Impact of survey geometry and super-sample covariance on future photometric galaxy surveys

S. Gouyou Beauchamps, F. Lacasa, I. Tutusaus, M. Aubert, P. Baratta, A. Gorce, Z. Sakr

A&A 659, A128 (2022)
https://doi.org/10.1051/0004-6361/202142052

Astronomy & Astrophysics

Impact of survey geometry and super-sample covariance on future photometric galaxy surveys

S. Gouyou Beauchamps[1], F. Lacasa[2], I. Tutusaus[3,4,5], M. Aubert[1,6], P. Baratta[1], A. Gorce[2,7,8], and Z. Sakr[5,9]

[1] Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France
e-mail: gouyou@cppm.in2p3.fr
[2] Université Paris-Saclay, CNRS, Institut d'astrophysique spatiale, 91405 Orsay, France
[3] Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
[4] Institut d'Estudis Espacials de Catalunya (IEEC), Carrer Gran Capità 2-4, 08034 Barcelona, Spain
[5] Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, 31400 Toulouse, France
[6] Univ Lyon, Univ Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, 69622 Villeurbanne, France
[7] Department of Physics, Blackett Laboratory, Imperial College, London SW7 2AZ, UK
[8] Department of Physics and McGill Space Institute, McGill University, Montreal, QC H3A 2T8, Canada
[9] Université St Joseph; UR EGFEM, Faculty of Sciences, Beirut, Lebanon

Received 19 August 2021 / Accepted 30 November 2021

ABSTRACT

Photometric galaxy surveys probe the late-time Universe where the density field is highly non-Gaussian. A consequence is the emergence of the super-sample covariance (SSC), a non-Gaussian covariance term that is sensitive to fluctuations on scales larger than the survey window. In this work, we study the impact of the survey geometry on the SSC and, subsequently, on cosmological parameter inference. We devise a fast SSC approximation that accounts for the survey geometry and compare its performance to the common approximation of rescaling the results by the fraction of the sky covered by the survey, $f_{\rm SKY}$, dubbed 'full-sky approximation'. To gauge the impact of our new SSC recipe, that we call 'partial-sky', we perform Fisher forecasts on the parameters of the (w_0, w_a)-CDM model in a 3×2 point analysis, varying the survey area, the geometry of the mask, and the galaxy distribution inside our redshift bins. The differences in the marginalised forecast errors –with the full-sky approximation performing poorly for small survey areas but excellently for stage-IV-like areas– are found to be absorbed by the marginalisation on galaxy bias nuisance parameters. For large survey areas, the unmarginalised errors are underestimated by about 10% for all probes considered. This is a hint that, even for stage-IV-like surveys, the partial-sky method introduced in this work will be necessary if tight priors are applied on these nuisance parameters. We make the partial-sky method public with a new release of the public code PySSC.

Key words. large-scale structure of Universe

1. Introduction

The large-scale structure (LSS) of the Universe is an excellent probe of cosmology, giving constraints, for example, on dark matter, dark energy, and the large-scale behaviour of gravity. Current galaxy surveys, such as the Kilo-Degree Survey (KiDS, Heymans et al. 2021) and the Dark Energy Survey (DES, DES Collaboration 2022), are starting to provide cosmological constraints competitive with those derived from the primordial cosmic microwave background (CMB) and its weak lensing (Planck Collaboration VI 2020). Additionally, the next generation of stage-IV surveys, like the *Vera C. Rubin* Observatory Legacy Survey of Space and Time (LSST, LSST Science Collaboration 2009), *Euclid* (Laureijs et al. 2011), and the Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration 2016), will make it possible to discriminate between dark energy and modified gravity models with unprecedented precision and as a result their observations will shed light on the origin of cosmic acceleration.

However, one key challenge when studying the late-time LSS is its non-Gaussian distribution, which results from the non-linear dynamics governing its time-evolution. This non-Gaussianity has a variety of consequences: at the level of the observables, information escapes the two-point correlation function to leak into higher orders (e.g. Carron 2012; Obreschkow et al. 2013); at the level of the likelihood, this latter becomes non-Gaussian although the impact on cosmological parameters appears to be weak (Lin et al. 2020; Upham et al. 2021); and lastly at the level of the covariance, which is the main focus of this article. Non-Gaussianity typically leads to an enhanced covariance, as the tail of extreme events becomes broader. For instance, the covariance of two-point functions receives contributions from a positive trispectrum (e.g. Lacasa 2018; Wadekar & Scoccimarro 2020, for galaxy clustering).

One of the non-Gaussian contributions to the covariance originates from long-wavelength super-sample modes, which shift the mean matter density inside the survey and coherently modulate all observables. This effect is often called sample variance in galaxy cluster analyses and super-sample covariance (SSC) when considering galaxy clustering and weak lensing. First discovered for cluster counts (Hu & Kravtsov 2003), a vast amount of literature has been devoted to the SSC in recent years (e.g. Takada & Hu 2013; Takada & Spergel 2014; Takahashi et al. 2014; Li et al. 2018;

Chan et al. 2018; Lacasa & Kunz 2017; Lacasa et al. 2018; Barreira et al. 2018a,b), and it is well known to have a large impact on constraints on the dark energy equation of state for future surveys, be it for clusters (Hu & Kravtsov 2003), weak lensing (Barreira et al. 2018a), or photometric galaxy clustering Lacasa (2020).

Given the fact that the SSC originates from the observation of a limited portion of the Universe, the angular mask of the survey, accounting for the unobserved regions of the sky (because of bright stars or the high luminosity of the Galactic plane for example), should be taken into account in the computation of the SSC. While the effect of the mask has been extensively studied for Gaussian covariance, especially for CMB analysis (Hivon et al. 2002; Efstathiou 2004), there has been no detailed study of whether or not it should be accounted for when including the SSC for future surveys. This is the aim of the present article.

In this work, we quantify and explore the impact of the SSC on the performance of future survey missions, and aim to constrain cosmology – and, in particular, dark energy. To model the SSC, we use the approximation presented in Lacasa & Grain (2019), extend it to account for partial-sky coverage, and compare it to the often-used flat-sky approximation. To forecast the impact on the inference power of the survey, we build on the *Euclid* forecast efforts of Euclid Collaboration (2020, hereafter EC-B2020).

The article is organised as follows: in Sect. 2, we compare three methods allowing to predict the SSC for full-sky, partial-sky, and flat-sky surveys. In Sect. 3, we describe our galaxy survey forecast methodology. In Sect. 4, we present the impact of including the SSC on the statistical power of various survey setups in terms of signal-to-noise and Fisher forecasts. We conclude in Sect. 5.

2. Method: SSC in full, partial, and flat sky

Let us consider two observables O_1 and O_2. These can be written as the integral over the line of sight of their density, respectively o_1 and o_2: $O_i = \int \mathrm{d}V\, o_i$, where $\mathrm{d}V = r^2(z)\frac{\mathrm{d}r}{\mathrm{d}z}\mathrm{d}z$ is the comoving volume per steradian and $r(z)$ the comoving distance. The SSC for these observables is then given by the general formula (Lacasa & Grain 2019):

$$\mathrm{Cov}_{\mathrm{SSC}}\left(O_1, O_2\right) = \iint \mathrm{d}V_1 \mathrm{d}V_2 \frac{\partial o_1}{\partial \delta_b}\left(z_1\right) \frac{\partial o_2}{\partial \delta_b}\left(z_2\right) \sigma^2\left(z_1, z_2\right), \quad (1)$$

where the quantity $\partial o_1/\partial\delta_b(z_1)$ describes how o_1 varies with changes in the background density δ_b.

The (co)variance of the background density is defined as

$$\sigma^2(z_1, z_2) = \langle \delta_b(z_1)\delta_b(z_2)\rangle = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \tilde{\mathcal{W}}\left(\mathbf{k}, z_1\right) \tilde{\mathcal{W}}^*\left(\mathbf{k}, z_2\right) P_m(k|z_{12}), \quad (2)$$

where $\tilde{\mathcal{W}}$ is the Fourier transform of the survey window function, whose expression will depend on whether we are looking at full- or partial-sky coverage. $P_m(k|z_{12}) = D(z_1)D(z_2)P_m(k|z=0)$ is the matter power spectrum at redshifts z_1 and z_2.

We take O_1 (resp. O_2) to be the angular power spectrum $C_{ij}^{AB}(\ell)$ cross-correlating two LSS tracers A and B (resp. C and D), typically galaxy clustering and galaxy shear. Each spectrum is measured for a redshift bin pair (respectively i j and k l) and this can be expressed using the Limber approximation as

$$C_{ij}^{AB}\left(\ell\right) = \int \mathrm{d}V\, W_i^A(z) W_j^B(z) P_{AB}\left(k_\ell|z\right), \quad (3)$$

where $P_{AB}(k_\ell|z)$ is the 3D power spectrum at $k_\ell \equiv (\ell + 1/2)/r(z)$ and $W_i^A(z)$ is the kernel of observable A corresponding to the redshift bin i. Subsequently, from Eq. (3), $o_{AB} = W_i^A(z)W_j^B(z)P_{AB}\left(k_\ell|z\right)$ (resp. o_{CD}), and if we assume that the derivatives $\partial o_i/\partial\delta_b$ vary slowly with redshift compared to $\sigma^2(z_1, z_2)$, we can rewrite Eq. (1) as

$$\mathrm{Cov}_{\mathrm{SSC}}(C_{ij}^{AB}(\ell), C_{kl}^{CD}(\ell)) \simeq \frac{\partial \bar{P}_{AB}}{\partial \delta_b}\left(z_1\right) \frac{\partial \bar{P}_{CD}}{\partial \delta_b}\left(z_2\right) \times \iint \mathrm{d}V_1\, \mathrm{d}V_2\, W_i^A(z_1) W_j^B(z_1) W_k^C(z_2) W_l^D(z_2)\, \sigma^2\left(z_1, z_2\right). \quad (4)$$

where we define

$$\frac{\partial \bar{P}_{AB}}{\partial \delta_b}\left(z\right) \equiv \frac{\int \mathrm{d}V\, W_i^A(z) W_j^B(z)\, \partial P_{AB}/\partial\delta_b\left(k_\ell|z\right)}{I^{AB}(i,j)}, \quad (5)$$

with $I^{AB}(i,j) \equiv \int \mathrm{d}V\, W_i^A(z)W_j^B(z)$. Let $R^{AB}(k)$ be the effective relative response of the considered power spectrum:

$$\frac{\partial P_{AB}}{\partial \delta_b}\left(k\right) \equiv R^{AB}(k)\, P_{AB}(k). \quad (6)$$

For the matter power spectrum, R is constant with redshift and can be computed from perturbation theory or estimated from simulations (Lacasa & Grain 2019). Then

$$\begin{aligned} \frac{\partial \bar{P}_{AB}}{\partial \delta_b}\left(z\right) \times I^{AB}(i,j) &= \int \mathrm{d}V\, W_i^A(z) W_j^B(z)\, \partial P_{AB}/\partial\delta_b\left(k_\ell|z\right) \\ &= \int \mathrm{d}V\, W_i^A(z) W_j^B(z)\, R^{AB}(k_\ell)\, P_{AB}(k_\ell, z) \\ &\equiv R_\ell^{AB}\, C_{ij}^{AB}(\ell). \end{aligned} \quad (7)$$

Finally, we define the matrix $S_{i,j;k,l}^{A,B;C,D}$, which is the dimensionless volume-averaged (co)variance of the background matter density contrast, by

$$S_{i,j;k,l}^{A,B;C,D} \equiv \int \mathrm{d}V_1 \mathrm{d}V_2\, \frac{W_i^A\left(z_1\right) W_j^B\left(z_1\right)}{I^{AB}\left(i,j\right)} \frac{W_k^C\left(z_2\right) W_l^D\left(z_2\right)}{I^{CD}\left(k,l\right)} \sigma^2\left(z_1, z_2\right), \quad (8)$$

and the covariance simply rewrites as:

$$\mathrm{Cov}_{\mathrm{SSC}}\left(C_{ij}^{AB}\left(\ell\right), C_{kl}^{CD}\left(\ell'\right)\right) \approx R_\ell^{AB} C_{ij}^{AB}\left(\ell\right) \times R_{\ell'}^{CD} C_{kl}^{CD}\left(\ell'\right) \times S_{i,j;k,l}^{A,B;C,D}. \quad (9)$$

We note that, here the **S** matrix has four indices, as it describes the covariance between all auto and cross-C_ℓ. We denote it S_{ijkl}, but for simplicity in the following we consider only the covariance between auto-C_ℓ, i.e. $\mathrm{Cov}_{\mathrm{SSC}}\left(C_{ii}^{AA}\left(\ell\right), C_{jj}^{BB}\left(\ell'\right)\right)$. In this case, the **S** matrix is denoted S_{ij}. We now derive the expression for this **S** matrix in different survey cases: full sky coverage (Sect. 2.1), partial sky coverage (Sect. 2.2) and flat sky approximation (Sect. 2.3).

2.1. Full-sky case

In the case of full sky coverage, the variance of the background density field is simply (Lacasa & Rosenfeld 2016)

$$\sigma^2(z_1, z_2) = \frac{1}{2\pi^2} \int k^2 \mathrm{d}k\, P_m(k|z_{12})\, j_0(kr_1)\, j_0(kr_2), \quad (10)$$

where r_i is the comoving distance to redshift z_i and j_0 is the spherical Bessel function of the first kind and order zero. Given that the angular matter power spectrum can be written as

$$C^m_{z_1,z_2}(\ell) = \frac{2}{\pi}\int k^2 \mathrm{d}k\, P_m(k|z_{12})\, j_\ell(kr_1)\, j_\ell(kr_2), \tag{11}$$

we can write σ^2 as its monopole:

$$\sigma^2 = \frac{1}{4\pi} C^m_{z_1,z_2}(\ell = 0). \tag{12}$$

By injecting this expression into Eq. (8), we can relate the **S** matrix to a power spectrum

$$S^{A,B;C,D}_{i,j;k,l} = \frac{1}{4\pi} C^{X,Y}(\ell = 0), \tag{13}$$

where

$$a^X_{\ell m} = \frac{\int \mathrm{d}V\, W^A_i(z) W^B_j(z)\, a^{\mathrm{matter}}_{\ell m}(z)}{\int \mathrm{d}V\, W^A_i(z) W^B_j(z)}, \tag{14}$$

$$a^Y_{\ell m} = \frac{\int \mathrm{d}V\, W^C_k(z) W^D_l(z)\, a^{\mathrm{matter}}_{\ell m}(z)}{\int \mathrm{d}V\, W^C_k(z) W^D_l(z)}, \tag{15}$$

so that

$$C^{X,Y}(\ell = 0) = \int \mathrm{d}V_1 \mathrm{d}V_2\, k^2 \mathrm{d}k\, \frac{W^A_i(z_1) W^B_j(z_1)}{I^{AB}(ij)} \frac{W^C_k(z_2) W^D_l(z_2)}{I^{CD}(k,l)} \times P_m(k|z_1, z_2)\, j_0(kr_1) j_0(kr_2). \tag{16}$$

The exquisite sensitivity of upcoming photometric surveys will lead to a shot-noise small enough for the SSC to be an important source of error on cosmological parameters. In particular, considering a full-sky SSC such as that outlined above will no longer be sufficient and the SSC associated with the limited size of the survey will need to be considered. Until now, authors have chosen the simple option of rescaling the full-sky covariance by a factor f^{-1}_{SKY} –where $f_{\mathrm{SKY}} \equiv \Omega_S/4\pi$ is the fraction of the sky covered by the survey and Ω_S its solid angle– in order to account for partial-sky coverage in the SSC (Lacasa et al. 2018). This latter is the first approximation we consider in the present work. Numerically, we use the Python implementation `PySSC` (Lacasa & Grain 2019)[1]

2.2. Partial-sky case

In this section, we derive the approach to SSC in the case of partial sky coverage. The coverage is represented by the angular survey window function or mask $\mathcal{W}(\hat{\mathbf{n}})$, where $\hat{\mathbf{n}}$ is the sky direction. Typically $\mathcal{W}(\hat{\mathbf{n}}) = 0$ for unobserved pixels and $\mathcal{W}(\hat{\mathbf{n}}) = 1$ for observed ones, though a non-binary mask is also possible, for example to represent the impact of inhomogeneous depth. In this article, we assume that the mask is the same at all considered redshifts. The covariance of the background mode is then given by (Lacasa et al. 2018)

$$\sigma^2(z_1, z_2) = \frac{1}{\Omega_S^2}\sum_\ell (2\ell+1)\, C^{\mathcal{W}}(\ell)\, C^m_{z_1,z_2}(\ell). \tag{17}$$

As for the full-sky case, we can see the S_{ijkl} matrix as a $C(\ell)$ of a non-physical field X, whose kernel is the product of the kernels $W^A W^B$. However, in this case, multipoles other than the monopole will contribute to the SSC:

$$S^{A,B;C,D}_{i,j;k,l} = \frac{1}{\Omega_S^2}\sum_\ell (2\ell+1) C^{X,Y}(\ell) C^{\mathcal{W}}(\ell). \tag{18}$$

It is interesting to note that, when considering the full-sky limit of Eq. (18) ($f_{\mathrm{SKY}} \to 1$ and $C^{\mathcal{W}}(\ell > 0) = 0$), we retrieve the full-sky matrix given in Eq. (13). Equation (18) is the second method for SSC estimation that we consider in this work.

Numerically, we have developed a Python implementation of Eq. (18), building on `PySSC`. We have validated this partial-sky implementation against a parallel implementation making use of the `AngPow` (Campagne et al. 2018) public code for angular power spectra. We present this validation in Appendix A, which shows a ∼6% agreement in the computation of the S_{ij} for surveys larger than ∼2.5% of the sky, while reaching ∼10% for smaller patches.

2.3. Flat-sky case

Another example of approximation used in the literature to simplify the computationally expensive estimation of the SSC is the flat-sky approximation. For a top-hat kernel of width δr in the flat-sky case, the $S_{i,j}$ matrix can indeed be simplified to (Hu & Kravtsov 2003; Lima & Hu 2007)

$$S_{i,j} = \frac{1}{2\pi^2}\int k_\perp \mathrm{d}k_\perp 4 \frac{J_1(k_\perp\theta_S r_1)}{k_\perp\theta_S r_1}\frac{J_1(k_\perp\theta_S r_2)}{k_\perp\theta_S r_2} \times \int \mathrm{d}k_\parallel j_0\left(\frac{k_\parallel \delta r_1}{2}\right) j_0\left(\frac{k_\parallel \delta r_2}{2}\right) \cos\left[k_\parallel (r_1 - r_2)\right] P_m(k \mid z_{12}), \tag{19}$$

for a cylindrical window function of radius θ_S delineating a survey solid angle $\Omega_S = 2\pi(1 - \cos\theta_S) \simeq \pi\theta_S^2$. The wave-vector $\boldsymbol{k} = (k_\parallel, k_\perp)$ is split into its components parallel and perpendicular to the line of sight. Here, J_1 is the Bessel function of the first kind and order one. The power spectrum $P_m(k \mid z_{12})$ is evaluated at the centre of the respective redshift bins, following Hu & Kravtsov (2003), Lima & Hu (2007).

This approximation, which is very efficient computationally, is the third and final approximation considered in this work. We use it as a point of comparison only, as it is limited to the case of top-hat kernels.

2.4. Method comparison

In this section, we compare the S_{ij} matrices obtained using the three different methods outlined above; the result of this comparison was used to choose a method on which to focus our interest for the remainder of the study.

We consider the S_{ij} matrices obtained for arbitrary top-hat kernels with the full-sky computation (Eq. (13)) rescaled by f^{-1}_{SKY}, the partial-sky computation (Eq. (18)), and with the flat-sky approximation (Eq. (19)) for circular masks of areas ranging from 1 deg^2 to 15 000 deg^2, as well as for ten non-overlapping redshift bins ranging from $z = 0.2$ to $z = 2.0$. The choice of top-hat kernels of width $\Delta z = 0.1$ is arbitrary as this example is intended for illustration purposes only. Overall, we find on one hand that the flat-sky approximation gives a S_{ij} matrix close to the true partial-sky computation for surveys with an area smaller than 5 deg^2 only. On the other hand, we find that the full-sky approximation, after rescaling, gives satisfying results for areas over 15 000 deg^2.

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

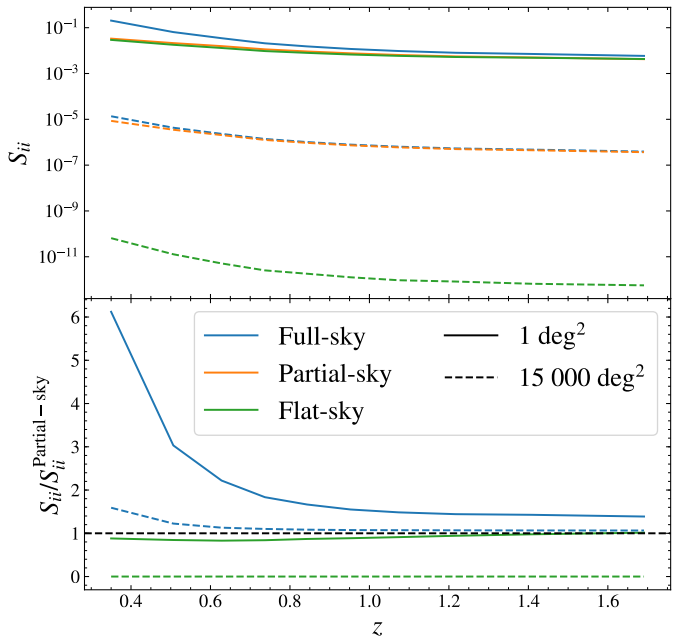


Fig. 1. Diagonal terms of the S_{ij} matrices obtained for surveys with an area of 1 deg^2 (plain lines) and 15 000 deg^2 (dashed lines). The colours correspond to the three different methods considered in this article: full sky in blue, partial sky in orange and flat sky in green. The lower panel shows the ratio to the partial-sky case.

In Fig. 1, we compare the diagonal terms of the S_{ij} matrices obtained for the 1 deg^2 and 15 000 deg^2 mask for the three treatments of the mask. We find that the flat-sky approximation performs better at higher redshift, with a relative difference below 10% for $z > 1.5$. For lower redshifts, the S_{ii} is underestimated by as much as 20% when using the flat-sky approximation, which corroborates the results of Lacasa et al. (2018). In particular, this approximation completely ignores cross-redshift bin correlations, which nevertheless exist when using the other two approaches. For survey masks larger than 5 deg^2, the flat-sky approximation underestimates the S_{ij} by as much as a factor two and by five orders of magnitude in the case of a wide 15 000 deg^2 mask, as can be seen in Fig. 1. This approximation should therefore not be considered in those cases. Conversely, the full-sky approach performs well for very wide survey masks. In this case, the full-sky approximation overestimates the diagonal of the S_{ij} by up to 1.5 for low redshifts ($z < 0.5$) but by less than 10% for $z > 1$.

Figure 2 shows the S_{ij} correlation matrices (i.e. divided by their diagonal elements) obtained for a circular survey mask of 15 000 deg^2. We can see that for non-overlapping redshift bins, the SSC results in anti-correlations, which decrease for distant bins. Such anti-correlations were already noted in the literature (Hu & Kravtsov 2003; Lacasa et al. 2018) and come from the fact that the matter correlation function becomes negative at large separations.

Therefore, we see that both approximations struggle to recover low-redshift correlations stemming from partial-sky SSC but can perform well for redshifts $z > 1$, where the correlations are weaker. Here, we only considered non-overlapping redshift bins, but we also compared the two matrices computed for overlapping bins. In that case, the correlation between the bins that overlap becomes positive and the structure of the matrix is more complex. This renders the interpretation of the impact of the partial-sky recipe on the SSC less evident if we only compare the S_{ij} matrices. For a more complete interpretation, we see in Sect. 4 how the differences between full sky and partial sky directly impact cosmological parameter inference. We do not consider the flat-sky approximation further because we are interested in large cosmological surveys, with wide survey areas. Instead, we compare results between the partial-sky computation, the full-sky approximation, and the Gaussian case in order to highlight contributions from the SSC to results.

3. Method: galaxy surveys forecast

In order to forecast the variation in constraining power of galaxy surveys depending on the covariance considered, we follow for the most part the forecast recipe presented in EC-B2020. We consider a Fisher matrix formalism and make use of the CosmoSIS[2] public code (Zuntz et al. 2015). In this section, we review the main aspects of the forecast and refer the reader to EC-B2020 for the remaining details.

3.1. Data

In these forecasts, we examine the constraining power of three cosmological probes: weak lensing (WL), photometric galaxy clustering (GCph), and their cross-correlation terms (XC), also known as galaxy-galaxy lensing. We refer to the full combination as GCph + WL + XC. We consider the tomographically binned projected angular power spectra as observables, $C_{ij}(\ell)$, where i, j label redshift pairs of tomographic bins. The angular spectra are presented in Eq. (3) using the Limber approximation. We use the same formalism for WL, GCph, and the XC terms. The main difference between the different probes appears through the different kernels used in the projection from the power spectrum of matter perturbations P_m to the spherical harmonic-space observable. Following EC-B2020, when computing the observables, we use the Limber and flat-sky approximations (Kitching et al. 2017; Kilbinger et al. 2017; Taylor et al. 2018), and we ignore reduced shear and magnification effects (Deshpande et al. 2020).

For the redshift distribution of galaxies, we follow EC-B2020 in considering ten tomographic redshift bins with the same number of galaxies in each bin. We assume a true underlying redshift distribution given by

$$n^{\text{true}}(z) \propto \left(\frac{z}{z_0}\right)^2 \exp\left[-\left(\frac{z}{z_0}\right)^{3/2}\right], \tag{20}$$

where $z_0 = 0.9/\sqrt{2}$. We then compute the photometric redshift distributions in each one of the bins by convolving the true distribution with a sum of two Gaussian distributions. One for the main dispersion of photometric redshift estimates and another one for the outliers. More specifically, the redshift distribution in the tomographic bin i is given by

$$n_i(z) = \frac{\int_{z_i^-}^{z_i^+} \mathrm{d}z_p n^{\text{true}}(z) p_{ph}(z_p|z)}{\int_{z_{\min}}^{z_{\max}} \mathrm{d}z \int_{z_i^-}^{z_i^+} \mathrm{d}z_p n^{\text{true}}(z) p_{ph}(z_p|z)}, \tag{21}$$

where (z_i^-, z_i^+) are the edges of the ith tomographic bin and set to the following values for the ten equally populated bins:

$$z_i = \{0.0010, 0.42, 0.56, 0.68, 0.79, 0.90, 1.02, 1.15, 1.32, 1.58, 2.50\}. \tag{22}$$

[2] Available at https://bitbucket.org/joezuntz/cosmosis/wiki/Home

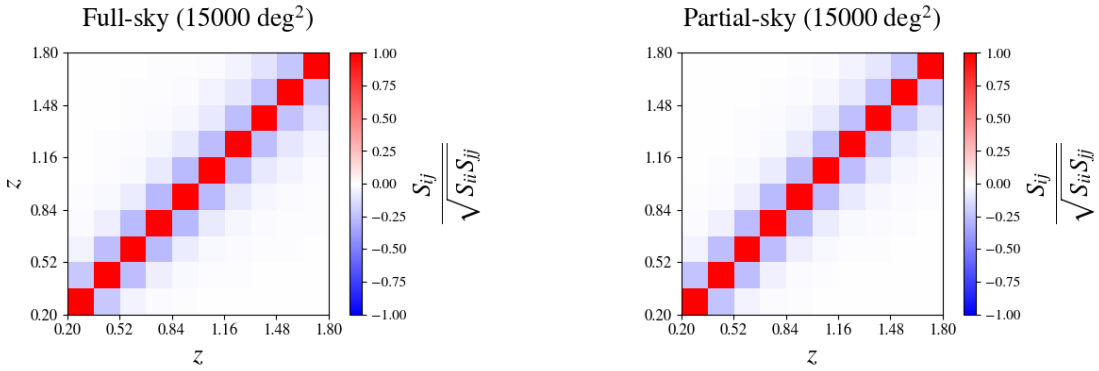


Fig. 2. S_{ij} correlation matrices obtained for a 15 000 deg^2 circular mask, and arbitrary, non-overlapping top-hat kernels for the full-sky computation of Eq. (13) (*left*) and the partial-sky computation of Eq. (18) (*right*).

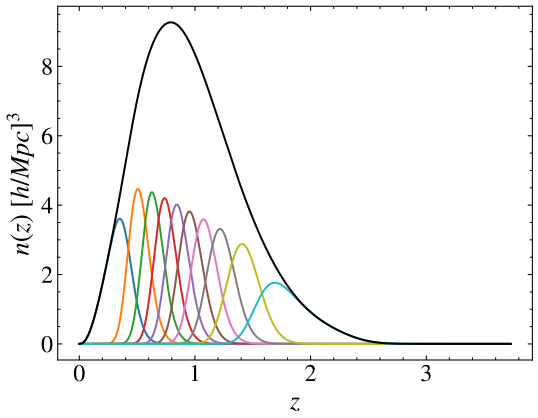


Fig. 3. Normalised galaxy number density distribution in the ten photometric redshift bins. The black line is the sum of all the redshift bins.

We follow EC-B2020 in parameterising the probability distribution function $p_{ph}(z_p|z)$ as

$$p_{ph}(z_p|z) = \frac{1-f_{\rm out}}{\sqrt{2\pi}\sigma_b(1+z)}\exp\left\{-\frac{1}{2}\left[\frac{z-c_b z_p - z_b}{\sigma_b(1+z)}\right]^2\right\} + \frac{f_{\rm out}}{\sqrt{2\pi}\sigma_o(1+z)}\exp\left\{-\frac{1}{2}\left[\frac{z-c_o z_p - z_o}{\sigma_o(1+z)}\right]^2\right\}. \quad (23)$$

For the outliers, we set the multiplicative bias to $c_o = 1$ and the additive bias to $z_o = 0.1$. For the remainder of the galaxies, we consider a multiplicative bias $c_b = 1$ and an additive bias $z_b = 0$. The uncertainty on the redshifts is assumed to be $\sigma_b = \sigma_o = 0.05$. We consider a default fraction of outliers $f_{\rm out} = 0.1$. We further assume a galaxy number density of thirty galaxies per arcmin^2. We show in Fig. 3 the normalised galaxy distribution considered in this analysis.

The WL power spectra contain contributions from cosmic shear and the intrinsic alignment of galaxies. We assume these intrinsic alignments are caused by a change in galaxy ellipticities that is linear in the density field. In this case, we can express the density–intrinsic and intrinsic–intrinsic three-dimensional power spectra, P_{mI} and P_{II}, as a linear function of the density power spectrum, with $P_{mI} = -A(z)P_m$, and $P_{\rm II} = [-A(z)]^2 P_m$. We follow EC-B2020 in parameterising A as

$$A(z) = \frac{\mathcal{A}_{\rm IA}C_{\rm IA}\Omega_m\mathcal{F}_{\rm IA}(z)}{D(z)}, \quad (24)$$

where $C_{\rm IA} = 0.0134$ is a normalisation constant, $D(z)$ is the growth factor, and $\mathcal{A}_{\rm IA}$ controls the amplitude of the IA contribution. We further model the redshift dependence as

$$\mathcal{F}_{\rm IA} = (1+z)^{\eta_{\rm IA}}\left[\frac{\langle L\rangle(z)}{L_*(z)}\right]^{\beta_{\rm IA}}, \quad (25)$$

where $\langle L\rangle(z)/L_*(z)$ is the ratio between the mean source luminosity and the characteristic scale of the luminosity function (Hirata et al. 2007; Bridle & King 2007). Following EC-B2020, we consider the following fiducial values for the intrinsic alignment nuisance parameters: $\{\mathcal{A}_{\rm IA}, \eta_{\rm IA}, \beta_{\rm IA}\} = \{1.72, -0.41, 2.17\}$.

With respect to GCph, one of the primary sources of uncertainty is the relation between the galaxy distribution and the underlying matter distribution, that is the galaxy bias. We consider a linear galaxy bias where the galaxy distribution δ_g is proportional to the matter distribution δ_m,

$$\delta_g(z) = b(z)\delta_m(z), \quad (26)$$

and the galaxy bias b only depends on redshift. We note that a linear galaxy bias is sufficiently accurate to analyse large scales (DES Collaboration 2022), while non-linear galaxy bias models are needed for the very small scales (see e.g. Sánchez et al. 2017; Desjacques et al. 2018). For simplicity, and in order not to mix the impact of the SSC with a non-linear galaxy bias modelling, we use the linear galaxy bias approximation in the following.

More specifically, and according to the approach used in EC-B2020, we consider a linear galaxy bias with a constant amplitude in each true redshift bin, that is

$$b(z_i \le z < z_{i+1}) = b_i, \quad (27)$$

where z_i and z_{i+1} stand for the boundaries of the ith redshift bin in true redshift. We choose a fiducial for the ten galaxy bias nuisance parameters given by $b_i = \sqrt{1+\bar{z}_i}$, where $\bar{z}_i$ is the mean redshift value of each redshift bin in true redshift.

For the full analysis, taking into account the correlations between GCph and WL, we consider both a Gaussian covariance alone and its combination with the SSC. The Gaussian covariance, accounting for all correlations between angular scales, redshift combinations, and different observables, can be expressed as:

$$\mathrm{Cov}_{\rm G}\left[C_{ij}^{AB}(\ell), C_{kl}^{CD}(\ell')\right] = \frac{\delta^{\rm K}_{\ell\ell'}}{(2\ell+1)f_{\rm SKY}\Delta\ell}\left\{\left[C_{ik}^{AC}(\ell)+N_{ik}^{AC}(\ell)\right]\left[C_{jl}^{BD}(\ell')+N_{jl}^{BD}(\ell')\right] + \left[C_{il}^{AD}(\ell)+N_{il}^{AD}(\ell)\right]\left[C_{jk}^{BC}(\ell')+N_{jk}^{BC}(\ell')\right]\right\}, \quad (28)$$

where A, B, C, and D stand for WL and GCph, i, j, k, and l run over all tomographic bins, $\delta^{\rm K}_{\ell\ell'}$ represents the Kronecker delta of ℓ and ℓ', and $\Delta\ell$ stands for the width of the multipole bins. We assume that $\Delta\ell$ is large enough so that the $f_{\rm SKY}$ approximation is valid, as shown by Hivon et al. (2002). The noise terms $N^{AB}_{ij}(\ell)$ are given by $\sigma^2_\epsilon \delta^{\rm K}_{ij}/\bar{n}_i$ for WL, where the variance of observed ellipticities is σ^2_ϵ, and $\delta^{\rm K}_{ij}/\bar{n}_i$ for GCph. We assume that the Poisson errors on WL and GCph are uncorrelated, yielding a null noise for XC.

We consider in the following the optimistic scenario presented in EC-B2020 concerning the multipole cuts used in the analysis. That is, we include all multipoles ranging from $\ell = 10$ to $\ell = 5000$ for WL and all multipoles ranging from $\ell = 10$ to $\ell = 3000$ for GCph and the XC terms. We note that we consider this optimistic case where we enter deep into the non-linear regime to study the impact of the SSC, where it is most relevant.

3.2. Cosmological models

When studying the impact of a partial-sky approach in the SSC, we consider a spatially flat Universe with cold dark matter and dark energy. We use the standard CPL parameterisation for the dark energy equation of state (Chevallier & Polarski 2001; Linder 2005):

$$w(z) = w_0 + w_a \frac{z}{1+z}. \tag{29}$$

In addition to the w_0 and w_a parameters describing dark energy, the cosmological model is described by the total matter density today, Ω_m, the dimensionless Hubble constant, h, the baryon density today, Ω_b, the slope of the primordial power spectrum, n_s, and the root-mean-square (RMS) of matter fluctuations on spheres of $8\,h^{-1}$ Mpc radius, σ_8. We further assume dark energy to be a minimally coupled scalar field with sound speed equal to the speed of light and no anisotropic stress. We therefore neglect dark energy fluctuations and use the parametrised post-Friedmann (PPF) framework (Hu & Sawicki 2007), which allows the dark energy equation of state to cross $w(z) = -1$ without developing instabilities in the perturbation sector.

We consider the following set as fiducial values for our cosmological parameters:

$$\begin{aligned} \boldsymbol{p} &= \{\Omega_m,\, \Omega_b,\, w_0,\, w_a,\, h,\, n_s,\, \sigma_8,\} \\ &= \{0.32,\, 0.05,\, -1.0,\, 0.0,\, 0.67,\, 0.96,\, 0.816\}, \end{aligned} \tag{30}$$

and we fix the sum of neutrino masses to $\sum m_\nu = 0.06\,\mathrm{eV}$, with one massive neutrino and two massless neutrinos.

3.3. Derived forecast quantities

To gauge the impact of the SSC computation on the forecast statistical power of the survey, we use two metrics. First, the signal-to-noise ratio (S/N) of the angular power spectrum of a given probe, which quantifies the strength of detection of the angular power spectrum in a model-independent way:

$$(S/N)^2 = \sum_{i,j,k,l} \sum_{\ell,\ell'} C^{AB}_{ij}(\ell) \mathrm{Cov}\left[C^{AB}_{ij}(\ell), C^{CD}_{kl}(\ell')\right]^{-1} C^{CD}_{kl}(\ell'), \tag{31}$$

where Cov is the (total) covariance matrix of the power spectrum, consisting of the sum of the Gaussian and SSC contributions: $\mathrm{Cov} = \mathrm{Cov_G} + \mathrm{Cov_{SSC}}$.

To quantify the impact on cosmological constraints, we use a second metric, the Fisher matrix:

$$F_{\alpha,\beta} = \sum_{i,j,k,l} \sum_{\ell,\ell'} \frac{\partial C^{AB}_{ij}(\ell)}{\partial\theta_\alpha} \mathrm{Cov}\left[C^{AB}_{ij}(\ell), C^{CD}_{kl}(\ell')\right]^{-1} \frac{C^{CD}_{kl}(\ell')}{\partial\theta_\beta}, \tag{32}$$

where θ_α and θ_β are two model parameters, such as the two parameters of the dark energy equation of state. From the Fisher matrix we can derive several quantities to quantify the constraints on cosmological parameters :

- The marginalised error on a given parameter θ_α is given by

$$\sigma_\alpha = \sqrt{(F^{-1})_{\alpha,\alpha}}. \tag{33}$$

  This expression implies that all the other parameters are marginalised over, meaning that their variation is taken into account when estimating the error.
- Instead, one could consider the unmarginalised error

$$\sigma^U_\alpha = \sqrt{1/F_{\alpha,\alpha}}, \tag{34}$$

  which is like effectively fixing the other parameters to their fiducial values. We see in the following that this distinction is important to understand how the difference in S/N between full-sky and partial-sky translates to the Fisher forecast.
- By considering the marginalised Fisher submatrix, F_{w_0,w_a}, of the Dark Energy parameter plane (w_0, w_a), we can define the Dark Energy Figure of Merit as

$$\mathrm{FoM}_{w_0,w_a} = \sqrt{\det(F_{w_0,w_a})}. \tag{35}$$

  This quantity is proportional to the inverse of the area delimited by the 2σ contour in the marginalised two-parameter plane.

Finally, it is interesting to note that the S/N is a particular case of the Fisher metric for a scaling parameter A defined as $C(\ell; A) = A \times C(\ell)^{\rm fid}$, where $C(\ell)^{\rm fid}$ is the angular power spectrum computed in the fiducial cosmology.

4. Results: impact on the statistical power

In this section, we compute the S/N and Fisher matrix, for the Gaussian and total (Gaussian + SSC) covariance. We focus on varying the survey masks (and $f_{\rm SKY}$) for the partial-sky (and full-sky) computation of the SSC in order to assess the impact of survey geometry on parameter inference. First, we vary the size of the survey, and then we consider different mask geometries for a fixed survey area before finally studying the dependence of the SSC computations on the survey $n(z)$. In order to simplify the notation, we denote the full-sky and partial-sky computation as fsky and psky, respectively.

4.1. Survey area

In this section, we study the impact of the SSC and its implementation through the fsky and psky recipes with respect to the survey area. For the psky computation we consider a circular mask with a $f_{\rm SKY}$ corresponding to an area ranging from 5 to $15\,000\,\mathrm{deg}^2$. We also gauge the importance of the SSC by comparing it to the Gaussian-only covariance.

First, we look at how the S/N of the angular power spectrum of the probes considered in this article, that is, WL, GCph, and GCph+WL+XC, evolves with the size of the survey. As stated in section 2, for overlapping redshift bins, the structure of the S_{ijkl}

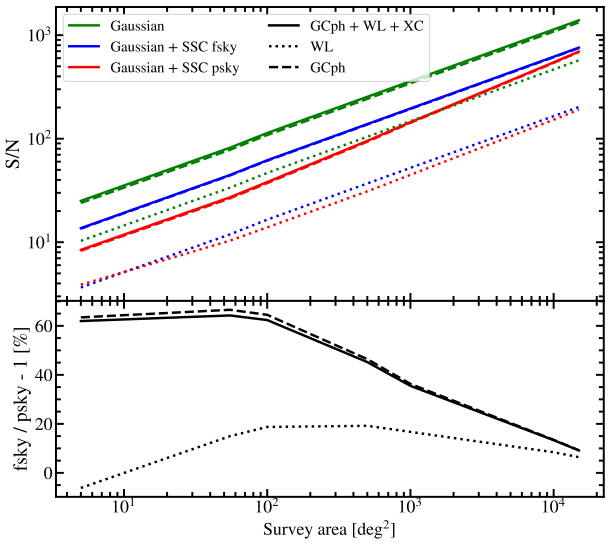


Fig. 4. *Top:* S/N of the angular power spectrum as a function of the survey area for WL (short dashes), GCph (long dashes), and GCph+WL+XC (plain line). The $f_{\rm SKY}$ approximation is shown in blue, the partial sky computation which accounts for the mask is shown in red, and the Gaussian case is shown in green. *Bottom:* relative difference in percent between the blue and the red lines.

matrix is more complex, and so to interpret the impact of fsky and psky on our observables, it is better to compare their S/N. We show the results of this comparison in Fig. 4, for the psky derivation, the fsky approximation, and the Gaussian covariance case. We see that all probes are significantly affected by the SSC regardless of the size of the survey. However, the relative difference between fsky and psky depends on the probe considered and the survey area. For WL, and for a small survey of 5 deg^2, the fsky approximation underestimates the S/N with respect to that of the more accurate psky computation by less than 10%. For larger surveys, fsky systematically overestimates the S/N, with a maximum of 20% for a 100 deg^2 survey and a minimum of 5% for a 15 000 deg^2 survey. For GCph+WL+XC, the behaviour of the S/N is similar to the GCph-only case, because the GCph dominates the signal with respect to WL. For these two probes, fsky leads to an overestimate of the S/N regardless of the survey area. The relative difference is maximum –with a 60% overestimation– for small surveys between 5 and 100 deg^2 and reaches a minimum of ∼7% for 15 000 deg^2, close to the WL case.

To understand how these results translate in terms of parameter constraints, we perform a Fisher forecast analysis following Sect. 3. Table 1 gathers the values of the Dark Energy FoM for all probes, considering the Gaussian covariance and the total Gaussian+SSC covariance in fsky and psky for survey areas of 5 and 15 000 deg^2. As expected from the S/N, the SSC has a large impact on the Dark Energy FoM, especially for WL and GCph+WL+XC, for which it is reduced by half for both small and large survey areas. For small survey areas, the fsky approximation underestimates the FoM, especially for auto-correlations of probes, which are the most impacted by the SSC. However, for the largest survey areas, the difference between fsky and psky is almost negligible.

To further understand these results, we present in Fig. 5 the marginalised and unmarginalised constraints on all cosmological parameters and survey areas of between 5 deg^2 and 15 000 deg^2. In this figure, we consider the Gaussian+SSC covariance using the psky derivation or the fsky approximation for different probes. In the unmarginalised case, the error bars resulting from using the fsky approximation or the psky computation follow the same evolution with the survey area as the S/N. That is, for all probes, when the fsky approximation leads to an overestimated S/N, it also results in underestimated error bars on cosmological parameters, and vice versa. Interestingly, when marginalising on all the varied parameters, that is cosmological and nuisance parameters, these results change. For GCph (see Fig 5a), the difference between fsky and psky is largely reduced, giving a relative difference of between 1 and −2% for all cosmological parameters and survey areas. For WL (Fig 5b), the marginalisation has the opposite effect and the error is overestimated with fsky, especially for the smallest surveys, going up to 50% increase for Ω_m, 35% for σ_8, and 20% for w_0, in the case of a 5 deg^2 survey. The relative difference is below 10% for surveys larger than 100 deg^2, and is close to zero for the largest areas. For GCph+WL+XC, Fig 5c, the situation is the same as for WL, with a smaller amplitude of the relative difference for Ω_m, σ_8, and w_0, except for w_a which shows a small underestimation of its error of ∼−2% on intermediate survey areas. An interesting point to note is that the most impacted cosmological parameters are the ones related to the amplitude of the power spectrum, namely Ω_m, σ_8, w_0, and w_a. This result is expected because the effect of the SSC appears on the amplitude of the power spectrum through the change in the background density δ_b.

Overall, we see that, for the marginalised constraints, the complete treatment of the mask in the derivation of the SSC is not necessary for large areas representative of upcoming stage-IV cosmological surveys. However, marginalising has an important effect on the impact of the SSC. Despite the fact that the S/Ns resulting from the fsky or psky computations largely differ, for GCph, the difference is seemingly absorbed in the nuisance parameters through marginalisation. In contrast, with WL, when marginalising, the difference is transferred to the cosmological parameters.

To confirm this interpretation, we look at the forecasted errors on all nuisance parameters, that is, the ten constant galaxy biases b_i and the three intrinsic alignment parameters $\mathcal{A}_{\rm IA}$, $\eta_{\rm IA}$ and $\beta_{\rm IA}$, when accounting for the SSC with fsky or psky. The results are shown in Fig. 6 for GCph+WL+XC. The galaxy biases show different constraints depending on the SSC recipe, with a maximum negative relative difference of between −2 and −5% for all b_i, except for b_2 which goes down to −10%. On the other hand, the errors on the intrinsic alignment parameters do not change whichever recipe is used.

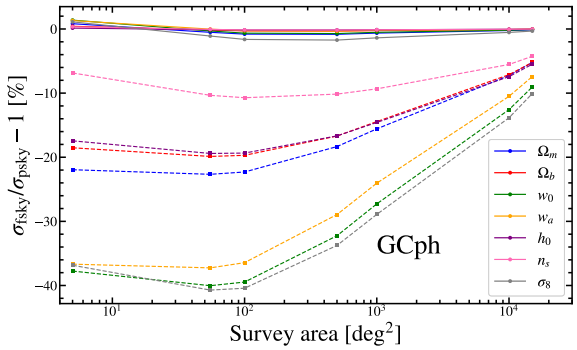
This is due to the fact that the simple model we use for the galaxy bias is just an amplitude factor on the power spectrum. Thus, through the marginalisation, the b_i nuisance parameters absorb the effect of SSC, which also modulates the amplitude of the power spectrum through the variation of the background density δ_b. The difference of SSC between fsky and psky is then mostly transferred to the linear galaxy bias, while the IA parameters are insensitive to SSC. We observe the same behaviour for GCph and WL alone. Interestingly, Wadekar et al. (2020) find a similar effect of the marginalisation when accounting for the full non-Gaussian covariance in a spectroscopic GC analysis.

The above results can be summarised in two important points. First, the S/N is a misleading metric when used to quantify the impact of a correct psky treatment of the SSC. Second, even if the difference in terms of marginalised cosmological constraints is close to zero between the fsky and psky approaches for large survey areas, for unmarginalised constraints the difference can be of the order of 10%.

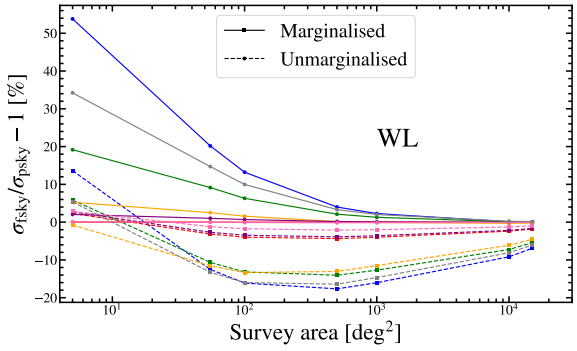
Table 1. Dark Energy Figure of Merit for all three probes when considering the Gaussian covariance or the full Gaussian+SSC covariance in fsky and psky.

Probe	Survey area [deg^2]	Gaussian	Gaussian + fsky SSC	Gaussian + psky SSC	fsky/psky − 1 [%]
WL	5	0.014	0.009	0.012	−30.87
	15 000	43.120	26.329	26.335	−0.02
GCph	5	0.035	0.029	0.029	−1.56
	15 000	103.714	88.636	88.455	0.20
GCph+WL+XC	5	0.346	0.150	0.166	−9.58
	15 000	1038.132	454.590	460.510	1.30

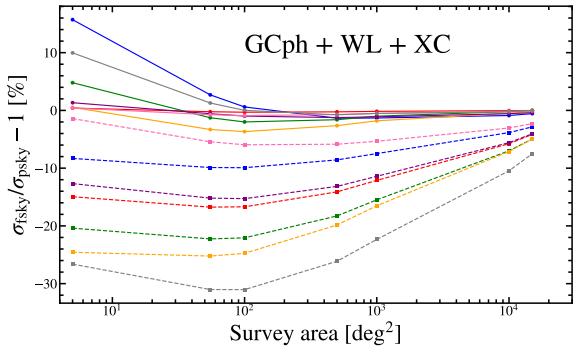
Notes. We show the results for the lowest and largest survey area.



(a)



(b)



(c)

Fig. 5. Forecast errors on cosmological parameters for the Gaussian+SSC covariance, using circular masks of different areas for GCph (*a*), WL (*b*), and GCph+WL+XC (*c*). For the three panels, we show the relative difference in % between fsky and psky for each cosmological parameter. The plain lines correspond to the constraints when marginalising on all parameters (cosmological and nuisance) and the dashed lines when there is no marginalisation.

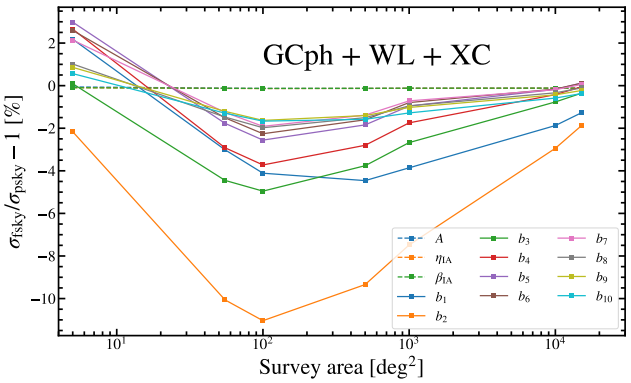


Fig. 6. Marginalised errors on nuisance parameters using circular masks of different areas for GCph+WL+XC. We show the relative difference as a percentage between fsky and psky for each nuisance parameter. The nuisance parameters associated to WL are shown in dashed lines and the ones associated with GCph in plain lines.

As the difference can be absorbed by the nuisance parameters, accounting for the full geometry of the mask when computing the SSC will be essential when tight priors on nuisance parameters are included in the analysis.

4.2. Survey geometry

In this section, we study the impact of the two different SSC recipes with respect to the geometry of the survey. We consider a survey with an area of 15 000 deg^2 and three different geometries: A single circular patch (such as the ones used in Sect. 4.1) dubbed '1 pole', two separated circular patches referred to as '2 poles', and a survey with a geometry close to future stage-IV surveys such as *Euclid*, where the Galactic and zodiacal plans have been removed. These three masks are represented in a Mollweide view in Fig. 7.

We perform a Fisher forecast in the same setting as described in Sect. 3 for the three masks considered. In Fig 8, we present the resulting marginalised and unmarginalised constraints in the psky and fsky cases for GCph+WL+XC. For the marginalised constraints (Fig 8a), as we already saw in the previous section, the relative difference between fsky and psky is close to zero for the simplest mask geometry. A more complex mask geometry leads to larger discrepancies between the two approaches, which nevertheless remain smaller than or close to 1% for all cosmological parameters. In the unmarginalised case (Fig 8b), we can see the same effect as was discussed in the previous section: the relative difference is larger when we do not marginalise.

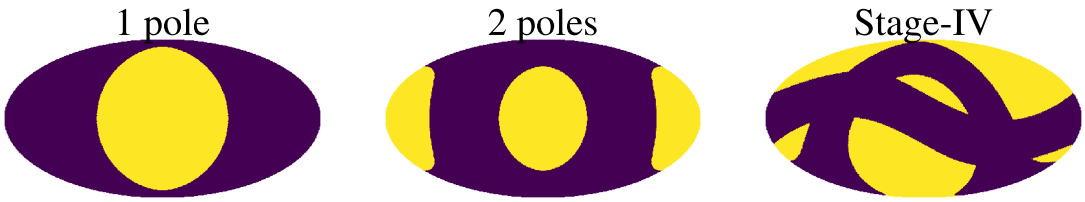


Fig. 7. Mollweide view of the three masks considered in Sect. 4.2. They all have an area of 15 000 deg^2, corresponding to $f_{\rm SKY} = 0.364$. The yellow area is the observable region.

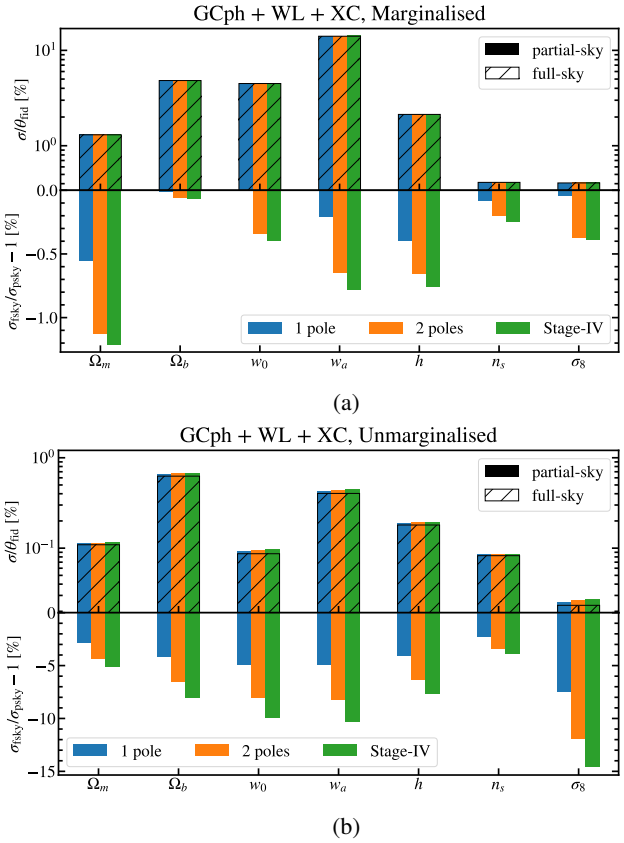


Fig. 8. Forecast errors on cosmological parameters for different mask geometry and a 15 000 deg^2 survey, for GCph+WL+XC, when marginalising (a) or not (b). *Top*: relative error as a percentage for each cosmological parameter. The filled coloured bars represent the constraints obtained with the psky derivation, using a circular mask (blue), a mask divided into two circles (orange), or a stage-IV-like mask (green). The hatched empty bars are the ones obtained in the fsky approximation. *Bottom*: relative difference between the fsky and the psky standard deviation for each mask geometry.

Additionally, regarding the impact of the geometry, we observe the same trend as in the marginalised constraints: a difference increases with the complexity of the mask. For the most complex stage-IV mask, the fsky approximation underestimates the error by almost 10% for w_0, w_a, and 15% for σ_8, in contrast to a 5% difference observed with the 1 pole mask. Similar results are obtained with GCph and WL alone.

Therefore, for large, stage-IV-like survey areas, the marginalised errors do not depend strongly on the mask geometry. However, as discussed in Sect. 4.1, this result will not hold when adding tight priors on nuisance parameters, which is equivalent to the unmarginalised errors case.

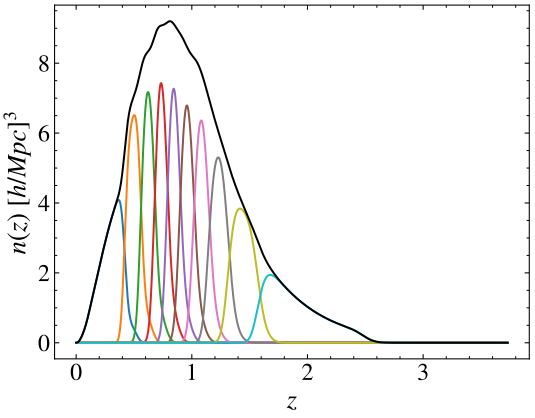


Fig. 9. Normalised galaxy number density distribution in the ten photometric redshift bins in the 'tight' case. The black line is the sum of all the redshift bins.

4.3. Survey galaxy distribution

In this section, we study the impact of the two different SSC recipes, comparing the forecasts obtained for different galaxy distributions $n(z)$. Indeed, as the S_{ijkl} matrix is computed from integrals of the kernels over z, it depends on the input $n(z)$. We use the forecast results from the previous sections and compare them to constraints obtained when changing some of the parameters of the assumed galaxy distribution (see Eqs. (20), (21) and (23)). We consider the same $n(z)$ as above and a new $n(z)$ with 'tighter' redshift bins. The new $n(z)$, dubbed 'tight', is shown in Fig. 9, whereas the original $n(z)$, dubbed 'wide', can be seen in Fig. 3. In addition, we also consider a third $n(z)$, which is similar to the wide one but with a different outlier fraction $f_{\rm out} = 0.25$, while it was 0.1 in the previous section. Indeed, the fraction of outlier redshifts is one of the most important issues to handle for photometric surveys. The wide $n(z)$ with $f_{\rm out} = 0.25$ is not shown as it is visually very similar to the case with $f_{\rm out} = 0.1$, but the interested reader can find a table in Appendix B with the values of the different parameters used to compute the three $n(z)$ considered in this section. Here, we use the 1 pole, 15 000 deg^2 mask introduced in Sect. 4.2 and visible in the left panel of Fig. 7 in all three cases.

We perform a Fisher forecast in the same setting as described in Sect. 4.1 for the three $n(z)$ considered. Figure 10 presents the resulting marginalised and unmarginalised constraints in the psky and fsky cases for GCph+WL+XC. Looking at the top panels of Figs. 10a,b, we see that for all parameters, the relative error is larger when using the wide $n(z)$. This can be related to the amount of overlap in the different $n(z)$ functions: for equal galaxy number densities, less overlap means less correlation between redshifts, and in turn more independent information in each bin. In addition, we see in Sect. 2 that for the pairs of bins that do not overlap, the SSC produces anti-correlations,

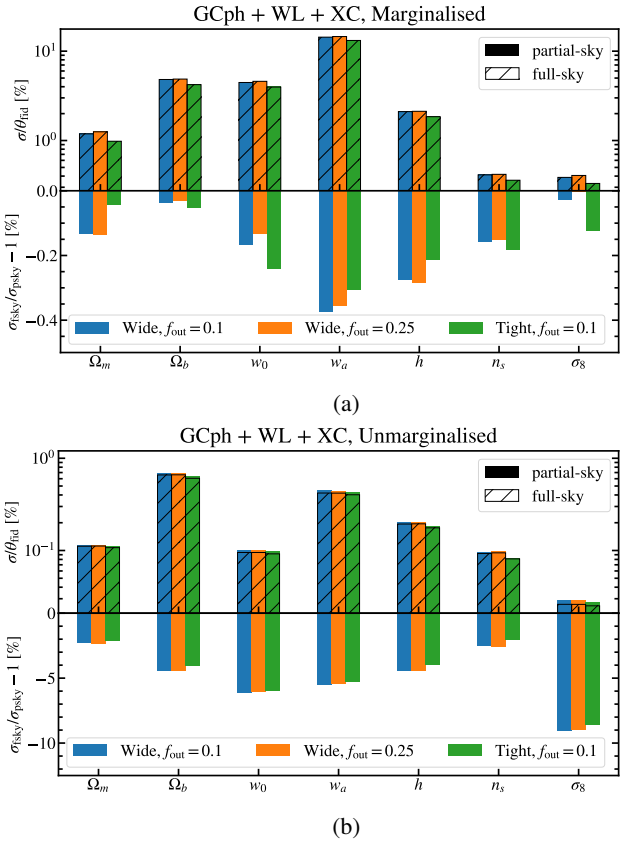


Fig. 10. Forecast errors on cosmological parameters for three $n(z)$: with wide redshift bins and an outlier fraction of $f_{out} = 0.1$ or $f_{mout} = 0.25$ and with tight redshift bins. The results are shown for GCph+WL+XC, when marginalising (a) or not (b). We suppose a circular 15 000 deg^2 survey in the three cases. *Top*: relative error as a percentage for each cosmological parameter. The filled coloured bars represent the constraints obtained with the psky derivation using the wide $n(z)$ with $f_{out} = 0.1$ (blue) or $f_{out} = 0.25$ (orange) and the tight $n(z)$ (green). The hatched empty bars are the ones obtained in the fsky approximation. *Bottom*: relative difference between the fsky and the psky standard deviation for each galaxy distribution.

which tends to increase the S/N compared to positive correlations. Increasing the outlier fraction to $f_{out} = 0.25$ does not seem to have a large impact on the constraints.

For the marginalised constraints (Fig. 8a), the difference between fsky and psky is less than 1% regardless of the $n(z)$ considered. Given the low relative difference, we do not find trends specific to a particular galaxy distribution. The relative difference can be larger for the wide or the tight case depending on the parameter considered. We find the same result for WL and GCph individually. However, for the unmarginalised constraints (Fig 10b), the difference between the two SSC recipes is slightly larger in the wide case than in the tight case. This could be explained by the larger error bars obtained with the former.

5. Conclusion

Many works have shown that the SSC can have a large impact on cosmological constraints coming from a variety of standard cosmological probes (Hu & Kravtsov 2003; Barreira et al. 2018a; Lacasa 2020). Most studies focused on describing the SSC using the approximation of flat-sky (Krause et al. 2017; Krause & Eifler 2017) or full-sky (Lacasa & Rosenfeld 2016; Lacasa & Grain 2019). A few more recent studies (Lacasa et al. 2018; Barreira et al. 2018a; Friedrich et al. 2021) investigated the more realistic case of a survey mask. In the present study, we extended the fast SSC approximation introduced in Lacasa & Grain (2019) to an accurate partial-sky treatment of the mask. This new SSC recipe retains the advantage of Lacasa & Grain (2019): it is fast to compute, mostly model-independent, and the code will be made available publicly when this article is published. Here, we use it to investigate the impact of the mask geometry on the SSC, and, in turn, on cosmological parameter inference for future surveys.

Comparing the standard flat-sky and full-sky approximations to our new, more accurate partial-sky computation, we find first of all that the flat-sky approximation only provides a satisfying estimate of SSC-induced correlations between redshift bins for survey areas smaller than 5 deg^2, and specifically for $z > 1$ (Sect. 2.4). On the other hand, the full-sky approach can recover the partial-sky SSC for wide surveys larger than 15 000 deg^2 with 10% precision, and, as expected, performs better at higher redshifts. Following these first results, we chose to focus our analysis (Sect. 4) on the full-sky and partial-sky methods to explore the impact of survey geometry on cosmological constraints. We considered the main photometric probes of the upcoming stage-IV surveys, the so-called 3x2pts analysis, composed of photometric galaxy clustering, weak lensing, and their cross-correlation (also called galaxy–galaxy lensing). We performed Fisher forecasts following the methodology of EC-B2020; see Sect. 3.

First, in Sect. 4.1, we consider a simple circular geometry for the mask and vary its area. Starting from a naive S/N comparison, we see that accounting for the SSC greatly reduces the S/N for all probes compared to the Gaussian covariance-only case. This result holds for the full-sky approximation, despite it overestimating the S/N, especially for small survey areas. Moving to parameter inference, we confirm that accounting for the SSC reduces the Dark Energy FoM for WL and GCph+WL+XC by half. Regarding the comparison between full-sky and partial-sky SSC, interpretation of the impact on cosmological constraints is more challenging. The unmarginalised errors are in agreement with the S/N: the full-sky approximation leads to underestimation of the error bars (because of overestimation of the S/N). After marginalisation, the difference in the errors induced by the two approximations is reduced, except for WL alone. That is because the large difference observed in the S/N is absorbed by the galaxy biases of GCph, while the intrinsic alignment parameters of WL are insensitive to the change in the SSC recipe. Overall, for both marginalised and unmarginalised constraints, the fsky approximation performs better as we increase the survey area, with a relative difference in cosmological errors of less than 10 % with psky for a 15 000 deg^2 survey.

Comparing surveys of areas of 15 000 deg^2 but with different footprints, we find that the more complex the mask geometry, the larger the difference between full-sky and partial-sky (see Sect. 4.2). For the marginalised constraints this difference is negligible as it is always below 1%. However, the difference cannot be neglected for unmarginalised errors, as for some parameters, including w_a, it ranges from 5% for the simplest mask to 10% for the most complex one. Finally, in Sect. 4.3, we compare constraints obtained with a circular 15 000 deg^2 mask with those obtained with the same mask but with a different galaxy number density distribution $n(z)$. For a $n(z)$ that has less overlap between redshift bins, the difference between full-sky and partial-sky in the unmarginalised constraints is found to decrease. That is because, for the same galaxy number density,

an $n(z)$ with less overlap between redshift bins leads to smaller error bars. However, the difference is small and the changes in the $n(z)$ we consider does not have a significant impact on the difference between full-sky and partial-sky for all probes.

Overall, these results show that, for wide surveys, such as the future stage-IV surveys and in particular *Euclid*, a complete treatment of the angular mask geometry when estimating the SSC is not crucial, and the full-sky approximation is sufficient. However, as also reported in Wadekar et al. (2020), marginalisation over the nuisance parameters can absorb the effect of non-Gaussian covariance, and if tight priors are applied to these parameters, the effect of the mask may no longer be negligible. As the partial-sky method we present in this article is made public[3], it can be used for further analysis to accurately account for the SSC in any cosmological survey.

Acknowledgements. S. G. B., F. L., I. T., M. A., P. B. and A. G. thank the organisers of the *Euclid* France 2019 summer school (`https://ecole-euclid.cnrs.fr`), where this project was initiated. F.L. was partially supported by a postdoctoral grant from Centre National d'Études Spatiales (CNES). I. T. acknowledges support from the Spanish Ministry of Science, Innovation and Universities through grant ESP2017-89838, and the H2020 programme of the European Commission through grant 776247. A. G. acknowledges financial support from the European Research Council under ERC grant number StG-638743 ("FIRSTDAWN") and from the Trottier Chair in Astrophysics, through the McGill Astrophysics Fellowship, as well as from the Canadian Institute for Advanced Research (CIFAR) Azrieli Global Scholars program and the Canada 150 Programme. This work also acknowledges support from the OCEVU LABEX (Grant No. ANR-11-LABX-0060) and the A*MIDEX project (Grant No. ANR-11-IDEX-0001-02) funded by the Investissements d'Avenir french government program managed by the ANR. This research made use of `astropy`, a community-developed core Python package for astronomy (Astropy Collaboration 2013, 2018); of `matplotlib`, a Python library for publication quality graphics (Hunter 2007); of `scipy`, a Python-based ecosystem of open-source software for mathematics, science, and engineering (Jones et al. 2001) – including `numpy` (Oliphant 2006); of `class`, a Boltzmann solver code for cosmology (Blas et al. 2011) and of the `Healpix` package (Górski et al. 2005).

References

Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33
Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, AJ, 156, 123
Barreira, A., Krause, E., & Schmidt, F. 2018a, JCAP, 10, 053
Barreira, A., Krause, E., & Schmidt, F. 2018b, JCAP, 6, 015
Blas, D., Lesgourgues, J., & Tram, T. 2011, JCAP, 07, 034
Bridle, S., & King, L. 2007, New J. Phys., 9, 444
Campagne, J.-E., Neveu, J., & Plaszczynski, S. 2018, Astrophysics Source Code Library [record ascl:1807.012]
Campagne, J. E., Neveu, J., & Plaszczynski, S. 2017, A&A, 602, A72
Carron, J. 2012, Phys. Rev. Lett., 108, 071301
Chan, K. C., Moradinezhad Dizgah, A., & Noreña, J. 2018, Phys. Rev. D, 97, 043532
Chevallier, M., & Polarski, D. 2001, Int. J. Mod. Phys. D, 10, 213
DES Collaboration (Abbott, T. M. C., et al.) 2022, Phys. Rev. D, 105, 023520
Deshpande, A. C., Kitching, T. D., Cardone, V. F., et al. 2020, A&A, 636, A95
DESI Collaboration (Aghamousa, A., et al.) 2016, ArXiv e-prints [arXiv:1611.00036]
Desjacques, V., Jeong, D., & Schmidt, F. 2018, Phys. Rep., 733, 1
Efstathiou, G. 2004, MNRAS, 349, 603
Euclid Collaboration (Blanchard, A., et al.) 2020, A&A, 642, A191
Friedrich, O., Andrade-Oliveira, F., Camacho, H., et al. 2021, MNRAS, 508, 3125
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Heymans, C., Tröster, T., Asgari, M., et al. 2021, A&A, 646, A140
Hirata, C. M., Mandelbaum, R., Ishak, M., et al. 2007, MNRAS, 381, 1197
Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2
Hu, W., & Kravtsov, A. V. 2003, ApJ, 584, 702
Hu, W., & Sawicki, I. 2007, Phys. Rev. D, 76, 104043
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open source scientific tools for Python
Kilbinger, M., Heymans, C., Asgari, M., et al. 2017, MNRAS, 472, 2126
Kitching, T. D., Alsing, J., Heavens, A. F., et al. 2017, MNRAS, 469, 2737
Krause, E., & Eifler, T. 2017, MNRAS, 470, 2100
Krause, E., Eifler, T. F., Zuntz, J., et al. 2017, ArXiv e-prints [arXiv:1706.09359]
Lacasa, F. 2018, A&A, 615, A1
Lacasa, F. 2020, A&A, 634, A74
Lacasa, F., & Grain, J. 2019, A&A, 624, A61
Lacasa, F., & Kunz, M. 2017, A&A, 604, A104
Lacasa, F., & Rosenfeld, R. 2016, JCAP, 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]
Li, Y., Schmittfull, M., & Seljak, U. 2018, JCAP, 2, 022
Lima, M., & Hu, W. 2007, Phys. Rev. D, 76, 123013
Lin, C.-H., Harnois-Déraps, J., Eifler, T., et al. 2020, MNRAS, 499, 2977
Linder, E. V. 2005, Phys. Rev. D, 72, 043529
LSST Science Collaboration (Abell, P. A., et al.) 2009, ArXiv e-prints [arXiv:0912.0201]
Obreschkow, D., Power, C., Bruderer, M., & Bonvin, C. 2013, ApJ, 762, 115
Oliphant, T. 2006, NumPy: A guide to NumPy (USA: Trelgol Publishing)
Planck Collaboration VI. 2020, A&A, 641, A6
Sánchez, A. G., Scoccimarro, R., Crocce, M., et al. 2017, MNRAS, 464, 1640
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
Taylor, P. L., Kitching, T. D., McEwen, J. D., & Tram, T. 2018, Phys. Rev. D, 98, 023522
Upham, R. E., Brown, M. L., & Whittaker, L. 2021, MNRAS, 503, 1999
Wadekar, D., & Scoccimarro, R. 2020, Phys. Rev. D, 102, 123517
Wadekar, D., Ivanov, M. M., & Scoccimarro, R. 2020, Phys. Rev. D, 102, 123521
Zonca, A., Singer, L., Lenz, D., et al. 2019, J. Open Source Softw., 4, 1298
Zuntz, J., Paterno, M., Jennings, E., et al. 2015, Astron. Comput., 12, 45

[3] Available at `https://github.com/fabienlacasa/PySSC`

Appendix A: Partial-sky code validation

In this Appendix we validate the outputs of the partial-sky code `PySSC.Sij_partsky`[1] allowing the computation of the S_{ij} matrix (in the case of a single observable; see Sect. 2.2) for a sample of different masks and radial window function configurations. We compare these results with the corresponding numerical predictions of the `AngPow`[4] (Campagne et al. 2017, 2018) public software.

As pointed out in Eq. (18), the S_{ijkl} matrix can be seen as the evaluation of the auto- and cross-angular power spectrum between observables and redshift bins (only redshift bins for the S_{ij} matrix). `AngPow` is specifically designed to compute the integral Eq. (16) for arbitrary multipoles ℓ, as long as the input radial window function is transformed such that $W(z) \rightarrow W^2(z)$ for a similar result between the two methods (see the `AngPow` integral in Campagne et al. 2017).

For this exercise, we choose several mask settings accounting for different SSC contribution levels, including circular angular patches of respectively [1, 5, 55, 100, 500, 1000, 10 000] deg^2 (corresponding to a fraction of the sky $f_{\rm SKY} = [0, 002, 0.01, 0.13, 0.24, 1.21, 2.49, 24.35]\times10^{-2}$) and a stage-IV mask (as featured in Fig. 7) equivalent to a full-sky survey without the Galactic plane and the zodiacal (15 000 deg^2 or equivalently $f_{\rm SKY} = 0.364$). For each of these masks, we use several arbitrary Gaussian radial window functions centred at redshifts $z = [0.5, 0.75, 1, 1.25, 1.5]$ with a variance of 0.01, allowing them to overlap.

Both codes are run in the context of a linear ΛCDM cosmology ($h = 0.67$, $\Omega_b = 0.05$, $\Omega_{cdm} = 0.27$, $n_s = 0.96$, $A_s = 2.1265.10^{-9}$) with an input power spectrum computed at $z = 0$ using the `CLASS` code[5] (Blas et al. 2011) in the range $k \in [10^{-5}, 3\times10^{-1}]h$/Mpc. The required maximum ℓ in the computation of the C_ℓ for a convergence of the variance within 5% is computed by the `PySSC.find_lmax` function for both codes. Here $\ell_{max} = [243, 218, 115, 90, 51, 36, 11]$ for the respective circular patches and $\ell_{max} = 20$ for the stage-IV one. Moreover, the mask spectrum $C_\ell(\mathcal{M})$ (see Eq. 18) is numerically computed using the healpix code[6](Zonca et al. 2019; Górski et al. 2005).

More specifically, the `Angpow` settings that we use are a trapezoidal radial quadrature with associated `radial_order` parameter set at 350 to maximize the accuracy, while performing an exact computation of the C_ℓ (no Limber approximation). Moreover, in order to freely modify the used radial window function and verify the condition $W(z) \rightarrow W^2(z)$, we use the input option `userfile`. The other computational parameters are left as default. On the `PySSC` setting side, we compute the wave mode integration using the precision parameter `precision` = 12.

In Fig. A.1, we compare the different S_{ij} output matrices predicted by the two methods in the context of the eight angular masks previously introduced. A compatibility of better than ∼ 6% accuracy can be seen between the two C_ℓ computational methods for all masks larger than 1000 deg^2 (> 2.5% of the sky), while keeping a better than ∼ 10% accuracy for smaller patches. We finally note that for each tested geometry, the deviation is damped for increasing redshifts.

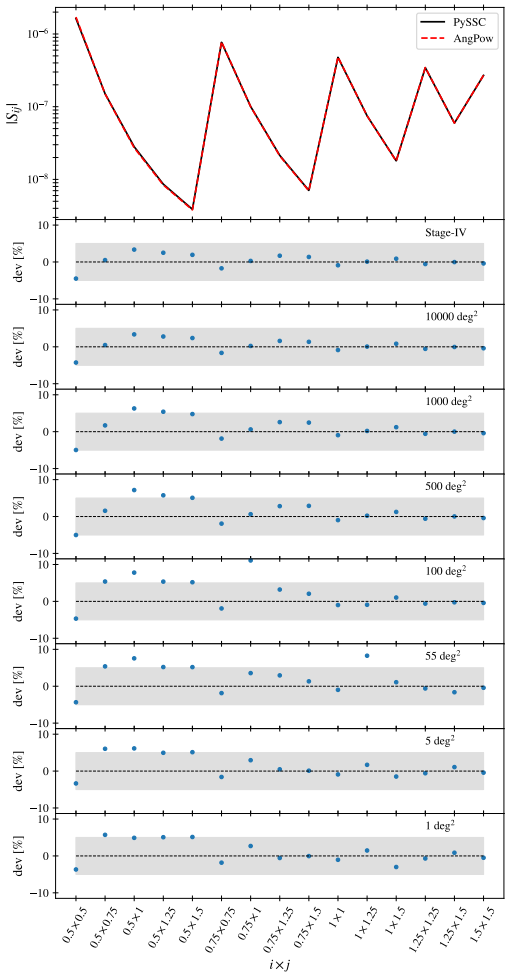


Fig. A.1. Top panel: Absolute value of the S_{ij} matrix elements in the case of a stage-IV angular mask, as predicted by `PySSC` in black solid line and `AngPow` in red dashed lines. The elements are ordered column by column of the lower half of the matrices, including the diagonal (i×i). The corresponding x-coordinates, $i \times j$, label the redshift bins that are cross-correlated. Panels 2 to 9: Relative deviations (in percent) between the two methods for decreasing survey area. The grey area depicts the 5% discrepancy.

Appendix B: Galaxy distribution parameters

In this Appendix, we show the values of the parameters of Eqs. (20), (21), and (23) that we used to compute the three $n(z)$ considered in Sect. 4.3.

Table B.1. Values of the parameters describing the photometric redshift distributions.

$n(z)$	c_b	z_b	σ_b	c_0	z_0	σ_0	f_{out}
Wide, $f_{out} = 0.1$	1.0	0.0	0.05	1.0	0.1	0.05	0.1
Wide, $f_{out} = 0.25$	1.0	0.0	0.05	1.0	0.1	0.05	0.25
Tight, $f_{out} = 0.1$	1.0	0.0	0.02	1.0	0.1	0.02	0.1

[4] gitlab.in2p3.fr/campagne/AngPow
[5] class-code.net
[6] healpix.sourceforge.net