Status & Plans for CMBXC Likelihood implementation

M. Lattanzi (on behalf of the BO-FE-PD groups)

Euclid CMBXC SWG Meeting – 04/11/19

Likelihood activity for Euclid CMBXC is coordinated by F. La Casa and P. Natoli

Two main (related but independent) activities

- Development and validation of spectral estimators (see talk by A. Gruppuso)
- Development and validation of a likelihood package (report in this talk)

Ongoing joint activity on the likelihood from the Bologna, Ferrara and Padova groups

Regular meetings, starting from past spring.

Three meetings so far, plan to increase frequency (~ 1 F2F meeting per month + regular telecons)

Bologna: M. Archidiacono, N. Mauri, L. Patrizii, G. Sirri, M. Tenti

Ferrara: A. Gruppuso (INAF-BO), M. Lattanzi, M. Migliaccio (RM2), D. Molinari, P. Natoli, L. Pagano, G. Polenta (ASI)

Padova: S. Dusini, A. Renzi, C. Sirignano, L. Stanco, G. Verza

Bayes' Theorem

parameters

likelihood

prior

$$\mathcal{P}(\theta|\mathbf{d}) = \frac{\mathcal{L}(\mathbf{d}|\theta)\mathcal{P}(\theta)}{\int \mathcal{L}(\mathbf{d}|\theta)\mathcal{P}(\theta)d\theta}$$

data

posterior

evidence

- The likelihood $L(d|\theta)$ plays a central role in all inferences that we make from observations
- In particular it is important for both parameter estimation and model selection
- It should describe how data depend on the model (this includes both the theoretical model, instrument properties, systematics…)
- Aim: build a likelihood for Euclid CMBXC (will focus on ISW in the following)
- When building a likelihood algorithm, we are faced with many choices

I. The data

parameters

likelihood

prior

$$\mathcal{P}(\theta|\mathbf{d}) = \frac{\mathcal{L}(\mathbf{d}|\theta)\mathcal{P}(\theta)}{\int \mathcal{L}(\mathbf{d}|\theta)\mathcal{P}(\theta)d\theta}$$

data

posterior

evidence

I. The data

- For the moment, we choose to develop a likelihood in harmonic space based on cross-spectra
- (...but is a real-space approach feasible?)
- Need to choose and validate an estimator (Pseudo-CL, QML?)
- Masks have to be built and validated
- Tomography? i.e. should we divide the data into redshift bins?

II. The model

parameters

likelihood

prior

$$\mathcal{P}(\theta|\mathbf{d}) = \frac{\mathcal{L}(\mathbf{d}|\theta)\mathcal{P}(\theta)}{\int \mathcal{L}(\mathbf{d}|\theta)\mathcal{P}(\theta)d\theta}$$

data

posterior

evidence

II. The model

The cross spectrum between the ISW induced CMB T fluctuations and a biased tracer of the mass distributions (e.g. a galaxy catalog) can be written in terms of an integral over the matter power spectrum (Ho et al. 2008):

$$C_\ell^{cT} = \frac{3\Omega_m \mathrm{H}_0^2}{c^3 \left(l + \frac{1}{2}\right)^2} \int dz\, b_c(z) \frac{dN}{dz} H(z) D(z) \frac{d}{dz} \left(\frac{D(z)}{a(z)} \right) \times P\left(k = \frac{l + \frac{1}{2}}{\chi(z)} \right).$$

(the Limber approximation has been used)
(theoretical assumption also enter the calculation e.g. massless neutrinos)

II. The model

- The matter power spectrum P(k) is a standard output of Boltzmann codes e.g. CAMB, CLASS…
- Should we compute the TG Cls inside the Boltzmann code…
- …. or get the P(k) from the Boltzmann code and compute the TG Cls in the likelihood code itself?
- The Limber approximation greatly speeds up the calculation but…
- …is it good enough for our purposes? In which ell range? Can we afford to do the exact integral?

III. The likelihood function

parameters

likelihood

prior

$$\mathcal{P}(\theta|\mathbf{d}) = \frac{\mathcal{L}(\mathbf{d}|\theta)\mathcal{P}(\theta)}{\int \mathcal{L}(\mathbf{d}|\theta)\mathcal{P}(\theta)d\theta}$$

data

posterior

evidence

III. The likelihood function

- We should choose a suitable approximation for the likelihood function
- Gaussian approximation can be a starting point…
- …but we might need to settle for a better approximation (e.g. Hamimeche-Lewis?)
- “Exact” likelihood for validation and forecasts?
- Whatever the choice, the likelihood approximation has to be thoroughly tested

IV. The Monte Carlo

parameters

likelihood

prior

$$\mathcal{P}(\theta|\mathbf{d}) = \frac{\mathcal{L}(\mathbf{d}|\theta)\mathcal{P}(\theta)}{\int \mathcal{L}(\mathbf{d}|\theta)\mathcal{P}(\theta)d\theta}$$

data

posterior

evidence

IV. The Monte Carlo

- Choice of the MC engine….
  - CosmoMC(+CAMB) has been the *de facto* standard until now (see the Planck analysis)
  - MontePython(+CLASS) is also widely used

- Beyond the choice of the MC engine:
  - Choice of nuisance parameters
  - Find sensible priors for the nuisance parameters

Status of activities (I)

- The meetings of the joint Bo-Fe-Pd group have been dedicated to the evaluation of available tools and to define an initial strategy – as well as to fix some implementation choices:
  - Harmonic-space likelihood
  - Boltzmann code+MC engine: CLASS + MontePython
  - TG Xspectra to be calculated in the likelihood code

Status of activities (II)

- Plan: start writing a simple likelihood CMBXC module for MontePython:
  - Gaussian likelihood
  - Limber approximation
  - No tomography

- and validate it on Gaussian signal + noise simulations
- … then start complicating things…..
- Currently studying how to go beyond the Limber approx
- Timescale for first version of the code: 6 months