Astronomy & Astrophysics

Fast and easy super-sample covariance of large-scale structure observables*

Fabien Lacasa[1] and Julien Grain[2]

[1] Département de Physique Théorique and Center for Astroparticle Physics, Université de Genève, 24 quai Ernest Ansermet, 1211 Geneva, Switzerland
e-mail: `fabien.lacasa@unige.ch`
[2] Institut d'Astrophysique Spatiale, CNRS (UMR8617) and Université Paris-Sud 11, Bâtiment 121, 91405 Orsay, France

ABSTRACT

We present a numerically cheap approximation to super-sample covariance (SSC) of large-scale structure cosmological probes, first in the case of angular power spectra. No new elements are needed besides those used to predict the considered probes, thus relieving analysis pipelines from having to develop a full SSC modeling, and reducing the computational load. The approximation is asymptotically exact for fine redshift bins $\Delta z \rightarrow 0$. We furthermore show how it can be implemented at the level of a Gaussian likelihood or a Fisher matrix forecast as a fast correction to the Gaussian case without needing to build large covariance matrices. Numerical application to a *Euclid*-like survey show that, compared to a full SSC computation, the approximation nicely recovers the signal-to-noise ratio and the Fisher forecasts on cosmological parameters of the wCDM cosmological model. Moreover, it allows for a fast prediction of which parameters are going to be the most affected by SSC and at what level. In the case of photometric galaxy clustering with *Euclid*-like specifications, we find that σ_8, n_s, and the dark energy equation of state w are particularly heavily affected. We finally show how to generalize the approximation for probes other than angular spectra (correlation functions, number counts, and bispectra) and at the likelihood level, allowing for the latter to be non-Gaussian if necessary. We release publicly a Python module allowing the implementation of the SSC approximation and a notebook reproducing the plots of the article.

Key words. large-scale structure of Universe – galaxies: statistics – methods: data analysis – methods: analytical

1. Introduction

The distribution of matter on large scales in the Universe is one of the main cosmological probes allowing for shading lights, for example on dark matter, dark energy, and gravity at cosmological scales. The current surveys of galaxies such as the Kilo-Degree Survey (KiDS; Hildebrandt et al. 2017) and the Dark Energy Survey (DES; Dark Energy Survey Collaboration 2018) recently provided cosmological constraints on the ΛCDM model from galaxy clustering and weak lensing, which are now competitive with constraints derived from the lensing of the Cosmic Microwave Background (CMB) and consistent with CMB primary anisotropies (for a recent comparison, see e.g., Planck Collaboration VI 2018). In the near future, large surveys such as the Large Synoptic Sky Telescope (LSST; LSST Science Collaborations 2009) and the *Euclid* satellite mission (Laureijs et al. 2011) will greatly improve our understanding of the structure of the Universe, the nature and the properties of dark energy, the potential modification of gravity at cosmological scales, and the initial conditions of cosmological perturbations (Amendola et al. 2013).

Unlike CMB primary anisotropies, however, late-time tracers of the large-scale structures (LSS) evolved through nonlinear dynamics, and as a result, the probability distribution function (pdf) of probes such as the galaxy distribution or weak lensing by LSS is no longer Gaussian, with deviation from a Gaussian distribution increasing at smaller scales. This first means that not all the information is compressed in the two-point statistics of the considered probes. Second, this means that the covariance of statistical observables built from LSS tracers (e.g., any n-point statistics) is increased by the presence of non-Gaussian contributions; as an example, the covariance on angular power spectra will be increased by contributions from a nonvanishing trispectrum. In the present context of preparing the cosmological interpretation of forthcoming datasets, and of forecasting the expected performance of future galaxy surveys that aim at precision cosmology from LSS tracers, it is now necessary to properly take into account the non-Gaussian contribution to the covariance for any inference of cosmological parameters from LSS observables.

Among the different non-Gaussian sources to the covariance (see Lacasa 2018 for a full derivation) is the super-sample covariance (SSC), first discovered for cluster counts by Hu & Kravtsov (2003) and to which a vast amount of literature has been devoted (e.g., Takada & Hu 2013; Takada & Spergel 2014; Takahashi et al. 2014; Li et al. 2018; Chan et al. 2018; Lacasa et al. 2018; Barreira et al. 2018a,b). This additional source of cosmic variance is inherent to all galaxy surveys due to the limited portion of the Universe that is observed, both in redshift depth and in sky fraction. SSC hence comes from the nonlinear impact of density fluctuations with wavelengths greater than the survey size. These super-survey modes modulate the local observables by making the background density averaged over the survey size to be nonrepresentative (either denser or

less dense than) of the averaged density in the Universe. Barring systematics, SSC is expected to be the dominant source of statistical error or cosmic variance for weak lensing (Barreira et al. 2018a) beyond the usual Gaussian covariance, although other terms may also be important for galaxy clustering (Lacasa 2018). It affects the whole set of statistical observables and correlates them. Contrary to intrasurvey sources of covariance, it can be shown that SSC cannot be reliably calibrated from the data itself or from classical simulations (Lacasa & Kunz 2017). This thus motivates the need for analytical or semi-analytical predictions of the effect, for use in the analysis of current and future galaxy surveys.

When analyzing such galaxy surveys, we usually deal with observables O_i being line-of-sight integrals of the form $O_i = \int dV_i \, \mathfrak{o}_i$, where $\mathfrak{o}_i$ is the comoving density of the observable (including selection effects such as redshift binning) and $dV = r^2(z)\frac{dr}{dz}dz$ is the comoving volume per steradian. Then the rigorous super-sample covariance for such observables is given by (e.g., Lacasa & Rosenfeld 2016)

$$\mathrm{Cov_{SSC}}\,(O_1, O_2) = \iint dV_1 dV_2 \, \frac{\partial \mathfrak{o}_1}{\partial \delta_b}(z_1) \, \frac{\partial \mathfrak{o}_2}{\partial \delta_b}(z_2) \, \sigma^2(z_1, z_2). \quad (1)$$

In the above, $\frac{\partial \mathfrak{o}_1}{\partial \delta_b}(z_1)$ is the response of the probe which amounts how a given probe varies with changes of the background density δ_b. The quantity $\sigma^2(z_1, z_2)$ reads (assuming full sky here for simplicity)

$$\sigma^2(z_1, z_2) = \frac{1}{2\pi^2} \int k^2 \, dk \, P_{\mathrm{m}}(k|z_{12}) \, j_0(kr_1) j_0(kr_2), \quad (2)$$

with $P_{\mathrm{m}}(k|z_{12})$ the linear matter cross-spectrum between redshifts z_1 and z_2, and j_0 the spherical Bessel functions. It basically amounts to the variation in background density on a given survey volume due to super-survey modes modulations.

Computing this SSC contribution to the covariance exactly, however, becomes costly very quickly. In practice, we take advantage of the separability in redshift (e.g., Lacasa et al. 2018; Barreira et al. 2018b) to reduce the cost of a covariance evaluation to that of an angular power spectrum evaluation. However the covariance needs to be evaluated at every pair of multipoles. For future surveys doing angular power spectra analysis with $\ell_{\max}$ of a few thousands, this induces a $O(10^3)$ slow-down of prediction pipelines, which can be increased by more orders of magnitude if we include tomography ($\Rightarrow$ pairs of redshift bins) and combine probes ($\Rightarrow$ pairs of probes).

Furthermore, an exact computation necessitates the knowledge of the probe's response $\frac{\partial \mathfrak{o}}{\partial \delta_b}$, either through analytical means or through simulations, for every redshift and multipoles, which is a barrier for analysts not already experts in the field of SSC. It is thus desirable to have instead simpler functions, if not fixed parameters, as we will find later on, with reference ansatzs that can be easily implemented by the community.

The aim of this article is thus to present an approximation for the SSC that allows fast numerical computation and ease of use by the community, and to assess its accuracy in a forecast analysis using the Fisher matrix approach.

The article is organized as follows. Our approximation is presented in Sect. 2 for the case of angular power spectra as our statistical observables. This approximation basically abolishes the above-mentioned numerical burden, and makes the computation of the super-sample covariance matrix as fast as the computation of the involved angular power spectra. Furthermore, we show in Sect. 3 that the resulting matrix form enables fast application to common uses of the covariance (i.e., in a Gaussian likelihood

or for computation of a signal-to-noise ratio, S/N, or a Fisher matrix) as a correction to the Gaussian case. Then in Sect. 4 we show numerical results validating the approximation and giving its range of applicability. Finally in Sect. 5 we generalize the approach to other statistics (number counts, correlation function, and bispectrum) and to the full likelihood, making the implementation of super-sample covariance feasible even if the likelihood is not Gaussian.

We release publicly a Python code that allows the easy implemention of SSC with our approach[1].

2. Approximating the SSC

We consider the case of the angular power spectra cross-correlating two LSS tracers, A and B. In the context of galaxy surveys, these two tracers typically are galaxy clustering and galaxy shear. This can, however, be extended to other LSS tracers such as lensing of the CMB or the integrated Sachs-Wolfe (iSW) effect. The signals are observed in some redshift bins indicated with indices i_z, j_z, etc., and with a given width. In full generality, the redshift bins may overlap[2].

We use the Limber approximation throughout the article, both for the power spectrum and the super-sample covariance. The approximation is accurate enough for the power spectrum on the range of scales of our later forecast ($\ell \geq 50$). Furthermore, it is even more adapted to super-sample covariance because SSC impacts the covariance on small scales $\ell \gtrsim 300$, as we show in Sect. 4.

With Limber approximation, the angular power spectrum between two signals can generally be written as

$$C_\ell^{AB}(i_z, j_z) = \int dV \, W_{i_z}^A(z) \, W_{j_z}^B(z) \, P_{AB}(k_\ell|z). \quad (3)$$

The weighting kernels $W_{i_z}^A(z)$, $W_{j_z}^B(z)$ are nonzero over the width of the redshift bin, and are expressed in units of [probe unit] · sr/(Mpc/h)^3. The quantity $P_{AB}(k_\ell|z)$ is the 3D power spectrum of the considered probe, evaluated at the Limber wavenumber $k_\ell = (\ell + 1/2)/r(z)$, with $r(z)$ the comoving distance. Weighting kernels and power spectra for the different probes of interest (galaxy clustering and shear, CMB lensing, iSW effect) are given in Appendix A.

For an angular power spectrum, the comoving density of the observable, i.e., $\mathfrak{o}_{AB}$ entering Eq. (1), is $\mathfrak{o}_{AB} = W_{i_z}^A(z) \, W_{i_z}^B(z) \, P_{AB}(k_\ell|z)$. Assuming that in Eq. (1) the responses, $\frac{\partial \mathfrak{o}_{AB}}{\partial \delta_b}$, vary slowly with redshift compared to $\sigma^2(z_1, z_2)$, we arrive at the approximation that is the basis of this article:

$$\mathrm{Cov_{SSC}}\left(C_\ell^{AB}(i_z, j_z), C_{\ell'}^{CD}(k_z, l_z)\right) \approx R_\ell^{AB} \, C_\ell^{AB}(i_z, j_z) \, R_{\ell'}^{CD} \, C_{\ell'}^{CD}(k_z, l_z)$$
$$\times S_{i_z, j_z; k_z, l_z}^{A,B;C,D}, \quad (4)$$

where the double integrals over redshift in Eq. (1) have been approximately performed. The matrix $S_{i_z, j_z; k_z, l_z}^{A,B;C,D}$ is the dimensionless volume-averaged (co)variance of the background matter density contrast

$$S_{i_z, j_z; k_z, l_z}^{A,B;C,D} = \int dV_1 \, dV_2 \, \frac{W_{i_z}^A(z_1) \, W_{j_z}^B(z_1)}{I^{AB}(i_z, j_z)} \, \frac{W_{k_z}^C(z_2) \, W_{l_z}^D(z_2)}{I^{CD}(k_z, l_z)} \, \sigma^2(z_1, z_2),$$
$$(5)$$

---

[1] https://github.com/fabienlacasa/PySSC

[2] This is the case for the shear signals and the iSW effect since they are integrated signals from the redshift of the source plane to the observer.

with

$$I^{AB}(i_z, j_z) = \int dV_1 \, W^A_{i_z}(z_1) \, W^B_{j_z}(z_1). \tag{6}$$

The quantity R_ℓ is the effective relative response of the considered power spectrum. In the context of second-order perturbation theory (hereafter 2PT), the growth-only response of the matter power spectrum is $\frac{\partial P(k)}{\partial \delta_b} = \frac{68}{21} P(k)$ (e.g., Takada & Hu 2013), i.e., $R_\ell = \frac{68}{21}$. Other nonlinear terms are present, however, which increase the total response. In Appendix C we detail these terms, and present a full computation of the response. Our formalism is valid for a general scale-dependent response. In numerical applications later in this article, we test the approximation both with the full response and with the simpler ansatz $R_\ell = 5 \equiv R$, which is the effective value found in Appendix C.

The key point is that starting from the approximation in Eq. (4) makes the computation of the SSC for angular power spectra have the same numerical cost as the computation of the power spectra themselves.

3. Application to parameter constraints

In this section we examine the consequences for data analysis or forecasts of the SSC covariance given by Eq. (4) as an update to the covariance, i.e., the total covariance is $C = C_{\text{noSSC}} + C_{\text{SSC}}$, where C_{noSSC} is the sum of all other contributions to the covariance matrix.

A common statistical use of a covariance matrix, C, is to compute scalar quantities of the form[3]

$$I = X^T \cdot C^{-1} \cdot Y, \tag{7}$$

where the dot "·" stands for matrix multiplication. For example, to compute the cumulative S/N we would have $X = Y = (C_\ell)_{\ell = \ell_{\min} \cdots \ell_{\max}} \equiv C$. In the exponent of a Gaussian likelihood, we would need $X = Y = \hat{C} - C(\boldsymbol{p})$, where $\hat{C} \equiv (\hat{C}_\ell)_{\ell = \ell_{\min} \cdots \ell_{\max}}$ is the estimated/measured power spectrum and $C(\boldsymbol{p}) \equiv (C_\ell(\boldsymbol{p}))_{\ell = \ell_{\min} \cdots \ell_{\max}}$ is the predicted power spectrum with model parameters $\boldsymbol{p}$. Finally for Fisher forecasts, computing the Fisher matrix, $F_{\alpha,\beta}$, requires $X = \partial C / \partial \boldsymbol{p}_\alpha \equiv \partial_\alpha C$ and $Y = \partial_\beta C$.

This last case is the primary aim of the article since it is a measure of the amount of information we have on cosmological parameters from the observables C. The Fisher matrix will be our figure of merit to gauge the quality of the approximation, with numerical results to be presented in Sect. 4.

Computing the scalar quantities given by Eq. (7) requires the inversion of the covariance matrix, $C = C_{\text{noSSC}} + C_{\text{SSC}}$. Using the approximation Eq. (4), adding the SSC corresponds to a rank 1 update of the covariance matrix of the angular power spectrum C_ℓ. Furthermore, in Appendix D, we detail the way to introduce binned power spectra in this approach, and we show that adding the SSC to the covariance of the binned spectra is also a rank 1 update of the covariance.

We thus make use of the Sherman–Morrison formula (Sherman & Morrison 1950; Bartlett 1951), which gives the

---

[3] For all these cases of interest, the matrix-vector products in Eq. (7) have to be understood as

$$I = \sum_{\ell, \ell' = \ell_{\min}}^{\ell_{\max}} X_\ell \, [C^{-1}]_{\ell \ell'} \, Y_{\ell'},$$

i.e., it corresponds to a cumulative quantity over multipole.

impact on matrix inversion of a rank 1 update,

$$\left(A + UV^T\right)^{-1} = A^{-1} - \frac{\left(A^{-1} \cdot U\right)\left(V^T \cdot A^{-1}\right)}{1 + V^T \cdot A^{-1} \cdot U}, \tag{8}$$

where A is any $n \times n$ square matrix, and U and V are two n-dimensional vectors, and T means the transpose.

3.1. Single probe and single redshift bin

If we neglect all non-Gaussian terms except SSC, the noSSC covariance reduces to the Gaussian term which is diagonal in full sky:

$$(C_G)_{\ell, \ell'} = G_\ell \, \delta_{\ell, \ell'} \qquad \text{with} \qquad G_\ell = \frac{2C_\ell^2}{2\ell + 1}. \tag{9}$$

This can simplify the inversion of the noSSC covariance later on. In partial sky observations, the Gaussian covariance will not be diagonal due to mask-induced couplings between different angular scales (i.e., ℓ-to-ℓ' couplings). It can nevertheless be made diagonal in practice by binning the power spectrum with bins wider than the typical width of the mask-induced couplings, as shown in Appendix D.

In the following we keep the covariance general throughout the derivation, and indicate when appropriate which expressions are simplified by the diagonal assumption. We also keep the same subscript, ℓ, to indicate either single multipoles or bins of multipoles.

The super-sample covariance has a separable form between the two multipoles so that we can write the total covariance as

$$C = C_{\text{noSSC}} + S_{i,i} \left(VV^T\right), \tag{10}$$

where V is a vector with size the number of multipoles, given by

$$V_\ell = R_\ell \, C_\ell, \tag{11}$$

and $S_{i,i}$ is just a number. The Sherman–Morrison formula Eq. (8) then gives the inverse covariance as

$$C^{-1} = C_{\text{noSSC}}^{-1} - \frac{S_{i,i} \left(C_{\text{noSSC}}^{-1} V\right)\left(V^T C_{\text{noSSC}}^{-1}\right)}{1 + S_{i,i} \left(V^T \cdot C_{\text{noSSC}}^{-1} \cdot V\right)}, \tag{12}$$

where $V^T \cdot C_{\text{noSSC}}^{-1} \cdot V$ is a scalar.

Thus, the scalar quantity defined in Eq. (7) is given by

$$I = I_{\text{noSSC}} - \frac{f_X^{\text{SSC}} \, f_Y^{\text{SSC}} \, S_{i,i}}{1 + V^T \cdot C_{\text{noSSC}}^{-1} \cdot V \, S_{i,i}}, \tag{13}$$

where we defined the scalar

$$f_X^{\text{SSC}} \equiv X^T \cdot C_{\text{noSSC}}^{-1} \cdot V \overset{\text{diagonal}}{=} \sum_\ell \frac{R_\ell \, C_\ell \, X_\ell}{G_\ell}. \tag{14}$$

The notation $\overset{\text{diagonal}}{=}$ in the above means that the assumption that the noSSC covariance matrix is diagonal has been used. In particular for Fisher matrices, SSC gives a negative correction to the noSSC case:

$$F_{\alpha,\beta} = F_{\alpha,\beta}^{\text{noSSC}} - \frac{f_\alpha^{\text{SSC}} \, f_\beta^{\text{SSC}} \, S_{i,i}}{1 + V^T \cdot C_{\text{noSSC}}^{-1} \cdot V \, S_{i,i}}, \tag{15}$$

with

$$f_\alpha^{\text{SSC}} \equiv \partial_\alpha C^T \cdot C_{\text{noSSC}}^{-1} \cdot V \overset{\text{diagonal}}{=} \sum_\ell \frac{R_\ell \, C_\ell \, \partial_\alpha C_\ell}{G_\ell}. \tag{16}$$

We finally note that all the above expressions for the impact of the SSC are easily extended to the case of binned spectra by replacing the vector V_ℓ by its binned version, V_b (see Appendix D).

3.2. Multi-probe and single redshift bin

A more complex case of interest is when we have spectra of different probes sharing the same redshift bin. One example is a multi-tracer analysis for galaxy clustering at a given redshift. Another is a weak-lensing analysis splitting galaxy types (e.g., red vs. blue galaxies) to better mitigate the effect of intrinsic alignments.

In the following we illustrate this with two probes (*A* and *B*) yielding three power spectra, although the results hold straightforwardly for more probes.

The three angular power spectra are generally correlated. So even in the Gaussian case, the covariance matrix is not diagonal as a function of probes, for example, even in full sky:

$$\text{Cov}_G(C_\ell^{AA}, C_\ell^{BB}) = \frac{2\left(C_\ell^{AB}\right)^2}{2\ell + 1} \neq 0. \tag{17}$$

However, it remains diagonal as a function of multipoles.

Given that they share the same redshift bin, and assuming the weighting kernels to have similar enough redshift dependence within the bin, the S matrix can be assumed independent of probes :

$$S_{i,i;i,i}^{W,X;Y,Z} = \text{cst} \equiv S_{i,i} \qquad \forall\, (W, X, Y, Z) \in \{A, B\}^4. \tag{18}$$

This property simplifies the super-sample covariance, allowing us to easily compute its impact on Fisher forecasts as we show below.

In the following we call n_c the number of spectra and n_ℓ the number of multipoles. It now becomes useful to arrange the $n_c \times n_\ell$ data vector $\boldsymbol{C}$ grouping probes together before multipoles, for example with two probes A and B:

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{C}_{\ell_{\min}} \\ \vdots \\ \boldsymbol{C}_{\ell_{\max}} \end{pmatrix} \qquad \text{with} \qquad \boldsymbol{C}_\ell = \begin{pmatrix} C_\ell^{AA} \\ C_\ell^{AB} \\ C_\ell^{BB} \end{pmatrix}. \tag{19}$$

The covariance matrix has a size $(n_c \times n_\ell) \times (n_c \times n_\ell)$. With $\boldsymbol{C}$ arranged as above, the covariance matrix is thus partitioned in $n_\ell \times n_\ell$ blocks, each of these blocks having a size $n_c \times n_c$.

In the Gaussian case, in full sky or with the f_{sky} approximation, the covariance matrix is block diagonal, and we call $\boldsymbol{G}_\ell$ these blocks of size $n_c \times n_c$ on the diagonal (see Appendix D for the case of binned spectra).

The formalism of Sect. 3.1 for a single probe is then easily adapted with only slight changes. We have the total covariance matrix

$$C = C_{\text{noSSC}} + C_{\text{SSC}} = C_{\text{noSSC}} + \boldsymbol{V}\boldsymbol{V}^T S_{i,i}, \tag{20}$$

where $\boldsymbol{V}$ is a $n_c \times n_\ell$ vector, given by

$$\boldsymbol{V}_\ell = R_\ell\, \boldsymbol{C}_\ell. \tag{21}$$

Then the inverse covariance follows

$$C^{-1} = C_{\text{noSSC}}^{-1} - \frac{S_{i,i}\left(C_{\text{noSSC}}^{-1}\boldsymbol{V}\right)\left(\boldsymbol{V}^T C_{\text{noSSC}}^{-1}\right)}{1 + S_{i,i}\left(\boldsymbol{V}^T \cdot C_{\text{noSSC}}^{-1} \cdot \boldsymbol{V}\right)}, \tag{22}$$

where $\boldsymbol{V}^T \cdot C_{\text{noSSC}}^{-1} \cdot \boldsymbol{V}$ is a scalar. For a block-diagonal covariance, it simplifies to

$$\boldsymbol{V}^T \cdot C_{\text{noSSC}}^{-1} \cdot \boldsymbol{V} \overset{\text{diagonal}}{=} \sum_\ell \boldsymbol{C}_\ell^T \cdot \boldsymbol{G}_\ell^{-1} \cdot \boldsymbol{C}_\ell, \tag{22}$$

with the inner matrix products being in the space of the n_c spectra, and appropriately reduced to the case of Sect. 3.1 when $n_c = 1$.

Thus, the scalar quantity defined in Eq. (7) is given by

$$I = I_{\text{noSSC}} - \frac{f_X^{\text{SSC}}\, f_Y^{\text{SSC}}\, S_{i,i}}{1 + \boldsymbol{V}^T \cdot C_{\text{noSSC}}^{-1} \cdot \boldsymbol{V}\, S_{i,i}}, \tag{23}$$

where we defined the scalar

$$f_X^{\text{SSC}} \equiv \boldsymbol{X}^T \cdot C_{\text{noSSC}}^{-1} \cdot \boldsymbol{V} \overset{\text{diagonal}}{=} \sum_\ell \boldsymbol{X}_\ell \cdot \boldsymbol{G}_\ell^{-1} \cdot R_\ell\, \boldsymbol{C}_\ell, \tag{24}$$

with again the inner matrix products in the space of the n_c spectra.

Finally, the total Fisher matrix is given by the noSSC Fisher matrix plus a negative SSC correction:

$$F_{\alpha,\beta} = F_{\alpha,\beta}^{\text{noSSC}} - \frac{f_\alpha^{\text{SSC}}\, f_\beta^{\text{SSC}}\, S_{i,i}}{1 + \boldsymbol{V}^T \cdot C_{\text{noSSC}}^{-1} \cdot \boldsymbol{V}\, S_{i,i}}, \tag{25}$$

with

$$f_\alpha^{\text{SSC}} \equiv \partial_\alpha \boldsymbol{C}^T \cdot C_{\text{noSSC}}^{-1} \cdot \boldsymbol{V} \overset{\text{diagonal}}{=} \sum_\ell \partial_\alpha \boldsymbol{C}_\ell^T \cdot \boldsymbol{G}_\ell^{-1} \cdot R_\ell\, \boldsymbol{C}_\ell. \tag{26}$$

3.3. Multi-probe and multiple redshift bins

A first case of interest is when we have probes in different nonoverlapping bins, or when the overlap is small enough to be neglected. This happens for instance for galaxies, cluster counts, or power spectra in sufficiently wide bins, i.e., larger than the photo-z error bars, and ≥ 0.1 to be larger than the width of the SSC $\sigma^2(z_1, z_2)$ (see Fig. 6 in Lacasa & Rosenfeld 2016).

In this case we can basically add up the bins independently,

$$I = \sum_i I_{i,i} = I_{\text{noSSC}} - \Delta I_{\text{SSC}}, \tag{27}$$

where $I_{i,i}$ is given by Eq. (23). The (negative) SSC correction is

$$\Delta I_{\text{SSC}} = \sum_i \frac{f_X^{\text{SSC}}\, f_Y^{\text{SSC}}\, S_{i,i}}{1 + \boldsymbol{V}^T \cdot C_{\text{noSSC}}^{-1} \cdot \boldsymbol{V}\, S_{i,i}}. \tag{28}$$

In particular for Fisher forecasts, the (negative) SSC correction reads

$$\Delta F_{\alpha,\beta}^{\text{SSC}} = \sum_i \frac{f_\alpha^{\text{SSC}}\, f_\beta^{\text{SSC}}\, S_{i,i}}{1 + \boldsymbol{V}^T \cdot C_{\text{noSSC}}^{-1} \cdot \boldsymbol{V}\, S_{i,i}}, \tag{29}$$

which is obtained as the sum over independent redshift bins of the SSC corrections derived for one single bin.

The second case of interest is when bins are overlapping. This happens for instance when analyzing galaxy shear, which integrates the signal from $z = 0$ to the sources, either alone or in combination with other probes. In that case, no simplifications can be carried out: the power spectra are correlated both as a function of multipoles and as a function of redshift bins. The covariance matrix must be built in full generality using Eq. (4), and then inverted numerically. We note that already at the Gaussian level inversion must be carried out numerically, due to the coupling between redshift bins.

3.4. Importance of SSC: an analytical rule of thumb

The importance of SSC can be gauged easily in an analytical way, if we assume a single redshift bin, and further approximate the response R_ℓ to be independent of scale $R_\ell \equiv R$. It is important to note that this scale independent assumption is not a requirement for numerical application, and may be relaxed as is be done in Sect. 4.

In this section, we gauge the importance of SSC analytically first for the S/N, then for the Fisher constraints.

3.4.1. Impact on the signal-to-noise ratio

The S/N of a set of angular power spectra, collected in a single data vector $\boldsymbol{C}$, is given by

$$(S/N)^2 = \boldsymbol{C} \cdot C^{-1} \cdot \boldsymbol{C}, \tag{30}$$

and we also introduce the noSSC version of it as

$$(S/N)_{\rm G}^2 = \boldsymbol{C} \cdot C_{\rm noSSC}^{-1} \cdot \boldsymbol{C}. \tag{31}$$

We recall that these are cumulative S/N values obtained as a summation over multipoles up to a maximum value $\ell_{\max}$. It is then a function of the maximum multipole up to which we integrate our observables. Let us also introduce the scalar quantity

$$Y \equiv \boldsymbol{V}^T \cdot C_{\rm noSSC}^{-1} \cdot \boldsymbol{V} \, S_{i,i}. \tag{32}$$

With the assumption of a single redshift bin and a scale-independent response, the scalar Y is shown to be proportional to the square of the noSSC S/N, i.e.,

$$Y = (S/N)_{\rm noSSC}^2 \, R^2 \, S_{i,i}. \tag{33}$$

Then the total (S/N) boils down to

$$(S/N)^2 = (S/N)_{\rm noSSC}^2 \left(1 - \frac{Y}{1+Y}\right) = \frac{(S/N)_{\rm noSSC}^2}{(1+Y)}. \tag{34}$$

It is thus obvious that the SSC decreases the S/N compared to the noSSC case as Y is by construction a positive number. The impact of the SSC is enhanced for higher values of Y, which is then an excellent indicator of its importance[4].

From Eq. (33), the impact of the SSC increases for higher S/N: the higher $(S/N)_{\rm noSSC}$, the higher Y. Enlarging the set of power spectra to smaller angular scales (i.e., increasing $\ell_{\max}$ to higher multipoles) increases the S/N, hence the impact of the SSC. By integrating to smaller scales, the entire S/N will thus reach a plateau at an asymptotic value

$$(S/N)_{\max}^2 = \frac{(S/N)_{\rm G}^2}{Y} = \frac{1}{R^2 \, S_{i,i}}. \tag{35}$$

This saturation is reached when $Y \sim 1$. For the case of a full sky cosmic variance-limited analysis of a single power spectrum up to a maximum multipole $\ell_{\max}$, and neglecting other non-Gaussian terms, we have $Y \sim \frac{\ell_{\max}^2}{2} R^2 S_{i,i}$. The typical angular scales above which the (S/N) starts to saturate because of the

SSC is defined by $Y \simeq 1$. We thus find that SSC becomes important when the analysis goes up to $\ell_{\max} \gtrsim \ell_{\rm SSC}$ given by

$$\ell_{\rm SSC} = \sqrt{\frac{2}{R^2 \, S_{i,i}}}. \tag{36}$$

Generalizing to the case of partial sky coverage and several cosmic variance-limited probes (in the same redshift bin and neglecting other non-Gaussian terms), the analysis is affected as soon as it reaches multipoles of order

$$\ell_{\rm SSC} = \sqrt{\frac{2}{N_p^{\rm eff} \, R^2 \, f_{\rm sky} \, S_{i,i}}}, \tag{37}$$

where

$$N_p^{\rm eff} = \frac{2 \, (S/N)_{\rm G,joint}^2}{f_{\rm sky} \, \ell_{\max}^2} \tag{38}$$

is the effective number of probes[5].

Finally, if the probes are not cosmic variance-limited, for example due to the presence of shot-noise (galaxy clustering) or shape noise (weak lensing), or if other non-Gaussian covariance terms are important, then we need a full computation of the noSSC S/N accounting for these additional sources of error; the criterion for the importance of SSC is $(S/N)_{\rm noSSC}^2 \gtrsim 1/(R^2 S_{i,i})$. We note that this critical value is that of the maximum S/N with the full covariance (Eq. (35)), this plateau being the same for single- and multi-probe cases (as long as all probes have the same response R); in other words, it is the maximum amount of information that can be extracted from matter fluctuations in a finite volume of the universe with probes with a given response, regardless of the number of probes.

3.4.2. Impact on Fisher constraints

The S/N is highly impacted by SSC, because in the SSC dominated regime and with a constant R_ℓ, Eq. (4) shows that C_ℓ measurements are 100% correlated and thus all information is lost on the overall amplitude. However, one may question the impact on cosmological parameters, if they are sensitive to other features of the power spectrum.

We first recall Eq. (25) for the Fisher information on model parameters α and β,

$$F_{\alpha,\beta} = F_{\alpha,\beta}^{\rm noSSC} - \frac{f_\alpha^{\rm SSC} \, f_\beta^{\rm SSC} \, S_{i,i}}{1+Y}, \tag{39}$$

rewritten here using Y. We further introduce two angles: first, the angle θ_α between the vectors $\boldsymbol{C}$ and $\partial_\alpha \boldsymbol{C}$,

$$\cos \theta_\alpha = \frac{\partial_\alpha \boldsymbol{C}^T \cdot C_{\rm noSSC}^{-1} \cdot \boldsymbol{V}}{\sqrt{\partial_\alpha \boldsymbol{C}^T \cdot C_{\rm noSSC}^{-1} \cdot \partial_\alpha \boldsymbol{C} \times \boldsymbol{V}^T \cdot C_{\rm noSSC}^{-1} \cdot \boldsymbol{V}}}$$

$$= \frac{f_\alpha^{\rm SSC}}{\sqrt{F_{\alpha,\alpha}^{\rm noSSC} \times Y/S_{i,i}}}, \tag{40}$$

---

[4] In the case of many *uncorrelated* redshift bins, Eq. (34) is generalized by summing over the redshift bins, i.e.,

$$(S/N)^2 = \sum_i \frac{(S/N)_{\rm G,i}^2}{(1+Y_i)}.$$

[5] It is exactly the number of probes if they are uncorrelated, but it goes down to 1 if they are totally correlated.

and second, the angle $\theta_{\alpha\beta}$ between $\partial_\alpha C$ and $\partial_\beta C$,

$$
\begin{aligned}
\cos\theta_{\alpha\beta} &= \frac{\partial_\alpha C^T \cdot C_{\mathrm{noSSC}}^{-1} \cdot \partial_\beta C}{\sqrt{\partial_\alpha C^T \cdot C_{\mathrm{noSSC}}^{-1} \cdot \partial_\alpha C \times \partial_\beta C^T \cdot C_{\mathrm{noSSC}}^{-1} \cdot \partial_\beta C}} \\
&= \frac{F_{\alpha,\beta}^{\mathrm{noSSC}}}{\sqrt{F_{\alpha,\alpha}^{\mathrm{noSSC}} \times F_{\beta,\beta}^{\mathrm{noSSC}}}}.
\end{aligned}
\tag{41}
$$

Let us roughly interpret these angles. The second, $\theta_{\alpha\beta}$, is easily interpreted as the noSSC correlation between the parameter α and the parameter β. The first angle, θ_α, can be interpreted as follows. We recall that V is proportional to the data vector C. Up to a normalization constant, C and thus V can be viewed as $\partial_A C$, where A is the normalization of the data vector. Since angles are obtained from normalized vectors, θ_α is thus a measure of the noSSC correlation between the parameter α and the normalization of the data vector.

The Fisher information matrix including SSC is then conveniently expressed as a function of the noSSC Fisher matrix, the parameters Y measuring the impact of the SSC on the S/N, and the angles θ_α, θ_β, and $\theta_{\alpha\beta}$:

$$
F_{\alpha,\beta} = F_{\alpha,\beta}^{\mathrm{noSSC}} \left(1 - \frac{\cos\theta_\alpha \cos\theta_\beta}{\cos\theta_{\alpha\beta}} \frac{Y}{1+Y} \right).
\tag{42}
$$

The impact of the SSC on the Fisher matrix is driven first by the impact of the SSC on the (S/N) through Y, and then by the angles θ_α, θ_β, and $\theta_{\alpha\beta}$. In particular for the diagonal elements, the change of the Fisher matrix is

$$
\frac{\delta F_{\alpha,\alpha}}{F_{\alpha,\alpha}^{\mathrm{noSSC}}} = -\cos^2\theta_\alpha \left(\frac{Y}{1+Y} \right).
\tag{43}
$$

This has a negative value, showing that the SSC lowers the amount of information on a given parameter.

Two conditions have to be met for the impact of the SSC to be important: Y should be greater than one, and $\cos^2\theta_\alpha$ should be close to one. Supposing $Y \gg 1$ and if $\theta_\alpha \neq 0, \pi$, i.e., $\cos\theta_\alpha \neq \pm 1$[6], then the Fisher information keeps increasing with $\ell_{\max}$, but at a reduced rate compared to the noSSC case, with the asymptote

$$
F_{\alpha,\alpha} \sim F_{\alpha,\alpha}^{\mathrm{noSSC}} \left(1 - \cos^2\theta_\alpha \right),
\tag{44}
$$

i.e., the unmarginalized error bar is increased as

$$
\sigma_\alpha \sim \frac{\sigma_\alpha^G}{|\sin\theta_\alpha|}.
\tag{45}
$$

A maximum impact of the SSC is thus obtained for θ_α close to zero, that is when the parameter α is at the noSSC level highly correlated with the normalization of the data vector.

When we have several parameters, the situation becomes more complex and cannot be judged with only rules of thumb. For example, a parameter α may seem unaffected by SSC because $\cos\theta_\alpha \ll 1$, but it may be correlated (already at the noSSC level) with a parameter β which is affected by SSC, so that α will be affected indirectly through marginalization. Another possibility is that we have two parameters that are uncorrelated at the noSSC level, but through Eq. (42) they

---

[6] The case $\cos\theta_\alpha = \pm 1$ corresponds to the parameter α being $\pm$ the amplitude of the power spectrum. In this case, up to a normalization, we go back to the case of the S/N studied in Sect. 3.4.1.

become correlated due to SSC ; in that case a large error on one parameter would rebound on the other, which did not happen in the noSSC case.

Summary. To decide the importance of SSC on parameter constraints, first compute the multipole above which the SSC dominated regime is entered, i.e.,

$$
\ell_{\mathrm{SSC}} = \sqrt{\frac{2}{N_p\, R_\ell^2\, f_{\mathrm{sky}}\, S_{i,i}}}.
\tag{46}
$$

If the analysis is restricted to scales such as $\ell \ll \ell_{\mathrm{SSC}}$, then it will not be affected. If the analysis enters the SSC dominated regime, then for each parameter of interest it is necessary to compute the angle

$$
\cos\theta_\alpha = \frac{\langle \partial_\alpha C, V \rangle}{\|\partial_\alpha C\|\, \|V\|}, \qquad \text{where} \quad \langle X, Y \rangle = X^T \cdot C_G^{-1} \cdot Y,
\tag{47}
$$

and the (unmarginalized) error bar on parameter α is increased asymptotically as

$$
\sigma_\alpha \sim \frac{\sigma_\alpha^G}{|\sin\theta_\alpha|}.
\tag{48}
$$

In the case with several cosmological parameters and/or nuisance parameters that need to be marginalized over, a full computation is necessary.

4. Numerical application to Fisher forecasts

To test the accuracy of the proposed SSC approximation, and to illustrate its impact on a cosmological analysis, we perform here an application to forecast the cosmological constraints from photometric galaxy clustering with the following specifications:
- single redshift bin with a top-hat window $0.9 < z < 1$, and galaxy numbers representative of *Euclid*: 28M galaxies in the bin, corresponding to a density of ~ 2.5 gal arcmin^{-2};
- full sky coverage and an analysis in the multipole range $50 < \ell < 2000$, in bins of constant width $\Delta\ell = 50$ (a realistic range for a conservative photometric galaxy cosmological analysis relying on a constant bias model);
- flat wCDM model with fiducial cosmological parameters from *Planck* 2013 ΛCDM constraints (Planck Collaboration XVI 2014):

$$
\{\Omega_b h^2, \Omega_c h^2, n_s, \sigma_8, H_0, w\} = \{0.022, 0.12, 0.96, 0.83, 67, -1\};
$$

- other non-Gaussian covariance terms beyond SSC are neglected.

With these specifications, we compute the galaxy angular power spectrum C_ℓ^{gal} using the halo model and halo occupation distribution as done in Lacasa & Rosenfeld (2016). On the one hand we find that the shot-noise level is completely negligible, so that we are signal-dominated over all the scales of interest. On the other hand, we find that the nonlinear part of the power spectrum is important on those scales, with the one-halo term dominating the two-halo term for $\ell > 800$.

For the S matrix, computed following Appendix B.2, we find the value

$$
S_{i,i} = 6.2 \times 10^{-7}.
\tag{49}
$$

Following Sect. 3.4.1 and assuming a scale independent response $R = 5$, this translates into a knee multipole $\ell_{\mathrm{SSC}} \sim 360$, and a plateau at $S/N \sim 250$.
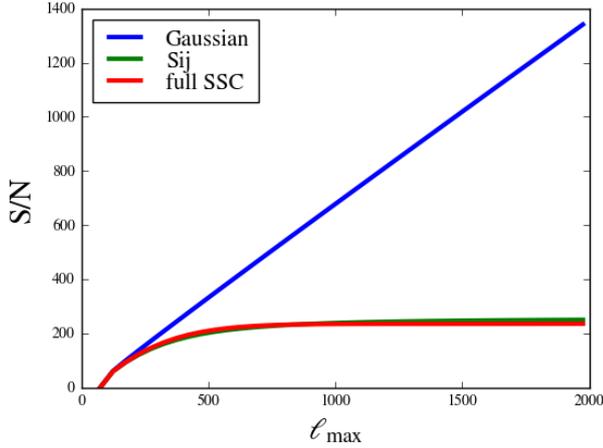
Fig. 1. Comparison of the cumulative S/N values up to a multipole $\ell_{\max}$ with different covariances: Gaussian only, with the $S_{i,j}$ approximation, and with a full SSC computation.

Fig. 2. $\cos^2 \theta$ for each cosmological parameter as a function of the maximum multipole $\ell_{\max}$. Parameters with $\cos^2 \theta$ close to 1 are the most affected by SSC when it starts to dominate the covariance (i.e., for multipoles $\ell \gtrsim \ell_{\mathrm{SSC}}$ with ℓ_{SSC} 360 in this specific case).

The cumulative S/N as a function of the maximum multipole of the analysis $\ell_{\max}$ is shown in Fig. 1. The (S/N) is computed with three different covariance matrices: Gaussian only, Gaussian plus SSC through Eq. (34), and Gaussian plus a full SSC computation following Lacasa & Rosenfeld (2016). We see that the Gaussian covariance matrix completely overestimates the significance of the angular power spectrum by a factor of ~ 5.7 at $\ell_{\max} = 2000$. However, the $S_{i,j}$ approximation with a constant response does recover precisely the full SSC computation over the whole multipole range with a precision better than 7%. We also see that the value of the knee multipole and the (S/N) plateau mentioned previously indeed capture the features of the full SSC curve. The S_{ij} approximation is thus validated at the level of the S/N.

The impact of the SSC on the cosmological parameter estimation is depicted in Figs. 2 and 3. Figure 2 shows the angle $\cos^2 \theta_\alpha$, defined by Eq. (40), for the main cosmological parameters of the wCDM model considered here. Interestingly, all curves are significantly different from zero. Following Sect. 3.4.2, this means that all parameters are going to be significantly impacted by SSC when reaching small scales (i.e., $\ell \gtrsim \ell_{\mathrm{SSC}}$). In this specific case, a value of $\cos^2 \theta_\alpha$ significantly different from zero will thus matter for an $\ell_{\max}$ greater than ~ 400. Some parameters should be affected less ($\Omega_b h^2$, $\Omega_c h^2$, and h), and others more (σ_8, n_s, and w_0).

Figure 3 confirms the qualitative conclusion of Fig. 2. It shows for each cosmological parameter the square root of the Fisher element as a function of the maximum multipole $\ell_{\max}$ of analysis. This coefficient has an evolution with $\ell_{\max}$ similar to the S/N[7], and its inverse is the error bar on the considered cosmological parameter when the other parameters are fixed[8]. As in the case of the (S/N), on large scales ($\ell_{\max} < \ell_{\mathrm{SSC}}$) the impact of the SSC is negligible. This happens in spite of the angle $\cos^2 \theta_\alpha$ being close to one for all the parameters, simply because the impact of the SSC on the (S/N) is subdominant: $Y \ll 1$. When

we integrate the signal to smaller scales, however ($\ell_{\max} \gtrsim \ell_{\mathrm{SSC}}$), the Gaussian covariance significantly overestimates the strength of the angular power spectrum: the Fisher elements become significantly smaller with the full covariance compared to the Gaussian case. Furthermore, the parameters most affected by SSC are indeed those that were identified in Fig. 2 (σ_8, n_s, and w_0), although the other parameters are still significantly affected. We note that this finding is in agreement with the recent analysis of Barreira et al. (2018a) in the different case of weak lensing.

We now briefly discuss the specific case of h, which is paradigmatic of how the SSC has an impact on the estimation of cosmological parameters via the interplay between Y and $\cos^2 \theta_\alpha$. For $\ell_{\max} < \ell_{\mathrm{SSC}}$, the Gaussian covariance and the full covariance gives the same results since $Y \ll 1$. Then for a range of multipoles roughly given by $500 \lesssim \ell_{\max} \lesssim 1500$, the Gaussian covariance and the full covariance do not give the same S/N; however, the impact of SSC on h is small in this range of multipole because the angle $\cos^2 \theta_h$ is close to zero, as shown in Fig. 2, which suppresses the impact of the SSC. Finally, for $\ell_{\max} \gtrsim 1500$, both $Y \gg 1$ and $\cos^2 \theta_h \sim 1$, and the impact of the SSC is clearly seen since the Gaussian covariance now overestimates $F_{h,h}$ compared to the full covariance.

Comparing the $S_{i,j}$ approximation with a constant response R to the full SSC computation, we find that the former reproduces the Fisher elements of the less affected parameters ($\Omega_b h^2$, $\Omega_c h^2$, and h) to 3% precision, the Fisher on n_s to better than 8% precision, but it is less precise for two of the heavily affected parameters (σ_8 and w_0), where the precision is around 30% at the highest $\ell_{\max}$. This is potentially an issue for application to surveys as *Euclid* with requirements of 10% precision of marginalized errors on cosmological parameters if the analysis is pursued to these small scales. We found that the approximation respects the 10% precision on SSC up to $\ell_{\max} = 1000$, but becomes less precise afterwards. We tracked the issue, and found that it originates from the assumption of a constant response $R_\ell = \mathrm{d} \ln C_\ell / \mathrm{d} \delta_b$. Using the proper scale-dependent response shown in Appendix C, we found that the $S_{i,j}$ approximation reproduces the Fisher elements of the full SSC computation to 5% precision over the whole multipole range.

The $S_{i,j}$ approximation is thus validated at the level of parameter constraints. In the constant response case, it reproduces the Fisher constraints to acceptable precision, except deep

---

[7] The unmarginalized S/N on a given parameter α is simply obtained by multiplying this coefficient with the input value of the parameter. It coincides with the S/N in the case of a model parameter being the amplitude A of the power spectrum: $C_\ell = A C_\ell^{\mathrm{template}}$.

[8] We do not attempt any marginalization or production of realistic forecasts. Our framework is different from the usual cosmological analyses since we use the halo model and HOD; for instance, we cannot marginalize over galaxy bias.
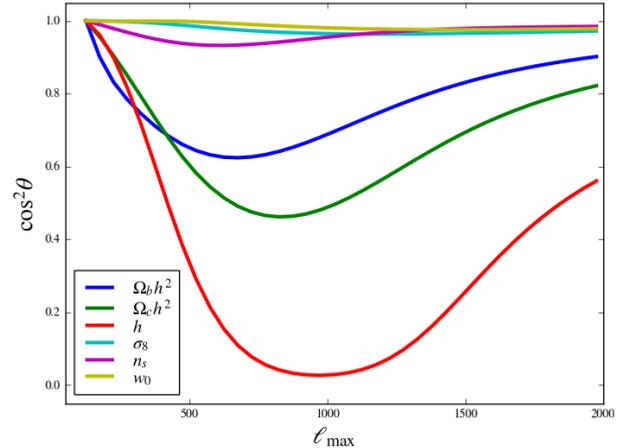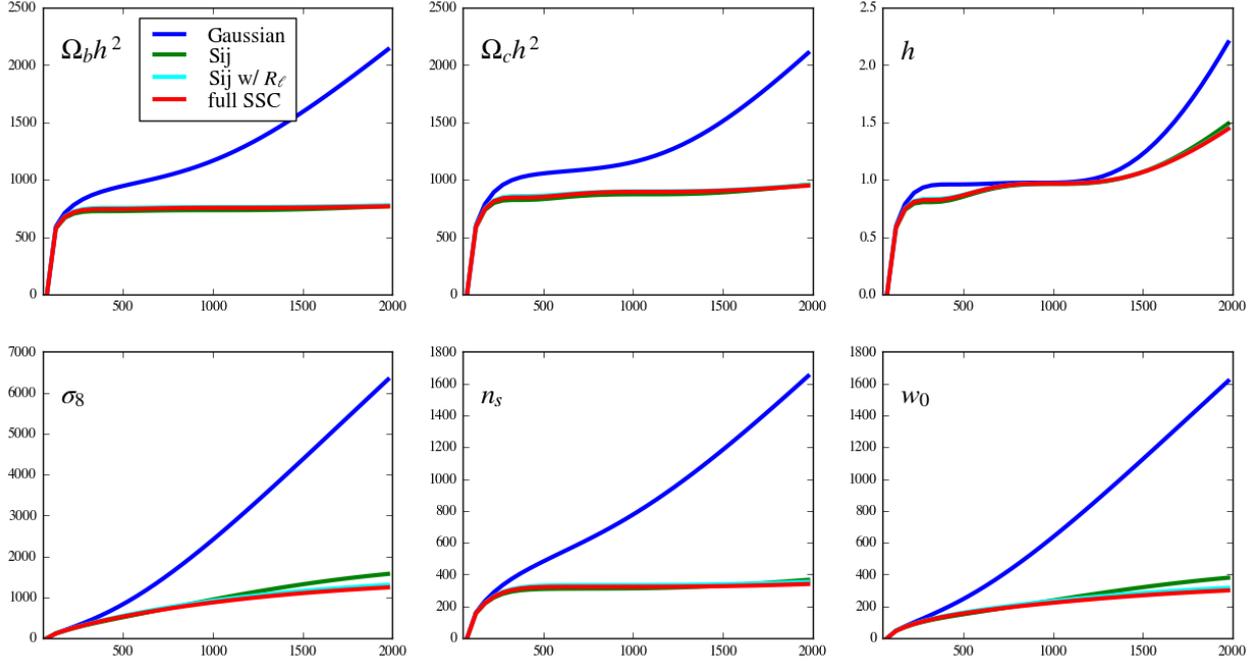
Fig. 3. Comparison of $\sqrt{F_{\alpha\alpha}}$ for an analysis up to a multipole ℓ_{max} with different covariances: Gaussian only (blue), $S_{i,j}$ approximation with a constant response (green), $S_{i,j}$ approximation with a scale-dependent response (cyan), full SSC computation (red).

in the SSC-dominated regime for the most affected parameters. Accounting for the scale-dependence of the response allows us to recover all parameter constraints to sufficient precision, if it is necessary to pursue the analysis to small scales.

5. Generalizations of the SSC approximation

5.1. Generalization to other statistics

A first note is that for 3D statistics there is no need for an approximation like that in Eq. (4). In such cases, analyses commonly use the (often implicit) assumption of no redshift evolution within the volume. And that assumption means that the SSC covariance already takes the same form as Eq. (4) (see, e.g., Takada & Hu 2013 for the 3D matter power spectrum $P(k)$).

Number counts. The case of cluster number counts is where the $S_{i,j}$ approximation actually first started, devised by Hu & Kravtsov (2003). Among other counts of interest for LSS surveys are those of galaxies and shear peaks. Generally, we note $N_\alpha(i_z)$ the number counts, with an index α specifying the type of object as well as the bin of the considered property (e.g., mass, luminosity, color, shear S/N). The response of these counts is the first order bias,

$$\frac{\partial N_\alpha}{\partial \delta_b} = b_\alpha N_\alpha, \tag{50}$$

i.e., $R_\alpha = b_\alpha$. The analog of Eq. (4) is then

$$\text{Cov}_{\text{SSC}}\left(N_\alpha(i_z), N_\beta(j_z)\right) \approx b_\alpha(i_z) N_\alpha(i_z) b_\beta(j_z) N_\beta(j_z) \times S^{\alpha;\beta}_{i_z;j_z}. \tag{51}$$

The SSC approximation is also extended to the cross-covariance with an angular power spectrum

$$\text{Cov}_{\text{SSC}}\left(N_\alpha(i_z), C^{CD}_{\ell'}(k_z, l_z)\right) \approx b_\alpha(i_z) N_\alpha(i_z) R^{CD}_{\ell'} C^{CD}_{\ell'}(k_z, l_z)$$
$$\times S^{\alpha;C,D}_{i_z;k_z,l_z}. \tag{52}$$

In the case of the angular power spectrum, we were able to consider the response as having weak scale dependence, and thus approximate R_ℓ = cst. In the case of counts this is generally not the case. For instance, for clusters the bias has a strong dependence on mass and will thus vary from bin to bin.

Correlation function. The 2D correlation function is a linear transform of the angular power spectrum

$$w(\theta) = \sum_\ell \frac{2\ell+1}{4\pi} C_\ell P_\ell(\cos\theta). \tag{53}$$

It is thus readily seen that its SSC covariance takes the form

$$\text{Cov}_{\text{SSC}}\left(w^{AB}_{i_z,j_z}(\theta), w^{CD}_{k_z,l_z}(\theta')\right) \approx \tilde{w}^{AB}_{i_z,j_z}(\theta)\, \tilde{w}^{CD}_{k_z,l_z}(\theta') \times S^{A,B;C,D}_{i_z,j_z;k_z,l_z}, \tag{54}$$

with

$$\tilde{w}(\theta) = \sum_\ell \frac{2\ell+1}{4\pi} R_\ell C_\ell P_\ell(\cos\theta) = (R * w)(\theta) \tag{55}$$

the convolution product (denoted $*$) of the original correlation function with the response $R(\theta)$. If the response $R_\ell \equiv R$ can be assumed constant, as in most of this article, then $R(\theta)$ reduces to a Dirac distribution $R(\theta) = R \times \delta(\theta)$. The convolution product of R with w thus simplifies to a standard product, i.e., $\tilde{w}(\theta) = R \times w(\theta)$, so that Eq. (54) takes the same form as Eq. (4).

Bispectrum. Analogously to Eq. (4), the SSC covariance for bispectra coefficients will take the form

$$\text{Cov}_{\text{SSC}}\left(b^{ABC}_{\ell_1\ell_2\ell_3}(i_z, j_z, k_z), b^{DEF}_{\ell'_1\ell'_2\ell'_3}(l_z, m_z, n_z)\right) \approx R^{ABC}_{\ell_1\ell_2\ell_3} b^{ABC}_{\ell_1\ell_2\ell_3}(i_z, j_z, k_z)$$
$$\times R^{DEF}_{\ell'_1\ell'_2\ell'_3} b^{DEF}_{\ell'_1\ell'_2\ell'_3}(l_z, m_z, n_z)\, S^{A,B,C;D,E,F}_{i_z,j_z,k_z;l_z,m_z,n_z}. \tag{56}$$

The 3D bispectrum SSC was studied extensively in Chan et al. (2018), who found that the growth-only response from perturbation theory is $R_{k_1k_2k_3} = \frac{433}{126}$, while the total response ranges between 4 and 6. Small BAO features are visible in these responses, but should we washed out in 2D projected quantities. A constant response $R = 5$ may thus give an acceptable first-order approximation, as we found in this article for C_ℓ.

5.2. Generalization to the likelihood

A usual assumption is that the likelihood of the observable vector O (e.g., the power spectrum $C = (C_\ell)_{\ell=\ell_{\min}\cdots\ell_{\max}}$, as in most of this article) given model parameters $\boldsymbol{p}$, is a multivariate Gaussian

$$P(O|\boldsymbol{p}) = \mathcal{N}(\bar{O}(\boldsymbol{p}), C_{\text{tot}}), \tag{57}$$

where $\mathcal{N}(X, \Sigma)$ denotes the Gaussian distribution with mean X and covariance Σ.

Given that $C_{\text{tot}} = C_{\text{std}} + C_{\text{SSC}}$ (e.g., for the power spectrum $C_{\text{std}} = C_{\text{G}}$) and using properties of Gaussian distributions, this can be rewritten (artificially for the moment) as the convolution

$$\begin{aligned} P(O|\boldsymbol{p}) &= \mathcal{N}(\bar{O}(\boldsymbol{p}), C_{\text{std}}) * \mathcal{N}(0, C_{\text{SSC}}) \\ &= P_{\text{std}}(O|\boldsymbol{p}) * \mathcal{N}(0, C_{\text{SSC}}), \end{aligned} \tag{58}$$

where P_{std} is the standard (no-SSC) likelihood. The second pdf can be interpreted physically as the probability that super-survey modes induce a shift of the observable vector.

We can then reformulate the probability in a form similar to that found for cluster counts by Lima & Hu (2004) (see also Appendix E),

$$P(O|\boldsymbol{p}) = \int \mathrm{d}(\delta O)\, P_{\text{std}}(O|\bar{O}(\boldsymbol{p}) + \delta O, C_{\text{G}})\, P_{\text{SSC}}(\delta O|0, C_{\text{SSC}}), \tag{59}$$

where the shift δO is a random variable with probability P_{SSC}, i.e., centered on zero and with covariance matrix C_{SSC}.

At first order, the observable reacts to the change of background δ_b induced by long wavelength modes, through the response $\frac{\partial O}{\partial \delta_b}$ (e.g., the power spectrum response discussed in Sect. 2 and Appendix C). Noting that

$$\bar{O}(\boldsymbol{p}, \delta_b) = \bar{O}(\boldsymbol{p}) + \delta O = \bar{O}(\boldsymbol{p}) + \frac{\partial O}{\partial \delta_b}\delta_b \tag{60}$$

is the (average) observable in a part of the universe with a background change δ_b, we can rewrite the likelihood as

$$P(O|\boldsymbol{p}) = \int \mathrm{d}\delta_b \underbrace{P_{\text{std}}(O|\bar{O}(\boldsymbol{p}, \delta_b), C_{\text{G}})\, P_{\text{SSC}}(\delta_b|0, S)}_{\equiv P(O|\boldsymbol{p}, \delta_b)}, \tag{61}$$

where the S matrix defined in Eq. (5) appears. We note that δ_b is not a simple scalar; it depends on the pair of probes and redshift bins of the observable considered.

Because δ_b is the density field smoothed over very large scales (the whole survey area), it is safe to assume that it has a Gaussian distribution, i.e., $P_{\text{SSC}} = \mathcal{N}(0, S)$. However, the same may not be true of $P_{\text{std}}(O|\boldsymbol{p})$, where the observable may not have a Gaussian likelihood. For instance in the case of cluster counts studied in Lima & Hu (2004), the observable follows a Poissonian distribution if $S = 0$. In the case of the angular power spectrum, which is a quadratic quantity, the observable follows a Wishart distribution on the full sky (e.g., Hamimeche & Lewis 2008), which has an important impact on inference from large angular scales. For galaxy lensing, this has been shown to be of importance by Sellentin et al. (2018). In the case of the bispectrum, it is also known that the likelihood should not be Gaussian (Chan & Blot 2017), although no numerical or analytical form exists for it at the moment.

This is where the rewriting Eq. (61) becomes useful in practice (beyond giving a nice physical interpretation) since we can now use for $P_{\text{std}}(O|\bar{O}(\boldsymbol{p}, \delta_b)$ a more realistic and possibly non-Gaussian pdf.

This means that SSC can be accounted for at the likelihood level through the hierarchical model

$$P(O|\boldsymbol{p}, \delta_b) = P_{\text{std}}\left(O|\bar{O}(\boldsymbol{p}, \delta_b)\right) \times \pi(\delta_b|\boldsymbol{p}), \tag{62}$$

where π is the prior on δ_b, i.e., $\mathcal{N}(0, S)$, where the S matrix depends implicitly on cosmological parameters (and potentially on other model parameters if they affect the weighting kernels). The value of δ_b then needs to be marginalized over to get constraints on the standard model parameters.

We note that in the separate universe approach, a region with a background change in a cosmology $\boldsymbol{p}$ can be simulated as a region with no background change but a different cosmology with parameters $\boldsymbol{p}'(\boldsymbol{p}, \delta_b)$ (Wagner et al. 2015). Thus, the likelihood Eq. (62) may be implemented with only small changes to current no-SSC likelihood pipelines,

$$P(O|\boldsymbol{p}, \delta_b) = P_{\text{std}}\left(O|\bar{O}\left(\boldsymbol{p}'(\boldsymbol{p}, \delta_b)\right)\right) \times \pi(\delta_b|\boldsymbol{p}), \tag{63}$$

which avoids having to model or measure the observable's response, and means that accounting for SSC is as easy as including extra nuisance parameters.

6. Conclusion

We presented a fast and easy approximation for the super sample covariance of 2D projected statistics; the study was mainly focused on the angular power spectrum C_ℓ and generalization to other statistics given later. In addition to the considered probe, this $S_{i,j}$ approximation relies on two ingredients:

- the S matrix, which is an integral of the (linear) matter power spectrum convolved with the survey window. In the flat sky limit, computable expressions are found in the literature (e.g., Aguena & Lima 2018; see Appendix B.2 for expressions for the full sky and partial sky cases).
- the probe's response. We found the simple ansatz $R = 5$ to perform very well for the case of C_ℓ. It is sufficient for *Euclid* precision requirements on parameter constraints for cosmological parameters of wCDM model, and up to $\ell_{\max} \sim 1000$. To push to smaller scales for σ_8 and w, it is necessary to account for the scale dependence of R_ℓ (see Appendix C, Table C.1).

Neither of these ingredients requires expensive computations or physical models additional to the usual cosmological tools necessary to predict the considered probes. The $S_{i,j}$ approximation can thus readily be implemented in cosmological prediction codes. Furthermore, we showed that SSC can be included in cosmological pipelines (for significance quantification, Fisher forecasts, or MCMC parameter estimation) through a simple correction to the Gaussian covariance case, not even spoiling the speed-up induced by a diagonal covariance.

The $S_{i,j}$ approximation also allows us to easily identify which cosmological parameters are going to be affected by SSC and at what level through the fast computation of the $\cos\theta_\alpha$ coefficients (Eq. (40)) and the scalar Y (Eq. (32)).

To facilitate the use of the $S_{i,j}$ approximation by the community, we publicly release a Python code that implements it, together with examples of applications[9].

In the case of photometric galaxy clustering in a redshift bin $0.9 < z < 1$ and with *Euclid*-like specifications, we found all cosmological constraints to be heavily impacted by SSC. This makes it necessary to include this effect in forecasts and analysis pipelines for future galaxy surveys, a task now largely eased

---

[9] https://github.com/fabienlacasa/PySSC

up by the $S_{i,j}$ approximation. Furthermore we showed how this approximation can be generalized beyond the angular power spectrum to other statistics such as number counts, the correlation function, and bispectrum, where we indicated the corresponding probe's response[10].

Finally, we showed that the $S_{i,j}$ approximation can be generalized at the likelihood level, which avoids having to assume a Gaussian likelihood, an assumption that is incorrect in many cases, for example cluster counts at high masses, C_ℓ at low ℓ or the correlation function on large scales, and the bispectrum. We will explore these likelihood developments in future works.

References

Aguena, M., & Lima, M. 2018, Phys. Rev. D, 98, 123529
Amendola, L., Appleby, S., Bacon, D., et al. 2013, Liv. Rev. Rel., 16, 6
Barreira, A., Krause, E., & Schmidt, F. 2018a, J. Cosmol. Astropart. Phys., 10, 053
Barreira, A., Krause, E., & Schmidt, F. 2018b, J. Cosmol. Astropart. Phys., 6, 015
Bartlett, M. S. 1951, Ann. Math. Statist., 22, 107
Chan, K. C., & Blot, L. 2017, Phys. Rev. D, 96, 023528
Chan, K. C., Moradinezhad Dizgah, A., & Noreña, J. 2018, Phys. Rev. D, 97, 043532
Dark Energy Survey Collaboration (Abbott, T. M. C., et al.) 2018, Phys. Rev. D, 98, 043526
Hamimeche, S., & Lewis, A. 2008, Phys. Rev. D, 77, 103013
Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, MNRAS, 465, 1454
Hu, W., & Kravtsov, A. V. 2003, ApJ, 584, 702
Kilbinger, M., Heymans, C., Asgari, M., et al. 2017, MNRAS, 472, 2126
Lacasa, F. 2018, A&A, 615, A1
Lacasa, F., & Kunz, M. 2017, A&A, 604, A104
Lacasa, F., & Rosenfeld, R. 2016, J. Cosmol. Astropart. Phys., 8, 005
Lacasa, F., Lima, M., & Aguena, M. 2018, A&A, 611, A83
Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints [arXiv:1110.3193]
Lewis, A., & Challinor, A. 2006, Phys. Rep., 429, 1
Li, Y., Schmittfull, M., & Seljak, U. 2018, J. Cosmol. Astropart. Phys., 2, 022
Lima, M., & Hu, W. 2004, Phys. Rev. D, 70, 043504
LSST Science Collaborations (Abell, P. A., et al.) 2009, ArXiv e-prints [arXiv:0912.0201]
Planck Collaboration XVI. 2014, A&A, 571, A16
Planck Collaboration XXI. 2016, A&A, 594, A21
Planck Collaboration VI. 2018, ArXiv e-prints [arXiv:1807.06209]
Sellentin, E., Heymans, C., & Harnois-Déraps, J. 2018, MNRAS, 477, 4879
Sherman, J., & Morrison, W. J. 1950, Ann. Math. Stat., 21, 124
Takada, M., & Hu, W. 2013, Phys. Rev. D, 87, 123504
Takada, M., & Spergel, D. N. 2014, MNRAS, 441, 2456
Takahashi, R., Soma, S., Takada, M., & Kayo, I. 2014, MNRAS, 444, 3473
Wagner, C., Schmidt, F., Chiang, C.-T., & Komatsu, E. 2015, MNRAS, 448, L11

---

[10] Namely the object's bias for counts, and $R = 5$ for the correlation function and the bispectrum.

Appendix A: Example of weighting kernels

Here we give the weighting kernels and 3D power spectrum needed to compute the angular power spectra, Eq. (3), in the convention used in this article for different LSS observables.

Galaxy clustering. The observable is the projected galaxy number density in a redshift bin. In this case the weighting kernel is

$$W_{i_z}^{GC}(z) = \frac{n_{gal}^{(i_z)}(z)}{N_{gal}(i_z)}, \tag{A.1}$$

with $n_{gal}(z)$ the 3D comoving galaxy density and $N_{gal}(i_z)$ the 2D number of galaxies per solid angle. The 3D power spectrum is the galaxy power spectrum, i.e., $P_{gg}(k) = P_{gal}(k)$, which on large scales is linked to the matter power spectrum via $P_{gg}(k) = b_g^2 P_m(k)$.

Weak lensing / shear. The observable is the galaxy shear averaged over a redshift bin. In this case the weighting is

$$W^{\kappa_{gal}}(z) = \frac{\mathcal{A}}{a(z)\, r(z)}\, q_{i_z}(z), \tag{A.2}$$

with $\mathcal{A} = \frac{3}{2}\Omega_m \left(\frac{H_0}{c}\right)^2$ and $q_{i_z}(z) = \int v z' \frac{n_{gal}^{(i_z)}(z')}{N_{gal}(i_z)} \frac{r'-r}{r'}$ the lensing efficiency (e.g., Kilbinger et al. 2017). The 3D power spectrum is the matter power spectrum $P_m(k)$.

CMB lensing. The observable is the distortion of the CMB temperature anisotropies. In this case the weighting is

$$W^{\kappa_{CMB}}(z) = \frac{\mathcal{A}}{a(z)\, r(z)} \frac{r_* - r}{r_*}, \tag{A.3}$$

with r_* the comoving distance to the CMB last scattering surface (Lewis & Challinor 2006), and the 3D power spectrum is the matter power spectrum $P_m(k)$.

Integrated Sachs-Wolfe effect. The observable is the iSW contribution to the temperature anisotropies of the CMB. In this case the weighting kernel is (Planck Collaboration XXI 2016)

$$W^{iSW}(k, z) = -\frac{2\mathcal{A}}{k^2} \frac{d(1+z)\, G(z)}{dz} \Bigg/ \left(G(z)\, r^2(z) \frac{dr}{dz} \right), \tag{A.4}$$

with $G(z)$ the linear growth function. The 3D power spectrum is the matter power spectrum $P_{mm}(k)$.

We note that this kernel does not depend only on redshift, but also on wavenumber k, but since the latter dependence is factorizable, it cancels out in the S matrix and thus does not impact the applicability of the approximation Eq. (4).

We also note that the Limber approximation, used throughout this article, is poorly adapted to the iSW signal because it peaks at low multipoles. Lacasa (2018) provided expressions for super-sample covariance without Limber's approximation. We leave the generalization of the present SSC approximation to this no-Limber case to future work. However, as the SSC impact peaks on small scales, we expect that its impact on iSW constraints should be small and that the present approximation should thus be good enough to gauge its level.

Appendix B: Particular cases for $\sigma^2(z_1, z_2)$ and S matrix

B.1. $\sigma^2(z_1, z_2)$

In full sky we have (Lacasa & Rosenfeld 2016)

$$\sigma^2(z_1, z_2) = \frac{1}{2\pi^2} \int k^2\, dk\, P_m(k|z_{12})\, j_0(kr_1) j_0(kr_2). \tag{B.1}$$

In the partial sky case, $\sigma^2(z_1, z_2)$ can be expanded in spherical harmonics (Lacasa et al. 2018) to get

$$\sigma^2(z_1, z_2) = \frac{1}{\Omega_S^2} \sum_\ell (2\ell + 1)\, C_\ell(W)\, C_\ell^m(z_1, z_2), \tag{B.2}$$

where $\Omega_S = 4\pi f_{sky}$ is the solid angle covered by the survey, $C_\ell(W)$ is the angular power spectrum of the survey mask, and C_ℓ^m is the angular power spectrum of matter given by

$$C_\ell^m(z_1, z_2) = \frac{2}{\pi} \int k^2\, dk\, j_\ell(kr_1)\, j_\ell(kr_2)\, P_m(k|z_{12}). \tag{B.3}$$

B.2. S matrix

Assuming that super-survey modes can be described by linear theory, the matter power spectrum writes $P_m(k|z_{12}) = G(z_1)G(z_2)P(k)$, where $G(z)$ is the growth function and we note simply $P(k)$ the power spectrum at $z = 0$.

It then results from its definition Eq. (5) that the S matrix is given by

$$S_{i_z, j_z; k_z, l_z}^{A,B;C,D} = \frac{1}{2\pi^2} \int k^2\, dk\, P(k)\, \mathcal{I}_{i_z, j_z}^{AB}(k)\, \mathcal{I}_{k_z, l_z}^{CD}(k), \tag{B.4}$$

where

$$\mathcal{I}_{i_z, j_z}^{AB}(k) = \int dV \frac{W_{i_z}^A(z)\, W_{j_z}^B(z)}{I^{AB}(i_z, j_z)}\, G(z)\, \mathrm{Win}(kr), \tag{B.5}$$

where we recall

$$I^{AB}(i_z, j_z) = \int dV\, W_{i_z}^A(z)\, W_{j_z}^B(z). \tag{B.6}$$

and the angle-averaged survey window is

$$\mathrm{Win}(kr) = \begin{cases} j_0(kr) & \text{full sky} \\ \frac{4\pi}{\Omega_S^2} \sum_\ell (2\ell + 1)\, C_\ell(W)\, j_\ell(kr) & \text{partial sky} \end{cases} \tag{B.7}$$

with $C_\ell(W)$ the power spectrum of the survey mask (Lacasa et al. 2018).

There is a special case where the S matrix can be further simplified analytically by assuming a full sky survey where the weighting kernel is constant within the redshift bins, and by approximating the growth function at the center of the redshift bin. If the weighting kernel is constant, we have

$$W_i(z) = \frac{3}{r_{max}^3(i) - r_{min}^3(i)}\, \mathbb{1}_{z \in [z_{min}(i), z_{max}(i)]}. \tag{B.8}$$

This happens, for instance, in the case of galaxy clustering with perfect redshift determinations and if the galaxy comoving density can be considered constant $n_{gal}(z) = $ cst. Then Eq. (B.6) simplifies to

$$I^{AB}(i_z, j_z) = \frac{3\, \delta_{i_z, j_z}}{r_{max}^3(i_z) - r_{min}^3(i_z)}. \tag{B.9}$$

This gives for Eq. (B.5)

$$\begin{aligned} \mathcal{I}_{i_z, j_z}^{AB}(k) &\approx G(z_{mean}(i_z)) \frac{3\, \delta_{i_z, j_z}}{r_{max}^3(i_z) - r_{min}^3(i_z)} \int_{r_{min}(i_z)}^{r_{max}(i_z)} r^2\, dr\, j_0(kr) \\ &= \frac{3\, G(z_{mean})\, \delta_{i_z, j_z}}{k\left(r_{max}^3 - r_{min}^3\right)} \left[r_{max}^2\, j_1(kr_{max}) - r_{min}^2\, j_1(kr_{min}) \right], \end{aligned} \tag{B.10}$$

which can be fed into Eq. (B.4) for the S matrix, leading to

$$S^{A,B;C,D}_{i_z,j_z;k_z,l_z} = \delta_{i_z,j_z} \, \delta_{k_z,l_z} \, \frac{3\,G\,(z_{\mathrm{mean}}(i_z))}{r^3_{\mathrm{max}}(i_z) - r^3_{\mathrm{min}}(i_z)} \, \frac{3\,G\,(z_{\mathrm{mean}}(k_z))}{r^3_{\mathrm{max}}(k_z) - r^3_{\mathrm{min}}(k_z)}$$
$$\times \frac{1}{2\pi^2} \int k^2 \, \mathrm{d}k \, P(k)$$
$$\times \left[r^2_{\mathrm{max}}(i_z) \, j_1 \, (k r_{\mathrm{max}}(i_z)) - r^2_{\mathrm{min}}(i_z) \, j_1 \, (k r_{\mathrm{min}}(i_z)) \right] / k$$
$$\times \left[r^2_{\mathrm{max}}(k_z) \, j_1 \, (k r_{\mathrm{max}}(k_z)) - r^2_{\mathrm{min}}(k_z) \, j_1 \, (k r_{\mathrm{min}}(k_z)) \right] / k. \tag{B.11}$$

We note that this expression has become independent of the considered probes (A, B, C, D). Furthermore, the Bessel function j_1 is a sum of sines and cosines, $j_1(x) = \frac{\sin x}{x^2} - \frac{\cos x}{x}$, and thus the S matrix can be expressed in terms of Fourier transforms of the matter power spectrum. Specifically defining

$$I^\pm_{c,n}(r_1, r_2) \equiv \int \mathrm{d}k \, P(k)/k^n \, \cos(k(r_1 \pm r_2)), \tag{B.12}$$

$$I^\pm_{s,n}(r_1, r_2) \equiv \int \mathrm{d}k \, P(k)/k^n \, \sin(k(r_1 \pm r_2)), \tag{B.13}$$

$$F(r_1, r_2) \equiv -I^+_{c,4}(r_1, r_2) + I^-_{c,4}(r_1, r_2) - (r_1 + r_2) I^+_{s,3}(r_1, r_2)$$
$$+ (r_1 - r_2) I^-_{s,3}(r_1, r_2) + r_1 r_2 \left[I^+_{c,2}(r_1, r_2) + I^-_{c,2}(r_1, r_2) \right], \tag{B.14}$$

and shortening

$$r_{-,i} \equiv r_{\mathrm{min}}(i_z), \quad r_{-,k} \equiv r_{\mathrm{min}}(k_z), \quad r_{+,i} \equiv r_{\mathrm{max}}(i_z), \quad r_{+,k} \equiv r_{\mathrm{max}}(k_z), \tag{B.15}$$

we have the long expression

$$S_{i_z,j_z;k_z,l_z} = \delta_{i_z,j_z} \, \delta_{k_z,l_z} \, \frac{3\,G\,(z_{\mathrm{mean}}(i_z))}{r^3_{+,i} - r^3_{-,i}} \, \frac{3\,G\,(z_{\mathrm{mean}}(k_z))}{r^3_{+,k} - r^3_{-,k}} \times \frac{1}{4\pi^2}$$
$$\times \left[F(r_{+,i}, r_{+,k}) - F(r_{+,i}, r_{-,k}) - F(r_{-,i}, r_{+,k}) \right.$$
$$\left. + F(r_{-,i}, r_{-,k}) \right]. \tag{B.16}$$

Formally, the $I^\pm_{c/s,n}$ are IR divergent integrals as when $k \to 0$ we have $P(k) \propto k^{n_s}$ with $n_s \sim 1$. However, for every $I^+_{c/s,n}$ there is an opposite $I^-_{c/s,n}$ that carries the same divergence, which is thus cancelled. Hence, when applying a lower cutoff k_{min} to all integrals, the full expression Eq. (B.16) is convergent when $k_{\mathrm{min}} \to 0$. Numerically, we need to apply this cutoff and not put it too low in order to avoid cases of large cancellations between large numbers where numerical errors could spoil the result. Inspecting Eq. (B.11) and recalling $j_1(x) \propto x$ when $x \to 0$, we see that the integrand is $\propto k^2 P(k) \propto k^{2+n_s}$ in the IR, meaning that the integral is quickly converging. A conservative choice is thus to ensure that the start of the integral is at least one decade before the matter-radiation equality, and that the Bessel functions are in the small x regime. Hence, we take $k_{\mathrm{min}} = \min\{k_{\mathrm{eq}}, 1/r_{\mathrm{max}}\}/10$. Numerically, an upper cutoff k_{max} also needs to be taken. The integrals are convergent in this limit so this is a less pressing issue. With the same type of argument as for k_{min}, we can see that Eq. (B.16) is well converged if we take the value $k_{\mathrm{max}} = 10 \times \max\{k_{\mathrm{eq}}, 1/r_{\mathrm{min}}\}$.

Appendix C: Angular power spectrum response

Takada & Hu (2013) showed that the 3D matter power spectrum reacts to a change in background through two separate effects:
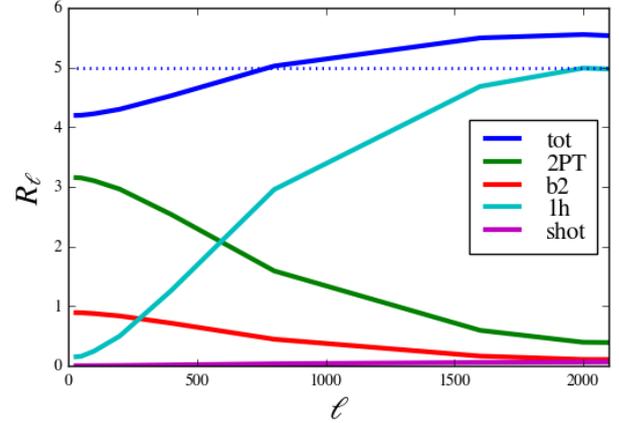
Fig. C.1. Power spectrum response R_ℓ and its different terms. The dashed line indicates the effective value taken in the analysis (see main text).

a term from second-order perturbation theory (2PT) that dominates on large scales, and a term from the one-halo part of the spectrum called halo sample variance that dominates on small scales. For the galaxy power spectrum, it was shown that the reaction also contains terms from second-order galaxy bias and shot-noise, and that the contribution from second-order nonlocal bias vanishes (Lacasa & Rosenfeld 2016; Lacasa 2018),

$$\frac{\partial P_{\mathrm{gal}}(k|z)}{\partial \delta_b} \equiv \left(\frac{68}{21} b^{\mathrm{gal}}_1(k, z)^2 + 2 b^{\mathrm{gal}}_2(k, z) \, b^{\mathrm{gal}}_1(k, z) \right) P_{\mathrm{m}}(k|z)$$
$$+ \frac{\partial P_{1\mathrm{h}}(k|z)}{\partial \delta_b} + b^{\mathrm{gal}}_1(k = 0, z)/\overline{n}_{\mathrm{gal}}(z), \tag{C.1}$$

where in the halo model

$$b^{\mathrm{gal}}_i(k, z) = \int \mathrm{d}M \, \frac{\mathrm{d}n_h}{\mathrm{d}M} \, \frac{\langle N_{\mathrm{gal}} \rangle}{\overline{n}_{\mathrm{gal}}(z)} \, u(k|M, z) \, b_i(M, z) \tag{C.2}$$

and

$$\frac{\partial P_{1\mathrm{h}}(k|z)}{\partial \delta_b} = \int \mathrm{d}M \, \frac{\mathrm{d}n_h}{\mathrm{d}M} \, \frac{\langle N_{\mathrm{gal}}(N_{\mathrm{gal}} - 1) \rangle}{\overline{n}_{\mathrm{gal}}(z)^2} \, u(k|M, z)^2 \, b_1(M, z) \tag{C.3}$$

with $\frac{\mathrm{d}n_h}{\mathrm{d}M}$ the halo mass function, $b_i(M, z)$ the i-th order halo bias, $u(k|M, z)$ the halo profile, and N_{gal} given by the halo occupation distribution.

We call the four terms in Eq. (C.1) respectively 2PT, b2, 1h, and shot. The reaction of the angular power spectrum then follows

$$\frac{\mathrm{d}C^{gg}_\ell(i_z, j_z)}{\mathrm{d}\delta_b} = \int \mathrm{d}V \, \frac{n^{(i_z)}_{\mathrm{gal}}(z)}{N_{\mathrm{gal}}(i_z)} \, \frac{n^{(j_z)}_{\mathrm{gal}}(z)}{N_{\mathrm{gal}}(j_z)} \, \frac{\mathrm{d}P_{\mathrm{gal}}(k_\ell|z)}{\mathrm{d}\delta_b}, \tag{C.4}$$

and it defines the (relative) response through

$$\frac{\mathrm{d}C^{gg}_\ell(i_z, j_z)}{\mathrm{d}\delta_b} = R^{gg}_\ell \, C^{gg}_\ell(i_z, j_z). \tag{C.5}$$

Figure C.1 shows this response and its different terms. On large scales the response is dominated by the 2PT and b2 terms, but quickly the 1h terms start to dominate. This switch between 2h and 1h terms appears at $\ell \sim 650$, i.e., earlier than the switch in C_ℓ which appears at $\ell \sim 800$. This happens because the nonlinear

Table C.1. Fits to the redshift dependence of the response of the galaxy power spectrum.

z	$\overline{R}$	R_0	R_1
0.1	4.06	5.00	−0.97
0.2	4.48	5.29	−0.83
0.3	4.79	5.24	−0.47
0.4	4.98	5.10	−0.13
0.5	5.08	4.95	0.14
0.6	5.12	4.78	0.34
0.7	5.10	4.62	0.50
0.8	5.06	4.45	0.63
0.9	4.98	4.29	0.72
1.0	4.89	4.13	0.78
1.1	4.79	4.00	0.82
1.2	4.70	3.91	0.82
1.3	4.63	3.86	0.80
1.4	4.59	3.85	0.77
1.5	4.61	3.89	0.75
1.6	4.70	3.97	0.76
1.7	4.87	4.09	0.80
1.8	5.13	4.26	0.90
1.9	5.49	4.46	1.06

Notes. Column 1 gives the start of the redshift bin, of width $\Delta z = 0.1$. Column 2 gives values for the scale independent model $R_\ell = \overline{R}$, while Cols. 3 and 4 give values for the linear model $R_\ell = R_0 + R_1 \times (\ell/\ell_0)$ with $\ell_0 = 1000$.

part of the power spectrum reacts more strongly to background change than the linear part : $\frac{\partial \ln P_{1h}}{\partial \delta_b} > \frac{\partial \ln P_{2h}}{\partial \delta_b}$.

The total angular response shows some scale dependence over the range of multipoles considered, ranging from $R_\ell \sim 4.2$ on large scales to $R_\ell \sim 5.5$ on small scales. In this article for simplicity we took a constant effective value $R_\ell = 5$ (dashed line in Fig. C.1). This has the advantage of allowing analytical calculations in Sect. 3.4, and we find in Sect. 4 that it reproduces adequately the S/N over all scales considered, and also reproduces the Fisher constraints on cosmological parameters, except deep in the SSC-dominated regime for the most affected parameters.

To go beyond this constant R_ℓ approximation, we need the scale dependence of the response. The redshift dependence is also needed, if we want to work on redshifts other than the one studied in this article ($0.9 < z < 1$). To answer both of these problems, we have computed numerically the full response through Eqs. (C.1) and (C.5) on scales $50 < \ell < 2000$ on a wide range of redshift ($0.1 < z < 2$ in bins $\Delta z = 0.1$). We then fitted the responses in each redshift bin either with a constant model $R_\ell = \overline{R}$ or a linear model $R_\ell = R_0 + R_1 \times (\ell/\ell_0)$ with $\ell_0 = 1000$. The values of the fitted parameters are given in Table C.1.

Another numerical approach that we anticipate is to calibrate the response through dedicated simulations, similarly to the work of Barreira et al. (2018b).

Appendix D: Impact of binning

Indicating bins of multipoles by the b subscript, the binned angular power spectrum, C_b, is defined as

$$C_b = \sum_{\ell \in b} \frac{S_{b\ell}}{\Delta_b} \times C_\ell, \tag{D.1}$$

where the summation is over multipoles within the bin b, Δ_b is the width of the bin, and $S_{b\ell}$ is a reshaping operator usually chosen to flatten the C_ℓ within bins of multipoles[11]. This reshaping operator is thus obtained assuming that within the bin b, the angular power spectrum is approximately given by $C_\ell \simeq \frac{C_b}{S_{b\ell}}$ with C_b a constant over the bin, and $S_{b\ell}$ a (usually theoretically) known function of ℓ[12]. Finally, the widths of the bins are usually chosen to be greater than the typical length in multipoles of the ℓ-to-ℓ' coupling induced by the mask.

The covariance of the binned spectrum, C_b, is related to the covariance of the full spectrum, C_ℓ, as

$$\mathrm{Cov}\,(C_b, C_{b'}) = \sum_{\ell \in b} \sum_{\ell' \in b'} \left(\frac{S_{b\ell}}{\Delta_b}\right)\left(\frac{S_{b'\ell'}}{\Delta'_b}\right) \mathrm{Cov}\,(C_\ell, C_{\ell'}). \tag{D.2}$$

By writing $\mathrm{Cov}\,(C_\ell, C_{\ell'}) = C_{\mathrm{G}} + C_{\mathrm{SSC}}$, the covariance of the binned spectra is then given by the sum of its Gaussian contribution and its super-sample contribution.

Choosing bins that are wider than the typical width of the mask-induced couplings leads to a diagonal Gaussian covariance $(C_{\mathrm{G}})_{b,b'} \simeq G_b\,\delta_{b,b'}$. An analytic expression for G_b can be obtained assuming the f_{sky} approximation, i.e.,

$$G_b = \sum_{\ell \in b} \left(\frac{S_{b\ell}}{\Delta_b}\right)^2 \times \left(\frac{2C_\ell^2}{(2\ell + 1)\,f_{\mathrm{sky}}}\right). \tag{D.3}$$

Since the reshaping function is chosen such that $S_{b\ell}C_\ell$ is roughly constant, we can simplify the above to get $G_b = \frac{2C_{\ell_b}^2}{(2\ell_b+1)\,f_{\mathrm{sky}}\,\Delta_b}$ by defining the average multipole in the bin ℓ_b with the identification $\frac{1}{(2\ell_b+1)\Delta_b} \equiv \sum_{\ell \in b} \frac{1}{(2\ell+1)\Delta_b^2}$.

For the SSC, we first recall that for a single probe and a single redshift bin, the covariance of the spectra is given by

$$[C_{\mathrm{SSC}}]_{\ell,\ell'} = S_{i,i}V_\ell V_{\ell'}, \tag{D.4}$$

with $V_\ell = R_\ell C_\ell$. It is then straightforward to show that for the binned spectra, we obtain

$$[C_{\mathrm{SSC}}]_{b,b'} = S_{i,i}V_b V_{b'}, \tag{D.5}$$

with the binned version of the vector V_ℓ, i.e.,

$$V_b = \sum_{\ell \in b} \frac{S_{b\ell}}{\Delta_b} \times V_\ell. \tag{D.6}$$

In the case where the response $R_\ell \equiv R$ is constant, this simplifies to $V_b = RC_b$. This shows that for a single probe and a single bin in redshift, adding the SSC to covariance of binned power spectra still corresponds to a rank 1 update of the Gaussian covariance.

The above is easily generalized to the other cases where it is enlarged to multi-probes and more than one redshift bin since it is exactly the same binning in multipoles which has to be used for the entire set of multi-probe and multi-redshift auto- and cross-spectra. The data vector is now built from the multi-probe binned angular power spectra, and the vector V_b in the SSC is obtained by binning the vector V_ℓ. Only the Gaussian covariance is slightly amended, being partitioned into $n_b \times n_b$ non-diagonal blocks of size $n_c \times n_c$ (n_b is the number of bins and

---

[11] This is easily related to the binning operator, $P_{b\ell}$, commonly used in the CMB context through $P_{b\ell} = S_{b\ell}/\Delta_b$.
[12] Typically $\ell(\ell + 1)$ in the CMB context.

n_c the number of auto- and cross-spectra). Using the f_{sky} approximation, it becomes block diagonal $C_G = G_b \, \delta_{b,b'}$. The blocks of size $n_c \times n_c$ read

$$
\begin{aligned}
[G_b]^{W,X;YZ} &\equiv \mathrm{Cov}_G \left(C_b^{WX}, C_b^{YZ} \right) \\
&= \sum_{\ell \in b} \left(\frac{S_{b\ell}}{\Delta_b} \right)^2 \times \left(\frac{C_\ell^{WY} C_\ell^{XZ} + C_\ell^{WZ} C_\ell^{XY}}{(2\ell + 1) f_{sky}} \right),
\end{aligned}
\tag{D.7}
$$

where (W, X, Y, Z) run over probes.

Appendix E: Likelihood of cluster counts

The purpose of this section is to recall the form of the full likelihood of cluster counts, a result which seems overlooked in the literature. Further, we extend the likelihood with the formulation developed in Sect. 5.2, which will ensure it to be well-defined analytically.

We call $N = (N_{i_M,i_z})_{i_M,i_z}$ the vector of cluster counts in all bins of mass (indexed by i_M) and redshift (indexed by i_z). Then without super-sample covariance (simply called sample variance in the cluster literature), the likelihood is a collection of independent Poisson distribution in each bin of mass and redshift,

$$
P(N|\mathbf{p}) = \prod_{i_M,i_z} \mathrm{Poiss}(N_{i_M,i_z}|\overline{N}_{i_M,i_z}(\mathbf{p})),
\tag{E.1}
$$

where $\overline{N}_{i_M,i_z}(\mathbf{p})$ is the model prediction for parameters $\mathbf{p}$.

The above likelihood is sufficient to describe small counts, i.e., at high mass. However for current and future surveys detecting an increasing number of clusters, it becomes necessary to account for the effect of sample variance (Hu & Kravtsov 2003). Lima & Hu (2004) found the full likelihood for cluster counts in different cells, which can be straightforwardly applied to our case with only one cell (the survey),

$$
P(N|\mathbf{p}) = \int \mathrm{d}^n \tilde{N} \left(\prod_{i_M,i_z} \mathrm{Poiss}(N_{i_M,i_z}|\tilde{N}_{i_M,i_z}) \right) \mathrm{Gauss}(\tilde{N} - \overline{N}, S'),
\tag{E.2}
$$

where S' is similar to the S matrix defined in Eq. (5) in the case of cluster counts, but also including the counts response, and was defined originally for a 3D survey neglecting redshift evolution (Lima & Hu 2004):

$$
S'_{i_M,i_z;j_M,j_z} = b_{i_M,i_z} N_{i_M,i_z} \, b_{j_M,j_z} N_{j_M,j_z} \times S_{i_z;j_z}.
\tag{E.3}
$$

The matrix $S_{i_z;j_z}$ reads

$$
S_{i_z;j_z} = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \, \tilde{W}_{i_z}^*(\mathbf{k}) \, \tilde{W}_{j_z}(\mathbf{k}) \, P(k),
\tag{E.4}
$$

where $\tilde{W}_{i_z}$ is the normalized ($\int \mathrm{d}^3 x \, \tilde{W}_{i_z}(\mathbf{x}) = 1$) window function in redshift bin i_z.

In the framework developed in Sect. 5.2, $\tilde{N}$ is interpreted as the average number counts in a region of the universe having a background change δ_b. Noting that the response of the cluster counts is $\frac{\partial N_{i_M,i_z}}{\partial \delta_b} = b_{i_M,i_z} N_{i_M,i_z}$, we can rewrite the likelihood as

$$
\begin{aligned}
P(N|\mathbf{p}) &= \int \mathrm{d}^n \delta N \left(\prod_{i_M,i_z} \mathrm{Poiss}\left(N_{i_M,i_z} | \overline{N}_{i_M,i_z} + \delta N_{i_M,i_z} \right) \right) \\
&\quad \times \mathrm{Gauss}\left(\delta N, \frac{\partial N}{\partial \delta_b}^T S \frac{\partial N}{\partial \delta_b} \right)
\end{aligned}
\tag{E.5}
$$

$$
\begin{aligned}
&= \int \mathrm{d}^{n_z} \delta_b \left(\prod_{i_M,i_z} \mathrm{Poiss}\left(N_{i_M,i_z} | \overline{N}_{i_M,i_z} + \frac{\partial N_{i_M,i_z}}{\partial \delta_b} \delta_b(i_z) \right) \right) \\
&\quad \times \mathrm{Gauss}\left(\delta_b, S \right).
\end{aligned}
\tag{E.6}
$$

Rigorously, we might be concerned about the edges of the integral in this likelihood: it does not make physical sense for δ_b to go to $-\infty$, as it corresponds to $\delta N \to -\infty$, i.e., $\tilde{N}$ becoming negative, which is impossible for a number of objects. In practice, this is unlikely to be a concern since for any reasonably sized survey the background change follows $\delta_b \ll 1$ at all redshifts, hence $\tilde{N} > 0$.

For the purpose of rigorousness, let us solve this physical concern nonetheless. When δ_b becomes of order 1, two approximations fail: (i) the pdf of δ_b being Gaussian, which is incorrect since for instance $\delta_b \geq -1$, and (ii) using a linear response ansatz $\tilde{N} = \overline{N} + \frac{\partial N}{\partial \delta_b} \delta_b$. Both failures can be cured formally with

$$
P(N|\mathbf{p}) = \int \mathrm{d}^{n_z} \delta_b \left(\prod_{i_M,i_z} \mathrm{Poiss}\left(N_{i_M,i_z} | \tilde{N}_{i_M,i_z}(\mathbf{p}, \delta_b) \right) \right) \times P(\delta_b|\mathbf{p}),
\tag{E.7}
$$

where $P(\delta_b|\mathbf{p})$ is the pdf of the background change (which has support on $\delta_b \in [-1, \infty[$) and $\tilde{N}_{i_M,i_z}(\mathbf{p}, \delta_b)$ is the average cluster count in a region with background change δ_b of a universe with cosmological parameters $\mathbf{p}$. For instance in the separate universe approach this $\tilde{N}_{i_M,i_z}(\mathbf{p}, \delta_b)$ could be computed thanks to a change of cosmological parameters $\mathbf{p}'(\mathbf{p}, \delta_b)$.