Beware of commonly used approximations II: estimating systematic biases in the best-fit parameters

José Luis Bernal, Nicola Bellomo, Alvise Raccanelli, Licia Verde

Abstract. Cosmological parameter estimation from forthcoming experiments promise to reach much greater precision than current constraints. As statistical errors shrink, the required control over systematic errors increases. Therefore, models or approximations that were sufficiently accurate so far, may introduce significant systematic biases in the parameter best-fit values and jeopardize the robustness of cosmological analyses. We present a general expression to estimate a priori the systematic error introduced in parameter inference due to the use of insufficiently good approximations in the computation of the observable of interest or the assumption of an incorrect underlying model. Although this methodology can be applied to measurements of any scientific field, we illustrate its power by studying the effect of modeling the angular galaxy power spectrum incorrectly. We also introduce Multi_CLASS, a new, public modification of the Boltzmann code CLASS, which includes the possibility to compute angular cross-power spectra for two different tracers. We find that significant biases in most of the cosmological parameters are introduced if one assumes the Limber approximation or neglects lensing magnification in modern galaxy survey analyses, and the effect is in general larger for the multi-tracer case, especially for the parameter controlling primordial non-Gaussianity of the local type, $f_{NL}$. 
An unprecedented experimental and theoretical effort during the last decades has established the Λ-Cold Dark Matter (ΛCDM) model as the standard model of cosmology, because of its striking precision in fitting most of the available observations. This effort has brought forward percent-level constraints on some of the parameters of the model (see e.g., [1–4]). Nonetheless, there are still some remaining tensions between experiments (see e.g., [5–8]) that, in the absence of non-accounted for systematic errors, might hint at the need for an extension of ΛCDM. Forthcoming and future galaxy surveys are expected to push the envelope of observational cosmology on the large scale structure side [9–14], but significant improvements are also expected for CMB experiments [15–17]. Moreover, the advent of line-intensity mapping experiments [18, 19] with its potential for cosmology (see e.g., [20]) promises to open a window to explore the Universe at higher redshift and close the gap between CMB and galaxy surveys observations (see e.g., [21, 22]).

The aim of these experiments is to achieve subpercent-level precision for parameter inference within ΛCDM and to further constrain (and possibly even detect) deviations from ΛCDM. The promising capabilities of these experiments will make possible a dramatic reduction of the statistical errors: expected statistical uncertainties are well below the current systematic error budget. Therefore, it is of crucial importance to maintain systematic biases below the statistical errors as well as a correct assessment of the final, true uncertainties. Accurate modeling of the target observables is one of the key ingredients needed in order to succeed in this challenge. However, higher accuracy often requires more complicated and expensive modeling. This can be mitigated by introducing careful approximations in the modeling of the signal and their covariances to speed up calculations without significant reductions in accuracy.

Using insufficiently accurate approximations in the analysis affects the inferred model parameters in two ways. First, it may shift the point in parameter space where the posterior distribution peaks, which would introduce a systematic bias in the best-fit parameters. Second, the shape of the posterior distribution may be affected, yielding a misestimate of the uncertainties. In summary, these two effects can be understood as errors in the parameters
values and in their error-bars. Therefore, the effects on the posterior (in terms of potential shifts of the best-fit parameters and their errors) of any approximation to be adopted must be estimated quantitatively in light of the forecasted experimental performance. Moreover, a theoretical systematic error might be introduced even if the modeling of the observable is accurate enough, when the underlying cosmological model assumed is incorrect: the inferred cosmological parameters assuming ΛCDM will be most likely biased if the Universe is better described by a different model.

We approach the estimation of the bias in the best-fit parameters in the same spirit of previous works focusing on specific problems in cosmology. Some studies focused on the bias introduced by assuming an incorrect cosmological model $M_1$ instead of the correct one $M_0$, where $M_1$ and $M_0$ are nested models [23]. The general expression for nested models can be found in Refs. [24, 25], where it was used to estimate the impact of not marginalizing over a subset of nuisance parameters. Others investigated the bias arising from an incorrect modeling of the target observable for several specific cases, including the impact of inhomogeneous reionization on CMB anisotropies [26], SN magnitudes [27], the modeling of redshift-space distortions of the galaxy power spectrum [28], the baryonic feedback on weak lensing measurements [29], emission-line galaxy power spectrum measurements without including the contamination due to line interlopers [30], the contribution of relativistic effects on galaxy surveys [31–33], and the neutrino-induced scale-dependence of the galaxy bias [34]. There are also studies focused on how parameter inference is affected by misestimations of the covariance matrices (see e.g., [35, 36]) or the use of an invalid likelihood (see e.g., [37, 38]).

Instead, in this work we aim to generalize the methodology to be applicable to any observable, not necessarily in cosmology. There are two different causes that lead to parameter biases: the assumption of an incorrect model (e.g., assuming a cosmological constant in the case dark energy is dynamical) and the use of an incorrect or incomplete modeling of the observable considered (e.g., using an inaccurate modeling of the redshift-space distortions). Although the previous studies mentioned above considered only one of them, these two causes may be interdependent. Our approach treats both of them on the same footing, accounting for any possible interaction between them.

Our approach can be summarized as follows. We consider data drawn from an underlying model. Then, we expand the theoretical prediction of the observable, computed according to a given model under study, around a fiducial set of parameters. Note that the model under study is not necessarily the same model the data are drawn from. Subsequently, we apply this expansion to the likelihood of the observable, maximize such likelihood, and obtain the best-fit value of the parameters. We follow this procedure in two cases: a correct description of the observable and an approximated one. From here, we estimate the systematic bias comparing the best-fit parameters obtained in both cases.

This technique, as is the case for Fisher matrix analyses, is conceived to be used prior to obtaining data. In this case, it is required to choose a model as a good representation of reality and compute the data according to it. Employing our approach before the data is obtained will allow for the determination of the level of accuracy needed in the data analysis of an experiment during its design. Moreover, it will help to quantify the significance of new contributions or corrections to the theoretical modeling of the signal in light of the potential of future experiments. Nonetheless, the methodology can also be applied to actual observations to ascertain whether a more numerically intensive modeling is needed.

We demonstrate the power of our approach by applying our methodology to the angular galaxy power spectrum. We primarily focus on the systematic shifts of the best-fit param-
eters, whereas the misestimation of the uncertainties in parameter inference is studied in a companion paper [39], hereinafter Paper I. Besides traditional, single-tracer analysis, we also consider multi-tracer analyses of the galaxy clustering. We present two practical examples: we explore how neglecting cosmic magnification or using the Limber approximation may introduce significant systematic biases ($\gtrsim 1 - 2\sigma$ in most of the parameters and cases under study) in the analyses of next-generation galaxy surveys.

We also present and publish Multi_CLASS. Multi_CLASS is a modification of the public Boltzmann code CLASS [40, 41] that allows for the computation of the angular cross-power spectra of two different tracers of the underlying density field, with their corresponding different redshift dependence of the tracer characteristics (bias, magnification bias, evolution bias and number density distribution). This possibility enables comprehensive theoretical multi-tracer analyses. Furthermore, it includes an implementation for primordial non-Gaussianities of the local type, parametrized by the $f_{NL}$ parameter, on the clustering observables. More details about the options and implementation of Multi_CLASS can be found in Appendix B of Paper I.

This paper is organized as follows: we present a general methodology to estimate the bias introduced in best-fit parameters by an inaccurate modeling of a given measurement in Section 2; introduce our test-case, the angular galaxy power spectrum, and our assumptions for the demonstration of the potential of our methodology in Section 3; show the estimated systematic bias introduced in cosmological parameter inference by using the angular galaxy power spectrum without modeling lensing magnification or applying the Limber approximation in Section 4; and conclude in Section 5.

2 Systematic shift of the best-fit parameters due to incorrect modeling

We begin by defining notation and conventions. We use vector operators for quantities in parameter space; operations involving the observable space (e.g., the data vector or its covariance) are explicitly written down as sums over the matrix elements. Unless otherwise stated, all vectors are column vectors. For the sake of clarity, the meaning of all symbols, superscripts and subscripts used in this section can be consulted in Table 1.

Let us consider a generic model $M$ specified by a set of parameters $\theta$ that, according to some theoretical modeling, determine the observable $\Psi$ and its covariance Cov. While in principle there is no need to distinguish between model and modeling, this is useful in some cases, like for instance, cosmology, where ‘model’ usually refers to the cosmological model (e.g., $\Lambda$CDM) and ‘modeling’ refers to the practical application of the model to describe the observables under study. In the case of cosmology, for example, there is a much wider variety of modeling approaches than of models. Assuming a Gaussian likelihood for $\Psi$ (which usually is a good approximation close to its maximum, or in case the central limit theorem applies), the logarithm of the likelihood given $M$ depends on $\theta$ as

$$-2 \log L(\Psi^d|M, \theta) = \sum_{i,j} \left(\Psi^d_i - \hat{\Psi}_i(\theta)\right) \left(\text{Cov}^{-1}(\theta)\right)_{ij} \left(\Psi^d_j - \hat{\Psi}_j(\theta)\right)^* + \log |\text{Cov}(\theta)|,$$

(2.1)

¹Multi_CLASS will be publicly available in https://github.com/nbellomo/Multi_CLASS upon the acceptance of this work.
where $\Psi^d$ and $\hat{\Psi}(\theta) \equiv \hat{\Psi}(\theta|M)$ are the data vector and the corresponding theoretical prediction of the observable, respectively, $i$ and $j$ denote vector and matrix indices, and the superscript ‘*’ indicates the complex conjugate operation. The constant terms in $\log L$, which are not included in Equation (2.1), do not affect parameter inference, hence we neglect them throughout this work. In the following, we assume real quantities for $\Psi$ in order to drop the notation referring to the complex conjugate. Nonetheless, we stress that our derivation is equally valid to complex observables.

Equation (2.1) considers the general case in which the covariance is not fixed but varies as function of the model parameters. For simplicity, in what follows we assume a fixed covariance (i.e., computed for a fiducial model with a specific choice of parameter values). Hence, the second term of the right hand side of Equation (2.1) is constant, and can be removed from the equation. It is straightforward to extend the methodology to the case of a parameter-dependent covariance matrix following the steps detailed below. Hereinafter we drop the explicit notation for the dependence on the parameters of the model.

For a specific set of parameters $\theta^\star$ within a given model, our target observable is $\hat{\Psi}^\star = \hat{\Psi}(\theta^\star)$. We can approximate the value of the observable under any small variation of the model parameters around $\theta^\star$ expanding to linear order as

$$
\hat{\Psi}_i(\theta^\star + \Delta \theta) \approx \hat{\Psi}_i^\star + \left( \nabla_\theta \hat{\Psi}_i^\star \right)^T \Delta \theta,
$$

(2.2)

where $\nabla_\theta$ denotes the gradient operator with respect to the model parameters, the superscript ‘$T$’ denotes the transpose operator, and $\Delta \theta$ is a small finite difference in the parameters space. This approximation is exact when $\Psi$ depends linearly on $\theta$; otherwise, its accuracy decreases as $\Delta \theta$ increases.

Let us assume that $M_0$, with parameters $\theta_0$, is the true underlying model for the phenomenon of interest. Then, $\Psi^d$ corresponds to a realization of $M_0$ given $\theta_0^\text{tr}$, with $\Psi(\theta_0^\text{tr}) = \langle \Psi^d \rangle$, where the superscript ‘$\text{tr}$’ refers to the parameter values corresponding to reality and and the brackets $\langle \cdot \rangle$ refer to the ensemble average. This is true in practice

$^2$In certain cases, like in cosmology, it is not always possible to obtain a meaningful ensemble average. In this scenario, we assume that $\theta^\text{tr}$ represent a faithful reproduction of the observed realization.
only if $\Psi^d$ is not contaminated with unaccounted-for systematics and an exact theoretical modeling is used to compute $\hat{\Psi}(\theta^d_M)$ and its covariance. On the other hand, we can consider a fiducial set of parameters $\theta^\text{fid}$ within $M$ guessed to be close to the point in parameter space where the likelihood peaks (e.g., inferred from prior or complementary experiments). Note that $M$ and $\theta^\text{fid}$ do not need to be the same model and parameters as $M_0$ and $\theta^\text{tr}_0$; then, in the most general case, we have $\hat{\Psi}(\theta^\text{fid}|M) = \hat{\Psi}^\text{fid} \neq \Psi^d$. We can expand $\hat{\Psi}^\text{fid}$ using Equation (2.2), apply it to Equation (2.1) and maximize the likelihood to obtain the best-fit parameters $\theta^\text{bf}$ for the model $M$. After the maximization, we have

$$
\sum_{i,j} \left( \nabla_{\theta} \hat{\Psi}^\text{fid}_i \left( \text{Cov}^{-1} \right)_{ij} \left[ \Psi^d_j - \hat{\Psi}^\text{fid}_j \right] \right) \Delta \theta = 0 \rightarrow
$$

$$
\sum_{i,j} \left( \nabla_{\theta} \hat{\Psi}^\text{fid}_i \left( \text{Cov}^{-1} \right)_{ij} \left( \Psi^d_j - \hat{\Psi}^\text{fid}_j \right) \right) = \sum_{i,j} \left( \nabla_{\theta} \hat{\Psi}^\text{fid}_i \left( \text{Cov}^{-1} \right)_{ij} \left( \nabla_{\theta} \hat{\Psi}^\text{fid}_j \right) \right) \Delta \theta,
$$

(2.3)

where we have neglected all derivatives of order higher than one. Solving Equation (2.3) for $\Delta \theta$ returns the step in parameter space needed in order to maximize the likelihood: $\Delta \theta \equiv \theta^\text{bf} - \theta^\text{fid}$. The factor multiplying $\Delta \theta$ in the right-hand side of Equation (2.3) can be identified as the Gaussian Fisher information matrix, whose elements are given by

$$
F_{a,b} = \left\langle \frac{\partial^2 \log L}{\partial \theta_a \partial \theta_b} \right\rangle = \sum_{i,j} \left( \frac{\partial \hat{\Psi}^\text{fid}_i}{\partial \theta_a} \right) \left( \text{Cov}^{-1} \right)_{ij} \left( \frac{\partial \hat{\Psi}^\text{fid}_j}{\partial \theta_b} \right),
$$

(2.4)

where $a$ and $b$ refer to indices of the parameters vector. Therefore, we can estimate the difference between the best-fit parameters and the fiducial parameters initially assumed as

$$
\Delta \theta \equiv \theta^\text{bf} - \theta^\text{fid} = F^{-1} \sum_{i,j} \left( \nabla_{\theta} \hat{\Psi}^\text{fid}_i \right) \left( \text{Cov}^{-1} \right)_{ij} \left( \Psi^d_j - \hat{\Psi}^\text{fid}_j \right).
$$

(2.5)

Of course, if $M$ is a good approximation of $M_0$ and the modeling used to compute $\hat{\Psi}^\text{fid}$ is accurate, $\theta^\text{bf}$ will be very close to $\theta^\text{tr}_0$.

We can apply the same procedure to a joint analysis of several likelihoods, corresponding to different observables, experiments or independent data sets. However, it is important to note that in general the global best-fit parameters $\theta^\text{bf}$ are different than the best-fit values $\theta^\text{bf}_\alpha$ for each independent likelihood. Let us consider independent likelihoods, so that the joint likelihood is the product of the individual likelihoods $L_\alpha$ (i.e., $\log L = \sum_\alpha \log L_\alpha$). In this scenario, Equation (2.5) is generalized as

$$
\Delta \theta = \left( \sum_{\alpha} F_\alpha \right)^{-1} \left[ \sum_{\alpha} \left( \nabla_{\theta} \hat{\Psi}^\text{fid}_{\alpha,i} \right) \left( \text{Cov}^{-1}_\alpha \right)_{ij} \left( \Psi^d_{\alpha,j} - \hat{\Psi}^\text{fid}_{\alpha,j} \right) \right],
$$

(2.6)

where we denote quantities referred to individual likelihoods with the subscript $\alpha$ (i.e., $\Psi_{\alpha,i}$ refers to the $i$-th element of the observable corresponding to the $\alpha$-th likelihood). If the likelihoods involved in the joint analysis have different nuisance parameters, $F_\alpha$ should be marginalized over the nuisance parameters not common between likelihoods, which will not be included in the parameters vector $\theta_\alpha$. Equation (2.6) can be straightforwardly generalized to non-independent likelihoods accounting for their covariance in the computation of both the Fisher matrix and the factor in square brackets.
Taking all this into account, we are now ready to compare the performance of a correct and an incorrect modeling of the observable, as well as the effects of assuming an incorrect underlying model. We assume, as before, that $\Psi^d$ is drawn from model $M_0$, but this model is unknown. We also consider two theoretical predictions of the observable, $\hat{\Psi}^C$ (correct) uses an accurate and precise modeling assuming a correct model $M_C$, while insufficiently good approximations are implemented, incorrect assumptions are adopted, or an imperfect model $M_I$ is used to compute $\hat{\Psi}^I$ (incorrect). We can estimate the systematic error $\Delta_{\text{syst}}$ induced by using $\hat{\Psi}^I$ instead of $\hat{\Psi}^C$ as

$$\Delta_{\text{syst}} \equiv \theta_{\text{bf},I} - \theta_{\text{bf},C},$$

where $\theta_{\text{bf},I}$ and $\theta_{\text{bf},C}$ are obtained using Equation (2.3) for the correct and incorrect cases introduced above.

Our approach to estimate the bias in the inferred parameters can be applied both to real measurements and before they are obtained. As said above, the latter case requires the assumption of a model $M_0$ with true parameters $\theta_{\text{tr}}^0$ as the perfect description of reality; in most forecasts, the assumed fiducial model for the analysis is considered to perfectly describe future observations, hence $\Psi^d = \hat{\Psi}^{\text{fid},C}$.

To make Equation (2.7) consistent, if $M_C$ and $M_I$ are nested models, the parameter space of the model with less parameters should be considered as a hyperplane of the parameter space of the other model, with the values of the extra parameters being kept fixed. Consider that $\theta_a$ is the extra parameter. In this case, for the model with less parameters we have $\partial \hat{\Psi} / \partial \theta_a = 0$. In order to model that this parameter is fixed in the Fisher matrix (and have the same number of parameters in the vectors that enter Equation (2.7)), we enforce $F_{*,a} = F_{a,*} \to \infty$, where the subscript * refer to all indices of the parameters vector. This is equivalent to have a perfect prior on $\theta_a$.

Equations (2.5), (2.6) and (2.7) can be applied to the analysis of any given observable, even prior to its measurement. This expression allows one to estimate the impact of modeling assumptions and approximations as well as incorrect choices of the underlying model on the inferred parameters. We envision that it will also be useful to single out possible sources of systematic errors affecting new or unexpected findings.

### 3 Observable and cosmological model considered

Although our methodology is valid for any measurement, here we focus on its application to cosmology. Given the vast amount of observed data and the complexity of some theoretical calculations, theoretical and numerical approximations are very common. Assessing the reliability of different assumptions might be challenging, but not doing it might have unacceptable consequences. Concretely, we consider the angular galaxy power spectrum as our target observable. The dramatic improvement in the quality of the observations will require a much better modeling in order to fully exploit coming galaxy surveys. Therefore, approximations that were commonly used in studies about galaxy clustering might not be accurate anymore. Some of these approximations include, but are not limited to: neglecting relativistic corrections, the Limber approximation, an incorrect estimation of the covariance matrix, a poor modeling of non-linear clustering and specific approximations used to model observational effects. For illustrative purposes, in this work we focus on the two first approximations of this list. In Paper I, where the focus is set on the misestimation of the parameter
uncertainties, we also study the consequences of neglecting the covariance between different redshifts bins. However, this is expected to have limited impact on the best-fit parameter values; hence, we do not consider this approximation in this study.

The modeling of the angular power spectrum and of the associated likelihood is discussed with detail in Paper I; here we describe it only briefly, and encourage the reader to refer to Paper I for a full description. The observed galaxy number count perturbations receive contributions from the intrinsic galaxy overdensities as well as from other effects, such as redshift-space distortions due to peculiar velocities or lensing effects caused by density perturbations along the line of sight. All these effects can be modeled in harmonic space introducing several transfer functions, the combined effect of which is given by the total transfer function $\Delta^X_\ell(k, z)$ as a function of wave number $k$ and redshift $z$, where $\ell$ is the corresponding multipole and $X$ refers to the tracer considered. The explicit form of the contributions from intrinsic clustering, peculiar velocities and relativistic corrections to $\Delta^X_\ell$ can be found in Appendix A of Paper I. $\Delta^X_\ell$ can be restricted to a given redshift bin applying a window function $W(z, z_X, \Delta z_X)$ centered at $z_X$, the width of which is controlled by $\Delta z_X$. For instance, $\Delta z_X$ often refers to the half-width or the standard deviation of a top-hat or a Gaussian window function, respectively. Accounting for the number density of galaxies per redshift $dN_X/dz$, the transfer function for a specific redshift bin can be expressed as

$$\Delta^{X,z_X}_\ell(k) = \int_0^\infty dz dN_X(z) W(z, z_X, \Delta z_X) \Delta^X_\ell(k, z),$$

where the integral of $W(z, z_X, \Delta z_X)dN_X/dz$ is equal to unity. With these conventions, the linear angular galaxy power spectrum for tracers $X$ and $Y$ and redshift bins $z_X$ and $z_Y$, respectively, is given by

$$C^{XY}_{\ell}(z_X, z_Y) = 4\pi \int \frac{dk}{k} P_0(k) \Delta^{X,z_X}_\ell(k) \Delta^{Y,z_Y}_\ell(k),$$

where $P_0(k) = k^3 P_0(k)/2\pi^2$ is the adimensional, almost scale-invariant, power-law primordial power spectrum of scalar curvature perturbations.

Galaxies are discrete tracers of the underlying density fluctuations, and therefore their power spectra are affected by shot noise. We assume a Poissonian scale-independent shot noise contribution to be added to the theoretical angular galaxy power spectrum computed in Equation (3.2), hence the total angular power spectrum can be defined as

$$\tilde{C}^{XY}_{\ell}(z_X, z_Y) = C^{XY}_{\ell}(z_X, z_Y) + \frac{\delta^K_{z_X z_Y} \delta^K_{XY}}{dN_X(z_X)/d\Omega},$$

where $dN_X(z_X)/d\Omega$ is the mean number density per steradian for tracer $X$ in the redshift bin centered at $z_X$, and $\delta^K$ is a Kronecker delta. Note that with these assumptions the shot noise term only contributes to the total power spectrum for the auto-power spectrum (i.e., same redshift bin and same tracer). Nonetheless, the shot noise might have non-Poissonian contributions (see e.g. [43, 44]). Furthermore, theoretical uncertainties can be added as noise, especially those regarding non-linear scales (see e.g., [45]).

We want to consider all possible combinations of tracers and redshift bins, and denote the number of redshift bins for tracer $X$ and $Y$ with $N_X$ and $N_Y$, respectively. As explained

---

3Observational effects (e.g., the observational mask) also modify the observed galaxy overdensity. However, their modeling is extremely case dependent, and will be presented elsewhere. Since we are interested in differential effects, neglecting this is not expected to affect our findings.
in detail in Paper I, one can consider the angular power spectra or the spherical harmonics coefficients of the galaxy number count fluctuations as the data vector. Depending on the choice, the theoretical $C_\ell$ are used in different manners. For the first option, the power spectra between different redshift bins and tracers at a given multipole are placed in a vector $C_\ell$ of size $N_X (N_X + 1)/2 + N_Y (N_Y + 1)/2 + N_X N_Y$. In turn, for the second option, they form matrices $C_\ell$ of size $(N_X + N_Y) \times (N_X + N_Y)$ which represent the covariance between the spherical harmonic coefficients. Hence, there is a $C_\ell$ vector and a $C_\ell$ matrix for each multipole $\ell$. This choice also determines how the elements of the Fisher matrix are computed [46]:

$$F_{ab} = \sum_{\ell,i,j} \left( \frac{\partial C_\ell}{\partial \theta_a} \right)_i \left( \frac{\partial C_\ell}{\partial \theta_b} \right)_j = \frac{1}{f_{\text{sky}} (2\ell + 1)} \left( C^{(i_1,j_1)}_\ell C^{(i_2,j_2)}_\ell + C^{(i_1,j_2)}_\ell C^{(i_2,j_1)}_\ell \right).$$

(3.4)

where $f_{\text{sky}}$ is the fraction of the sky probed by the survey, $M_\ell$ is a matrix representing the covariance between the elements of $C_\ell$, indicated by the indices $i$ and $j$. In the second line of Equation (3.4), $p, q, r$ and $s$ refer to the indices of $C_\ell$; in the sum over these indices, one can recognise the trace of the product of the four matrices involved. We refer the interested reader to Appendix A of Ref. [47] for further details on the derivation of this expression and the matrix properties used. The index $i$ of $C_\ell$ corresponds to $\tilde{C}_\ell^{(i_1,i_2)} \equiv \tilde{C}_\ell^{XY} (z_X, z_Y)$, where $i_1$ and $i_2$ specify each unique combination of $X$ and $z_X$, $Y$ and $z_Y$, respectively, of the transfer functions involved in Equation (3.2). Then, each element $i,j$ of the covariance matrix $M_\ell$ is given by

$$\left( M_\ell \right)_{ij} = \frac{1}{f_{\text{sky}} (2\ell + 1)} \left( \tilde{C}^{(i_1,j_1)}_\ell \tilde{C}^{(i_2,j_2)}_\ell + \tilde{C}^{(i_1,j_2)}_\ell \tilde{C}^{(i_2,j_1)}_\ell \right).$$

(3.5)

Now that we have specified the target observable, its covariance, and its Fisher matrix, we can explicitly apply the formalism derived in Section 2. In this case, we can identify the data and theory vector $\Psi$ as $C_\ell$, and its covariance is given by $M_\ell$. Specifying the general expression in Equation (2.5) to the case of the angular galaxy power spectra we obtain that the difference between the best-fit parameters and the initially assumed fiducial parameters within a general model is

$$\Delta \theta = \theta^{\text{bf}} - \theta^{\text{fid}} = F^{-1} \sum_{\ell,i,j} \left( \nabla_{\theta} C^{\text{fid}}_\ell \right)_i \left( M_\ell^{-1} \right)_{ij} \left( C^{\text{fid}}_\ell - C^{\text{fid}}_\ell \right)_j = \frac{1}{f_{\text{sky}} (2\ell + 1)} \sum_{\ell, p, q, r, s} \left( \nabla_{\theta} C^{\text{fid}}_\ell \right)_{pq} \left( \left( C^{\text{fid}}_\ell \right)^{-1} \right)_{qr} \left( C^{\text{fid}}_\ell - C^{\text{fid}}_\ell \right)_{rs} \left[ \left( C^{\text{fid}}_\ell \right)^{-1} \right]_{sp}.$$

(3.6)

Equation (3.6) can be applied to the correct description of the angular galaxy power spectra and to an approximated one. The substitution of these results in Equation (2.7) yields the systematic bias introduced in parameter inference due to the approximated description, $\Delta_{\text{syst}}$.

As pointed out in Paper I, Equation (3.4) is only valid in the case that the same multipole range is used for all redshift bins. This is because both $M_\ell$ and $C_\ell$ would be singular otherwise: the matrices corresponding to the multipoles that are not used in all redshift bins would contain complete rows and columns filled with zeros. Therefore, Equation (3.4)
cannot be applied when the maximum multipole used depends on redshift. This would be the case, for instance, of modeling the redshift dependence of the scales for which non-linear clustering becomes significant. However, one can consider different likelihoods using a different multipole range for each of them in order to overcome this limitation. Each of these likelihoods includes only the power spectra between the redshift bins that cover the corresponding multipole range. We refer the interested reader to Paper I for more details. Taking this into account, it is straightforward to generalize Equation (3.6) to this case comparing Equations (2.5) and Equation (2.6).

3.1 Cosmological model under study and straw-man survey examples

Given the promising prospects of future galaxy surveys to constrain primordial non-Gaussianities (see e.g., [11, 13, 48–52]), we choose ΛCDM+ fNL to be the cosmological model under study, where fNL parametrizes the amplitude of primordial non-Gaussianity of the local type controlling the amplitude of the quadratic contributions of a Gaussian random field to the Bardeen potential. The effect of this type of non-Gaussianity on the clustering of halos can be modeled with a strong scale-dependence of the galaxy bias on large scales [53–56]. Denoting the standard, scale-independent Gaussian galaxy bias as b_G and using the large-scale structure convention for fNL, the total galaxy bias is given by

\[ b_{\text{tot}}(k, z) = b_G + (b_G - 1)f_{\text{NL}}\delta_{\text{crit}} \frac{3\Omega_mH_0^2}{c^2k^2T(k, z)}, \]  

where \( \delta_{\text{crit}} = 1.68 \) is the critical density related with spherical gravitational collapse in an Einstein-de Sitter cosmology, \( \Omega_m \) is the matter density parameter today, \( H_0 \) is the Hubble constant, \( c \) is the speed of light, and \( T(k, z) \) is the transfer function of matter.\(^4\)

Including the Gaussian galaxy bias of each tracer as a model parameter, the set of parameters considered in this work to model ΛCDM+ fNL is \( \theta = \{b, \omega_b, \omega_c, n_s, b_X, b_Y, f_{\text{NL}}\} \), \(^5\) where \( \omega_b \) and \( \omega_c \) are the physical densities of baryons and cold dark matter, respectively, \( n_s \) is the spectral index of \( P_0(k) \), and \( b_X \) (\( b_Y \)) is the Gaussian galaxy bias for the tracer \( X \) (\( Y \)). Note that for analyses with only one tracer, \( b_X \) and \( b_Y \) become simply \( b_G \).

As done in Paper I, we consider three straw-man galaxy surveys to study if the systematic biases depend qualitatively on the survey parameters, as well as to study the multi-tracer case. First, we consider a survey with galaxies uniformly distributed in redshift and a galaxy density per unit redshift and square degree \( d^2N_g/dzd\Omega = 1070 \text{ gal/deg}^2 \). We also consider two other more realistic surveys, inspired by the galaxy redshift distribution expected for Euclid [11] (in a pessimistic scenario) and SPHEREx [14], which we approximate by

\[ \frac{d^2N}{dzd\Omega} = A \left( \frac{z}{z_0} \right)^\alpha e^{-(z/z_0)^{1.5}} \text{ gal/deg}^2, \]  

with \( A_{\text{Eu}} = 2.4 \times 10^5 \), \( z_{0\text{Eu}} = 0.54 \), and \( \alpha_{\text{Eu}} = 4.0 \) for Euclid; and \( A_{\text{SP}} = 2.93 \times 10^4 \), \( z_{0\text{SP}} = 0.53 \) and \( \alpha_{\text{SP}} = 1.1 \). We denote these two galaxy distributions as Euclid-like and

\(^4\)Detailed comparison with N-body simulations indicate that there might be a correction factor of order unity to Equation (3.7) (see e.g., [57, 58]), which for simplicity we omit here. Hence, \( f_{\text{NL}} \), in Equation (3.7) should be considered as an effective primordial non-Gaussianity parameter of about the same magnitude of the true underlying \( f_{\text{NL}} \).

\(^5\)We do not consider the amplitude of \( P_0 \) as a free parameter in our analysis because it is almost completely degenerate with \( b_X^2 \) and \( b_Y^2 \), although we are aware that it should be included in an analysis of real observations.
under study matches the true cosmology, so that \( M \) only from incorrect modeling. Therefore, in what follows we consider that the only model apply the methodology to a case where the bias introduced in the best-fit parameters arises but also to act as a warning for future measurements. For illustration purposes, we choose to The results shown in this section aim to be an example of the performance of our methodology, respectively. In addition, we consider a conservative case in which \( \ell \) take \( \theta \) take \( \theta \) considered significantly biased if the corresponding component of \( C \) use the Limber approximation. In turn, \( \theta \) does not change the qualitative results of the examples under study.

## 4 Systematic bias induced by different approximations

The results shown in this section aim to be an example of the performance of our methodology, but also to act as a warning for future measurements. For illustration purposes, we choose to apply the methodology to a case where the bias introduced in the best-fit parameters arises only from incorrect modeling. Therefore, in what follows we consider that the only model under study matches the true cosmology, so that \( M^C = M_0 \), while \( M^I \) refers to the same underlying model but when an incorrect modeling of the observable is used. Furthermore, we take \( \theta^{\text{fid}, I} = \theta^{\text{fid}, C} = \theta^{\text{fid}}_0 \). This means that the data drawn from \( M_0 \) is equal to the prediction for \( M^C \) using the correct modeling: \( C^{\text{fid}}_\ell \equiv C^{\text{fid}}_\ell (\theta^{\text{fid}, C} | M^C) \). We choose \( M_0 \) to be \( \Lambda \) CDM + \( f_{\text{NL}} \), with parameter values \( \theta^{\text{fid}}_0 \): \( h = 0.6727, \omega_b = 0.02225, \omega_{\text{cdm}} = 0.1198, \sigma_8 = 0.9645 \), \( f_{\text{NL}} = 0 \), an amplitude of the primordial power spectrum \( \ln 10^{10} A_s = 3.0940 \), and we consider three massive degenerate neutrinos with mass \( m_\nu = 0.02 \) eV each. We assume scale- and redshift-independent Gaussian galaxy bias for the sake of simplicity: \( b^\text{unif}_G = b^{\text{unif}}_G = 2 \) for the uniform and Euclid-like surveys and \( b^{\text{unif}}_G = 2 \) for the SPHEREx-like survey; this assumption does not change the qualitative results of the examples under study.

The modeling we use to compute \( C^C_\ell \) and \( C^d_\ell \) includes relativistic corrections and redshift-space distortions due to peculiar velocities and lensing magnification, and does not use the Limber approximation. In turn, \( C^i_\ell \) differs from them in one aspect of the modeling: in Section 4.1 \( C^I_\ell \) does not include the contribution from lensing magnification, while in Section 4.2 \( C^I_\ell \) uses the Limber approximation. Both cases are explored using a multi-tracer analysis of the angular galaxy power spectra in Section 4.3.

In all cases considered below, the estimated systematic biases correspond to the full vector of the shift in the multidimensional parameter space. Afterwards, we assess the significance of the estimated biases in the marginalized constraints on each of the parameters by comparing them with the 68% confidence level of their respective marginalized uncertainties, obtained using the same approximations. As a general rule of thumb, the parameter \( \theta_a \) would be considered significantly biased if the corresponding component of \( \Delta_{\text{syst}}, \Delta_{\text{syst,a}} \), is larger
than the 68% confidence level marginalized uncertainty in the inference of $\theta_a$: $\Delta_{\text{syst},a}/\sigma_{\theta_a} \gtrsim 1$. We refer the interested reader to Paper I for a detailed discussion on the effects of approximations on the parameter uncertainties.

4.1 Ignoring lensing magnification

Matter density perturbations along the line of sight affect how we observe the galaxy density distribution [59]. Therefore, the observed galaxy number count perturbations are determined by the intrinsic clustering, redshift-space distortions due to peculiar velocities and relativistic corrections. These corrections can be separated into contributions from lensing magnification, doppler, and gravitational potential effects such as time-delays and integrated Sachs-Wolfe effect [60–65].

Lensing magnification is a subdominant contribution to the observed galaxy overdensities except when the redshift separation between bins is large enough so that the correlation due to intrinsic clustering is negligible. However, it is normally the largest relativistic correction, especially when cross-correlating two different redshift bins using the angular power spectrum [66]. Some arguments against including lensing magnification include its subdominant relative contribution to the galaxy power spectrum at the scales explored so far, the large computational expenses required for its calculation and the difficulties in obtaining an accurate determination of the magnification bias parameter (see e.g., [67]). Nonetheless, the magnification contribution contains cosmological information complementary to weak-lensing shear [25, 68, 69].

To understand and model the effect of lensing magnification consider that the gravitational lensing contribution to the galaxy overdensity consists of two competing effects: on the one hand, it stretches the volume behind the lens; on the other, it magnifies individual sources and promotes faint galaxies above the magnitude limit of the survey [70]. This changes the observed galaxy number density $n_{\text{obs}}$ in a flux-limited survey:

$$n_{\text{obs}} = n [1 + (5s - 2)\kappa], \quad (4.1)$$

where $n$ is the intrinsic galaxy number density, $s$ is the magnification bias parameter, and $\kappa$ is the convergence [71]. Note that $s = 0.4$ corresponds to a vanishing contribution from lensing magnification. Since the magnification bias parameter depends on the tracer used, we distinguish between $s^X$ and $s^Y$. We do not consider $s$ as a nuisance parameter (as should be done in a more quantitative analysis of actual observations); instead, we study several cases with different constant values of $s$ in order to study the dependence of the systematic biases on this parameter.

We show the significance of the estimated biases when lensing magnification is neglected as a function of the magnification bias parameter in Figure 1. Considering only single-tracer analyses, we show results for the uniform and Euclid-like galaxy surveys, and normalize the estimated biases with the forecasted 68% confidence level marginalized constraints in order to show the bias significance. We find that the size of the systematic bias grows as $|s - 0.4|$ increases, and that it grows faster for magnified populations (i.e., $s > 0.4$), than for demagnified. Results for different redshift distributions are qualitatively very similar. The only difference is that, for $\ell_{\text{max}} = 200$, the estimated biases are larger for the uniform survey than for the Euclid-like survey when the redshift bins overlap, and vice versa when the redshift bins do not overlap. While for the Euclid-like survey the biases are always larger when the redshift bin do not overlap, this is only true for the uniform survey when $s > 0.4$. 


Figure 1: Ratio of the estimated bias in the model parameters over the forecasted 68% confidence level marginalized constraints for cases when the lensing magnification is included (blue) and not as a function of the assumed magnification bias parameter. We show results for uniform (green) and Euclid-like (red) galaxy redshift distributions, and with non-overlapping (top) and overlapping (bottom) redshift bins; note the change of scale in the $y$-axis between them. In all cases, the case for $\ell_{\text{max}}(z)$ is shown with wide solid lines with circle markers, while thin solid lines without markers correspond to the case with constant $\ell_{\text{max}} = 200$. Dotted lines mark $|\Delta_{\text{syst},\alpha}/\sigma_\alpha| = 1$.

In general, the significance of the bias is much larger for the case in which $\ell_{\text{max}} = 200$ than when $\ell_{\text{max}}$ varies with redshift (reaching higher values). This may be counter-intuitive, since the signal-to-noise ratio of the magnification contribution increases at smaller scales.
However, the higher-$\ell$ part is computed only for the highest-z bins (see Section 3.1), so that the relative contribution of lensing magnification is smaller at these scales (remember that lensing magnification dominates the angular power spectra for very separated redshift bins). Moreover, the constraints on the cosmological parameters also improve for increasing $\ell_{\text{max}}$, and $\Delta \theta \propto F^{-1}$. The reduction of the significance of the bias with higher $\ell_{\text{max}}$ can be understood as the information beyond lensing magnification encoded in the angular galaxy power spectra having more weight in the final parameters constraints, with respect to the $\ell_{\text{max}} = 200$ case.

Assessing the significance of the biases exploring only one-dimensional marginalized parameter constraints might be misleading. The bias might be much more significant for systematic shifts in perpendicular directions to parameter degeneracies than what would be inferred from one-dimensional projections. We show forecasted two-dimensional marginalized constraints for all parameter combinations from a single-tracer Euclid-like survey in the $\ell_{\text{max}}(z)$ case including the estimated biases with respect to the fiducial cosmology in Figure 2. We compare the case including magnification (with $s = 0.8$) and the case without including magnification (shown in blue and red, respectively). We show results for the non-overlap (overlap) redshift bin configuration in dark (light) colors. As found in Paper I, the parameter degeneracies obtained neglecting lensing magnification are not necessarily the same as for the correct analysis. Furthermore, we find the systematic bias to be aligned with the degeneracies of the incorrectly estimated constraints. Whether the alignment is general or depends on the observable is beyond the scope of this work. As expected from Figure 1, the biases shown in Figure 2 are larger for the case with non-overlapping redshift bins. Finally, while the correct uncertainties do not depend on whether the redshift bins overlap or not, the resulting incorrect confidence level regions (i.e., neglecting magnification) are slightly larger when the redshift bins overlap (which further reduces the significance of the bias). The results using $\ell_{\text{max}} = 200$ are qualitatively similar (with a higher significance of the bias, as shown in Figure 1).

### 4.2 Using the Limber approximation

As shown in Equation (3.2), the angular power spectrum is the projection along the line-of-sight of the spatial, three-dimensional power spectrum. The transfer functions that drive this projection include spherical Bessel functions. Therefore, integration over two spherical Bessel functions is needed to compute $C_{\ell}$ for each multipole. Given the oscillatory nature of the spherical Bessel functions, it is very computationally expensive to ensure the convergence of these integrals, which slows down the calculation of $C_{\ell}$. The Limber approximation [72–74] aims to alleviate this problem by approximating the spherical Bessel functions as Dirac delta functions. However, the transformation of low-order spherical Bessel functions (i.e., low $\ell$) into Dirac delta functions is not accurate, which means that the Limber approximation breaks down for $C_{\ell}$ at large scales.

We show two-dimensional forecasted marginalized constraints on the cosmological parameters from the angular power spectrum of a single-tracer Euclid-like galaxy survey in Figure 3. We show results using $\ell_{\text{max}}(z)$ both with and without the Limber approximation (red and blue, respectively), and show the results with overlapping redshift bins in lighter colors. We assume $s = 0.4$ in order to avoid contributions from lensing magnification. As noted in Paper I, parameter uncertainties are underestimated when the Limber approximation is used, especially for non-overlapping redshift bins. This makes the (generally) small shifts in the best-fit values more significant when compared to the small forecasted errors. The bias
is especially worrisome for $f_{\text{NL}}$, the only example where the shift in the best fit is very large. This is because the signature of local primordial non-Gaussianities manifests at large scales, which is where the Limber approximation breaks down. The estimated systematic bias in $f_{\text{NL}}$ is $\gtrsim 18(14)\sigma$ for the case with non-overlapping (overlapping) redshift bins. This result indicates that using the Limber approximation at large scales potentially leads to a false positive detection of primordial non-Gaussianity. In general, biases are smaller when the
Figure 3: Same as Figure 2, but comparing the results with and without using the Limber approximation and assuming $s = 0.4$.

redshift bins overlap, as it was the case for the non-inclusion of lensing magnification of the signal. The results using $\ell_{\text{max}} = 200$ are qualitatively similar, with the weaker constraints (due to the use of a narrower multipole range) but with a comparable significance of the bias due to the use of the Limber approximation. This is because the Limber approximation breaks down at low $\ell$, which is a regime covered in both cases.

4.3 Multi-tracer galaxy power spectrum

One of the most exciting prospects offered by the next-generation galaxy surveys is the possibility to perform multi-tracer analyses [75, 76]. Instead of using all galaxies as a single tracer, using various tracers of the underlying density field at the same time and
fully accounting for their cross-correlations, provides additional information. The gain comes from probing the same volume more than once, each time with a different galaxy bias, which reduces the cosmic variance for quantities that are related to the galaxy bias. Besides tightening the constraints on all cosmological parameters, especially those related with the galaxy bias, the use of multi-tracer approaches are expected to be especially effective to constrain physics that affect the large scales, such as local primordial non-Gaussianities.

Given the reduced cosmic variance, systematic errors may produce more significant biases in parameter inference than for single-tracer studies. We apply the methodology described in Section 2 to multi-tracer analyses of the angular galaxy clustering in order to assess the impact of the approximations discussed above. We use Multi_CLASS to compute the angular cross-power spectrum of two different tracers (with their own redshift distribution, galaxy bias and magnification bias parameter). We consider two different galaxy surveys (or galaxy populations observed by the same survey) following a Euclid-like and a SPHEREx-like redshift distribution and with different galaxy bias as specified in Section 3.1.

We show forecasted two-dimensional marginalized constraints of the cosmological parameters under study with and without modeling lensing magnification and using $\ell_{\text{max}}(z)$ in the signal and covariance in Figure 4. For illustrative purposes, we focus just on the case with magnification bias parameters $s_{\text{Eu}}^{-1} = s_{\text{SP}}^{-1} = 0.6$ for the Euclid-like and SPHEREx-like surveys. The figure also includes the estimated bias in the best-fit parameters. We can appreciate that, as reported in Paper I, ignoring lensing magnification overestimates the uncertainties in the parameters, except for $f_{\text{NL}}$ (and $\omega_b$, for which there is practically no effect). Moreover, there are still significant biases ($\sim 1 - 2\sigma$) in all parameters when the redshift bins do not overlap (shown in darker colors). Surprisingly, the systematic bias is much less significant ($\lesssim 1\sigma$) when the redshift bins do overlap. As in Figure 2, the estimated bias is aligned with the degeneracy between the parameters, which almost does not change whether lensing magnification is included or not. The alignment between the bias and the parameter degeneracies is most likely not generic, but very population selection (and thus, survey) dependent. With a slightly different set up (or different fiducial $s$ for the two populations) this alignment may not hold. Such scenario may lead to great impact in final results: a change in the degeneracies would greatly exacerbate the effect of the systematic bias introduced by approximations when combining the angular galaxy power spectra with other cosmological probes.

Similarly to the single-tracer case, the estimated bias found when using $\ell_{\text{max}} = 200$ is larger than using $\ell_{\text{max}}(z)$, but the degeneracies and direction of the shifts remain unchanged. The increase of the significance of the bias using $\ell_{\text{max}} = 200$ instead of $\ell_{\text{max}}(z)$ is smaller than for the single-tracer case. This is because the multi-tracer approach reduces the cosmic variance, so that the relative contribution to the constraints from large scales in the $\ell_{\text{max}}(z)$ case is higher than in the single-tracer case and the effect of using $\ell_{\text{max}} = 200$ is smaller.

Figure 5 shows the analogous results for using or not the Limber approximation (and considering lensing magnification in both cases). The uncertainties of the parameters are underestimated at approximately the same level as for the single-tracer case, but the estimated biases are larger in this case. Note also that in some cases the biases is not aligned with the parameter degeneracies. As for the Euclid-like only analysis shown in the previous section, the biases are smaller for overlapping bins (lighter colors) than for non-overlapping bins, because when the redshift bins overlap biases are smaller and incorrect uncertainties are larger. In this case $f_{\text{NL}}$ is again by far the most affected parameter, with biases of $19\sigma$ ($14\sigma$) for non-overlapping (overlapping) redshift bins. In this case, as for the single-tracer case, the
Multi tracer: Euclid-like ($s^{Eu} = 0.6$) + SPHEREx-like ($s^{SP} = 0.6$) Surveys

$\ell_{\text{max}}(z) = \{180, 500, 1100, 1900, 3000\}$

Figure 4: Same as Figure 2, but for a multi-tracer analysis of the Euclid-like and SPHEREx-like surveys considered in this work, assuming $s = 0.6$ for both surveys.

results using $\ell_{\text{max}} = 200$ are very similar to those shown in Figure 5.

Finally, we can compare the estimated marginalized biases for the Euclid-like only analyses and for the multi-tracer analysis combining the Euclid-like and the SPHEREx-like surveys. We show the ratio of the estimated biases over the forecasted 68% confidence level marginalized constraints for both cases (using both $\ell_{\text{max}}(z)$ and $\ell_{\text{max}} = 200$) in Figure 6. We show results both for overlapping and non-overlapping bins, and using separately the two approximations considered in this work: neglecting lensing magnification (left panel), and using the Limber approximation (right panel). Figure 6 shows in clearer way the comparison between the single-tracer and multi-tracer cases and using different criteria for $\ell_{\text{max}}$ discussed above. When lensing magnification is not included the biases are approximately the same for both
Figure 5: Same as Figure 3, but for a multi-tracer analysis of the Euclid-like and SPHEREx-like surveys considered in this work, assuming $s = 0.6$ for both surveys.

Single- and multi-tracer analyses if $\ell_{\text{max}}(z)$, while they are larger for a single-tracer analysis if $\ell_{\text{max}} = 200$. In all cases, the estimated biases are larger for non-overlapping redshift bins. Regarding the use of the Limber approximation, the biases are larger for all parameters for the multi-tracer case. The dependence of the estimated biases on the criterion used for $\ell_{\text{max}}$ is not significant, except for $\omega_b$ and $f_{\text{NL}}$. Finally, the estimated biases are again always larger for the non-overlapping redshift bins, with a special mention to $f_{\text{NL}}$, the difference of which is more than $4\sigma$.

Although exploiting higher multipoles of the angular galaxy power spectra returns smaller biases for the specific sources of theoretical systematic errors explored in this work, we emphasize that there are other key features in the modeling of the observable suscep-
Figure 6: Ratio of the estimated bias in the cosmological parameters over the forecasted 68% confidence level marginalized constraints for cases when the lensing magnification is not modeled (left panel) and the Limber approximation is used (right panel). We show results considering a Euclid-like survey (diamonds), and its combination with an SPHEREx-like survey performing a multi-tracer analysis (circles), both for overlapping (orange) and non-overlapping (red) redshift bins. In all cases we consider $s = 0.6$ for both galaxy populations. Filled (hollow) markers refer to results using $\ell_{\text{max}}(z)$ ($\ell_{\text{max}} = 200$), respectively. Note the different scale in the $y$-axis in each panel, and that the $y$-axis in the right panel is broken. Dotted lines mark $|\Delta_{\text{syst},a}/\sigma_a| = 1$.

tible to induce a bias in the best-fit parameters. The modeling of non-linear clustering is probably the most important one, and it arises at small scales. Therefore, we advocate for a comprehensive estimation of potential biases in the inferred best-fit parameters using the methodology described in this work, accounting for all possible sources of systematics or approximations adopted, before freezing the analysis pipeline.
Cosmology needs to transition from the precision to the accuracy era. Reducing the systematic error budget below the statistical uncertainties represents a crucial step in that direction. Besides controlling observational systematics, improved theoretical models of cosmological observables will be needed. Approximations that have been accurate enough so far, may introduce significant systematic errors for forthcoming experiments.

There are two kinds of errors that can be introduced into an analysis: a modification of the shape of the posterior distribution and a shift of the location of its peak. These produce a misestimation of the model parameters covariance and a systematic bias on their best-fit values. While the former has been studied on a companion paper [39], here we have focused on the latter.

Expanding upon previous works, we have presented a completely general methodology to estimate the systematic bias introduced in parameter inference when the theoretical model for a given measurement is not accurate enough or the assumed underlying model is not incorrect. The derived methodology is equally applicable to any measurement, even beyond the field of cosmology. Equations (2.5) and (2.7) are the main result of this paper and, given its complete flexibility and easy to use, we advocate its implementation whenever different approximations are under consideration.

Our methodology can also be useful to optimize analyses that rely on an assumed fiducial cosmology. Since this may introduce a systematic bias (see e.g., [77] for a detailed study in the case of the BAO analysis), our methodology can be iteratively applied to find a fiducial cosmology with a prediction more concordant with the measurements. Our methodology assumes Gaussian posteriors since it is based on the Fisher matrix formalism, but this assumption can be relaxed, following [78–81].

To show its performance, we have applied our methodology to the angular galaxy power spectra as observed by next generation galaxy surveys. Instead of considering specific examples, we have used straw-man, yet realistic, galaxy-survey specifications and have shown how neglecting lensing magnification or using the Limber approximation can bias cosmological parameter inference. We have found significant biases (most of them $\gtrsim 1\sigma$) in all the cases explored in this work. Moreover, we have also included examples of multi-tracer analyses, using Multi_CLASS, a modified version of CLASS which allows to compute the angular cross-power spectra for two different tracers of the matter distribution. In general, we have found that the estimated systematic biases due to the considered approximations are more significant for multi-tracer analyses.

We stress that our results cannot be taken quantitatively at face value because the significance of the biases are expected to be very dependent on the survey and galaxy population specifications. Nonetheless, the risk of introducing significant systematic biases in the results is general, and our work must be considered as a warning for analyses of future experiments. Although, for illustration, we have focused in just two sources of theoretical systematic errors, there are many other potential origins of systematic errors, both observational and theoretical. Therefore, it is of crucial importance that any source of systematics is reduced well below the $1\sigma$ statistical uncertainty to ensure the robustness of the results. This is because the joint contribution of small systematic errors can significantly bias the inferred constraints otherwise, even if the sources of systematics are unrelated between them.

Considering the dramatic experimental upgrades that many fields of physics will experience in the coming years, it is of paramount importance to fully exploit the potential of
new experiments and to obtain unbiased results. To do so, we need to estimate potential biases introduced by approximations under consideration. In this work, we have provided a methodology to do that and estimate the significance of systematic biases introduced in parameter inference. Finally, we envision that our methodology will also be useful to single out possible sources of systematic errors affecting new or unexpected findings.

Acknowledgments

We would like to thank Alan Heavens for comments on last stages of this manuscript. JLB is supported by the Allan C. and Dorothy H. Davis Fellowship, and has been supported by the Spanish MINECO under grant BES-2015-071307, co-funded by the ESF, during part of the development of this work. Funding for this work was partially provided by the Spanish MINECO under projects AYA2014-58747-P AEI/FEDER, UE, and MDM-2014-0369 of IC-CUB (Unidad de Excelencia María de Maeztu). NB is supported by the Spanish MINECO under grant BES-2015-073372. AR has received funding from the People Programme (Marie Curie Actions) of the European Union H2020 Programme under REA grant agreement number 706896 (COSMOFLAGS). LV acknowledges support by European Union’s Horizon 2020 research and innovation programme ERC (BePreSySe, grant agreement 725327).

References

[1] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, et al., “Planck 2018 results. VI. Cosmological parameters”, ArXiv e-prints (July, 2018), arXiv:1807.06209.

[2] The SDSS-III BOSS, S. Alam et al., “The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample”, MNRAS 470 (Sept., 2017) 2617–2652, arXiv:1607.03155.

[3] D. M. Scolnic, D. O. Jones, A. Rest, Y. C. Pan, et al., “The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample”, Astrophys. J. 859 no. 2, (Jun, 2018) 101, arXiv:1710.00845 [astro-ph.CO].

[4] T. M. C. Abbott, F. B. Abdalla, A. Alarcon, J. Aleksić, S. Allam, S. Allen, A. Amara, J. Annis, J. Asorey, S. Avila, D. Bacon, E. Balbinot, and Dark Energy Survey Collaboration, “Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing”, PRD 98 no. 4, (Aug, 2018) 043526, arXiv:1708.01530 [astro-ph.CO].

[5] L. Verde, T. Treu, and A. G. Riess, “Tensions between the early and late Universe”, Nature Astronomy 3 (Sep, 2019) 891–895, arXiv:1907.10625 [astro-ph.CO].

[6] A. G. Riess, S. Casertano, W. Yuan, L. M. Macri, and D. Scolnic, “Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM”, Astrophys. J 876 no. 1, (May, 2019) 85, arXiv:1903.07603 [astro-ph.CO].

[7] K. C. Wong, S. H. Suyu, G. C. F. Chen, C. E. Rusu, M. Millon, D. Sluse, V. Bonvin, C. D. Fassnacht, S. Taubenberger, M. W. Auger, S. Birrer, J. H. H. Chan, F. Courbin, S. Hilbert, O. Tihhonova, T. Treu, A. Agnello, X. Ding, I. Jee, E. Komatsu, A. J. Shajib, A. Sonnenfeld, R. D. Bland ford, L. V. E. Koopmans, P. J. Marshall, and G. Meylan, “H0LiCOW XIII. A 2.4% measurement of $H_0$ from lensed quasars: 5.3σ tension between early and late-Universe probes”, arXiv e-prints (Jul, 2019) arXiv:1907.04869, arXiv:1907.04869 [astro-ph.CO].
H. Hildebrandt, F. Köhlinger, J. L. van den Busch, B. Joachimi, C. Heymans, A. Kannawadi, A. H. Wright, M. Asgari, C. Blake, H. Hoekstra, and S. Joudaki, “KiDS+VIKING-450: Cosmic shear tomography with optical+infrared data”, arXiv e-prints (Dec, 2018) arXiv:1812.06076.

R. P. Norris et al., “EMU: Evolutionary Map of the Universe”, PASA 28 (Aug., 2011) 215–248.

DESI Collaboration, A. Aghamousa, and et al., “The DESI Experiment Part I: Science,Targeting, and Survey Design”, ArXiv e-prints (Oct., 2016) arXiv:1611.00036.

L. Amendola, S. Appleby, A. Avgoustidis, D. Bacon, T. Baker, et al., “Cosmology and fundamental physics with the Euclid satellite”, Living Reviews in Relativity 21 no. 1, (Apr, 2018) arXiv:1606.00180.

The LSST Science Collaborations, P. A. Abell et al., “LSST Science Book, Version 2.0”, arXiv:0912.0201.

Square Kilometre Array Cosmology Science Working Group, D. J. Bacon, et al., “Cosmology with Phase 1 of the Square Kilometre Array Red Book 2018: Technical specifications and performance forecasts”, PASA 37 (Jan., 2020) e007, arXiv:1811.02743.

O. Doré, J. Bock, M. Ashby, P. Capak, A. Cooray, et al., “Cosmology with the SPHEREX All-Sky Spectral Survey”, arXiv e-prints (Dec, 2014) arXiv:1412.4872.

D. J. Watts, B. Wang, A. Ali, J. W. Appel, C. L. Bennett, D. T. Chuss, S. Dahal, J. R. Eimer, T. Essinger-Hileman, K. Harrington, G. Hinshaw, J. Iuliano, T. A. Marriage, N. J. Miller, I. L. Padilla, L. Parker, M. Petroff, K. Rostem, M. B. Silva, E. J. Wollack, and Z. Xu, “A Projected Estimate of the Reionization Optical Depth Using the CLASS Experiment’s Sample Variance Limited E-mode Measurement”, Astrophys. J. 863 (Aug., 2018) 121.

P. Ade, and The Simons Observatory collaboration, “The Simons Observatory: science goals and forecasts”, JCAP 2019 no. 2, (Feb, 2019) 056, arXiv:1808.07445.

K. N. Abazajian et al., “CMB-S4 Science Book, First Edition”, arXiv e-prints (Oct, 2016) arXiv:1610.02743.

E. D. Kovetz, M. P. Viero, A. Lidz, L. Newburgh, M. Rahman, E. Switzer, M. Kamionkowski, et al., “Line-Intensity Mapping: 2017 Status Report”, arXiv e-prints (Sep, 2017) arXiv:1709.09066.

E. D. Kovetz, P. C. Breyssse, A. Lidz, J. Bock, C. M. Bradford, T.-C. Chang, S. Foreman, H. Padmanabhan, A. Pullen, D. Riechers, M. B. Silva, and E. Switzer, “Astrophysics and Cosmology with Line-Intensity Mapping”, arXiv e-prints (Mar., 2019) arXiv:1903.04496.

J. L. Bernal, P. C. Breyssse, H. Gil-Marín, and E. D. Kovetz, “A User’s Guide to Extracting Cosmological Information from Line-Intensity Maps”, arXiv e-prints (Jul, 2019) arXiv:1907.10067.

J. L. Bernal, P. C. Breyssse, and E. D. Kovetz, “The Cosmic Expansion History from Line-Intensity Mapping”, arXiv e-prints (Jul, 2019) arXiv:1907.10065.

J. B. Muñoz, “A Standard Ruler at Cosmic Dawn”, arXiv e-prints (Apr, 2019) arXiv:1904.07868.

A. F. Heavens, T. D. Kitching, and L. Verde, “On model selection forecasting, dark energy and modified gravity”, MNRAS 380 no. 3, (Sep, 2007) 1029–1035.
24] A. N. Taylor, T. D. Kitching, D. J. Bacon, and A. F. Heavens, “Probing dark energy with the shear-ratio geometric test”, MNRAS 374 no. 4, (Feb., 2007) 1377–1403, arXiv:astro-ph/0606416 [astro-ph].

25] C. A. J. Duncan, B. Joachimi, A. F. Heavens, C. Heymans, and H. Hildebrandt, “On the complementarity of galaxy clustering with cosmic shear and flux magnification”, MNRAS 437 no. 3, (Jan., 2014) 2471–2487, arXiv:1306.6870 [astro-ph].

26] L. Knox, R. Scoccimarro, and S. Dodelson, “Impact of Inhomogeneous Reionization on Cosmic Microwave Background Anisotropy”, Phys. Rev. Letters 81 no. 10, (Sept., 1998) 2004–2007, arXiv:astro-ph/9805012 [astro-ph].

27] A. G. Kim, E. V. Linder, R. Miquel, and N. Mostek, “Effects of systematic uncertainties on the supernova determination of cosmological parameters”, MNRAS 347 no. 3, (Jan., 2004) 909–920, arXiv:astro-ph/0304509 [astro-ph].

28] A. Taruya, T. Nishimichi, and S. Saito, “Baryon acoustic oscillations in 2D: Modeling redshift-space power spectrum from perturbation theory”, Phys. Rev. D 82 no. 6, (Sep, 2010) 063522, arXiv:1006.0699 [astro-ph.CO].

29] A. Natarajan, A. R. Zentner, N. Battaglia, and H. Trac, “Systematic errors in the measurement of neutrino masses due to baryonic feedback processes: Prospects for stage IV lensing surveys”, Phys. Rev. D 90 no. 6, (Sep, 2014) 063516, arXiv:1405.6205 [astro-ph.CO].

30] A. R. Pullen, C. M. Hirata, O. Doré, and A. Raccanelli, “Interloper bias in future large-scale structure surveys”, Pub. of the Astron. Soc. of Japan 68 no. 1, (Feb, 2016) 12, arXiv:1507.05092 [astro-ph.CO].

31] S. Camera, R. Maartens, and M. G. Santos, “Einstein’s legacy in galaxy surveys.”, MNRAS 451 (July, 2015) L80–L84, arXiv:1412.4781 [astro-ph.CO].

32] W. Cardona, R. Durrer, M. Kunz, and F. Montanari, “Lensing convergence and the neutrino mass scale in galaxy redshift surveys”, Phys. Rev. D 94 no. 4, (Aug., 2016) 043007, arXiv:1603.06481 [astro-ph.CO].

33] G. Jelic-Cizmek, F. Lepori, C. Bonvin, and R. Durrer, “On the importance of lensing for galaxy clustering in photometric and spectroscopic surveys”, arXiv e-prints (Apr., 2020) arXiv:2004.12981, arXiv:2004.12981 [astro-ph.CO].

34] A. Raccanelli, L. Verde, and F. Villaescusa-Navarro, “Biases from neutrino bias: to worry or not to worry?”, MNRAS 483 no. 1, (Feb, 2019) 734–743, arXiv:1704.07837 [astro-ph.CO].

35] A. Taylor, B. Joachimi, and T. Kitching, “Putting the precision in precision cosmology: How accurate should your data covariance matrix be?”, MNRAS 432 no. 3, (July, 2013) 1928–1946, arXiv:1212.4359 [astro-ph.CO].

36] D. Kodwani, D. Alonso, and P. Ferreira, “The effect on cosmological parameter estimation of a parameter dependent covariance matrix”, The Open Journal of Astrophysics 2 no. 1, (Mar., 2019) 3, arXiv:1811.11584 [astro-ph.CO].

37] E. Sellentin and A. F. Heavens, “On the insufficiency of arbitrarily precise covariance matrices: non-Gaussian weak-lensing likelihoods”, MNRAS 473 no. 2, (Jan., 2018) 2355–2363, arXiv:1707.04488 [astro-ph.CO].

38] M. S. Wang, W. J. Percival, S. Avila, R. Crittenden, and D. Bianchi, “Cosmological inference from galaxy-clustering power spectrum: Gaussianization and covariance decomposition”, MNRAS 486 no. 1, (June, 2019) 951–965, arXiv:1811.08155 [astro-ph.CO].

39] N. Bellomo, J. L. Bernal, G. Scelfo, A. Raccanelli, and L. Verde, “Beware of commonly used approximations I: errors in forecasts”, In prep. (2020).

40] J. Lesgourgues, “The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview”,

– 23 –
[41] D. Blas, J. Lesgourgues, and T. Tram, “The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes”, JCAP 1107 (2011) 034, arXiv:1104.2933 [astro-ph.CO].

[42] R. A. Fisher, “The Fiducial Argument in Statistical Inference”, Annals Eugen. 6 (1935) 391–398.

[43] M. Schmittfull, M. Simonović, V. Assassi, and M. Zaldarriaga, “Modeling Biased Tracers at the Field Level”, arXiv e-prints (Nov, 2018) arXiv:1811.10640, arXiv:1811.10640 [astro-ph.CO].

[44] D. Ginzburg, V. Desjacques, and K. C. Chan, “Shot noise and biased tracers: A new look at the halo model”, Phys. Rev. D 96 no. 8, (Oct, 2017) 083528, arXiv:1706.08738 [astro-ph.CO].

[45] F. Montanari and R. Durrer, “Measuring the lensing potential with tomographic galaxy number counts”, JCAP 2015 no. 10, (Oct., 2015) 070, arXiv:1506.01369 [astro-ph.CO].

[46] M. Tegmark, A. N. Taylor, and A. F. Heavens, “Karhunen-Loève Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets?”, Astrophys. J. 480 (May, 1997) 22–35, astro-ph/9603021.

[47] S. Hamimeche and A. Lewis, “Likelihood analysis of CMB temperature and polarization power spectra”, Phys. Rev. D 77 (May, 2008) 103013, arXiv:0801.0554.

[48] A. Raccanelli, O. Doré, and N. Dalal, “Optimization of spectroscopic surveys for testing non-Gaussianity”, JCAP 8 (Aug., 2015) 034, arXiv:1409.1927.

[49] A. Raccanelli, M. Shiraishi, N. Bartolo, D. Bertacca, M. Liguori, S. Matarrese, R. P. Norris, and D. Parkinson, “Future constraints on angle-dependent non-Gaussianity from large radio surveys”, Physics of the Dark Universe 15 (Mar., 2017) 35–46.

[50] E.-M. Mueller, W. J. Percival, and R. Ruggeri, “Optimizing primordial non-Gaussianity measurements from galaxy surveys”, MNRAS 485 (May, 2019) 4160–4166, arXiv:1702.05088.

[51] J. L. Bernal, A. Raccanelli, E. D. Kovetz, D. Parkinson, R. P. Norris, G. Danforth, and C. Schmitt, “Probing ΛCDM cosmology with the Evolutionary Map of the Universe survey”, JCAP 2019 no. 2, (Feb, 2019) 030, arXiv:1810.06672 [astro-ph.CO].

[52] D. Karagiannis, A. Lazanu, M. Liguori, A. Raccanelli, N. Bartolo, and L. Verde, “Constraining primordial non-Gaussianity with bispectrum and power spectrum from upcoming optical and radio surveys”, MNRAS 478 no. 1, (July, 2018) 1341–1376, arXiv:1801.09280 [astro-ph.CO].

[53] S. Matarrese, L. Verde, and R. Jimenez, “The Abundance of High-Redshift Objects as a Probe of Non-Gaussian Initial Conditions”, Astrophys. J. 541 (Sept., 2000) 10–24, astro-ph/0001366.

[54] N. Dalal, O. Doré, D. Huterer, and A. Shirokov, “Imprints of primordial non-Gaussianities on large-scale structure: Scale-dependent bias and abundance of virialized objects”, Phys. Rev. D 77 no. 12, (June, 2008) 123514, arXiv:0710.4560.

[55] S. Matarrese and L. Verde, “The Effect of Primordial Non-Gaussianity on Halo Bias”, Astrophys. J. Letters 677 (Apr., 2008) L77, arXiv:0801.4826.

[56] V. Desjacques and U. Seljak, “Primordial non-Gaussianity from the large-scale structure”, Classical and Quantum Gravity 27 no. 12, (June, 2010) 124011, arXiv:1003.5020.

[57] M. Grossi, L. Verde, C. Carbone, K. Dolag, E. Branchini, F. Iannuzzi, S. Matarrese, and L. Moscardini, “Large-scale non-Gaussian mass function and halo bias: tests on N-body simulations”, MNRAS 398 no. 1, (Sept., 2009) 321–332, arXiv:0902.2013 [astro-ph.CO].

[58] C. Wagner and L. Verde, “N-body simulations with generic non-Gaussian initial conditions II: halo bias”, JCAP 2012 no. 3, (Mar., 2012) 002, arXiv:1102.3229 [astro-ph.CO].
[59] J. Yoo, A. L. Fitzpatrick, and M. Zaldarriaga, “New perspective on galaxy clustering as a cosmological probe: General relativistic effects”, Phys. Rev. D 80 no. 8, (Oct., 2009) 083514, arXiv:0907.0707 [astro-ph.CO].

[60] J. Yoo, “General relativistic description of the observed galaxy power spectrum: Do we understand what we measure?”, Phys. Rev. D 82 no. 8, (Oct., 2010) 083508, arXiv:1009.3021.

[61] C. Bonvin and R. Durrer, “What galaxy surveys really measure”, Phys. Rev. D 84 (Sept., 2011) 063505.

[62] A. Challinor and A. Lewis, “Linear power spectrum of observed source number counts”, Phys. Rev. D 84 (Aug., 2011) 043516.

[63] D. Jeong, F. Schmidt, and C. M. Hirata, “Large-scale clustering of galaxies in general relativity”, Phys. Rev. D 85 no. 2, (Jan., 2012) 023504, arXiv:1107.5427.

[64] D. Bertacca, R. Maartens, A. Raccanelli, and C. Clarkson, “Beyond the plane-parallel and Newtonian approach: wide-angle redshift distortions and convergence in general relativity”, JCAP 10 (Oct., 2012) 025, arXiv:1205.5221.

[65] A. Raccanelli, F. Montanari, D. Bertacca, O. Doré, and R. Durrer, “Cosmological measurements with general relativistic galaxy correlations”, JCAP 2016 (May, 2016) 009.

[66] A. Raccanelli, D. Bertacca, R. Maartens, C. Clarkson, and O. Doré, “Lensing and time-delay contributions to galaxy correlations”, General Relativity and Gravitation 48 (July, 2016) 84.

[67] H. Hildebrandt, “Observational biases in flux magnification measurements”, MNRAS 455 no. 4, (Feb., 2016) 3943–3951, arXiv:1511.01352 [astro-ph.GA].

[68] L. van Waerbeke, “Shear and magnification: cosmic complementarity”, MNRAS 401 no. 3, (Jan., 2010) 2093–2100, arXiv:0906.1583 [astro-ph.CO].

[69] B. Casaponsa, A. F. Heavens, T. D. Kitching, L. Miller, R. B. Barreiro, and E. Martínez-González, “Size magnification as a complement to cosmic shear”, MNRAS 430 no. 4, (Apr., 2013) 2844–2853, arXiv:1209.1646 [astro-ph.CO].

[70] E. L. Turner, J. P. Ostriker, and J. R. Gott, III, “The statistics of gravitational lenses - The distributions of image angular separations and lens redshifts”, Astrophysical Journal 284 (Sep, 1984) 1–22.

[71] M. Bartelmann and P. Schneider, “Weak gravitational lensing”, Physics Report 340 (Jan., 2001) 291–472.

[72] D. N. Limber, “The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field”, Astrophysical Journal 117 (1953) 134.

[73] N. Kaiser, “Weak Lensing and Cosmology”, The Astrophysical Journal 498 no. 1, (1998) 26, arXiv:astro-ph/9610120.

[74] M. LoVerde and N. Afshordi, “Extended Limber approximation”, Phys. Rev. D 78 (Dec, 2008) 123506, arXiv:0809.5112.

[75] U. Seljak, “Extracting Primordial Non-Gaussianity without Cosmic Variance”, Physical Review Letters 102 no. 2, (Jan., 2009) 021302, arXiv:0807.1770.

[76] P. McDonald and U. Seljak, “How to evade the sample variance limit on measurements of redshift-space distortions”, JCAP 10 (Oct., 2009) 007, arXiv:0810.0323.

[77] J. L. Bernal, T. L. Smith, K. K. Boddy, and M. Kamionkowski, “Robustness of baryon acoustic oscillations constraints to beyond-ΛCDM cosmologies”, arXiv e-prints (Apr., 2020) arXiv:2004.07263, arXiv:2004.07263 [astro-ph.CO].

[78] B. Joachimi and A. N. Taylor, “Forecasts of non-Gaussian parameter spaces using Box–Cox transformations”, Monthly Notices of the Royal Astronomical Society 416 no. 2, (2011)
[79] E. Sellentin, M. Quartin, and L. Amendola, “Breaking the spell of Gaussianity: forecasting with higher order Fisher matrices”, MNRAS 441 (June, 2014) 1831–1840, arXiv:1401.6892.

[80] E. Sellentin, “A fast, always positive definite and normalizable approximation of non-Gaussian likelihoods”, Monthly Notices of the Royal Astronomical Society 453 no. 1, (2015) 893–898, arXiv:1506.04866.

[81] L. Amendola and E. Sellentin, “Optimizing parameter constraints: a new tool for Fisher matrix forecasts”, Monthly Notices of the Royal Astronomical Society 457 no. 2, (2016) 1490–1495, arXiv:1602.01746.