What is the super-sample covariance? A fresh perspective for second-order shear statistics

Laila Linke1,2, Pierre A. Burger1, Sven Heydenreich1,3, Lucas Porth1, and Peter Schneider1

1 Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany
2 Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, 6020 Innsbruck, Austria
e-mail: laila.linke@uibk.ac.at
3 Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

Received 23 February 2023 / Accepted 25 October 2023

ABSTRACT

Cosmological analyses of second-order weak lensing statistics require precise and accurate covariance estimates. These covariances are impacted by two sometimes neglected terms: a negative contribution to the Gaussian covariance due to a finite survey area, and the super-sample covariance (SSC), which for the power spectrum contains the impact of Fourier modes larger than the survey window. We show here that these two effects are connected and can be seen as correction terms to the ‘large-field-approximation’, the asymptotic case of an infinitely large survey area. We describe the two terms collectively as “finite-field terms”. We derive the covariance of second-order shear statistics from first principles. For this, we use an estimator in real space without relying on an estimator for the power spectrum. The resulting covariance does not scale inversely with the survey area, as might naively be assumed. This scaling is only correct under the large-field approximation when the contribution of the finite-field terms tends to zero. Furthermore, all parts of the covariance, not only the SSC, depend on the power spectrum and trispectrum at all modes, including those larger than the survey. We also show that it is generally impossible to transform an estimate of the power spectrum covariance into the covariance of a real-space statistic. Such a transformation is only possible in the asymptotic case of the large-field approximation. Additionally, we find that the total covariance of a real-space statistic can be calculated using correlation function estimates on spatial scales smaller than the survey window. Consequently, estimating covariances of real-space statistics, in principle, does not require information on spatial scales larger than the survey area. We demonstrate that this covariance estimation method is equivalent to the standard sample covariance method.

Key words. gravitation – gravitational lensing: weak – methods: analytical – methods: statistical – cosmological parameters – large-scale structure of Universe

1. Introduction

Second-order statistics of cosmic shear are essential tools for cosmological analyses (Heymans et al. 2021; Hikage et al. 2019; Abbott et al. 2022). Inference of cosmological parameters from these statistics requires a robust understanding of their covariances. While covariance models for second-order shear statistics have been derived and validated for over a decade (Joachimi et al. 2008; Takada & Hu 2013), two effects have only recently garnered more attention; specifically, the super-sample covariance (SSC) and the impact of the survey window on the Gaussian covariance. Our goal here is to show that these two effects are related and caused by the same approximation on the survey window function.

Analytic models for the covariance of second-order statistics are usually expressed in terms of the power spectrum covariance. The power spectrum covariance can be divided into three terms: a Gaussian and an intra-survey non-Gaussian part, which depend only on the power spectrum and trispectrum at ℓ-modes within a survey area, and the SSC, which also depends on ℓ-modes outside the survey (Takada & Hu 2013). While the first two terms scale linearly with the inverse survey area, the SSC shows a complicated dependence on the survey window function (Lacasa & Grain 2019; Gouyou Beauchamps et al. 2022). An analytic model for the SSC was derived in Barreira & Schmidt (2017).

Additionally, N-body simulations with small box sizes and periodic boundary conditions cannot fully reproduce the SSC of the matter power spectrum, as they do not include the small ℓ-modes (large spatial scales) on which the SSC depends (de Putter et al. 2012; Takahashi et al. 2009). Instead, one needs to use either large boxes, from which only a small region is taken to estimate the power spectrum (e.g. Bayer et al. 2023), or ‘separate universe simulations’, where multiple realisations with varying mean densities are simulated (Li et al. 2014). As shown in Barreira et al. (2018), the SSC is usually larger than the intra-survey non-Gaussian term and therefore the dominating contribution to the non-diagonal elements of the covariance. Estimating it correctly is therefore vital for cosmic shear analyses.

However, while the SSC for the power spectrum can be interpreted as capturing the clustering information at ℓ-modes outside a survey, the same interpretation is not necessarily accurate for statistics in real space. These real-space statistics, such as shear correlation functions (Kaiser 1992; Amon et al. 2022) or complete orthogonal sets of E-/B-integrals (COSTEBIs; Schneider et al. 2010; Asgari et al. 2020) are preferred for cosmological analyses, as they can be directly estimated from survey data. Here, we will derive the covariance for the estimator of general second-order shear statistics in real space to find an interpretation of the SSC of these statistics. We will show the following four key findings.

A33, page 1 of 10
First, for a ‘localised’ \(^1\) second-order shear statistic \(\Xi\), the full covariance \(C_\Xi\) can be obtained from correlation functions of the convergence field smoothed according to the chosen statistic. These correlation functions need to be known only on scales smaller than the survey area. Correlations on spatial scales larger than the survey area do not impact the covariance, including the Gaussian finite-field and the SSC term.

Second, the exact covariances of both \(\Xi\) and the power spectrum \(P\) do not scale inversely with the survey area. Instead, they also depend on the survey geometry via the SSC and a suppression term of the Gaussian covariance term. Under the assumption of an infinitely broad survey window function, these terms vanish and we recover the naive scaling with survey area.

Third, all parts of \(C_\Xi\) depend on the power spectrum and trispectrum at \(\ell\)-modes within and outside of the survey area. Therefore, the SSC term alone does not capture the clustering information at \(\ell\)-modes outside a survey.

Finally, it is, in general, not possible to transform the covariance of the power spectrum to the covariance of a real-space statistic. Such a conversion requires the assumption of an infinitely broad survey window.

The present paper is structured as follows: in Sect. 2, we discuss the covariance of the power spectrum and show the origin of the finite-field terms. In Sect. 3, we introduce an estimator for a real-space statistic \(\Xi\) and show how its covariance is related to correlation functions of \(\Xi\). We demonstrate in Sect. 4 that using correlation function estimates gives the same result (up to a prefactor) as the usual sample covariance approach. In Sect. 5, we connect the covariance of \(\Xi\) to the power spectrum and trispectrum and show that the finite-field terms for \(\Xi\) are given by the difference between the exact covariance and an approximation of the covariance for an infinitely broad survey window function. We conclude in Sect. 6.

Throughout this paper, we are working in the flat-sky limit. Any figures and calculations are performed using the parameters and simulations described in Appendix B. We note that we are not explicitly giving the dependence of the covariances on shape noise as its effect can be included by replacing the power spectrum \(P\) by \(P + \sigma_T^2/2n\) in all covariance expressions, where \(n\) is the galaxy number density and \(\sigma_T^2\) the two-component ellipticity dispersion.

2. Power spectrum covariance

Before we consider the covariance of a real space statistic, we first give an overview of the covariance of the power spectrum and the origin of the SSC based on Takada & Hu (2013). We are considering here the power spectrum of the weak-lensing convergence, which is the normalised surface mass density, which itself is related to the density contrast \(\delta\). In a flat universe, and at angular position \(\theta\) and comoving distance \(\chi\), the convergence is

\[
\kappa(\theta) = \frac{3H_0^2\Omega_m}{2\pi^2} \int_0^\infty dq q(\chi') \frac{\delta(\chi, \chi')}{\chi},
\]

where

\[
q(\chi') = \int_0^\infty dq' p(\chi') \frac{\chi'-\chi}{\chi'},
\]

with the Hubble constant \(H_0\), the matter density parameter \(\Omega_m\), the cosmic scale factor \(a(\chi)\) at \(\chi\), normalised to unity today, and the probability distribution \(p(\chi')\) of source galaxies with comoving distance.

The convergence power spectrum \(P(\ell)\) is defined by

\[
(2\pi)^2 \delta_D(\ell + \ell') P(\ell) = \langle \kappa(\ell) \kappa'(\ell') \rangle,
\]

where \(\kappa\) is the convergence and the tilde denotes Fourier transform. We assume a sample of simple geometry (i.e. continuous and without small-scale masks) with a window function \(W\), which is either zero or one, and survey area \(A\), given by

\[
A = \int d^2q W(\theta),
\]

Here, \(P\) can be estimated with the estimator

\[
P_\ell = \frac{1}{A^2} \int A_{A}(\ell_1) \int A_{A}(\ell_2) P(\ell_1) P(\ell_2) \delta_D(\ell_1 + \ell_2) + C_P^{SSC}(\ell_1, \ell_2),
\]

where \(A_{A}(\ell)\) denotes the size of the \(\ell\)-bin.

The usual form used for the covariance of \(P\) in the literature (Takada & Hu 2013; Krause & Eifler 2017) is

\[
P_\ell = \frac{1}{A^2} \int A_{A}(\ell_1) \int A_{A}(\ell_2) P(\ell_1) P(\ell_2) \delta_D(\ell_1 + \ell_2) + C_P^{SSC}(\ell_1, \ell_2),
\]

where \(T\) is the convergence trispectrum and \(C_P^{SSC}\) is the SSC of the power spectrum. Takada & Hu (2013) derive the SSC to be

\[
C_P^{SSC}(\ell_1, \ell_2) = \frac{1}{A^2} \int A_{A}(\ell_1) \int A_{A}(\ell_2) \int A_{A}(\ell_3) \int A_{A}(\ell_4) \int \frac{d^2q}{(2\pi)^2} T_{SSC}(\ell_1, \ell_2, q) \tilde{W}(q) \tilde{W}(-q),
\]

where \(T_{SSC}\) is part of the convergence trispectrum and given by Eq. (32) in Takada & Hu (2013). However, Eq. (5) is incomplete. To see this, we start from Eq. (4), so \(P_\ell\) is

\[
P_\ell = \frac{1}{A^2} \int A_{A}(\ell_1) \int A_{A}(\ell_2) \int A_{A}(\ell_3) \int A_{A}(\ell_4) \int \frac{d^2q}{(2\pi)^2} \tilde{W}(q),
\]

\[
\times \left[ \langle \kappa(\ell_1 + q_1) \kappa(-\ell_1 + q_2) \kappa(\ell_2 + q_3) \kappa(-\ell_2 + q_4) \rangle - \langle \kappa(\ell_1 + q_1) \kappa(-\ell_1 + q_2) \rangle \langle \kappa(\ell_2 + q_3) \kappa(-\ell_2 + q_4) \rangle \right] + \text{Permut.}
\]

The four-point function of \(\tilde{\kappa}\) can be decomposed into its connected and unconnected parts and can be written in terms of the power spectrum and trispectrum as

\[
\langle \kappa(\ell_1) \kappa(\ell_2) \kappa(\ell_3) \kappa(\ell_4) \rangle = \langle \kappa(\ell_1) \kappa(\ell_2) \kappa(\ell_3) \kappa(\ell_4) \rangle + \text{Permut.,}
\]

\[
= (2\pi)^3 T_4(\ell_1, \ell_2, \ell_3, \ell_4) \delta_D(\ell_1 + \ell_2 + \ell_3 + \ell_4) + \text{Permut.,}
\]

\[
= (2\pi)^3 T_4(\ell_1, \ell_2, \ell_3, \ell_4) \delta_D(\ell_1 + \ell_2 + \ell_3 + \ell_4) + \text{Permut.}
\]
Therefore, after suitably renaming of the $q_i$, using $q = q_1 + q_2$, and after evaluating the Dirac-functions and simplifying,

$$C_p(\ell_1, \ell_2) = \frac{2}{A^2} \int_{A_{\ell_1}} \frac{d^2 \ell_1'}{A_{\ell_1}} \int_{A_{\ell_2}} \frac{d^2 \ell_2'}{A_{\ell_2}} P(\ell_1') P(\ell_2')$$

$$\times \left[ W(\ell_1') W(\ell_2') W(-\ell_1' - \ell_2') \right]$$

$$\times \left[ 2 \int \frac{d^2 q}{(2\pi)^2} T(\ell_1', -\ell_1' + q, \ell_2', -\ell_2' - q) \hat{W}(q) \hat{W}(-q) \right].$$

(9)

Considering only $\ell$ at small spatial scales well within the survey area and ignoring the impact of masks, $\hat{W}(q)$ gives significant contributions only for $q \ll \ell$, and therefore we can approximate $P(\ell + q) = P(\ell)$. Then, as $\ell$ is either one or zero,

$$\int \frac{d^2 q}{(2\pi)^2} W(\ell + q) \hat{W}(q) = \int d^2 \alpha W^2(\alpha) e^{-i \alpha \ell}$$

$$= \int d^2 \alpha W(\alpha) e^{-i \alpha \ell} = \hat{W}(\ell),$$

(10)

we find

$$C_p(\ell_1, \ell_2) = \frac{2}{A^2} \int_{A_{\ell_1}} \frac{d^2 \ell_1'}{A_{\ell_1}} \int_{A_{\ell_2}} \frac{d^2 \ell_2'}{A_{\ell_2}} P(\ell_1') P(\ell_2')$$

$$\times \left[ W(\ell_1') W(\ell_2') W(-\ell_1' - \ell_2') \right]$$

$$\times \left[ 2 \int \frac{d^2 q}{(2\pi)^2} T(\ell_1', -\ell_1' + q, \ell_2', -\ell_2' - q) \hat{W}(q) \hat{W}(-q) \right].$$

(12)

where we introduced the geometry factor $G_A$, defined as

$$G_A(q) = \frac{1}{A^3} \int d^2 \alpha_1 d^2 \alpha_2 W(\alpha_1) W(\alpha_2) e^{-i q \cdot (\alpha_1 - \alpha_2)}.$$ 

(13)

The geometry factor contains the full dependence of $C_p$ on the survey area.

However, $C_p$ in Eq. (12) is not the same as $C_p^{\mathrm{lin}}$ in Eq. (5). To convert $C_p$ to $C_p^{\mathrm{lin}}$, we need to perform the ‘large-field approximation’. For this approximation, we note that $G_A$ is related to the function $E_A(q)$, which for a point $\alpha$ inside $A$ gives the probability that a point $\alpha + \eta$ is also inside $A$, and is given by (Heydenreich et al. 2020; Linke et al. 2023)

$$E_A(q) = \frac{1}{A} \int_A d^2 \alpha W(\alpha + \eta).$$

(14)

With $E_A$, $G_A$ is

$$G_A(q) = \frac{1}{A} \int d^2 \alpha E_A(\alpha) e^{i q \cdot \alpha}.$$ 

(15)

The large-field approximation now assumes that $A$ is infinitely large so that $E_A(\alpha)$ is unity for all $\alpha$. Then,

$$G_A(q) \to \frac{(2\pi)^2}{A} \delta_D(q).$$

(16)

We define the result of $C_p$ under this approximation as $C_p^{\text{lin}}$, given as

$$C_p^{\text{lin}}(\ell_1, \ell_2) = \frac{1}{A^2} \int_{A_{\ell_1}} \frac{d^2 \ell_1'}{A_{\ell_1}} \int_{A_{\ell_2}} \frac{d^2 \ell_2'}{A_{\ell_2}} P(\ell_1') P(\ell_2')$$

$$\times \left[ 2 \int \frac{d^2 q}{(2\pi)^2} T(\ell_1', -\ell_1' + q, \ell_2', -\ell_2' - q) \hat{W}(q) \hat{W}(-q) \right].$$

These are exactly the first two terms of $C_p^{\text{lin}}$. Consequently, the full covariance $C_p$ can be written as the sum of $C_p^{\text{lin}}$ and the finite-field terms $C_p^{\text{FF}}$. The finite-field terms are given by

$$C_p^{\text{FF}} = 2 \int_{A_{\ell_1}} \frac{d^2 \ell_1'}{A_{\ell_1}} \int_{A_{\ell_2}} \frac{d^2 \ell_2'}{A_{\ell_2}} P(\ell_1') P(\ell_2')$$

$$\times \left[ G_A(\ell_1' + \ell_2') - \frac{(2\pi)^2}{A} \delta_D(\ell_1' + \ell_2') \right]$$

$$\times \left[ \frac{1}{A} \int \frac{d^2 q}{(2\pi)^2} T(\ell_1', -\ell_1' + q, \ell_2', -\ell_2' - q) \hat{W}(q) \hat{W}(-q) \right]$$

$$+ \frac{1}{A} \int \frac{d^2 q}{(2\pi)^2} T(\ell_1, -\ell_1 + q, \ell_2, -\ell_2 - q) \hat{W}(q) \hat{W}(-q).$$

(17)

Takada & Hu (2013) showed that the trispectrum can be approximated by

$$T(\ell_1, -\ell_1 + q, \ell_2, -\ell_2 - q) \approx T(\ell_1, -\ell_1 + q, \ell_2, -\ell_2 - q) + T_{\text{SSC}}(\ell_1, \ell_2, q),$$

(19)

with the individual parts given as Limber integrals over the trispectrum expressions in their Eq. (32). Notably, $T_{\text{SSC}}$ is proportional to the linear matter power spectrum $P_l(q/\chi)$ at wavevector $q/\chi$ and comoving distance $\chi$. Therefore, $T_{\text{SSC}}$ vanishes for $|q| = 0$. With Eq. (19), one obtains

$$C_p^{\text{FF}} = 2 \int \frac{d^2 \ell_1'}{A_{\ell_1}} \int \frac{d^2 \ell_2'}{A_{\ell_2}} P(\ell_1') P(\ell_2')$$

$$\times \left[ \frac{1}{A} \hat{W}(\ell_1') \hat{W}(\ell_2') \hat{W}(-\ell_1' - \ell_2') - (2\pi)^2 \delta_D(\ell_1' + \ell_2') \right]$$

$$\times \left[ \frac{1}{A} \int \frac{d^2 q}{(2\pi)^2} T_{\text{SSC}}(\ell_1, \ell_2, q) \hat{W}(q) \hat{W}(-q) \right]$$

$$= C_p^{\text{FF}} + C_p^{\text{SSC}}.$$
Gaussian covariance. As becomes apparent in our derivation, the SSC and the Gaussian suppression term can both be seen as corrections for the large-field approximation, which is why we collectively refer to them as finite-field terms. We note that \( C^\xi_0(\ell_1, \ell_2) \) only depends on the power spectrum and trispectrum at \( \ell \) close to \( \ell_1 \) and \( \ell_2 \), because the integrals in Eq. (17) only go over the \( A_{\ell}(\ell) \). In contrast, the SSC depends on the trispectrum at all \( \ell \)-modes because the integral over the \( q \) extends over all \( \mathbb{R}^2 \). Therefore, for the power spectrum, the SSC captures the dependence of the covariance on modes outside the survey area. We will see in the following sections that this intuition does not hold for real space statistics. For this, in the following section we derive the covariance for an estimator of \( \Xi \) in real space.

3. Covariance of a real space statistic

We now consider a localised second-order shear statistic \( \Xi \) in real space that can be written as

\[
\Xi(\theta, \alpha) = \int d^2\theta_1 \int d^2\theta_2 U_1(\theta, \theta_1 - \alpha) U_2(\theta, \theta_2 - \alpha) \langle \kappa(\theta_1) \kappa(\theta_2) \rangle,
\]

where the \( U_i \) are two filter functions of scale \( \theta \), and \( \kappa \) is the weak-lensing convergence. Examples of such a statistic are the COSEBIs \( \langle M_{AB}^\gamma \rangle \) (Schneider et al. 2010) or the second-order aperture statistics \( \langle M_{AB}^\gamma \rangle \) (Schneider et al. 1998). We note that while we are writing the statistic here in terms of the (unobservable) convergence for simplicity, \( \Xi \) can also be written in terms of the weak lensing shear \( \gamma \), which is observable if \( U_1 \) and \( U_2 \) are compensated filter functions. This is the case for both COSEBIs and \( \langle M_{AB}^\gamma \rangle \).

To estimate the statistic from a convergence field \( \kappa \) of size \( A' \), we can convolve \( \kappa \) with the filter functions and then average over the pixel values. However, due to the finiteness of the survey area, the convolution result at the survey boundaries is biased. Therefore, the average needs to be taken over a smaller area \( A \), which excludes the borders of the field (see Fig. 1). This leads to the estimator

\[
\hat{\Xi}(\theta) = \frac{1}{A} \int d^2\alpha \int_{i=1}^{2} d^2 \theta_i U_i(\theta, \theta_i - \alpha) \kappa_i,
\]

(22)

where \( \kappa_i = \kappa(\theta_i) \). Under the assumption that \( U_i(\theta, \theta_i - \alpha) \) vanishes for \( \theta_i \) outside of \( A' \) for all \( \alpha \in A \), we can replace the integral over \( A' \) by an integral over the whole \( \mathbb{R}^2 \). With this, and the survey window function \( W(\theta) \), which is one for \( \theta \) inside \( A \) and zero otherwise,

\[
\hat{\Xi}(\theta) = \frac{1}{A} \int d^2\alpha W(\alpha) \int_{i=1}^{2} d^2 \theta_i U_i(\theta, \theta_i - \alpha) \kappa_i.
\]

(23)

The covariance of \( \hat{\Xi} \) is

\[
C_{\Xi}(\theta_1, \theta_2) = \langle \hat{\Xi}(\theta_1) \hat{\Xi}(\theta_2) \rangle - \langle \hat{\Xi}(\theta_1) \rangle \langle \hat{\Xi}(\theta_2) \rangle.
\]

(24)

With Eq. (23),

\[
\langle \hat{\Xi}(\theta_1) \hat{\Xi}(\theta_2) \rangle = \frac{1}{A^2} \int d^2\alpha_1 \int d^2\alpha_2 W(\alpha_1) W(\alpha_2) \times \int_{i=1}^{2} d^2 \theta_i \int d^2 \theta_j \int d^2 \theta_k \langle \kappa(\theta_1) \kappa(\theta_2) \rangle
\]

\[
\times U_1(\theta_1, \theta_1 - \alpha_1) U_2(\theta_2, \theta_2 - \alpha_2) \times U_1(\theta_1', \theta_1' - \alpha_1) U_2(\theta_2', \theta_2' - \alpha_2)
\]

\[
= \frac{1}{A^2} \int d^2\alpha \int d^2\alpha_2 W(\alpha_1) W(\alpha_2) \times \langle \kappa(\theta_1, \alpha_1) \kappa(\theta_2, \alpha_1) \kappa(\theta_2', \alpha_2) \kappa(\theta_2', \alpha_2) \rangle,
\]

(26)

where we introduced the smoothed convergence field \( \kappa_\alpha \),

\[
\kappa_\alpha(\theta, \alpha) = \int d^2\theta U(\theta, \theta - \alpha) \kappa(\theta).
\]

(27)

For ease of notation we define the field \( \nu^2(\theta_1, \alpha) := \kappa_\alpha(\theta_1, \alpha) \kappa_\alpha(\theta_2, \alpha) \). The expectation value in Eq. (26) is a second-order correlation function \( \xi_{\nu^2} \) of this field and is defined as

\[
\xi_{\nu^2}(\theta_1, \theta_2, \eta) = \langle \nu^2(\theta_1, \alpha) \nu^2(\theta_2, \alpha + \eta) \rangle
\]

\[
= \langle \kappa(\theta_1, \alpha) \kappa(\theta_2, \alpha) \kappa(\theta_2', \alpha + \eta) \kappa(\theta_2', \alpha + \eta) \rangle.
\]

(28)

With \( \xi_{\nu^2} \) and the function \( E_{\alpha} \), defined in Eq. (14),

\[
C_{\Xi}(\theta_1, \theta_2) = \frac{1}{A^2} \int d^2\alpha_1 \int d^2\alpha_2 W(\alpha_1) W(\alpha_2) \times \xi_{\nu^2}(\theta_1, \theta_2, |\alpha_1 - \alpha_2|) - \langle \hat{\Xi}(\theta_1) \rangle \langle \hat{\Xi}(\theta_2) \rangle
\]

\[
= \int d^2\eta E_{\alpha}(\eta) \langle \xi_{\nu^2}(\theta_1, \theta_2, \eta) - \langle \hat{\Xi}(\theta_1) \rangle \langle \hat{\Xi}(\theta_2) \rangle \rangle.
\]

(29)

In this expression, the covariance of \( \Xi \) can be inferred from a two-point correlation function \( \xi_{\nu^2} \). Notably, \( \xi_{\nu^2} \) needs to be known only for \( \eta \) inside the survey area, as \( E_{\alpha} \) vanishes outside. Consequently, \( C_{\Xi} \) does not depend on any information on spatial scales larger than the survey area.

This finding might appear to contradict Eq. (12), which shows that the power spectrum covariance (in particular, the SSC) clearly depends on Fourier modes larger than the survey. However, Eq. (29) takes a ‘real space’ view. The real-space correlation function \( \xi_{\nu^2} \) at spatial scales within the survey window is impacted by the power spectrum and trispectrum at Fourier modes.
modes larger than the survey window. Therefore, while the covariance estimation in Fourier space requires all modes, in real space, we can limit ourselves to spatial scales within the survey window. Thus, while information from ‘super-survey’ Fourier modes is needed, no information on ‘super-survey’ spatial scales is required.

4. Covariance estimation from correlation functions

An interesting aspect of Eq. (29) is that it allows the covariance of \( \hat{\xi} \) to be quickly evaluated for various survey geometries. Given a single, full-sky simulation of the convergence field, \( \xi_{\omega} \) and \( \langle \hat{\xi}(\theta_1) \rangle \) can be measured with high accuracy. Then, \( C_2 \) can be estimated for any survey geometry, simply by adapting the \( E_4(\eta) \) in Eq. (29).

Another application of Eq. (29) is to estimate \( \xi_{\omega} \) as the average of \( \xi_{\omega}^{(i)} \) measured for each realisation \( i \) of an ensemble of simulations. This approach delivers a covariance estimate that coincides with the standard sample covariance estimate. To see this, we use

\[
\xi_{\omega}^{(i)}(\theta_1, \theta_2; \eta) = \frac{1}{A} \frac{1}{E_4(\eta)} \int d^2 \alpha \sum_{n=1}^{N} W_A(\alpha) \xi_{\omega}^{(i)}(\alpha) \text{W}(\alpha + \eta),
\]

and

\[
\xi_{\omega}(\theta_1, \theta_2; \eta) = \frac{1}{N} \sum_{i=1}^{N} \xi_{\omega}^{(i)}(\theta_1, \theta_2; \eta),
\]

where \( N \) is the number of realisations. We further use

\[
\langle \hat{\xi}(\theta) \rangle = \frac{1}{A} \frac{1}{N} \sum_{i=1}^{N} \int d^2 \alpha W_A(\alpha) \xi_{\omega}^{(i)}(\alpha, \theta)
\]

\[
= \frac{1}{N} \sum_{i=1}^{N} \hat{\xi}(\theta).
\]

Then, with Eq. (29), the covariance \( C_{2}^{\text{corr}} \) from the correlation functions is

\[
C_{2}^{\text{corr}}(\theta_1, \theta_2) = \frac{1}{N A^2} \sum_{i=1}^{N} \int d^2 \eta \int d^2 \alpha W_A(\alpha) W_A(\alpha + \eta)
\]

\[
\times \xi_{\omega}^{(i)}(\alpha, \theta_1) \xi_{\omega}^{(i)}(\alpha + \eta, \theta_2) - \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \hat{\xi}(\theta_1) \hat{\xi}(\theta_2) - \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \hat{\xi}(\theta_1) \hat{\xi}(\theta_2)
\]

\[
= \frac{1}{N} \sum_{i=1}^{N} \hat{\xi}(\theta_1) \hat{\xi}(\theta_2) - \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \hat{\xi}(\theta_1) \hat{\xi}(\theta_2)
\]

\[
= \frac{N}{N-1} C_{2}^{\text{sim}}(\theta_1, \theta_2),
\]

with the sample covariance \( C_{2}^{\text{sim}} \) defined as

\[
C_{2}^{\text{sim}} = \frac{1}{N-1} \sum_{i=1}^{N} \left( \hat{\xi}(\theta_1) - \langle \hat{\xi}(\theta_1) \rangle \right) \left( \hat{\xi}(\theta_2) - \langle \hat{\xi}(\theta_2) \rangle \right).
\]

Consequently, estimating the covariance from the average \( \kappa^2 \) correlation function of an ensemble of simulations is equivalent to estimating it from the \( \Xi \) measured individually on each realisation.

We validate this finding by specifying our statistic in the remainder of this section to the second-order aperture mass \( \langle M_{2}^{\theta} \rangle(\theta) \) with the filter function by Crittenden et al. (2002):

\[
U_i(\theta, \hat{\theta}) = U_2(\theta, \hat{\theta}) = U(\theta, \hat{\theta}) = \frac{1}{2\pi \theta^2} \left( 1 - \frac{\theta^2}{2\eta^2} \right) \exp \left( -\frac{\theta^2}{2\eta^2} \right)
\]

(37)

Then, setting \( \kappa^2 = M_{2}^{\theta}, \) which is given as

\[
M_{2}^{\theta}(\theta, \hat{\theta}) = \int d^2 \alpha_1 \int d^2 \alpha_2 U(\theta, |\theta - \alpha_1|) U(\theta, |\theta - \alpha_2|) s(\alpha_1) s(\alpha_2),
\]

and \( \xi_{\omega} \) becomes

\[
\xi_{\omega}(\theta_1, \theta_2; \eta) = \langle M_{2}^{\theta}(\theta_1; \theta, \eta) M_{2}^{\theta}(\theta_2; \theta + \eta) \rangle.
\]

(39)

This expression can be generalised to higher orders of \( M_{2}^{\theta} \) with \( n > 2. \) Porth & Smith (2021) give the expressions for the variance of \( M_{n}^{\theta} \) estimated with the so-called direct estimator as a function of the correlation functions of \( M_{n}^{\theta} \) for a single aperture. We validate Eq. (29) for the aperture statistics by measuring the covariance \( C_{M_{2}^{\theta}}^{\text{corr}} \) of \( M_{2}^{\theta} \) in convergence maps from the SciNet Lightcone Simulations (SLICS; Harnois-Déraps & van Waerbeke 2015), the details of which are described in Appendix B.1. We estimate the covariance of \( \langle M_{2}^{\theta} \rangle \) using the sample covariance and correlation function-based approach (see Appendix B.2).

The covariance estimates and their difference are shown in the first two panels of Fig. 2. The two estimates almost coincide. Accordingly, the correlation-function-based approach captures the full covariance, even though \( \xi_{\omega} \) is known only for scales within the survey area. Consequently, as expected from Eq. (29), no information on spatial scales outside the survey is needed for an accurate covariance estimate.

5. Connection between real and Fourier space statistics

The statistic \( \Xi \) defined in Eq. (21) can be expressed in terms of the power spectrum as

\[
\Xi(\theta) = \int \frac{d^2 \ell}{(2\pi)^2} \hat{U}_1(\theta, \ell) \hat{U}_2(\theta, \ell) P(\ell),
\]

(40)

where the \( \hat{U}_i \) are the Fourier transforms of the filter functions \( U_i. \) A common strategy (e.g., Joachimi et al. 2021; Friedrich et al. 2021) to model the covariance of \( \Xi \) is to use

\[
C_{\Xi}(\theta_1, \theta_2) = \int \frac{d^2 \ell_1}{(2\pi)^2} \int \frac{d^2 \ell_2}{(2\pi)^2} \hat{U}_1(\theta_1, \ell_1) \hat{U}_2(\theta_2, \ell_2) C_P(\ell_1, \ell_2) \times \hat{U}_1(\theta_1, \ell_2) \hat{U}_2(\theta_2, \ell_1)
\]

(41)

However, this approach is not necessarily correct. To show this, we relate \( C_{\Xi} \) to the power- and trispectrum to compare it to previous expressions of second-order shear covariances and to discuss

Linke, L., et al.: A&A proofs, manuscript no. aa46225-23
the SSC for real-space statistics. As shown in Appendix A, $C_2(\theta_1, \theta_2)$ can be expressed as

$$C_2(\theta_1, \theta_2) = \frac{d^2 \ell_1}{(2\pi)^2} \frac{d^2 \ell_2}{(2\pi)^2} G_A(\ell_1 + \ell_2) P(\ell_1) P(\ell_2)$$

(42)

\[
\times \left( \hat{U}_1(\theta_1, \ell_1) \hat{U}_1(\theta_2, \ell_1) \hat{U}_2(\theta_1, \ell_2) \hat{U}_2(\theta_2, \ell_2) + \hat{U}_1(\theta_1, \ell_1) \hat{U}_1(\theta_2, \ell_2) \hat{U}_2(\theta_1, \ell_2) \hat{U}_2(\theta_2, \ell_1) \right) \\
\times \int \frac{d^2 \ell_1}{(2\pi)^2} \frac{d^2 \ell_2}{(2\pi)^2} G_A(\ell_1 + \ell_2) \\
\times T(\ell_1, \ell_2, \ell_3, -\ell_1 - \ell_2 - \ell_3) \hat{U}_1(\theta_1, \ell_1) \hat{U}_2(\theta_1, \ell_2) \\
\times \hat{U}_1(\theta_2, \ell_3) \hat{U}_2(\theta_2, \ell_1 + \ell_2 + \ell_3).
\]

One notices a Gaussian and a non-Gaussian part of the covariance, with the Gaussian part depending on the power spectrum and the non-Gaussian part depending on the trispectrum, similar to the exact covariance of the power spectrum in Eq. (9). However, $C_2$ is not in the form of Eq. (41) which is a weighted integral over $C_P$. We also notice three differences when directly comparing Eq. (42) to Eqs. (E.1) and (E.7) in Joachimi et al. (2021). First, neither the Gaussian term nor the non-Gaussian term scales with the inverse of the survey area $A$, and instead both show a more complicated dependence on survey geometry via $G_A$. Second, the non-Gaussian term depends on the trispectrum for all $\ell$-configurations, not simply for parallelograms with $\ell_3 = \ell_1$. Third, the SSC does not appear as an additional term.

To reconcile Eq. (42) with the expressions in Joachimi et al. (2021), we need to perform the large-field approximation. As mentioned in Sect. 2, under this approximation, $G_A$ is proportional to a Dirac delta, so the covariance becomes

$$C_2^\infty(\theta_1, \theta_2) = \frac{2}{A} \int \frac{d^2 \ell_1}{(2\pi)^2} P^2(\ell_1) \hat{U}_1(\theta_1, \ell_1) \hat{U}_1(\theta_2, \ell_1) \hat{U}_2(\theta_1, \ell_1) \hat{U}_2(\theta_2, \ell_1)$$

(43)

\[
+ \frac{1}{A} \int \frac{d^2 \ell_1}{(2\pi)^2} \int \frac{d^2 \ell_2}{(2\pi)^2} G_A(\ell_1 + \ell_2) \\
\times T(\ell_1, -\ell_1, \ell_2, -\ell_2) \\
\times \hat{U}_1(\theta_1, \ell_1) \hat{U}_2(\theta_1, \ell_1) \hat{U}_1(\theta_2, \ell_2) \hat{U}_2(\theta_2, \ell_2).
\]

This is equivalent to the commonly used expressions for the Gaussian and intra-survey non-Gaussian covariance for a second-order statistic (Eqs. (E.1) and (E.7) in Joachimi et al. 2021). In particular, we recover the scaling with the inverse survey area and the dependence on parallelogram $\ell$-configurations for the non-Gaussian part. Consequently, the sum of the Gaussian and intra-survey non-Gaussian covariance can be considered as an approximation of the exact covariance in Eq. (42) for very large survey windows.

A comparison of Eq. (43) to Eq. (17) shows that $C_2^\infty$ is given by the integral over the large-field approximation of the power spectrum covariance $C_P^\infty$. Therefore, the approach to obtain the covariance of the real-space statistic from the power spectrum covariance is correct if the large-field approximation holds.
However, the large-field approximation neglects a significant part of the covariance. We can see this for $\langle M_{\ell}^2 \rangle$, for which we calculate Eq. (43) according to Appendix B.3. The third panel of Fig. 2 shows the covariance for $\langle M_{\ell}^2 \rangle$, which is modelled with Eq. (43), along with the fractional difference from the sample covariance estimate from the SLICS (see Appendix B). The approximation $C_{\ell}^{\infty}$ is significantly too small, with deviations of more than five times the statistical uncertainty on the sample covariance. This large difference is not surprising given that the survey area $A$ here is only $62 \text{ deg}^2$. However, even for the Kilo-degree survey (KiDS) data release KiDS-1000, $C_{\ell}$ cannot describe the covariance of the cosmic shear band powers (Joachimi et al. 2021).

In analogy to the power spectrum case, we define the difference between $C_{\ell}$ and $C_{\ell}^{\infty}$ as finite-field terms $C_{\ell}^{\text{FF}}$ for $\hat{\Sigma}$. To calculate this term, we can perform the same approximation as that performed for the power spectrum covariance, namely that the modes $\ell, \ell_1$, and $\ell_2$ are large compared to the modes $q$ on which $G_A$ varies. With this,

$$C_{\ell}^{\text{FF}}(\theta_1, \theta_2) = \int \frac{d^2 \ell_1}{(2\pi)^2} \int \frac{d^2 \ell_2}{(2\pi)^2} \tilde{U}_1(\ell_1, \ell_1) \tilde{U}_2(\ell_2, \ell_2) \tilde{U}_2(\ell_1, \ell_2) \tilde{U}_2(\ell_2, \ell_1) \times \left\{ P(\ell_1) P(\ell_2) \left[ G_A(\ell_1 + \ell_2) - \frac{(2\pi)^2}{A} \delta_0(\ell_1 + \ell_2) \right] \right\}$$

$$= C_{\ell}^{\text{FF,G}}(\theta_1, \theta_2) + C_{\ell}^{\text{SSC}}(\theta_1, \theta_2),$$

where $T_{\text{SSC}}$ is the same as in Eq. (19). To show that $C_{\ell}$ is indeed composed of $C_{\ell}^{\infty}$ and $C_{\ell}^{\text{FF}}$, we show in Fig. 2 the modelled covariance of $\langle M_{\ell}^2 \rangle$, including the finite-field terms and the sample covariance estimate in the SLICS. We see that the covariances coincide within the bootstrap uncertainties of the sample covariance.

The first summand $C_{\ell}^{\text{FF,G}}$ in Eq. (44) depends on the power spectrum and therefore already occurs for Gaussian fields. In general, it is negative and decreases the magnitude of the other Gaussian covariance term. This effect has already been noticed for the covariance of shear correlation functions in Sato et al. (2011).

In Joachimi et al. (2021), only the second summand $C_{\ell}^{\text{SSC}}$ in Eq. (44) is considered (see their Eq. (E.10)), while the first summand $C_{\ell}^{\text{FF}}$ is neglected. However, at least for the second-order aperture statistics, this neglect has only a small impact, because the $C_{\ell}^{\text{FF,G}}$ is small compared to $C_{\ell}^{\text{SSC}}$. This can be seen in Fig. 3, where we compare the full finite-field term to $C_{\ell}^{\text{SSC}}$ for the $\hat{M}_{\ell}^2$ covariance in the SLICS. The first finite-field term accounts for less than 5% of the total $C_{\ell}^{\text{FF}}$. Therefore, using just the SSC is accurate enough to describe the sample covariance in the SLICS.

It is important to note here that both $C_{\ell}^{\infty}$ and $C_{\ell}^{\text{FF}}$ depend on the power spectrum and trispectrum throughout the $\ell$-space. Consequently, modes larger than the survey area impact not only the SSC term but also $C_{\ell}^{\infty}$. The common interpretation that the SSC captures the full impact of ‘super-survey’ $\ell$-modes is incorrect for real-space statistics.

6. Conclusion

We derived the full covariance $C_{\ell}$ for a localised, second-order statistic $\Sigma$ in real space and compared it to the covariance $C_{\ell}$ of the lensing power spectrum. Both covariances depend on the exact survey geometry. Under the large-field approximation, which is the limit for a broad window function, these covariances reduce to approximated terms $C_{\ell}^{\infty}$ and $C_{\ell}^{\text{FF}}$, which scale with the inverse survey area. While we define $\Sigma$ in terms of the convergence $\kappa$, we note that for compensated filter functions, $\Sigma$ can be equivalently written in terms of the weak lensing shear, so all our conclusions are valid for such shear statistics as well.

We find that the difference between $C_{\ell}$ and $C_{\ell}^{\infty}$ gives rise to two terms, which we collectively refer to as finite-field terms $C_{\ell}^{\text{FF}}$. While $C_{\ell}^{\text{FF}}$ scales linearly with inverse survey area, the finite-field terms show a complicated dependence on the survey geometry. One of these terms is the SSC $C_{\ell}^{\text{SSC}}$, which is in contrast to the other terms, depends on the matter trispectrum on all $\ell$-modes including those larger than the survey area.

The covariance $C_{\ell}$ can also be written as the sum of a large-field approximation $C_{\ell}^{\infty}$ and finite-field terms $C_{\ell}^{\text{FF}}$. However, both $C_{\ell}^{\infty}$ and $C_{\ell}^{\text{FF}}$ depend on the power spectrum and trispectrum at all $\ell$-modes, including those larger than the survey area. Therefore, the label ‘super-sample’ is slightly misleading for the SSC of a real-space statistic, as it is not the only term containing super-sample information.
The second finite field term in addition to the SSC depends on the power spectrum and is already present in Gaussian fields. This term essentially decreases the Gaussian covariance, an effect already noted for shear correlation functions by Sato et al. (2011). While we show here that neglecting this term (e.g. Joachimi et al. 2021) is accurate for the second-order aperture statistics \( \hat{C}_3 \), found that for the shear correlation functions \( \xi_+ \) and \( \xi_- \) ignoring this effect leads to a significant overestimation of the covariance. We suspect this difference occurs because the aperture mass filter function is compensated. Therefore, the average aperture mass vanishes and the integral in the Gaussian finite field term is essentially ‘cut-off’ at small \( \ell \). However, for the correlation functions, small \( \ell \) need to be taken into account.

We show that the covariance \( C_3 \) of the real-space statistic cannot be obtained from the power spectrum covariance without the large-field approximation. The commonly used transformation between power spectrum covariance and real-space covariance only holds for \( C_3^\text{ap} \) and \( C_3^\infty \). This finding is not surprising. A linear transform between the \( C_3 \) and \( C_3^\infty \) is mathematically only possible if the estimators \( \hat{P} \) and \( \hat{\Xi} \) are related linearly. While \( \hat{P} \) and \( \hat{\Xi} \) are indeed related by a Fourier transform, the estimators are generally not. Consequently, one would not expect to simply transform one covariance into the other.

Finally, we demonstrate that \( C_3 \) can be fully determined from correlation functions of smoothed convergence maps known only at spatial scales smaller than the survey area. We show that if one estimates the correlation functions from an ensemble of simulations, this approach gives the same result (up to a prefactor) as estimating the statistics directly on each realisation and taking the sample covariance. This indicates that correlations outside the survey area do not influence the covariance of a second-order shear statistic in real space.

This finding is not surprising given that Schneider et al. (2002) already showed that the Gaussian covariance of shear correlation functions is given by second-order correlation functions of galaxy ellipticities known inside the survey area. However, it provides an alternative method to estimate covariances for varying survey geometries quickly. If the \( \kappa' \)-correlation function is known accurately, for example from a single, high-precision simulation, the covariance of the statistics \( \hat{\Xi} \) can be calculated by specifying the geometry function \( E_A \) and evaluating Eq. (29). Consequently, for this case, changing the survey geometry would not require new simulations to be produced for covariance estimates, provided the redshift distribution of source galaxies remains the same.

While this paper is concerned with second-order statistics only, finite-field terms also play a significant role for higher-order statistics. Linke et al. (2023) showed that for third-order aperture statistics \( M_3^\text{ap} \), several finite-field terms occur, with one of them dominating the Gaussian covariance. In contrast, for the convergence probability distribution Uhlemann et al. (2023) showed that the large-field approximation (in their notation \( P_\ell(\theta) \approx \theta \)) leads to model covariances in agreement with simulations, and so the SSC is less important for their statistic. This indicates that the impact of SSC depends on the considered statistic.

Acknowledgements. We thank the anonymous referee, as well as Alex Barreira and Alex Hall, for their thoughtful and valuable comments. Funded by the TRAMatter (University of Bonn) as part of the Excellence Strategy of the federal and state governments. This work has been supported by the Deutsche Forschungsgemeinschaft through the project SCHN 342/15-1 and DFG SCHN 342/13. PAB and SH acknowledge support from the German Academic Scholarship Foundation. LP acknowledges support from the DLR grant 50OE2002. We would like to thank Joachim Harnois-Déraps for making public the SLICS mock data, which can be found at http://slics.roe.ac.uk/. We thank Oliver Friedrich and Nick Westler for helpful discussions.

References
Uhlemann, C., Friedrich, O., Boyle, A., et al. 2023, Open J. Astrophys., 6, 1

A&A 681, A33 (2024)
Appendix A: Expressing $C_{\hat{\kappa}}$ in terms of the power- and trispectrum

In this Appendix, we relate $C_{\hat{\kappa}}$ to the power spectrum $P$ and the trispectrum $T$. We start from Eq. (25) and rewrite it in terms of the Fourier transform $\tilde{\kappa}$ of the convergence, using

$$\int d^2\vartheta \, \kappa(\vartheta) \, U_{1/2}(\vartheta, \vartheta - \alpha) = \int \frac{d^2\ell}{(2\pi)^2} \tilde{\kappa}(\ell) \, \tilde{U}_{1/2}(\theta, \ell). \quad (A.1)$$

This leads to

$$\left\langle \tilde{\xi}(\theta_1) \, \tilde{\xi}(\theta_2) \right\rangle = \frac{1}{A^2} \int d^2\alpha \, W_{\alpha}(\alpha) \, W_{\alpha}(\alpha_2) \int \frac{d^2\ell_1}{(2\pi)^2} \int \frac{d^2\ell_2}{(2\pi)^2} \int \frac{d^2\ell_3}{(2\pi)^2} \int \frac{d^2\ell_4}{(2\pi)^2} \times \tilde{U}_{1}(\theta_1, \ell_1) \, \tilde{U}_{2}(\theta_2, \ell_2) \, \tilde{U}_{1}(\theta_3, \ell_3) \, \tilde{U}_{2}(\theta_4, \ell_4) \, e^{i(\ell_1 + \ell_2)\alpha_1 + i(\ell_3 + \ell_4)\alpha_2} \left(\tilde{\kappa}(\ell_1) \tilde{\kappa}(\ell_2) \tilde{\kappa}(\ell_3) \tilde{\kappa}(\ell_4)\right) \quad (A.2)$$

We decompose the four-point function as suggested by Eq. (8). This leads to

$$\left\langle \tilde{\xi}(\theta_1) \, \tilde{\xi}(\theta_2) \right\rangle = \frac{1}{A^2} \int d^2\alpha \, W_{\alpha}(\alpha) \, W_{\alpha}(\alpha_2) \int \frac{d^2\ell_1}{(2\pi)^2} \int \frac{d^2\ell_2}{(2\pi)^2} \int \frac{d^2\ell_3}{(2\pi)^2} \int \frac{d^2\ell_4}{(2\pi)^2} \times \tilde{U}_{1}(\theta_1, \ell_1) \, \tilde{U}_{2}(\theta_2, \ell_2) \, \tilde{U}_{1}(\theta_3, \ell_3) \, \tilde{U}_{2}(\theta_4, \ell_4) \, e^{i(\ell_1 + \ell_2)\alpha_1 + i(\ell_3 + \ell_4)\alpha_2} \times \left[P(\ell_1) \, P(\ell_2) \, (2\pi)^4 \, \delta_0(\ell_1 + \ell_2) \, \delta_0(\ell_3 + \ell_4) + P(\ell_1) \, P(\ell_2) \, (2\pi)^4 \, \delta_0(\ell_1 + \ell_3) \, \delta_0(\ell_2 + \ell_4) \right. \\
+ \left.P(\ell_1) \, P(\ell_2) \, (2\pi)^4 \, \delta_0(\ell_1 + \ell_4) \, \delta_0(\ell_2 + \ell_3) + T(\ell_1, \ell_2, \ell_3, \ell_4)(2\pi)^2 \, \delta_0(\ell_1 + \ell_2 + \ell_3 + \ell_4)\right]\right) \quad (A.3)$$

Given that

$$\left\langle \tilde{\xi}(\theta) \right\rangle = \frac{1}{A} \int d^2\alpha \, W_{\alpha}(\alpha) \int \frac{d^2\ell}{(2\pi)^2} P(\ell_1) \, \tilde{U}_{1}(\theta, \ell) \, \tilde{U}_{2}(\theta, \ell) \quad (A.4)$$

the covariance can be written

$$C_{\hat{\kappa}}(\theta_1, \theta_2) = \frac{1}{A} \int d^2\alpha \, W_{\alpha}(\alpha) \, W_{\alpha}(\alpha_2) \left[ \int \frac{d^2\ell_1}{(2\pi)^2} \int \frac{d^2\ell_2}{(2\pi)^2} \int \frac{d^2\ell_3}{(2\pi)^2} \int \frac{d^2\ell_4}{(2\pi)^2} e^{i(\ell_1 + \ell_2)\alpha_1 \alpha_2} P(\ell_1) \, P(\ell_2) \times \left(\tilde{U}_{1}(\theta_1, \ell_1) \, \tilde{U}_{2}(\theta_2, \ell_2) \, \tilde{U}_{1}(\theta_3, \ell_3) \, \tilde{U}_{2}(\theta_4, \ell_4) \right) \right. \\
+ \left. \int \frac{d^2\ell_1}{(2\pi)^2} \int \frac{d^2\ell_2}{(2\pi)^2} \int \frac{d^2\ell_3}{(2\pi)^2} \int \frac{d^2\ell_4}{(2\pi)^2} \delta_0(\ell_1 + \ell_2 - \ell_3 - \ell_4) \right) \quad (A.5)$$

$$= \int \frac{d^2\ell_1}{(2\pi)^2} \int \frac{d^2\ell_2}{(2\pi)^2} G_A(\ell_1 + \ell_2) P(\ell_1) \, P(\ell_2) \times \left(\tilde{U}_{1}(\theta_1, \ell_1) \, \tilde{U}_{2}(\theta_2, \ell_2) \, \tilde{U}_{1}(\theta_3, \ell_3) \, \tilde{U}_{2}(\theta_4, \ell_4) \right) \left. \right) \quad (A.6)$$

where we introduce the geometry factor $G_A$ defined in Eq. (13).

Appendix B: Details on validation measurements

All calculations in this work are performed for a flat $\Lambda$CDM cosmology with dimensionless Hubble-parameter $h = 0.69$, clustering parameter $\sigma_8 = 0.83$, matter density $\Omega_m = 0.29$, baryon density $\Omega_b = 0.047$, and power spectrum scale index $n_s = 0.969$.

Appendix B.1: Validation data

We use the same mock shear catalogues from the Scinet Lightcone Simulations (SLICS, Harnois-Déraps & van Waerbeke 2015) as used in Heydenreich et al. (2023) and Linke et al. (2023). The simulations use a flat $\Lambda$CDM cosmology with our fiducial cosmological parameters and contain 1536$^3$ particles inside a 505 h$^{-1}$ Mpc box. We use shear catalogues from 924 pseudo-independent lines of sight, each with a square area of 100 deg$^2$. The source galaxies are distributed with a redshift distribution of

$$n(z) \propto z^2 e^{-z/z_0} \beta,$$

with $z_0 = 0.637$, $\beta = 1.5$, and normalisation such that the overall galaxy density is 30 arcmin$^{-2}$, as expected for a Stage IV lensing survey. The shear includes shape noise, which is infused by adding random ellipticities from a Gaussian distribution. The two-component ellipticity dispersion $\sigma_\epsilon^2$ of the shape noise is $(0.37)^2$. 

Linke, L., et al.: A&A proofs, manuscript no. aa46225-23
Appendix B.2: Covariance measurement

We estimate the covariance in the simulation in two different ways. In the first approach, we replace the expectation value of $\langle M_{ap}^2 \rangle$ by its spatial average for each realisation of the simulation and use the sample covariance between the realisation as covariance estimate. To do so, we use that the aperture mass filter $U_\theta$ is related to a filter $Q_\theta$, for which
\begin{equation}
M_{ap}(\theta; \bar{\theta}) = \int d^2 \theta' \; Q_\theta(\theta - \theta') \gamma_\theta(\theta'; \bar{\theta}) ,
\end{equation}
where $Q_\theta$ is given by
\begin{equation}
Q_\theta(\theta) = \frac{2}{\theta^2} \int_0^\theta d\theta' \; \theta' U_\theta(\theta') - U_\theta(\theta) .
\end{equation}

We calculate Eq. (B.2) using Fast Fourier Transform (FFT) for the aperture scale radii $\theta \in \{4', 8', 16'\}$, which gives us an aperture mass map $M_{ap}(i; \theta)$ for each realisation $i$ and scale radius $\theta$. To remove border effects, we cut off a border of four times the largest aperture radius (i.e. $4 \times 16' = 64'$) from each side of the aperture mass maps. Then we square the aperture mass maps for each scale radius and subsequently take the average over all pixel values to get an estimate for $\langle M_{ap}^2 \rangle$ for each realisation. We take the sample covariance of these realisations as our covariance estimate.

For the sample covariance, we estimate uncertainties using bootstrapping. For this, we create 10,000 lists of 924 randomly drawn integers between 1 and 924, where each list can contain any integer multiple times. We calculate a sample covariance for each list using the $\langle M_{ap}^2 \rangle$ for which $i$ is a list element, leading to 10,000 covariance estimates. Our uncertainty estimate is the standard deviation of these 10,000 estimates.

The second approach to estimate $C_{\langle M_{ap}^2 \rangle}$ in the simulations uses Eq. (29). For this, we use the squared aperture mass maps created before. We then calculate the correlation function $\xi_{M_{ap}}^{(i)}$ for each realisation $i$ using
\begin{equation}
\xi_{M_{ap}}^{(i)}(\theta_1, \theta_2, \eta) = \frac{1}{A \epsilon_A(\eta)} \int_{A} d^2 \alpha \; M_{ap}^2(\theta_1, \alpha) M_{ap}^2(\theta_2, \alpha + \eta) W(\alpha + \eta)
\end{equation}
We evaluate this equation using FFT, as implemented in the scipy routine correlate. The mean of all $\xi_{M_{ap}}^{(i)}$ is then plugged into Eq. (29) to yield the covariance estimate $C_{\langle M_{ap}^2 \rangle}^{corr}$.

We measure $\xi_{\langle M_{ap}^2 \rangle}$ as a function of the two-dimensional vector $\eta$, not just its magnitude $\eta$. Therefore, the measurement method allows us to test the statistical anisotropy of the convergence field by comparing, for example, $\xi_{\langle M_{ap}^2 \rangle}(\eta)$ with $\xi_{\langle M_{ap}^2 \rangle}(-\eta)$.

Appendix B.3: Polyspectra modelling

We model the covariance once with the large-field approximation $C_{\xi}^{corr}$ from Eq. (43) and once including the SSC term in Eq. (44). For this, we use the revised halofit prescription for the power spectrum (Takahashi et al. 2012). The trispectrum is modelled with the one-halo term of the halo model (Cooray & Sheth 2002), using a Sheth–Tormen halo mass function and halo bias (Sheth & Tormen 1999), and Navarro–Frenk–White halo profiles (Navarro et al. 1996).