{11}$$

degradation $\sim 50\%$.

Primordial non-Gaussianity constrain

Findings!!

- Uncertainty on $f_{\rm NL}^{\rm loc}$ and $f_{\rm NL}^{\rm equi}$ obtained from the Fisher matrix for a fiducial cosmology with Gaussian initial conditions.

	Real Space	RSD Plane Parallel	RSD Wide Angle	$+v_{\rm spec}$-error	$+v_{\rm photo}$-error
$\Delta f_{\rm NL}^{\rm loc}$	13.9	16.5	16.2	17.3	22.8
$\Delta f_{\rm NL}^{\rm equi}$	36.3	43.0	41.6	51.8	63.4

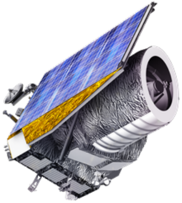
Future work and Summary

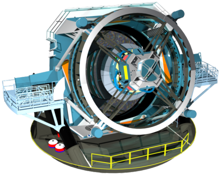
Towards a realistic covariance matrix

- Need of a large simulation volume. Computational cost.
- Do small-box simulations, since topological features are $\sim 10 - 40$ Mpc$/h$.
- Fast simulations. Does not fully reproduce Halo distribution of large scale simulations.

Real data constraints

Euclid





LSST

- Make contact with observation, including window functions, survey geometry.
- Response of the pipeline against small scale physics.

Summary I

- Persistence Homology provides a complementary analysis to conventional approach.

- For a conservative approach we forecast $\Delta f_{\rm NL}^{\rm loc} \sim 16$ and $\Delta f_{\rm NL}^{\rm equi} \sim 41$ in redshift space without the distant observer approximation

- The wide angle observer is trivial to implement compared with Powerspectrum prediction [6].

- Redshift Space Distortion and Uncertainties are easy to implement and leads to a small degradation of the constraints

Summary II

.

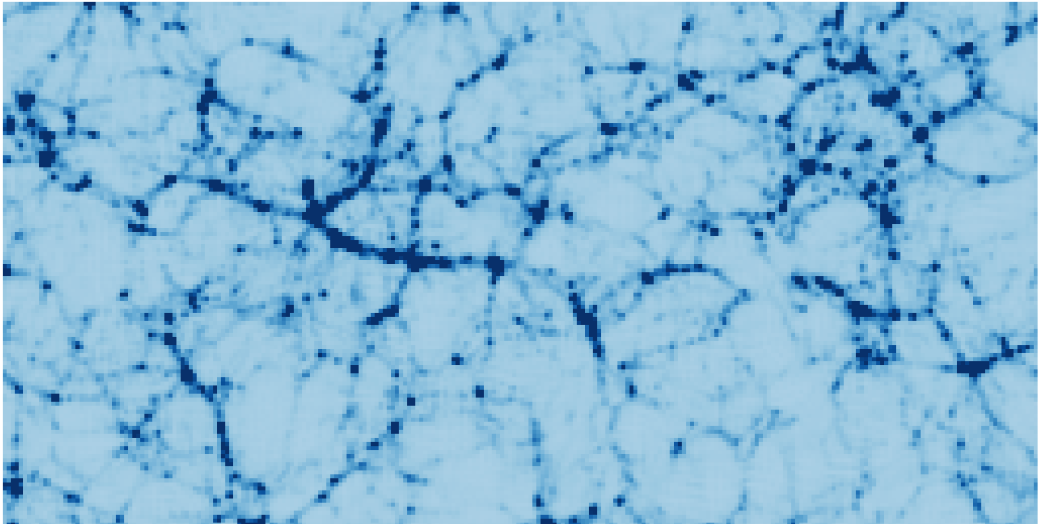
- Spectroscopic sample: $\Delta f_{\rm NL}^{\rm loc} \sim 17.3$ and $\Delta f_{\rm NL}^{\rm equi} \sim 51.8$
- Photometric sample: $\Delta f_{\rm NL}^{\rm loc} \sim 22.8$ and $\Delta f_{\rm NL}^{\rm equi} \sim 63.4$
- Competitive constrain compared with the ones from Plannk's and future LSS forescast. [5][3]

	Planck	Radio continuum, 1 μJy	Radio continuum, 10 μJy	Spectroscopic	Photometric
Local	5.0	0.2	0.6	1.3	0.3
Equilateral	43	244	274	57	184
Orthogonal	21	18	29	18	38

Table 12. Summary of 1σ limits for the three PNG types considered, from radio continuum and optical surveys derived from combining the power spectrum and bispectrum and accounting for RSD, the trispectrum term and theoretical errors. See text for more details.

[3] Karagiannis et al, 1801.09280

Thank you



References I

- M. Biagetti, J. Calles, L. Castiblanco, A. Cole and J. Noreña, *Fisher Forecasts for Primordial non-Gaussianity from Persistent Homology*, 2203.08262.
- M. Biagetti, A. Cole and G. Shiu, *The Persistence of Large Scale Structures I: Primordial non-Gaussianity*, *JCAP* **04** (2021) 061 [2009.04819].
- Planck collaboration, *Planck 2018 results. IX. Constraints on primordial non-Gaussianity*, *Astron. Astrophys.* **641** (2020) A9 [1905.05697].
- E. Castorina et al., *Redshift-weighted constraints on primordial non-Gaussianity from the clustering of the eBOSS DR14 quasars in Fourier space*, *JCAP* **09** (2019) 010 [1904.08859].

References II

- D. Karagiannis, A. Lazanu, M. Liguori, A. Raccanelli, N. Bartolo and L. Verde, *Constraining primordial non-Gaussianity with bispectrum and power spectrum from upcoming optical and radio surveys*, *Mon. Not. Roy. Astron. Soc.* **478** (2018) 1341 [1801.09280].
- E. Castorina and M. White, *Wide angle effects for peculiar velocities*, *Mon. Not. Roy. Astron. Soc.* **499** (2020) 893 [1911.08353].