

# Error Computation for Emulation

April 7, 2020

## 1 Introduction and Notation

This document outlines the error measurements used in our emulation procedure.

The test set contains mock catalogs with  $C = 7$  different cosmologies  $c$ . There are  $B = 5$  boxes (or realizations)  $b$  for each cosmology. These are each populated with  $H = 100$  different halo occupation distribution (HOD) models  $h$ , the same 100 for each cosmology. (There are also 10 random seeds used per HOD, but for now we only use one. We'll worry about this later.)

We are concerned with a clustering statistic  $y$ . We use  $y_{obs}$  to mean the statistic computed on a mock or data. We use  $y_{pred}$  to mean the emulator prediction for that statistic. We take  $y$  to have  $P$  dimensions (bins in our case), indexed by  $p$  (or  $p'$ ).

All of the errors in this document, denoted by  $\sigma$ , are fractional errors. To obtain the physical error, they must be multiplied by the observation.

The statistic is computed on all  $7 \times 5 \times 100 = 3500$  mocks. The statistic is averaged over the 5 boxes, so our final test set contains  $N = 700$  statistics. We will denote these with index  $n$ .

## 2 Test set error

The error on the test set is computed as follows. For each cosmology, compute the mean statistic over the boxes. Compute the deviation from this mean for each of the boxes, defined as the fractional error between the box statistic and the mean statistic for that cosmology. Do this for every HOD. The test set error is the standard deviation of all of these deviations from the mean.

Breaking this down: For each cosmology and HOD, compute the mean statistic  $\bar{y}_{c,h}$  of the  $B$  boxes.

$$\bar{y}_{c,h} = \frac{1}{B} \sum_b^B y_{b,c,h} \quad (1)$$

For each cosmology and HOD, compute the deviations from the mean for each box  $b$ .

$$d_{b,c,h} = \frac{y_{b,c,h} - \bar{y}_{c,h}}{\bar{y}_{c,h}} \quad (2)$$

Compute the standard deviation of all of the deviations from the mean.

$$\bar{d} = \frac{1}{BCH} \sum_b^B \sum_c^C \sum_h^H d_{b,c,h} \quad (3)$$

$$\sigma_{test} = \sqrt{\frac{1}{BCH} \sum_b^B \sum_c^C \sum_h^H (d_{b,c,h} - \bar{d})^2} \quad (4)$$

Note that  $\sigma_{test}$  is a vector of length  $P$ , with a scalar error for each bin. This test set error  $\sigma_{test}$  is used as an input to the Gaussian Process to represent the error on the *training set*. We do this because we have a better handle on this error due to the multiple realizations of the test set. We note that the procedure outlined here estimates the sample variance contribution to the error. The random seeds we have not yet included estimate the shot noise contribution; we will update with this later.

### 3 Emulator performance and error

The overall emulator performance  $\sigma_{perf}$  is defined as

$$\sigma_{perf}^2 = \sigma_{emu}^2 + \sigma_{test}^2 \quad (5)$$

We obtain this with the standard deviation of the fractional error of the test set emulation:

$$f_n = \frac{y_{n,pred} - y_{n,obs}}{y_{n,obs}} \quad (6)$$

$$\sigma_{perf} = \sqrt{\frac{1}{N} \sum_n^N (f_n - \bar{f})^2} \quad (7)$$

We care about the emulator error  $\sigma_{emu}$  separately, so we use

$$\sigma_{emu} = \sqrt{\sigma_{perf}^2 - \sigma_{test}^2} \quad (8)$$

Both  $\sigma_{emu}$  and  $\sigma_{perf}$  are vectors of length  $P$ , with a scalar error for each  $p$ -bin.

## 4 Covariance and likelihood

The covariance between the bins of the statistic is obtained with the Minerva mock catalogs. We compute the statistic on  $M = 100$  Minerva mocks, which we index by  $m$ . The covariance matrix  $C$  is a  $P \times P$  matrix, given by (in vector notation)

$$C = \frac{1}{N-1} \sum_m (y_m - \bar{y})(y_m - \bar{y})^T \quad (9)$$

In element notation, this is

$$C_{pp'} = \frac{1}{N-1} \sum_m (y_{m,p} - \bar{y}_p)(y_{m,p'} - \bar{y}_{p'}) \quad (10)$$

The correlation matrix, or reduced covariance matrix,  $R$  is the normalized covariance matrix, with elements

$$R_{pp'} = \frac{C_{pp'}}{\sqrt{C_{pp}C_{p'p'}}} \quad (11)$$

We combine all of these errors to obtain a final covariance matrix  $\mathcal{C}$  for our likelihood. We denote the measurement covariance matrix  $\mathcal{C}^{meas}$ ; this includes the covariance between bins in the Minerva mocks and in the test set. We need to multiply our test set fractional error  $\sigma_{test}$  by the measured statistic on our data  $y_{obs}$  to obtain the test set physical error before combining it with the correlation matrix.

$$\mathcal{C}_{pp'}^{meas} = R_{pp'}(\sigma_{test,p}y_{obs,p})(\sigma_{test,p'}y_{obs,p'}) \quad (12)$$

We assume the emulator error is independent of the measurement error, and there is no covariance of the emulator error between bins. Then the

emulator covariance matrix  $\mathcal{C}^{emu}$  is, once again multiplying our fractional error by the observed statistic,

$$\mathcal{C}^{emu} = diag(\sigma_{emu}^2 y_{obs,p}^2) \quad (13)$$

The final covariance matrix is

$$\mathcal{C} = \mathcal{C}^{meas} + \mathcal{C}^{emu} \quad (14)$$

The log-likelihood  $\mathcal{L}$  is then defined as

$$\mathcal{L} = -\frac{1}{2}(y_{pred} - y_{obs})\mathcal{C}^{-1}(y_{pred} - y_{obs})^T \quad (15)$$