

# ***Euclid*: Covariance of weak lensing pseudo- $C_\ell$ estimates**

## **Calculation, comparison to simulations, and dependence on survey geometry\***

R. E. Upham<sup>1</sup>, M. L. Brown<sup>1</sup>, L. Whittaker<sup>2,1</sup>, A. Amara<sup>3</sup>, N. Auricchio<sup>4</sup>, D. Bonino<sup>5</sup>, E. Branchini<sup>6,7</sup>, M. Brescia<sup>8</sup>, J. Brinchmann<sup>9</sup>, V. Capobianco<sup>5</sup>, C. Carbone<sup>10</sup>, J. Carretero<sup>11,12</sup>, M. Castellano<sup>13</sup>, S. Cavuoti<sup>14,8,15</sup>, A. Cimatti<sup>16,17</sup>, R. Cledassou<sup>18,19</sup>, G. Congedo<sup>20</sup>, L. Conversi<sup>21,22</sup>, Y. Copin<sup>23</sup>, L. Corcione<sup>5</sup>, M. Cropper<sup>24</sup>, A. Da Silva<sup>25,26</sup>, H. Degaudenzi<sup>27</sup>, M. Douspis<sup>28</sup>, F. Dubath<sup>27</sup>, C. A. J. Duncan<sup>29</sup>, X. Dupac<sup>21</sup>, S. Dusini<sup>30</sup>, A. Ealet<sup>23</sup>, S. Farrens<sup>31</sup>, S. Ferriol<sup>23</sup>, P. Fosalba<sup>32,33</sup>, M. Frailis<sup>34</sup>, E. Franceschi<sup>4</sup>, M. Fumana<sup>10</sup>, B. Garilli<sup>10</sup>, B. Gillis<sup>20</sup>, C. Giocoli<sup>35,36</sup>, F. Grupp<sup>37,38</sup>, S. V. H. Haugan<sup>39</sup>, H. Hoekstra<sup>40</sup>, W. Holmes<sup>41</sup>, F. Hormuth<sup>42</sup>, A. Hornstrup<sup>43</sup>, K. Jahnke<sup>44</sup>, S. Kermiche<sup>45</sup>, A. Kiessling<sup>41</sup>, M. Kilbinger<sup>31</sup>, T. Kitching<sup>24</sup>, M. Kümmel<sup>38</sup>, M. Kunz<sup>46</sup>, H. Kurki-Suonio<sup>47</sup>, S. Ligorì<sup>5</sup>, P. B. Lilje<sup>39</sup>, I. Lloro<sup>48</sup>, O. Marggraf<sup>49</sup>, K. Markovic<sup>41</sup>, F. Marulli<sup>16,4,50</sup>, M. Meneghetti<sup>4,50</sup>, G. Meylan<sup>51</sup>, M. Moresco<sup>16,4</sup>, L. Moscardini<sup>16,4,50</sup>, E. Munari<sup>34</sup>, S. M. Niemi<sup>52</sup>, C. Padilla<sup>12</sup>, S. Paltani<sup>27</sup>, F. Pasian<sup>34</sup>, K. Pedersen<sup>53</sup>, V. Pettorino<sup>31</sup>, S. Pires<sup>31</sup>, M. Poncet<sup>18</sup>, L. Popa<sup>54</sup>, F. Raison<sup>37</sup>, J. Rhodes<sup>41</sup>, E. Rossetti<sup>16</sup>, R. Saglia<sup>37,38</sup>, B. Sartoris<sup>55,34</sup>, P. Schneider<sup>49</sup>, A. Secroun<sup>45</sup>, G. Seidel<sup>44</sup>, C. Sirignano<sup>56,30</sup>, G. Sirri<sup>50</sup>, L. Stanco<sup>30</sup>, J.-L. Starck<sup>31</sup>, P. Tallada-Crespi<sup>57,11</sup>, D. Tavagnacco<sup>34</sup>, A. N. Taylor<sup>20</sup>, I. Tereno<sup>25,58</sup>, R. Toledo-Moreo<sup>59</sup>, F. Torradeflot<sup>57,11</sup>, L. Valenziano<sup>4,50</sup>, Y. Wang<sup>60</sup>, G. Zamorani<sup>4</sup>, J. Zoubian<sup>45</sup>, S. Andreon<sup>61</sup>, M. Baldi<sup>62,4,50</sup>, S. Camera<sup>63,64,5</sup>, V. F. Cardone<sup>13,65</sup>, G. Fabbian<sup>66,67</sup>, G. Polenta<sup>68</sup>, A. Renzi<sup>56,30</sup>, B. Joachimi<sup>2</sup>, A. Hall<sup>20</sup>, A. Loureiro<sup>69,2,20</sup>, and E. Sellentin<sup>70,40</sup>

*(Affiliations can be found after the references)*

Received 14 December 2021 / Accepted 5 February 2022

### **ABSTRACT**

An accurate covariance matrix is essential for obtaining reliable cosmological results when using a Gaussian likelihood. In this paper we study the covariance of pseudo- $C_\ell$  estimates of tomographic cosmic shear power spectra. Using two existing publicly available codes in combination, we calculate the full covariance matrix, including mode-coupling contributions arising from both partial sky coverage and non-linear structure growth. For three different sky masks, we compare the theoretical covariance matrix to that estimated from publicly available  $N$ -body weak lensing simulations, finding good agreement. We find that as a more extreme sky cut is applied, a corresponding increase in both Gaussian off-diagonal covariance and non-Gaussian super-sample covariance is observed in both theory and simulations, in accordance with expectations. Studying the different contributions to the covariance in detail, we find that the Gaussian covariance dominates along the main diagonal and the closest off-diagonals, but farther away from the main diagonal the super-sample covariance is dominant. Forming mock constraints in parameters that describe matter clustering and dark energy, we find that neglecting non-Gaussian contributions to the covariance can lead to underestimating the true size of confidence regions by up to 70 per cent. The dominant non-Gaussian covariance component is the super-sample covariance, but neglecting the smaller connected non-Gaussian covariance can still lead to the underestimation of uncertainties by 10–20 per cent. A real cosmological analysis will require marginalisation over many nuisance parameters, which will decrease the relative importance of all cosmological contributions to the covariance, so these values should be taken as upper limits on the importance of each component.

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

### **1. Introduction**

There are currently many unanswered questions in cosmology, including the origin of the accelerating expansion of the universe and apparent tensions within the dominant  $\Lambda$  cold dark matter model (e.g. Di Valentino et al. 2021, and references therein). One of the most promising tools with which to make progress on these questions in the coming years is the analysis of weak gravitational lensing of distant galaxies by large-scale structure, also known as cosmic shear. The upcoming ESA *Euclid* space mission<sup>1</sup> (Laureijs et al. 2011), as well as other surveys such as those with the *Vera C. Rubin* Observatory in Chile (the Legacy Survey of Space

and Time<sup>2</sup>; Ivezić et al. 2019) and the Square Kilometre Array (SKA) radio observatory in Australia and South Africa<sup>3</sup> (Square Kilometre Array Cosmology Science Working Group 2020), will observe over a billion galaxies, which is expected to lead to unprecedented precision on cosmological constraints – a more than an order of magnitude increase over the previous generation of experiments (Harrison et al. 2016). In order to obtain reliable results, this precision is necessarily accompanied by a requirement to understand all elements of an analysis pipeline to an equally unprecedented degree, including the interplay between the likelihood and estimator effects.

In this paper we are concerned specifically with pseudo- $C_\ell$  estimators, which have been used previously for the analysis

\* This paper is published on behalf of the Euclid Consortium.

<sup>1</sup> <https://www.euclid-ec.org>

<sup>2</sup> <https://www.lsst.org>

<sup>3</sup> <https://www.skatelescope.org>

of weak lensing data from the Hyper-Suprime Cam (HSC) Subaru Strategic Program<sup>4</sup> (Miyazaki et al. 2012) in Hikage et al. (2019) and the Dark Energy Survey<sup>5</sup> (DES; Dark Energy Survey Collaboration 2005) in Camacho et al. (2021) and will be used in the analysis of future *Euclid* data (Loureiro et al. 2021). It was recently shown in Upham et al. (2021) that a Gaussian likelihood is sufficient to obtain accurate cosmological results from weak lensing pseudo- $C_\ell$  estimates. An important ingredient for a Gaussian likelihood is the covariance matrix, so in this paper we focus on the calculation of a cosmic shear pseudo- $C_\ell$  covariance matrix.

The problem of calculating covariance matrices for weak lensing has been extensively discussed in the literature, ranging from analytic or semi-analytic approaches (Cooray & Hu 2001; Schneider et al. 2002; Joachimi et al. 2008; Takada & Jain 2009; Pielorz et al. 2010; Hilbert et al. 2011; Barreira et al. 2018a; Hall & Taylor 2019; Gouyou Beauchamps et al. 2022) through to estimation from simulations (Sato et al. 2011; Harnois-Déraps & Van Waerbeke 2015; Sellentin & Heavens 2016a,b; Harnois-Déraps et al. 2018, 2019; Sgier et al. 2019; Schneider et al. 2020). Here we extend this work to focus specifically on the covariance of pseudo- $C_\ell$  estimates, for which coupling between modes occurs due to the effect of incomplete sky coverage. This effect is in addition to the non-Gaussian mode coupling that is inherent in weak lensing data as a result of non-linear structure growth, which is known to be important for parameter inference (Sato & Nishimichi 2013; Barreira et al. 2018b).

In Sect. 2 we discuss the different Gaussian and non-Gaussian components of the cosmic shear pseudo- $C_\ell$  covariance and their implementation in existing publicly available code. We compare this theoretical covariance to that measured from publicly available weak lensing simulations in Sect. 3. In Sect. 4 we examine the relative importance of the different covariance contributions and how this depends on the mask, which describes the details of sky coverage. This part of our analysis shares some similarities with that of Barreira et al. (2018b), who also studied the relative importance of the different contributions to the cosmic shear covariance for a *Euclid*-like survey and concluded that the ‘connected non-Gaussian’ component (see Sect. 2) can be neglected for only a  $\lesssim 5$  per cent underestimation in single-parameter  $1\sigma$  errors. However, in this paper we specifically study pseudo- $C_\ell$  estimates, for which the survey mask mixes power between all multipoles and induces correlations even for Gaussian fields, which for many covariance elements dominate over other sources of correlation (see Sect. 3). This effect was not included in the analysis of Barreira et al. (2018b), who assumed a diagonal Gaussian covariance, and its inclusion may lead to different conclusions about the relative importance of the different contributions to the covariance. We discuss our conclusions in Sect. 5.

## 2. Cosmic shear power spectrum covariance contributions

We begin this section by summarising the different contributions to the cosmic shear power spectrum covariance. We refer the reader to Barreira et al. (2018a) and the other references provided both therein and below for a thorough theoretical background and derivation.

Starting in three-dimensional space, the covariance of the matter power spectrum receives two contributions: one that

depends on the matter power spectrum itself, and one that depends on a particular (‘parallelogram’) configuration of the matter trispectrum that corresponds to the Fourier transform of the connected four-point correlation function. For a Gaussian matter distribution, only the first contribution is non-vanishing, and hence it is commonly referred to as the ‘Gaussian covariance’, which we will use here. (It is also sometimes referred to as the ‘disconnected’ covariance.) We follow Barreira et al. (2018a) in referring to the second contribution as the ‘connected non-Gaussian’ component.

However, for any realistic finite-volume survey such as *Euclid*, the observed matter power spectrum is convolved with a three-dimensional window function. While the Gaussian and non-Gaussian terms remain distinct, this has the effect of introducing additional non-Gaussian couplings between large-scale modes outside the survey and small-scale modes within the survey. This is commonly known as ‘super-sample’ (originally ‘beat-coupling’; Hamilton et al. 2006) covariance, and physically can be explained by the fact that unobservable large-scale modes within which the survey is embedded can influence the rate of small-scale non-linear structure growth, and therefore also the strength of coupling between small-scale modes (Takada & Hu 2013; Barreira et al. 2018a). Perhaps counter-intuitively, it turns out that this is generally the dominant source of non-Gaussian covariance (Hamilton et al. 2006; Barreira et al. 2018b).

Progressing to projected two-point statistics such as cosmic shear angular power spectra, the same three components – Gaussian, super-sample and connected non-Gaussian – contribute to the covariance. Strictly speaking, the separation of the super-sample and connected non-Gaussian components is only exact under the Limber approximation (Barreira et al. 2018a), but the inaccuracy of the Limber approximation is only relevant on very large scales (very low multipoles,  $\ell \lesssim 20$ ) where non-Gaussian correlations are small.

We will now discuss the calculation of each of the three cosmic shear covariance components in turn.

### 2.1. Gaussian covariance

To calculate the Gaussian covariance, we used the ‘improved narrow kernel approximation’ method (Nicola et al. 2021) implemented using the publicly available code NaMaster<sup>6</sup> (Alonso et al. 2019; García-García et al. 2019). We now provide further details and some background on this method.

The Gaussian covariance component of a general statistically isotropic field on the sphere is equivalent to the total covariance of a Gaussian field with the same power spectrum. The analytic covariance of pseudo- $C_\ell$  estimates on Gaussian fields has been well studied in the cosmic microwave background literature (Efstathiou 2004; Challinor & Chon 2005; Brown et al. 2005; Upham et al. 2020, for a non-Gaussian likelihood method) as well as in the context of weak lensing (García-García et al. 2019; Nicola et al. 2021). The exact Gaussian pseudo- $C_\ell$  covariance can be written down analytically, and includes terms of the following form (e.g. Brown et al. 2005):

$$\text{Cov}(\bar{C}_\ell, \bar{C}_{\ell'}) = \sum_{\substack{m, m' \\ \ell_1, \ell_2 \\ m_1, m_2}} W_{\ell\ell_1}^{mm_1} (W_{\ell'\ell_1}^{m'm_1})^* W_{\ell\ell_2}^{m'm_2} (W_{\ell\ell_2}^{mm_2})^* C_{\ell_1} C_{\ell_2} \quad (1)$$

+ similar terms,

<sup>4</sup> <https://hsc.mtk.nao.ac.jp>

<sup>5</sup> <https://www.darkenergysurvey.org>

<sup>6</sup> <https://github.com/LSSTDESC/NaMaster>

where the harmonic space window functions  $W$  are given in Eq. (8) of Brown et al. (2005), and the ‘similar terms’ involve different combinations of power spectra depending on the situation and spins being considered (Hansen & Górski 2003; Challinor & Chon 2005).

The evaluation of Eq. (1) requires  $\mathcal{O}(\ell_{\max}^6)$  operations per term, so it is impractical to evaluate exactly and in practice approximations are used. These commonly involve substitutions of the following kind (Efstathiou 2004; Brown et al. 2005; García-García et al. 2019):

$$C_{\ell_1} C_{\ell_2} \rightarrow C_\ell C_{\ell'}, \quad (2)$$

which allows the power spectrum dependence to be brought out of the sums in Eq. (1). This means that the coefficients in the similar terms are now all the same (except for any possible spin dependence, or if different fields use different masks). Symmetry properties of the harmonic space window function allow the calculation of these coefficients to be further simplified, to the point where the covariance can be evaluated in a reasonable time. In essence, the approximation in Eq. (2) assumes that the power spectrum is constant over the region around a given  $\ell$  in which the window function is non-negligible. This will be accurate as long as the window function is sufficiently sharply peaked, and therefore this approximation is often known as the ‘narrow kernel approximation’ (NKA). A generalised version of the NKA is described in García-García et al. (2019) and implemented in NaMaster, which supports an arbitrary number of correlated spin-0 and spin-2 fields and has both curved-sky and flat-sky support. We used the curved-sky spin-2 version, which naturally accounts for  $E$ - $B$  leakage (we assume noise-only  $B$ -modes). By default, NaMaster provides the covariance of deconvolved pseudo- $C_\ell$  estimates, but we set the `coupled=True` option to instead obtain the covariance of ‘raw’, un-deconvolved estimates such as those produced by the pseudo- $C_\ell$  estimator developed for *Euclid*, which is described in Loureiro et al. (2021). We did not apply any  $E$ - $B$  purification or noise de-biasing either.

Nicola et al. (2021) introduced a small modification to the NKA that turns out to significantly increase its accuracy, which they refer to as the improved NKA. It involves simply replacing each  $C_\ell$  in the standard NKA by its mode-coupled counterpart,

$$C_\ell \rightarrow \frac{\langle \tilde{C}_\ell \rangle}{f_{\text{sky}}} = \frac{\sum_{\ell'} M_{\ell\ell'} C_{\ell'}}{f_{\text{sky}}}, \quad (3)$$

where  $M$  is the usual pseudo- $C_\ell$  mixing or mode-coupling matrix (see Brown et al. 2005 for a full derivation of its calculation), and the division by the sky fraction  $f_{\text{sky}}$  is to avoid double-counting the loss of power on the cut sky. NaMaster includes the functionality to calculate the mixing matrix and to apply it to power spectra to calculate  $\langle \tilde{C}_\ell \rangle$ , so the extension to the improved NKA is trivial. The recent pseudo- $C_\ell$  analysis of DES Year 1 observations in Camacho et al. (2021) provided the first application of the improved NKA to real data, but takes the approach of deconvolution and noise subtraction to obtain unbiased estimates of the underlying power spectrum, unlike the forward-modelling approach taken in this paper.

## 2.2. Super-sample covariance

To calculate the super-sample covariance contribution we used the publicly available code CosmoLike<sup>7</sup> (Krause & Eifler 2017).

<sup>7</sup> <https://github.com/CosmoLike>

Specifically, we used an adapted version of the CosmoCov<sup>8</sup> correlation function covariance package (Fang et al. 2020), which obtains the real-space non-Gaussian covariance as a transform of the harmonic space covariance. CosmoCov has been used for the DES Year 1 and Year 3 cosmological analyses (Krause et al. 2017, 2021; Friedrich et al. 2021), as well as for the non-Gaussian covariance in the Year 1 pseudo- $C_\ell$  analysis in Camacho et al. (2021). Our adapted code to expose the harmonic space covariance directly is available online<sup>9</sup>.

CosmoLike calculates the super-sample covariance using the approach introduced in Takada & Hu (2013), by considering the response of the small-scale non-linear matter power spectrum to changes in the mean linear density field (see also Chiang et al. 2014; Li et al. 2014). The response of the non-linear matter power spectrum is evaluated using a halo model, with the details given in Eqs. (A7)–(A11) of Krause & Eifler (2017). We note that since that paper was published, the calculation of the survey variance  $\sigma_b(z)$  in their Eq. (A8) has been replaced within CosmoLike with the following calculation:

$$\sigma_b(z) = \frac{1}{4\pi r^2} \frac{1}{C_{\ell=0}^{\text{mask}}} \sum_{\ell=0}^{1000} (2\ell+1) C_\ell^{\text{mask}} P_{\text{lin}}\left(\frac{\ell+1/2}{r}, z\right), \quad (4)$$

where  $r = f_\kappa(\chi(z))$ , and  $C_\ell^{\text{mask}}$  is the power spectrum of the mask. This treatment of the cut sky using the mask power spectrum was derived in Barreira et al. (2018a). The value of  $\ell_{\max} = 1000$  in Eq. (4) is arbitrary, but has a negligible impact in practice because the power spectrum of both masks is negligible above  $\ell \sim 100$ . We note that CosmoLike also provides an alternative implementation of both the super-sample and connected non-Gaussian covariance using the ‘response approach’, which additionally accounts for a tidal contribution to the super-sample covariance and has been found to be more accurate than the standard implementation (Wagner et al. 2015; Barreira & Schmidt 2017a,b; Barreira et al. 2018a; Schmidt et al. 2018). Here we use the standard CosmoLike implementation as used by the DES Collaboration (Krause et al. 2017, 2021; Friedrich et al. 2021; Camacho et al. 2021).

As stated previously, we wish to forward-model the effect of the mask to obtain the covariance of raw un-deconvolved pseudo- $C_\ell$  estimates following Loureiro et al. (2021). The non-Gaussian covariance output from CosmoLike corresponds to the covariance of unbiased estimates of the underlying power spectrum  $\tilde{C}_\ell$ , and does not account for any estimator effects such as cut-sky mode coupling. Within the context of the pseudo- $C_\ell$  method, the closest way to interpret the covariance output from CosmoLike is as the covariance of deconvolved pseudo- $C_\ell$  estimates. This is how it is interpreted in Camacho et al. (2021) and how we chose to interpret it here. In this case, the unbiased estimates  $\tilde{C}_\ell$  may in principle be obtained from raw pseudo- $C_\ell$  estimates  $\tilde{C}_\ell^{\text{raw}}$  as

$$\tilde{C}_\ell = \sum_{\ell'} M_{\ell\ell'}^{-1} \tilde{C}_{\ell'}^{\text{raw}}. \quad (5)$$

It follows that the covariance matrices of  $\tilde{C}_\ell$  and  $\tilde{C}_\ell^{\text{raw}}$ , which we denote respectively as  $\tilde{\Sigma}$  and  $\tilde{\Sigma}^{\text{raw}}$ , are related as

$$\tilde{\Sigma} = (M^{-1}) \tilde{\Sigma}^{\text{raw}} (M^{-1})^\top. \quad (6)$$

We interpret the covariance as output from CosmoLike as  $\tilde{\Sigma}$ . This choice is necessarily an approximation since CosmoLike

<sup>8</sup> <https://github.com/CosmoLike/CosmoCov>

<sup>9</sup> [https://github.com/robinupham/CosmoCov\\_CIConv](https://github.com/robinupham/CosmoCov_CIConv)



does not account for any estimator effects, including pseudo- $C_\ell$  mode coupling, but it is a necessary choice and is equivalent to that made in the DES Y1 pseudo- $C_\ell$  analysis in [Camacho et al. \(2021\)](#). In both that paper and this one (Sect. 3), the resulting covariance is compared to simulations, with good agreement. A full general non-Gaussian pseudo- $C_\ell$  covariance is presented in [Shirasaki et al. \(2015\)](#), but in practice an approximation such as the one we make here is necessary. We therefore inverted Eq. (6) to obtain the relation that we used to transform  $\widehat{\Sigma}$  to the raw pseudo- $C_\ell$  covariance,  $\widetilde{\Sigma}$ :

$$\widetilde{\Sigma} = \mathbf{M}\widehat{\Sigma}\mathbf{M}^\top. \quad (7)$$

We obtained the spin-2 mixing matrix  $\mathbf{M}$  for this calculation using NaMaster.

### 2.3. Connected non-Gaussian covariance

We also calculated the connected non-Gaussian covariance component using CosmoLike (referred to in [Krause & Eifler 2017](#) as the ‘non-Gaussian covariance in the absence of survey window effects’), using the same adapted version of the CosmoCov package that we make available at the URL provided in Sect. 2.2.

CosmoLike calculates the connected non-Gaussian covariance as the projected matter trispectrum. The trispectrum is calculated using a halo model, with the details given in Eqs. (A3)–(A6) of [Krause & Eifler \(2017\)](#) (see also [Cooray & Sheth 2002](#); [Takada & Jain 2009](#)). We find this to be suitably accurate in Sect. 3.2, and we also find in Sect. 4 that the contribution from the connected non-Gaussian component to the total parameter posterior error is no more than 10–20 per cent. As with the super-sample covariance component, we multiplied the connected non-Gaussian covariance matrix by the mixing matrix using Eq. (7).

The calculation of the connected non-Gaussian covariance component is much slower than the other two components. As a result, in Sect. 4.2 we use an approximation to directly estimate the connected non-Gaussian covariance of power spectrum bandpowers (i.e. power spectra that have been binned in multipole space), which we subsequently use to obtain mock parameter constraints. The approximation is described in that section. However, the results provided in Sects. 3 and 4.1 are obtained using the full (i.e. per-multipole) connected non-Gaussian covariance matrix for a single redshift bin. This took 36 days to calculate on 55 CPUs for 32 million elements<sup>10</sup> By contrast, the equivalent Gaussian and super-sample covariance matrices each took around an hour to calculate on 12 CPUs.

## 3. Comparison to simulations

### 3.1. Method

We used the publicly available<sup>11</sup> full-sky simulated spin-2 shear maps of [Takahashi et al. \(2017\)](#), which were produced by ray tracing through cosmological  $N$ -body simulations. The simulations use a maximum box size of  $6300 h^{-1}\text{Mpc}$ . These simulated maps are quite versatile, not only because they cover the full

sky, but also because they describe the underlying shear field with no shape noise or shot noise (which could be added if required, but we chose not to here). A total of 108 realisations were performed, which is a relatively small number considering that very large numbers of realisations may be required for full convergence of a simulated covariance ([Blot et al. 2016](#)), but we find that a useful comparison between theory and simulations is still possible. In addition, finite box simulations will necessarily underestimate the super-sample covariance ([Hamilton et al. 2006](#); [Li et al. 2014](#)), but we do not detect a significant deficit. We refer the reader to [Takahashi et al. \(2017\)](#) for further details and validation of the simulations, and we note that density maps, halo catalogues and lensed cosmic microwave background maps are also made available at the same URL.

The shear maps are provided at 38 redshift slices from  $z = 0.05$  to  $5.3$ . For each realisation, we combined these into five redshift bins, following a Gaussian redshift distribution centred at  $z = 0.65, 0.95, 1.25, 1.55, 1.85$  with a standard deviation of  $0.3$ . The combined shear map for each redshift bin is formed as a weighted average over all 38 slices, with the weights given by the probability density of a Gaussian distribution with the appropriate mean and standard deviation. We comment on this choice of redshift distribution below. We then took three copies: one full-sky with no mask, one with a full *Euclid*-like mask including the survey footprint and a bright star mask ( $f_{\text{sky}} = 0.31$ ), and one with a *Euclid* DR1-like footprint but no bright star mask ( $f_{\text{sky}} = 0.06$ ). The full *Euclid*-like and *Euclid* DR1-like masks are shown in Fig. 1. These masks approximate the coverage of the *Euclid* Wide Survey at different stages but do not exactly correspond to what will be observed, which is described in [Euclid Collaboration \(2022\)](#). We also assume that the masks are uncorrelated with the signal, which may not be the case in practice (e.g. [Fabbian et al. 2021](#)). Finally, we used the healpy<sup>12</sup> ([Zonca et al. 2019](#)) interface to the HEALPIX<sup>13</sup> ([Górski et al. 2005](#)) software to measure the spin-2 shear power spectra for each realisation. The comparisons shown in this section are for the  $E$ -mode auto-power in the lowest redshift bin.

For the theoretical covariance components described in Sect. 2, we used the same cosmology and redshift distribution as the simulations. We used  $\ell_{\text{max}} = 8000$  in intermediate calculations to fully account for all relevant mode coupling, but we limited the comparison to  $\ell \leq 3000$  because the  $n_{\text{side}} = 4096$  maps that we used experience distortion from limited angular resolution above this point, as documented in [Takahashi et al. \(2017\)](#). We note that  $n_{\text{side}} = 8192$  maps are also available, so the  $\ell$  range could in principle be extended, albeit with significantly increased computational requirements.

### Choice of redshift distribution

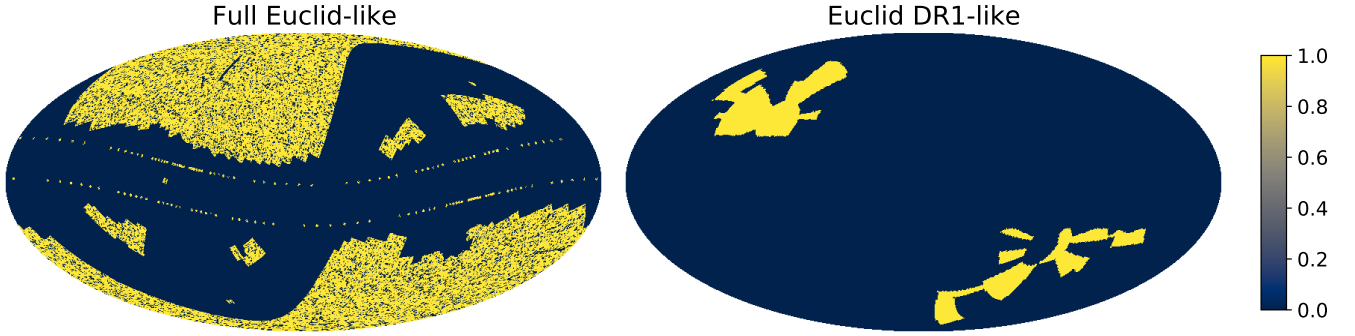
The choice of redshift distribution used here – five Gaussian bins, centred on  $z = 0.65, 0.95, 1.25, 1.55, 1.85$  with a standard deviation of  $0.3$  – was made for simplicity, with the relatively low number of redshift bins (five) also chosen for computational efficiency. Since we have the freedom to enforce agreement between the redshift distributions in the simulations and theory, there is little additional value in choosing a more complicated, more realistic distribution. Future cosmological analyses of real *Euclid* data are likely to use a larger number of bins with less overlap than is used here (e.g. [Euclid Collaboration 2021](#)). There is no reason that this will affect the results presented in

<sup>10</sup> This number comes from a data vector running from  $\ell_{\text{min}} = 2$  to  $\ell_{\text{max}} = 8000$ , which has a length of  $n = 8000 - 2 + 1 = 7999$ , leading to a number of unique covariance elements equal to  $n(n+1)/2 = 31\,996\,000$ .

<sup>11</sup> [http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky\\_raytracing](http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing)

<sup>12</sup> <https://github.com/healpy/healpy>

<sup>13</sup> <http://healpix.sourceforge.net>



**Fig. 1.** Full *Euclid*-like and *Euclid* DR1-like masks, which we use in Sects. 3 and 4 to quantify the effects of different masks on the power spectrum covariance.

this paper, although we note that marginalising over nuisance parameters describing photometric redshift uncertainties will reduce the importance of all cosmological contributions to the covariance.

### 3.2. Results

In Fig. 2 we show correlation matrices for the simulated covariance compared to the total theoretical covariance for each mask, as well as the individual components of the theory covariance. There appears to be good agreement between the simulated and total theoretical correlation matrices for all three masks. We discuss the relative contributions from the three covariance components in Sect. 4.

In Fig. 3 we show a detailed comparison of certain diagonals of the covariance matrix. For the main diagonal ( $\Delta\ell = 0$ ) we show the variance divided by the square of the power spectrum, to remove any effects coming from disagreement in the power spectrum between the simulations and theory, which is not our focus here. For the off-diagonals ( $\Delta\ell = 2, 10, 100$ ) we show the correlation. In all panels the simulated line is a rolling average over 50  $\ell$ s and the shaded region is the standard deviation over this range. For the full-sky and full *Euclid*-like masks, we find excellent agreement between the theory and simulations. For the more extreme *Euclid* DR1-like mask there is a slightly worse, but still generally good, level of agreement. In particular, the super-sample covariance component is clearly correctly increasing with the severity of the sky cut to match the additional correlation found in the simulations. We discuss the relative sizes and importance of the three theory contributions further in Sect. 4. We conclude from Figs. 2 and 3 that CosmoLike’s non-Gaussian covariance calculations appear to be suitably accurate, to the degree that can be assessed using these simulations.

We also repeat this comparison for purely Gaussian fields, which we simulated using healpy. The results are shown in Fig. 4, where we find an excellent level of agreement between the Gaussian field simulations and the prediction from the improved NKA, which is a significant improvement over the standard NKA (not shown).

## 4. Importance of covariance components and dependence on mask

In this section we discuss the size and importance of the different components of the cosmic shear pseudo- $C_\ell$  covariance, and how these properties depend on the mask.

### 4.1. Relative sizes of components

#### 4.1.1. Without shape noise

We first discuss the case without shape noise, which has already been shown in Figs. 2 and 3. We find from the full correlation matrices plotted in Fig. 2 that the main diagonal of the matrix is dominated by the Gaussian component, which is purely diagonal in the full-sky case and visibly broadens slightly as the sky cut is increased. The super-sample covariance is the dominant off-diagonal component, particularly at higher  $\ell$  but extending down visibly even to  $\ell < 500$  in the case of the most extreme *Euclid* DR1-like mask. The connected non-Gaussian contribution is barely visible on the colour scale, other than for the *Euclid* DR1-like mask at low  $\ell$ .

A more detailed comparison of the relative sizes of the different components is possible with the selected diagonals shown in Fig. 3. We find again that the main diagonal ( $\Delta\ell = 0$ ) is dominated by the Gaussian component, but the extent to which this is the case is reduced as the sky cut is increased, as the contribution increases from both non-Gaussian components. Moving away from the main diagonal, at  $\Delta\ell = 10$  the Gaussian is still the largest component (except on the full sky, where its contribution is purely diagonal), but by  $\Delta\ell = 100$  the super-sample component is dominant and increasing towards higher  $\ell$ . It is clear that the super-sample covariance contribution increases with a more severe sky cut, though notably it is still visibly non-zero even for full-sky observations. The connected non-Gaussian component is the subdominant non-Gaussian contribution at all values of  $\ell$  and  $\Delta\ell$  for all masks.

#### 4.1.2. With shape noise

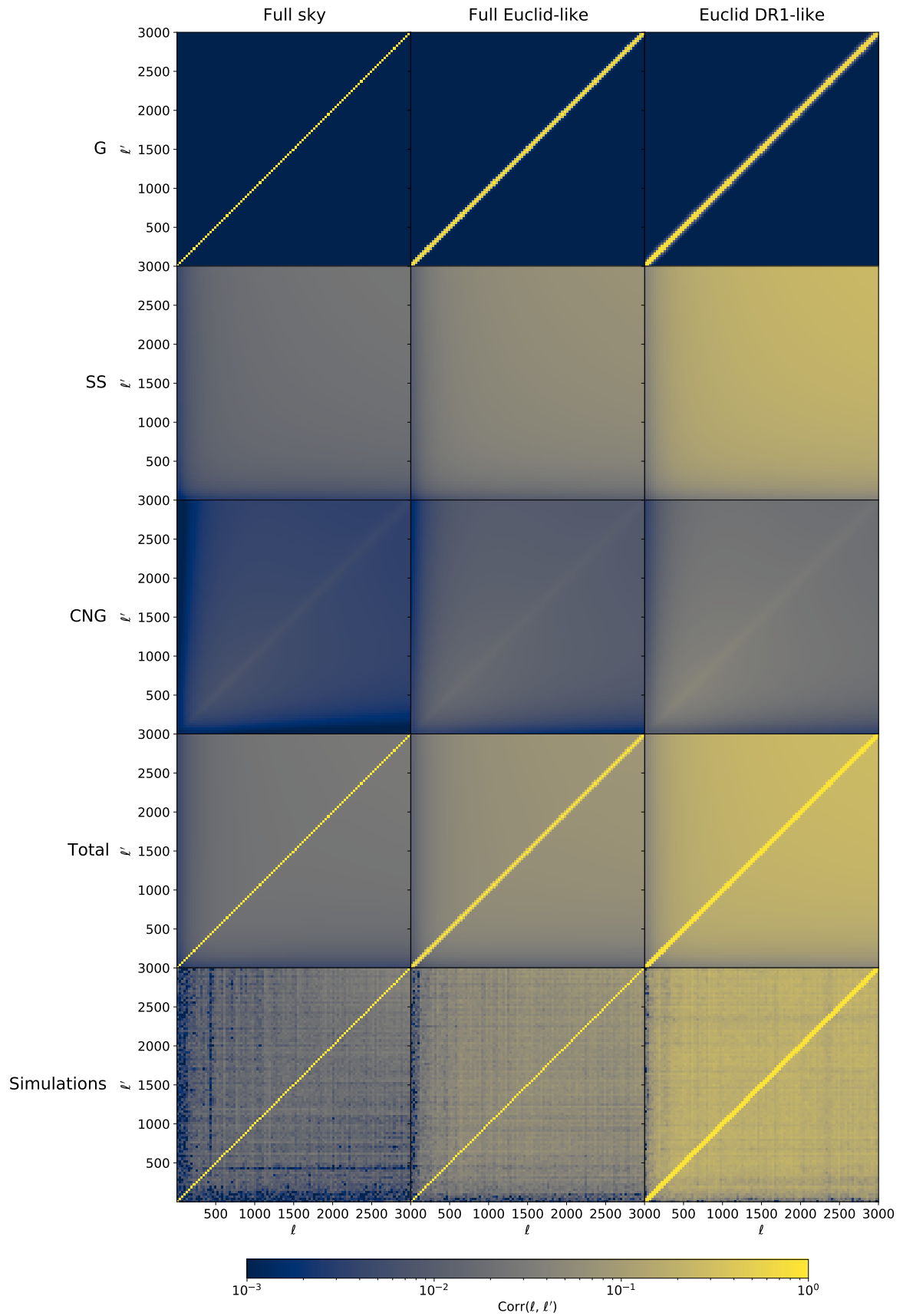
In Fig. 5 we show an equivalent comparison of the sizes of the theoretical covariance components with shape noise included. We assumed Gaussian shape noise and included it as a contribution to the power spectrum,

$$C_\ell \rightarrow C_\ell + N_\ell; \quad (8)$$

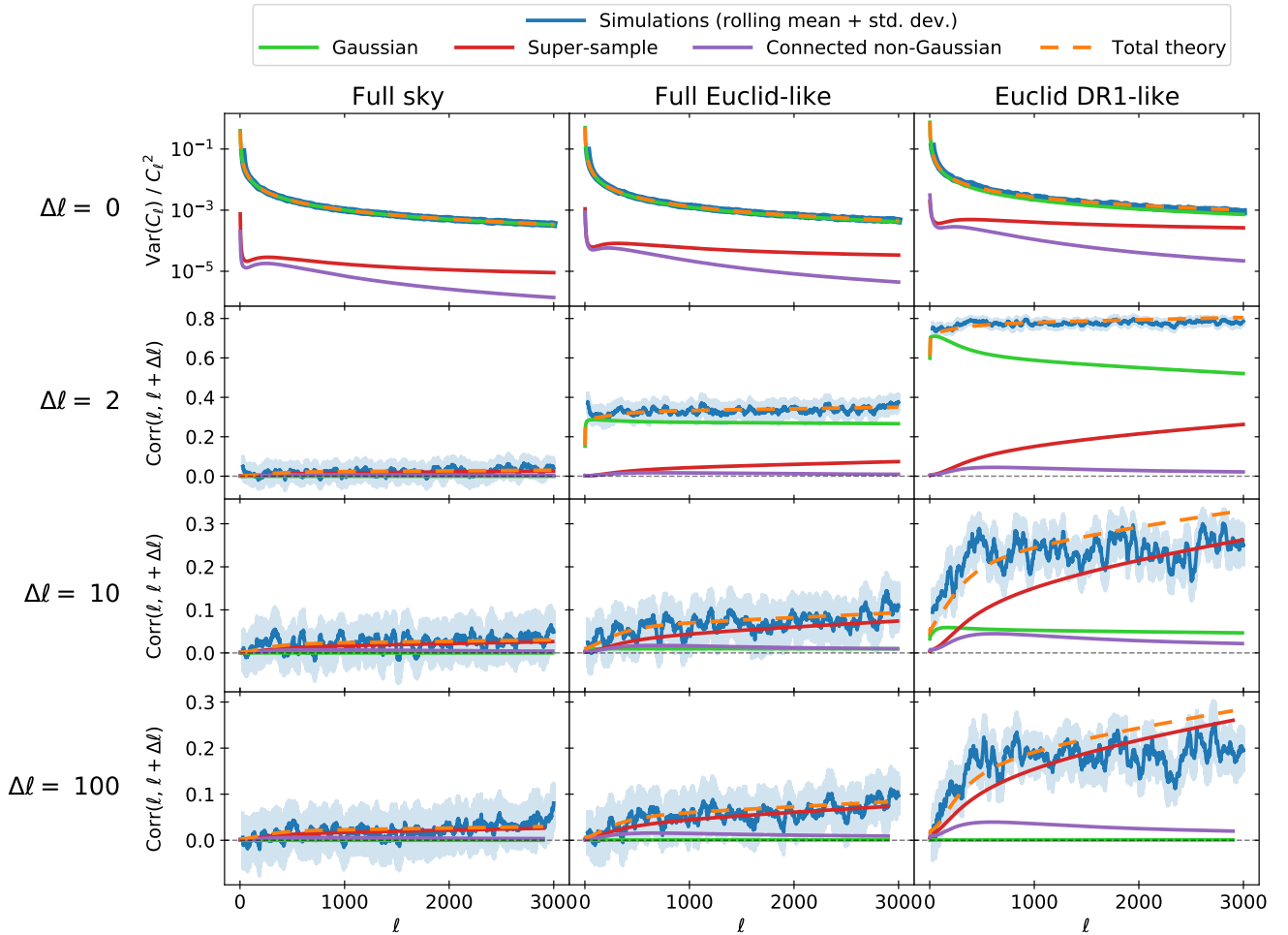
$$N_\ell = \frac{\sigma_\varepsilon^2}{N_i}, \quad (9)$$

where  $\sigma_\varepsilon$  is the intrinsic galaxy shape dispersion per component and  $N_i$  is the number of galaxies per steradian per redshift bin. We assumed a *Euclid*-like galaxy number density of  $30 \text{ arcmin}^{-2}$  equally split over five redshift bins, and a value of  $\sigma_\varepsilon = 0.3$ .

With shape noise included, we find quite different behaviour to the no-noise case. The Gaussian-dominated main diagonal



**Fig. 2.** Correlation matrices for the simulated covariance compared to the total theoretical covariance for each mask and for the individual components of the theory covariance: Gaussian (G), super-sample (SS), and connected non-Gaussian (CNG). The covariance shown here is for the shear  $E$ -mode power spectrum in the lowest redshift bin, without shape noise.



**Fig. 3.** Comparison between the covariance predicted by theory and measured from simulations, for the three masks. *Top row*: variance divided by the power spectrum squared, and *lower three rows*: correlation. In all panels the simulated line is a rolling average over 50  $\ell$ s, and the shaded region is the standard deviation over this range. The covariance shown here is for the shear  $E$ -mode power spectrum in the lowest redshift bin, without shape noise.

is substantially increased, especially at higher  $\ell$ , resulting in non-Gaussian off-diagonal correlations being significantly suppressed. The result is that the Gaussian component is dominant at all  $\ell$  as far away from the main diagonal as  $\Delta\ell = 10$ . By  $\Delta\ell = 100$ , the Gaussian component is no longer dominant at lower  $\ell$ , but continues to account for the largest contribution at higher  $\ell$ : above  $\ell \sim 1000$  for the full *Euclid*-like mask and  $\ell \sim 1500$  for the *Euclid* DR1-like mask. This suggests that once shape noise is included, the Gaussian component is more important than the no-noise results in Fig. 3 suggest.

#### 4.2. Importance for parameter constraints

While the relative size of the different covariance components studied in Sect. 4.1 offers interesting insight into their behaviour, it says relatively little about the actual importance of each component. In particular, it is unclear to what extent the dominance of the Gaussian component on and close to the main diagonal is offset by its sub-dominance farther away from the main diagonal. Here, to gain more insight into this, we produce mock parameter constraints. Shape noise is included in this section, following Eq. (9).

Here we used five redshift bins, including all auto- and cross-spectra ( $E$ -modes only), giving 15 power spectra in total. We

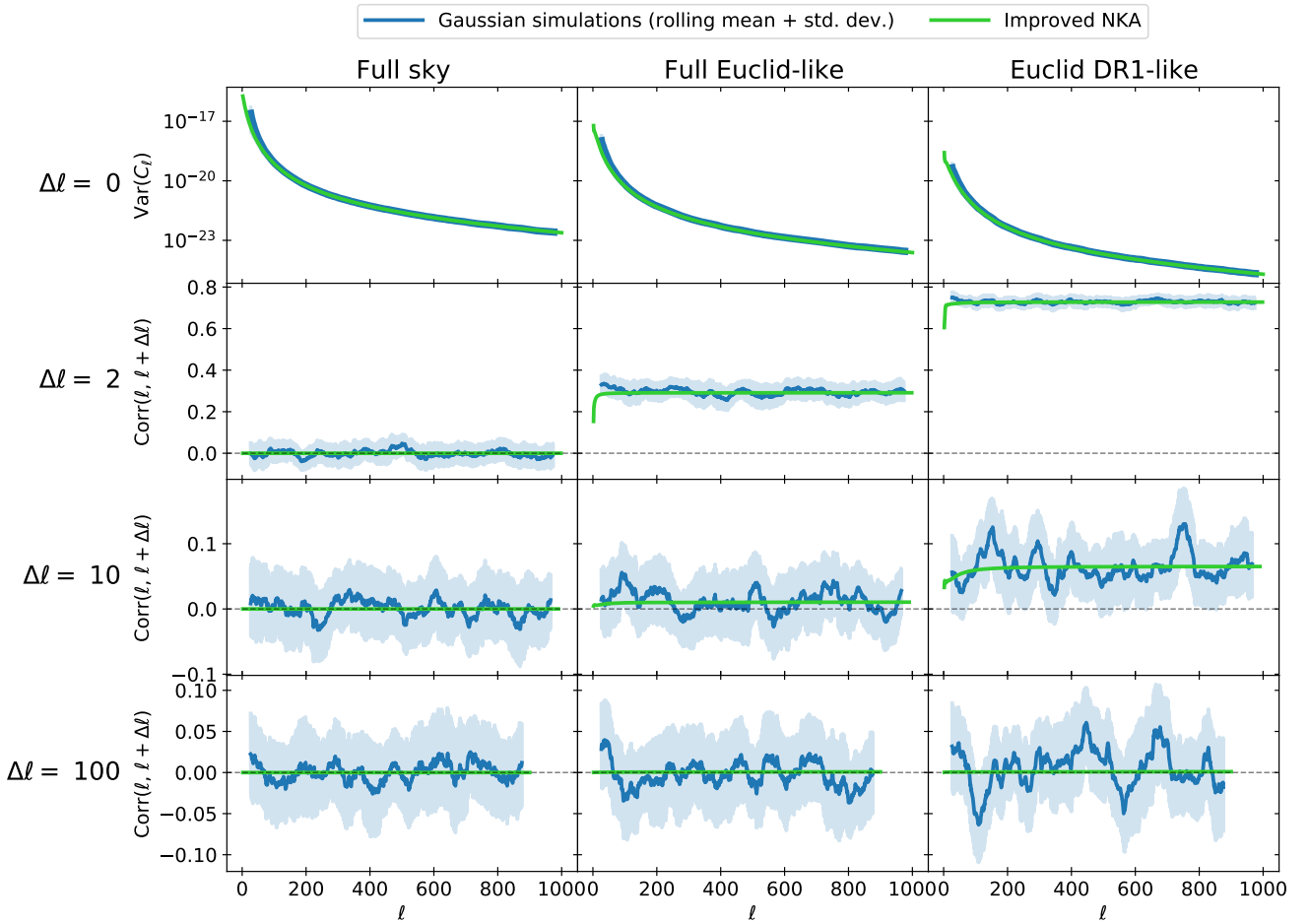
used scales up to  $\ell_{\max} = 5000$ . The full data vector for this setup would have  $n = 75\,000$  elements ( $15 \times 5000$ ), which gives  $n(n+1)/2 = 2.8$  billion unique covariance elements. Due to the time needed to evaluate the projected matter trispectrum, it would be unfeasible to calculate the connected non-Gaussian contribution in full. As a result, we elected to bin into 12 logarithmically spaced bandpowers and use an approximation to obtain the connected non-Gaussian bandpower component from a more modest number of per- $\ell$  covariance calculations. We describe this approximation in Sect. 4.2.1. The Gaussian and super-sample covariance components were calculated in full, using scales up to  $\ell_{\max} = 8000$  for intermediate calculations, before being binned into bandpowers as

$$P_b = \sum_{\ell} P_{b\ell} C_{\ell}, \quad (10)$$

where  $P$  is the bandpower binning matrix whose elements are given by

$$P_{b\ell} = \begin{cases} \frac{\ell(\ell+1)}{2\pi} [\ell_{\min}^{b+1} - \ell_{\min}^b]^{-1} & \text{for } \ell_{\min}^b \leq \ell < \ell_{\min}^{b+1}, \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

where  $\ell_{\min}^b$  is the lower edge of bin  $b$ .



**Fig. 4.** Comparison between covariance measured from Gaussian field simulations and predicted using the improved NKA method, for the three masks. In all panels the simulated line is a rolling average over 50  $\ell$ s, and the shaded region is the standard deviation over this range. No shape noise is included.

We obtained a mock observation by sampling from a Gaussian likelihood with the total covariance. The input mean was the fiducial theory power spectra, plus noise for auto-spectra given by Eq. (9), mixed using the mixing matrix obtained using NaMaster, then binned following Eq. (10). This random sampling process replicates the randomness of cosmic variance that is present in a real observation, and means that – as with real data – the resulting posterior distributions are not centred on the ‘true’ input parameters. We verified that bandpowers measured from the Takahashi et al. (2017) simulations used in Sect. 3 are no more non-Gaussian than those measured from Gaussian field simulations, and therefore since a Gaussian likelihood was shown to be sufficiently accurate for Gaussian fields in Upham et al. (2021), it is a suitable choice here. To obtain parameter constraints, we iterated over two-parameter grids produced using CosmoSIS<sup>14</sup> (Zuntz et al. 2015), following the pipeline described in Sect. 2.2.1 of Upham et al. (2021)<sup>15</sup>. All other parameters were held fixed. At each point in parameter space, we compared theory bandpowers – calculated in the same way as the input mean to the observation described above – to the observed bandpowers using a Gaussian likelihood with different combinations of covariance components. All combinations necessarily include

<sup>14</sup> <https://bitbucket.org/joezuntz/cosmosis>

<sup>15</sup> The pipeline includes the CAMB (Lewis et al. 2000; Howlett et al. 2012) and Halofit-Takahashi (Smith et al. 2003; Takahashi et al. 2012) modules.

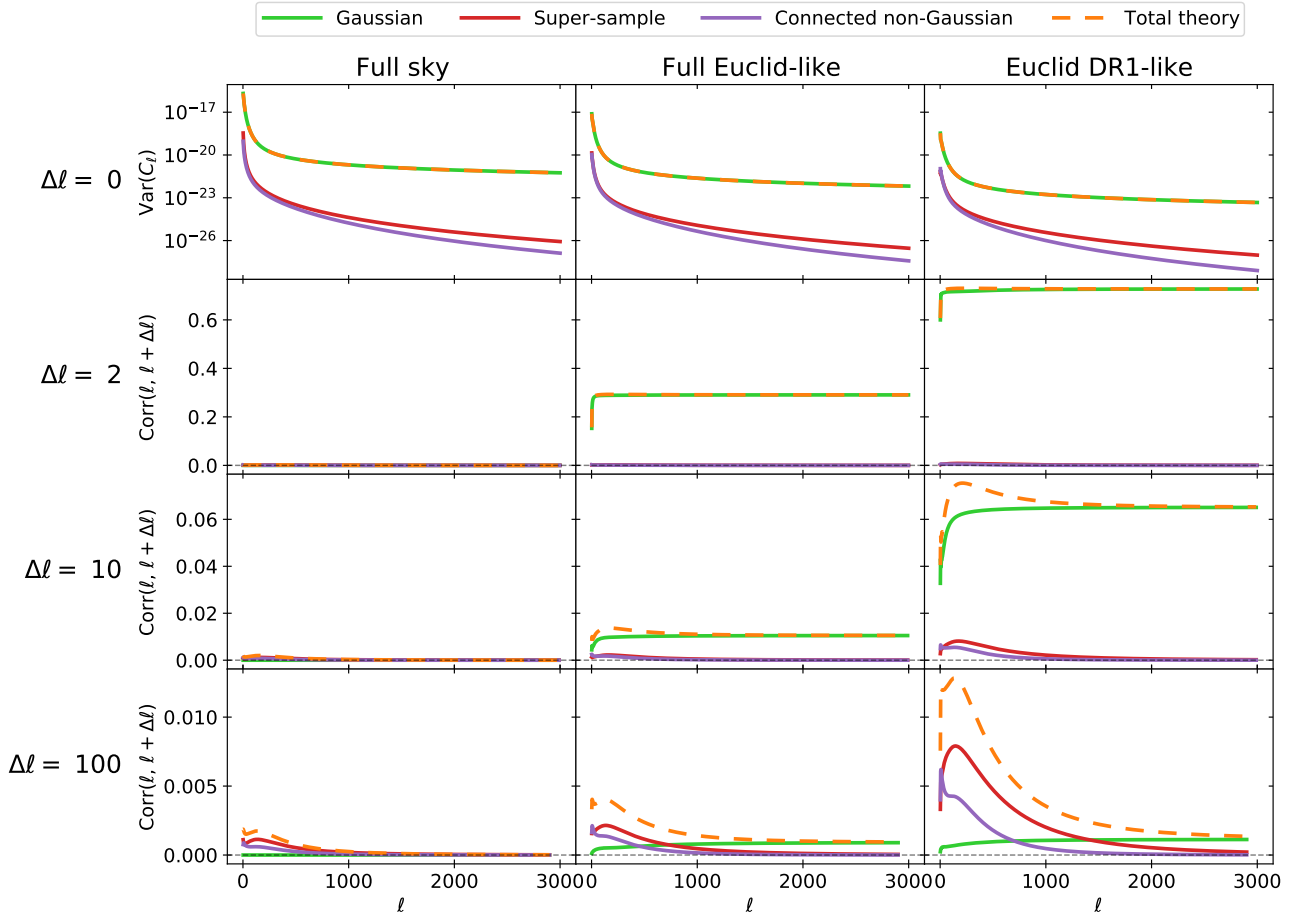
the Gaussian component, since this on its own is a valid positive definite covariance matrix, unlike the super-sample and connected non-Gaussian components.

#### 4.2.1. Connected non-Gaussian approximation

As noted above, it is impractical to calculate the connected non-Gaussian component for all 2.8 billion unique elements of the full covariance matrix. Instead we used an approximation to directly obtain its contribution to the bandpower covariance. This approximation was designed to mimic two effects: the mixing of power by the survey mask, and the binning of individual multipoles into bandpowers. In our analysis, both of these processes are cosmology-independent: each shear field uses the same mask, and all power spectra use the same binning scheme. As a result, both processes should have approximately the same effect on every power spectrum, and consequently also every covariance block. Therefore, the approximation that we make used two sets of weights – one to mimic binning and the other to mimic mixing – which were calibrated for the covariance of the shear auto-power spectrum in the lowest redshift bin and then applied to all further blocks.

First, we evaluated in full the connected non-Gaussian component for a single covariance element per bandpower pair, for all combinations of power spectra. This was chosen to be for the weighted average  $\ell$  in each bandpower, with the weights given by





**Fig. 5.** Comparison between contributions to the theoretical covariance for the three masks, with shape noise included following Eq. (9). *Top row*: variance, and *lower three rows*: correlation.

$P_{b\ell}$  (Eq. (11)), rounded to the nearest integer. This vastly reduced the number of projected trispectrum calculations, to 16 290<sup>16</sup>. We then re-weighted the result using the weights calibrated using the covariance of the shear auto-power spectrum in the lowest redshift bin, which was calculated in full for the previous sections.

The weighting can be understood as a two-step process. First, we applied a ‘binning’ weighting, designed to mimic the effect of taking the full unbinned covariance matrix and binning it into bandpowers. Then we applied a ‘mixing’ weighting, designed to mimic the effect of the mixing matrix. In both cases, the weights were obtained by carrying out the process in full for the shear auto-power spectrum in the lowest redshift bin.

This can be illustrated using equations as follows. For the first block (covariance of shear auto-power in the lowest redshift bin), the following procedure was used to transform the full unbinned covariance  $\text{Cov}_{\text{unbinned}}$  into a final binned and mixed block  $\text{Cov}_{\text{mixed}}$ , via a binned and unmixed stage  $\text{Cov}_{\text{binned}}$ :

$$\text{Cov}_{\text{binned}} = \mathbf{P} \text{Cov}_{\text{unbinned}} \mathbf{P}^T; \quad (12)$$

$$\text{Cov}_{\text{mixed}} = \mathbf{M} \text{Cov}_{\text{binned}} \mathbf{M}^T, \quad (13)$$

where  $\mathbf{P}$  and  $\mathbf{M}$  are the bandpower binning and pseudo- $C_\ell$  mixing matrices, respectively. Elements were selected from  $\text{Cov}_{\text{unbinned}}$  corresponding to the weighted average  $\ell$  within each bandpower

<sup>16</sup> This number comes from a reduced data vector of 12 bandpowers and 15 power spectra, giving a data vector of length  $n = 12 \times 15 = 180$  and a number of unique covariance elements of  $n(n+1)/2 = 16\,290$ .

to give  $\text{Cov}_{\text{sampled}}$ . The matrices of binning weights  $\mathbf{w}_{\text{bin}}$  and mixing weights  $\mathbf{w}_{\text{mix}}$  were then calculated as

$$\mathbf{w}_{\text{bin}} = \frac{\text{Cov}_{\text{binned}}}{\text{Cov}_{\text{sampled}}} \quad (\text{elementwise}); \quad (14)$$

$$\mathbf{w}_{\text{mix}} = \frac{\text{Cov}_{\text{mixed}}}{\text{Cov}_{\text{binned}}} \quad (\text{elementwise}). \quad (15)$$

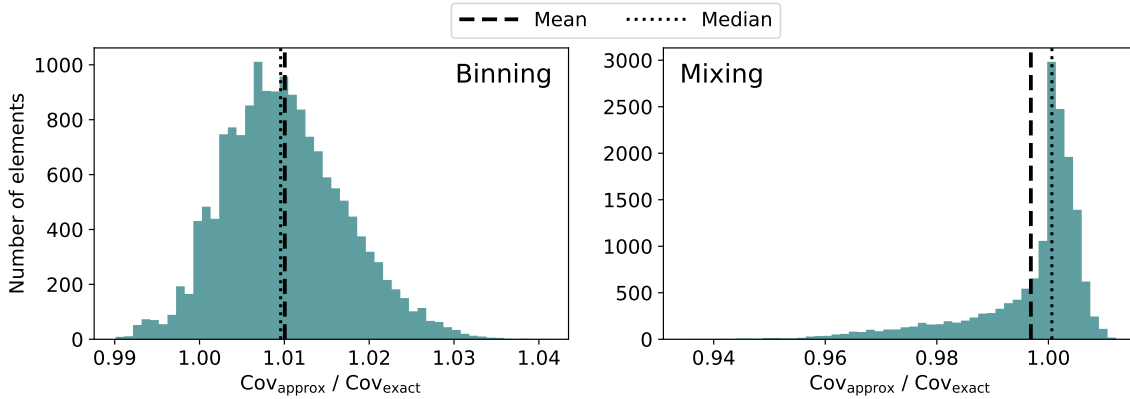
Finally, the sampled covariance blocks  $\text{Cov}_{\text{sampled}}$  were calculated for every block in the full covariance matrix, and transformed to give approximate binned and mixed covariance blocks as

$$\text{Cov}_{\text{binned\_approx}} = \mathbf{w}_{\text{bin}} * \text{Cov}_{\text{sampled}} \quad (\text{elementwise}); \quad (16)$$

$$\text{Cov}_{\text{mixed\_approx}} = \mathbf{w}_{\text{mix}} * \text{Cov}_{\text{binned\_approx}} \quad (\text{elementwise}). \quad (17)$$

While this two-step weighting process could be equivalently formulated as a single step, we chose to separate the effect of the binning and mixing approximations for additional insight into their respective effects.

We were able to validate these approximations by carrying out an equivalent process for the super-sample covariance matrix and comparing the results to those obtained using the full correct treatment. Histograms of the ratios between the approximate and exact covariance for each step, for all elements of the bandpower covariance across all redshift bins, are shown in Fig. 6. We find that each step introduces a bias of order per cent on average (although conveniently in opposite directions) with a



**Fig. 6.** Validation of the connected non-Gaussian approximation used to obtain the mock parameter constraints in Sect. 4.2, which is described in Sect. 4.2.1. Histograms of the ratio of the approximate to exact covariance are shown, for the ‘binning’ (left) and ‘mixing’ (right) steps, for all elements of the bandpower covariance matrix across all redshift bins, measured using the super-sample covariance. The results in all other sections are obtained using the connected non-Gaussian component, calculated in full.

spread of a few per cent. We find this to be sufficiently accurate for our purposes here, especially considering that the connected non-Gaussian component is the smallest of the three, but nevertheless this small potential error should be borne in mind when interpreting the results. Since the super-sample and connected non-Gaussian covariance contributions were shown to be similarly smooth in Sect. 3, the fact that this approximation works well for the former suggests that it should too for the latter.

#### 4.2.2. Results

Two-parameter constraints for different combinations of covariance components are shown in Fig. 7. The top row shows dark energy equation of state parameters ( $w_0$ ,  $w_a$ ), where  $w(a) = w_0 + w_a(1 - a)$ . The bottom row shows the matter density  $\Omega_m$  and the amplitude of the matter power spectrum at  $z = 0$  on the scale of  $8 h^{-1}$  Mpc,  $\sigma_8$ . The three columns are for the three different masks. In each panel, all parameters other than the two shown are held fixed. Only the 1 and  $3\sigma$  confidence regions are marked, which respectively contain the highest 68.3 and 99.7 per cent of the posterior probability mass. We list the relative areas of the  $3\sigma$  confidence region for each combination of parameters and mask in Table 1.

We find that the Gaussian contribution ( $G$ ) alone only covers 30–38 per cent of the full  $3\sigma$  region for  $(w_0, w_a)$  and 37–49 per cent for  $(\Omega_m, \sigma_8)$ . The Gaussian and connected non-Gaussian components combined ( $G + \text{CNG}$ ) cover 51–63 and 54–70 per cent for  $(w_0, w_a)$  and  $(\Omega_m, \sigma_8)$ , respectively, while the Gaussian and super-sample components combined ( $G + \text{SS}$ ) cover 82–84 and 90–92 per cent. These results are broadly in line with the single-parameter error bar ratios obtained in Barreira et al. (2018b).

There is some amount of apparent mask dependence: we find that as the sky cut is increased, the relative area of  $G + \text{SS}$  sees a very small increase (by 2–3 per cent) whereas  $G + \text{CNG}$  and  $G$  see larger decreases (by 16–22 and 8–12 per cent, respectively). This is consistent with the expectation that super-sample covariance should become more important as the sky is cut further, since this excludes more modes from the survey.

We also find some small shifts in the posterior means between the different composite covariance results for a given mask, despite the fact that the mean of the Gaussian likelihood used in each case is identical and only the covariance differs. This demonstrates how an incorrect covariance leads to an incor-

rect weighting of the random scatter present in the data due to cosmic variance, and therefore to posterior constraints having not only the wrong size but an erroneous position too, although this is a small effect.

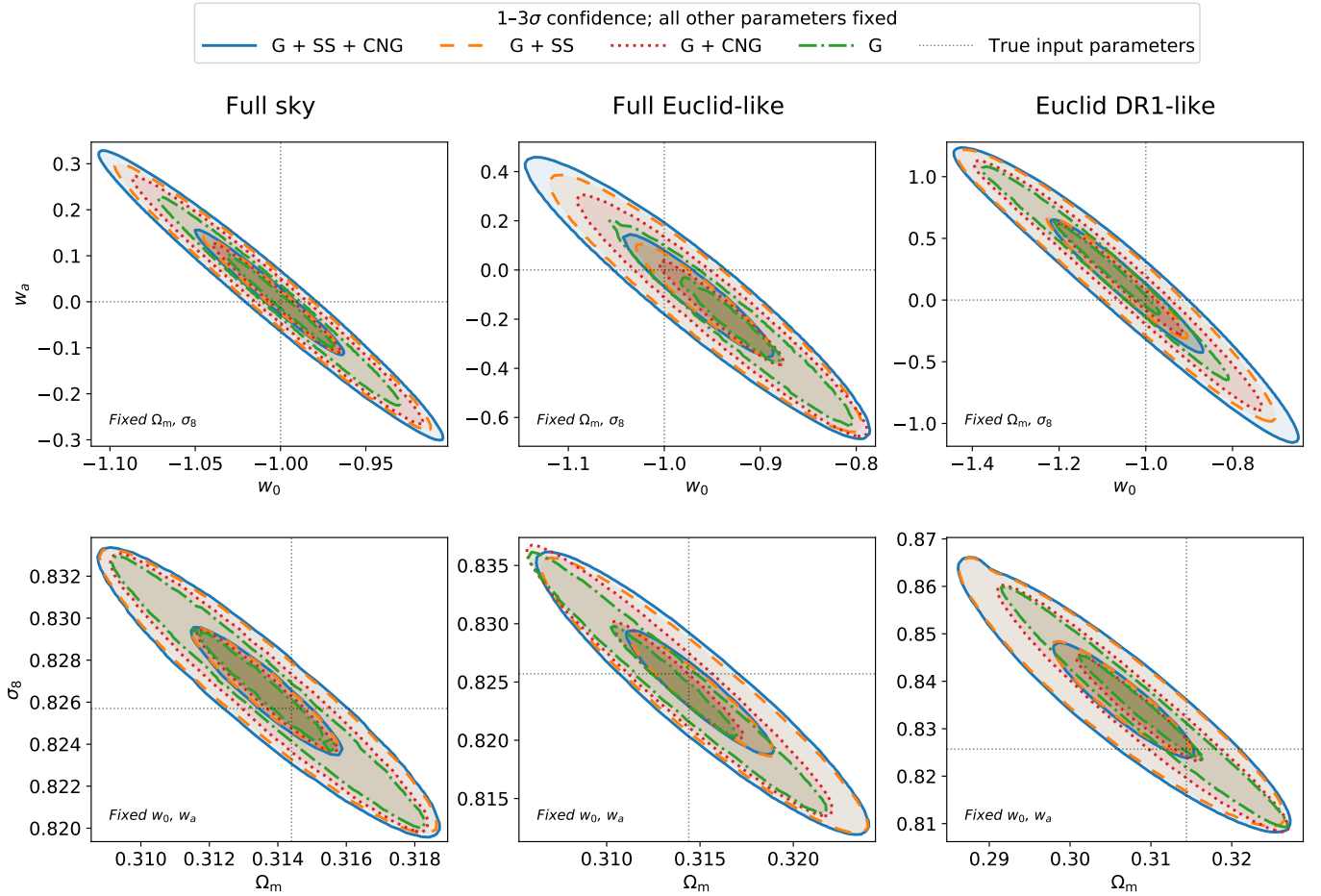
#### 4.2.3. Effect of marginalisation over additional parameters

We are motivated to explore the impact of marginalisation over additional parameters on the relative importance of each covariance component by the results of Barreira et al. (2018b). In that paper, the authors produced mock parameter constraints with different covariance contributions included, both for a single parameter at a time with all others fixed and for five parameters with all five allowed to vary simultaneously. For the latter case, they display marginalised two-parameter constraints (Barreira et al. 2018b, their Fig. 3). For  $(w_0, w_a)$  in particular (and to a lesser extent with some other parameter pairs), the  $2\sigma$  confidence region obtained with the Gaussian covariance alone appears roughly the same as that for the total covariance, to within the sampling noise. This is in contrast to their single-parameter constraints, for which the Gaussian covariance only produced  $\sim 50$  per cent of the  $1\sigma$  uncertainty on  $w_0$  obtained using the full covariance (see Fig. 2 of Barreira et al. 2018b). This raises the question of whether marginalisation might reduce the differences between the constraints obtained using different combinations of covariance components.

We explore this question by performing a three-parameter likelihood analysis using the full *Euclid*-like mask over  $(w_0, w_a, \Omega_m)$ , with all other parameters still held fixed. We show one- and two-parameter marginalised constraints in Fig. 8. In Table 2 we compare relative areas and widths of one- and two-parameter  $3\sigma$  confidence regions before and after marginalisation over a third parameter.

While the relative areas in Fig. 8 appear qualitatively similar to those in Fig. 7, we find that there are in fact some substantial quantitative differences, as shown by the values in Table 2. In particular, there is almost a doubling in the area of the constraints on  $(w_0, w_a)$  from the Gaussian covariance only ( $G$ ) relative to the total covariance ( $G + \text{SS} + \text{CNG}$ ) – from 36 to 69 per cent – when marginalising over  $\Omega_m$  rather than holding it fixed. A similar but smaller increase is seen for the other subsets ( $G + \text{SS}$ ,  $G + \text{CNG}$ ) of the total covariance for the same parameters.

For constraints on  $w_a$  alone, we find that the width of the  $3\sigma$  confidence region for  $G$  relative to  $G + \text{SS} + \text{CNG}$  increases



**Fig. 7.** Two-parameter constraints for different masks and different combinations of covariance contributions: Gaussian (G), super-sample (SS), and connected non-Gaussian (CNG). Shape noise is included, following Eq. (9). In each panel, all parameters other than the two shown are held fixed. Only the 1 and  $3\sigma$  confidence regions are marked, which respectively contain the highest 68.3 and 99.7 per cent of the posterior probability mass. The relative areas of each  $3\sigma$  confidence region are listed in Table 1. We note that the axis ranges differ between panels.

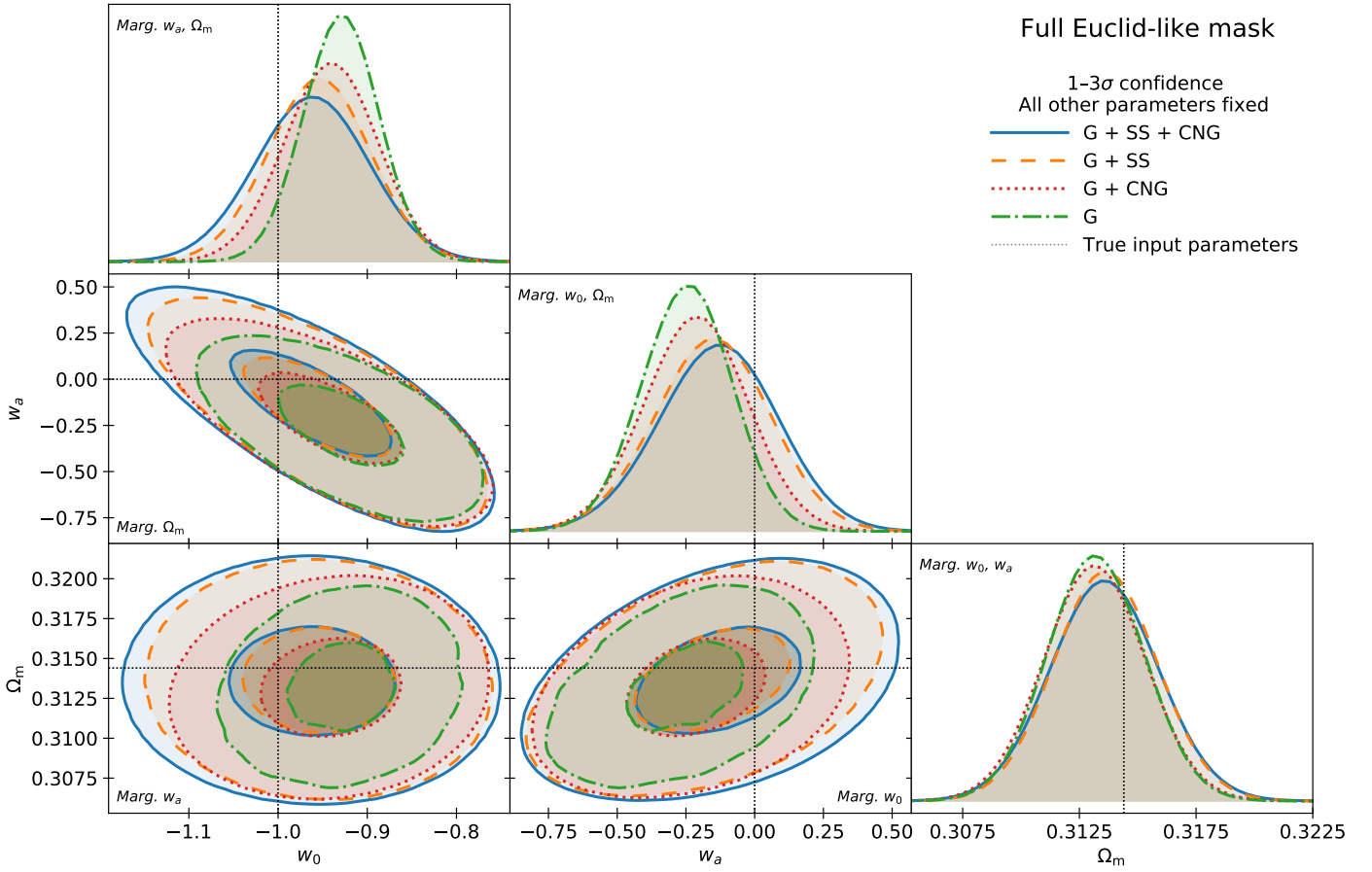
**Table 1.** Relative areas of  $3\sigma$  confidence regions in Fig. 7.

Parameters	Mask	Relative area of $3\sigma$ confidence region (%)			
		G + SS + CNG	G + SS	G + CNG	G
$(w_0, w_a)$	Full sky	100	82	63	38
	Full <i>Euclid</i> -like mask	100	84	61	36
	<i>Euclid</i> DR1-like mask	100	85	51	30
$(\Omega_m, \sigma_8)$	Full sky	100	90	70	49
	Full <i>Euclid</i> -like mask	100	90	66	45
	<i>Euclid</i> DR1-like mask	100	92	54	37

**Notes.** The table shows the relative area of the  $3\sigma$  confidence region in Fig. 7 for each combination of parameters and mask and for different combinations of covariance components: Gaussian (G), super-sample (SS), and connected non-Gaussian (CNG). The relative  $1\sigma$  areas are similar to the  $3\sigma$  areas.

slightly when marginalising over both  $w_0$  and  $\Omega_m$  compared to only marginalising over  $w_0$ , and a similar slight increase is seen for G + SS and G + CNG. However, for constraints on  $w_0$ , we find a small decrease in relative widths when marginalising over both  $w_a$  and  $\Omega_m$  rather than only  $w_a$ . One reason for this difference in behaviour between  $w_0$  and  $w_a$  may be that – as seen in Fig. 8 – there is clearly a much stronger correlation between  $w_a$  and  $\Omega_m$  than between  $w_0$  and  $\Omega_m$ . Marginalisation over a strongly correlated parameter should broaden constraints more than marginalisation over a more weakly correlated param-

eter (indeed, marginalisation over a truly independent parameter should have no effect at all), but it is not obvious that this should change the ratio of relative areas rather than simply broadening all constraints by the same factor. Regardless of the origin of this behaviour, it does appear to be the case that marginalisation over additional parameters – particularly those with which the constrained parameters are correlated – affects the relative importance of the difference covariance contributions. This is in agreement with the findings of Barreira et al. (2018b).



**Fig. 8.** Two- and one-parameter marginalised constraints obtained from a joint three-parameter analysis of  $(w_0, w_a, \Omega_m)$  for the full *Euclid*-like mask, including different combinations of covariance contributions: Gaussian (G), super-sample (SS), and connected non-Gaussian (CNG). Shape noise is included, following Eq. (9). The constraints in each panel have been obtained by marginalising over one or two parameters in the joint three-parameter posterior; for example, the panel marked ‘Marg.  $w_a, \Omega_m$ ’ has been marginalised over  $w_a$  and  $\Omega_m$ . All other parameters are held fixed. Only the 1 and 3  $\sigma$  confidence regions are marked.

**Table 2.** Impact of marginalisation on 3 $\sigma$  confidence regions.

Parameter(s)	Marginalised over	Relative area or width of 3 $\sigma$ confidence region (%)			
		G + SS + CNG	G + SS	G + CNG	G
$(w_0, w_a)$	–	100	84	61	36
	$\Omega_m$	100	89	84	69
$w_0$	$w_a$	100	91	88	72
	$(w_a, \Omega_m)$	100	90	83	69
$w_a$	$w_0$	100	88	84	70
	$(w_0, \Omega_m)$	100	96	85	74

**Notes.** The table shows the relative areas (for two-parameter constraints) and widths (for one-parameter constraints) of the 3 $\sigma$  confidence regions obtained using different combinations of covariance contributions, Gaussian (G), super-sample (SS), and connected non-Gaussian (CNG), for the full *Euclid*-like mask. Each row contains two sub-rows: the top sub-row is based on a two-parameter fit, which is marginalised over zero or one parameters; the bottom sub-row is based on a three-parameter fit, which is marginalised over one or two parameters. The relative 1 $\sigma$  areas are similar to the 3 $\sigma$  areas.

## 5. Conclusions

As the era of next-generation weak lensing surveys such as *Euclid* rapidly approaches, it is increasingly important to understand the properties of all steps of an analysis pipeline, including the covariance used in the likelihood. We have described in Sect. 2 how existing publicly available codes can be used in combination to calculate the full covariance matrix of cosmic shear pseudo- $C_\ell$  estimates, including the full details of an

arbitrary mask. We have further shown in Sect. 3 that existing simulations can be used to verify the accuracy of a theoretical covariance, and we found a high degree of agreement and consistency between theory and simulations. This agreement persists for different masks, showing that the theoretical covariance contributions correctly account for both the cut-sky mode coupling that is inherent to the pseudo- $C_\ell$  method and the non-Gaussian mode coupling, including additional cut-sky super-sample covariance.



This is encouraging for the use of pseudo- $C_\ell$  estimators in weak lensing, whose convenience and speed make them an attractive choice of analysis framework for future surveys. However, an outstanding challenge with such estimators is the need to understand their statistical properties sufficiently well such that they can be used to deliver reliable cosmological constraints to the precision and accuracy needed by future high-precision weak lensing surveys. This challenge has now to a large degree been addressed since we now know that not only is a Gaussian likelihood sufficient (Upham et al. 2021), but a full covariance can be evaluated and validated using the methods shown in this paper.

We have found that it is essential to include the non-Gaussian contributions to the covariance, even though cut-sky mode coupling means that the Gaussian covariance component dominates off-diagonal modes close to the main diagonal. The relative size and importance of the Gaussian component increases when including shape noise, but we have shown in Sect. 4 that only including the Gaussian component in parameter inference can lead to an underestimation of uncertainties by up to 70 per cent. The dominant non-Gaussian covariance component is the super-sample covariance, but neglecting the subdominant connected non-Gaussian covariance component can still lead to uncertainty underestimation on the scale of 10–20 per cent. In addition, neglecting some covariance contributions can lead to biases in the position of posterior parameter constraints as well as their size. However, a real cosmological analysis will require marginalisation over many nuisance parameters, which will decrease the relative importance of all cosmological contributions to the covariance, so these values should be taken as upper limits on the importance of each component. Perhaps for this reason it was found in the analysis of DES Year 3 data in Friedrich et al. (2021) that the connected non-Gaussian contribution could be entirely neglected, but our results suggest that we cannot necessarily extend this conclusion to a *Euclid*-like survey. However, this need not be an inconvenience, since we have shown that approximations of the kind described in Sect. 4.2.1 can be used to obtain the connected non-Gaussian component in a manageable amount of time with a loss of accuracy of only a few per cent on an already subdominant term. Finally, we have found that marginalisation over additional cosmological parameters may have a substantial effect on the relative importance of the different covariance components. We can conclude from this that it is important to take all marginalisation into account when, for example, determining the required accuracy of theoretical results for a particular science goal or producing forecasts of parameter constraints. The consistency of our results with those of Barreira et al. (2018b) implies that cut-sky mode coupling has relatively little impact on the respective importance of the covariance components.

The  $N$ -body weak lensing simulations are available online<sup>17</sup> and are described in Takahashi et al. (2017). We additionally make available online<sup>18</sup>.

- tomographic pseudo- $C_\ell$  power spectra measured from these simulations and
- the full connected non-Gaussian covariance matrix for the lowest redshift bin

All other data used in this article can be quickly generated using code that we have made available online<sup>19</sup>.

**Acknowledgements.** We thank the internal and external referees and colleagues for helpful feedback, which has improved the manuscript. This work would not have been possible without the contributions to the community from the developers of NaMaster (Alonso et al. 2019; García-García et al. 2019), CosmoLike (Krause & Eifler 2017; Fang et al. 2020) and the simulations of Takahashi et al. (2017). REU acknowledges a studentship from the UK Science and Technology Facilities Council. The Euclid Consortium acknowledges the European Space Agency and a number of agencies and institutes that have supported the development of *Euclid*, in particular the Academy of Finland, the Agenzia Spaziale Italiana, the Belgian Science Policy, the Canadian Euclid Consortium, the French Centre National d’Etudes Spatiales, the Deutsches Zentrum für Luft- und Raumfahrt, the Danish Space Research Institute, the Fundação para a Ciência e a Tecnologia, the Ministerio de Economía y Competitividad, the National Aeronautics and Space Administration, the National Astronomical Observatory of Japan, the Nederlandse Onderzoekschool Voor Astronomie, the Norwegian Space Agency, the Romanian Space Agency, the State Secretariat for Education, Research and Innovation (SERI) at the Swiss Space Office (SSO), and the United Kingdom Space Agency. A complete and detailed list is available on the *Euclid* web site (<http://www.euclid-ec.org>).

## References

- Alonso, D., Sanchez, J., & Slosar, A. 2019, *MNRAS*, **484**, 4127
- Barreira, A., & Schmidt, F. 2017a, *JCAP*, **2017**, 53
- Barreira, A., & Schmidt, F. 2017b, *JCAP*, **2017**, 51
- Barreira, A., Krause, E., & Schmidt, F. 2018a, *JCAP*, **2018**, 015
- Barreira, A., Krause, E., & Schmidt, F. 2018b, *JCAP*, **2018**, 053
- Blot, L., Corasaniti, P. S., Amendola, L., & Kitching, T. D. 2016, *MNRAS*, **458**, 4462
- Brown, M. L., Castro, P. G., & Taylor, A. N. 2005, *MNRAS*, **360**, 1262
- Camacho, H., Andrade-Oliveira, F., Troja, A., et al. 2021, ArXiv e-prints [arXiv:2111.07203]
- Challinor, A., & Chon, G. 2005, *MNRAS*, **360**, 509
- Chiang, C.-T., Wagner, C., Schmidt, F., & Komatsu, E. 2014, *JCAP*, **2014**, 48
- Cooray, A., & Hu, W. 2001, *ApJ*, **554**, 56
- Cooray, A., & Sheth, R. 2002, *Phys. Rep.*, **372**, 1
- Dark Energy Survey Collaboration 2005, ArXiv e-prints [arXiv:astro-ph/0510346]
- Di Valentino, E., Anchordoqui, L. A., Akarsu, Ö., et al. 2021, *Astropart. Phys.*, **131**, 102606
- Efstathiou, G. 2004, *MNRAS*, **349**, 603
- Euclid Collaboration (Pocino, A., et al.) 2021, *A&A*, **655**, A44
- Euclid Collaboration (Scaramella, R., et al.) 2022, *A&A*, in press, <https://doi.org/10.1051/0004-6361/202141938>
- Fabbian, G., Carron, J., Lewis, A., & Lembo, M. 2021, *Phys. Rev. D*, **103**, 43535
- Fang, X., Eifler, T., & Krause, E. 2020, *MNRAS*, **497**, 2699
- Friedrich, O., Andrade-Oliveira, F., Camacho, H., et al. 2021, *MNRAS*, **503**, 3125
- García-García, C., Alonso, D., & Bellini, E. 2019, *JCAP*, **2019**, 043
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, **622**, 759
- Gouyou Beauchamps, S., Lacasa, F., Tutusaus, I., et al. 2022, *A&A*, **659**, A128
- Hall, A., & Taylor, A. 2019, *MNRAS*, **483**, 189
- Hamilton, A. J., Rimes, C. D., & Scoccimarro, R. 2006, *MNRAS*, **371**, 1188
- Hansen, F. K., & Górski, K. M. 2003, *MNRAS*, **343**, 559
- Harnois-Déraps, J., & Van Waerbeke, L. 2015, *MNRAS*, **450**, 2857
- Harnois-Déraps, J., Amon, A., Choi, A., et al. 2018, *MNRAS*, **481**, 1337
- Harnois-Déraps, J., Giblin, B., & Joachimi, B. 2019, *A&A*, **631**, A160
- Harrison, I., Camera, S., Zuntz, J., & Brown, M. L. 2016, *MNRAS*, **463**, 3674
- Hikage, C., Oguri, M., Hamana, T., et al. 2019, *PASJ*, **71**, 43
- Hilbert, S., Hartlap, J., & Schneider, P. 2011, *A&A*, **536**, A85
- Howlett, C., Lewis, A., Hall, A., & Challinor, A. 2012, *JCAP*, **2012**, 027
- Ivezić, Z., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, **873**, 111
- Joachimi, B., Schneider, P., & Eifler, T. 2008, *A&A*, **477**, 43
- Krause, E., & Eifler, T. 2017, *MNRAS*, **470**, 2100
- Krause, E., Eifler, T. F., Zuntz, J., et al. 2017, ArXiv e-prints [arXiv:1706.09359]
- Krause, E., Fang, X., Pandey, S., et al. 2021, ArXiv e-prints [arXiv:2105.13548]
- Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints [arXiv:1110.3193]
- Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, **538**, 473
- Li, Y., Hu, W., & Takada, M. 2014, *Phys. Rev. D*, **89**, 83519
- Loureiro, A., Whittaker, L., Spurio Mancini, A., et al. 2021, *A&A*, submitted, [arXiv:2110.06947]
- Miyazaki, S., Komiyama, Y., Nakaya, H., et al. 2012, *Proc. SPIE*, **8446**, 84460Z
- Nicola, A., García-García, C., Alonso, D., et al. 2021, *JCAP*, **2021**, 067
- Pielorz, J., Rödiger, J., Tereno, I., & Schneider, P. 2010, *A&A*, **514**, A79
- Sato, M., & Nishimichi, T. 2013, *Phys. Rev. D*, **87**, 123538

<sup>17</sup> [http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky\\_raytracing](http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing)

<sup>18</sup> <https://doi.org/10.5281/zenodo.5163132>

<sup>19</sup> [https://github.com/robinupham/shear\\_pcl\\_cov](https://github.com/robinupham/shear_pcl_cov)

- Sato, M., Takada, M., Hamana, T., & Matsubara, T. 2011, *ApJ*, **734**, 76
- Schmidt, A. S., White, S. D., Schmidt, F., & Stücker, J. 2018, *MNRAS*, **479**, 162
- Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002, *A&A*, **396**, 1
- Schneider, A., Stoira, N., Refregier, A., et al. 2020, *JCAP*, **2020**, 19
- Sellentin, E., & Heavens, A. F. 2016a, *MNRAS*, **456**, L132
- Sellentin, E., & Heavens, A. F. 2016b, *MNRAS*, **464**, 4658
- Sgier, R. J., Réfrégier, A., Amara, A., & Nicola, A. 2019, *JCAP*, **2019**, 44
- Shirasaki, M., Hamana, T., & Yoshida, N. 2015, *MNRAS*, **453**, 3043
- Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, *MNRAS*, **341**, 1311
- Square Kilometre Array Cosmology Science Working Group (Bacon, D. J., et al.) 2020, *PASA*, **37**, E007
- Takada, M., & Hu, W. 2013, *Phys. Rev. D*, **87**, 123504
- Takada, M., & Jain, B. 2009, *MNRAS*, **395**, 2065
- Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, *ApJ*, **761**, 152
- Takahashi, R., Hamana, T., Shirasaki, M., et al. 2017, *ApJ*, **850**, 24
- Upham, R. E., Whittaker, L., & Brown, M. L. 2020, *MNRAS*, **491**, 3165
- Upham, R. E., Brown, M. L., & Whittaker, L. 2021, *MNRAS*, **503**, 1999
- Wagner, C., Schmidt, F., Chiang, C.-T., & Komatsu, E. 2015, *JCAP*, **2015**, 42
- Zonca, A., Singer, L., Lenz, D., et al. 2019, *J. Open Source Softw.*, **4**, 1298
- Zuntz, J., Paterno, M., Jennings, E., et al. 2015, *Astron. Comput.*, **12**, 45
- <sup>1</sup> Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK  
e-mail: robin.upham@manchester.ac.uk
- <sup>2</sup> Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
- <sup>3</sup> Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
- <sup>4</sup> INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy
- <sup>5</sup> INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese (TO), Italy
- <sup>6</sup> Department of Mathematics and Physics, Roma Tre University, Via della Vasca Navale 84, 00146 Rome, Italy
- <sup>7</sup> INFN-Sezione di Roma Tre, Via della Vasca Navale 84, 00146 Roma, Italy
- <sup>8</sup> INAF-Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131 Napoli, Italy
- <sup>9</sup> Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, 4150-762 Porto, Portugal
- <sup>10</sup> INAF-IASF Milano, Via Alfonso Corti 12, 20133 Milano, Italy
- <sup>11</sup> Port d'Informació Científica, Campus UAB, C. Albareda s/n, 08193 Bellaterra (Barcelona), Spain
- <sup>12</sup> Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona), Spain
- <sup>13</sup> INAF-Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monteporzio Catone, Italy
- <sup>14</sup> Department of Physics "E. Pancini", University Federico II, Via Cinthia 6, 80126 Napoli, Italy
- <sup>15</sup> INFN section of Naples, Via Cinthia 6, 80126 Napoli, Italy
- <sup>16</sup> Dipartimento di Fisica e Astronomia "Augusto Righi" – Alma Mater Studiorum Università di Bologna, Via Piero Gobetti 93/2, 40129 Bologna, Italy
- <sup>17</sup> INAF-Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125 Firenze, Italy
- <sup>18</sup> Centre National d'Etudes Spatiales, Toulouse, France
- <sup>19</sup> Institut national de physique nucléaire et de physique des particules, 3 rue Michel-Ange, 75794 Paris Cédex 16, France
- <sup>20</sup> Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
- <sup>21</sup> ESAC/ESA, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, 28692 Villanueva de la Cañada, Madrid, Spain
- <sup>22</sup> European Space Agency/ESRIN, Largo Galileo Galilei 1, 00044 Frascati, Roma, Italy
- <sup>23</sup> Univ Lyon, Univ Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, 69622 Villeurbanne, France
- <sup>24</sup> Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK
- <sup>25</sup> Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, 1749-016 Lisboa, Portugal
- <sup>26</sup> Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
- <sup>27</sup> Department of Astronomy, University of Geneva, ch. d'Ecogia 16, 1290 Versoix, Switzerland
- <sup>28</sup> Université Paris-Saclay, CNRS, Institut d'astrophysique spatiale, 91405 Orsay, France
- <sup>29</sup> Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH, UK
- <sup>30</sup> INFN-Padova, Via Marzolo 8, 35131 Padova, Italy
- <sup>31</sup> AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris, 91191 Gif-sur-Yvette, France
- <sup>32</sup> Institut d'Estudis Espacials de Catalunya (IEEC), Carrer Gran Capità 2-4, 08034 Barcelona, Spain
- <sup>33</sup> Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
- <sup>34</sup> INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34131 Trieste, Italy
- <sup>35</sup> Istituto Nazionale di Astrofisica (INAF) – Osservatorio di Astrofisica e Scienza dello Spazio (OAS), Via Gobetti 93/3, 40127 Bologna, Italy
- <sup>36</sup> Istituto Nazionale di Fisica Nucleare, Sezione di Bologna, Via Irnerio 46, 40126 Bologna, Italy
- <sup>37</sup> Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, 85748 Garching, Germany
- <sup>38</sup> Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstrasse 1, 81679 München, Germany
- <sup>39</sup> Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029 Blindern, 0315 Oslo, Norway
- <sup>40</sup> Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands
- <sup>41</sup> Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
- <sup>42</sup> von Hoerner & Sulger GmbH, Schloßplatz 8, 68723 Schwetzingen, Germany
- <sup>43</sup> Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark
- <sup>44</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
- <sup>45</sup> Aix-Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France
- <sup>46</sup> Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, 1211 Genève 4, Switzerland
- <sup>47</sup> Department of Physics and Helsinki Institute of Physics, Gustaf Hällströmin katu 2, 00014 University of Helsinki, Finland
- <sup>48</sup> NOVA optical infrared instrumentation group at ASTRON, Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands
- <sup>49</sup> Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
- <sup>50</sup> INFN-Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
- <sup>51</sup> Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
- <sup>52</sup> European Space Agency/ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
- <sup>53</sup> Department of Physics and Astronomy, University of Aarhus, Ny Munkegade 120, 8000 Aarhus C, Denmark
- <sup>54</sup> Institute of Space Science, Bucharest 077125, Romania
- <sup>55</sup> IFPU, Institute for Fundamental Physics of the Universe, Via Beirut 2, 34151 Trieste, Italy
- <sup>56</sup> Dipartimento di Fisica e Astronomia "G. Galilei", Università di Padova, Via Marzolo 8, 35131 Padova, Italy

- <sup>57</sup> Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Avenida Complutense 40, 28040 Madrid, Spain
- <sup>58</sup> Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal
- <sup>59</sup> Universidad Politécnica de Cartagena, Departamento de Electrónica y Tecnología de Computadoras, 30202 Cartagena, Spain
- <sup>60</sup> Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA
- <sup>61</sup> INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy
- <sup>62</sup> Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
- <sup>63</sup> Dipartimento di Fisica, Università degli Studi di Torino, Via P. Giuria 1, 10125 Torino, Italy
- <sup>64</sup> INFN-Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy
- <sup>65</sup> INFN-Sezione di Roma, Piazzale Aldo Moro, 2 – c/o Dipartimento di Fisica, Edificio G. Marconi, 00185 Roma, Italy
- <sup>66</sup> Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, 10010 New York, NY, USA
- <sup>67</sup> School of Physics and Astronomy, Cardiff University, The Parade, Cardiff CF24 3AA, UK
- <sup>68</sup> Space Science Data Center, Italian Space Agency, Via del Politecnico snc, 00133 Roma, Italy
- <sup>69</sup> Astrophysics Group, Blackett Laboratory, Imperial College London, London SW7 2AZ, UK
- <sup>70</sup> Mathematical Institute, University of Leiden, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands