Astronomy
&
Astrophysics

A new data compression method and its application to cosmic shear analysis

Marika Asgari[1] and Peter Schneider[2]

[1] SUPA, Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK
   e-mail: ma@roe.ac.uk
[2] Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany

ABSTRACT

Context. Future large scale cosmological surveys will provide huge data sets whose analysis requires efficient data compression. In particular, the calculation of accurate covariances is extremely challenging with an increasing number of observables used in the statistical analysis.

Aims. The primary aim of this paper is to introduce a formalism for achieving efficient data compression, based on a local expansion of the observables around a fiducial cosmological model. We specifically apply and test this approach for the case of cosmic shear statistics. In addition, we study how well band powers can be obtained from measuring shear correlation functions over a finite interval of separations.

Methods. We demonstrate the performance of our approach, using a Fisher analysis on cosmic shear tomography described in terms of E-/B-mode separating statistics (COSEBIs).

Results. We show that our data compression is highly effective in extracting essentially the full cosmological information from a strongly reduced number of observables. Specifically, the number of observables needed decreases by at least one order of magnitude relative to the COSEBIs, which already compress the data substantially compared to the shear two-point correlation functions. The efficiency appears to be affected only slightly if a highly inaccurate covariance is used for defining the compressed data vector, showing the robustness of the method. In addition, we show the strong limitations on the possibility of constructing top-hat filters in Fourier space, for which the real-space analog has a finite support, yielding strong bounds on the accuracy of band power estimates.

Conclusions. We conclude that efficient data compression is achievable and that the number of compressed data points depends on the number of model parameters. Furthermore, a band convergence power spectrum inferred from a finite angular range cannot be accurately estimated. The error on an estimated band power is larger for a narrower filter and a narrower angular range, which for relevant cases can be as large as 10% for $\Delta \ell / \ell \sim 0.1$.

Key words. gravitational lensing: weak – methods: numerical – methods: statistical – cosmology: observations

1. Introduction

Future cosmological surveys are faced with the difficulty of extracting cosmological parameters from their wealth of observables. Taking *Euclid*[1] as an example, statistics to be obtained from the data include second-order shear statistics across several populations of source galaxies, which – using the common usage – will be termed "redshift bins" throughout this paper. As shown in Schneider et al. (2010); and Asgari et al. (2012), the COSEBIs (Complete Orthogonal E-/B-mode Integrals) form appropriate combinations of the shear two-point correlation functions $\xi_\pm(\theta)$ that cleanly separate E- and B-mode shear (see, e.g., Crittenden et al. 2002; Schneider et al. 2002b). In addition, COSEBIs are highly efficient in terms of data compression, since essentially all cosmological information is contained in a few COSEBIs modes (see, e.g., Kilbinger et al. 2013; Huff et al. 2014, for applications of COSEBIs to cosmic shear data sets).

The efficiency of data compression decreases, however, if several populations of sources are used. For example, with about ten redshift bins, the total number of COSEBIs that should be used to extract cosmological information is approximately 500. Furthermore, higher order shear information contains additional,

valuable information – both regarding cosmological parameters and for calibrating the shear data – and should be taken into account. Since third-order shear statistics depend on three variables (say, three sides of a triangle), and combinations of three red-shift bins, the number of observables for third-order shear statistics that need to be considered is almost certainly considerably larger than for second-order shear statistics. Furthermore, shear-peak statistics have been shown to yield powerful constraints and should likewise be considered (see, e.g., Marian et al. 2013, and references therein). Therefore, the number of pure shear observables will be several thousand, although the number of cosmological parameters to be determined is around a dozen.

In practice, issues are even more complicated in that astrophysical and other systematics need to be accounted for. For example, the effects of intrinsic alignments (see, e.g., Joachimi & Bridle 2010, and references therein) need to be mitigated by including further observables, i.e., the galaxy-galaxy lensing signal and the galaxy correlation functions. Even if one uses a COSEBIs-like data compression for them (e.g., Eifler et al. 2014), the number of redshift combinations will still lead to a strongly enhanced number of observables.

One of the major difficulties in analyzing this data is to determine the expectation values for these observables as a function of the parameters and, in particular, to estimate their covariance

---

matrix. If one determines the covariance as a sample variance of different numerical realizations, one needs many more realizations than the dimension of the data vector in order to get a reliable estimate of the covariance matrix and its inverse (see, e.g., Hartlap et al. 2007). Because of this difficulty, data compression is mandatory for any analysis of survey data. Taylor & Joachimi (2014) have found relations between the accuracy of the estimated covariance, the number of simulations, and observables for Gaussian distributed data and underlying parameters. Their work also suggests that data compression is essential for reducing the number of simulations required for reaching a certain accuracy in the covariance.

In this paper, we suggest a form of data compression that is based on the sensitivity of the various observables to the parameters that are to be estimated. The cosmological parameters currently are and, until the launch of *Euclid*, will be even more strongly constrained, and thus only a relatively small volume in parameter space needs to be explored[2]. We therefore assume that the relevant parameter region is small, which allows us to define linear combinations of observables based on a low-order Taylor expansion of the dependence of these observables on parameters, which should contain almost all the cosmological information in the data.

In the following section, we introduce our data compression formalism for general observables (statistics). We then apply this method to COSEBIs in Sect. 3 to study how this strategy works for compressing it. In Sect. 4 we specify our cosmological model, which is used for the results section. In Sect. 5 we first illustrate the weight functions for the compressed statistics made of COSEBIs, and then use a Fisher formalism to explore the efficiency of the compressed versus regular COSEBIs. Section 6 is dedicated to mimicking a band power spectrum using linear combinations of COSEBIs. Finally we conclude in Sect. 7.

2. Formalism

We let $\hat{X}_n$ be the statistics obtained from the data, $1 \le n \le N$, with expectation value $\langle \hat{X}_n \rangle = X_n(\phi_\mu)$, where the ϕ_μ, $1 \le \mu \le P$, denotes the parameters of the model, including the cosmological parameters and others. Assuming that the uncertainty in the parameters is "small", we consider an expansion of the functions $X_n(\phi_\mu)$ around the fiducial value $X_n^{\mathrm{f}} = X_n(\phi_\mu^{\mathrm{f}})$,

$$X_n(\phi_\mu) = X_n^{\mathrm{f}} + \mathsf{D}_{n\mu} p_\mu + \frac{1}{2}\mathsf{Z}_{n\mu\nu} p_\mu p_\nu, \tag{1}$$

where $p_\mu = \phi_\mu - \phi_\mu^{\mathrm{f}}$, and

$$\mathsf{D}_{n\mu} = \left(\frac{\partial X_n}{\partial \phi_\mu}\right)_{|\phi_\kappa^{\mathrm{f}}}; \quad \mathsf{Z}_{n\mu\nu} = \left(\frac{\partial^2 X_n}{\partial \phi_\mu \, \partial \phi_\nu}\right)_{|\phi_\kappa^{\mathrm{f}}} \tag{2}$$

are the first and second derivatives of the expectation values with respect to the model parameters, taken at the fiducial point in parameter space. Here and below, summation over repeated indices is implied, unless noted otherwise.

We assume that the likelihood $\mathcal{L}(\chi^2)$ is a monotonically decreasing function of

$$\chi^2 = \left[\hat{X}_m - X_m(\phi_\mu)\right] \left(\mathsf{C}^{-1}\right)_{mn} \left[\hat{X}_n - X_n(\phi_\mu)\right], \tag{3}$$

[2] Additional parameters, needed to parametrize intrinsic alignment effects, may be less well constrained from independent data sets or theoretical models.

where C is the covariance matrix of the observables $\hat{X}_n$. Maximizing the likelihood then requires finding the minimum of χ^2 with respect to the parameters; using Eq. (1), we obtain

$$\begin{aligned}\frac{\partial \chi^2}{\partial p_\kappa} = &-2\left[\mathsf{D}_{m\kappa} + \mathsf{Z}_{m\kappa\mu} p_\mu\right]\left(\mathsf{C}^{-1}\right)_{mn} \\ &\times \left[\hat{X}_n - X_n^{\mathrm{f}} - \mathsf{D}_{n\mu} p_\mu - \frac{1}{2}\mathsf{Z}_{n\mu\nu} p_\mu p_\nu\right] = 0.\end{aligned} \tag{4}$$

In this equation, we have neglected the dependence of the covariance matrix on the parameters, either because C is determined from the data itself or because the dependence of C on the parameters is assumed to be weak. From Eq. (4), we see that the determination of the parameters p_μ does involve the observables $\hat{X}_n$ but only in the linear combinations

$$\hat{F}_\kappa := \mathsf{D}_{m\kappa}\left(\mathsf{C}^{-1}\right)_{mn}\hat{X}_n; \quad \hat{S}_{\kappa\nu} := \mathsf{Z}_{m\kappa\nu}\left(\mathsf{C}^{-1}\right)_{mn}\hat{X}_n, \tag{5}$$

with expectation value

$$F_\kappa := \mathsf{D}_{m\kappa}\left(\mathsf{C}^{-1}\right)_{mn}X_n; \quad S_{\kappa\nu} := \mathsf{Z}_{m\kappa\nu}\left(\mathsf{C}^{-1}\right)_{mn}X_n. \tag{6}$$

Thus, the expansion of the expectation values of the original observables $\hat{X}_n$ around a fiducial model motivates the definition of linear combinations of observables that contain all the information about the parameters ϕ_κ, provided the second-order expansion is accurate. The set (5) of $P + P(P+1)/2 = P(P+3)/2$ observables thus is expected to allow for an efficient data compression (note that $\hat{S}_{\mu\nu} = \hat{S}_{\nu\mu}$).

To obtain the new observables $\hat{F}_\kappa$ and $\hat{S}_{\kappa\nu}$, one first needs to estimate the covariance C of the original observables that, owing to the high dimensionality in future cosmological surveys, provides a real challenge. However, the covariance C is needed here for the definition of appropriate combinations of observables, and not for parameter estimates. As a result, an approximation for C may be expected to be sufficient for this purpose. Disregarding the parameter dependence of C in the derivation of Eq. (4) provides such an approximation, which avoids the need to obtain a large covariance matrix for more than one cosmological model. If the approximation for C deviates substantially from the true covariance, we expect that the new observables do not contain the full information about the parameters, since they deviate from the "optimal" combination of the original $\hat{X}_n$. Hence, the better the initial estimate of C, the more efficient the new observables will be.

We therefore propose a strategy for first obtaining an approximation for the covariance C, based on which of the new observables $\hat{F}_\kappa$ and $\hat{S}_{\kappa\nu}$ are defined. The number of these observables is substantially smaller than the original ones, so an accurate estimation of their covariance can be obtained from fewer simulations compared to C. On the other hand, the number of new observables is substantially larger than the number of parameters, which is expected to mitigate the choice of non-optimal combinations from an approximate form of C. It is for this reason that we consider the second-order derivatives of the original observables; the first-order ones coincide with that of the Karhunen–Loève method for the case of known covariance (see, e.g., Tegmark et al. 1997).

We now combine the new observables $\hat{F}_\kappa$ and $\hat{S}_{\kappa\nu}$ into the $N' = P(P+3)/2$ compressed quantities $\hat{X}_i^{\mathrm{c}}$. According to Eq. (5), we can write

$$\hat{X}^{\mathrm{c}} = \mathbf{H}\mathsf{C}^{-1}\hat{X} \equiv \mathbf{B}\hat{X}, \tag{7}$$

where we use vectorial notation for the $\hat{X}_i^c$ and $\hat{X}_n$. The $N' \times N$ (rows × columns) matrix $\mathbf{H}$ is given in terms of first and second partial derivatives of the functions $X_n(\phi_\kappa)$ at the fiducial point in parameter space and $\mathbf{B} = \mathbf{H}\mathbf{C}^{-1}$ is the compression matrix. Accordingly, the covariance matrix of $\hat{X}^c$ is given as

$$\mathbf{C}^c = \mathbf{B}\mathbf{C}\mathbf{B}^t, \qquad (8)$$

where the superscript "t" denotes the transpose of a matrix. The χ^2-function in terms of the new observables is

$$\chi^2 = \left[\hat{X}^c - X^c\right]^t (\mathbf{C}^c)^{-1} \left[\hat{X}^c - X^c\right]. \qquad (9)$$

From what was discussed above, the covariance $\mathbf{C}^c$ should be calculated from $\mathbf{C}$ only if an accurate estimate of the latter can be obtained; in general, it will be much more practical to determine $\mathbf{C}^c$ directly, such as from simulations.

For completeness, we mention for the special case that the covariance $\mathbf{C}$ *can indeed* be determined accurately, which we can solve Eq. (4) for the parameters $\phi_\mu = p_\mu + \phi_\mu^f$. Writing it in terms of the new observables, Eq. (4) becomes

$$\Delta F_\kappa + \Delta S_{\kappa\nu} p_\nu = (D_{m\kappa} + Z_{m\kappa\nu} p_\nu)\left(C^{-1}\right)_{mn}\left(D_{n\nu} p_\nu + \frac{1}{2} Z_{n\mu\nu} p_\mu p_\nu\right), \qquad (10)$$

with $\Delta F_\kappa = \hat{F}_\kappa - F_\kappa^f$, $\Delta S_{\kappa\nu} = \hat{S}_{\kappa\nu} - S_{\kappa\nu}^f$. If we then expand $p_\mu = p_\mu^{(1)} + p_\mu^{(2)}$, where $p_\mu^{(1)}$ ($p_\mu^{(2)}$) is first (second) order in the ΔF_κ, $\Delta S_{\kappa\nu}$, we obtain to first order

$$\Delta F_\kappa = D_{m\kappa}\left(C^{-1}\right)_{mn} D_{n\nu} p_\nu^{(1)} \equiv U_{\kappa\nu} p_\nu^{(1)}, \qquad (11)$$

from which we can easily obtain $p_\nu^{(1)}$ from the inverse of the symmetric matrix $\mathbf{U}$, $p_\nu^{(1)} = \left(U^{-1}\right)_{\nu\mu} \Delta F_\mu$. The second-order terms lead to the equation

$$\Delta S_{\kappa\nu} p_\nu^{(1)} = U_{\kappa\nu} p_\nu^{(2)} + \left(\frac{1}{2} G_{\kappa\mu\nu} + G_{\mu\kappa\nu}\right) p_\mu^{(1)} p_\nu^{(1)}, \qquad (12)$$

where we defined

$$G_{\kappa\mu\nu} := D_{m\kappa}\left(C^{-1}\right)_{mn} Z_{n\mu\nu}. \qquad (13)$$

With the foregoing solution for $p_\mu^{(1)}$ and the inverse of $\mathbf{U}$, this can be immediately solved for $p_\mu^{(2)}$.

3. Application to COSEBIs

We now apply the method of the previous section to specific statistics for cosmic shear measurements, the COSEBIs (see Schneider et al. 2010). They provide a complete representation of the shear two-point correlation functions (2PCFs) in a given finite interval of angular scales, chosen such that they cleanly separate between E and B modes (Crittenden et al. 2002; Schneider et al. 2002b). In our previous work (Asgari et al. 2012), we showed that COSEBIs also provide an efficient means of data compression, since the full cosmological information contained in the 2PCFs can be recovered with a small number of COSEBIs. However, in the case of several redshift bins for the source galaxies, the number of components grows with the number of tomographic redshift bins, r, by a factor of $r(r + 1)/2$. In this section we use the formalism explained in Sect. 2 to obtain a way to compress the number of relevant statistical quantities and

compare the results with a full COSEBIs analysis. The E-mode COSEBIs are related to the 2PCFs via

$$E_n^{(ij)} = \frac{1}{2} \int_{\theta_{\min}}^{\theta_{\max}} d\vartheta\,\vartheta\,\left[T_{+n}(\vartheta)\,\xi_+^{(ij)}(\vartheta) + T_{-n}(\vartheta)\,\xi_-^{(ij)}(\vartheta)\right], \qquad (14)$$

where $1 \leq i, j \leq r$ label the redshift bins considered. The COSEBIs are defined for a given range of angular separations, $[\theta_{\min}, \theta_{\max}]$, i.e., the $T_{\pm n}(\vartheta)$ are zero outside this interval. They form a complete basis for all filter functions that are defined on a finite angular range and satisfy the conditions

$$\int_{\theta_{\min}}^{\theta_{\max}} d\vartheta\,\vartheta\,T_+(\vartheta) = 0 = \int_{\theta_{\min}}^{\theta_{\max}} d\vartheta\,\vartheta^3\,T_+(\vartheta), \qquad (15)$$

which are the necessary and sufficient conditions for separating the E- and B-modes obtained from the shear two-point correlation function measured on a finite interval and for removing ambiguous E-/B-modes (Schneider & Kilbinger 2007). As a result any allowed filter function is a linear combination of them. E-mode COSEBIs are related to the power spectrum by

$$E_n^{(ij)} = \int_0^\infty \frac{d\ell\,\ell}{2\pi} P_E^{(ij)}(\ell) W_n(\ell), \qquad (16)$$

where $P_E^{(ij)}$ is the E-mode convergence cross-power spectra of redshift bins i and j and

$$W_n(\ell) = \int_{\theta_{\min}}^{\theta_{\max}} d\vartheta\,\vartheta\,T_+(\vartheta)\,J_0(\vartheta\ell), \qquad (17)$$

where J_0 is the zeroth-order Bessel function of the first kind (see Schneider et al. 2010; and Asgari et al. 2012, where the filters are defined and shown).

In the following, we use the logarithmic COSEBIs, which yield a more efficient data compression than the linear COSEBIs. The Log-COSEBIs $T_{+n}(\vartheta)$ filters are polynomials in $\ln(\vartheta)$ (see Schneider et al. 2010); i.e., they have more oscillations on small scales so are more sensitive to variations in the shear 2PCFs on those scales. As it turned out, an approximately uniform distribution of roots of the weight function on logarithmic angular scales covers the cosmological information in the shear 2PCFs with a smaller number of components.

To apply the method of the past section for obtaining a compressed version of COSEBIs, we need to find their (approximate) covariance matrix for a given cosmology, in addition to their first- and second-order derivatives with respect to the cosmological parameters. The new set of statistics are related to the COSEBIs via the compression matrix, $\mathbf{B}$, defined before in Eq. (7),

$$E_\mu^c = B_{\mu\mathcal{N}} E_{\mathcal{N}} = B_{\mu nij} E_n^{(ij)}, \qquad (18)$$

where the new index

$$\mathcal{N} = \left[(i-1)r - \frac{(i-1)(i-2)}{2} + (j-1)\right] n_{\max} + n \qquad (19)$$

is a combination of the three indices i, j and n, and $n_{\max}$ is the maximum order of COSEBIs considered, and r is the total number of redshift bins.

Table 1. Fiducial cosmological parameters consistent with the WMAP 7-years results, and the underlying true parameters consistent with *Planck*.

	σ_8	$\Omega_{\rm m}$	Ω_Λ	w_0	$n_{\rm s}$	h	$\Omega_{\rm b}$
Fiducial	0.8	0.27	0.73	−1.0	0.97	0.70	0.045
True	0.83	0.31	0.68	−1.1	0.96	0.67	0.049

Notes. The normalization of the power spectrum, σ_8, is the standard deviation of perturbations in a sphere of radius $8\,h^{-1}$ Mpc today. $\Omega_{\rm m}$, Ω_Λ, and $\Omega_{\rm b}$ are the matter, the dark energy and the baryonic matter density parameters, respectively. w_0 is the dark energy equation of state parameter, which is equal to the ratio of dark energy pressure to its density. The spectral index $n_{\rm s}$ is the power of the initial power spectrum. The dimensionless Hubble parameter, h, characterizes the rate of expansion today.

Table 2. Parameters of a fiducial large future survey.

z-distribution parameters					Survey parameters		
α	β	z_0	$z_{\min}$	$z_{\max}$	A	σ_ϵ	$\bar{n}$
2.0	1.5	0.71	0.0	2.0	20000	0.3	35

Notes. α, β, and z_0 characterize the total redshift distribution of sources, while $z_{\min}$ and $z_{\max}$ indicate the minimum and the maximum redshifts of the sources use for the cosmic shear analysis. Here A is the survey area in units of deg^2, σ_ϵ is the galaxy intrinsic ellipticity dispersion, and $\bar{n}$ is the mean number density of sources per square arcminute in the field.

4. Cosmological model, survey parameters, and covariance

A cold dark matter (CDM) cosmological models with dynamical dark energy, characterized by its equation-of-state parameter, w_0, is used throughout this work (for references to wCDM models, see Peebles & Ratra 2003, and references therein). Table 1 contains the two sets of parameter values considered here. The fiducial model is used for obtaining the compressed COSEBIs (CCOSEBIs hereafter), while the assumed "true" underlying cosmology is slightly different. That means that we calculate the CCOSEBIs according to the equations of Sect. 2, using the covariance and parameter derivatives of the COSEBIs, **C**, **D**, and **Z**, for the fiducial cosmology, but these new observables $E_\mu^{\rm c}$ are applied using the "true" cosmological model. The linear matter power spectrum is calculated using the Bond & Efstathiou (1984) transfer function and a primordial power-law power spectrum with spectral index $n_{\rm s}$. The halo fit formula of Smith et al. (2003) is adopted for nonlinear scales.

For a cosmic shear analysis, we need the survey parameters and the redshift distribution of the galaxies. The latter is characterized by (see, e.g., Brainerd et al. 1996)

$$p(z) \propto \left(\frac{z}{z_0}\right)^\alpha \exp\left[-\left(\frac{z}{z_0}\right)^\beta\right], \tag{20}$$

for $z_{\min} \le z \le z_{\max}$ where the parameters, α, β, z_0, $z_{\min}$ and $z_{\max}$ depend on the survey. Table 2 summarizes the survey and redshift parameters assumed in our analysis.

We assume Gaussian shear fields to find the covariances needed for obtaining CCOSEBIs and also for the Fisher analysis (see Joachimi et al. 2008). The relation between the E-COSEBIs

covariance and the convergence power spectrum for redshift bin pairs ij and kl is

$$
\begin{aligned}
C_{mn}^{(ij)(kl)} &\equiv \left\langle E_m^{(ij)} E_n^{(kl)} \right\rangle - \left\langle E_m^{(ij)} \right\rangle \left\langle E_n^{(kl)} \right\rangle \\
&= \frac{1}{2\pi A} \int_0^\infty {\rm d}\ell\,\ell\, W_m(\ell) W_n(\ell) \\
&\quad \times \left(\bar{P}_{\rm E}^{(ik)}(\ell) \bar{P}_{\rm E}^{(jl)}(\ell) + \bar{P}_{\rm E}^{(il)}(\ell) \bar{P}_{\rm E}^{(jk)}(\ell) \right),
\end{aligned}
\tag{21}
$$

where

$$\bar{P}_{\rm E}^{(ik)}(\ell) := P_{\rm E}^{(ik)}(\ell) + \delta_{ik} \frac{\sigma_\epsilon^2}{2\bar{n}_i}, \tag{22}$$

and A is the survey area, σ_ϵ is the galaxy intrinsic ellipticity dispersion and $\bar{n}_i$ is the average galaxy number density in redshift bin i. The overall shape of the COSEBIs covariance is shown in Asgari et al. (2012).

5. Results

This section is dedicated to our results. The filter functions of the CCOSEBIs for the fiducial cosmology are shown first, followed by a figure-of-merit analysis. We compare the figure-of-merit values for cases where the covariance is known versus the use of a wrong covariance in constructing the CCOSEBIs.

5.1. Weight functions of compact COSEBIs

Inserting Eqs. (14) and (16) into Eq. (18) results in relations between $\boldsymbol{E}^{\rm c}$ and the COSEBIs filters,

$$E_\mu^{\rm c} = \frac{1}{2} \int_{\theta_{\min}}^{\theta_{\max}} {\rm d}\vartheta\, \vartheta \left[\mathsf{B}_{\mu n i j} T_{+n}(\vartheta)\, \xi_+^{(ij)}(\vartheta) + \mathsf{B}_{\mu n i j} T_{-n}(\vartheta)\, \xi_-^{(ij)}(\vartheta) \right], \tag{23}$$

and

$$E_\mu^{\rm c} = \int_0^\infty \frac{{\rm d}\ell\,\ell}{2\pi} \mathsf{B}_{\mu n i j} W_n(\ell)\, P_{\rm E}^{(ij)}(\ell). \tag{24}$$

For each redshift bin pair, ij, a set of $N' = P(P+3)/2$ (P is the number of free parameters) filters exist. The new filters in real and Fourier space, respectively, are

$$
\begin{aligned}
T_{\pm\mu ij}^{\rm c}(\vartheta) &= \mathsf{B}_{\mu n i j} T_{\pm n}(\vartheta) \\
W_{\mu ij}^{\rm c}(\ell) &= \mathsf{B}_{\mu n i j} W_n(\ell).
\end{aligned}
\tag{25}
$$

With the above definitions we can rewrite the compressed statistics, $\boldsymbol{E}^{\rm c}$, in terms of the compressed filter functions,

$$E_\mu^{\rm c} = \frac{1}{2} \int_{\theta_{\min}}^{\theta_{\max}} {\rm d}\vartheta\, \vartheta \left[T_{+\mu ij}^{\rm c} \xi_+^{(ij)}(\vartheta) + T_{-\mu ij}^{\rm c} \xi_-^{(ij)}(\vartheta) \right], \tag{26}$$

and

$$E_\mu^{\rm c} = \int_0^\infty \frac{{\rm d}\ell\,\ell}{2\pi} W_{\mu ij}^{\rm c}(\ell)\, P_{\rm E}^{(ij)}(\ell). \tag{27}$$

Multiplying each $E_\mu^{\rm c}$ by a constant has no effect on the information level. We can therefore normalize the filter functions for each compressed statistic separately, so that

$$\sum_{ij} \frac{1}{\Delta\vartheta} \int_{\theta_{\min}}^{\theta_{\max}} {\rm d}\vartheta\, [T_{+\mu ij}^{\rm c}(\vartheta)]^2 = 1. \tag{28}$$

Figure 1 shows the first-order filter functions, $T^F_{+\mu ij}(\vartheta)$ and $W^F_{+\mu ij}(\ell)$, with $1 \le \mu \le P = 7$ for our fiducial cosmology described in Sect. 4. Here we assume three redshift bins and seven free fiducial parameters, using 20 COSEBIs filters defined between $\theta_{\min} = 1'$ and $\theta_{\max} = 400'$. The redshift bins were chosen such that they contain an equal number of galaxies. Since the filters are designed to maximize the information obtained, their shape shows where most of the information in $\xi_+(\vartheta)$ or $P_E(\ell)$ lies. Here we choose to only show the first-order filters, although later on we use the first and second order, as well as the combination of both, to obtain the figure-of-merit. The general trend of T^F shows that there is more information about all of the parameters in the higher-redshift bins and on smaller angular scales.

However, each individual parameter shows a different pattern for each of the redshift pairs. For example, the real space filters, T^F, for Ω_Λ have significantly higher amplitudes for combinations of redshift bins two and three compared to combinations with the lowest redshift bin. The σ_8 and Ω_m filters follow each other closely, although in Fourier space, i.e. W^F, the differences are more pronounced. A closer look at the plots shows that, since these curves are not exactly the same and also evolve with redshift, it is possible to break their degeneracies, which is present in single-redshift studies (e.g., van Waerbeke et al. 2001; Hoekstra et al. 2002; Jarvis et al. 2003; Hetterscheidt et al. 2007; Kilbinger et al. 2013). The oscillations of the W^F are a real feature of the CCOSEBIs weights and do not vanish when more COSEBIs are incorporated in calculating them.

Tables 3 and 4 show the elements of the compression matrix, **B**, for CCOSEBIs. According to Eq. (18), each row of **B** corresponds to the coefficients for making one of the CCOSEBIs statistics, E^c_μ, by linearly combining the COSEBIs E_n. The value of the elements of **B** show how important each COSEBIs mode is for building a CCOSEBIs mode. In both tables, the element values are much lower for large n compared to smaller n. As a result we can safely (conservatively) take just the first 20 COSEBIs to build the compressed statistics.

5.2. Fisher analysis

The Fisher matrix depends on the data, E^c_μ, via

$$\mathsf{F}_{\mu\nu} = \frac{1}{2}\mathrm{Tr}\left[(\mathbf{C}^c)^{-1}\,\mathbf{C}^c_{,\mu}\,(\mathbf{C}^c)^{-1}\,\mathbf{C}^c_{,\nu} + (\mathbf{C}^c)^{-1}\,\mathbf{M}_{\mu\nu}\right], \quad (29)$$

where $\mathbf{C}^c$ is the data covariance, $\mathbf{M}_{\mu\nu} = E^c_{,\mu}\,(E^c_{,\nu})^t + E^c_{,\nu}\,(E^c_{,\mu})^t$ and the commas followed by subscripts indicate partial derivatives with respect to the cosmological parameters (see Tegmark et al. 1997, for example). We use the same figure-of-merit, f, which gives a measure of the mean error on parameters, as in Asgari et al. (2012),

$$f = \left(\frac{1}{\sqrt{\det \mathbf{F}}}\right)^{1/P}, \quad (30)$$

where P is the number of free parameters considered. Furthermore, we have shown in Asgari et al. (2012) that for a sufficiently large survey one can neglect the first term in Eq. (29) since it does not depend on the survey area (see Eq. (21)), while the second term is proportional to the survey area. Therefore, we neglect the first term in this study.

Figure 2 shows the dependence of f on the number of COSEBIs modes, $n_{\max}$, for eight redshift bins and seven free parameters. The constraints get tighter as $n_{\max}$ increases and reach a saturation level for all cases. The solid curve shows how much information can be gained if the $8n_{\max} \times 9/2 = 36n_{\max}$ COSEBIs are used, i.e., the maximum information, or minimum f value. The points show the amount of information in the first- (F) and second-order (S) CCOSEBIs, as well as the combination of both, denoted by E^c as before. The parameters used for calculating the covariance matrix to build F, S, and E^c are that of the fiducial cosmology, which are slightly different from the assumed true parameters (see Table 1). Nevertheless, the first-order CCOSEBIs are sufficient to reach a similar Fisher information level. The F data vector for this case have seven components, S have 28, and E^c have 35, while for $n_{\max} = 10$, the COSEBIs have 360 components; i.e., there is at least an order-of-magnitude difference between the number of observables for CCOSEBIs and COSEBIs. As a result, we can obtain the same accuracy of derived parameters with a highly significant reduction of observables.

The strong reduction of observables needed to cover all the cosmological information is of great interest for obtaining accurate covariances, hence reliable confidence regions for cosmological parameters. Whereas analytical methods may be able to obtain approximate covariances (see, e.g., Takada & Jain 2009; Sato et al. 2009; Pielorz et al. 2010; Hilbert et al. 2011; Takahashi et al. 2011, 2014, and references therein), an accurate covariance accounting for the complex survey geometry will probably require extensive simulations. Obtaining the covariance as sample variance from independent realizations of the simulated cosmology requires a number of realizations, which is about proportional to the number of observables (see Hartlap et al. 2007; Taylor & Joachimi 2014). Even a modest reduction in the number of relevant observables is therefore useful. As we have seen above, the CCOSEBIs serve this purpose very well.

Whereas the construction of the CCOSEBIs requires information about the covariance, this does not have to be very accurate. To show how using a substantially wrong covariance in defining the compressed data vector affects the constrains, we artificially change the value of σ_ϵ which affects the diagonals of the covariance matrix according to Eqs. (21) and (22). Figure 3 shows f for seven free parameters, five redshift bins, and 20 COSEBIs modes, as a function of the change in the parameter σ_ϵ. The parameter f is normalized by its minimum value, i.e., using COSEBIs with their true covariance, while σ_ϵ is normalized by its true value. The first-order statistics, F, which has the same dimension as the parameter space, rapidly diverges from the true Fisher information limit, while the second order, S, and E^c, which span a larger dimensional space, are much less sensitive to the errors of the COSEBIs covariance, used for constructing the CCOSEBIs. Even for a 16 times larger σ_ϵ, the fractional difference between the optimal f and the measured one from E^c is small. As a result, the consideration of the second-order statistics indeed provides a powerful mitigation for inaccurate covariances.

We stress again that the accuracy with which the original covariance can be determined affects the efficiency of our data compression, in the sense that what fraction of the cosmological information contained in the original COSEBIs is preserved in the compressed COSEBIs. However, it does not bias the parameter determination, because the CCOSEBIs are linear combinations of the original ones.

We therefore conclude that the method proposed here – constructing new observables using an approximate covariance, and employing them for cosmological parameter studies – yields a very promising tool for effectively reducing the necessary efforts for constructing accurate covariances. This data compression will also be of great help if the covariances are to be obtained
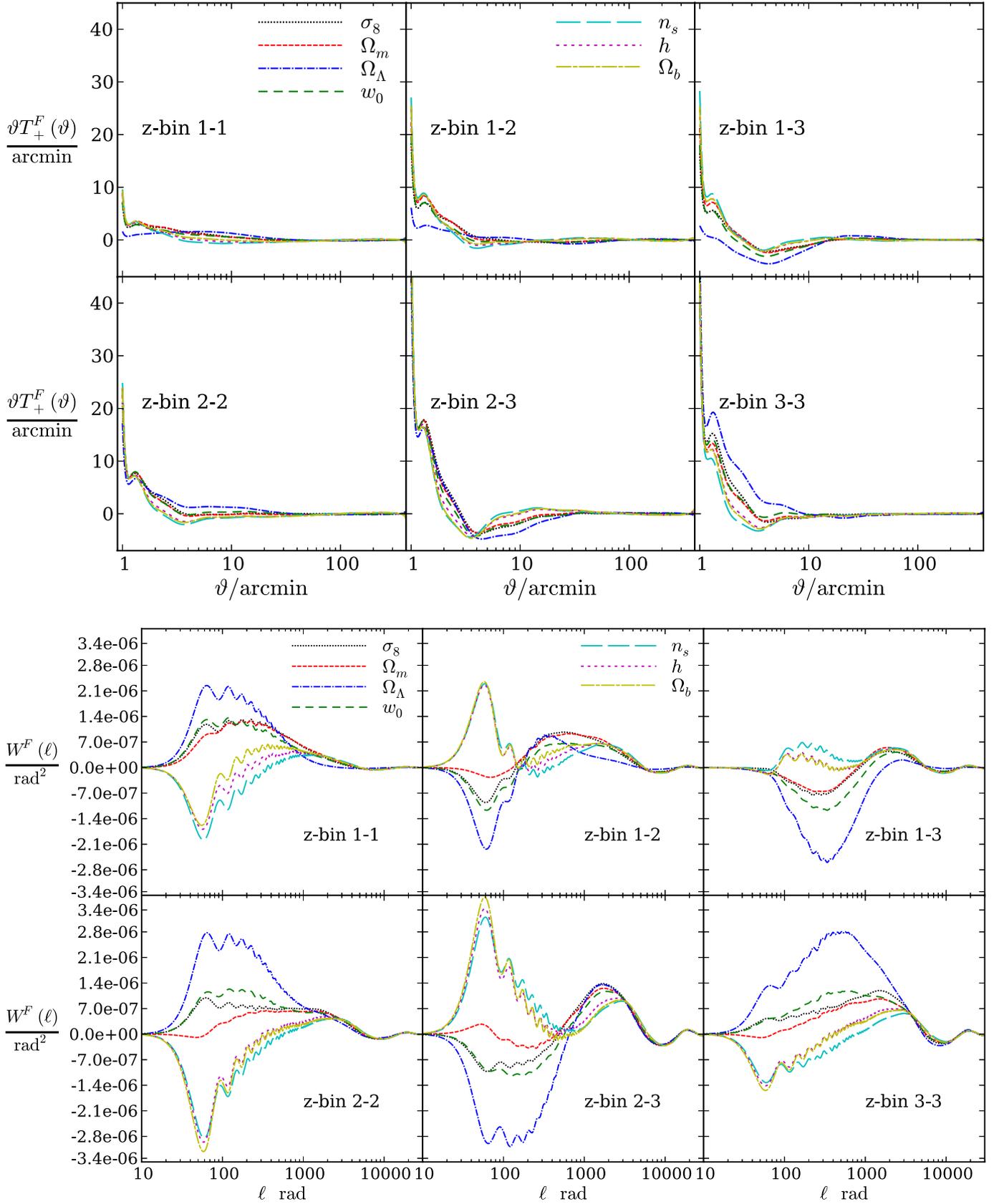
Fig. 1. Filter functions $T^F_{+\mu ij}(\vartheta)$ (*top panel*) and $W^F_{\mu ij}(\ell)$ (*bottom panel*) related to their progenitors via Eq. (25). For clarity we show $\vartheta T^F_{+\mu ij}(\vartheta)$ on a logarithmic ϑ-scale. 3 redshift bins and 20 COSEBIs filters defined between $\theta_{\min} = 1'$ and $\theta_{\max} = 400'$ are considered. Each parameter has a different filter function for each redshift pair bin (*z*-bin). A comparison between these functions and their progenitors (see Schneider et al. 2010; and Asgari et al. 2012) show that they have considerably fewer structures and oscillations. The filters for each parameter are normalized according to Eq. (28). Hence one can compare how effective each *z*-bin pair is for constraining the parameters.

Table 3. Elements of the normalized compression matrix in percentage, $100 \times B_{\mu n}$, for 1 redshift bin and 3 parameters.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
σ_8	43.7	−59.2	53.4	−36.0	17.2	−4.8	0.7	−1.9	4.5	−5.9	5.8	−4.6	2.8	−1.3	0.3
Ω_m	42.1	−57.8	53.5	−37.9	20.1	−7.2	1.8	−1.9	4.2	−5.9	6.0	−4.8	3.0	−1.3	0.3
Ω_Λ	48.5	−61.4	51.0	−31.3	13.4	−3.1	0.3	−1.8	4.3	−5.7	5.6	−4.4	2.7	−1.2	0.3
σ_8, σ_8	46.2	−61.2	52.8	−32.5	12.5	−1.1	−1.0	−1.7	4.7	−6.0	5.6	−4.2	2.5	−1.1	0.3
σ_8, Ω_m	38.2	−54.5	53.9	−41.9	25.4	−11.3	3.7	−2.2	3.9	−5.8	6.2	−5.2	3.3	−1.5	0.4
σ_8, Ω_Λ	44.9	−61.0	53.8	−33.2	11.8	0.9	−3.2	−0.3	4.5	−6.3	5.8	−4.1	2.3	−0.9	0.2
Ω_m, Ω_m	47.6	−62.0	52.1	−30.8	10.9	−0.3	−1.2	−1.7	4.7	−5.8	5.4	−4.1	2.5	−1.1	0.3
Ω_m, Ω_Λ	49.8	−63.7	51.1	−26.2	4.5	5.2	−3.8	−1.7	5.4	−6.1	4.9	−3.3	1.9	−0.8	0.2
$\Omega_\Lambda, \Omega_\Lambda$	48.9	−62.8	51.5	−28.7	8.4	1.8	−2.3	−1.4	4.7	−5.9	5.3	−3.8	2.3	−1.0	0.2

Notes. The first column shows the subscript of E_μ^c, where one parameter subscripts belong to the first-order statistics, F, and double subscripts belong to the second-order statistics S. The first row shows the value of n.

Table 4. Normalized compression matrix elements in percentage, $100 \times \mathbf{B}$, for 2 redshift bins and 3 parameters.

	z-bin 1-1						z-bin 1-2						z-bin 2-2					
	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
σ_8	17.8	−17.4	10.4	−4.4	1.3	−0.2	22.2	−33.3	31.3	−20.3	8.4	−1.8	36.8	−48.2	41.4	−25.8	10.8	−2.3
Ω_m	19.4	−19.9	12.8	−6.1	2.2	−0.4	25.7	−37.9	35.7	−23.5	10.1	−2.2	33.2	−43.4	37.1	−23	9.6	−2.1
Ω_Λ	13.1	−8.2	1.7	0.2	0.1	−0.2	1.0	−13.6	20.4	−16.2	7.3	−1.6	50.7	−59.7	45.0	−24.3	8.9	−1.7
σ_8, σ_8	12.1	−10.6	4.5	−0.5	−0.5	0.2	13.3	−20.8	20.5	−14.1	6.2	−1.4	43.1	−56.7	48.5	−29.6	12.1	−2.5
σ_8, Ω_m	24.9	−26.2	17.9	−9.4	3.7	−0.8	37.4	−51.4	44.5	−26.5	10.0	−1.8	15.0	−21.6	21.1	−15.3	7.5	−1.9
σ_8, Ω_Λ	4.3	−4.0	1.1	1.1	−1.2	0.4	5.0	−8.0	10.2	−9.5	5.7	−1.7	44.9	−60.7	52.5	−31.9	12.7	−2.5
Ω_m, Ω_m	13.6	−12.8	6.3	−1.3	−0.3	0.2	17.7	−25.2	23.5	−15.7	6.9	−1.6	41.7	−54.8	46.4	−28	11.2	−2.3
Ω_m, Ω_Λ	9.4	−9.5	5.1	−1.1	−0.4	0.3	13.7	−19.4	18.1	−12.3	5.6	−1.3	44.1	−57.9	49.1	−29.5	11.6	−2.3
$\Omega_\Lambda, \Omega_\Lambda$	9.5	−7.5	2.1	1.0	−1.1	0.3	7.8	−14.2	15.8	−11.9	5.7	−1.4	46.5	−59.8	49.4	−29.1	11.3	−2.2

Notes. The first column shows the subscript of E_μ^c, where one-parameter subscripts belong to the first-order statistics, F, and double subscripts belong to the second-order statistics S. The absolute values of the columns of $\mathbf{B}$ for each redshift pair decreases rapidly after $n = 3$.
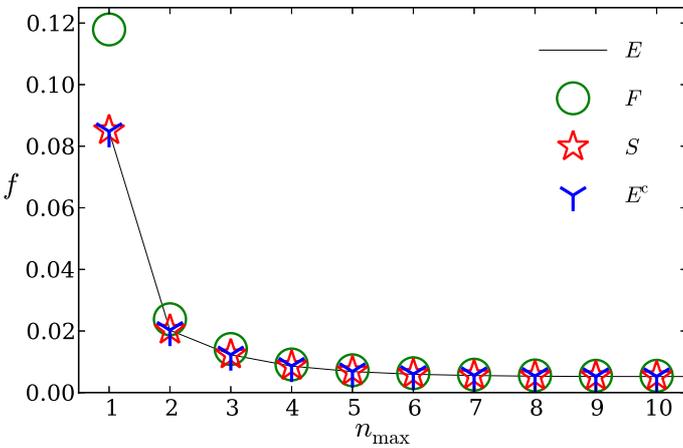

Fig. 2. Figure-of-merit, f, as a function of the number of COSEBIs, n_{max}, used. 7 free parameters listed in Table 1, and 8 tomographic redshift bins are considered here. The solid line shows the result for using Log-COSEBIs with the true underlying cosmology. It also represents the maximum information level for a given n_{max}. The circles, stars and the Y-shaped symbols represent the f-values for First order, Second order, and their combination E^c, where n_{max} COSEBIs modes with the fiducial cosmological parameters are utilized in making them.
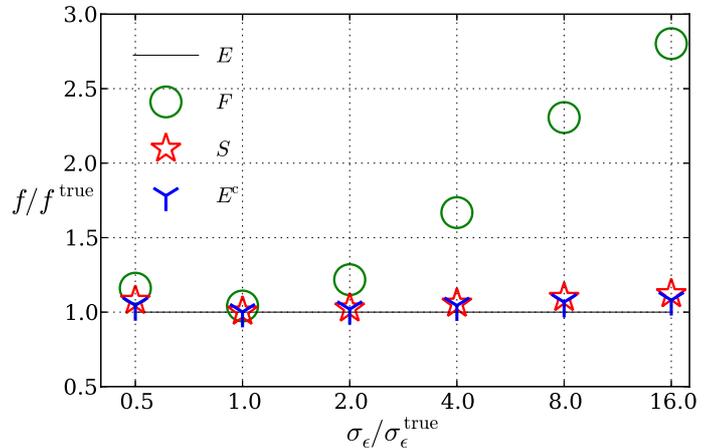
from the data themselves, e.g., by subdividing the survey region, calculating the sample variance on each subsurvey, and scaling the result with the ratio of subsurvey to survey area.

Fig. 3. Figure-of-merit, f, as a function of σ_ϵ. f is normalized by its minimum value which corresponds to using COSEBIs with the correct covariance (the solid line). The intrinsic ellipticity dispersion of galaxies, σ_ϵ, is varied with respect to its true value, 0.3, to show the effects of using a wrong covariance. The markers show the value of f for first order, F, second order, S and the combination of both E^c CCOSEBIs.

6. Band power

As mentioned in Sect. 3, any filter function defined on a finite angular interval that satisfies the constraints (15) can be expressed in terms of the COSEBIs filters. A particular filter one might be interested in is a top-hat function in Fourier space, corresponding to a band power (e.g., Brown et al. 2003; Hikage et al. 2011). In this section we study how well band powers can be

approximated from correlation functions measured on a finite interval with clean E-/B-mode separation.

Thus, let $\hat{W}(\ell)$ be a target filter function in Fourier space, and let us design a filter that approximates $\hat{W}(\ell)$ as closely as possible. That means we want to find a filter that minimizes

$$\Delta = \int d\ell\, \ell\, \left[W(\ell) - \hat{W}(\ell) \right]^2, \tag{31}$$

where $W(\ell)$ is a linear combination of the $W_n(\ell)$,

$$W(\ell) = \sum_n c_n W_n(\ell). \tag{32}$$

The Δ integral can be defined with a different weighting of ℓ; for example, one can replace $d\ell\, \ell$ in Eq. (31) by $d \ln \ell$ or simply $d\ell$. Doing so does not affect the final estimation accuracy of $\hat{W}(\ell)$ significantly, but may be numerically advantageous.

We want to find the coefficients c_n that minimize Δ; setting the derivatives of Δ with respect to c_m to zero, we find

$$\sum_n c_n \int d\ell\, \ell\, W_n(\ell) W_m(\ell) = \int d\ell\, \ell\, W_m(\ell) \hat{W}(\ell). \tag{33}$$

By defining the matrix

$$\mathsf{Q}_{nm} \equiv \int d\ell\, \ell\, W_n(\ell) W_m(\ell), \tag{34}$$

and the vector

$$V_n \equiv \int d\ell\, \ell\, W_n(\ell) \hat{W}(\ell), \tag{35}$$

we can rewrite Eq. (33) in matrix form,

$$\mathsf{Q} c = V \implies c = \mathsf{Q}^{-1} V. \tag{36}$$

The minimum of Δ for this solution is

$$\Delta_{\min} = \int d\ell\, \ell\, \hat{W}^2(\ell) - V \mathsf{Q}^{-1} V, \tag{37}$$

and we quantify the relative deviation of the closest filter W to $\hat{W}$ by

$$\delta_{\min} = \sqrt{\Delta_{\min} \left[\int d\ell\, \ell\, \hat{W}^2(\ell) \right]^{-1}}. \tag{38}$$

Filters that satisfy Eq. (15) and vanish outside of the angular range $[\theta_{\min}, \theta_{\max}]$ are the only ones that can be represented by COSEBIs. Applying any filter that does not satisfy these conditions, on either power spectra or 2PCFs, therefore results in spillage from outside of the measured angular range. A top-hat function in Fourier space is an example of a filter that is not easily represented by weight functions that correspond to a finite range in real space. A top-hat function in Fourier is defined as $\hat{W}(\ell) = 1$ between $\ell_{\min}$ and $\ell_{\max}$, and zero otherwise. The real space version, $\hat{T}_+(\vartheta)$, of such a function is

$$\hat{T}_+(\vartheta) = \int_0^\infty d\ell\, \ell\, \mathrm{J}_0(\ell\vartheta) \hat{W}(\ell) \tag{39}$$

$$= \int_{\ell_{\min}}^{\ell_{\max}} d\ell\, \ell\, \mathrm{J}_0(\ell\vartheta) = \frac{1}{\vartheta} \left[\ell_{\max} \mathrm{J}_1(\ell_{\max}\vartheta) - \ell_{\min} \mathrm{J}_1(\ell_{\min}\vartheta) \right].$$

This equation shows that the function $\hat{T}_+(\vartheta)$ is manifestly non-zero for $\vartheta < \theta_{\min}$ and $\vartheta > \theta_{\max}$, which implies that the band

power cannot be represented in terms of E-/B-mode separating combinations of correlation functions over a finite angular interval.

Using Parseval's theorem, we can find a lower bound for Δ,

$$\Delta = \int_0^\infty d\ell\, \ell\, \left[W(\ell) - \hat{W}(\ell) \right]^2 = (2\pi)^2 \int_0^\infty d\vartheta\, \vartheta\, \left[T_+(\vartheta) - \hat{T}_+(\vartheta) \right]^2$$

$$= (2\pi)^2 \left\{ \int_0^{\theta_{\min}} d\vartheta\, \vartheta\, \hat{T}_+^2(\vartheta) + \int_{\theta_{\max}}^\infty d\vartheta\, \vartheta\, \hat{T}_+^2(\vartheta) \right. \tag{40}$$

$$\left. + \int_{\theta_{\min}}^{\theta_{\max}} d\vartheta\, \vartheta\, \left[T_+(\vartheta) - \hat{T}_+(\vartheta) \right]^2 \right\},$$

where $T_+(\vartheta)$ is the real space form of $W(\ell)$. The sum of the first two integrals in (40) is the absolute lower bound on Δ, since the last integral in that equation is non-negative, irrespective of the choice of $T_+(\vartheta)$. Hence, the lower bound for $\delta_{\min}$ is

$$\delta_{\min} \geq \delta_{\mathrm{LB}} \equiv \sqrt{\frac{8\pi^2}{\ell_{\max}^2 - \ell_{\min}^2} \left\{ \int_0^{\theta_{\min}} d\vartheta\, \vartheta\, \hat{T}_+^2(\vartheta) + \int_{\theta_{\max}}^\infty d\vartheta\, \vartheta\, \hat{T}_+^2(\vartheta) \right\}}. \tag{41}$$

To reach the absolute lower bound, the last integral in (40) should vanish. It is necessary and sufficient for $\hat{T}_+(\vartheta)$ to satisfy the conditions (15) for that to happen, since then $\hat{T}_+(\vartheta)$ in the range $\theta_{\min} \leq \vartheta \leq \theta_{\max}$ can be represented as a sum over the COSEBIs weights T_{+n}. Inserting the analytic form of $\hat{T}_+(\vartheta)$ from Eq. (39) into Eq. (15) results in the following two conditions:

$$\mathrm{I} := \mathrm{J}_0(\ell_{\min}\theta_{\max}) - \mathrm{J}_0(\ell_{\min}\theta_{\min}) - \mathrm{J}_0(\ell_{\max}\theta_{\max}) + \mathrm{J}_0(\ell_{\max}\theta_{\min}) = 0, \tag{42}$$

and

$$\mathrm{II} := \theta_{\min}^2 \mathrm{J}_0(\ell_{\min}\theta_{\min}) - \theta_{\max}^2 \mathrm{J}_0(\ell_{\min}\theta_{\max}) \tag{43}$$

$$+ \frac{2}{\ell_{\min}} [\mathrm{J}_1(\ell_{\min}\theta_{\max}) - \mathrm{J}_1(\ell_{\min}\theta_{\min})]$$

$$- \theta_{\min}^2 \mathrm{J}_0(\ell_{\max}\theta_{\min}) + \theta_{\max}^2 \mathrm{J}_0(\ell_{\max}\theta_{\max})$$

$$- \frac{2}{\ell_{\max}} [\mathrm{J}_1(\ell_{\max}\theta_{\max}) - \mathrm{J}_1(\ell_{\max}\theta_{\min})] = 0,$$

which should be simultaneously true for $\delta_{\min} = \delta_{\mathrm{LB}}$. However, most combinations of $\theta_{\min}, \theta_{\max}, \ell_{\min}$, and $\ell_{\max}$ do not satisfy these conditions. As a result in general $\delta_{\min} > \delta_{\mathrm{LB}}$ for most cases.

Once the COSEBIs, E_n, are measured from the data, the band power can be estimated by linearly combining them,

$$\hat{E} = \frac{1}{2\pi} \int d\ell\, \ell\, \hat{W}(\ell) P_{\mathrm{E}}(\ell) \approx \frac{1}{2\pi} \int d\ell\, \ell\, W(\ell) P_{\mathrm{E}}(\ell) \tag{44}$$

$$= \sum_n c_n \frac{1}{2\pi} \int d\ell\, \ell\, W_n(\ell) P_{\mathrm{E}}(\ell) = \sum_n c_n E_n.$$

Figure 4 depicts the convergence to an estimated $W(\ell)$ by increasing the number of COSEBIs modes. The number of COSEBIs needed for convergence is substantially higher than the number needed for constraining parameters with one redshift bin.

Figure 5 shows the dependence of $\delta_{\min}$ on the number $n_{\max}$ of COSEBIs. Here we demonstrate that for all ranges considered in ℓ, a saturation level is reached; i.e., adding more COSEBIs filters will not lead to a smaller difference between the estimated W and the top hat. Furthermore, in Table 5 we show the value of $\delta_{\min}$ for 80 COSEBIs that can be compared to its lower bound, δ_{LB} (see Eq. (41)), and the relative difference between the estimated band power and its true value, $\delta_{\mathrm{band}} = |(\hat{E} - E)/\hat{E}|$. The saturated $\delta_{\min}$
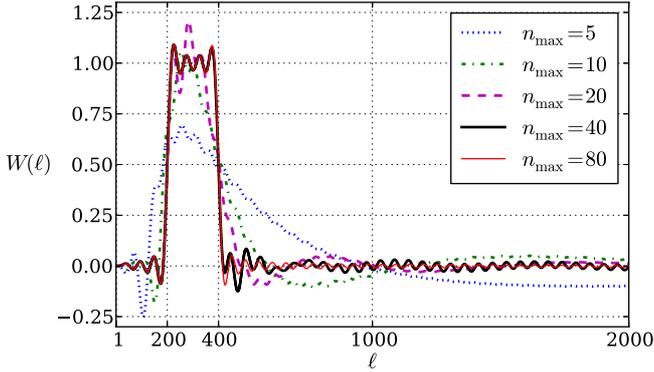
Fig. 4. Estimated top-hat filter function with $\ell_{min} = 200$ and $\ell_{max} = 400$, from n_{max} COSEBIs filters defined on $1' < \vartheta < 400'$. The changes between using 40 and 80 COSEBIs filters are small, so that no better representation is obtained by using an even higher value of n_{max}.

values are higher than but close to δ_{LB}. The difference between the two arises from violating conditions (42) and (43). In the table, we use three ℓ-weighting schemes, which do not change the δ_{min} values significantly. However, the estimated band power deviations, δ_{band}, can vary by much more than a few percent between the cases. This is due to the spillage of the estimated band power and the fact that the ℓ-weighting scheme decides which way the spillage is directed to. The δ_{band} values are cosmology dependent and can be very different for a power spectrum with more features.

It is interesting to note that the deviations δ_{band} of the estimated band powers from their true values are in some cases smaller than the relative deviation δ_{min} between the top-hat filter and the best representation of the top hat by COSEBIs weight functions. This, however, is an effect of the properties of the power spectrum in our assumed cosmological model: the power spectrum is smooth enough that the spilling caused by the effective weight $W(\ell)$ out of, and into the range of the top hat, can compensate for each other (see Schneider et al. 2002a, for a related discussion on band powers in cosmic shear analysis). That δ_{band} is relatively small is therefore not a statement about the accuracy of the method of band power estimates, but rather a consequence of the properties of the power spectrum. But the latter should be probed by estimating the band power. Thus, it would be strongly misleading to judge the accuracy of the method on presumed properties that ought to be investigated instead. Indeed, it is the quantity δ_{min} that yields an estimate of the accuracy with which band powers can be obtained.

Becker & Rozo (2014) constructed slightly different band powers, with a weight function in ℓ-space of a log-normal form. As expected, their smoother weight $\hat{W}$, which does not impose strict ℓ cuts in Fourier space, compared to the top-hat considered here, leads to less spilling of the best-fitting weight function in ϑ-space. In other words, the log-normal weight $\hat{W}$ can be more accurately represented by the Fourier transform of the shear correlation function on a finite interval, at the price of a non-diagonal covariance of band powers even in the case of Gaussian shear fields.

7. Summary and discussion

Data compression is an important challenge for future cosmological surveys, and it is essential for estimating accurate covariances. Current cosmological surveys, such as *Planck*, provide us with tight constraints on most cosmological parameters.
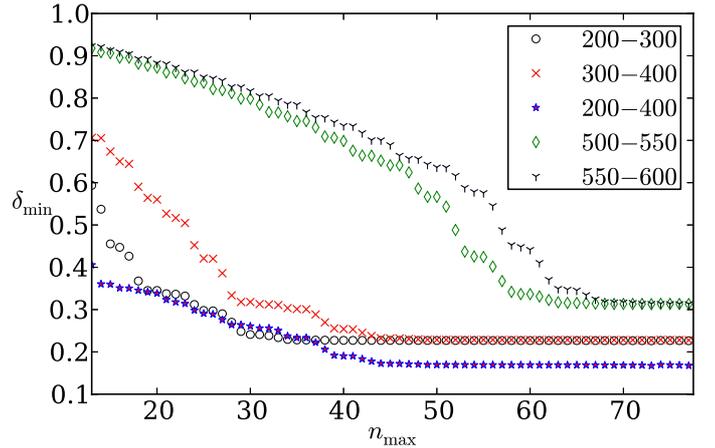

Fig. 5. Relative difference, δ_{min}, between the estimated top hat and the input as a function of the number of COSEBIs filters, n_{max}, utilized for a few ℓ-ranges. In all cases the saturation level is reached before $n_{max} = 80$. The minimum value of δ_{min} is shown in Table 5. In general, a higher number of modes is needed for a narrower top-hat filter, which is due to the spillage beyond the observed angular range (see Eq. (41)).

Table 5. Examples of band power estimation from 80 COSEBIs for $\xi_\pm(\vartheta)$ in the interval $\vartheta \in [1', 400']$.

$l_{min} - l_{max}$	δ_{min}	δ_{LB}	$\delta_{band}, \mathrm{d}\ell\,\ell$	$\delta_{band}, \mathrm{d}\ell$	$\delta_{band}, \mathrm{d}\ell/\ell$
200−300	2.28	2.27	3.31	0.72	2.29
300−400	2.30	2.29	6.78	1.74	2.99
200−400	1.72	1.71	4.94	1.20	2.62
500−550	3.14	3.14	10.24	3.03	4.27
550−600	3.14	3.14	14.90	7.83	9.14

Notes. The first column shows the ℓ-range of the top-hat function $\hat{W}(\ell)$. The rest of the columns show percentage values for minimum relative difference between the estimated and the top hat, δ_{min}, the absolute minimum (Lower Bound) value for δ_{min}, δ_{LB} (see Eq. (41)), and the relative difference between the estimated band power and its true value, $\delta_{band} = |(\hat{E} - E)/\hat{E}|$, for different ℓ weightings, respectively. δ_{min} and δ_{LB} values correspond to the $\mathrm{d}\ell\,\ell$ weighting. The values of these quantities for other cases are similar. The cosmological model used here is the fiducial model from Table 1, with one redshift bin.

This motivated us to define combinations of statistics inspired by their low-order Taylor expansion around a fiducial cosmological model. The strategy for finding the compressed statistics involves first- and second-order derivatives of parent statistics with respect to the free parameters, as well as their covariance. The statistics corresponding to the first-order derivatives, F, have the same dimension as the parameter space, while the statistics derived from second-order derivatives, S, provide a possibility to span a larger dimensional space. Consequently, F is more sensitive to the choice of the fiducial cosmology and covariance. The combination of F and S enables one to use well-defined and motivated sets of statistics that alleviate many of the data analysis problems. In total the number of compressed statistics is $P(P + 3)/2$, where P is the number of free parameters in the model.

In the case of a cosmic shear analysis, the COSEBIs already provide an effective compression compared to other two-point statistics, such as the shear two-point correlation functions. However, adding tomographic bins, which is necessary for intrinsic alignment corrections, substantially increases the number of observables. As a result, further data compression

is required. We applied our compression formalism to Log-COSEBIs to study its properties. We found that for a well-estimated COSEBIs covariance matrix, the first-order compressed statistics are sufficient. However, as mentioned above, the accuracy of covariance estimations from simulations depend on the number of observables incorporated. The higher this number is, the more simulations are needed, which rapidly becomes too expensive. Consequently, we used highly inaccurate covariances for defining the compressed COSEBIs (CCOSEBIs), to test their efficiency for such cases. We found that the figure-of-merit obtained from the first-order CCOSEBIs deviates substantially from the optimal information level as the difference between the assumed COSEBIs covariance and their true covariance increases. In contrast, the set of second-order CCOSEBIs is far less sensitive to the choice of covariance, owing to its larger dimensionality. The combination of both is basically insensitive to the accuracy of the covariance, at least in the framework of the simple model that we have tested here. Consequently, we propose that this strategy is applicable to the future data analysis. We note that our first-order CCOSEBIs is equivalent to the Karhunen–Loève data compression (with parameter-independent covariance) in that the covariance is accurately known (Tegmark et al. 1997).

In this paper we used a Fisher analysis, which assumes the parameters have a normal distribution, to compare the constrains from COSEBIs and CCOSEBIs. Both Fisher matrix and $\boldsymbol{F}$, the first-order compressed statistics depend only on the first-order derivatives and the covariance. If the fiducial cosmology coincides with the truth and the covariance is exact, then the $\boldsymbol{F}$ is equivalent to a Fisher formalism, since in this case the derivative matrix of $\boldsymbol{F}$ is equal to its covariance matrix, which is consequently equal to the Fisher matrix. However, when the covariance deviates from the truth, the differences become visible. For our future studies we plan to use likelihood analysis, which does not make assumptions about the Gaussianity of the likelihood with respect to the model parameters.

The COSEBIs filter functions form a complete basis for any filter that satisfies Eq. (15), which are necessary and sufficient conditions for a clean E-/B-separation on a finite interval, together with the condition that the filters should also vanish outside of the finite angular range. Consequently, any filter that satisfies these conditions can be represented by a linear combination of the COSEBIs filters. In this paper we showed how any given weight function can be mimicked by COSEBIs weights. In particular, we tried to represent top-hat filters in Fourier space using this strategy. We found that, owing to the infinite support of a Fourier top-hat in real space, an accurate representations of them is impossible. This task becomes harder as the top hat and the angular range get narrower. Consequently, band convergence power spectra estimated from finite angular range information will suffer from spillage, so they will be inaccurate and biased, in a way that depends on the power spectrum – the quantity to be probed. We therefore caution against using narrow-band power spectra for cosmic shear analysis. The estimated powers are relatively accurate if the power spectra are instead smooth functions of ℓ. However, for such smooth functions, there are better ways to characterize them than using band powers, such as presenting them by a set of basis functions. We thus see no advantage in using power spectra for cosmic shear analysis on a finite angular range.

References

Asgari, M., Schneider, P., & Simon, P. 2012, A&A, 542, A122
Becker, M. R., & Rozo, E. 2014 [arXiv:1412.3851]
Bond, J. R., & Efstathiou, G. 1984, ApJ, 285, L45
Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, ApJ, 466, 623
Brown, M. L., Taylor, A. N., Bacon, D. J., et al. 2003, MNRAS, 341, 100
Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2002, ApJ, 568, 20
Eifler, T., Krause, E., Schneider, P., & Honscheid, K. 2014, MNRAS, 440, 1379
Hartlap, J., Simon, P., & Schneider, P. 2007, A&A, 464, 399
Hetterscheidt, M., Simon, P., Schirmer, M., et al. 2007, A&A, 468, 859
Hikage, C., Takada, M., Hamana, T., & Spergel, D. 2011, MNRAS, 412, 65
Hilbert, S., Hartlap, J., & Schneider, P. 2011, A&A, 536, A85
Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2002, ApJ, 577, 595
Huff, E. M., Eifler, T., Hirata, C. M., et al. 2014, MNRAS, 440, 1322
Jarvis, M., Bernstein, G. M., Fischer, P., et al. 2003, AJ, 125, 1014
Joachimi, B., & Bridle, S. L. 2010, A&A, 523, A1
Joachimi, B., Schneider, P., & Eifler, T. 2008, A&A, 477, 43
Kilbinger, M., Fu, L., Heymans, C., et al. 2013, MNRAS, 430, 2200
Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints
    [arXiv:1110.3193]
Marian, L., Smith, R. E., Hilbert, S., & Schneider, P. 2013, MNRAS, 432, 1338
Peebles, P. J., & Ratra, B. 2003, Rev. Mod. Phys., 75, 559
Pielorz, J., Rödiger, J., Tereno, I., & Schneider, P. 2010, A&A, 514, A79
Sato, M., Hamana, T., Takahashi, R., et al. 2009, ApJ, 701, 945
Schneider, P., & Kilbinger, M. 2007, A&A, 462, 841
Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002a, A&A, 396, 1
Schneider, P., van Waerbeke, L., & Mellier, Y. 2002b, A&A, 389, 729
Schneider, P., Eifler, T., & Krause, E. 2010, A&A, 520, A116
Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311
Takada, M., & Jain, B. 2009, MNRAS, 395, 2065
Takahashi, R., Yoshida, N., Takada, M., et al. 2011, ApJ, 726, 7
Takahashi, R., Soma, S., Takada, M., & Kayo, I. 2014, MNRAS, 444, 3473
Taylor, A., & Joachimi, B. 2014, MNRAS, 442, 2728
Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, ApJ, 480, 22
van Waerbeke, L., Mellier, Y., Radovich, M., et al. 2001, A&A, 374, 757