Testing the accuracy of likelihoods for cluster abundance cosmology

C. Payerne,1* C. Murray,1 C. Combet,1 C. Doux,1 A. Fumagalli,2,3,4,5 M. Penna-Lima6

1 Université Grenoble Alpes, CNRS, LPSC-IN2P3, 38000 Grenoble, France
2 Dipartimento di Fisica - Sezione di Astronomia, Università di Trieste, Via Tiepolo 11, I-34131 Trieste, Italy
3 INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, I-34131 Trieste, Italy
4 IFPU, Institute for Fundamental Physics of the Universe, via Beirut 2, 34151 Trieste, Italy
5 INFN, Sezione di Trieste, Via Valerio 2, I-34127 Trieste TS, Italy
6 Instituto de Física, Universidade de Brasília, 70910-900, Brasília, DF, Brazil

ABSTRACT

The abundance of galaxy clusters is a sensitive probe to the amplitude of matter density fluctuations, the total amount of matter in the Universe as well as its expansion history. Inferring correct values and accurate uncertainties of cosmological parameters requires accurate knowledge of cluster abundance statistics, encoded in the likelihood function. In this paper, we test the accuracy of cluster abundance likelihoods used in the literature, namely the Poisson and Gaussian likelihoods as well as the more complete description of the Gauss-Poisson Compound likelihood. This is repeated for a variety of binning choices and analysis setups. In order to evaluate the accuracy of a given likelihood, this work compares individual posterior covariances to the covariance of estimators over the 1000 simulated dark matter halo catalogs obtained from PINOCCHIO algorithm. We find that for Rubin/LSST and Euclid-like surveys the Gaussian likelihood gives robust constraints over a large range of binning choices. The Poisson likelihood, that does not account for sample covariance, always underestimates the errors on the parameters, even when the sample volume is reduced or only high-mass clusters are considered. We find no benefit in using the more complex Gauss-Poisson Compound likelihood as it gives essentially the same results as the Gaussian likelihood, but at a greater computational cost. Finally, in this ideal setup, we note only a small gain on the parameter error bars when using a large number of bins in the mass-redshift plane.

Key words: galaxies: clusters: general – cosmology: cosmological parameters – methods: statistical

1 INTRODUCTION

Galaxy clusters have been essential in the construction of the standard model of cosmology, providing some of the first evidence of dark matter (Zwicky 1933) through the motions of galaxies within galaxy clusters and their spatial distribution provided evidence for the primordial origin of density fluctuations (Kaiser 1984). Since their formation is sensitively dependent on the properties of both dark energy and dark matter, future measurements of the masses of galaxy clusters and their abundance will hopefully deliver deep insights on the nature of these elusive components of the Universe.

The primary cosmological probe for galaxy clusters is the counting of their abundance as a function of mass and redshift. The Rubin Observatory Legacy Survey of Space and Time (LSST Science Collaboration et al. 2009) and the Euclid mission survey (Laureijs et al. 2011) will mark a new era in the study of galaxy clusters for cosmology due to an order of magnitude increase of the number of galaxy clusters with precision weak gravitational lensing information. In order to fully take advantage of these datasets it is important to have an accurate statistical description of the abundance of clusters in the Universe.

Galaxy clusters form in the densest regions of the cosmic web. Their formation is a complex, non-linear process, dominated by gravitational forces although it is important also to consider baryonic effects (Henson et al. 2017). They form from the collapse of overdense regions of the initially Gaussian matter density field of the early Universe. Early attempts at describing their statistics were with the excursion-set model (Press & Schechter 1974; Bond et al. 1991; Sheth & Tormen 1999). Such models must simplify greatly the complexity of cluster formation; assuming spherical or ellipsoidal collapse of over-dense regions. Therefore to understand in more detail the distribution of clusters in mass and redshift cosmological simulations are essential (Jenkins et al. 2001; Evrard et al. 2002; Tinker et al. 2008).

Much of this existing work however has focused on the prediction of the mean of the cluster abundance for a given mass and redshift, rather than the distribution of the number of clusters at a given mass and redshift. Essentially this is the challenge of this study.

In practice three cluster likelihoods have been used to perform cluster count cosmology: a Poissonian likelihood (e.g., Mantz et al. 2015; Bocquet et al. 2019; Planck Collaboration et al. 2016), a Gaussian likelihood (e.g., Abbott et al. 2020; Lesci et al. 2022a) and, more recently, a combination of the two (Lesci et al. 2022b). Each of these likelihoods are described in detail later in the article.

The latent likelihood of the distribution of cluster abundances in the Universe is not Poissonian, Gaussian or a combination of the...
two. By the latent likelihood here we refer to the unknown underlying likelihood which would perfectly describe the distribution of galaxy clusters in the Universe given a cosmological model. The halo model of structures in the Universe (in which each of these likelihoods is to some extent built upon) is a model, which at some level will break down. The distribution of galaxy clusters in a universe predicted by a cosmological model will be more complicated than any of these likelihoods. Therefore the assumption is that the cluster abundance distribution is well described by these models. If not this will impact the constraint of cosmological parameters, possibly leading to mis-estimation of parameters errors and even bias the results.

This work aims to answer a very general question: is our statistical description of the abundance of clusters sufficiently accurate for future surveys? Within the cosmology community at large this has been assessed in the context of mis-estimation of the covariance matrices for cosmological analyses (Dodelson & Schneider 2013; Taylor & Joachimi 2014; Sellentin & Heavens 2017). Here we move beyond the Gaussian approximation in the specific case of cluster abundance cosmology, to answer to what extent are the Poissonian and Gaussian models insufficient. The insufficiency of Gaussian likelihoods in the case of the CFHTLenS cosmic shear survey was considered by Sellentin & Heavens (2018), where they note since that one-point statistic of the correlations are skewed, the Gaussian approximation is not necessarily accurate.

In addition, we seek to answer multiple related questions within this work. They are: (i) which likelihoods can sufficiently describe the abundance of galaxy clusters as a function of mass and redshift and provide robust constraints, (ii) among those, does a given likelihood provides stronger cosmological constraints than the others?, and (iii) is there an optimal binning regime that allows us to extract the most cosmological information?

To answer these questions we present a comparison of the cluster distributions using a set of cosmological simulations and compare each theoretical model to that measured in the distribution. Subsequently, we show the effects on the obtained cosmological parameters and the estimated posterior of the cosmological parameters.

In Section 2, we give an overview of the three different likelihoods for cluster abundance cosmology used in this work. In Section 3, we present a framework for testing the accuracy of each of these likelihoods. The 1000 PINOCCHIO simulations that we use in this paper are presented in Section 4. Section 5 describes the covariance matrices used in the Gaussian and Gauss-Poisson Compound likelihoods and Section 6 presents the methodology of the work and the technical details of the posterior sampling. The main results of this articles are given in Section 7, in which using 1000 PINOCCHIO simulations we obtain the posteriors with each likelihood. By comparing the individual posteriors to the distribution of parameters estimated from the 1000 simulations we can test the accuracy and precision of each of our likelihoods. Finally, we discuss the conclusions of our work in Section 8.

## 2 LIKELIHOODS FOR CLUSTER ABUNDANCE COSMOLOGY

In the following Section we describe the three different likelihoods used in this work; the Gauss-Poisson Compound likelihood, the Poissonian likelihood and the Gaussian likelihood.

### 2.1 Gauss-Poisson Compound likelihood (GPC) likelihood

Galaxy clusters trace the matter density field in the universe. The abundance, or number density, of galaxy clusters follows spatial fluctuations of the underlying matter density field. This relation can be parameterised by taking clusters to be biased tracers of the underlying density field (see Cooray & Sheth 2002, for a review of the halo model). The overdensity of galaxy clusters, \( \delta_h(x) \), in a region centered at \( x \).

\[
\delta_h(x) = h(m, z) \delta(x),
\]

where \( h(m, z) \) is the linear bias parameter which is a function of both cluster mass and redshift and \( \delta(x, z) \) is the underlying matter density field. Using Eq. (1) the local halo number density, \( n \), can be expressed as,

\[
n(m, z, x) = \bar{n}(m, z)(1 + b(m, z) \delta(x, z)),
\]

where \( \bar{n} \) is the predicted mean number density of clusters at a given mass and redshift for a given cosmological model. As we are concerned with fluctuations on the size of the survey, the matter density field \( \delta(x, z) \) will be well described by a Gaussian random field; the same would not be true on smaller scales.

The abundance of galaxy clusters is the count of clusters in a specific spatial region, such as the observed number of galaxy clusters will be a Poisson realisation of the local halo number density, \( n \), which in turn is a Gaussian realisation of \( \bar{n} \) depending on the underlying density field of the survey region.

Hu & Kravtsov (2003) proposed a Gauss-Poisson Compound likelihood for cluster abundance cosmology that combines these two statistical effects; the average number of clusters \( \lambda_{ij} \) within a given \( i \)-th mass and \( j \)-th redshift is calculated from the halo mass function \( \frac{dn(m, z)}{dm} \) predicting the average number density of halos per mass range and the comoving partial volume \( \frac{d^2V(z)}{d\delta d\Omega} \) as follows,

\[
\lambda_{ij} = \Omega_S \int_{z_j}^{z_{j+1}} dz \int_{m_{ij}}^{m_{i+1}} dm \frac{dn(m, z)}{dm} \frac{d^2V(z)}{d\delta d\Omega},
\]

where \( \Omega_S \) is the survey solid angle. In this work, we use letters of the \textit{Latin} alphabet \((i, j, k, \ldots)\) to refer data vector indices in the data space. Given that the statistics of the underlying matter density field \( \delta(x, z) \) are Gaussian, \( n(m, z, x) \) is a random Gaussian realisation of \( \bar{n}(m, z) \).

The observed abundance of clusters \( N_{ij} \) in the \( i \)-th mass and \( j \)-th redshift bins would be a Poisson realisation of the Gaussian random variable

\[
\bar{\lambda}_{ij} = \lambda_{ij} + \delta \lambda_{ij},
\]

where the abundance perturbation \( \delta \lambda_{ij} \) writes in terms of the halo bias \( b \) and the matter overdensity field \( \delta \) such as

\[
\delta \lambda_{ij} = \int_{\Omega_S} d^2\theta \int_{z_j}^{z_{j+1}} dz \int_{m_{ij}}^{m_{i+1}} dm \frac{dn(m, z)}{dm} \frac{d^2V(z)}{d\delta d\Omega} \times b(m, z) \delta(f_k(w) \theta, w),
\]

where \( \delta(f_k(w) \theta, w) \) is the three-dimensional matter overdensity field at the comoving position \((x, y, z) = (f_k(w) \theta, w)\), where \( \theta \) denotes the two-dimensional position on the sky, and \( f_k(w) \) and \( w \) are respectively the transverse and line-of-sight comoving distances. The line-of-sight comoving distance depends on the input redshift \( w \equiv w(z) \).
For a set of observed cluster abundance \( \mathcal{N} = \{N_{ij}\} \) we have the Gauss-Poisson Compound cluster likelihood likelihood,

\[
\mathcal{L}(\mathcal{N}) = \int d\lambda \mathcal{N}(\lambda, \Sigma_{SV}) \times \prod_{i,j=1}^{c} \mathcal{P}(N_{ij}|\lambda_{ij}),
\]

where \( c = N_z \times N_m \) is the total number of mass-redshift bins. In the above equation, \( \mathcal{N}(\lambda, \Sigma_{SV}) \) is the multivariate Gaussian distribution with mean \( \lambda = \{\lambda_{ij}\} \) and covariance \( \Sigma_{SV} \). The distribution \( \mathcal{P} \) is the single variate Poisson distribution given by

\[
\mathcal{P}(N_{ij}|\lambda_{ij}) = \frac{\lambda_{ij}^{N_{ij}} e^{-\lambda_{ij}}}{N_{ij}!}.
\]

The quantity \( \Sigma_{SV} \) is the sample covariance, i.e. the covariance of the \( \delta \lambda_{ij} \) such as \( \Sigma_{SV}[ij][kl] = \langle \delta \lambda_{ij} \delta \lambda_{kl} \rangle \).

For a given binning scheme, if we neglect correlations between two different bins of redshift (which in general are small, see Fig. A1), the GPC likelihood in Eq. (6) can be written as (Takada & Spergel 2014)

\[
\mathcal{L}(\mathcal{N}) = \prod_{j=1}^{N_z} \mathcal{L}_j([N_{ij}]_{1 \leq i \leq N_m}),
\]

where each term \( \mathcal{L}_j \) in Eq. (8) are given by the integral form

\[
\mathcal{L}_j(\mathcal{N}) = \int d\delta \mathcal{N}(\delta[0,S_j]) \prod_{i=1}^{N_m} \mathcal{P}(N_{ij}|\lambda_{ij}(1+b_{ij}\delta)).
\]

In Eq. (9), \( \lambda_{ij} \) is the cluster abundance prediction and \( b_{ij} \) is the average halo bias in the \( i \)-th mass and \( j \)-th redshift bin given by \( \langle \lambda b \rangle_{ij}/\lambda_{ij} \) in Eq. (25). \( S_j = \langle \delta \rangle \) is fluctuation amplitude of the matter density perturbation \( \delta \) in the \( j \)-th redshift bin, as given in Eq. (A6). Here, the GPC likelihood is approximated as the product of \( N_z \) single variate integrals (instead of computing a \( N_z \times N_m \) multivariate integrals).

The GPC satisfies the two limit cases depending on the relative contribution between average cluster abundance shot noise (see Section 2.2) and sample variance \( \Sigma_{SV} \), where the GPC likelihood takes the form of a Poissonian distribution or of a Gaussian distribution. The two next sections briefly presents these two limiting cases, which are the most widely used in the literature.

### 2.2 Poisson likelihood

When the cluster abundance is dominant relative to the sample variance, i.e. \( \lambda_{ij} \gg \Sigma_{SV}[ij][i=j] \) the GPC likelihood in Eq. (6) simplifies to the form (Lima & Hu 2004).

\[
\mathcal{L}(\mathcal{N}) \approx \prod_{i,j=1}^{c} \mathcal{P}(N_{ij}|\lambda_{ij}),
\]

where \( \mathcal{P} \) is the Poisson distribution in Eq. (7), and \( \lambda_{ij} \) is the average number of cluster in Eq. (3). In this approximation, cluster counts in different mass-redshift bins are independent sampling of the Poisson mean \( \langle N_{ij} \rangle = \lambda_{ij} \) and each have variance \( \sigma_{ij}^2 = \lambda_{ij} \). This variance \( \lambda_{ij} \) is referred to as Poisson shot noise. It denotes the intrinsic scatter of an uncorrelated count experiment.

### 2.3 Gaussian likelihood

When the shot noise \( \lambda_{ij} \) is no more dominant relative to the sample variance, and \( \lambda_{ij} \gg 1 \), the GPC likelihood takes the form of a multivariate Gaussian likelihood (Lima & Hu 2004)

\[
\mathcal{L}(\mathcal{N}) \approx \frac{1}{\sqrt{(2\pi)^N |\Sigma|}} \exp \left[ -\frac{1}{2} \left( \mathcal{N} - \lambda \right)^T \Sigma^{-1} \left( \mathcal{N} - \lambda \right) \right].
\]

In the above, the full covariance matrix \( \Sigma \) for cluster abundance is given by

\[
\Sigma = \Sigma_{SN} + \Sigma_{SV},
\]

where \( \Sigma_{SN} = \text{diag}(\lambda) \) is the Poisson shot noise covariance and is completely diagonal as there is no correlation between different mass-redshift bins. \( \Sigma_{SV} \) is the sample covariance. In this approximation, the cluster counts are correlated Gaussian variables with extra-variance given by the sample variance.

We used the Core Cosmology Library (Chisari et al. 2019) to predict the halo mass function from Despali et al. (2016) and the halo bias from Tinker et al. (2010).

### 3 TESTING likelihood ACCURACY

To take an explicit example, we can consider multiple realisations of our Universe which have at their origin the exact same cosmological model. For each realisation, we use a cluster abundance likelihood to estimate a set of cosmological parameters. In the case that we use the true model describing the statistics of cluster abundance in the Universe, the posteriors we estimate in each Universe will be consistent, after accounting for the expected statistical fluctuations. If our cluster abundance likelihood is lacking, for example, the Poisson likelihood missing the sample variance contribution, our estimated parameters may be biased and/or misestimate the parameter errors. In this case, the analysis of two realisations of the same cosmological model may lead to the estimation of inconsistent cosmological parameters.

Therefore we can develop a test for the likelihood model by comparing the dispersion of the posterior means over many realisations of the Universe against the predicted dispersion from an individual realisation. And more precisely, we can compare the average posterior covariance from an ensemble of the simulations to the covariance of an ensemble of posterior means.

The posterior distribution for a set of parameters \( \theta = (\theta_{\alpha}) \) is given in terms of the likelihood \( \mathcal{L}_Y \) and the prior \( \pi(\theta) \) by the Bayes theorem, namely

\[
\mathcal{P}_Y(\theta|\mathcal{N}) \propto \mathcal{L}_Y(\mathcal{N}|\theta) \pi(\theta).
\]

where \( \mathcal{L}_Y \) is the likelihood used for cosmological inference, i.e. the one chosen for the analysis. For an individual dataset \( N \), the posterior covariance for a set of parameters \( \theta \) estimated from \( \mathcal{P}_Y \) is defined as

\[
C_{\theta\theta}^{\text{ind}} = \mathbb{E}_{\mathcal{P}_Y}[(\theta_{\alpha} - \bar{\theta}_{\alpha})(\theta_{\beta} - \bar{\theta}_{\beta})],
\]

where \( \bar{\theta}_{\alpha} \) is the average parameter given the posterior \( \mathcal{P}_Y \) namely

\[
\bar{\theta}_{\alpha} = \mathbb{E}_{\mathcal{P}_Y}[\theta_{\alpha}].
\]

In this work, we will consider greek symbols (\( \alpha, \beta, \gamma, \ldots \)) to refer to parameter indices in the cosmological parameter space. A lower bound for the posterior covariance, \( C_{\theta\theta}^{\text{ind}} \), can be set using the Fisher matrix formalism Fisher (1935) (see also Tegmark et al. 1997 and Heavens et al. 2014) and the Cramer-Rao inequality (Cramer 1952)

\[
C_{\alpha\beta}^{\text{ind}} \geq (F^{-1})_{\alpha\beta},
\]

where \( F \) is the Fisher matrix given by (Fisher 1935),

\[
F_{\alpha\beta} = \mathbb{E}_{\mathcal{L}_Y} \left[ \left( \frac{\partial}{\partial \theta_{\alpha}} \ln \mathcal{L}_Y \right) \left( \frac{\partial}{\partial \theta_{\beta}} \ln \mathcal{L}_Y \right) \right].
\]
where the expectation runs over the realisation of the data that follows the likelihood distribution. The Fisher information matrix for multivariate Gaussian and Poisson likelihoods (see Appendix B2) is,

\[ F_{\alpha\beta} = \lambda_{\alpha}^{-1} \Sigma_{\gamma}^{-1} \lambda_{\beta}, \tag{18} \]

with \( \Sigma_{\gamma} \) the data covariance matrix assumed to be independent of the cosmological parameters, and where \( \lambda_{\alpha} \) is the derivative of the model that predict the signal with respect to the cosmological parameter \( \alpha \). The Fisher forecast estimates the first order of the posterior covariance. When the posterior is not Gaussian in the parameter space, the first order Fisher forecast is not sufficient and higher order forecasts needs to be considered (Wolz et al. 2012; Sellentin et al. 2014). In our case, we will see that posteriors are almost Gaussian in the parameter space, then \( F^{-1}_{\alpha\beta} \approx C^{ind}_{\alpha\beta} \).

We now define \( C^{ens} \) as the covariance of the ensemble of posterior means obtained for data generated from the latent likelihood \( L_X \).

Therefore,

\[ C^{ens}_{\alpha\beta} = \mathbb{E}_{L_X} \left[ \left( \hat{\theta}_\alpha - \bar{\theta}_\alpha \right) \left( \hat{\theta}_\beta - \bar{\theta}_\beta \right) \right], \tag{19} \]

where \( \bar{\theta}_\alpha \) is the individual posterior mean defined in Eq. (15), derived from the individual posterior \( P_Y \) in Eq. (13). We have introduced the likelihood \( L_X \), which is the true underlying likelihood which describes the data. By definition, \( C^{ens} \) defines the frequentist covariance of a Bayesian estimator, that depends on both analysis and simulation. Therefore, the accuracy of the PINOCCHIO algorithm is within \( \sim 5-10\% \) for reproducing the halo mass function, halo two-point statistics, and the halo bias at all relevant redshifts used in this study (Munari et al. 2017).

Finally, if we substitute Eq. (18) into Eq. (20) and take the two covariance matrices to be equal, \( \Sigma_{\gamma} = \Sigma_{X} \) (see again Appendix B), then the two posteriors are indeed equivalent,

\[ C^{ens} = C^{ind}, \tag{21} \]

We can therefore justify the comparison of these two different posterior covariance matrices as a method to test the accuracy of a given likelihood. Additionally in the case that \( \Sigma_{\gamma} \neq \Sigma_{X} \) we have,

\[ C^{ind} \neq C^{ens}. \tag{22} \]

In the following we assume this result also holds for the Gaussian-Poisson likelihood (although not formally demonstrated). Since we find the latent likelihood to be close to Gaussian this is likely true in our specific case.

Note that this equivalence is discussed more general terms in Efron (2015), and also in Penna-Lima et al. (2014) for an unbinned Poisson likelihood. Similar calculations are made in Percival et al. (2022) in order to obtain a prior which gives equivalent results for Frequentist and Bayesian analyses.

In order to perform such tests we require many simulations of the universe. These simulations are described in the following section.

---

1 This is the case when the latent likelihood and the likelihood used for the posterior estimation parameters are the same.
mass-redshift bins, is computed using (Lacasa et al. 2018)
\[ S^{\text{th}}_{ij} = \Omega \int z_j dz \int m_i dm \frac{dV(z)}{d\Omega} b(m,z)^{ij}, \]
where \( S_{ij} \) is the matter density fluctuation covariance between the two redshift bins \( j \) and \( i \). In practice, we use the PySSC package\(^2\) to compute \( S_{ij} \) (see Appendix A for details). The quantity \( \langle b \lambda \rangle_{ij} \) is the average halo bias in the \( i \)-th mass and \( j \)-th redshift bin such as
\[ \langle b \lambda \rangle_{ij} = \Omega \int z_j dz \int m_i dm \frac{dV(z)}{d\Omega} b(m,z)^{ij}. \]

The cluster count covariance matrix may also be estimated from the data. To validate our use of the PySSC code, Appendix A presents the comparison between the covariance matrix estimated from the 1000 simulations to the theoretical prediction computed with PySSC.

6 METHODOLOGY

The method we use to test each likelihood is the following:

- for each of the 1000 mock catalogs we obtain a posterior (using importance sampling);
- from each of these posteriors we calculate the posterior mean and posterior covariance;
- from the ensemble of the posteriors we calculate the mean of the posterior means and the covariance of the posterior means;
- we then compare the results of the individual posteriors against the ensemble of the posteriors;
- we compute the Fisher forecast from Eq. (18) and ensemble forecasts from Eq. (20) for comparison.

This test is performed for each of the likelihoods for three different binning schemes. The binning schemes in the mass-redshift plane are given in Section 6.1. The method used, importance sampling, is detailed in Section 6.2. In Section 6.3 we show how the ensemble posterior covariance is calculated.

6.1 Mass and redshift ranges, and binning setups

For this analysis, we adopt a binning scheme of the \((\log_{10})\text{mass-redshift}\) plane within the redshift range \([z_{\text{min}}, z_{\text{max}}] = [0.2, 1.2]\) and the mass range \([\log_{10}(m_{\text{min}}/M_\odot), \log_{10}(m_{\text{max}}/M_\odot)] = [14.2, 15.6]\). We have three different binnings: a coarse, medium, and fine binning which are detailed in Table 1. Both redshift and \((\log_{10})\text{mass}\) bins are equally spaced.

Figure 2 shows the predicted abundance for the fiducial cosmology for the three binning schemes. For the coarse case, the number of expected halos per bin never drops below one and the bins are generally well-populated (there is an average of 5000 halos per mass and redshift over all the binning scheme, ranging from 1 to 25 000). For the medium and fine grids a significant portion of the mass-redshift plane are expected to have only a few or no halos. From this we expect the Gaussian approximation to be valid only for the coarse binning scheme, while this may be a less accurate statistical description for the two other cases. The different binning schemes allow us to explore a variety of regimes, from the shot noise-dominated regime to the non-negligible sample variance regime, and to test the performance of the likelihoods described in Section 2 in each regime.

6.2 Posterior estimation using importance sampling

To undertake this analysis, from the cluster abundance data vector \( \vec{N} \), we perform a cosmological analysis for the 1000 simulations, three binning schemes and three likelihoods. Therefore we wish to obtain 9000 posteriors in the parameter space \( \theta = (\Omega_m, c_g) \). This would be computationally expensive for a standard Markov Chain Monte Carlo method. Instead, we use importance sampling (IS)\(^3\) to efficiently sample the joint posterior distribution of the parameters. We briefly recap the basic framework of importance sampling below (see Elvira & Martino 2021 for a recent review).

Importance sampling can be used to estimate the statistical properties of the posterior distribution \( P_Y(\theta|N) \) in Eq. (13) from random

\(^2\) https://github.com/fabienlacasa/PySSC

\(^3\) Although not shown here, we did check on a single simulation that the two approaches yield the same results. Importance sampling is however much faster to run, since count prediction is evaluated once at each location in the parameter space. In particular, it is much faster for GPC where the likelihood function is computationally taxing.

\[ \]
sample of points that follows a proposal distribution \( q \). Considering a random variable \( X \sim p \), the expectation of \( X \) is given by

\[
E_p[X] = \int_{-\infty}^{+\infty} dx p(x) x.
\]  

(26)

The above average can be estimated from sampling another random variable \( Y \) following a proposal distribution \( q \). We call \( \hat{X} \) the unbiased estimator of \( E_p[X] \),

\[
\hat{X} = \frac{1}{N_q} \sum_{k=1}^{N_q} w_k y_k,
\]  

(27)

where \( N_q \) is the number of samples and the weights, \( w_k \) are defined as,

\[
w_k = \frac{p(y_k)}{q(y_k)}.
\]  

(28)

The mean of \( X \) is therefore expressed as a weighted mean over the sample \( Y \). So the estimated individual average parameter \( \hat{\theta}_i^{\text{ind}} \) is given by,

\[
\hat{\theta}_i^{\text{ind}} = \sum_{k=1}^{N_q} w_k [\theta_\alpha]_k,
\]  

(29)

where \([\theta_\alpha]_k\) is the sampled parameter value \( \theta_\alpha \) at position \( k \). From this, one may also estimate the unbiased weighted covariance matrix of \( \theta = (\Omega_m, \sigma_8) \) as

\[
\hat{C}^{\text{ind}}_{\alpha\beta} = \sum_{k=1}^{N_q} \tilde{w}_k ((\theta_\alpha)_k - \hat{\theta}_i^{\text{ind}})((\theta_\beta)_k - \hat{\theta}_i^{\text{ind}}).
\]  

(30)

with

\[
\tilde{w}_k = w_k \frac{\sum_{k=1}^{N_q} w_k}{\left(\sum_{k=1}^{N_q} w_k\right)^2 - \sum_{k=1}^{N_q} w_k^2}.
\]  

(31)

For the proposal \( q \), we use a bi-variate Gaussian distribution centered on the \((\Omega_m, \sigma_8)\) input values of the PINOCCHIO simulations. The parameter covariance matrix of the proposal distribution \( q \) is chosen as the Fisher forecast (given in Eq. (18)) for the coarse binning scheme, multiplied by a factor 36. This ensures that the proposal distribution encompasses a large enough region of the parameter space covered by the ensemble of \( \sigma \) contours of each individual posterior.

To compute the posterior we use the different likelihoods \( L \) presented in Section 2, and we use flat priors for the parameters, namely \( \Omega_m \in [0.293, 0.321] \) and \( \sigma_8 \in [0.820, 0.836] \). These ranges were chosen as 12 times the square root of the Fisher variances on each parameter given in Eq. (18). The proposal has to be wide enough to contain 2\( \sigma \) contours for each posterior.

As an illustration, the light blue points in Fig. 3 show \( N_q = 10^4 \) samples of the proposal distribution \( q \) for the coarse \( 4 \times 4 \) binning. The four coloured contours corresponds to the 1 and 2\( \sigma \) confidence intervals obtained using the binned Poissonian likelihood, i.e. \( L(N|\theta) \) given by Eq. (10), for four of the PINOCCHIO simulations\(^4\).

\(^4\) The codes used to compute the cluster abundance and covari-
These are obtained as the weighted histograms of the $q$-sample. The unsmoothed contours have been obtained directly from the $10^4$ samples.

### 6.3 Estimating the ensemble posterior covariance

After the 1000 posteriors have been estimated using importance sampling for each of the 1000 simulations we can estimate the ensemble posterior covariance. The estimator of the cosmological parameter mean value $\hat{\theta}_\alpha$ in Eq. (15) is noted $[\hat{\theta}_\alpha^{\text{ind}}]_k$ for the $k$-th simulation and is calculated as in Eq. (29). The mean cosmological parameters over the ensemble $\langle \hat{\theta}_\alpha \rangle$ that appear in Eq. (19) are calculated from these mean values,

$$
\hat{\theta}_\alpha^{\text{ens}} = \frac{1}{N_{\text{sim}}} \sum_{k=1}^{N_{\text{sim}}} [\hat{\theta}_\alpha^{\text{ind}}]_k.
$$

The corresponding ensemble covariance matrix is estimated as,

$$
\hat{C}_{\alpha\beta}^{\text{ens}} = \frac{1}{N_{\text{sim}} - 1} \sum_{k=1}^{N_{\text{sim}}} ([\hat{\theta}_\alpha^{\text{ind}}]_k - \hat{\theta}_\alpha^{\text{ens}}) (\hat{\theta}_\beta^{\text{ens}} - \hat{\theta}_\beta^{\text{ens}}).
$$

From this covariance matrix we can extract the ensemble variance and correlation coefficient in order to compare these results against those of the individual posteriors calculated in Eq. (30).

### 7 RESULTS

In this Section we present the results of testing the accuracy of likelihoods using importance sampling and the set of the 1000 PINOCCHIO simulations. In Section 7.1 we compare the individual posterior means to the ensemble posterior mean, in order to check for a bias in the likelihoods. In Section 7.2 we compare the individual posterior variances and individual posterior correlation coefficients against the ensemble posterior variance and ensemble correlation coefficients, in order to show the robustness of the estimated posteriors. In Section 7.3 and Section 7.4 we respectively reduce the sky area and increase the lower mass-cut value of the analysis in order to identify if a regime exists in which the Poisson likelihood or the GPC likelihood obtain tighter, yet still robust, cosmological constraints than the Gaussian likelihood.

#### 7.1 Bias on the posterior mean

Figure 4 shows, for the coarse $4 \times 4$ binning, the histograms of the 1000 means of the posteriors for $\Omega_m$ (top row) and $\sigma_8$ (bottom row), computed with Eq. (29). From left to right, the columns corresponds to the three likelihoods; Poissonian, Gaussian and GPC. The input values of the parameters are given by the vertical solid lines. The distributions of the estimated cosmological parameters are very similar for each likelihood and they are roughly normally distributed about the input cosmology.

In Fig. 5 we summarise the results for the bias on the mean estimated cosmological parameters for each of the three binning and the three likelihoods. The black dots show the mean of the 1000 estimated cosmological parameters. The black solid lines shows the standard deviation of the mean. The red intervals shows the error on the mean of the 1000 estimated cosmological parameters (the standard deviation divided by $\sqrt{1000}$). The black dotted line indicates the input cosmological parameters to the simulations. We see that there is a small consistent bias between the input value of the cosmological parameters and the mean of the estimated parameters (the dotted black line does not fall within the red error bars for any of the binnings or the likelihoods). There are many potential explanations...
for this bias. It suggests that the likelihood used and therefore the underlying halo model is on some level inaccurate. As the bias is consistent between the Poisson likelihood which does not use the halo bias and the other two likelihoods which do this bias is not because of the halo bias model used. The bias is nonetheless small, as it is sub-percent and much smaller than the variance of an individual posterior, which we present in the following subsection.

7.2 Robustness of the posterior covariance

Now we consider the errors and correlations of the parameters obtained from individual simulations by comparing them to the dispersion of the mean values of the posteriors over the 1000 simulations. As discussed in Section 3, this provides a useful test to check the robustness of the error bars.

Figure 6 summarizes the error estimations for the three binning schemes (columns) and the two parameters ($\sigma_8$ and $\Omega_m$ on the top and bottom row respectively). In each panel, the likelihood used for the analysis is given on the x-axis. The blue dots with errorbars correspond to the mean and dispersion of the individual errors that were computed using the importance sampling covariance matrix in Eq. (30). The red crosses are the values of the standard deviation of the ensemble of mean values, calculated from the ensemble covariance matrix presented in Eq. (33). The red crosses are the same errors as the black error bars in Fig. 5.

The errors slightly decrease when using more bins, regardless of the likelihood, with an improvement of 10%, between the coarse (left) and medium (middle) binning scheme. There might therefore be a small benefit to consider a finer binning when dealing with the upcoming survey data of Rubin LSST or Euclid. Pushing things further, there seems however to be no gain (or only a very marginal one) in using the fine binning compared to the medium one.

The Poisson likelihood underestimates the individual errors by up to 30%. This is evident as the blue dots are below the red crosses (the individual errors are smaller than the ensemble errors). This is true even when a significant fraction of the mass-redshift space has low number counts for medium and fine binnings. See for example Fig. 2. The errors are underestimated for the Poisson likelihood as this likelihood does not include the sample variance contribution to the cluster abundance.

We see that the agreement is much better for the Gaussian and GPC descriptions, that account for sample variance. The individual errors are only slightly under- or over-estimated (~ 2 – 5% for the coarse and medium cases) compared to the dispersion of the 1000 means. It seems likely that this slight underestimation arises from our use of only the block-diagonal terms in the Gaussian and GPC covariance matrices (we ignore correlations between redshift bins).

In Fumagalli et al. (2021) it is also evident that the parameter constraints for the average individual posteriors for the Poisson likelihood are tighter than that of the Gaussian likelihood. In this work we were able to show explicitly that this is not because the Poisson uses information beyond the Gaussian likelihood but rather that the Poisson neglects the sample variance contribution.

We can also better understand the results using two different forecasts, motivated by the discussion presented in Section 3. We compare the individual posterior results and the ensemble posterior results against two different forecasts. The first is the often used Fisher forecast (e.g., Sartoris et al. 2016; Artis et al. 2021), introduced here in Eq. (18). This forecast should always agree with the estimated individual errors since the likelihood used in the analysis (which we call the analysis likelihood) is assumed to be equivalent to the latent likelihood driving the data. These forecasts are the black circles in Fig. 6. There is agreement between the blue dots and the black circles therefore these forecasts provide a consistency check on the analysis. We did not compute the Fisher forecast of the GPC individual parameter covariance, since it requires the computation of as many multidimensional integrals as the total number of mass-redshift bins in Eq. (17), so we did not investigate this complex computation in this work. But regarding the difference between Gaussian and GPC individual errors for all binning scheme, we can assume that the GPC Fisher constraints are roughly the same as for the Gaussian case.

Secondly we look at the ensemble forecast introduced in Eq. (20). This gives a prediction of the errors on the cosmological parameters where the analysis likelihood differs from the latent likelihood. As we do not know the latent likelihood we use the Gaussian likelihood as a substitute for the latent likelihood in this forecast, therefore stating explicitly that $C_{\text{ens}} = C_{\text{full}}$ for the Gaussian analysis likelihood. These forecast are shown for the Poisson analysis likelihood with the black boxes in Fig. 6. We found that for the Poisson analysis likelihood these forecasts predict that the ensemble variance is higher than individual errors. Still for the Poisson likelihood, both ensemble and Fisher forecast are respectively larger and smaller than the correct errors we should obtain if the Gaussian likelihood is the latent one (black circle in the Gaussian column). We do not show the ensemble forecast for the Gaussian analysis likelihood, since it overlaps with the forecasted individual error (black circles). These forecasts give good agreement with the calculated ensemble errors, therefore we can show that the underestimate of the Poisson errors does indeed appear to be due to lacking the sample variance contribution to the likelihood.

In addition to the errors on $\Omega_m$ and $\sigma_8$, we also compare the correlation coefficients using the three likelihoods and three binnings and present the results in Fig. 7. For each binning setup we find a negative correlation coefficient between the two cosmological parameters, which is the expected degeneracy between $\Omega_m$ and $\sigma_8$ with cluster abundance cosmology. Additionally we compute the correlation from the Fisher forecast and the ensemble forecasts. We find that the correlation is accurate for the Poisson likelihood, the blue dots are close to the red crosses. The correlation is slightly under-estimated for the GPC and Gaussian likelihoods (at most on the order of 5% in the case of medium binning). This is likely because only the block-diagonal terms are used in our covariance matrices for these likelihoods. This correlation would be important when calculating $S_8 = \sigma_8 \sqrt{\Omega_m}/0.3$, which is often used to quantify the tension between CMB and large-scale structure cosmological probes (see Di Valentino et al. (2021) for a recent review). A 5% bias on the correlation coefficient, like we see for the GPC and Gaussian cases, would lead to a bias of 1% on the error of $S_8$ and is therefore unimportant.

Finally, in the three binning regimes where the tests above were undertaken, the Gaussian and GPC likelihood always obtain the same parameter constraints. The GPC provides, in principle, a more accurate description of the statistical behaviour of the data, but the Gaussian approximation is clearly sufficient in the three use cases.

The Gaussian likelihood remains an accurate description of the data, up to a very large number of bins, when the cluster analysis is performed with clusters above $10^{14} \, M_\odot$ in a Rubin or Euclid-like survey. This is despite the bins being poorly populated. We would have expected that at a sufficiently fine binning that the Gaussian likelihood would no longer be valid, which we do not see in our results. But even with the fine binning, there are still a significant number of bins containing $\sim 100$ halos (see right panel in Fig. 2). The answer then is that if we chosen an even finer binning that we would reach a regime in which the Gaussian approximation was no longer valid. However as we see little improvement between the cos-
Likelihoods for cluster abundance cosmology

Figure 6. Errors on $\Omega_m$ (top) and $\sigma_8$ (bottom) for the different binning setups in Table 1 (columns). For each plots, the chosen likelihood is given on the x-axis while errors are on the y-axis. Blue dots correspond to the average of individual errors, and the error-bars represent the standard deviation of these individual errors. Red crosses correspond to the standard deviation of the 1000 best fits. Black circles correspond to Fisher forecasts in Eq. (18) (for the Poissonian and Gaussian likelihoods) and black boxes to Ensemble forecast in Eq. (20) (only for the Poisson likelihood, see text for details).

Figure 7. Correlation coefficients between $\Omega_m$ and $\sigma_8$ for the different binning setups in Table 1 (left to right). For each plot, the chosen likelihood is indicated on the x-axis while correlation is on the y-axis. Blue dots correspond to the average of individual correlations, and the error-bars represent to the standard deviation of these individual correlations. Red crosses correspond to the correlation of the 1000 two-dimensional best fits. Black circles correspond to Fisher forecasts in Eq. (18) (for the Poissonian and Gaussian likelihoods) and black boxes to the ensemble forecast in Eq. (20) (only for the Poisson likelihood, see text for details).

7.3 Reduced volume sample

As shown in Section 2, the Gaussian and Poisson likelihoods are two limit cases of the GPC description. In the previous section, where the number of clusters per bin is large, the GPC and Gaussian descriptions behave similarly and give robust constraints, conversely...
to the Poisson case. In this section and the next, we perform the following exercise: we reduce the number of clusters per bin to see if we can reach a regime where GPC and Poisson would give similar results and if those results could be more robust than the Gaussian case. To do so, we perform here another analysis in which we reduce the sky fraction by a factor 10, therefore reducing the total number of halo to $\sim 10^4$ per simulation.

We show the results of the three likelihoods to this reduced volume sample using the fine binning in Fig. 8. By reducing the sky fraction, errors on $\Omega_m$ and $\sigma_8$ increased by a factor of $\sim 1/f_{sky}$ compared to the constraints obtained using the full PINOCCHIO sky area in Section 7.2. The Poisson likelihood still underestimates errors by $\sim 25-30\%$, and the Gaussian and GPC individual errors are closer to the ensemble errors, but negatively biased by a factor $\sim 5\%$.

### 7.4 Considering high mass clusters only

In the same spirit of Section 7.3, we now reduce the number of objects by selecting high mass clusters only by selecting $M > 5.10^{14} M_\odot$, but keeping the full PINOCCHIO sky area. We repeat the analysis with 100 mass bins this defined from $5.10^{14} M_\odot$ to $10^{15.6} M_\odot$ (we keep the upper limit the same as defined in Section 7), and 100 redshift bins defined in Table 1. The total number of halos used in the analysis now drops to $\sim 5 \times 10^3$. We show in Fig. 8 (second line) the results of this work. We see that the Poisson likelihood still underestimates error on cosmological parameters from $5-20\%$ (depending on the cosmological parameter). With this result, we see that off-diagonal terms of the sample covariance and the sample variance are still important. The Gaussian and GPC individual errors are fully consistent with the ensemble errors. In this regime and under our model assumptions, Gaussian and GPC describe correctly the latent likelihood.

### 8 CONCLUSIONS

Cluster abundance cosmology relies on a likelihood to extract cosmological information from the data that has been assumed in the past to follow a Poisson, Gaussian or the Gauss-Poisson Compound distribution. In the context of the forthcoming cluster surveys, such as the Rubin Legacy Survey of Space and Time or the Euclid survey, it is important to consider if these likelihoods are sufficiently accurate and robust.

To do so, we make use of the PINOCCHIO simulation set, that provides us with 1000 realisations of $10,313 \text{deg}^2$ (i.e. 1/4th off the full sky), for a given cosmology. For each likelihood and for each simulation, we perform a cosmological analysis using importance sampling to recover posteriors for $\sigma_8$ and $\Omega_m$. The comparison of the average and covariance of the 1000 posterior means to that of the individual posteriors, provides us with the metric to assess the reliability of the likelihood under scrutiny. This is repeated for three binning schemes in the mass-redshift plane, ranging from 16 to $10^4$ bins. Therefore in this work we have sampled a total of 9000 posteriors. For each binning and each likelihood, we find a consistent bias in the recovered parameters despite this idealized set up. This bias is small with respect to the individual posterior variance and will therefore be negligible for future cosmological analyses.

For the three binning schemes we have explored, over the full survey area, the Poisson likelihood always underestimates the errors on the parameters by $\sim 30\%$; this is unsurprising as the sample variance.
cannot be accounted for in that description. Reducing the sky fraction to ~ 1030 deg\(^2\) or limiting the analysis to the most massive clusters does not change that conclusion despite yielding bins populated by less than five clusters. The Gaussian and Gauss-Poisson Compound likelihoods perform similarly in all cases, underestimating the errors on the recovered parameters by at most 5%.

Increasing the number of bins from the coarse (16 bins) to fine (10\(^4\) bins) cases only yields a marginal ~10% improvement on \(\sigma_8\) and \(\Omega_m\) error bars. The advantage of a finer binning may therefore disappear in a real analysis when using a Gaussian likelihood (Penna-Lima et al. 2014; Artis et al. 2021; Sartoris et al. 2016). This may become clearer in the future as we are also performing a study about the impact of binned versus unbinned cluster count likelihoods on cosmological constraints (complementary to this work), focusing on the calibration of mass-proxy relations. Unbinned cluster count likelihoods and finer binnings may also be useful when studying the correlations between different observable properties of galaxy clusters; such as the cluster richness, weak lensing mass, X-ray mass and SZ masses (for example see Rozo et al. 2009; Mahdavi et al. 2013; Umetsu et al. 2020; Murray et al. 2022).

This analysis is also the opportunity to revisit the question of forecasts. We compute the Fisher forecasts for the Poisson and Gaussian likelihoods, finding them (as expected) in agreement with the individual posterior covariances. We also develop a forecast approach that, conversely to the Fisher forecast, allows us to account for a latent likelihood that is not necessarily the same as the analysis likelihood. In our work this allowed us to show that the difference between the Poisson individual posterior errors and the ensemble errors can be explained by the non-inclusion of the sample covariance in the Poisson likelihood.

In more general terms, this new type of forecast, which in this work we call the ensemble forecast, would allow us to assess the reduction of the constraining power of an approximate likelihood in the case where the latent likelihood is known. This would be useful where the latent likelihood is computationally expensive to sample, for example, with many off-diagonal terms in the covariance matrix.

To conclude, from this work and for the general purpose of cluster abundance cosmology with future large scale surveys, the Gaussian likelihood remains a sufficient description of the data and we find no clear benefit to make the analysis more complex and slower using the GPC description or a large number of bins.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewer for their insightful comments and suggestions. We thank Pierluigi Monaco and Alex Saro for providing the PINOCCHIO catalogs, discussions of the work and providing useful comments on the manuscript. We thank Fabien Lacasa for his help in using the PySSC code. CM acknowledges funding from the French ”Programme d’investissements d’avenir” through the Enigmass Labex. We gratefully acknowledge support from the CNRS/IN2P3 Computing Center (Lyon - France) for providing computing and data-processing resources needed for this work.

DATA AVAILABILITY

The PINOCCHIO simulations used in this article cannot be shared publicly. The data will be shared on reasonable request to Alessandra Fumagalli (alessandra.fumagalli@inaf.it).

REFERENCES

Abbott T. M. C., et al., 2020, Phys. Rev. D, 102, 023509
Artis E., Melin J.-B., Bartlett J. G., Murray C., 2021, A&A, 649, A47
Bocquet S., et al., 2019, ApJ, 878, 55
Bond J. R., Myers S. T., 1996, ApJS, 103, 1
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bouchet F. R., Colombi S., Hivon E., Juszkiewicz R., 1995, A&A, 296, 575
Buchert T., 1992, MNRAS, 254, 729
Chisari N. E., et al., 2019, The Astrophysical Journal Supplement Series, 242, 2
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Cramer R., 1952, Advanced statistical methods in biometric research. John Wiley & Sons, Inc.
Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2016, MNRAS, 456, 2486
Di Valentino E., et al., 2021, Astroparticle Physics, 131, 102604
Dodelson S., Schneider M. D., 2013, Phys. Rev. D, 88, 063537
Efron B., 2015, Journal of the Royal Statistical Society, 77, 617
Eisenstein D. J., Loeb A., 1995, ApJ, 439, 520
Elvira V., Martino L., 2021, arXiv e-prints, p. arXiv:2102.05407
Evrard A. E., et al., 2002, ApJ, 573, 7
Fisher R. A., 1935, Journal of the royal statistical society, 98, 39
Fumagalli A., et al., 2021, A&A, 652, A21
Gouyou Beauchamps S., Lacasa F., Tutusaus I., Aubert M., Baratta P., Gorce A., Sakr Z., 2022, A&A, 659, A128
Heavens A. F., Seikel M., Nord B. D., Aich M., Bouffanais Y., Bassett B. A., Hobson M. P., 2014, MNRAS, 445, 1687
Henson M. A., Barnes D. N., Kay S. T., McCarthy I. G., Schaye J., 2017, MNRAS, 465, 3361
Hu W., Kravtsov A. V., 2003, The Astrophysical Journal, 584, 702–715
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Kaiser N., 1984, ApJ, 284, L9
Kh H. H., 1996, NIST, 70, 263
LSST Science Collaboration et al., 2009, arXiv e-prints, p. arXiv:0912.0201
Lacasa F., Lima M., Aguena M., 2018, Astronomy & Astrophysics, 611, A83
Laureijs R., et al., 2011, arXiv e-prints, p. arXiv:1110.3193
Lesco G. F., et al., 2022a, arXiv e-prints, p. arXiv:2203.07398
Lesco G. F., et al., 2022b, A&A, 659, A88
Lima M., Hu W., 2004, Phys. Rev. D, 70, 043504
Mahdavi A., Hoekstra H., Babul A., Bildfell C., Jeltema T., Henry J. P., 2013, ApJ, 767, 116
Mantz A. B., et al., 2015, MNRAS, 446, 2205
Monaco P., 1997, MNRAS, 287, 753
Monaco P., Theuns T., Taffoni G., 2002, MNRAS, 331, 587
Moutarde F., Alimi J. M., Bouchet F. R., Pellat R., Ramani A., 1991, ApJ, 382, 377
Munari E., Monaco P., Sefusatti E., Castorina E., Mohammad F. G., Anselmi S., Borgani S., 2017, MNRAS, 465, 4658
Murray C., Bartlett J. G., Artis E., Melin J.-B., 2022, MNRAS, 512, 4785
Penna-Lima M., Makler M., Wuensche C. A., 2014, J. Cosmology Astropart. Phys., 2014, 039
Percival W. J., Friedrich O., Sellentin E., Heavens A., 2022, MNRAS, 510, 3207
Planck Collaboration et al., 2014, A&A, 571, A16
Planck Collaboration et al., 2016, A&A, 594, A24
Press W. H., Schechter P., 1974, ApJ, 187, 425
Raveri M., Hu W., 2019, Physical Review D, 99
Rozo E., et al., 2009, ApJ, 699, 768
Sartoris B., et al., 2016, MNRAS, 459, 1764
Sellentin E., Heavens A. F., 2017, MNRAS, 464, 4658
Sellentin E., Heavens A. F., 2018, MNRAS, 473, 2355
Sellentin E., Quarinni M., Amendola L., 2014, MNRAS, 441, 1831
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Takada M., Spergel D. N., 2014, Monthly Notices of the Royal Astronomical Society, 441, 2456–2475
Taylor A., Joachimi B., 2014, MNRAS, 442, 2728
APPENDIX A: SAMPLE VARIANCE WITH PYSSC

The sample covariance $\Sigma_{SV}$ between two counts $N_{ij}$ and $N_{kl}$ is given by (Hu & Kravtsov 2003)

$$\Sigma_{SV}[ij][kl] = \int_{z_j}^{z_{j+1}} dV_1 \int_{z_k}^{z_{k+1}} dV_2 \frac{\partial O_i(z_1)}{\partial \delta} \frac{\partial O_k(z_2)}{\partial \delta} \sigma^2(z_1, z_2),$$

(A1)

where the quantity

$$\frac{\partial O_i(z)}{\partial \delta} = \int_{m_i}^{m_{i+1}} dm b(m, z) \frac{dn(m, z)}{dm}$$

(A2)

is the response of the probe $O_i$ defined as

$$O_i(z) = \int_{m_i}^{m_{i+1}} dm \frac{dn(m, z)}{dm} (1 + b(m, z) \delta)$$

(A3)

relative to the background density field. The quantity $\sigma^2(z_1, z_2)$ is the amplitude fluctuation of the matter power spectrum between two different redshifts and depends on the survey window function. Considering the full sky area, $\sigma^2(z_1, z_2)$ is defined as

$$\sigma^2(z_1, z_2) = \frac{1}{2\pi} \int k^2 dk j_0(kw(z1))j_0(kw(z2))P_m(k|z_1, z_2),$$

(A4)

where $w$ is the line-of-sight comoving distance, and $j_0$ is the spherical Bessel function of order 0, $P_m(k|z_1, z_2)$ is the matter power spectrum at redshift $z_1$ and $z_2$. Assuming $\frac{\partial O_i(z)}{\partial \delta}$ varies slowly with redshift compared to the matter fluctuation amplitude, we get the approximation (Lacasa et al. 2018)

$$\Sigma_{SV}^{th}[ij][kl] \approx \langle b \lambda \rangle_{ij} \langle b \lambda \rangle_{kl} S_{ij},$$

(A5)

where $\langle b \lambda \rangle_{ij}$ is given in Eq. (25). The matrix $S_{ij}$ is the average amplitude fluctuation over the two separate volume and is defined by

$$S_{ij} = \frac{1}{V_j V_i} \int_{z_j}^{z_{j+1}} dV_1 \int_{z_k}^{z_{k+1}} dV_2 \sigma^2(z_1, z_2).$$

(A6)

In practice, we use the PySSC package to compute $S_{ij}$ matrix in Eq. (A6) (Lacasa et al. 2018). We then rescale it by the sky fraction and use $S_{ij} = S_{ij}^{fullsky} / f_{sky} = S_{ij}^{fullsky} \times 4\pi / \Omega_S$ (Lacasa et al. 2018).

While this is only an approximation, Gouyou Beauchamps et al. (2022) shows that this performs well for sufficiently large sky areas, such as the one covered by the PINOCCHIO simulations. To check our use of PySSC, we compare the full theoretical prediction of the covariance, given by $\Sigma = \Sigma^{th} + \Sigma^{SN}$ ($\Sigma^{SN}$ is given in Eq. (25)), to the covariance directly estimated from the data. The estimated full covariance $\hat{\Sigma}_{ij}[kl]$ between 2D mass-redshift bins $ij$ and $kl$ calculated from the 1000 simulations,

$$\hat{\Sigma}_{ij}[kl] = \frac{1}{N_{\text{sim}} - 1} \sum_{n=1}^{N_{\text{sim}}} \left( \hat{N}_{ij} - \langle \hat{N}_{ij} \rangle \right) \left( \hat{N}_{kl} - \langle \hat{N}_{kl} \rangle \right),$$

(A7)

where $N_{\text{sim}}$ is the number simulations, $\hat{N}_{ij}$ is the measured number of clusters in bin $ij$ for the $n$-th PINOCCHIO simulation, and $\langle \hat{N}_{ij} \rangle$ is the cluster abundance measured in bin $ij$ average over the 1000 simulations.

The left panel in Fig. A1 shows the correlation matrix, defined as,

$$R_{ij}[kl] = \frac{\Sigma_{ij}[kl]}{\sqrt{\Sigma_{ij}[ij] \Sigma_{kl}[kl]}}$$

(A8)

for the data-derived covariance (lower triangle) and the theoretical covariance prediction (upper triangle). This is shown for the coarsest binning scheme (see Section 6.1 and Fig. 2) with the correlation matrix organised in four redshift bins, appearing as block diagonal, inside which the four mass bins are added.

We find the correlation between bins to be well-represented by the theoretical prediction for the block diagonal terms and the first off-diagonal terms. However, for more distant correlations, statistical noise dominates the data-derived covariance matrix (lower triangle) while the correlation drops to zero in the theoretical estimation (upper triangle).

To further compare the two estimations, we compute the relative difference of the variance (diagonal elements) between the two approaches, defined as

$$b_{ij} = \frac{\Sigma_{ij}^{data}[ij][ij] - \Sigma_{ij}^{th}[ij][ij]}{\Sigma_{ij}^{th}[ij][ij]}$$

(A9)

The results are represented as a function of the mean abundance, for the 16 diagonal elements of the $4 \times 4$ coarse binning scheme, in the right panel in Fig. A1. We find a bias of about 10%, which is expected according to the precision of PINOCCHIO two-point statistics.

This check validates our computation of the theoretical covariance, which we favour in our analysis as the data-derived covariance is affected by statistical noise due to the limited number of samples.

APPENDIX B: VARIANCE OF ESTIMATORS

In this appendix, we forecast the ensemble covariance matrix in Eq. (19) in terms of both statistical properties of the data encoded in the input and latent likelihoods. We also show that the ensemble and individual covariance matrices are equivalent when the input and latent likelihoods are equivalent when we consider the simple case of Gaussian likelihoods.

From Efron (2015), we have that the gradient of the estimator $\nabla \hat{\theta}_\alpha$ with respect to the data is given by

$$[\nabla \hat{\theta}_\alpha]_i = \text{Cov}_{\phi_Y \theta} \left( \{ \theta_\alpha, \nabla \ln L_{\gamma}(N|\theta) \}_i \right) |N|.$$  

(B1)

From Eq. (B1), we can derive the forecast for the ensemble covariance matrix, depending on the properties of the analysis likelihood. We first study the case of a Gaussian analysis likelihood, and the case of a Poisson likelihood.
B1 Gaussian analysis likelihood

The gradient $\nabla \ln L_Y$ can be expressed for the particular case of a Gaussian likelihood $L_Y$ with an analysis covariance matrix $\Sigma_Y$, and we have that

$$\nabla \ln L_Y[N(\theta) = \Sigma_Y^{-1}[\lambda(\theta) - N]].$$

(B2)

Using Eq. (B2) in Eq. (B1), we get the general expression, which is valid for an input Gaussian likelihood $L_Y$ that

$$\nabla \tilde{\theta}_\alpha(N) = \Sigma_Y^{-1}[\lambda(N) - N].$$

(B3)

Second, we consider the parameter estimator $\tilde{\theta}(N)$, at fixed observed cluster count $N$. We now consider that the observed cluster count $N$ corresponds to the latent average cluster count $N_0$ with small additive deviation, i.e. we can write $N = N_0 + \delta N$. In this formalism, we state that $\theta_0$ are the latent cosmological parameters giving $N_0 = \lambda(\theta_0)$, and we state that $N_0 = \mathbb{E}_Y(N|\theta_0) = \mathbb{E}_X(N|\theta_0)$. Using the expression of the posterior in Eq. (13), of the estimator in Eq. (15) and the posterior variance in Eq. (14), we consider the Taylor expansion of the estimator $\tilde{\theta}(N_0 + \delta N)$ considering the deviation $\delta N = N - N_0$, so we get for the first order that

$$\tilde{\theta}_\alpha(N) - \tilde{\theta}_\alpha(N_0) \approx \nabla \tilde{\theta}_\alpha |_{\theta_0} \cdot [N - N_0],$$

(B4)

where in Eq. (B4), $\nabla \tilde{\theta}_\alpha |_{\theta_0}$ corresponds to the gradient defined in Eq. (B1) evaluated at $N = N_0$. For Eq. (B4) to be valid, it is assumed that the estimator $\tilde{\theta}(N)$ is sufficiently linear around $N_0$ with respect to the data, as it is considered for linear error propagation from data to estimated parameter (Ku 1996). We now assume that the model is well approximated by a linear dependence on the parameter of interest, e.g. Raveri & Hu (2019), and taking the first order in the Taylor expansion of $\lambda(\theta)$ at $\theta_0$, we get that

$$\nabla \tilde{\theta}_\alpha |_{\theta_0} = \Sigma_Y^{-1} \lambda_\gamma C^{-1}_{\gamma \alpha}(\lambda_\gamma |_\theta_0).$$

(B5)

where $C^{-1}_{\gamma \alpha} = \text{Cov}_{\gamma \alpha}([\theta_{\gamma}, \theta_\alpha]_0)$ is the posterior covariance matrix as defined in Eq. (14), and $\lambda_\gamma$ is the derivative of the model prediction evaluated at $\theta = \theta_0$. Note that we subsequently preserved the first order moments of 2 Taylor expansions respectively in the data and parameter space, that need to be valid to follow the proof. We will consider these approximations correct for the following steps. Re-writing the expression of the ensemble covariance in Eq. (19), we find that

$$C^{\text{ens}}_{\alpha \beta} = \text{Cov}_{X_X}([\tilde{\theta}_\alpha(N), \tilde{\theta}_\beta(N)]|\theta_0)$$

(B6)

$$= (C^{\text{ind}}\lambda_\alpha)^T \Sigma_{\gamma}^{-1} \Sigma_{\gamma} \Sigma_{\gamma}^{-1} (C^{\text{ind}}\lambda_\beta).$$

(B7)

In Eq. (B7), the statistical properties of the data assumed by choosing $L_Y$ appears in $\Sigma_{\gamma}^{-1}$, as well as in the individual parameter covariance computed using $\mathcal{P}_Y$. Whereas the latent statistical properties of data appears in $\Sigma_X$. The Eq. (B7) is an application of the delta method approximation method for error propagation (Efroymson 2015, see their Eq. (2.8)) for the Gaussian case, but also a generalization since we take account of covariance between parameters and different latent and analysis data covariances respectively $\Sigma_X$ and $\Sigma_Y$. We can explore the solution when, in this formalism, we get $C^{\text{ens}} = C^{\text{ind}} = C$, i.e. when the covariance of the estimator when considering an infinite realisation of the data $N$ is consistent with the individual parameter covariance obtained from a single dataset $N$. Following the Gaussian approximation of individual posteriors in the parameter space, we can use the Fisher forecast defined in Eq. (18) for $C^{\text{ind}}$, since the Eq. (B7) is evaluated at $N = N_0$. Using this, we can simplify the problem by trying to find the relation between the data covariances $\Sigma_X$ and $\Sigma_Y$ to show that

$$C_{\alpha \beta}^{\text{F}} = (C\lambda_\alpha)^T \Sigma^{-1} (C\lambda_\beta)$$

(B8)

$$C_{\alpha \beta}^{\text{F}} = \lambda_{\alpha \beta}^{-1} \Sigma_{\alpha \beta}^{-1}$$

(B9)

where we defined the modified covariance matrix $\Sigma^{-1} = \Sigma_Y^{-1} \Sigma_X^{-1}$ in Eq. (B8), and Eq. (B9) is given by Fisher forecasting the individual parameter covariance in Eq. (18), that is always true for Gaussian and Poisson likelihoods. Multiplying Eq. (B8) by $C^{-1}_{\gamma \alpha}$ and summing over $\alpha$, we get

$$\delta_{\gamma \beta} = \lambda_{\gamma \beta}^{-1} (C\lambda_\gamma)^T \Sigma^{-1} \lambda_\beta.$$

(B10)
By re-multiplying the above equation by $C^{-1}_{\alpha\beta}$, we get the result
\[ C^{-1}_{\alpha\beta} = \Lambda^T \Sigma^{-1} \Lambda, \]  
which is the same as Eq. (B9) when setting $\Sigma_X = \Sigma_Y$. So we found that for the ensemble parameter covariance to be equal to the individual parameter covariance, this formalism requires that the data covariance of the analysis and latent likelihoods are the same. In this appendix, we considered only the first order in the Taylor expansion in Eq. (B4). When posteriors are not Gaussian, this approximation is not valid.

**B2 Poisson analysis likelihood**

In the previous section, we found that the ensemble parameter and individual parameter covariances are equal when $\Sigma_X = \Sigma_Y$, when the analysis likelihood $L_Y$ is a multivariate Gaussian distribution defined in Eq. (11) with data covariance $\Sigma_Y$, and where the latent likelihood is not specified, but with data covariance matrix $\Sigma_X$. In this section, we consider the case where the analysis likelihood $L_Y$ is the Poissonian as defined in Eq. (10). To compute the ensemble forecast, we first need to evaluate $[\nabla \tilde{\theta}_\alpha]_i$ in Eq. (B1). First, we have that the Poisson log-likelihood is,
\[ \ln L_Y(N|\theta) = -\lambda_{tot}(\theta) + \sum_{k=1}^c N_k \ln \lambda_k(\theta) - \ln(N_k!), \]  
where $\lambda_{tot}(\theta)$ is the total predicted abundance over all the redshift bins. As given in Eq. (B2), we take the first derivative of the log-likelihood with respect to the data vector component $N_i$, giving
\[ \nabla_i \ln L_Y(N|\theta) = \ln \lambda_i(\theta) - \psi(N_i), \]  
where $\psi$ is known as the digamma function and is defined as
\[ \psi(N_i) = -\gamma + \sum_{k=1}^{N_i} \frac{1}{k}, \]  
where $\gamma \approx 0.577$ is the Euler–Mascheroni constant, and the sum on the right side is the $m$-th harmonic number. We note that the quantity $\ln(N_k) - \psi(N_k)$ is independent of $\theta$, so it can be removed from the calculation of the covariance in Eq. (B1). $[\nabla \tilde{\theta}_\alpha]_i$ now verifies the general form
\[ [\nabla \tilde{\theta}_\alpha]_i = \text{Cov}_{\gamma}[\{\theta_\alpha, \ln \lambda_i(\theta)\}|N]. \]  
Next, taking the first order of the Taylor expansion of $\lambda_i(\theta)$ with respect to $\theta$, we get $\lambda_i(\theta) = (\lambda_0)_i + (\lambda_i)_i(\theta - \theta_0)$. Using $\ln(1 + x) \approx x + o(x^2)$, we get finally that
\[ [\nabla \tilde{\theta}_\alpha]_i = \frac{1}{(\lambda_0)_i} C^{\text{ind}}_{\alpha\gamma}(\lambda, \gamma) = [\Sigma_Y^{-1}]_i C^{\text{ind}}_{\alpha\gamma}(\lambda, \gamma)_i, \]  
where $\Sigma_Y = \text{diag}[[\{\lambda_0\}_i]_{1 \leq i \leq c}]$ is the covariance matrix of the data considering the Poisson distribution in Eq. (10) (in that case, abundances are not correlated and the variance of each $k$-th cluster count is the shot noise $(\lambda_0)_k$), and $C^{\text{ind}}_{\alpha\gamma} = \text{Cov}_{\gamma}[\{\theta_\alpha, \theta_\beta\}|\lambda_0]$ is the posterior covariance matrix as defined in Eq. (14). We find the same expression as for the Gaussian case in Eq. (B5). To test the condition $C^{\text{ens}} = C^{\text{ind}}$ for the Poisson analysis likelihood, we first need to compute the Fisher forecast for the Poisson case. From the Poisson likelihood in Eq. (B12), we have that the second derivative of the Poisson log-likelihood writes
\[ (\ln L)_\alpha\beta = -(\lambda_{tot})_{\alpha\beta} + \sum_{k=1}^c N_k \left( \frac{\lambda_k(\lambda_\beta)_k}{\lambda_k^2} - \frac{(\lambda_\alpha)_k(\lambda_\beta)_k}{\lambda_k^2} \right), \]  
Following the Fisher matrix general formula in Eq. (17) and using the equality
\[ (\ln L)_\alpha\beta = \frac{L_{\alpha\beta}}{L} - (\ln L)_\alpha (\ln L)_\beta, \]  
we get after few calculations the Fisher information matrix for the Poisson likelihood that writes
\[ F_{\alpha\beta} = \sum_{k=1}^c \frac{(\lambda_k)_k(\lambda_\beta)_k}{(\lambda_0)_k^2} = \Lambda^T \Sigma_Y^{-1} \Lambda, \]  
where $C^{\text{ind}}$ can be forecasted as $F^{-1}$. We see that the forecasted ensemble and individual parameter covariances for the Poisson analysis likelihood are obtained using the diagonal data covariance $\Sigma_Y$ with only shot noise in Eq. (18) for $C^{\text{ind}}$ and in Eq. (20) for $C^{\text{ens}}$. As shown in Appendix B1, the equality $C^{\text{ind}} = C^{\text{ens}} = C$ is then verified when $\Sigma_X = \Sigma_Y$ (result obtained following the derivation from Eq. (B8) to Eq. (B11)).

This paper has been typeset from a T\textsc{ex}/L\textsc{tex} file prepared by the author.