MNRAS **000**, 1–2 (2018) Preprint 7 March 2018 Compiled using MNRAS LATEX style file v3.0

Notes On Abacus BAO Analysis

Yutong Duan [1]★

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

This serves as detailed notes on the procedures of BAO analysis with AbacusCosmos to accompany the code. Part 1 is on the calculation of correlation functions and covariance matrices. Part 2 is on the fitting methods of the BAO fitter.

Key words: keyword1 – keyword2 – keyword3

1 INTRODUCTION

Precision measurement of the BAO signal from galaxy correlation functions requires understanding the systematics. The standard approach is producing many mock catalogues of galaxies/quasars, test the analysis pipelines, and evaluate the sensitivity to systematics. The pipelines largely consist of two parts, statistics and fitting. The statistics of galaxy catalogues are correlation functions and covariance matrices. Then the statistics are fed to the fitting procedures, yielding best-fit α, the BAO scale parameter with respect to the fiducial model.

2 CORRELATION FUNCTIONS AND COVARIANCE MATRIX

2.1 Reading Halo Catalogue

Given a cosmology and redshift, there are 16 boxes with different phases. For each phase, load the halo catalogue, and in the halo table, add the following columns:

$$m_{\text{halo, vir}} = m_{\text{halo}} \tag{1}$$

$$r_{\text{halo, vir}} = r_{\text{halo}} \tag{2}$$

$$c_{\text{NFW}} = \frac{R_{\text{vir}}}{r_{\text{R, klypin}}} \tag{3}$$

This is to keep the Abacus catalogues compatible with halotools, where these fields are assumed present. The NFW profile $\rho(r)$ describes dark matter density as a function of radius in the halo. NFW formula is just a very simple way of defining the halo concentration c_{NFW}, which fixes $\rho(r)$. The Klypin definition of the halo scale radius R_{s} is considered more stable than the usual R_{s}.

2.2 Populating Halos with Galaxies

There are many Halo Occupation Distribution (HOD) models available which put synthetic galaxies in the dark matter halos. Some mainstream ones are `['zheng07', 'leauthaud11', 'tinker13', 'hearin15', 'zu_mandelbaum15', 'zu_mandelbaum16', 'cacciato09']`.

There are a number of parameters that can be varied in each model. For a $(1100\,\text{Mpc/h})^3$ simulation box and `zheng07` model with magnitude threshold = -18, particle number cut = 1, and other default parameters, typical numbers are

$$N_{\text{halos}} = 8 \times 10^6 \tag{4}$$

$$N_{\text{galaxies}} = 1.5 \times 10^7. \tag{5}$$

For comparison, in BOSS DR12 there are 1.2×10^6 galaxies in three redshift bins, 0.38, 0.51, and 0.61.

2.3 Correlation Functions

All counting is done in fine (s, μ) bins. s bin edges are from 0 to 150 Mpc/h at 5 Mpc/h steps, and $\mu = \cos\theta$ bin edges are from 0 to 1 at 0.01 steps, meaning a total of (150, 100) bins. For the PH natural estimator, DD and RR are all we need.

Auto-correlation DD pair counts of the simulation box is saved as `paircount-DD.npy` in original Currfunc count format, a structured array. Note that `c_api_timer` needs to be turned off for this to be saved and recovered properly and not as "object" type.

RR pair counts can be calculated analytically as given in the appendix.

3 BAO FITTER

A recent, functional fitter is one by Ross https://github.com/ashleyjross/LSSanalysis used in BOSS DR12 for the paper on BAO in correlation functions. We rewrite the fitter in a modern style in Python conforming to PEP standards.

REFERENCES

Wang Y., Yang X., Mo H. J., van den Bosch F. C., Weinmann S. M., Chu Y., 2008, ApJ, 687, 919

★ E-mail: dyt@physics.bu.edu

APPENDIX A: AUTO-CORRELATION FUNCTION ESTIMATOR

In this implementation, raw pair counts are saved, and $N_{\rm D}$, $N_{\rm R}$ normalisation is done only at the estimator step.

For auto-correlation of a sample in periodic box, the Peebles & Hauser (1974) estimator (natural estimator) is

$$\xi = \frac{N_{\rm R}(N_{\rm R}-1)}{N_{\rm D}(N_{\rm D}-1)}\frac{DD}{RR} - 1 \tag{A1}$$

where DD is the auto-correlation data-data pair count and RR is the random-random pair count.

One way to obtain RR is to generate a random sample of particle number $N_{\rm R}$ in the same volume and calculate its auto-correlation pair counts. Alternatively, RR can be calculated analytically as follows. Let dV be the volume of the (s,μ) bin and $V = L_{\rm box}^3$ be the volume the data sample occupies. The random sample must have the same number density as the data sample, $n_{\rm R} = n_{\rm D}$. For simple survey geometry, such as a cube, we may well let the random sample have the same number count, volume, and number density as the data sample.

$$RR(s,\mu) = \frac{N_{\rm R}(N_{\rm R}-1)}{V}\,\mathrm{d}V(s,\mu) \tag{A2}$$

where $\mathrm{d}V(s,\mu)$ is the volume of the (s,μ) bin.

APPENDIX B: CROSS-CORRELATION FUNCTION ESTIMATOR

For cross-correlation between a sample D_1 in the periodic simulation box and a subsample D_2 in its subvolume, the Davis & Peebles (1983) estimator is

$$\xi = \frac{\bar{n}_{\rm R1}}{\bar{n}_{\rm D1}}\frac{D_1D_2}{R_1D_2} - 1 \tag{B1}$$

where $\bar{n}_{\rm D1}$ is the mean number density of data sample D_1, $\bar{n}_{\rm R1}$ is the mean number density of R_1, the random sample corresponding to D_1, D_1D_2 is the cross-correlation pair count between two data samples, and R_1D_2 is the cross-correlation pair count between a random and a data sample. Usually a random sample is about 7 times the size of the corresponding data sample (Wang et al. 2008).

Again one may generate a random sample R_1 in the same volume and do the cross counting, or alternatively calculate it analytically,

$$R_1D_2 = \frac{N_{\rm R1}N_{\rm D2}}{V_1}\,\mathrm{d}V(s,\mu) = \bar{n}_{\rm R1}N_{\rm D2}\,\mathrm{d}V(s,\mu). \tag{B2}$$

This paper has been typeset from a TEX/LATEX file prepared by the author.