Draft version July 31, 2018
Typeset using LaTeX **twocolumn** style in AASTeX61

COVARIANCE TESTING

DESC Members

ABSTRACT

There are a number of codes that compute covariance matrices analytically; the plan is to use these to build TJPCov. In this project, we start along the path of comparing these different codes, building up a suite of tools that can be used to compare covariance matrices. We expect these tools to be useful not only for converging on a single accurate code for computing covariance matrices but also more generally for understanding which parts of the covariance matrix carry the most information (and therefore need the most attention to get right) and which are not relevant (so for example matrices that are not positive definite may still be usable if the negative eigenmodes are not relevant).

1. INTRODUCTION

2. METHODS

There are several ways to tests covariance matrices. We will illustrate each of these on cosmic shear statistics $\xi_{\pm}(\theta)$, focusing for the most part of the Year 1 results of the Dark Energy Survey Abbott et al. (2017). One of the codes will be `Cosmolike` Krause & Eifler (2017); another will be the one used to analyze the KiDS-450 survey Khlinger et al. (2017).

2.1. *One-to-one Comparison*

This is a simple matter of comparing elements of a covariance matrix, usually starting with diagonal elements. Figure 1 shows an example.

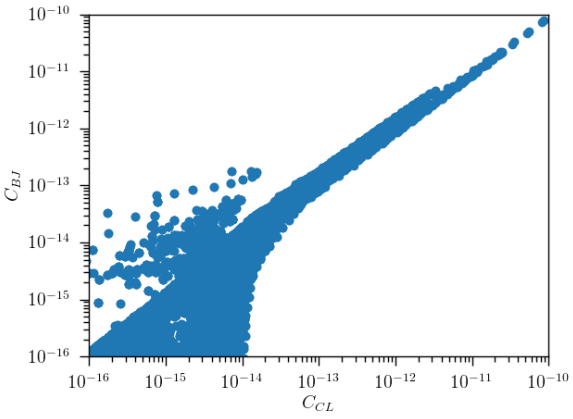

Figure 1. A simple scatter plot of elements of covariance matrices produced by two separate halo model codes.

2.2. *Eigenvalues and eigenvectors*

This is slightly more sophisticated: diagonalize the covariance matrix and examine the eigenvalues and also the associated eigenvectors. Figure 2 shows an example of the eigenvalues from two different covariance codes.

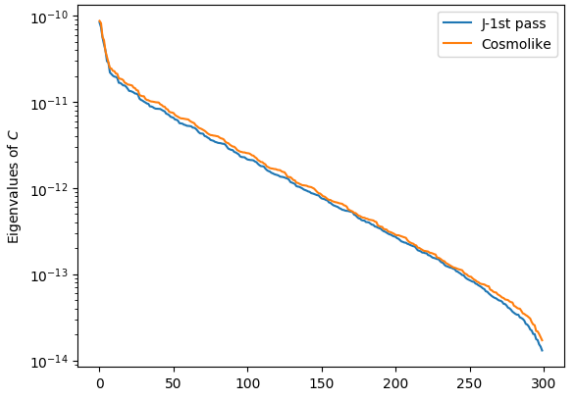

Figure 2. A simple scatter plot of elements of covariance matrices produced by two separate halo model codes.

Figure 3 shows an example of one of the eigenvectors, the one associated with the smallest eigenvalue. This low-eigenvalue mode picks up the differences between the correlation function at different angular scales (each vertical line delineates between two-point functions of shears in different tomographic bin pairs).

2.3. *Parameter Estimation*

Ultimately, what matters is how well the likelihood does at extracting parameter constraints. Since most analyses assume a Gaussian likelihood, this boils down to how well the contours in parameter space agree when computing the χ^2 using two different covariance matrices.

2.4. *Shrinkage*

There have been several methods proposed in the literature to compress the data vectors, extracting as much

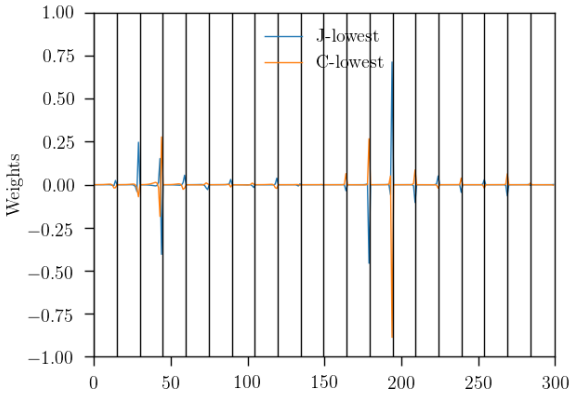

Figure 3. A simple scatter plot of elements of covariance matrices produced by two separate halo model codes.

information as possible. Here we consider two: first compression at the map level Alonso (2018), where linear combination of the tomographic maps are used. If there are 4 tomographic bins, an uncompressed analysis would require ten separate 2-point functions (or 20 for cosmic shear), whereas a compression scheme leads to just a few uncorrelated maps. If there were 3 such maps, then only three 2-point functions would need to be used for the likelihood analysis.

We characterize a given element in the data vector by its angular and tomographic indices: $a_i = a_{lm,\alpha}$ where l, m denote the indices corresponding to given spherical harmonics and α is a given tomographic bin, or equivalently $a_i = a_\alpha(\vec{\theta})$ in real space. The compression occurs in terms of tomographic bins, so that

$$b_\mu(\vec{\theta}) = \sum_\alpha F_{\mu\alpha} a_\alpha(\theta) \tag{1}$$

where the F's are chosen[1] so that the correlation functions of the b's are diagonal in tomographic space:

$$\begin{aligned} w_{\mu\nu}(\theta) &= \langle b_\mu(\vec{\theta}_1) b_\nu(\vec{\theta}_2)\rangle|_{|\vec{\theta}_1 - \vec{\theta}_2|\in\theta} \\ &= \delta_{\mu,\nu} w_\mu(\theta). \end{aligned} \tag{2}$$

Usually with N_t tomographic bins, one must consider $N_t(N_t+1)/2$ correlation functions, but – due to the transformation that diagonalizes the elements – there are only N_t correlation functions to consider (for a given θ). Even better, these can be ordered by the information they contain, so fewer than $N_t$can be used. In the example used in Alonso (2018), 16 tomographic bins were assumed, so that the standard treatment would require 136 correlation functions, but only 3 were needed in order to extract accurate constraints.

[1] Note that this definition of F differs from that in Alonso (2018) in that it includes the inverse of the noise matrix.

This method suggests a way of extracting the most important pieces of the full covariance matrix C. We simply compute the covariance matrix of the w_μ in terms of C and keep only the most important terms. That is,

$$\begin{aligned} C^b_{\mu\nu} &\equiv \langle (w_\mu - \bar{w}_\mu)(w_\nu - \bar{w}_\nu)\rangle \\ &= \sum_{\alpha\alpha'\beta\beta'} F_{\mu\alpha} F_{\mu\alpha'} F_{\nu\beta} F_{\nu\beta'} C_{\alpha\alpha'\beta\beta'}. \end{aligned} \tag{3}$$

Here the angular indices have been suppressed (there are two of them, one for each w_μ), but the full tomographic complexity has been retained. The covariance matrix on the right includes a total of $N_t^2 \times N_t^2$ terms (some of which are equal because of symmetry) corresponding to all possible pairs of two-point functions. But C^b on the right contains only N_t^2 elements, and again these are ordered, so we can use only a subset of them. In the simplest case, where only one linear combination is needed so $\mu = \nu = 1$, a single number (for each pair of angular bins) captures all the relevant information from the full covariance matrix.

The second compression takes place at the 2-point level Zablocki & Dodelson (2016), with the compressed data vector containing linear combinations of the many 2-point functions. In principle, this might work with only N_p 2-point functions where N_p is the number of parameters varied, and each mode, or linear combination, contains all the information necessary about the parameter of interest.

For each parameter p_α that is varied, one captures a single linear mode

$$y_\alpha = U_{\alpha,i} D_i \tag{4}$$

where D_i are the data points and the coefficients are defined as

$$U_{\alpha,i} \equiv \frac{\partial T_j}{\partial p_\alpha} C^{-1}{}_{ji} \tag{5}$$

where T_j is the theoretical prediction for the data point D_j. The now much smaller data set $\{y_\alpha\}$, which contains as few as N_p data points carries with it its own covariance matrix, with which the χ^2 can be computed for each point in parameter space. Propagating through shows that this covariance matrix is related to the original C_{ij} via

$$C_{\alpha\beta} = U_{\alpha i} C_{ij} U^t_{j\beta}. \tag{6}$$

To take an example from DES-Y1, C_{ij} is a 400×400 matrix, while the number of parameters needed to specify the model is only 16, so $C_{\alpha\beta}$ is a 16×16 matrix. We have apparently captured from the initial set of $400 \times 401/2 = 80,200$ independent elements of the covariance matrix a small subset (only 136 in this case) of linear combinations of these 80k elements that really

matter. If two covariance matrices give the same set of $C_{\alpha\beta}$, we do not care whether any of the other eighty thousand elements differ from one another.

3. RESULTS

4. DISCUSSION

5. CONCLUSION

ACKNOWLEDGMENTS

The DESC acknowledges ongoing support from the Institut National de Physique Nucléaire et de Physique des Particules in France; the Science & Technology Facilities Council in the United Kingdom; and the Department of Energy, the National Science Foundation, and the LSST Corporation in the United States. DESC uses resources of the IN2P3 Computing Center (CC-IN2P3–Lyon/Villeurbanne - France) funded by the Centre National de la Recherche Scientifique; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231; STFC DiRAC HPC Facilities, funded by UK BIS National E-infrastructure capital grants; and the UK particle physics grid, supported by the GridPP Collaboration. This work was performed in part under DOE Contract DE-AC02-76SF00515.

REFERENCES

Abbott, T. M. C., et al. 2017, arXiv:1708.01530

Alonso, D. 2018, Mon. Not. Roy. Astron. Soc., 473, 4306

Krause, E., & Eifler, T. 2017, Mon. Not. Roy. Astron. Soc., 470, 2100

Khlinger, F., et al. 2017, Mon. Not. Roy. Astron. Soc., 471, 4412

Zablocki, A., & Dodelson, S. 2016, Phys. Rev., D93, 083525