Astronomy & Astrophysics

Super-sample covariance approximations and partial sky coverage

Fabien Lacasa[1], Marcos Lima[2], and Michel Aguena[2]

[1] Département de Physique Théorique, Université de Genève, 24 quai Ernest Ansermet, 1211 Geneva, Switzerland
 e-mail: fabien.lacasa@unige.ch
[2] Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, CEP 05314-970, São Paulo-SP, Brazil

ABSTRACT

Super-sample covariance (SSC) is the dominant source of statistical error on large scale structure (LSS) observables for both current and future galaxy surveys. In this work, we concentrate on the SSC of cluster counts, also known as sample variance, which is particularly useful for the self-calibration of the cluster observable-mass relation; our approach can similarly be applied to other observables, such as galaxy clustering and lensing shear. We first examined the accuracy of two analytical approximations proposed in the literature for the flat sky limit, finding that they are accurate at the 15% and 30–35% level, respectively, for covariances of counts in the same redshift bin. We then developed a harmonic expansion formalism that allows for the prediction of SSC in an arbitrary survey mask geometry, such as large sky areas of current and future surveys. We show analytically and numerically that this formalism recovers the full sky and flat sky limits present in the literature. We then present an efficient numerical implementation of the formalism, which allows fast and easy runs of covariance predictions when the survey mask is modified. We applied our method to a mask that is broadly similar to the Dark Energy Survey footprint, finding a non-negligible negative cross-z covariance, i.e. redshift bins are anti-correlated. We also examined the case of data removal from holes due to, for example bright stars, quality cuts, or systematic removals, and find that this does not have noticeable effects on the structure of the SSC matrix, only rescaling its amplitude by the effective survey area. These advances enable analytical covariances of LSS observables to be computed for current and future galaxy surveys, which cover large areas of the sky where the flat sky approximation fails.

Key words. large-scale structure of Universe – methods: analytical

1. Introduction

The large scale structure (LSS) of the Universe results from the growth of local density perturbations induced by gravitational collapse within an expanding background (e.g. Peebles 1980). The multi-point correlations and associated spectra that characterise this structure can be measured in real data and combined with theory predictions to constrain cosmological models, including gravity theories and relative amounts of dark matter and dark energy (Astier et al. 2006; Percival et al. 2010; Riess et al. 2009; Hildebrandt et al. 2017; Beutler et al. 2017; Hinton et al. 2017; DES Collaboration 2017). However, likelihood methods of parameter inference also require reliable estimates of the covariances associated with structure observables. Estimates of the covariances can be obtained from the actual data via methods such as jackknife and bootstrap, from simulations and Monte Carlo realisations, and from theoretical predictions that account for the known observational effects (Dodelson & Schneider 2013; Giannantonio et al. 2016; Crocce et al. 2016; O'Connell et al. 2016; Singh et al. 2017; Shirasaki et al. 2017; Blot et al. 2016; Escoffier et al. 2016; Pearson & Samushia 2016; Camacho et al. in prep.). In this work we focus on the latter approach.

Within the so-called halo model paradigm (Cooray & Sheth 2002), the LSS can be characterised by the statistical properties of the Universe building blocks, i.e. dark matter haloes. These haloes are characterised by their abundance, bias, and profiles, all of which can be studied from dark matter N-body simulations

and also from high-quality data sets. As galaxy clusters develop within dark matter haloes, they trace the highest density peaks. Their number counts and covariances are very sensitive probes of structure growth and expansion of the Universe (Lima & Hu 2004, 2005, 2007; Schmidt et al. 2009; Aguena & Lima 2016). A number of past and current surveys have detected clusters in multiple wavelengths with hopes to use such detections for cosmological purposes (Miller et al. 2005; Koester et al. 2007; Soares-Santos et al. 2011; Dietrich et al. 2014; Rykoff et al. 2014, 2016; Bleem et al. 2015; Planck Collaboration XXVII 2016; Bayliss et al. 2016).

As we consider scales close to the survey maximum size, the number of modes available decreases significantly, which represents an intrinsic source of uncertainty for the inferred structure properties. The cosmic variance quantifies these uncertainties and may have contributions from scales much larger than the survey itself, in which case we refer to it as super-sample covariance (SSC), whose effects have been studied recently in multiple contexts (e.g. Takada & Hu 2013; Li et al. 2014a,b, 2016; Takahashi et al. 2014; Shirasaki et al. 2017; Hu et al. 2016). SSC has been shown to be particularly important for probe combinations as it couples observables together (Takada & Bridle 2007; Takada & Spergel 2014; Krause & Eifler 2017; Lacasa & Rosenfeld 2016). Within a local patch, large scale modes change the effective average density, which can differ significantly and unpredictably from the true background density, affecting the inferred correlations. In fact, SSC may be the dominant source of errors in Jackknife covariance estimations (Shirasaki et al. 2017).

Finally, the survey geometry or footprint and its selection function, which is characterised by, for example masks and depth maps, also affect the estimation of correlations and covariances (Takahashi et al. 2014). Not only the geometry and selection must be known to good precision, but their properties must be properly propagated into the measured and predicted correlations and covariances. Some predictions can only be directly made under certain approximations, for example for full sky calculations, while for the more realistic case of partial sky coverage, further complicating assumptions need to be made. As we attempt to extract maximum information from observations, we may end up with a complicated survey mask containing holes, for example owing to saturated pixels, asteroids, contamination from stars, or simply pixels that do not satisfy a depth criterion. All these effects must be accounted for in a proper cosmological analysis.

In this article, we study the effect of partial sky coverage and arbitrary masks in the estimation of the SSC of cluster counts. We propose a method to estimate this covariance efficiently for an arbitrary mask, and compare this approach to a number of approximated calculations in the literature. We show that our general calculation reduces to the approximated computations in the appropriate limits, but differs from such computations in general. Although beyond the scope of this work, we expect our theoretical estimation can be compared to other methods of estimating the covariance, which also attempt to account for the survey geometry and mask effects such as with simulations. One of the advantages of our method, however, is that it allows for the covariance to be computed as a function of cosmological and nuisance parameters at each step of a Monte Carlo Markov chain, allowing, for example for self-calibration of cluster observable-mass distribution in general cosmological analyses of cluster samples (Lima & Hu 2005). It is also unbiased, contrary to internal covariance estimators (Lacasa & Kunz 2017). Finally, it is numerically much cheaper than running thousands of N-body simulations, the latter needing to be much larger than the survey to capture super-survey modes efficiently.

This article is organised as follows. In Sect. 2.2, we introduce the cluster counts covariance and the formalism for an exact SSC computation, and we then consider the case of full sky and flat sky (small angles) limit, providing some comparisons for the covariance kernels between those cases. In Sect. 3 we consider the flat sky limit under various approximations that have been proposed in the literature for computing the SCC and compare the results to the exact computation. In Sect. 4, we propose a method to numerically compute SSC in the case of partial sky coverage with an arbitrary survey mask. In Sect. 5 we present our results, applying the method to the case of a geometry similar to that of the Dark Energy Survey (DES), discussing the mask effects and recovering the flat sky limit. Finally, in Sect. 6 we present our conclusions and perspectives.

In all numerical computations, we take a cosmology for a flat ΛCDM Universe with parameter values $h = 0.67$, $\Omega_b h^2 = 0.022$, $\Omega_c h^2 = 0.12$, $w = -1$, $n_S = 0.96$, $\sigma_8 = 0.83$. Cluster counts are computed in two bins of redshift in the range $z \in [0.4, 0.6]$ with $\Delta z = 0.1$, and four bins of mass in the range $\log[M/(h^{-1} M_\odot) \in [14, 16]$ with $\Delta \log[M/(h^{-1} M_\odot)] = 0.5$. The halo mass function is from a fit to simulations from Tinker et al. (2008), the halo bias is from Tinker et al. (2010), and the linear matter power spectrum is from the transfer function by Eisenstein & Hu (1998). We set the following notational conventions: $r(z)$ is the comoving distance, that is $dr = c\, dz/H(z)$ with $H(z)$ the Hubble parameter, and $dV = r^2 dr$ is the comoving volume element per steradian. We use short cuts such as $dX_{12} = dX_1 dX_2$ and $P_m(k|z_{12}) = P_m(k|z_1, z_2) = G(z_1)G(z_2)P_m(k)$ with $G(z)$ the linear

growth function. Often the limits of redshift or mass integrals are implicitly those of the redshift or mass bin considered.

2. Covariance of large scale structure observables and super-sample covariance

2.1. Covariance of large scale structure observables

LSS observables have a covariance that can be composed of many different terms (e.g. Lacasa & Rosenfeld 2016; Lacasa 2018). The particular case of interest to this article is that of cluster counts, which can be viewed as a halo 1-point correlation function summed over the survey area.

We address the covariance as given by a decomposition in two specific terms: 1-halo and 2-halo

$$\mathrm{Cov}(N_{\mathrm{cl}}(i_M, i_z), N_{\mathrm{cl}}(j_M, j_z)) = \mathrm{Cov}^{1h} + \mathrm{Cov}^{2h}. \qquad (1)$$

These terms are denoted respectively as shot-noise variance and sample (co)variance in the literature (e.g. Hu & Kravtsov 2003).

The 1-halo or shot-noise variance term is given by,

$$\mathrm{Cov}^{\mathrm{shot}}(N_{\mathrm{cl}}(i_M, i_z), N_{\mathrm{cl}}(j_M, j_z)) = \frac{N_{\mathrm{cl}}(i_M, i_z)}{\Omega_S} \, \delta_{i_M, j_M} \, \delta_{i_z, j_z}, \qquad (2)$$

where Ω_S is the solid angle covered by the survey. In the following discussion, this term is accounted for when showing the total counts covariance. Its dependence on the survey footprint is a trivial scaling by Ω_S [1] and does not require any particular formalism to be predicted for a survey with large and/or complicated coverage.

We refer to the 2-halo term as SSC, which is described in more detail in Sect. 2.2.

For observables other than cluster number counts (for instance the galaxy angular power spectrum; see e.g. Lacasa & Rosenfeld 2016; Lacasa 2018) the covariance may contain many more terms, whose dependence on the survey geometry may be far from trivial. In this article, we restrict ourselves to studying the geometry dependence of SSC, as needed for cluster number counts analyses.

2.2. Super-sample covariance

SSC is a source of uncertainties for LSS observables coming from modes of size larger than the survey. The effect of these large scale modes on structure comes from the fact that the effective or local matter density $\widetilde{\rho}_m(z)$ averaged within the survey can be different from the true background density $\bar{\rho}_m(z)$ averaged over the whole Universe or an ensemble of Universes written as

$$\widetilde{\rho}_m(z) = \bar{\rho}_m(z)[1 + \delta_b(z)], \qquad (3)$$

where $\delta_b(z)$ is a background density perturbation induced by large scale modes. Equivalently, defining the density contrasts

$$\delta_m(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}_m}{\bar{\rho}_m}, \qquad (4)$$

---

[1] We note that the factor Ω_S appears because in our convention N_{cl} is a number of objects per steradian. In data analysis it is sometimes more convenient to use the absolute number of objects $\tilde{N}_{\mathrm{cl}} = N_{\mathrm{cl}} \times \Omega_S$, for which we have the simpler formula $\mathrm{Var}(\tilde{N}_{\mathrm{cl}}(i_M, i_z)) = \tilde{N}_{\mathrm{cl}}(i_M, i_z)$.

the ensemble average is $\langle \delta_m(\mathbf{x}) \rangle = 0$ by definition, however the spatial average over the survey is $\langle \delta_m(\mathbf{x}) \rangle_{\text{survey}} = \delta_b$.

We split the survey window function $W(\mathbf{x})$ into its radial $W_r(z)$ and angular $W(\hat{\mathbf{n}})$ pieces, i.e. $W(\mathbf{x}) = W_r(z)W(\hat{\mathbf{n}})$. The radial part simply specifies the redshift binning and we do not explicitly state this in our description below anymore because it is effectively included with implicit bounds on redshift integrals. On the other hand, the angular window $W(\hat{\mathbf{n}})$ is the main object whose effect on the covariances we want to consider. In principle $W(\hat{\mathbf{n}})$ may depend on redshift as well. In fact, this is very often the case for surveys with significant depth variations across the sky. For simplicity, we keep the angular window only as a function of the angular vector, although the formalism can be easily generalised (see Appendix A).

If the survey angular window $W(\hat{\mathbf{n}})$ subtends a solid angle Ω_S in the sky, and we denote the position vector $\mathbf{x} = r(z)\hat{\mathbf{n}}$, the background perturbation δ_b is given by

$$\delta_b(z) = \frac{1}{\Omega_S} \int \mathrm{d}^2\hat{\mathbf{n}}\, W(\hat{\mathbf{n}})\, \delta_m(r(z)\hat{\mathbf{n}}, z). \tag{5}$$

As a result, all LSS observables respond to such change of background density, becoming correlated. Explicitly, the SSC term of the cross-covariance between two observables O_1 and O_2 is given by (e.g. Lacasa & Rosenfeld 2016)

$$\text{Cov}^{\text{SSC}}(O_1, O_2) = \int \mathrm{d}V_{12}\, \frac{\partial \mathfrak{o}_1}{\partial \delta_b}\, \frac{\partial \mathfrak{o}_2}{\partial \delta_b}\, \sigma^2(z_1, z_2), \tag{6}$$

where $\mathfrak{o}_i$ is the density of observable O_i (per comoving volume $\mathrm{d}V$), $\partial \mathfrak{o}_i / \partial \delta_b$ is its reaction to the change in background δ_b, and $\sigma^2(z_1, z_2)$ is the covariance of δ_b, i.e.

$$\sigma^2(z_1, z_2) = \langle \delta_b(z_1)\, \delta_b(z_2) \rangle = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\, \tilde{W}(\mathbf{k}, z_1)\tilde{W}^*(\mathbf{k}, z_2)$$
$$\times P_m(k|z_{12}), \tag{7}$$

where $\tilde{W}$ is the Fourier transform of the survey window function.

In this article, we are mostly interested in the case of cluster number counts $N_{\text{cl}}(i_M, i_z)$ per steradian within bins of mass i_M and bins of redshift i_z[2],

$$N_{\text{cl}}(i_M, i_z) = \int_{z \in \text{bin}(i_z)} \mathrm{d}V \int_{M \in \text{bin}(i_M)} \mathrm{d}M\, \frac{\mathrm{d}n_h}{\mathrm{d}M}(M, z), \tag{8}$$

where $\mathrm{d}n_h/\mathrm{d}M$ is the halo mass function. The SSC of these counts is traditionally denoted simply as sample (co)variance in the cluster literature (Hu & Kravtsov 2003). It is important to notice though that super-horizon modes affect even full sky surveys and must always be accounted for. The SSC for clusters counts is given by (e.g. Lacasa & Rosenfeld 2016)

$$\text{Cov}^{\text{SSC}}(N_{\text{cl}}(i_M, i_z), N_{\text{cl}}(j_M, j_z))$$
$$= \int \mathrm{d}V_{12}\, \frac{\partial n_h}{\partial \delta_b}(i_M, z_1)\, \frac{\partial n_h}{\partial \delta_b}(i_M, z_2)\, \sigma^2(z_1, z_2), \tag{9}$$

where the response of the halo number density $n_h = \mathrm{d}N_{\text{cl}}/\mathrm{d}V$ to a change of background density is given by the number-weighted halo-bias averaged within the bin (e.g. Schmidt et al. 2013) as follows:

$$\frac{\partial n_h}{\partial \delta_b}(i_M, z) = \int_{M \in \text{bin}(i_M)} \mathrm{d}M\, \frac{\mathrm{d}n_h}{\mathrm{d}M}\, b(M, z), \tag{10}$$

---

[2] Throughout this article N_{cl} is the number of clusters per steradian. For application to a survey, N_{cl} and its covariance may be rescaled by the observed area.
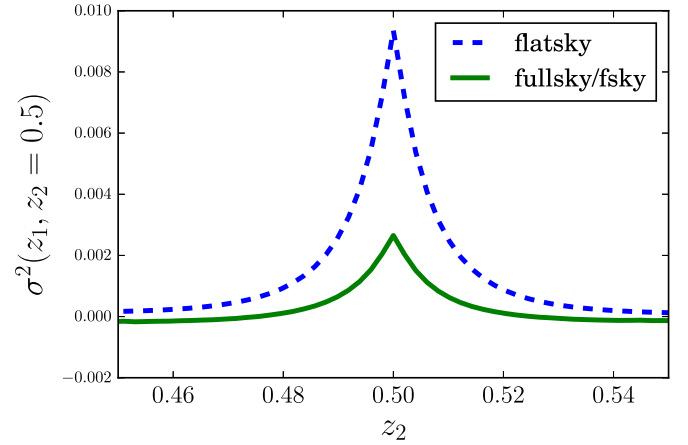

Fig. 1. Comparison of $\sigma^2(z_1, z_2)$ for $z_1 = 0.5$ in different cases. In blue the flat sky formula Eq. (12) for a survey angular radius $\theta_S = 5$ deg. In green the full sky formula Eq. (11) rescaled by a factor $1/f_{\text{SKY}}$.

where $b(M, z)$ is the (first order) halo bias.

Most of the elements for the computation of the covariance pose no particular numerical problem. A notable exception is $\sigma^2(z_1, z_2)$, as it depends on the survey geometry. Numerically tractable formulae are known only for a few cases with simple geometries. First, for the case of full sky, we have (Lacasa & Rosenfeld 2016)

$$\sigma^2_{\text{fullsky}}(z_1, z_2) = \frac{1}{2\pi^2} \int k^2\, \mathrm{d}k\, j_0(kr_1)\, j_0(kr_2)\, P_m(k|z_{12}), \tag{11}$$

where $r_i = r(z_i)$. Second, an expression for $\sigma^2(z_1, z_2)$ is also known in the flat sky limit (e.g. Hu & Kravtsov 2003; Lima & Hu 2007; Aguena & Lima 2016)

$$\sigma^2_{\text{flatsky}}(z_1, z_2) = \frac{1}{2\pi^2} \int k_\perp\, \mathrm{d}k_\perp\, 4\, \frac{J_1(k_\perp \theta_S r_1)}{k_\perp \theta_S r_1}\, \frac{J_1(k_\perp \theta_S r_2)}{k_\perp \theta_S r_2}$$
$$\times \int \mathrm{d}k_\parallel\, \cos\left[k_\parallel(r_1 - r_2)\right]\, P_m(k|z_{12}), \tag{12}$$

for a cylindrical window function of radius θ_S delineating a survey solid angle $\Omega_S = 2\pi(1 - \cos\theta_S) \approx \pi\theta_S^2$, and the wavevector $\mathbf{k} = (k_\parallel, k_\perp)$ is split into its components parallel and perpendicular to the line of sight, where $k^2 = k_\perp^2 + k_\parallel^2$.

Figure 1 shows the comparison of σ^2 resulting from the two expressions above for an angular window of radius $\theta_S = 5$ deg. The full sky covariance was rescaled by a factor f_{SKY}^{-1}, where $f_{\text{SKY}} = \Omega_S/4\pi$ and $\Omega_S \simeq \pi\theta_S^2$ for θ_S translated in radians.

We see that both covariances share the behaviour of peaking at $z_1 = z_2$ and decreasing to zero as $|z_1 - z_2|$ increases[3]. Regarding the amplitudes, we see that rescaling the full sky covariance by the usual $1/f_{\text{SKY}}$ factor underpredicts the covariance by a factor ~ 3.4 in this case[4]. Even when rescaling the covariances to the same peak amplitude by hand, we find that their shapes are broadly similar but differ in details. The full sky covariance is more strongly peaked at the centre, then decreases and crosses zero, reaching a negative minima of height $\sim 7\%$ of its peak, before slowly asymptoting zero. By contrast, the flat sky covariance is a bit broader in its positive part, but its negative minimum

---

[3] By eye, both functions seem symmetric around $z_1 = z_2$, but in detail this is not true.

[4] An artificial factor $1/f_{\text{SKY}}^2$ would not fare better, this time overpredicting the covariance.

is only ∼0.8% of its peak, and it asymptotes faster to zero. This clearly indicates that SSC is non-trivially related to sky coverage, and a more general approach is required for its accurate computation.

3. Comparison of super-sample covariance approximations

The full Eq. (9) for SSC can be numerically expensive, as it requires a double redshift integral. Furthermore, for each pair of redshift, a double $(k_\perp, k_\parallel)$ integral is required (e.g. in the flat sky case). To tame down this burden, several approximations have been devised in the literature. Below we consider two approximations in the flat sky regime.

First, in Eq. (9) we can assume that $\partial n_h / \partial \delta_b$ varies slowly with redshift within the bins, such that it can be approximated by its bin-averaged value and taken out of the integral. In fact this may be a good approximation for sufficiently narrow bins. In this case, Eq. (9) takes the form (Hu & Kravtsov 2003)

$$\mathrm{Cov}^{\mathrm{SSC}}(N_{\mathrm{cl}}(i_M, i_z), N_{\mathrm{cl}}(j_M, j_z))$$
$$\approx N_{\mathrm{cl}}(i_M, i_z)\, b(i_M, i_z)\, N_{\mathrm{cl}}(j_M, j_z)\, b(j_M, j_z)\, S_{i_z, j_z}, \quad (13)$$

where

$$S_{i_z, j_z} = \int \frac{\mathrm{d}V_{12}}{V_{12}}\, \sigma^2(z_1, z_2) \quad (14)$$

and the normalised number-weighted halo bias within the bin is

$$b(i_M, i_z) = \frac{1}{N_{\mathrm{cl}}(i_M, i_z)} \int_{\substack{z \in \mathrm{bin}(i_z) \\ M \in \mathrm{bin}(i_M)}} \mathrm{d}M\, \mathrm{d}V\, \frac{\mathrm{d}n_h}{\mathrm{d}M}\, b(M, z). \quad (15)$$

This approximation implicitly assumes that $b(M, z)$ varies slowly with redshift within the bin i_z, compared to $\sigma^2(z_1, z_2)$.
For a cylindrical window function of height δr in the flat sky case, the sample variance matrix S_{i_z, j_z} takes the form (Hu & Kravtsov 2003; Lima & Hu 2007; Aguena & Lima 2016)

$$S_{i_z, j_z} = \frac{1}{2\pi^2} \int k_\perp\, \mathrm{d}k_\perp\, 4\, \frac{J_1(k_\perp \theta_S r_1)}{k_\perp \theta_S r_1}\, \frac{J_1(k_\perp \theta_S r_2)}{k_\perp \theta_S r_2}$$
$$\times \int \mathrm{d}k_\parallel\, j_0\left(\frac{k_\parallel \delta r_1}{2}\right) j_0\left(\frac{k_\parallel \delta r_2}{2}\right) \cos\left[k_\parallel (r_1 - r_2)\right] P_m(k|z_{12}), \quad (16)$$

where the power spectrum $P_m(k|z_{12})$ is evaluated at the centre of the respective redshift bins. A nice feature of this approximation is that it removes the need to compute a double redshift integral. Furthermore the double integral over $(k_\perp, k_\parallel)$ only needs to be computed n_z^2 times (where n_z is the number of redshift bins), instead of all redshift pairs (z_1, z_2) needed to compute the redshift integral otherwise. This speeds up considerably the computation of SSC. We call this approximation the Sij method.

Second, another approximation used (e.g. Krause & Eifler 2017) is that $\sigma^2(z_1, z_2)$ is a Dirac delta function at $z_1 = z_2$, so that the double redshift integral collapses into a single integral

$$\mathrm{Cov}^{\mathrm{SSC}}(N_{\mathrm{cl}}(i_M, i_z), N_{\mathrm{cl}}(j_M, j_z))$$
$$\approx \delta_{i_z, j_z} \int \mathrm{d}V\, r^2(z)\, \frac{\partial n_h}{\partial \delta_b}(i_M, z)\, \frac{\partial n_h}{\partial \delta_b}(j_M, z)\, \sigma_b(\Omega_S, z), \quad (17)$$

where, following notation from Krause & Eifler (2017), we have

$$\sigma_b(\Omega_S, z) = \int \frac{k_\perp\, \mathrm{d}k_\perp}{2\pi}\, P_m(k_\perp, z) \left[\frac{2 J_1(k_\perp r(z)\theta_S)}{k_\perp r(z)\theta_S}\right]^2, \quad (18)$$
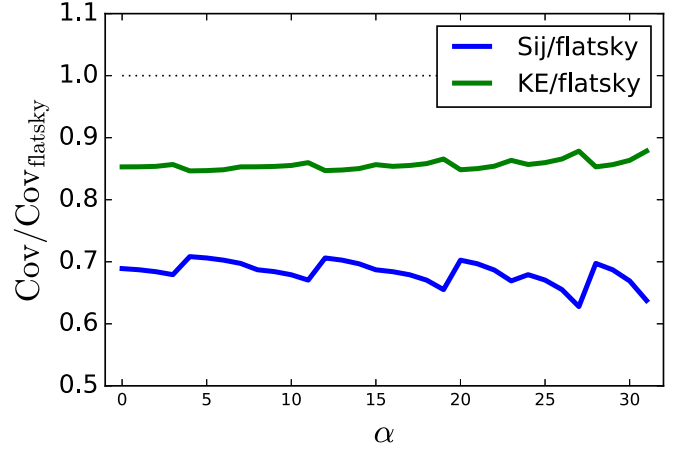
Fig. 2. Ratio of the SSC covariances from the Sij approximation and KE approximation to the full computation. Auto-z covariances are ordered as a function of $\alpha = j_M + n_M\, i_M + n_M^2\, i_z$ (see text for details).

where $P_m(k_\perp, z) = P_m(k = k_\perp, z)$, and we note that the $k_\parallel$ integral has disappeared and that σ_b has units of Mpc/h.

The assumption behind this approximation is that the 3D window function $W(\mathbf{x})$ is much wider in the radial direction than in the transverse direction. Thus super-survey modes have $k_\parallel \ll k_\perp$, and $P_m(k)$ can be taken as approximately constant within the $k_\parallel$ integral of Eq. (12). We thus expect the approximation to fare better for wide redshift bins and small angles. By limiting the computation to equal redshift bins $i_z = j_z$, and reducing the multiple integrals to a single redshift and a single wave-vector $k_\perp$ integral, the approximation speeds up the SSC computation considerably. Hereafter, we call this approximation the KE method.

We implemented all three SSC methods numerically: (1) full computation from Eqs. (9) and (12), (2) Sij approximation, and (3) KE approximation. Figure 2 shows the ratio of the auto-z covariances to the full computation, plotted as a function of $\alpha = j_M + n_M\, i_M + n_M^2\, i_z$; i.e. the first four points show the covariance ratio for $i_M = i_z = 0$ ($14 < \log M < 14.5$ and $0.4 < z < 0.5$) and varying j_M, the next four points are for $i_M = 1$, $i_z = 0$, etc. This order is simply convenient to plotting this multi-variable function. As visible in Fig. 2 both the Sij and KE approximations underpredict the amplitude of the covariance by 30–35% and by ∼15%, respectively. The fact that the Sij approximation underpredicts the covariance conforms to our intuition; indeed, by taking values at the centre of the redshift bins, the approximation neglects the fact that the mass function and matter power spectrum both grow with time, which makes their integral larger[5]. We checked that the Sij method and full computation are in much better agreement for smaller redshift bins, and indeed it can be seen analytically that they reduce to each other in the limit $\Delta z \to 0$.

In Fig. 3, we show the correlation matrices for cluster counts obtained from each of the three SSC computations. The data points are ordered with increasing mass then increasing redshift, i.e. the number count index is $i_M + n_M\, i_z$[6].

As visible in Fig. 3, for cross-redshift covariances the full computation gives a positive correlation between our two

---

[5] E.g. $1/3 = \int_0^1 x^2 \mathrm{d}x > 1/2 * \int_0^1 x \mathrm{d}x = 1/4$.
[6] For example 0 refers to the lowest mass bin ($14 < \log M < 14.5$) in the first redshift bin ($0.4 < z < 0.5$), 3 refers to the highest mass bin ($15.5 < \log M < 16$) still in the first redshift bin, and then we go to the second redshift bin ($0.5 < z < 0.6$).
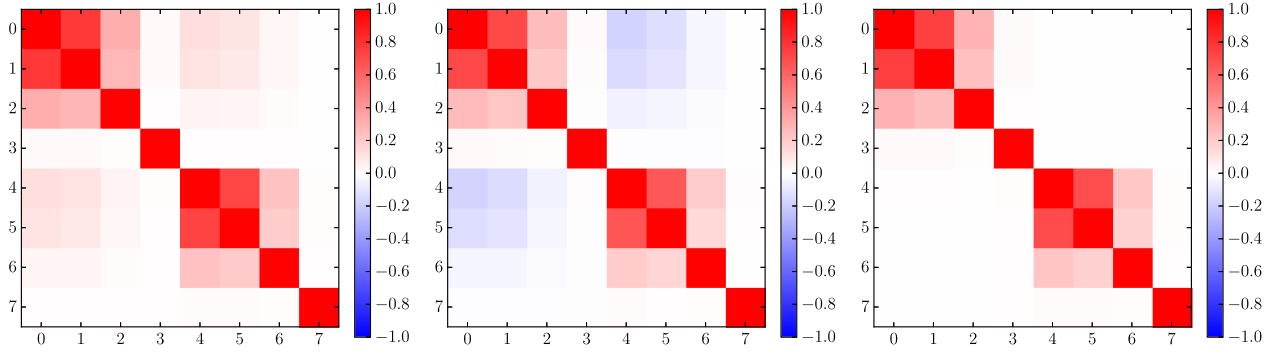
Fig. 3. Comparison of the cluster counts correlation matrix for different SSC computations, for a survey with angular radius $\theta_S = 5$ deg. The SSC matrix is shown in 2 blocks for the redshift bins, and each block has 4 entries for the logarithmic mass bins. *From left to right*: full numerical computation from Eq. (9), Sij approximation, and KE approximation.

redshift bins, reaching ~13% at maximum for $\Delta z = 0.1$. On the other hand, the Sij approximation predicts anti-correlation between the bins, reaching at maximum -16%, while the KE approximation by construction predicts zero correlation. We checked that the Sij method performs better as the redshift bin width decreases. For instance at $\Delta z = 0.01$, the agreement with the full computation is very good.

As a conclusion of this section, we see that the full computation is necessary to faithfully predict the SSC covariance matrix in the case $\Delta z = 0.1$. In fact, this is a representative binning for current photometric galaxy surveys whose present photo-z errors are of this order (e.g. Sánchez et al. 2014).

4. Super sample covariance for partial sky coverage

In this section, we assume for simplicity that the survey mask is independent of redshift, as we in fact already did implicitly in Eq. (5). It is however straightforward to generalise the formalism to a redshift-dependent mask (see Appendix A).

4.1. Formal derivation

We want to compute the covariance of the background mode

$$\sigma^2(z_1, z_2) = \langle \delta_b(z_1)\, \delta_b(z_2) \rangle. \tag{19}$$

Given that the mask has zero value outside the survey, and given the normalisation of spherical harmonics such that $Y_{00} = 1/\sqrt{4\pi}$, the background mode of the survey is related to the monopole of the masked matter density as

$$\delta_b(z) = \frac{1}{\sqrt{4\pi}\, f_{\text{SKY}}}\, a_{00}^{\text{masked}}(z), \tag{20}$$

such that we have

$$\sigma^2(z_1, z_2) = \frac{1}{4\pi f_{\text{SKY}}^2}\, \hat{C}_0^{\text{masked}}(z_1, z_2), \tag{21}$$

where $\hat{C}_0^{\text{masked}}$ is the angular power spectrum at $\ell = 0$ of the masked matter density field. From C_ℓ pseudo-spectrum methods (e.g. Hivon et al. 2002), we know that the masked power spectrum is related to the true spectrum via the so-called coupling matrix, such that

$$\hat{C}_0^{\text{masked}}(z_1, z_2) = \sum_\ell \mathcal{M}_{0,\ell}\, C_\ell^{\text{true}}(z_1, z_2), \tag{22}$$

where C_ℓ^{true} (hereafter simply C_ℓ^{m}) is the angular spectrum of matter density in an infinitesimal redshift shell. The coupling matrix is given by

$$\mathcal{M}_{\ell_1, \ell_2} = (2\ell_2 + 1) \sum_{\ell_3} \frac{(2\ell_3 + 1)}{4\pi} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2 C_{\ell_3}(W), \tag{23}$$

where $C_\ell(W)$ is the angular power spectrum of the angular window $W(\hat{\mathbf{n}})$. In our case, given the properties of Wigner symbols, this simplifies to

$$\mathcal{M}_{0,\ell} = \frac{(2\ell + 1)^2}{4\pi} \begin{pmatrix} 0 & \ell & \ell \\ 0 & 0 & 0 \end{pmatrix}^2 C_\ell(W). \tag{24}$$

Furthermore we have $\begin{pmatrix} 0 & \ell & \ell \\ 0 & 0 & 0 \end{pmatrix} = (2\ell + 1)^{-1/2}$ and thus

$$\sigma^2(z_1, z_2) = \frac{1}{\Omega_S^2} \sum_\ell (2\ell + 1)\, C_\ell(W)\, C_\ell^{\text{m}}(z_1, z_2), \tag{25}$$

where $\Omega_S = 4\pi f_{\text{SKY}}$ and the matter angular power spectrum C_ℓ^{m} is given by (e.g. Campagne et al. 2017)

$$C_\ell^{\text{m}}(z_1, z_2) = \frac{2}{\pi} \int k^2\, \mathrm{d}k\; j_\ell(kr_1)\, j_\ell(kr_2)\, P_m(k|z_{12}). \tag{26}$$

The combined Eqs. (25) and (26) represent the main analytical result of this article.

4.2. Limiting cases and remarks

It is interesting to examine a few limiting cases of Eqs. (25) and (26). In the full sky case, we have $C_\ell(W) = 4\pi\, \delta_{0,\ell}$ and $f_{\text{SKY}} = 1$, and therefore thus we have

$$\sigma^2(z_1, z_2) = \sigma_{\text{fullsky}}^2(z_1, z_2) = C_0^{\text{m}}(z_1, z_2) \big/ 4\pi\,, \tag{27}$$

recovering indeed Eq. (11).

In partial sky, $C_0(W) = 4\pi f_{\text{SKY}}^2$, such that

$$\sigma^2(z_1, z_2) = \sigma_{\text{fullsky}}^2(z_1, z_2) + \frac{1}{\Omega_S^2} \sum_{\ell \geq 1} (2\ell + 1)\, C_\ell(W)\, C_\ell^{\text{m}}(z_1, z_2), \tag{28}$$

i.e. the full sky covariance term is the first in the sum contributing to the partial-sky covariance.

If $C_\ell^m(z_1, z_2)$ is scale independent, i.e. $C_\ell^m(z_1, z_2) = C_0^m(z_1, z_2)$ independently of ℓ (e.g. for $P_m(k|z_{12}) = \text{const.}$, see Sect. 5.2), we have

$$\sigma^2(z_1, z_2) = \sigma_{\text{fullsky}}^2(z_1, z_2) \times \frac{1}{f_{\text{SKY}}^2} \sum_\ell \frac{2\ell+1}{4\pi} C_\ell(W)$$
$$= \sigma_{\text{fullsky}}^2(z_1, z_2)/f_{\text{SKY}}, \qquad (29)$$

and therefore $\text{Cov}^{\text{SSC}} = \text{Cov}_{\text{fullsky}}^{\text{SSC}}/f_{\text{SKY}}$, i.e. we obtain the usual f_{SKY} approximation to partial sky covariance. Conversely the reverse is also true: if the f_{SKY} approximation holds for any mask, then $C_\ell^m(z_1, z_2)$ is constant. Given that for viable cosmological models $C_\ell^m(z_1, z_2)$ is not constant, this shows that SSC is a non-trivial source of covariance that cannot be treated by classical approximations.

Now we notice that Eq. (25) can be rewritten as

$$\sigma^2(z_1, z_2) = \frac{1}{f_{\text{SKY}}^2} \sum_\ell \frac{2\ell+1}{4\pi} C_\ell(W) \sigma_\ell^2(z_1, z_2), \qquad (30)$$

with

$$\sigma_\ell^2(z_1, z_2) = \frac{C_\ell^m(z_1, z_2)}{4\pi}$$
$$= \frac{1}{2\pi^2} \int k^2 \, dk \, j_\ell(kr_1) \, j_\ell(kr_2) \, P_m(k|z_{12}), \qquad (31)$$

such that in full sky $\sigma^2(z_1, z_2) = \sigma_0^2(z_1, z_2)$. Applying this to the SSC covariance of two observables, for example number counts, we have

$$\text{Cov}^{\text{SSC}}(O_1, O_2) = \frac{1}{f_{\text{SKY}}^2} \sum_\ell \frac{2\ell+1}{4\pi} C_\ell(W) \, \text{Cov}_\ell^{\text{SSC}}(O_1, O_2), \quad (32)$$

where

$$\text{Cov}_\ell^{\text{SSC}}(O_1, O_2) = \int dV_{12} \frac{\partial \mathfrak{o}_1}{\partial \delta_b} \frac{\partial \mathfrak{o}_1}{\partial \delta_b} \sigma_\ell^2(z_1, z_2) \qquad (33)$$
$$= \frac{1}{4\pi} \int dV_{12} \frac{\partial \mathfrak{o}_1}{\partial \delta_b} \frac{\partial \mathfrak{o}_1}{\partial \delta_b} C_\ell^m(z_1, z_2).$$

Therefore the combined Eqs. (30)–(33), along with the window spectrum $C_\ell(W)$, allow for the computation of the SSC in partial sky for a general window and binning choice.

Numerically, we can pre-compute and tabulate $\text{Cov}_\ell^{\text{SSC}}(O_1, O_2)$, and then just change $C_\ell(W)$ as the mask varies. This enables easier studies of the mask effect such as optimisations of survey strategy (Takahashi et al. 2014), forecasts for improvements as a survey area grows, and comparisons between different surveys. Conversely, for a fixed survey with a well-defined mask, we can pre-compute $C_\ell(W)$ for the specific survey mask, and change $\text{Cov}_\ell^{\text{SSC}}(O_1, O_2)$ as a function of cosmology within likelihood cosmological analyses, taking full account of geometry, mask, selection, and cosmological dependencies in the covariances, and therefore deriving reliable parameter uncertainties.

4.3. Implementation

When estimating the SSC from Eq. (32), one numerical difficulty is the evaluation of the integral in Eq. (31), given that the Bessel functions highly oscillate with slow damping as $kr_i \to \infty$.

For the first multipoles, we may express the Bessel functions in terms of sine and cosine. Through trigonometrical identities, the integrals with products of Bessel function can thus be expressed as a sum of Fourier transforms. Derivations and expressions for the first three multipoles are given in Appendix C. For illustrative purposes we give below the expression for $\ell = 0$

$$C_0^m(z_1, z_2) = \frac{I_0^c(r_1 - r_2) - I_0^c(r_1 + r_2)}{2}, \qquad (34)$$

where

$$I_0^c(r) = \frac{2}{\pi} \int k^2 \, dk \, \cos(kr) \, P_m(k|z_{12})/k^2 \qquad (35)$$

is a continuous cosine transform that can be efficiently approximated numerically with a discrete fast Fourier transform (FFT). However, as argued in Appendix C, this method becomes too cumbersome at high ℓ and may also become numerically unstable. We could carry it out only for $\ell = 0, 1, 2$.

As an alternative, we decided not to evaluate $\sigma^2(z_1, z_2)$, but instead to switch the order of the integrals over k and z, to compute the covariance directly as

$$\text{Cov}_\ell^{\text{SSC}}(N_{\text{cl}}(i_M, i_z), N_{\text{cl}}(j_M, j_z)) = \frac{1}{2\pi^2} \int k^2 dk \, P_m(k|z = 0)$$
$$\times \Psi_\ell^{n_h}(k|i_M, i_z) \, \Psi_\ell^{n_h}(k|j_M, j_z), \qquad (36)$$

where the kernel $\Psi_\ell^{n_h}$ is given by

$$\Psi_\ell^{n_h}(k|i_M, i_z) \equiv \int_{z \in \text{bin}(i_z)} dV \, G(z) \frac{\partial n_h}{\partial \delta_b}(i_M, z) \, j_\ell(kr). \qquad (37)$$

As can be seen in Fig. 4, this redshift integral effectively damps out the Bessel oscillations on scales $k > k_{\text{peak}} + 2\pi/\Delta r(i_z)$, where $\Delta r(i_z) = r(z_{\text{max}}) - r(z_{\text{min}})$ is the width of the redshift bin in terms of comoving distance. This damping makes the k integral Eq. (36) much easier to deal with numerically compared to Eq. (26), since the integrand support is now more compact. So numerically we could just compute the later integrals by brute force, and this is the method we use for all numerical results shown hereafter in this article. We note that this method does not compute $\sigma^2(z_1, z_2)$ as an intermediary product, and thus prevents from comparison of this quantity in partial sky with its alter ego in the full sky or flat sky limits, as in Fig. 1.

We remark that alternatively, Eq. (37) can be expressed as a Hankel transform

$$\Psi_\ell^{n_h}(k|i_M, i_r) = \int_{r \in \text{bin}(i_r)} dr \, r^2 \, G(r) \frac{\partial n_h}{\partial \delta_b}(i_M, r) \, j_\ell(kr), \qquad (38)$$

via the inversion of the $r(z)$ relation and defining a radial bin i_r from i_z. Equation (38) can be efficiently evaluated using Fourier transform methods such as FFTLog (e.g. Hamilton 2000; Camacho et al. in prep.). Although this is not the numerical method we used in this article, we note it can be a useful approach for future high-precision applications.

The results presented in this article consider the case of cluster counts, and the equations can be straightforwardly generalised to other probes of the LSS; see Appendix B.
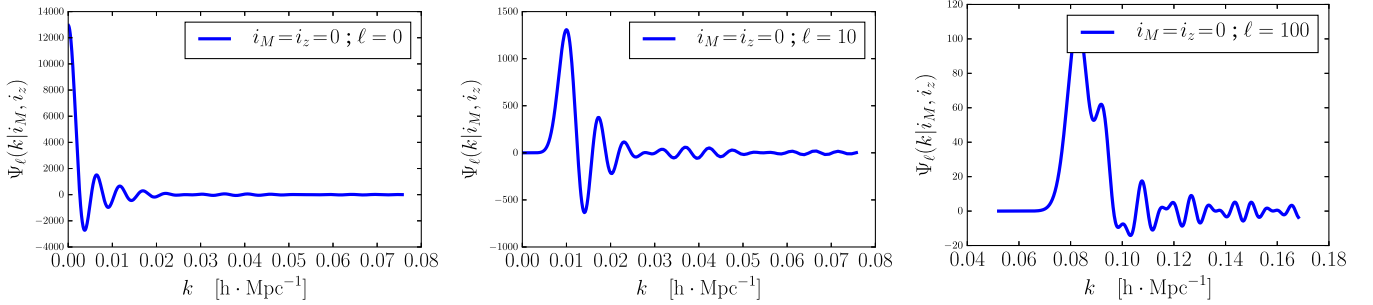
Fig. 4. $\Psi_\ell^{n_h}(k|i_M, i_z)$ as a function of k, for the representative case $i_M = i_z = 0$ (log M = 14–14.5, z = 0.4–0.5). *Left*: at the lowest multipole $\ell = 0$. *Centre*: at $\ell = 10$. *Right*: at $\ell = 100$.

5. Results

First, as a consistency test, we checked $\mathrm{Cov}_{\ell=0}^{\mathrm{SSC}}$ against the full sky covariance matrix computed via $\sigma_{\mathrm{fullsky}}^2(z_1, z_2)$ given in Eq. (11) (e.g. Lacasa & Rosenfeld 2016). We find good agreement, to 0.8% precision on the auto-redshift covariance, and 7% precision on the cross-redshift covariance.

Second, we show in Fig. 5 the general results for $\mathrm{Cov}_\ell^{\mathrm{SSC}}(N_{\mathrm{cl}}(i_M, i_z), N_{\mathrm{cl}}(j_M, j_z))$ in two representative cases: auto-covariance of the same redshift bin and cross-correlation (i.e. cross-covariance normalised by the corresponding auto-covariances) between two redshift bins. Both plots are for the lowest mass bin ($i_M = j_M = 0$, corresponding to log M = 14–14.5), although the shape of the curves does not change significantly when taking other mass bins (even with $i_M \neq j_M$), only their amplitude changes.

For the auto-covariance (left plot), we see that it first rises with ℓ, to a maximum corresponding to the angular scale of the matter-radiation equality (i.e. the peak of $P_m(k)$) at that redshift, and then decreases monotonically. This behaviour conforms to our expectations, since the scale dependence is that of a projection of the 3D power spectrum $P_m(k)$. This means that $C_\ell^m(z_1, z_2)$ is not constant with ℓ, and thus, as already discussed in Sect. 4, it means that the f_{SKY} approximation does not hold for SSC.

For the cross-covariance (right plot), the situation is interesting as we see that the covariance is first negative, increases, crosses zero, and reaches a maximum towards $\ell = 25$, and then decreases asymptotically to zero. This means that small surveys have a robustly negligible covariance between redshift bins (at least for this bin width $\Delta z = 0.1$), and can thus use block-diagonal covariance matrices to speed up their likelihoods estimations. However for surveys with a large sky coverage, the cross-covariance is non-negligible and depends on the survey area and shape, becoming either positive or negative depending on the mask. In those cases, careful estimation of the SSC is thus critical.

5.1. Application to a realistic survey mask

In order to illustrate the SSC method in a realistic case, we used two Healpix (Górski et al. 2005) masks visible in Fig. 6. The first mask is binary and is broadly similar to the footprint of DES[7,8], although we warn that it does not in any way represent the actual DES survey area, and we do not attempt to draw any particular conclusion for that survey. The second mask represents a more pessimistic case, where we considered that 15% of the survey area had to be discarded due to bright stars, satellite trails, or other systematics. In order to simulate this effect, we

---

[7] www.darkenergysurvey.org
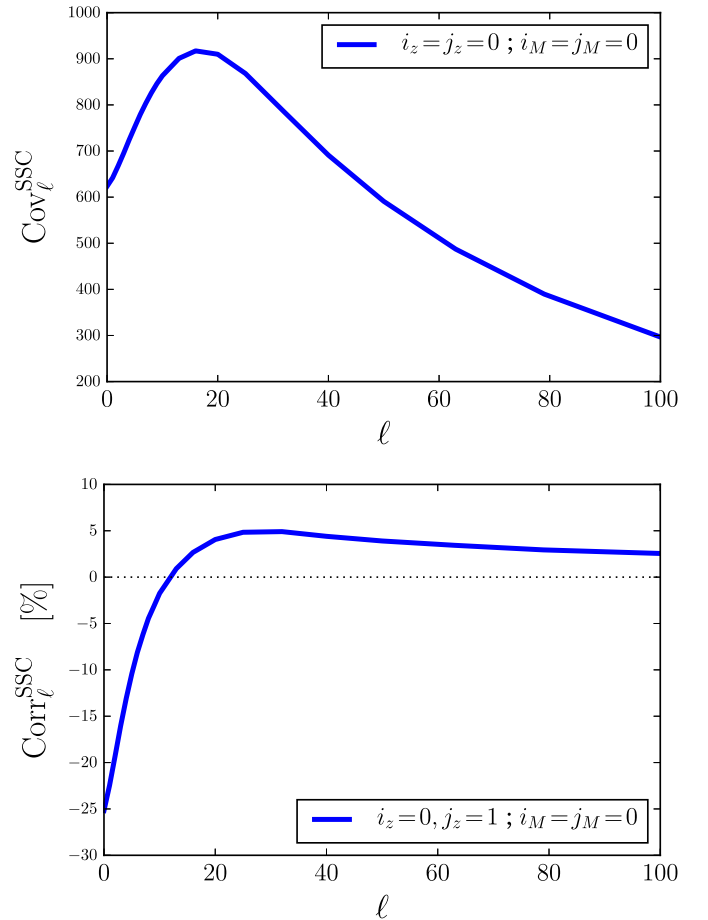[8] This mask was created and kindly provided to us by Flavia Sobreira.

Fig. 5. $\mathrm{Cov}_\ell^{\mathrm{SSC}}$ as a function of ℓ, in the representative case $i_M = j_M = 0$ (log M = 14–14.5). *Top*: same redshift $i_z = j_z = 0$ ($z = 0.4$–0.5). *Bottom*: cross-redshift $i_z = 1 \times j_z = 2$ ($z = [0.4$–$0.5] \times [0.5$–$0.6]$).

simply upgraded the mask to high resolution (nside = 4096), poked random holes in it, then degraded the mask back to the original resolution (nside = 1024).

The angular power spectra $C_\ell(W)$ for these two masks can be seen on Fig. 7. At low multipoles, the two spectra differ by a constant multiplicative factor 0.85^2, which is simply the ratio of the respective sky coverages. A second component in the spectrum of the second mask appears on smaller scales, which is a constant shot noise due to the random holes. However we see that this component is very subdominant, hence we can already expect that the only difference between the covariance of the two masks is due to the different sky coverage.

When implementing the sum over multipoles of Eq. (32), we find that we reached 1% convergence already at $\ell_{\max} = 50$. This

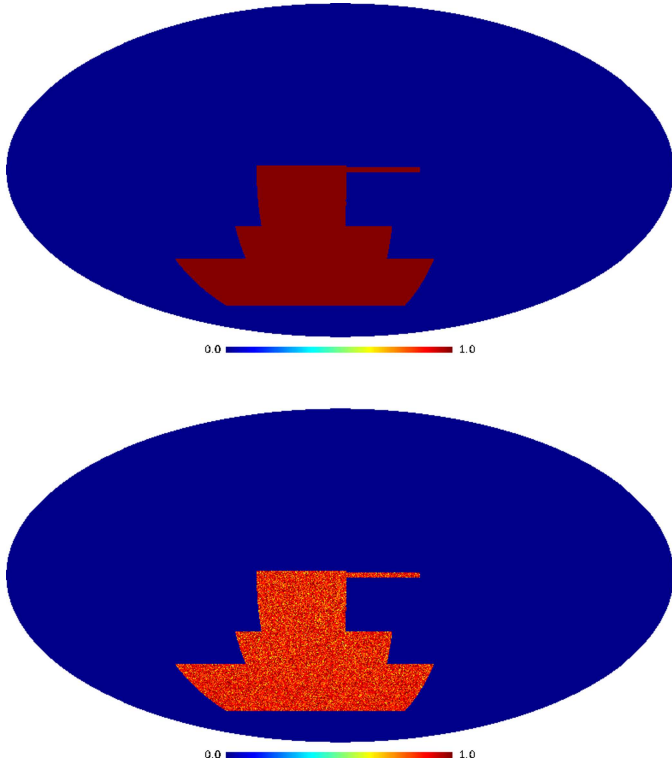**Fig. 6.** Masks used in the analysis. *Top*: simple footprint, assumed observed uniformly. *Bottom*: same but with simulated 15% rejection of observations due to systematics (see text for details).
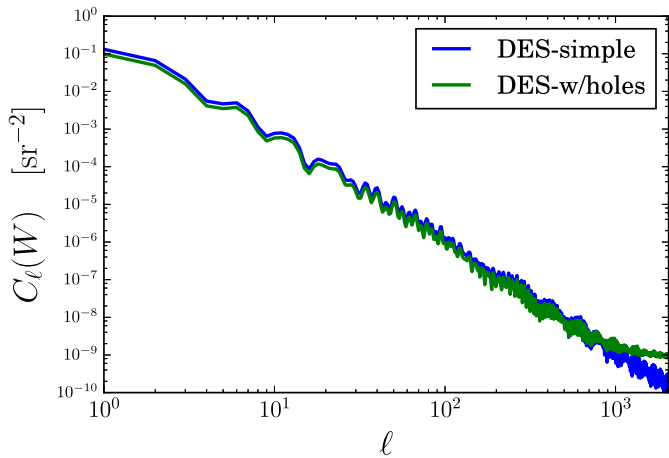

Fig. 7. Angular power spectra of the two masks used in the analysis and shown in Fig. 6.

means that $\mathrm{Cov}_\ell^{\mathrm{SSC}}$ need only be computed on a small number of multipoles for current and future surveys with large sky coverage, rendering the partial sky formalism developed here even more computationally efficient. Comparing the SSC covariances of the two masks, we find that they indeed only differ due to the different sky fraction, but when renormalised by f_{SKY} they are identical to numerical precision (0.01% in our case). The total correlation matrix (including shot-noise) for the first mask is shown in Fig. 8.

We see that the SSC is an important contribution to the covariance matrix, dominating the error bars at $\log M \leq 14.5$. Concerning cross-redshifts, the SSC yields an anti-correlation between the adjacent bins reaching up to -10%. The cross-covariance is thus far from negligible for such a large survey.
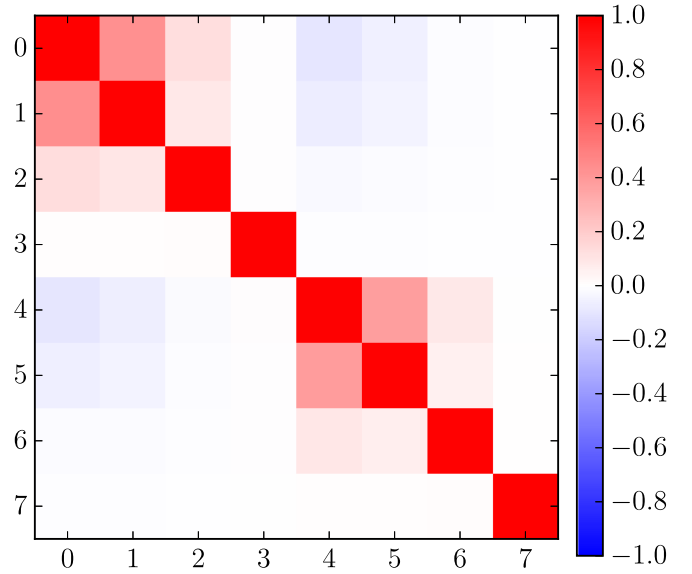
Fig. 8. Total correlation matrix (SSC + shot-noise) of cluster counts for the first mask. The matrix is organised, as in Fig. 3, in two redshift blocks of increasing mass.

The formalism presented in Sect. 4 is thus perfectly adapted to the numerical prediction of SSC, even in the case of a complex survey geometry. In fact we see that such computation is indeed necessary to reproduce the non-trivial behaviour of SSC, yielding, for example non-negligible anti-correlation of redshift bins, in the case presented above.

5.2. Flat sky limit

One remaining question is that of the link between the partial sky approach developed in Sect. 4 and the flat sky approximation used previously in the literature and shown in Eq. (12). The two approaches cannot be compared or related at the level of the formalism because the $(k_\perp, k_\parallel)$ splitting does not apply for sufficiently large angles. However, we can compare what covariance the two formalisms predict in some limits.

First, we can compare the equations analytically in the case of a constant spectrum $P_m(k|z_{12}) = \mathrm{cst} \equiv 1$. In the partial sky formalism we have

$$C_\ell^m(z_1, z_2) = \frac{2}{\pi} \int k^2 \mathrm{d}k\, j_\ell(kr_1) j_\ell(kr_2) \times 1 = \frac{\delta(r_1 - r_2)}{r_1^2}, \quad (39)$$

and

$$\mathrm{Cov}_\ell^{\mathrm{SSC}}(N_{\mathrm{cl}}(i_M, i_z), N_{\mathrm{cl}}(j_M, j_z))$$
$$= \frac{\delta_{i_z, j_z}}{4\pi} \int \mathrm{d}V\, \frac{\partial n_h}{\partial \delta_b}(i_M, z) \frac{\partial n_h}{\partial \delta_b}(j_M, z) \quad (40)$$

is independent of ℓ. Thus the SSC is given by

$$\mathrm{Cov}^{\mathrm{SSC}}(\cdots) = \frac{\delta_{i_z, j_z}}{4\pi} \int \mathrm{d}V\, \frac{\partial n_h}{\partial \delta_b}(i_M, z) \frac{\partial n_h}{\partial \delta_b}(j_M, z)$$
$$\times \frac{1}{f_{\mathrm{SKY}}^2} \underbrace{\sum_\ell \frac{(2\ell + 1)}{4\pi} C_\ell(W)}_{= f_{\mathrm{SKY}}}$$
$$= \frac{\delta_{i_z, j_z}}{\Omega_S} \int \mathrm{d}V\, \frac{\partial n_h}{\partial \delta_b}(i_M, z) \frac{\partial n_h}{\partial \delta_b}(j_M, z), \quad (41)$$
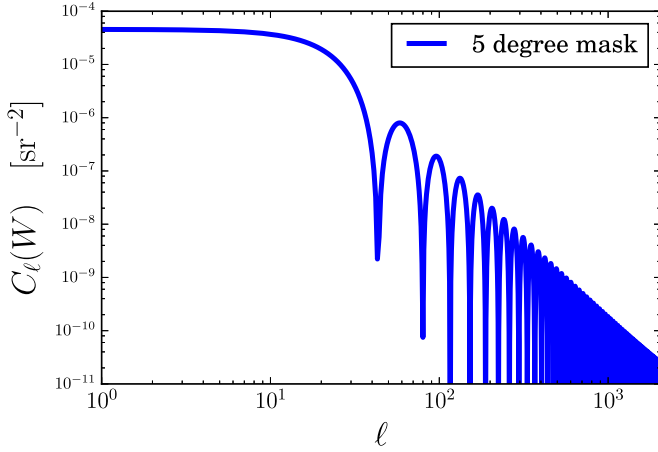
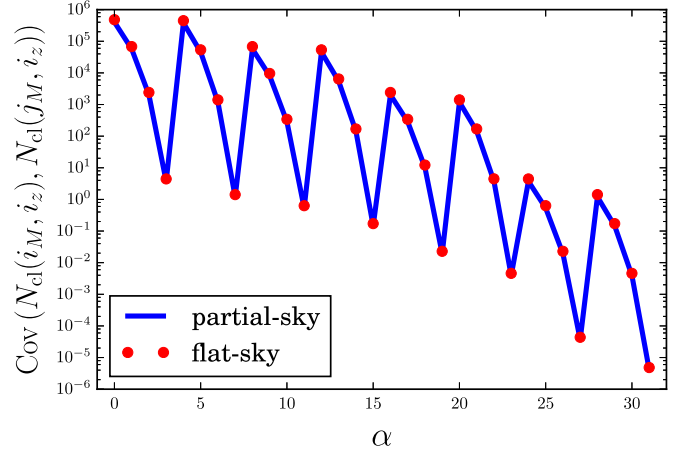**Fig. 9.** Angular power spectrum of the 5 deg polar cap mask.

Fig. 10. Comparison of the SSC covariances derived from the partial sky formalism and the flat sky formula for a 5 deg circular sky patch. Auto-z covariances are shown as a function of $\alpha = j_M + n_M i_M + n_M^2 i_z$ (see text for details).

where we made the counts $N_{\rm cl}$ implicit for the sake of clarity.

Now in the flat sky formalism we have

$$
\sigma^2(z_1, z_2) = \frac{1}{2\pi^2} \int_0^\infty k_\perp \, dk_\perp \, 4 \frac{J_1(k_\perp \theta_S r_1)}{k_\perp \theta_S r_1} \frac{J_1(k_\perp \theta_S r_2)}{k_\perp \theta_S r_2}
$$

$$
\times \underbrace{\int_0^\infty dk_\parallel \, \cos\left[k_\parallel (r_1 - r_2)\right] \times 1}_{=(2\pi)\,\delta(r_1-r_2)/2}
$$

$$
= \frac{2}{\pi} \, \delta(r_1 - r_2) \, \frac{1}{(\theta r_1)^2} \underbrace{\int x \, dx \, (J_1(x)/x)^2}_{=1/2}
$$

$$
= \frac{\delta(r_1 - r_2)}{r_1^2} \times \frac{1}{\Omega_S}, \tag{42}
$$

where we changed variables $x = k_\perp \theta_S r_1$ and recall $\Omega_S \approx \pi \theta_S^2$. Inserting this into Eq. (9), we again obtain Eq. (41). Therefore the two approaches indeed agree in the flat sky limit for a constant spectrum.

Second, we can compare the results numerically for a mask with small enough sky coverage. To do so, we created a polar cap mask of radius 5 deg[9]. The power spectrum of this mask is shown in Fig. 9.

Compared to Fig. 7, we see that power extends to smaller angular scales or higher multipoles. In this case, we found that we had to extend the sum in Eq. (32) to higher multipoles, $\ell_{\rm max} = 250$ to reach 1% level convergence of the covariance prediction. This stays however numerically tractable, through the observation that the ℓ dependence of ${\rm Cov}_\ell^{\rm SSC}$ is very smooth, especially after $\ell_{\rm peak} = 25$. Thus we can sample this dependence only for a small number of logarithmically spaced multipoles and interpolate when computing the sum in Eq. (32).

We found good agreement between the covariance derived from the partial sky formalism and that derived from the flat sky limit Eq. (12). This is visible, for example in Fig. 10 showing both auto-z covariances as a function of $\alpha = j_M + n_M i_M + n_M^2 i_z$, i.e. the same ordering as in Fig. 2.

The partial sky formalism thus successfully recovers the flat sky approximation in the flat sky limit. Numerically, we even found the partial sky formalism to be faster than the full computation of Eq. (12), and only three times slower than the Sij approximation Eq. (13).

---

[9] As the power spectrum of a map is invariant under SO(3), centring the mask on the north pole is simply a convenient choice.

6. Conclusion

SSC, also often called sample variance, is the dominant error for cluster counts at low cluster masses. For instance, Hu & Cohn (2006) have shown that even for a survey radius $\theta = 2.4$ deg, the sample covariance is of the same order of magnitude as shot noise for cluster counts above $\log M = 14.2$, although shot noise dominates for a threshold $\log M = 14.4$. As shot-noise decreases faster with survey area than SSC, careful predictions of SSC become crucial for current and future surveys covering ever larger sky areas. It becomes even more crucial as these surveys are able to probe lower cluster masses through a higher density of galaxy detections. SSC is also crucial to the analysis of galaxy clustering and lensing shear, where it dominates statistical errors on small scales, and for probe combinations as it has been shown to couple probes (Takada & Bridle 2007; Takada & Spergel 2014; Lacasa & Rosenfeld 2016).

In the case of cluster counts covariance, we examined theoretical SSC computation methods in the flat sky limit, comparing two analytical approximations proposed in the literature to a full computation. We found that both approximations underpredict the auto-z covariance by 15% to 30–35%.

We then presented a harmonic expansion method for efficiently and accurately computing SSC for an arbitrary survey window function. We developed the method in the case of cluster counts, but it can be straightforwardly generalised to other probes such as galaxy clustering or lensing shear. Our derived expressions generalise previous full sky and flat sky equations found in the literature, properly reducing to these equations in the corresponding limits. We have cast the final covariance expression from the partial sky formalism in a way that allows easy modification of the survey mask. This is particularly suitable for comparison of surveys, design of survey strategy, and tests of data cuts due to quality selection criteria or systematics.

When applying our partial sky formalism to a mask broadly similar to the DES footprint with $f_{\rm SKY} \sim 10\%$, we found a $\sim -15\%$ cross-z covariance, meaning that the observables in the two redshift bins considered are anti-correlated. Hence the covariance matrix cannot be taken as block diagonal, as is the case for the approximation by Krause & Eifler (2017), which is however restricted to the flat sky limit. We also examined the possibility that the survey area is further reduced by pixel

removal due to bright stars or systematics, for example. We found that this does not have important effects on the structure of the SSC matrix, only rescaling its amplitude by the effective survey area.

The results presented in this article thus render possible the theoretical computation of LSS covariances that account for selection and mask effect and also vary as a function of model parameters, as is the case in CMB analyses, for example. The latter parameter dependence is important in the case of likelihood analysis of cluster constraints, as it allows for self-calibration of the cluster observable-mass relation (Lima & Hu 2005; Hu & Cohn 2006; Baxter et al. 2016). For general probes, it can also improve cosmological parameter constraints from likelihood inference analyses compared to methods that either neglect these effects or fix the covariance from data or simulations, thereby avoiding the risk of fixing the covariance at a potentially incorrect cosmology.

References

Aguena, M., & Lima, M. 2016, Phys. Rev. D, submitted [arXiv:1611.05468]
Astier, P., Guy, J., Regnault, N., et al. 2006, A&A, 447, 31
Baxter, E. J., Rozo, E., Jain, B., Rykoff, E., & Wechsler, R. H. 2016, MNRAS, 463, 205
Bayliss, M. B., Ruel, J., Stubbs, C. W., et al. 2016, ApJS, 227, 3
Beutler, F., Seo, H.-J., Ross, A. J., et al. 2017, MNRAS, 464, 3409
Bleem, L. E., Stalder, B., de Haan, T., et al. 2015, ApJS, 216, 27
Blot, L., Corasaniti, P. S., Amendola, L., & Kitching, T. D. 2016, MNRAS, 458, 4462
Campagne, J.-E., Neveu, J., & Plaszczynski, S. 2017, A&A, 602, A72
Cooray, A., & Sheth, R. 2002, Phys. Rep., 372, 1
Crocce, M., Carretero, J., Bauer, A. H., et al. 2016, MNRAS, 455, 4301
DES Collaboration. 2017, ArXiv e-prints [arXiv:1708.01530]
Dietrich, J. P., Zhang, Y., Song, J., et al. 2014, MNRAS, 443, 1713
Dodelson, S., & Schneider, M. D. 2013, Phys. Rev. D, 88, 063537
Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
Escoffier, S., Cousinou, M.-C., Tilquin, A., et al. 2016, ArXiv e-prints [arXiv:1606.00233]
Giannantonio, T., Fosalba, P., Cawthon, R., et al. 2016, MNRAS, 456, 3213
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Hamilton, A. J. S. 2000, MNRAS, 312, 257
Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, MNRAS, 465, 1454
Hinton, S. R., Kazin, E., Davis, T. M., et al. 2017, MNRAS, 464, 4807
Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2
Hu, W., & Cohn, J. D. 2006, Phys. Rev. D, 73, 067301
Hu, W., & Kravtsov, A. V. 2003, ApJ, 584, 702
Hu, W., Chiang, C.-T., Li, Y., & LoVerde, M. 2016, Phys. Rev. D, 94, 023002
Koester, B. P., McKay, T. A., Annis, J., et al. 2007, ApJ, 660, 221
Krause, E., & Eifler, T. 2017, MNRAS, 470, 2100
Lacasa, F. 2018, A&A, in press, DOI:10.1051/0004-6361/201732343
Lacasa, F., & Kunz, M. 2017, A&A, 604, A104
Lacasa, F., & Rosenfeld, R. 2016, JCAP, 8, 005
Li, Y., Hu, W., & Takada, M. 2014a, Phys. Rev. D, 89, 083519
Li, Y., Hu, W., & Takada, M. 2014b, Phys. Rev. D, 90, 103530
Li, Y., Hu, W., & Takada, M. 2016, Phys. Rev. D, 93, 063507
Lima, M., & Hu, W. 2004, Phys. Rev. D, 70, 043504
Lima, M., & Hu, W. 2005, Phys. Rev. D, 72, 043006
Lima, M., & Hu, W. 2007, Phys. Rev. D, 76, 123013
Miller, C. J., Nichol, R. C., Reichart, D., et al. 2005, AJ, 130, 968
O'Connell, R., Eisenstein, D., Vargas, M., Ho, S., & Padmanabhan, N. 2016, MNRAS, 462, 2681
Pearson, D. W., & Samushia, L. 2016, MNRAS, 457, 993
Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe (Princeton: Princeton University Press)
Percival, W. J., Reid, B. A., Eisenstein, D. J., et al. 2010, MNRAS, 401, 2148
Planck Collaboration XXVII. 2016, A&A, 594, A27
Riess, A. G., Macri, L., Casertano, S., et al. 2009, ApJ, 699, 539
Rykoff, E. S., Rozo, E., Busha, M. T., et al. 2014, ApJ, 785, 104
Rykoff, E. S., Rozo, E., Hollowood, D., et al. 2016, ApJS, 224, 1
Sánchez, C., Carrasco Kind, M., Lin, H., et al. 2014, MNRAS, 445, 1482
Schmidt, F., Lima, M., Oyaizu, H., & Hu, W. 2009, Phys. Rev. D, 79, 083518
Schmidt, F., Jeong, D., & Desjacques, V. 2013, Phys. Rev. D, 88, 023515
Shirasaki, M., Takada, M., Miyatake, H., et al. 2017, MNRAS, 470, 3476
Singh, S., Mandelbaum, R., Seljak, U., Slosar, A., & Vazquez Gonzalez, J. 2017, MNRAS, 471, 3827
Soares-Santos, M., de Carvalho, R. R., Annis, J., et al. 2011, ApJ, 727, 45
Takada, M., & Bridle, S. 2007, New J. Phys., 9, 446
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
Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709
Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878

Appendix A: Redshift-dependent mask

In the case where the survey angular mask W depends on redshift (e.g. when there are significant depth variations in the sky), the definition of the background mode Eq. (5) is simply changed to

$$\delta_b(z) = \frac{1}{\Omega_S(z)} \int d^2\hat{\mathbf{n}}\, W(\hat{\mathbf{n}}, z)\, \delta(r\hat{\mathbf{n}}, z), \tag{A.1}$$

where now the survey solid angle Ω_S depends on redshift.

Then Eq. (25) for $\sigma^2(z_1, z_2)$ is changed to

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

where $C_\ell[W(z_1), W(z_2)]$ is the angular cross-spectrum between the two masks at the two redshifts. This equation can then be integrated over (z_1, z_2) through Eq. (6) to yield the SSC covariance.

We note that in this redshift-dependent case, we can no longer permute the k and z integrals as was done in Sect. 4.3, which yielded a numerically efficient method. Recent advances by Campagne et al. (2017) may however render Eq. (A.2) numerically computable through the AngPow software.

Appendix B: Super-sample covariance for other probes

The main results of this article were derived for the covariance of cluster number counts. However, it is straightforward to generalise the equations to other LSS probes, such as galaxy clustering and lensing shear. In those cases, the covariance equations are simpler for the angular power spectrum C_ℓ than for the angular correlation function. For instance, using Limber's approximation, Lacasa & Rosenfeld (2016) have shown that the SSC of the galaxy angular power spectrum is given by

$$\mathrm{Cov}\left(C_{\ell_1}^{\mathrm{gal}}(i_z), C_{\ell_2}^{\mathrm{gal}}(j_z)\right) = \int dV_{12}\, \frac{\overline{n}_{\mathrm{gal}}(z_1)\,\overline{n}_{\mathrm{gal}}(z_2)}{\Delta N_{\mathrm{gal}}(i_z)\,\Delta N_{\mathrm{gal}}(j_z)}$$
$$\times \frac{\partial P_{\mathrm{gal}}(k_{\ell_1}, z_1)}{\partial \delta_b} \frac{\partial P_{\mathrm{gal}}(k_{\ell_2}, z_2)}{\partial \delta_b}$$
$$\times \sigma^2(z_1, z_2), \tag{B.1}$$

where $k_{\ell_i} = (\ell_i + 1/2)/r(z_i)$.

For lensing shear, we would get a similar equation, replacing the galaxy number density by the lensing selection function and the 3D galaxy power spectrum by the matter power spectrum. The equation can thus be generalised to

$$\mathrm{Cov}\left(C_{\ell_1}^{\alpha}(i_z), C_{\ell_2}^{\alpha}(j_z)\right) = \int dV_{12}\, W_\alpha(z_1)\, W_\alpha(z_2)$$
$$\times \frac{\partial P_\alpha(k_{\ell_1}, z_1)}{\partial \delta_b} \frac{\partial P_\alpha(k_{\ell_2}, z_2)}{\partial \delta_b}$$
$$\times \sigma^2(z_1, z_2), \tag{B.2}$$

where the α index refers to either lensing or galaxy, and W_α is the corresponding selection function. In that case, it is straightforward to generalise Eq. (36), even including the possibility of the cross-covariance between galaxy and shear,

$$\mathrm{Cov}_\ell\left(C_{\ell_1}^{\alpha}(i_z), C_{\ell_2}^{\beta}(j_z)\right) = \frac{1}{2\pi^2} \int k^2\, dk\, P_m(k|z=0)$$
$$\times \Psi_\ell^{\ell_1, \alpha}(k|i_z)\, \Psi_\ell^{\ell_2, \beta}(k|j_z), \tag{B.3}$$

where

$$\Psi_\ell^{\ell_i, \alpha}(k|i_z) \equiv \int_{z\in\mathrm{bin}(i_z)} dV\, G(z)\, W_\alpha(z)\, \frac{\partial P_\alpha(k_{\ell_i}, z)}{\partial \delta_b}\, j_\ell(kr), \tag{B.4}$$

and the cross-covariance between cluster counts and either galaxy or shear,

$$\mathrm{Cov}_\ell\left(N_{\mathrm{cl}}(i_M, i_z), C_{\ell_2}^{\beta}(j_z)\right) = \frac{1}{2\pi^2} \int k^2\, dk\, P_m(k|z=0)$$
$$\times \Psi_\ell^{n_h}(k|i_M, i_z)\, \Psi_\ell^{\ell_2, \beta}(k|j_z), \tag{B.5}$$

which gives us all the equations needed to compute the auto and cross-covariances of cluster counts, galaxy angular power spectrum, and lensing shear power spectrum, i.e. the three main cosmological probes of current and future photometric galaxy surveys.

Appendix C: First multipoles

The spherical Bessel functions $j_n(x)$ obey the recurrence relation

$$j_{n+1}(x) = \frac{2n+1}{x}\, j_n(x) - j_{n-1}(x), \tag{C.1}$$

such that they can be written analytically in terms of sines, cosines and polynomials, given the initial conditions

$$j_{-1}(x) = \frac{\cos(x)}{x}, \tag{C.2}$$

$$j_0(x) = \frac{\sin(x)}{x}. \tag{C.3}$$

The first few spherical Bessel functions are given by

$$j_0(x) = \frac{\sin(x)}{x}, \tag{C.4}$$

$$j_1(x) = \frac{1}{x}\frac{\sin(x)}{x} - \frac{\cos(x)}{x}, \tag{C.5}$$

$$j_2(x) = \left(\frac{3}{x^2} - 1\right)\frac{\sin(x)}{x} - \frac{3}{x}\frac{\cos(x)}{x}. \tag{C.6}$$

We are trying to compute the following integrals:

$$C_\ell^{\mathrm{m}}(z_{12}) = \frac{2}{\pi} \int k^2 dk\, j_\ell(kr_1)\, j_\ell(kr_2)\, P_m(k|z_{12}). \tag{C.7}$$

For $j_\ell(x)$ of the form $A_\ell(x)\sin(x)/x - B_\ell(x)\cos(x)/x$, this yields

$$C_\ell^{\mathrm{m}}(z_{12}) = \frac{2}{\pi} \int k^2 dk\, \frac{P_m(k|z_{12})}{k^2 r_1 r_2}$$
$$\times \frac{1}{2}\Big[(A_\ell(kr_1)A_\ell(kr_2) + B_\ell(kr_1)B_\ell(kr_2))$$
$$\times \cos(k(r_1 - r_2))$$
$$+ (-A_\ell(kr_1)A_\ell(kr_2) + B_\ell(kr_1)B_\ell(kr_2))$$
$$\times \cos(k(r_1 + r_2))$$
$$- (A_\ell(kr_1)B_\ell(kr_2) - B_\ell(kr_1)A_\ell(kr_2))$$
$$\times \sin(k(r_1 - r_2))$$
$$- (A_\ell(kr_1)B_\ell(kr_2) + B_\ell(kr_1)A_\ell(kr_2))$$
$$\times \sin(k(r_1 + r_2))\Big], \tag{C.8}$$

where the four integrals are Fourier (sine or cosine) transforms and can be computed numerically with FFTs. Let us note

$$I_c^+(f(k, r_1, r_2)) \equiv \frac{2}{\pi} \int k^2 \mathrm{d}k \, \frac{P_m(k|z_{12})}{k^2 r_1 r_2} f(k, r_1, r_2)$$
$$\times \cos(k(r_1 + r_2)), \tag{C.9}$$

$$I_c^-(f(k, r_1, r_2)) \equiv \frac{2}{\pi} \int k^2 \mathrm{d}k \, \frac{P_m(k|z_{12})}{k^2 r_1 r_2} f(k, r_1, r_2)$$
$$\times \cos(k(r_1 - r_2)), \tag{C.10}$$

$$I_s^+(f(k, r_1, r_2)) \equiv \frac{2}{\pi} \int k^2 \mathrm{d}k \, \frac{P_m(k|z_{12})}{k^2 r_1 r_2} f(k, r_1, r_2)$$
$$\times \sin(k(r_1 + r_2)), \tag{C.11}$$

$$I_s^-(f(k, r_1, r_2)) \equiv \frac{2}{\pi} \int k^2 \mathrm{d}k \, \frac{P_m(k|z_{12})}{k^2 r_1 r_2} f(k, r_1, r_2)$$
$$\times \sin(k(r_1 - r_2)). \tag{C.12}$$

Then Eq. (C.8) can be rewritten as

$$C_\ell^{\mathrm{m}}(z_{12}) = \frac{1}{2} \Big[I_c^-(A_{\ell,1}A_{\ell,2} + B_{\ell,1}B_{\ell,2})$$
$$+ I_c^+(-A_{\ell,1}A_{\ell,2} + B_{\ell,1}B_{\ell,2})$$
$$- I_s^-(A_{\ell,1}B_{\ell,2} - B_{\ell,1}A_{\ell,2})$$
$$- I_s^+(A_{\ell,1}B_{\ell,2} + B_{\ell,1}A_{\ell,2}) \Big], \tag{C.13}$$

with $A_{\ell,i} = A_\ell(kr_i)$. For $\ell = 0$ we have $A_0 = 1$ and $B_0 = 0$, thus

$$C_0^{\mathrm{m}}(z_{12}) = \frac{1}{2} \Big[I_c^-(1) - I_c^+(1) \Big]. \tag{C.14}$$

For $\ell = 1$ we have $A_1 = 1/x$ and $B_1 = 1$. Therefore

$$C_1^{\mathrm{m}}(z_{12}) = \frac{1}{2} \Big[I_c^- \Big(\frac{1}{k^2 r_1 r_2} + 1 \Big) + I_c^+ \Big(-\frac{1}{k^2 r_1 r_2} + 1 \Big)$$
$$- I_s^- \Big(\frac{1}{kr_1} - \frac{1}{kr_2} \Big) - I_s^+ \Big(\frac{1}{kr_1} + \frac{1}{kr_2} \Big) \Big]. \tag{C.15}$$

If we define $I_{c/s,n}^{+/-} \equiv I_{c/s}^{+/-}(1/k^n)$, the above equations yield

$$C_0^{\mathrm{m}}(z_{12}) = \frac{1}{2} \Big[I_{c,0}^- - I_{c,0}^+ \Big], \tag{C.16}$$

$$C_1^{\mathrm{m}}(z_{12}) = \frac{1}{2} \Big[I_{c,0}^- + I_{c,0}^+ - I_{s,1}^- \Big(\frac{1}{r_1} - \frac{1}{r_2} \Big)$$
$$- I_{s,1}^+ \Big(\frac{1}{r_1} + \frac{1}{r_2} \Big) + \frac{I_{c,2}^- - I_{c,2}^+}{r_1 r_2} \Big]. \tag{C.17}$$

For $\ell = 2$ we have $A_2 = 3/x^2 - 1$ and $B_2 = 3/x$, thus

$$C_2^{\mathrm{m}}(z_{12}) = \frac{1}{2} \Big[I_c^- \Big(\frac{9}{k^4 r_1^2 r_2^2} + \frac{3}{k^2} \Big(\frac{3}{r_1 r_2} - \frac{1}{r_1^2} - \frac{1}{r_2^2} \Big) + 1 \Big)$$
$$+ I_c^+ \Big(-\frac{9}{k^4 r_1^2 r_2^2} + \frac{3}{k^2} \Big(\frac{3}{r_1 r_2} + \frac{1}{r_1^2} + \frac{1}{r_2^2} \Big) - 1 \Big)$$
$$- I_s^- \Big(\frac{9}{k^3 r_1 r_2} + \frac{3}{k} \Big) \times \Big(\frac{1}{r_1} - \frac{1}{r_2} \Big)$$
$$- I_s^+ \Big(\frac{9}{k^3 r_1 r_2} - \frac{3}{k} \Big) \times \Big(\frac{1}{r_1} + \frac{1}{r_2} \Big) \Big] \tag{C.18}$$

$$= \frac{1}{2} \Big[I_{c,0}^- - I_{c,0}^+ - 3\, I_{s,1}^- \Big(\frac{1}{r_1} - \frac{1}{r_2} \Big)$$
$$+ 3\, I_{s,1}^+ \Big(\frac{1}{r_1} + \frac{1}{r_2} \Big) + 3\, I_{c,2}^- \Big(\frac{3}{r_1 r_2} - \frac{1}{r_1^2} - \frac{1}{r_2^2} \Big)$$
$$+ 3\, I_{c,2}^+ \Big(\frac{3}{r_1 r_2} + \frac{1}{r_1^2} + \frac{1}{r_2^2} \Big) - \frac{9\, I_{s,3}^-}{r_1 r_2} \Big(\frac{1}{r_1} - \frac{1}{r_2} \Big)$$
$$- \frac{9\, I_{s,3}^+}{r_1 r_2} \Big(\frac{1}{r_1} + \frac{1}{r_2} \Big) + \frac{9 \big(I_{c,4}^- - I_{c,4}^+ \big)}{r_1^2 r_2^2} \Big]. \tag{C.19}$$

We note that A_ℓ and B_ℓ follow the same recurrence relation Eq. (C.1) as the spherical Bessel functions j_n. As such we can look at whether there is a recurrence relation that would allow us to get the analytical formula for a general $C_\ell^{\mathrm{m}}(z_{12})$. Our efforts in this direction have shown only partially fruitful and are described below. We have

$$C_{\ell+1}^{\mathrm{m}}(z_{12}) = \frac{1}{2} \Big[I_c^-(A_{\ell+1,1}A_{\ell+1,2} + B_{\ell+1,1}B_{\ell+1,2})$$
$$+ I_c^+(-A_{\ell+1,1}A_{\ell+1,2} + B_{\ell+1,1}B_{\ell+1,2})$$
$$- I_s^-(A_{\ell+1,1}B_{\ell+1,2} - B_{\ell+1,1}A_{\ell+1,2})$$
$$- I_s^+(A_{\ell+1,1}B_{\ell+1,2} + B_{\ell+1,1}A_{\ell+1,2}) \Big]$$
$$\equiv \frac{1}{2} \big(C_{\ell+1}^{c,-} + C_{\ell+1}^{c,+} - C_{\ell+1}^{s,-} - C_{\ell+1}^{s,+} \big), \tag{C.20}$$

and then

$$C_{\ell+1}^{c,-} = I_c^- \Big[\Big(\frac{2\ell+1}{kr_1} A_{\ell,1} - A_{\ell-1,1} \Big) \Big(\frac{2\ell+1}{kr_2} A_{\ell,2} - A_{\ell-1,2} \Big)$$
$$+ \Big(\frac{2\ell+1}{kr_1} B_{\ell,1} - B_{\ell-1,1} \Big) \Big(\frac{2\ell+1}{kr_2} B_{\ell,2} - B_{\ell-1,2} \Big) \Big] \tag{C.21}$$

$$= C_{\ell-1}^{c,-} + \frac{(2\ell+1)^2}{r_1 r_2} I_c^- \big[(A_{\ell,1}A_{\ell,2} + B_{\ell,1}B_{\ell,2})/k^2 \big]$$
$$- (2\ell+1)\, I_c^- \big[(A_{\ell-1,1}A_{\ell,2} + B_{\ell-1,1}B_{\ell,2})/kr_1$$
$$+ (1 \leftrightarrow 2) \big]. \tag{C.22}$$

Now we define the following quantities

$$D_\ell^{c,-} \equiv I_c^- \big[(A_{\ell-1,1}A_{\ell,2} + B_{\ell-1,1}B_{\ell,2})/kr_1 + (1 \leftrightarrow 2) \big], \tag{C.23}$$

$$E_\ell^{c,-} \equiv I_c^- \big[(A_{\ell-1,1}A_{\ell,2} + B_{\ell-1,1}B_{\ell,2})/kr_2 + (1 \leftrightarrow 2) \big], \tag{C.24}$$

and decompose $C_\ell^{c,-}$, $D_\ell^{c,-}$ and $E_\ell^{c,-}$ onto the basis $I_{c,i}^-$

$$C_\ell^{c,-} = \sum_{i=0}^{+\infty} \alpha_{\ell,i}^{c,-}(r_1, r_2) I_{c,i}^-, \tag{C.25}$$

$$D_\ell^{c,-} = \sum_{i=0}^{+\infty} \beta_{\ell,i}^{c,-}(r_1, r_2) I_{c,i}^-, \tag{C.26}$$

$$E_\ell^{c,-} = \sum_{i=0}^{+\infty} \gamma_{\ell,i}^{c,-}(r_1, r_2) I_{c,i}^-. \tag{C.27}$$

Then Eq. (C.22) gives a first recurrence relation

$$\alpha_{\ell+1,i}^{c,-} = \alpha_{\ell-1,i}^{c,-} + \frac{(2\ell+1)^2}{r_1 r_2} \alpha_{\ell,i-2}^{c,-} - (2\ell+1)\beta_{\ell,i}^{c,-}. \tag{C.28}$$

By computing $D_{\ell+1}^{c,-}$ and $E_{\ell+1}^{c,-}$, we can see that we have the two other recurrence relations

$$\beta_{\ell+1,i}^{c,-} = (2\ell+1)\frac{2}{r_1 r_2}\alpha_{\ell,i-2}^{c,-} - \gamma_{\ell,i}^{c,-}, \qquad (C.29)$$

$$\gamma_{\ell+1,i}^{c,-} = (2\ell+1)\left(\frac{1}{r_1^2}+\frac{1}{r_2^2}\right)\alpha_{\ell,i-2}^{c,-} - \beta_{\ell,i}^{c,-}. \qquad (C.30)$$

The system is closed through the initial conditions

$$\alpha_{0,0}^{c,-} = 1 \qquad \alpha_{0,2}^{c,-} = 0 \qquad \alpha_{0,i}^{c,-} = 0 \text{ for } i \geq 3, \qquad (C.31)$$

$$\alpha_{1,0}^{c,-} = 1 \qquad \alpha_{1,2}^{c,-} = \frac{1}{r_1 r_2} \qquad \alpha_{1,i}^{c,-} = 0 \text{ for } i \geq 3, \qquad (C.32)$$

$$\beta_{1,0}^{c,-} = 0 \qquad \beta_{1,2}^{c,-} = \frac{2}{r_1 r_2} \qquad \beta_{1,i}^{c,-} = 0 \text{ for } i \geq 3, \qquad (C.33)$$

$$\gamma_{1,0}^{c,-} = 0 \qquad \gamma_{1,2}^{c,-} = \left(\frac{1}{r_1^2}+\frac{1}{r_2^2}\right) \qquad \gamma_{1,i}^{c,-} = 0 \text{ for } i \geq 3. \qquad (C.34)$$

The following properties can easily be shown by recurrence:

$$\forall \ell \qquad \alpha_{\ell,0}^{c,-} = 1, \qquad \beta_{\ell,0}^{c,-} = 0, \qquad \gamma_{\ell,0}^{c,-} = 0. \qquad (C.35)$$

$$\forall \ell \qquad \alpha_{\ell,i}^{c,-} = 0, \qquad \text{if } i \text{ odd or } i > 2\ell. \qquad (C.36)$$

$$\forall \ell \qquad \alpha_{\ell,2\ell}^{c,-} = \frac{\prod_{n=0}^{\ell-1}(2n+1)^2}{(r_1 r_2)^\ell} = \frac{(2\ell-1)!^2}{(4r_1 r_2)^\ell \, \ell!^2}. \qquad (C.37)$$

With a bit more work, we can in principle solve for $\beta_{\ell,2}^{c,-}$ and $\gamma_{\ell,2}^{c,-}$; since $\alpha_{\ell,0}^{c,-} = 1$,

$$\beta_{\ell+1,2}^{c,-} = (2\ell+1)\frac{2}{r_1 r_2} - \gamma_{\ell,2}^{c,-}, \qquad (C.38)$$

$$\gamma_{\ell+1,2}^{c,-} = (2\ell+1)\left(\frac{1}{r_1^2}+\frac{1}{r_2^2}\right) - \beta_{\ell,2}^{c,-}. \qquad (C.39)$$

This can be put in the form

$$X_{\ell+1} = MX_\ell + V_\ell, \qquad (C.40)$$

for

$$X_\ell = \begin{pmatrix} \beta_{\ell,2}^{c,-} \\ \gamma_{\ell,2}^{c,-} \end{pmatrix}, \qquad (C.41)$$

$$M = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \qquad (C.42)$$

$$V_\ell = (2\ell+1)\begin{pmatrix} \frac{2}{r_1 r_2} \\ \frac{1}{r_1^2}+\frac{1}{r_2^2} \end{pmatrix} = (2\ell+1)X_1. \qquad (C.43)$$

Finally, Eq. (C.40) has the following solution:

$$X_\ell = \sum_{k=0}^{\ell-1} M^{\ell-k-1}(2k+1)X_1, \qquad (C.44)$$

which in principle gives the solution for $\beta_{\ell,2}^{c,-}$ and $\gamma_{\ell,2}^{c,-}$. We can insert this into Eq. (C.28) for $i = 2$ to get a closed recurrence relation for $\alpha_{\ell,2}^{c,-}$ and solve for it. In turn, this can be inserted into Eqs. (C.29) and (C.30) for $i = 4$ to get a closed recurrence relation for $(\beta_{\ell,4}^{c,-}, \gamma_{\ell,4}^{c,-})$, and so on. So this represents a procedure for solving this set of equations and get analytical expressions for $C_\ell^{c,-}$. Similarly, we could derive recurrence relations for $C_\ell^{c,+}$, $C_\ell^{s,-}$, and $C_\ell^{s,+}$. In practice this represents a daunting task, as even the formula for $\ell = 2$ (Eq. (C.19)) is already cumbersome.

Moreover, we may expect that these analytical formulae are doomed to present numerical instability at high ℓ. Indeed, the expansion of the spherical Bessel functions in terms of powers of $1/x$ as in Eqs. (C.4)–(C.6) is ill-advised at high ℓ, as it leads to delicate cancellations of the numerous terms, in particular for $x \leq \ell$. We thus expect that our analytical formulae for C_ℓ also leads to delicate cancellations that may be numerically unstable. Another way of seeing this is to notice that, at high ℓ, we have from Limber's approximation that $C_\ell \propto P_m[k = (\ell+1/2)/r(z)]$, and is thus a decreasing function of ℓ. However we saw that $\alpha_{\ell,0}^{c,-} = 1$ and thus C_ℓ always contains a term $I_{c,0}^-/2 \approx C_0$. This term hence needs to be (at least partially) cancelled by higher order terms, and this cancellation needs to be increasingly precise at high ℓ, since $P_m(k)$ is a steep decreasing function of k.

Given these issues, in practice, we implemented this low multipole method only for $\ell = 0, 1, 2$, and used these to check our results obtained from the numerical method described in Sect. 4.3.