The Integrated 3-Point Correlation Functions of Cosmic Shear and Projected Galaxy Density Fields

Anik Halder

München 2023

The Integrated 3-Point Correlation Functions of Cosmic Shear and Projected Galaxy Density Fields

Anik Halder

Contents

Zusammenfassung

Mithilfe des schwachen Gravitationslineneffekts versuchen Kosmologen, die Natur der dunklen Materie und dunklen Energie zu verstehen. Programme zur Auswertungen großskaliger Himmelsdurchmusterungen verfolgen dieses Ziel zurzeit hauptsächlich durch Messung und Analyse der 2-Punkt-Korrelationsfunktion (2PKF) des kosmischen Scherungsfelds, d.h. der geringfügigen Verzerrungen von Bildern von Hintergrundgalaxien durch gravitative Gezeitenfelder im Vordergrund. Die großskalige Struktur unseres Universums folgt jedoch *keiner Gaußverteilung*, und signifikante Anteile kosmologischer Informationen sind in höheren Momenten des Scherungsfeldes enthalten, welche nicht von diesen traditionellen 2-Punkt-Statistiken erfasst werden. Das zuverlässige Extrahieren dieser Information höherer Ordnung ist insbesondere daher erstrebenswert, weil sie Messungen kosmologischer Parameter stark verbessern kann. Im Laufe des letzten Jahrzehnts wurden eine Vielzahl von Statistiken höherer Ordnung vorgeschlagen, um höhere Momente des Scherungsfelds zu quantifizieren und zu nutzen. Solche Statistiken stehen aber in der Regel vor einer Reihe von Herausforderungen, insbesondere dem hohen numerischen Aufwand für ihre Schätzungen sowie schwerwiegenden Mängel in den theoretischen Modellen. Daher ist die Anwendung solcher Methoden bisher begrenzt geblieben. Um diese Hürden zu überwinden, entwickele und analysiere ich in dieser Arbeit eine neuartige Statistik höherer Ordnung des kosmischen Scherungsfelders namens *integrierte 3-Punkt-Korrelationsfunktion* (3PKF) $\zeta_\pm$. Diese kann direkt anhand von kosmischem Scherungsdaten gemessen werden, indem lokale Messungen der Scherungs-2PKF mit der mittleren Scherungs-Aperturmasse innerhalb von lokalen sub-Volumen einer Himmelsdurchmusterung korreliert werden. Zusätzlich zu Details und Validierung dieser Methode präsentiere ich eine gemeinsame Analyse von $\zeta_\pm$ und den traditionellen Scherungs-2PKF $\xi_\pm$ in verblindeten Daten des drei-Jahres-Datensatzes des sogenannten Dark Energy Survey (DES). Unsere Ergebnisse zeigen, dass die von $\zeta_\pm$ zu $\xi_\pm$ addierte Information signifikante Verbesserungen der Messungen kosmologischer Parameter ermöglicht, insbesondere eine Verbesserung von etwa 40% für den Zustandsgleichungsparameter der dunklen Energie, w_0.

Ermutigt von diesen Ergebnissen analysiere ich auch das Potential der integrierten 3-Punkt-Kreuzkorrelationen zwischen kosmischem Scherungs- und den Dichtefeldern. Damit erweitere ich das weitgenutzte Schema der 3×2-Punkt-Korrelationsfunktionen von Scherungs- und Galaxiendichtefeld auf ein praktisches Schema höherer Ordnung: die integrierten 6×3-Punkt-Korrelationsfunktionen. Damit können nicht nur kosmologische, sondern auch Parameter der Beziehung von Materie- und Galaxiendichte um 20-40% genauer bestimmt

werden. Diese Ergebnisse motivieren daher zukünftige Anwendungen der Galaxie-Scherung
integrierten 3PKF in Bedobachtungsdaten. Neben der integrierten 3PKF präsentiere ich
auch meine Beteiligung an Fortschritten bei der theoretischen Modellierung einer anderen
Statistik, welche ebenfalls in der Lage ist, einen umfassenderen Blick auf die großskalige
Struktur des kosmischen Dichtefeldes zu werfen als die 2PKF: die gemeinsame Wahrschein-
lichkeitsdichtefunktion $p(\delta_m, \delta_g)$ des lokalen Materiedichtekontrasts δ_m und des Galaxien-
dichtekontrasts δ_g. Auch diese Analysemethode kann nicht nur kosmologische Parameter
einschränken, sondern auch detaillierte Informationen über die Verbindung zwischen der
unsichtbaren dunklen Materie und dem beobachteten Galaxiendichtefeld extrahieren. Die
Beiträge dieser Arbeit ebnen somit den Weg für eine effiziente und effektive Ausnutzung
der nicht-Gaußschen Information in aktuellen und zukünftigen Himmelsdurchmusterungen.

Abstract

A central goal of weak gravitational lensing cosmology is to understand the nature of dark matter and dark energy. The key program of ongoing lensing surveys involves 2-point correlation function (2PCF) analyses of the cosmic shear field, minute distortions of background galaxy images by intervening foreground large-scale structure (LSS) of our Universe, to constrain cosmological parameters. However, the LSS is *non-Gaussian* distributed with important information in its higher-order moments, not captured by these traditional 2-point statistics. Reliably extracting this higher-order information can therefore enable tighter cosmological parameter constraints. Over the past decade, a plethora of higher-order statistics have hence been proposed to harness this information from lensing data, but several challenges including computationally expensive estimation, significant deficiencies in theoretical models etc. have kept their analyses rather few and standalone. To overcome these hurdles, in this thesis, I have developed and analysed a novel weak lensing higher-order statistic called the *integrated 3-point correlation function* (3PCF) $\zeta_\pm$ which can be directly measured on cosmic shear data by correlating the local measurements of shear 2PCF with the mean lensing aperture mass signal within patches on the sky-survey. In addition to modelling and validating this statistic, me and my collaborators have also performed a joint analysis of $\zeta_\pm$ alongside the traditional shear 2PCF $\xi_\pm$ in the blinded cosmic shear Year 3 dataset of the Dark Energy Survey. Our results show that the addition of $\zeta_\pm$ to $\xi_\pm$ yields significant tightening of cosmological parameter constraints, specifically $\sim 40\%$ improvement on the dark energy equation of state parameter w_0.

Encouraged by these promising results I have also proposed the integrated 3-point cross-correlations between cosmic shear and the foreground galaxy density fields thus extending the popular 3×2-point galaxy-shear correlation functions to the practical higher-order framework: integrated 6×3-point correlations which have the potential to bring further 20-40% improvements on not only cosmological but also on galaxy bias parameters which describe the connection between the invisible dark matter and the observed galaxy density field that traces it. These results therefore motivate future applications of the galaxy-shear integrated 3PCFs in real data. In addition to the integrated 3PCF, I also present my contributions to advances in the theoretical modelling of another statistic capable of obtaining a more comprehensive view of the LSS than 2PCFs: the joint probability density function $p(\delta_m, \delta_g)$ of local matter δ_m and galaxy density δ_g fluctuations which can not only constrain cosmological parameters but also extract detailed information about the galaxy bias parameters. The contributions of this thesis will thus pave the way for a efficient and

effective utilisation of the non-Gaussian information from current and upcoming galaxy imaging sky-surveys.

Outline of this thesis

This thesis broadly consists of two parts. The first part, chapters 1 and 2 describe the context within which this work is based. These two chapters together discuss some basics of cosmology, the smooth background universe, matter perturbations on the homogeneous background, their evolution, and motivate the emergence of the non-Gaussian distribution of the large-scale structure (LSS) of the Universe. They also describe the observed projected fields in galaxy imaging surveys namely the weak lensing and galaxy clustering fields and the traditional 2-point statistical methods used to analyse these fields. The topics discussed in these chapters contain standard material which can be found in most cosmology textbooks and articles. Besides Secs. 2.5.2 and 2.5.3, the only original content in these chapters are my errors.[1]

The second part of the thesis contains the novel contributions made by me together with my collaborators, on two higher-order statistics which can extract the non-Gaussian information content of the LSS that is not accessible to traditional 2-point statistics. These are the *integrated 3-point correlation functions (3PCF)* where the material is reprinted from the articles Halder et al. (2021); Halder & Barreira (2022); Gong et al. (2023); Halder et al. (2023) in Chapters 3, 4, 5 and 7, respectively. In chapter 6, I present (Halder and DES collaboration (in prep.)) the *blinded* cosmological constraints from the analysis of the integrated 3PCF on real cosmic shear data from the Year 3 data release of the Dark Energy Survey. In chapter 8 I reprint the article published in Friedrich et al. (2022) where we have advanced the development of the joint 1-point *probability density function (PDF) of the galaxy and matter density fluctuations*, another higher-order LSS statistic.

For the quick reader, I present a short illustrative summary of the key contributions of this thesis (integrated 3PCFs and PDF) in chapter 9. I conclude that chapter with some future directions that can be envisioned with the frameworks established in this work.

Thank you for reading!

---

[1]This joke is borrowed from MacKenzie (2000).

Chapter 1

Background

In the current picture of modern cosmology, our Universe emerged around 13.8 billion years ago with a homogeneous and isotropic distribution of matter at a very high temperature and density. Since then, the Universe has expanded, cooled down and developed a magnificent network of structures consisting of stars, galaxies, galaxy clusters, voids and filaments that we collectively call the cosmic web. It is understood that tiny fluctuations embedded in the matter density of the early Universe (which were initially quantum fluctuations stretched to macroscopic scales by an extremely brief inflationary epoch at very early times) have evolved to form the cosmic web through a myriad of interesting physical processes playing out on every length scale. One of the goals in cosmology is to understand the picture of our Universe on large scales and understand the nature of the constituents and the forces which shape the cosmic web.

In this chapter we briefly discuss some of the fundamental ideas in cosmology and summarize our current understanding of the growth of structures in the Universe. We discuss the commonly observed cosmological fields which are extensively studied in order to decipher the large-scale distribution of structures with statistical methods. The material reviewed in this chapter can be found in most graduate level cosmology textbooks. We have adapted the material mainly from the following textbooks - Mukhanov (2005), Hobson et al. (2006), Dodelson & Schmidt (2020), Coles & Lucchin (2002).

1.1   Basics of cosmology

In this section we briefly summarise some of the concepts and equations which are at the heart of modern cosmology.

Two of the fundamental assumptions in cosmology are (Mukhanov, 2005):

- the Universe is **spatially** homogeneous and isotropic when viewed on ('smoothed' over) spatial scales larger than 100 Mpc. This is known as the **cosmological principle** and has been verified to be true with observational tests such as from the cosmic microwave background (CMB) map Planck Collaboration et al. (2018) or from the

distribution of galaxies from galaxy redshift surveys (Ntelis et al., 2017). On smaller scales there are highly inhomogeneous (non-linear) structures such as galaxies, galaxy clusters etc.

- On large scales, the Universe evolves through gravitational interactions which are described by the theory of General Relativity (GR).

Ever since Edwin Hubble (Hubble, 1929) found that distant galaxies appear to to be receding from us, we have known that the Universe is not static, but changing with time. Therefore it is important to emphasise that in the context of Einstein's GR where we view our Universe as a 4-dimensional (4D) *spacetime*[1], the idea of homogeneity and isotropy applies only to the 3-dimensional (3D) spatial part of this spacetime. This condition that space is homogeneous and isotropic restricts the geometry of the spatial part to only three general possibilities with constant curvature: $K > 0$ (**closed** 3D sphere with constant positive curvature), $K = 0$ (**flat** 3D space), $K < 0$ (**open** 3D hyperbolic space with constant negative curvature) (Mukhanov, 2005). We shall briefly discuss below why this is the case.

We can think of *slicing* spacetime into a time-ordered sequence of 3D spatial *hyper-surfaces*[2]. The cosmological principle requires each of these 3D spatial hyper-surfaces to be homogeneous and isotropic (Hobson et al., 2006) or, in other words, to be *maximally symmetric* (such a space has the largest possible symmetry). On the other hand, any given space has an intrinsic curvature (geometry) which is characterised by a quantity called the *Riemann curvature tensor* $R_{\alpha\beta\gamma\delta}$ whose components define the geometric properties of the space. In 3D, this curvature tensor has, in general, six independent components — three independent translations and three rotations — each of which is a function of the spatial coordinates. Hence, one needs six functions to define these components for such a general 3D space. However, the more symmetric a space is, the fewer are the number of functions required to specify these components (or fewer the number of independent components). Therefore, for a maximally symmetric space, one requires only a single constant number — the curvature K which is independent of the spatial coordinates — to characterise the space (Hobson et al., 2006). This is why there are only three possibilities for the curvature: $K > 0$, $K = 0$, $K < 0$.

Hence, the line element $\mathrm{d}l^2$ of a 3D homogeneous and isotropic space of constant curvature K when expressed in spherical coordinates (r, θ, ϕ) is given by (Mukhanov, 2005):

$$\begin{aligned} \mathrm{d}l^2 &= \gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j \\ &= a^2\left(\frac{\mathrm{d}r^2}{1 - Kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right) \end{aligned} \tag{1.1}$$

---

[1] 3 dimensions of space $(x^1, x^2, x^3) \equiv (x, y, z)$ and 1 dimension of time $x^0 \equiv t$.

[2] Each spatial hyper-surface is spanned by the spacelike coordinates $\boldsymbol{x} = (x^1, x^2, x^3)^{\mathrm{T}}$ where T stands for transpose. A particular value of time t refers to a specific spatial hyper-surface.

where a^2 is a positive quantity and defines the curvature scale of the 3D space (e.g. $\frac{a}{\sqrt{K}}$ is the radius of the 3D sphere when $K > 0$) and γ_{ij} is the *metric tensor* for the given space. The metric tensor defines the notion of distance between a pair of points located in the space. Mathematically, its form will look different when expressed in different coordinate systems. In the above equation we have chosen to express γ_{ij} in spherical coordinates. Note that in the first line of the equation we are using the Einstein summation convention with the indices i, j running from 1 to 3 (over the three spatial coordinates).

We can now write down the line element $\mathrm{d}s^2$ for 4D spacetime of our Universe which consists of a time-ordered sequence of homogeneous and isotropic spatial hypersurfaces. Using natural units in which the speed of light $c = 1$ and a $(+, -, -, -)$ signature for the 4D spacetime metric $g_{\mu\nu}$, the line element can be written as (Mukhanov, 2005):

$$
\begin{aligned}
\mathrm{d}s^2 &= g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu \\
&= \mathrm{d}t^2 - \mathrm{d}l^2 \\
&= \mathrm{d}t^2 - a^2(t)\left(\frac{\mathrm{d}r^2}{1 - Kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right)
\end{aligned}
\tag{1.2}
$$

where a is now a function of time and is known as the scale factor — it indicates 'how big' the spacelike hyper-surface is at a given moment of time t. Making a time-dependent is the only way to preserve the homogeneity and isotropy of 3D space and still allow for its time evolution (Mukhanov, 2005). The metric tensor for 4D spacetime is denoted as $g_{\mu\nu}$ where μ, ν runs from 0 to 3 i.e. $(x^0, x^1, x^2, x^3) \equiv (t, r, \theta, \phi)$. This 4D metric tensor for the Universe under the assumption of the cosmological principle is called the Friedmann-Leimatre-Robertson-Walker (FLRW) metric. Given the metric tensor, one can compute the Riemann curvature tensor $R_{\alpha\beta\gamma\delta}$ of the 4D spacetime and find its contracted forms — the *Ricci tensor* $R_{\alpha\beta}$ and the *Ricci scalar* R. With this, one can compute the Einstein tensor $G_{\mu\nu}$ which represents the geometry of spacetime (Hobson et al., 2006):

$$
G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} \ .
\tag{1.3}
$$

So far, we have discussed only about the geometry of the Universe, but we have not talked about its content. Another underlying assumption in modern cosmology is that:

- On large scales, the Universe is uniformly filled with matter and energy which can be described as a perfect fluid (Mukhanov, 2005).

A perfect fluid is characterised by an energy density ρ, pressure p, and 4-velocity u_μ. The energy-momentum tensor $T_{\mu\nu}$ of this perfect fluid is given by (Mukhanov, 2005):

$$
T_{\mu\nu} = (\rho + p)u_\mu u_\nu + pg_{\mu\nu}
\tag{1.4}
$$

where the pressure p is related to the energy density ρ by the **equation of state**. In most cases that are studied in cosmology, the fluids of interest are so called *barotropic* fluids

for which the sound speed is given by $c_s^2 = \mathrm{d}p/\mathrm{d}\rho$ (Dodelson & Schmidt, 2020) and the equation of state reads as (Mukhanov, 2005):

$$p = w\rho \tag{1.5}$$

where w is known as the equation of state parameter. The energy-momentum tensor represents the matter/energy content of spacetime.

Einstein's theory of gravity relates the geometry of spacetime with the distribution of matter within it. Mathematically, this is expressed through the field equations of GR (Hobson et al., 2006):

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{1.6}$$

where G is Newton's gravitational constant. Equation (1.6) can be interpreted as a set of equations which dictate how matter/energy determines the curvature (geometry) of spacetime.

Thus, from our perfect fluid energy-momentum tensor and geometry of our Universe governed by the FLRW metric, we can now write the equations for $a(t)$ and K using the field equations. These equations are known as the Friedmann equations and they describe the evolution of the Universe (Mukhanov, 2005):

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} \ , \tag{1.7}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + 3p\right) \ . \tag{1.8}$$

In words, the equations relate the constant scalar curvature K of the 3D spatial hypersurfaces and the time evolution of the scale factor $a(t)$ of these hyper-surfaces to the pressure and energy of the matter (for a single matter species) contained in the Universe. The scale factor of the Universe today (at time t_0) is often chosen as a reference point and is set to 1 i.e. $a_0 \equiv a(t_0) = 1$.

Differentiating the first Friedmann equation with respect to time and combining with the second equation we get (Mukhanov, 2005):

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0 \ . \tag{1.9}$$

This is the energy conservation equation (which is also contained in the field equations) and it relates the time evolution of the energy density of a given species of matter. Using the equation of state (1.5), one can solve the above equation to find that the energy density of a particular barotropic matter species with equation of state parameter w relates to the scale factor as

$$\rho(t) = \rho_0 a^{-3(1+w)} \tag{1.10}$$

where ρ_0 is the energy density of that particular matter/energy species today.

In the discussions so far, the equations have been formulated in terms of the scale factor $a(t)$. However, it should be noted that a cannot be measured directly. An important quantity associated to the scale factor is the notion of **redshift** which is a more direct observable than a. The redshift z of a distant luminous object (e.g. galaxy) is defined as (Coles & Lucchin, 2002):

$$z = \frac{\lambda_0 - \lambda_e}{\lambda_e} \tag{1.11}$$

where λ_e is the wavelength of the photon emitted by the object at time t_e (when the scale factor of the Universe is $a(t_e)$) and λ_0 is the wavelength of the photon when observed by us at time t_0 (when the scale factor of the Universe is $a(t_0)$). In terms of the scale factor the redshift can be written as (Coles & Lucchin, 2002):

$$z = \frac{a(t_0)}{a(t_e)} - 1 \ . \tag{1.12}$$

The larger the redshift, the farther away an object is and the further back in time we look (as the photon naturally takes longer to traverse the Universe if the object is more distant).

1.2   Matter and energy content of the Universe

For multiple non-interacting species of matter ρ_i in the Universe, the total energy-momentum tensor in (1.6) is given by a sum of the separate energy-momentum tensors of different species. Hence, we can write the Friedmann equations (equations (1.7), (1.8)) as

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3} \sum_i \rho_i - \frac{K}{a^2} \ , \tag{1.13}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} \sum_i (\rho_i + 3p_i) \ . \tag{1.14}$$

In cosmology, the *Hubble* parameter H, characterises the rate of expansion, is defined as

$$H(t) \equiv \frac{\dot{a}}{a} \ . \tag{1.15}$$

The value of the Hubble parameter today is called the Hubble constant H_0. It is usually written as

$$H_0 = 100 \ h_{100} \ \mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$$

where h_{100} is the dimensionless form of the Hubble constant divided by $100 \ \mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$. Its current value is approximately $h_{100} \approx 0.7$ (Dodelson & Schmidt, 2020). In terms of the H, we can rewrite the first Friedmann equation as

$$\frac{8\pi G}{3H^2} \sum_i \rho_i - \frac{K}{a^2 H^2} = 1 \ . \tag{1.16}$$

The quantity

$$\rho_c(t) \equiv \frac{3H^2}{8\pi G} \tag{1.17}$$

can be identified as the *critical density*. This is because from equation (1.16) we can see that if the value of $\sum_i \rho_i = \rho_c$, then $K = 0$ (i.e. a flat Universe). Otherwise, if $\sum_i \rho_i$ is either larger or smaller than ρ_c, then the Universe has either a closed ($K > 0$) or an open ($K < 0$) geometry, respectively. The curvature of the Universe (spatial part) depends completely on the total energy density of matter/energy species contained within it.

Defining the *density parameter* Ω_i for a given species of matter:

$$\Omega_i(t) = \frac{\rho_i}{\rho_c} \tag{1.18}$$

and for the spatial curvature:

$$\Omega_K(t) = -\frac{K}{a^2 H^2} \tag{1.19}$$

we see that the first Friedmann equation (equation (1.13)) is written as

$$\sum_i \Omega_i + \Omega_K = 1 \ . \tag{1.20}$$

Inferring the amount of matter and energy content of different species present in the Universe allows us to understand the geometry and fate of the Universe. Hence, it is a topic of intense investigation to find and constrain the values of these density parameters $\Omega_{i,0}$ (i.e. at the current time t_0) for different matter/energy species and also the values for the Hubble constant H_0 and the curvature scalar K, with ever increasing precision.

We will now briefly discuss about the different matter/energy species which make up our Universe.

Matter

The first constituent of the Universe is the so called *dust* (or matter) — collisionless, nonrelativistic matter, with pressure[3] $p = 0$ and thus, $w = 0$. The energy density for dust evolves as $\rho_m \propto a^{-3}$ (see equation (1.10)) which can simply be interpreted as the decrease in the number density of particles as the Universe expands (because volume increases as a^3). However, there is a distinction even between the type of dust. Firstly, there is the luminous **baryonic matter** which we experience in our every day lives — objects which are made of normal atomic matter — and this matter interacts not only gravitationally but also through electromagnetic (that is why we can see these objects), strong and weak nuclear forces. Examples include ordinary rocks, planets, stars, star clusters and galaxies for which the

---

[3]In cosmology, when the pressure of a fluid $p \ll \rho c^2$ i.e. the pressure is gravitationally insignificant one denotes it as pressure-free matter.

pressure is negligible in comparison with their energy density. However, observations (e.g. nucleosynthesis of light elements after the Big Bang, from the CMB (Dodelson & Schmidt, 2020)) show that $\Omega_{b,0} \approx 0.05$ which is a very tiny budget of the total energy density content of the Universe. Most of the matter which is present in the Universe is in a non-luminous form known as **dark matter** which is thought to only interact gravitationally with the remaining energy in the Universe. The evidence for the presence of dark matter has been confirmed e.g. from gravitational lensing observations around galaxy clusters (for a review see Massey et al. (2010)). It is an active field of research to understand what makes up this dark matter. It is also currently understood that dark matter is presumably 'cold' implying that it does not have high velocities, allowing it to clump together. This cold dark matter is currently estimated to have a density parameter of $\Omega_{cdm,0} \approx 0.25$ (Mukhanov, 2005).

Radiation

The other form of energy that pervades the Universe is **radiation** and it can describe either actual electromagnetic radiation (with an equation of state parameter $w_\gamma = 1/3$), or massive particles moving at very high velocities (relativistic — close to the speed of light) such that they are indistinguishable from photons (i.e. their $w_\gamma \approx 1/3$). The energy density in radiation evolves as $\rho_\gamma \propto a^{-4}$ (see equation (1.10)) which can be interpreted as the usual decrease in the number density of photons as the Universe expands (as a^{-3}) and the stretching of the wavelength of the radiation (or the decrease in frequency/energy of the radiation — another a^{-1} factor) due to the expansion. So the energy density of radiation falls off faster than that of matter (which evolves as a^{-3}). Today the total energy density in the radiation component is insignificant and the density parameter is estimated to be $\Omega_{\gamma,0} \approx 0.005$ (Dodelson & Schmidt, 2020). From this, we can realise that further back in time, there was once an epoch in the early Universe when the energy density of radiation was much larger than that of dust. Thus, the very early Universe can be thought of as a radiation dominated era. After the radiation energy density became insignificant, the matter component started dominating the energy budget of the Universe.

Dark energy

If the Universe would only contain matter (dark matter and baryons) and radiation, the total energy density of the Universe would act to slow down the expanding Universe that Edwin Hubble had observed in 1920s (irrespective of the geometry of the Universe i.e. whether the Universe is open, flat or closed). However, recent observations (e.g. distance measurements of Type Ia supernovae (Riess et al., 1998)) have shown that the current expansion of the Universe is rather speeding up (accelerating i.e. $\ddot{a} > 0$) than slowing down! This phenomenon can only be explained in a Universe where $\rho + 3p < 0$ (see equation (1.8)). This is the case when the dominant contribution to the energy density budget of the Universe is by a so called **dark energy** component with an equation of state parameter $w_{DE} < -\frac{1}{3}$. Although we do not know what constitutes this dark energy, observations (Planck Collaboration et al., 2018) have shown that $w_{DE,0} \approx 0.7$ with $\omega_{DE} \approx -1$. The

equation of state $\omega_{\text{DE}} = -1$ can be interpreted as adding a **cosmological constant** Λ to the field equations:

$$G_{\mu\nu} - \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{1.21}$$

which in turn can be thought of as a constant energy density (in the Friedmann equations) given by $\rho_{\text{DE}} = -p_{\text{DE}} = \frac{\Lambda}{8\pi G}$ associated to the energy density of the vacuum of space itself. The scale factor of a spatially flat, cosmological constant dominated Universe grows exponentially with time $a(t) \propto \exp(H_0 t)$ (Mukhanov, 2005).

Adding up all these measured density parameters for matter/energy (excluding curvature) it turns out that the total energy density of our Universe $\Omega_{\text{tot},0} \approx 1$ resulting in a universe with a spatial geometry close to being flat. These parameters taken together comprise the **Lambda Cold Dark Matter** (ΛCDM) model of cosmology. It is currently accepted to be the standard model of modern cosmology. Besides ΛCDM, it is also of great interest to constrain other cosmological models such as wCDM where the dark energy equation of state $w_0 \equiv w_{\text{DE}}$ is allowed to be different from 1. Such a model with $w_0 \neq 1$ portrays dark energy to be different from being a cosmological constant. Testing these different cosmological models and constraining their parameters with observations is one of the main goals of current cosmological research.

1.3   Structures in the Universe

According to our current theoretical understanding (Mukhanov, 2005), at the very beginning, the Universe seems to have underwent a period of exponential expansion (much like today's era of dark-energy domination but for a very small period of time) for a tiny fraction of a second after which it stopped growing exponentially. The energy from inflation was converted into radiation, matter particles and the Universe entered the normal expanding phase as supported by the standard Big Bang model. The quantum fluctuations which were present in the primordial Universe were stretched to macroscopic scales due to this inflationary expansion and became the tiny fluctuations δ embedded in the dense hot plasma of particles and radiation. δ, often called the **density contrast** or **density fluctuation**, in any energy density field ρ (e.g. radiation, matter etc.) is defined as

$$\delta = \frac{\rho - \bar{\rho}}{\bar{\rho}} \;. \tag{1.22}$$

Although the energy density is positive i.e. $\rho \geq 0$, the density contrast can take negative values with a lower limit of -1 i.e. $\delta \geq -1$. The ρ can thus be expressed in terms of δ by writing $\rho = \bar{\rho}(1 + \delta)$. It is important to note that from its definition, the mean value of the density contrast is zero i.e. $\langle \delta \rangle = 0$.

These tiny fluctuations δ in the hot plasma manifested as small over and underdensities in the gravitational potential landscape of the matter-energy content of the early Universe.

The dark matter particles which only interacted gravitationally, started falling into these small overdensities (where the gravitational potential was larger) and clumped together to make the potential wells even deeper (overdense regions got more overdense). On the other hand, the underdensities became even more underdense as material started moving towards the overdensities. Nevertheless, the baryonic matter (protons, electrons) was tightly coupled to the energetic radiation (e.g. through scattering of light by free electrons) as the temperature of the Universe was still very high. As the Universe expanded, the hot plasma cooled down enough for atoms to form for the very first time i.e. photons were not energetic enough to prevent a proton and an electron coming together to form a Hydrogen atom (this is known as *recombination*). This happened almost 380000 years after the Big Bang and it led to the first radiation being emitted that could freely travel through the Universe. The emitted light from that period which reaches us today from all directions is a faint glow which has been redshifted to the microwave range — hence, the name Cosmic Microwave Background (CMB). During recombination, the radiation which escaped from the over and underdensities of the dark matter density field left corresponding imprints in the observed CMB in the form of temperature fluctuations which are seen to be of the order of $\delta T \sim 10^{-5}$ (Mukhanov, 2005). These temperature fluctuations are observed (on the CMB sky-map) to be spatially distributed almost according to a Gaussian random field i.e. there was a complete symmetry in the abundance and amplitude of underdense and overdense regions in the early Universe.

After recombination, the baryons which had decoupled from the photons started falling into the overdense regions of the dark matter density field. In these regions, the Universe started coalescing material to form objects such as stars and eventually galaxies. Where the overdensities in the dark matter field were larger, more number of objects (e.g. galaxies) formed. Hence, it is thought that the observed galaxies and galaxy clusters trace the underlying dark matter density field but they do so in a *biased*[4] manner in the sense that the galaxy density fluctuations are enhanced or suppressed over those of the matter fluctuations. The spatial distribution of the initial matter density fluctuations which was close to a Gaussian, preserved its shape at the beginning when $|\delta| \ll 1$ i.e. as long as the gravitational collapse was in the *linear* regime (Mo et al., 2010). However, with time, as the amplitude of the fluctuations became $|\delta| \geq 1$, the distribution became non-symmetrical (or non-Gaussian) due to the *non-linear*[5] gravitational collapse of dark matter and baryonic material. This eventually led to the formation of large overdense structures such as galaxies and galaxy clusters over the next 10 billion years. However, for the last 3 billion years, the Universe entered the dark-energy dominated era which started another accelerated phase

---

[4]The bias is usually different for different tracers of the matter density field. For example, galaxies with different intrinsic properties (morphological type, luminosity etc.) are seen to have different clustering properties and hence different bias relations to the underlying matter field. This indicates that the bias contains information about the physics of galaxy formation (Baugh, 2013).

[5]The term linear collapse reflects the regime when the different Fourier modes of the density fluctuation δ evolves independently. On the other hand, non-linear gravitational collapse implies the coupling of different Fourier modes of the fluctuation.

of expansion of the Universe resulting in a suppressed rate of formation of structures on the largest scales. The observed cosmic web of Large Scale Structure (LSS) that we see today is thus a result of the interplay of these physical processes which have been taking place since the beginning of our Universe.

1.4   Cosmology with galaxy imaging surveys

Observationally, there are multiple ways to test and constrain the parameters of the ΛCDM or any other cosmological model. Depending on the particular parameter one wants to constrain, different 'observational probes' are studied. For example in order to constrain H_0 one can perform a gravitational lensing study of multiple-imaged quasars (Wong et al., 2019) or measure distances to Type Ia supernovae (Riess et al., 1998). Many of the density parameters can be measured accurately by observing the CMB (Planck Collaboration et al., 2018). Another way of constraining the density parameters is by studying the distribution of structures in the cosmic web of LSS. If the Universe has indeed evolved gravitationally in the way portrayed above (we shall discuss the gravitational evolution of density fluctuations in more detail in section 2.2.1), then studying the LSS can give us a lot of information about the underlying behaviour of our Universe. In order to do so, one has to observe millions of galaxies over large areas of the sky and map out the patterns of the cosmic web. Once the map of the field of specific types of tracers (e.g. luminous red galaxies have been studied to have a particular bias relation with the underlying matter field in a specific redshift range (Cabré & Gaztañaga, 2009)) is known, one can extract information from this map using various cosmological probes (e.g. galaxy clustering, gravitational lensing, redshift space distortions etc.) and compare the observed signals of these probes with their corresponding theoretical models and infer the values of the underlying model parameters. The goal of such galaxy surveys (LSS surveys) is to not only infer about the large scale cosmological behaviour of our Universe but to also understand the nature of galaxy and structure formation through the study of different biased tracers.

An example of such a survey of the LSS is the Dark Energy Survey (DES) (Dark Energy Survey Collaboration et al., 2016), whose primary goal is to infer whether 70% of the Universe is filled with a dark-energy within the context of GR, or whether gravity behaves differently on large scales than what GR predicts — thereby giving rise to the accelerated phase of expansion that we observe today (Abbott et al., 2018). In order to do so, the DES collaboration has photometrically imaged in 5 different optical filter bands (grizY) more than 100 million distant galaxies and tens of thousands of galaxy clusters in an area of approximately 5000 square degree of the sky (see Figure 1.1) using the 570 megapixel Dark Energy Camera (DECam), at the Blanco telescope of the Cerro Tololo Inter-American Observatory (CTIO) in Chile. Analysis of the DES data — (i) of the clustering of hundreds of thousands of galaxies and of the (ii) gravitational weak lensing distortion of the images of millions of background galaxies by the foreground cosmic structure — has given tight constraints on the values for the cosmological parameters (e.g. density parameters) asso-
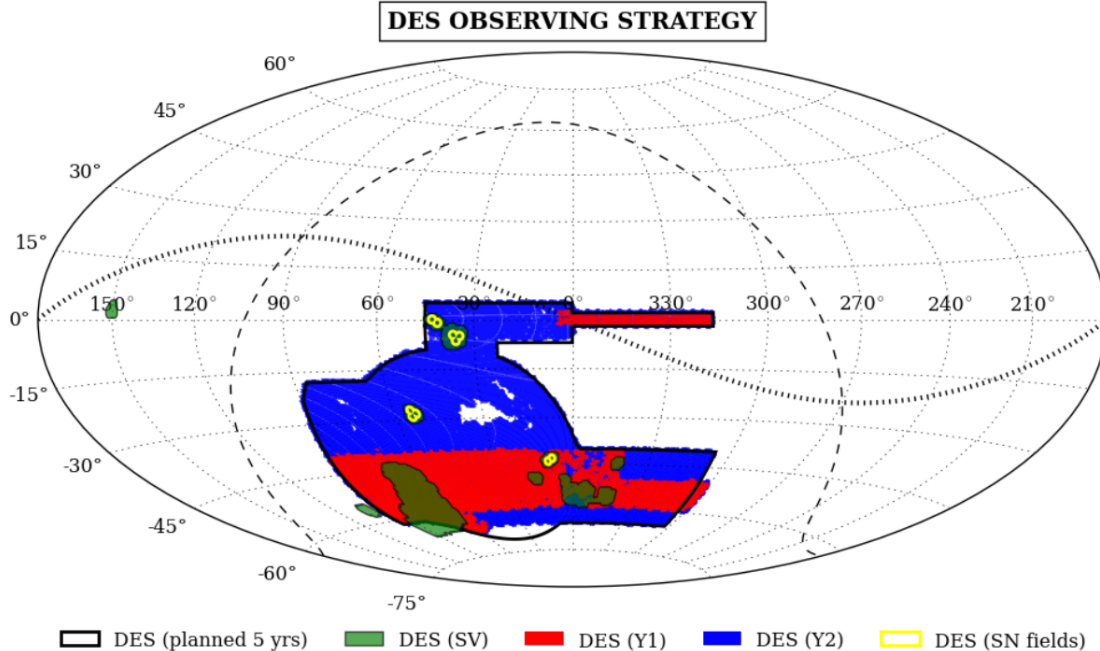
Figure 1.1: Observing strategy of the Dark Energy Survey shown in Hammer projection of the celestial sky in equatorial coordinates. The dashed and dotted lines indicate the Galactic plane and the ecliptic plane respectively. The area indicated on the map shows the Dark Energy Survey's observed footprint for different stages/programs of the survey: Science Verification (SV), Year 1 (Y1), Year 2 (Y2), Year 5 (Y5), and Supernova Program (SN). Image adapted from (Dark Energy Survey Collaboration et al., 2016).

ciated with the ΛCDM or wCDM models. This has been performed via the measurement and Markov-Chain-Monte-Carlo (MCMC) data inference of a particular summary statistic called the *2-point correlation function* of these probes (DES Collaboration, 2022) (see Sec. 2.5.1).

However, the 2-point correlation function is only the lowest order statistic that can be measured from the data of a density field. Evaluating higher-order correlation functions can provide a wealth of information on structure formation scenarios that cannot be probed by investigating only the second moment of the galaxy and matter density field distributions. This is the main theme of this thesis where we conceptualise, model and investigate a particular higher-order statistic called the *integrated 3-point correlation function* for weak lensing datasets and analyse this statistic's constraining power in constraining cosmological model parameters when complementing the 2-point analyses in Year 3 DES (DESY3) cosmic shear data. We also study another higher-order statistic called the *probability density function* of the cosmic density fields which is sensitive to all higher-order 1-point moments of the field thereby enabling a more encompassing view of the cosmos than just the information in the variance of the fields.

Chapter 2

Cosmic density fields and statistics

According to our current understanding, the quantum mechanical origin of the density fluctuations leads us to interpreting all the observed fields in cosmology as instances or realisations of random fields. CMB temperature fluctuations, CMB polarisation, dark matter density fluctuations etc. are a few examples of such observed random fields. In this chapter we will look at the mathematical definition of a random field and discuss the Gaussian random field which is a very good descriptor for the density fields in the early Universe. We shall also study the concept of power spectrum and correlation functions which will be used extensively in later chapters of the thesis. We also discuss the need for studying practical higher-order statistical methods which are required to extract more information out of the late-time cosmic density fields which are inherently non-Gaussian, and can therefore enable tighter constraints on the parameters of our cosmological models relative to 2-point statistics alone. We will then discuss some physical cosmological fields in the late Universe such as the weak lensing and galaxy contrast fields. We will conclude this chapter by discussing the traditionally analysed 2-point statistical methods on these two fields as well as higher-order statistics, namely the integrated 3-point correlation functions and the probability density function of the density fields, which we develop and study in the later chapters.

2.1   Random fields

A spatial random field F is a random process (defined in a volume V in a given D-dimensional space) which takes up random values at each location. In other words, at each point $\mathbf{x}_i \in V$, the value of the field is a random variable $F_i \equiv F(\mathbf{x}_i)$. Hence, the whole volume V can be regarded as a collection of random variables and an **instance** of this random field would be the values *realised* by these random variables (Xavier et al., 2016)

$$\boldsymbol{F} = \Big(F_1, F_2, F_3, ..., F_N \Big)^{\mathrm{T}} , \tag{2.1}$$

where the subscript $i = 1$ to N denote N mutually distinct points (locations $\mathbf{x}_i$) in V (Xavier et al., 2016) such that the value *realised* by the random variable at that location

is $F_i \equiv F(\mathbf{x}_i)$. This vector of random variables $\mathbf{F}$ can be described in terms of a joint multi-variate N-point probability density function (PDF) $p(\boldsymbol{F})$. In other words, this joint N-point PDF gives the probability of realising a certain instance[1] $\boldsymbol{F}$ of the random field. The notion of the random field can be also extended to fields described not only in 3D (like the matter density contrast) but also to random fields in 2D space i.e. on the surface of the celestial sphere - a sky map (e.g. the projected matter density as probed via weak lensing fields).

2.1.1 Gaussian random field

Observations of the temperature fluctuations in the CMB sky-map (Planck Collaboration et al., 2018) show that the temperature fluctuation field (hence, also the matter density contrast field) in the early Universe is extremely close to Gaussian. This property is still preserved in the density field when it evolves only through linear gravitational instability. Hence, modelling the density field as Gaussian is a good test-bed for theories of the early Universe.

For a random field F which is a Gaussian (normal) random field, the joint N-point PDF $p(\boldsymbol{F})$ of the field at any finite set of locations $\mathbf{x}_i \in V$ is given by (Xavier et al., 2016):

$$p(\boldsymbol{F}) = \frac{1}{\sqrt{(2\pi)^N \det(\mathbf{C})}} \exp\left(-\frac{1}{2}(\boldsymbol{F} - \boldsymbol{\mu})^{\mathrm{T}} \mathbf{C}^{-1}(\boldsymbol{F} - \boldsymbol{\mu})\right) \ . \tag{2.2}$$

Hence, a Gaussian field evaluated at N locations is a multi-variate Gaussian distribution and is characterised firstly by the mean vector

$$\boldsymbol{\mu} = \left(\mu_1, \mu_2, \mu_3, ..., \mu_N\right)^{\mathrm{T}} \equiv \left(\langle \mathrm{F}_1 \rangle, \langle \mathrm{F}_2 \rangle, \langle \mathrm{F}_3 \rangle, ..., \langle \mathrm{F}_N \rangle\right)^{\mathrm{T}} \ , \tag{2.3}$$

with $\langle . \rangle$ denoting the expectation value or the 'first moment' of a random variable. The expectation value is also often denoted as $E\{.\}$ and can be interpreted as an ensemble average (average over many independent realisations) or a volume average. The joint PDF is also characterised by the $N \times N$ covariance matrix[2] $\mathbf{C}$ where the entry to the i-th row and j-th column of the matrix is given by the covariance $\mathbf{cov}\,(\mathrm{F}_i, \mathrm{F}_j)$ of the random variables F_i and F_j. Equivalently, it is also called the 'second *central* moment' or the *connected part* of the '2-point correlation function' $\xi_{\mathrm{F}}\,(\mathbf{x}_i, \mathbf{x}_j)$ of the field F at points $\mathbf{x}_i$ and $\mathbf{x}_j$ (Friedrich et al., 2018):

$$\mathbf{C}_{ij} = \mathbf{cov}\,(\mathrm{F}_i, \mathrm{F}_j) = \xi_{\mathrm{F}}\,(\mathbf{x}_i, \mathbf{x}_j) = \left\langle\, (\mathrm{F}_i - \langle \mathrm{F}_i \rangle)\,(\mathrm{F}_j - \langle \mathrm{F}_j \rangle)\, \right\rangle \tag{2.4}$$

---

[1]A certain instance means a collection of 'realisations' of the N RVs as shown in equation (2.1). For example, if N identical coins are tossed simultaneously at the N locations, an instance of this coin toss random field might be $\boldsymbol{F} = \left(H, T, T, H, T, H, ..., T\right)^{\mathrm{T}}$ where H stands for Heads and T for Tails — the 2 sides of a coin.

[2]The covariance matrix is a general concept and has more usage than just the purpose of defining the joint PDF of a multi-variate Gaussian distribution as shown in this particular case.

given that the random variables F_i and F_j have a finite mean and variance (e.g. this is not true for the Cauchy distribution). The variance σ_i^2 of the random variable F_i is the C_{ii} element of the covariance matrix (the diagonal entries)

$$\sigma_i^2 = \mathbf{C}_{ii} = \left\langle (F_i - \langle F_i \rangle)^2 \right\rangle . \tag{2.5}$$

The positive square root of the variance is referred to as the standard deviation σ_i.

An interesting property of this multi-variate Gaussian distribution is that its marginal distribution for the random variable F_i characterised by the 1-point PDF at a given location $\mathbf{x}_i$, is also Gaussian (Xavier et al., 2016). This 1-point PDF $p(F_i)$ can be easily deduced from equation (2.2) to be:

$$p(F_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2}(F_i - \mu_i)^2 \right) . \tag{2.6}$$

Hence, a single Gaussian random variable is characterised by its mean μ_i and variance σ_i^2 (analogous to the multivariate case where the PDF is characterised by the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\mathbf{C}$ as discussed above).

One can go on to define the 3-point correlation function (third central moment) for any random field in general (not just for Gaussian) similar to equation (2.4) but for 3 points $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\} \in V$:

$$\xi_{\mathrm{F}}\left(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\right) = \left\langle (F_1 - \langle F_1 \rangle)(F_2 - \langle F_2 \rangle)(F_3 - \langle F_3 \rangle) \right\rangle . \tag{2.7}$$

Higher order correlation functions (central moments) like the 4-point, 5-point and so on till the N-point correlation (or N-point central moment) follow similarly:

$$\xi_{\mathrm{F}}\left(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, ..., \mathbf{x}_N\right) = \left\langle (F_1 - \langle F_1 \rangle)(F_2 - \langle F_2 \rangle)(F_3 - \langle F_3 \rangle)....(F_N - \langle F_N \rangle) \right\rangle . \tag{2.8}$$

One thing to note is that for a *zero-mean* field i.e. $\langle F_i \rangle = 0 \quad \forall \quad \mathbf{x}_i \in V$, equation (2.8) can be simply written as

$$\xi_{\mathrm{F}}\left(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, ..., \mathbf{x}_N\right) = \left\langle F_1 F_2 F_3 .... F_N \right\rangle . \tag{2.9}$$

It is interesting to note that for a uni-variate Gaussian distribution the only two interesting moments are the mean and the variance because all the higher order central moments are either 0 (all odd-numbered moments) or are described in terms of the variance (all even-numbered moments). This is known as **Isserlis' theorem** (Isserlis, 1918). In equation form, these higher order moments for a single Gaussian variable F_i reads (Xavier et al., 2016):

$$\langle (F_i - \langle F_i \rangle)^n \rangle = \begin{cases} 0 & n \text{ is odd,} \\ \frac{n!}{(n/2)!}\left(\frac{\sigma_i^2}{2}\right)^{n/2} & n \text{ is even.} \end{cases} \tag{2.10}$$

For $n = 2$, one retrieves σ_i^2 as expected from equation (2.5). This theorem can be generalised to the higher order moments (see equation (2.8)) for a multi-variate Gaussian

distribution. The same argument holds in that case as well and all the higher order even-numbered central moments such as the kurtosis are described in terms of the elements of the covariance matrix (if N is even in equation (2.8)). On the other hand, the odd-numbered central moments such as the *connected part* of the three-point correlation function (or the skewness) are 0 (if N is odd in equation (2.8)).

2.1.2   Homogeneity and isotropy of random fields

For any random field F, if the N-point correlation function has the following form

$$\xi_{\mathrm{F}}\left(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, ..., \mathbf{x}_N\right) = \xi_{\mathrm{F}}\left(\mathbf{R} \cdot \mathbf{x}_1 + \mathbf{a}, \mathbf{R} \cdot \mathbf{x}_2 + \mathbf{a}, \mathbf{R} \cdot \mathbf{x}_3 + \mathbf{a}, ..., \mathbf{R} \cdot \mathbf{x}_N + \mathbf{a}\right) \ , \quad (2.11)$$

where $\mathbf{a}$ is a translation vector and $\mathbf{R}$ is a rotational matrix, then the field F is said to be *statistically homogeneous* and *statistically isotropic*.

This can be understood easily from the case for a Gaussian field G where one has to only look at the 2-point correlation function (consequence of Isserlis' theorem) to describe the field. If the 2-point function $\xi_{\mathrm{G}}\left(\mathbf{x}_i, \mathbf{x}_j\right)$ at points $\mathbf{x}_i$ and $\mathbf{x}_j$ is only dependent on the separation vector $\mathbf{r} = \mathbf{x}_i - \mathbf{x}_j$ between the two points and not on their precise locations i.e.

$$\xi_{\mathrm{G}}\left(\mathbf{x}_i, \mathbf{x}_j\right) = \xi_{\mathrm{G}}\left(\mathbf{x}_i - \mathbf{x}_j\right) \equiv \xi_{\mathrm{G}}(\mathbf{r}), \quad (2.12)$$

and the entries of the mean vector $\boldsymbol{\mu}$ are all the same i.e. the expectation value of G is independent of the location and is a constant:

$$\langle \mathrm{G}(\mathbf{x}_i)\rangle = \text{constant} \quad \forall \quad \mathbf{x}_i \in V, \quad (2.13)$$

then the Gaussian field G is said to be *statistically homogeneous*. The variance in the case of a homogeneous random field is often referred to as the value of the correlation function (or the covariance) at 'zero-lag' $\xi_{\mathrm{G}}\left(\mathbf{0}\right)$ i.e. when $\mathbf{x}_i - \mathbf{x}_j = \mathbf{0}$.

On top of this, if the 2-point correlation between any two locations is independent of the direction and only depends on the magnitude of the separation vector i.e.

$$\xi_{\mathrm{G}}\left(\mathbf{x}_i, \mathbf{x}_j\right) = \xi_{\mathrm{G}}\left(\mathbf{x}_i - \mathbf{x}_j\right) = \xi_{\mathrm{G}}\left(|\mathbf{x}_i - \mathbf{x}_j|\right) \equiv \xi_{\mathrm{G}}(r), \quad (2.14)$$

then the Gaussian field G is said to be *statistically homogeneous* and *statistically isotropic*. $\xi_{\mathrm{G}}\left(\mathbf{x}_i, \mathbf{x}_j\right)$ tells us how correlated are any two points of the field which lie at a separation $r = |\mathbf{x}_i - \mathbf{x}_j|$.

Thus, one can infer that homogeneity can be associated with translational invariance, and isotropy with rotational invariance of a random field as shown in equation (2.11). For non-Gaussian fields one needs to check whether (2.11) holds for all N-point correlations and not just for $N = 2$ (which is the case for a Gaussian field as discussed above) to find whether the field is homogeneous and isotropic.

2.1.3 Spectral representation of random fields

A zero-mean, homogeneous random field F can be decomposed into its constituent frequencies or wave-functions (this is also known as the spectral representation theorem (Martinez & Saar, 2001)). In flat Euclidean space this is the Fourier series expansion of the function in terms of sines and cosines (the basis wave-functions). The Fourier decomposition of the random field F defined in volume V in flat space (in the continuum limit) is given by (Dodelson & Schmidt, 2020):

$$F(\mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \, F(\mathbf{k}) \, e^{i\mathbf{k}.\mathbf{x}} \tag{2.15}$$

where $F_{\mathbf{k}} \equiv F(\mathbf{k})$ is the Fourier coefficient (is a complex quantity in general) at a given wave-vector $\mathbf{k}$ in Fourier space[3]. $|F_{\mathbf{k}}|$ describes the amplitude of that Fourier mode. If the field $F(\mathbf{x})$ is real (which is generally true for all cosmological fields), then

$$F(-\mathbf{k}) = F(\mathbf{k})^* \, . \tag{2.16}$$

On the other hand, $F(\mathbf{k})$ is related to $F(\mathbf{x})$ by the inverse Fourier transform:

$$F(\mathbf{k}) = \int d^3\mathbf{x} \, F(\mathbf{x}) \, e^{-i\mathbf{k}.\mathbf{x}} \, . \tag{2.17}$$

Using the Fourier expansion of a zero-mean field F, one can express the 2-point correlation function $\xi_F(\mathbf{x}, \mathbf{x}') \equiv \langle F(\mathbf{x})F(\mathbf{x}') \rangle$ in Fourier space. When one does this for a real, homogeneous and isotropic field:

$$\langle F(\mathbf{x})F(\mathbf{x}') \rangle = \langle F(\mathbf{x})F^*(\mathbf{x}') \rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \int \frac{d^3\mathbf{k}'}{(2\pi)^3} \langle F(\mathbf{k})F^*(\mathbf{k}') \rangle \, e^{i\mathbf{k}.\mathbf{x}} \, e^{-i\mathbf{k}'.\mathbf{x}'}, \tag{2.18}$$

using the assumptions of homogeneity and isotropy of the field, it can be shown that the 2-point correlation function $\langle F(\mathbf{x})F(\mathbf{x}') \rangle$ is a Fourier transform of a quantity known as the 'power spectrum' $P_F(k)$ which is defined as follows (Dodelson & Schmidt, 2020)

$$\langle F(\mathbf{k})F(\mathbf{k}') \rangle = (2\pi)^3 P_F(k)\delta^D(\mathbf{k} + \mathbf{k}') \, , \tag{2.19}$$

where $\delta^D(\mathbf{k}+\mathbf{k}')$ is a Dirac-delta function which ensures the translational invariance (homogeneity) of the field in Fourier space depicting that different Fourier modes are uncorrelated (this is one of the main advantages of working in Fourier space). The rotational invariance (isotropy) is seen through the directional independence of the mode $\mathbf{k}$ in the argument of the power spectrum $P_F(k)$. It is often convenient to define a dimensionless form of the power spectrum as (Peacock, 1998):

$$\Delta_F^2(k) \equiv \frac{k^3}{2\pi^2}P_F(k) \, . \tag{2.20}$$

---

[3]A large magnitude of the Fourier mode $\mathbf{k}$ corresponds to a small physical length scale and vice versa.

Using equation (2.19) in equation (2.18) and then evaluating the integrals in spherical coordinates, one finds a simplified expression for the 2-point correlation function in terms of the dimensionless power spectrum $\Delta_{\mathrm{F}}^2(k)$ (Peacock, 1998)

$$\xi_{\mathrm{F}}\left(\mathbf{x}, \mathbf{x}'\right) \equiv \left\langle \mathrm{F}(\mathbf{x})\mathrm{F}(\mathbf{x}')\right\rangle = \int \frac{\mathrm{d}k}{k} \Delta_{\mathrm{F}}^2(k) j_0(k|\mathbf{x} - \mathbf{x}'|) \ , \tag{2.21}$$

where $j_0(\alpha) = \sin \alpha/\alpha$ is the spherical Bessel function of order zero. The expression for the variance of the field (2-point correlation at zero-lag i.e. $\mathbf{x} = \mathbf{x}'$) reads:

$$\xi_{\mathrm{F}}\left(0\right) \equiv \xi_{\mathrm{F}}\left(\mathbf{x}, \mathbf{x}\right) = \int \mathrm{d}\ln k \ \Delta_{\mathrm{F}}^2(k) \ . \tag{2.22}$$

When the dimensionless power-spectrum $\Delta_{\mathrm{F}}^2(k) = \mathrm{constant}$ i.e. scale-invariant, the variance gets equal contributions from every decade in k.

As an aside, one should note that $\mathrm{F}(\mathbf{k})$ is related to $\mathrm{F}(\mathbf{x})$ as an integral which is a linear operator — see equation (2.17). This means that if the PDF of the field $\mathrm{F}(\mathbf{x})$ is a multi-variate Gaussian, then the PDF of $\mathrm{F}(\mathbf{k})$ is also a multi-variate Gaussian. However, there is an extra advantage in Fourier space than in real space. Due to translational invariance, in Fourier space, the different $\mathbf{k}$ modes are uncorrelated (see discussion in the previous paragraph). This means that for a Gaussian random field, the Fourier modes are *statistically independent* of each other (this is only for the Gaussian case) (Mo et al., 2010).

2.2 Matter density field

In section 1.3 we have sketched the idea how the initial matter density perturbations δ evolve into the late-time density fluctuations field of large scale structures. In this section we will briefly discuss the equations which govern the evolution of the perturbations in the linear regime and how that results in observables of statistical measures of the density field, namely the power spectrum (or the second connected moment of the density field). We will then chalk out in a perturbation theory picture, how nonlinear evolution of the initial Gaussian distributed perturbations results in the generation of higher-order moments such as the bispectrum (third connected moment of the field), which is a key signature of the non-Gaussianity of the evolved field and is central to the ideas presented later on in the thesis. We will then discuss how one can characterise the 2-point correlation or the power spectrum of these density fluctuations (section 2.2) and project the fluctuations onto the celestial sphere (section 2.4). In subsection 2.4.1 we will briefly discuss the gravitational lensing convergence/shear fields and their associated power spectra.

2.2.1 Dynamics of gravitational instability

In section 1.1 we talked about a homogeneous and isotropic universe which can be described using the FLRW metric. To consider how initial inhomogeneities δ embedded within the pressure-free matter density field (with pressure $p(\mathbf{r}, t) = 0$) in the early universe evolved

into the large inhomogeneous structures that we see today, we need to consider perturbations to the FLRW metric and in the energy-momentum tensor of a perfect fluid. Putting them into the Einstein equations (1.6) we can get the equations describing the evolution of these fluctuations against an expanding homogeneous and isotropic background Universe. The interested reader can find the general relativistic derivation of the equations for the evolution of the perturbations from most graduate level cosmology textbooks Mukhanov (2005), Dodelson & Schmidt (2020). Here we will simply motivate the use of the equations.

At scales much smaller than the curvature radius of the universe (the Hubble radius) and in the regime that local gravitational potentials are small, general relativity effects can be considered negligible. Hence, using Newtonian theory, one can derive equations governing the evolution of perturbations in the density $\rho(\mathbf{r}, t)$ and velocity $\mathbf{v}(\mathbf{r}, t)$ of the cold-dark matter component (under the approximation that it is a non-relativistic collisionless fluid) under the influence of a gravitational field with potential $\phi(\mathbf{r}, t)$ where t is the time coordinate and $\mathbf{r}$ is the physical spatial coordinate which is related to the comoving coordinate $\mathbf{x}$ via $\mathbf{r}(t) = a(t)\mathbf{x}$. To achieve this, similar to the density contrast δ being the perturbation component to the density field ρ, we also need to identify the perturbation components $\mathbf{u}$ and Φ to the velocity and the gravitational potential of the fluid, respectively. Working in comoving frame $\mathbf{x}$ and expressing the time in conformal coordinate τ such that $\mathrm{d}t = a(\tau)\mathrm{d}\tau$, we can define these various perturbation components $\delta, \mathbf{u}, \Phi$ as (Bernardeau et al., 2002):

$$\rho(\mathbf{x}, \tau) \equiv \bar{\rho}(\tau) \left[1 + \delta(\mathbf{x}, \tau)\right] , \tag{2.23}$$

$$\mathbf{v}(\mathbf{x}, \tau) \equiv \mathcal{H}\mathbf{x} + \mathbf{u}(\mathbf{x}, \tau) , \tag{2.24}$$

$$\phi(\mathbf{x}, \tau) \equiv -\frac{1}{2}\frac{\partial \mathcal{H}}{\partial \tau}x^2 + \Phi(\mathbf{x}, \tau) , \tag{2.25}$$

where $\mathcal{H} = aH$ is the *conformal* Hubble parameter. In the expanding Universe this gravitational potential perturbation Φ is sourced by the density contrast δ and is therefore described by the *Poisson equation*:

$$\nabla^2 \Phi(\mathbf{x}, \tau) = 4\pi G a^2(\tau)\bar{\rho}(\tau)\delta(\mathbf{x}, \tau) \qquad (\text{Poisson}) . \tag{2.26}$$

Subsequently, the description of the evolution of δ and $\mathbf{u}$ can be derived from the *Vlasov equation* which describes the time evolution of the distribution function $f(\mathbf{x}, \mathbf{u}, \tau)$ of the fluid in phase space. This equation (also called the *collisionless Boltzmann equation*) states that under the conservation of particle number density as well as the phase-space volume of the fluid (Liouville's theorem), the total time derivative of $\mathrm{d}f/\mathrm{d}\tau$ vanishes:

$$\frac{\mathrm{d}f}{\mathrm{d}\tau} = \frac{\partial f}{\partial \tau} + \mathbf{u} \cdot \nabla f - \nabla \Phi \cdot \frac{\partial f}{\partial \mathbf{u}} = 0 \qquad (\text{Vlasov}) . \tag{2.27}$$

Equations (2.27) and (2.26) together form a system called the *Vlasov-Poisson equation* which are the main equations from which all calculations of gravitational instability can be

derived. Taking the first and second momentum moments (i.e. integrating out the particle conjugate momentum $\mathbf{p} = am\mathbf{u}$) of the Vlasov equation, we can derive the *continuity* (describing mass conservation) and *Euler* (describing conservation of momentum) equations:

$$\frac{\partial \delta}{\partial \tau} + \nabla \cdot ([1 + \delta]\mathbf{u}) = 0 \qquad \text{(continuity)} \tag{2.28}$$

$$\frac{\partial \mathbf{u}}{\partial \tau} + \mathcal{H}\mathbf{v} + (\mathbf{v} \cdot \nabla)\mathbf{u} + \nabla \Phi = 0 \qquad \text{(Euler)} \tag{2.29}$$

where in the Euler equation we have assumed that there is no contribution from the velocity dispersion of the dark-matter fluid (or from anisotropic pressure) which is a good enough approximation in the first stages of gravitational instability when structures did not have time to collapse and virialize. For a given cosmology (e.g. in the context of the ΛCDM model, the cosmological parameters such as Ω_i and equation of state specify the evolution of a, $\mathcal{H}$), one can in principle solve this set of nonlinear equations which describe the evolution of dark matter perturbations through gravitational instability.

2.2.2 Standard perturbation theory

However, these equations are actually difficult to solve. One has to rely on approximations such as *linearising* the equations or considering highly symmetric perturbations (e.g. spherical collapse) in order to find solutions. Here we first discuss the linearised version of the equations by considering the perturbations to be very small ($|\delta| \ll 1$) and neglecting all terms which are of second order in the perturbations. Combining the above equations under such considerations gives us the following *linear* second-order differential equation for δ which we denote for clarity as δ_L:

$$\frac{\partial^2 \delta_L}{\partial \tau^2} + \mathcal{H}\frac{\partial \delta_L}{\partial \tau} - 4\pi G a^2 \bar{\rho}\delta_L = 0 \ . \tag{2.30}$$

The solution of δ_L in this equation has a growing and a decaying mode. Assuming that the perturbation observed at location $\mathbf{x}$ and at conformal time τ is influenced by the growing mode only (as the decaying mode vanishes quickly; Bernardeau et al. (2002)), the growing solution of this equation can be written in terms of $D(\tau)$ (called the linear *growth factor* which solves the above equation and relates how the matter density perturbation $\delta_L(\mathbf{x}, \tau)$ evolves due to linear growth over time):

$$\delta_L(\mathbf{x}, \tau) = \frac{D(\tau)}{D(\tau_i)}\delta(\mathbf{x}, \tau_i) \tag{2.31}$$

where τ_i is an initial conformal time and $\delta(\mathbf{x}, \tau_i)$ is the initial density perturbation. This equation leaves apparent that when the initial density contrast $\delta_i \equiv \delta(\mathbf{x}, \tau_i)$ is Gaussian distributed, linear evolution will preserve the Gaussian shape of the distribution.

Using the definition of the 2-point correlation function (see equation (2.4)) (of a statis-

tically homogeneous and isotropic field)

$$\xi_\delta(|\Delta\mathbf{x}|, \tau) \equiv \langle \delta(\mathbf{x}_1, \tau)\delta(\mathbf{x}_2, \tau) \rangle \ , \tag{2.32}$$

where $\Delta\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$, and using only the leading-order (linear) PT computation of the density field $\delta^{(1)} \equiv \delta_L$, one finds that the corresponding 2-point correlation of the linear density field which consists of Fourier space correlations $\langle \delta^{(1)}\delta^{(1)} \rangle$, can be written as

$$\xi_L(|\Delta\mathbf{x}|, \tau) = \left(\frac{D(\tau)}{D(\tau_i)}\right)^2 \xi_\delta(|\Delta\mathbf{x}|, \tau_i) \ . \tag{2.33}$$

Correspondingly, the linearised power spectrum (see equation (2.19)) is given by

$$P_L(k, \tau) = \left(\frac{D(\tau)}{D(\tau_i)}\right)^2 P_\delta(k, \tau_i) \ . \tag{2.34}$$

According to the theory of inflation — which is the most popular understanding of the origin of density perturbations in our Universe — the primordial power spectrum $P_\delta(k, \tau_i)$ of the statistically homogeneous and isotropic matter density fluctuations field δ right after the very end of the inflationary phase of the Universe is predicted to be a power-law as a function of the amplitude of the Fourier modes k (Mukhanov, 2005; Dodelson & Schmidt, 2020; Peacock, 1998)

$$P_\delta(k, \tau_i) = A_\delta k^{n_s} \ , \tag{2.35}$$

where n_s is the so called spectral index (can be thought of as the tilt/slope of the power spectrum in a log-log plot of $\ln P_\delta(k, \tau_i)$ vs. $\ln k$) of the power spectrum and is predicted to be almost a constant with only small variation with scale. A_δ is the amplitude of this primordial matter power spectrum and it is often written in terms of the variance $\sigma_{8/h}^2$ of the density fluctuations at the current epoch (τ_0). This can be done by first defining the smoothed linear matter density contrast $\delta_{L(R)}$ in a spherical aperture of radius R

$$\delta_{L(R)}(\mathbf{x}, \tau) \equiv \frac{3}{4\pi R^3} \int\limits_{|\mathbf{y}_1 - \mathbf{x}| < R} d^3y_1 \ \delta_L(\mathbf{y}, \tau) \ . \tag{2.36}$$

Computing the variance $\sigma_{L(R)}^2(\tau) \equiv \langle \delta_{L(R)}(\mathbf{x}, \tau)^2 \rangle$ of this statistically homogeneous and smoothed linear matter density contrast field gives

$$\sigma_{L(R)}^2(\tau) = \left(\frac{3}{4\pi R^3}\right)^2 \int\limits_{|\mathbf{y}_1| < R} d^3y_1 \int\limits_{|\mathbf{y}_2| < R} d^3y_2 \ \xi_{L,\delta}(|\mathbf{y}_1 - \mathbf{y}_2|, \tau) \ . \tag{2.37}$$

The above equation can also be written in terms of the amplitude of the matter power spectrum $P_{L,\delta}(\mathbf{k}, \tau)$ (as it is related to the $\xi_{L,\delta}(\mathbf{y}_1 - \mathbf{y}_2, \tau)$ through a Fourier transform). Thus, the amplitude A_δ of the primordial power spectrum can be described in terms of

the variance $\sigma^2_{L(R)}(\tau)$ of the smoothed density fluctuations at a given conformal time τ. In LSS studies, it is common to define the quantity

$$\sigma^2_{8/h} \equiv \sigma^2_{L(8\ \text{Mpc}/h)}(\tau_0) \, , \tag{2.38}$$

which describes this variance of the linear density contrast in an 8 Mpc/h radius at the current epoch τ_0 (as this value is seen to be $\mathcal{O}(1)$ at τ_0).

In the above discussion we only studied the linear growth of the matter density perturbations through gravitational instability. However, in the early universe there were other processes such as the tight coupling between the photon-baryon plasma (pre-recombination) which also left its imprint on the power spectrum of matter density fluctuations. This is encoded by the so called linear *Transfer function* $T(k)$ which relates the amplitudes of the Fourier modes of the density fluctuations in the post-recombination era to the initial curvature perturbations in the spatial hypersurfaces (Mo et al., 2010). The transfer function can be modelled (e.g. Eisenstein & Hu (1998)) or computed accurately via Boltzmann codes such as `CLASS` (Blas et al., 2011). Including this effect of the Transfer function, equation (2.34) is modified to

$$P_{L,\delta}(k, \tau) = T(k)^2 \left(\frac{D(\tau)}{D(\tau_i)} \right)^2 P_\delta(k, \tau_i) \, . \tag{2.39}$$

This is the late-time linear matter density fluctuations power spectrum. However, the linearised evolution of density perturbations is not a realistic approximation as it greatly underestimates the growth of structures on small-scales where the amplitude of the density perturbations $|\delta| \geq 1$. In such a scenario one has to solve the set of nonlinear equations. Theoretically, this is a challenging task and one can use e.g. *standard perturbation theory* (SPT) techniques to solve these equations *perturbatively* i.e. in increasingly higher order of the linear density contrast (Bernardeau et al., 2002). In SPT, one assumes that the density field (and also velocity field but not shown here) can be expanded as a series about the linear solutions when $\delta \ll 1$:

$$\delta(\mathbf{x}, \tau) = \sum_{i=1}^{n} \delta^{(n)}(\mathbf{x}, \tau) \, , \tag{2.40}$$

where $\delta^{(1)} = \delta_L$ is linear in the initial density field, $\delta^{(2)}$ is quadratic in the initial density field and so on. By plugging in the solution of a lower order term (e.g. $\delta^{(1)}$) into the continuity, Euler and Poisson equations allows one to reiteratively solve for the next higher-order term (e.g. $\delta^{(2)}$) in this series expansion. Working in Fourier space

$$\delta(\mathbf{k}, \tau) = \int \mathrm{d}^3\mathbf{x} \, \delta(\mathbf{x}, \tau) e^{-i\mathbf{k}\cdot\mathbf{x}} \, , \tag{2.41}$$

for example, one finds that the second-order PT solution to the density field is given by (suppressing the time component for clarity of notation) Dodelson & Schmidt (2020):

$$\delta^{(2)}(\mathbf{k}) = \int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} \, F_2(\mathbf{q}, \mathbf{k} - \mathbf{q}) \, \delta^{(1)}(\mathbf{q})\delta^{(1)}(\mathbf{k} - \mathbf{q}), \tag{2.42}$$

where[4]

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{2}{7}\frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} + \frac{1}{2}\,\mathbf{k}_1 \cdot \mathbf{k}_2\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right)\;. \tag{2.43}$$

The above equation states that the evolution of $\delta^{(2)}(\mathbf{k})$ is determined by the mode coupling, through the kernel F_2, of the $\delta^{(1)}$ fields at all pairs of wavevectors $\mathbf{q}$ and $\mathbf{k} - \mathbf{q}$ whose sum is $\mathbf{k}$. In such a perturbative manner one can compute further higher-order terms in the expansion of δ with corresponding mode-coupling kernels for the corresponding order in perturbation theory. The real matter density fluctuations power spectrum $P_\delta(k, \tau)$ is therefore a modified form of equation (2.44) after including the non-linear corrections Dodelson & Schmidt (2020):

$$P_\delta(k, \tau) = P_L(k, \tau) + P_{NLO}(k, \tau) + \dots\,, \tag{2.44}$$

where P_{NLO} is the next-to-leading order correction (also called *1-loop terms*) to the power spectrum by including the next higher-order terms in the perturbative expansion of δ i.e. Fourier space correlations $\langle \delta^{(1)}\delta^{(3)}\rangle$ and $\langle \delta^{(2)}\delta^{(2)}\rangle$. Further higher-order terms can be added by including more higher-order perturbation term correlations. However, this becomes increasingly more complex. Hence, in order to obtain the full nonlinear cold dark matter power spectrum one calibrates accurate fitting formula to measurements of the power spectrum from high-resolution N-body simulations (e.g. `halofit` by Takahashi et al. (2012)).

A consequence of the nonlinear nature of the evolution equations is that even if we start with a Gaussian distribution of δ at an early time, the nonlinear evolution will distort the distribution away from Gaussian at a later time. This gives rise to the non-Gaussian distribution of the late-time density fluctuation field of the LSS and hence for example, a non-zero signature for the connected part of the three-point correlation function of the density field (see Eq. 2.7) and equivalently its Fourier space counterpart, the bispectrum. The bispectrum for the matter density field is defined as (suppressing the time component):

$$\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle = (2\pi)^3 \delta^D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)B_\delta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)\,, \tag{2.45}$$

where we have assumed that the field is homogeneous i.e. the bispectrum is non-zero only for closed triangles of vectors $\mathbf{k}_1$, $\mathbf{k}_2$ and $\mathbf{k}_3$. At leading-order in nonlinear perturbation theory (also called *tree-level)*, the bispectrum receives contributions from Fourier space correlations $\langle \delta^{(1)}\delta^{(1)}\delta^{(2)}\rangle$ terms which give:

$$B_{tree}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2F_2(\mathbf{k}_1, \mathbf{k}_2)P_L(k_1)P_L(k_2) + \text{cycl. permutations}\,. \tag{2.46}$$

As discussed for the power spectrum, one can do the same and compute higher-order perturbative correction to the tree-level bispectrum. The full nonlinear matter bispectrum

---

[4]Note that this particular form of the F_2 kernel is only for an Einstein-de Sitter universe (completely matter dominated universe with $\Omega_m = 1$). However, the dependence of F_2 on cosmological parameters is

can also be obtained using accurate fitting formula from N-body simulations as has been recently done using the `bihalofit` fitting formula by Takahashi et al. (2020).

Since the late-time matter density contrast field is non-Gaussian distributed, the power spectrum $P_\delta(k)$ does not contain full information about the nonlinear evolution of the matter density field. Hence, to extract further information about the field (especially about the non-linearities in either Fourier or real space, one needs to look at higher order correlations such as the 3-point correlation function or the bispectrum (as well as other higher-order N-point correlation functions).

2.3   Galaxy density field

The 3D matter density fluctuations $\delta(\mathbf{x}, \tau)$ cannot be directly observed and we can only probe it through tracers such as the galaxy density field (which is a biased tracer) or the gravitational lensing cosmic shear field (see Sec. 2.4.1). Here, we discuss the galaxy density contrast field δ_g^{3D} (we specifically denote the superscript '3D' here to distinguish it from the projected galaxy density contrast that we shall discuss later in section 2.4.2) and also introduce the so called *galaxy bias* parameters.

The comoving number density [units of Mpc^{-3}] of galaxies[5] $n_g^{3D}(\boldsymbol{x}, \tau)$ at comoving position $\boldsymbol{x}$ and temporal epoch τ can be written in terms of the number density contrast of galaxies as $\delta_g^{3D}(\boldsymbol{x}, \tau)$ as

$$n_g^{3D}(\boldsymbol{x}, \tau) = \bar{n}_g^{3D}(\tau)[1 + \delta_g^{3D}(\boldsymbol{x}, \tau)] \ , \tag{2.47}$$

where $\bar{n}_g^{3D}(\tau)$ is the cosmic mean comoving number density of galaxies at time τ. Note that unlike the mean comoving density of matter in the Universe $\bar{\rho}_m^{3D}$ which is independent of time, the $\bar{n}_g^{3D}(\tau)$ is not constant since the number of galaxies (which live inside collapsed halos) within a given comoving volume evolves with time.

A better way of expressing the local number density (in a deterministic manner) is by Taylor expanding it about the cosmic mean density with respect to some specific operator $\mathcal{O}$ each of which can involve derivatives of the local gravitational potential Φ (e.g. $\mathcal{O}$ can be the 3D matter density contrast field $\delta \equiv \delta_m^{3D}$, tidal field K_{ij}, their time derivatives, higher-order derivatives of the gravitational potential, their powers and combinations etc.

---

very weak, making the EdS form an extremely good approximation in other cosmologies as well.

[5]These equations hold for any tracer of the dark matter field e.g. halos and not only galaxies. For tracers other than galaxies we will always explicitly specify the tracer subscript for all relevant quantities (number density, bias parameters etc.). However, for galaxies we will sometimes drop the corresponding subscript as it is the general tracer of interest for us in this thesis.

i.e. $\mathcal{O} \in \{\delta, \delta^2, ..., \delta^n, K_{ij}K^{ij}, ...\}$), as follows:

$$
n_g^{3D}(\boldsymbol{x}, \tau) = \bar{n}_g^{3D}(\tau) + \sum_{\mathcal{O}} \left.\frac{\partial\, \bar{n}_g^{3D}(\tau)}{\partial\, \mathcal{O}}\right|_{\mathcal{O}=0} \mathcal{O}(\boldsymbol{x}, \tau)
$$

$$
= \bar{n}_g^{3D}(\tau) \left[1 + \sum_{\mathcal{O}} \underbrace{\left.\frac{\partial \ln\, \bar{n}_g^{3D}(\tau)}{\partial\, \mathcal{O}}\right|_{\mathcal{O}=0}}_{\equiv b_{\mathcal{O}}(\tau)} \mathcal{O}(\boldsymbol{x}, \tau) \right] .
$$

Therefore, the smoothed[6] local galaxy number density contrast $\delta_g^{3D}(\boldsymbol{x}, \tau)$ can be expressed using a series expansion of these operators $\mathcal{O}(\boldsymbol{x}, \tau)$ with accompanying *bias* coefficients $b_{\mathcal{O}}$ of these operators (Desjacques et al., 2018):

$$
\delta_g^{3D}(\boldsymbol{x}, \tau) = \sum_{\mathcal{O}} b_{\mathcal{O}}(\tau)\, \mathcal{O}(\boldsymbol{x}, \tau)\; + \left[\epsilon(\boldsymbol{x}, \tau) + \sum_{\mathcal{O}} \epsilon_{\mathcal{O}}(\boldsymbol{x}, \tau)\, \mathcal{O}(\boldsymbol{x}, \tau) \right] . \tag{2.48}
$$

Physically, the bias parameters $b_{\mathcal{O}}(\tau)$ can be interpreted as the response of the local number density of galaxies $n_g^{3D}(\boldsymbol{x}, \tau)$ to changes in the amplitude of the operators $\mathcal{O}(\boldsymbol{x}, \tau)$. Practically, these bias parameters absorb the complicated details of galaxy and halo formation and evolution and describe the connection between the underlying dark matter density field and the observed galaxy field (Desjacques et al., 2018; Barreira et al., 2021). The terms inside the square brackets denote the non-deterministic (stochastic) part of the relation which arise due to perturbations on scales smaller than the smoothing scale of the galaxy halo and the underlying $\mathcal{O}$ fields. Similar to the deterministic bias relation, the stochastic relation comes with its own free parameters $\epsilon_{\mathcal{O}}$ and an offset term ϵ. Note that these operators $\mathcal{O}$ are the so-called *renormalized* operators and hence the associated bias (and stochastic) parameters are the *physical, renormalized* parameters which are the bias coefficients one would associate with an observed tracer sample (Desjacques et al., 2018).

2.3.1 Galaxy bias in the Halo Occupation Distribution framework

A widely used framework to obtain analytical expressions for these galaxy bias parameters are through the Halo Model approach. In short, the halo model assumes that all the matter in the Universe resides inside dark matter halos and galaxies are hosted inside those halos. In this framework, the effective bias of halos at redshift z (corresponding to comoving distance χ) in a halo mass bin $[M_{h,min}, M_{h,max}]$ is given by

$$
b_{\mathcal{O},h}(\chi) = \frac{1}{\bar{n}_h^{3D}(\chi)} \int_{M_{h,min}}^{M_{h,max}} \mathrm{d}M\, \frac{\mathrm{d}\bar{n}_h^{3D}}{\mathrm{d}M_h}(M_h, \chi)\, b_{\mathcal{O},h}(M_h, \chi)\, , \tag{2.49}
$$

---

[6]over some length scale e.g. typical Lagrangian radius of a galaxy halo.

where $\mathrm{d}\bar{n}_h^{\mathrm{3D}}/\mathrm{d}M_h$ is the global halo mass function (3D number density of dark matter halos in an infinitesimal mass bin $\mathrm{d}M_h$ around halos of mass M_h), $b_{\mathcal{O},h}(M_h, \chi)$ is the bias parameter of dark matter halos of mass M_h, and $\bar{n}_h^{\mathrm{3D}}$ is the 3D global number density of halos inside the mass bin at redshift z:

$$\bar{n}_h^{\mathrm{3D}}(\chi) = \int_{M_{h,min}}^{M_{h,max}} \mathrm{d}M_h \ \frac{\mathrm{d}\bar{n}_h^{\mathrm{3D}}}{\mathrm{d}M}(M_h, \chi) \tag{2.50}$$

Now, in order to get analytical expressions for the galaxy bias parameters $b_{\mathcal{O}}(\chi) \equiv b_{\mathcal{O},g}(\chi)$ one uses the widely used Halo Occupation Distribution (HOD) models (Berlind & Weinberg, 2002) which specify the expected number of galaxies $\bar{N}_g(M_h, \chi)$ residing inside dark matter halos of mass M_h at redshift z. Briefly, an HOD model describes a probability distribution $P(N_g|M_h)$ for a halo mass M_h to host N_g galaxies. Concretely, the HOD model separates the contribution from central and satellite galaxies, and has the following functional forms for the mean values of the central and satellite galaxies and therefore the total mean number of galaxies hosted by halos of mass M_h is

$$\langle N_g | M_h \rangle = \langle N_{cen} | M_h \rangle + \langle N_{sat} | M_h \rangle. \tag{2.51}$$

Hence, in this framework the effective 3D global number density of galaxies at redshift z which are hosted in halos of mass $M_h \in [M_{h,min}, M_{h,max}]$ is given by:

$$\bar{n}_g^{\mathrm{3D}}(\chi) = \int_{M_{h,min}}^{M_{h,max}} \mathrm{d}M_h \ \frac{\mathrm{d}\bar{n}_h^{\mathrm{3D}}}{\mathrm{d}M}(M_h, \chi) \ \bar{N}_g(M_h, \chi) \ . \tag{2.52}$$

The galaxy bias parameters are in turn expressed as:

$$b_{\mathcal{O}}(\chi) = \frac{1}{\bar{n}_g^{\mathrm{3D}}(\chi)} \int_{M_{h,min}}^{M_{h,max}} \mathrm{d}M \ \frac{\mathrm{d}\bar{n}_h^{\mathrm{3D}}}{\mathrm{d}M_h}(M_h, \chi) \ \bar{N}_g(M_h, \chi) \left[b_{\mathcal{O},h}(M_h, \chi) + R_{\mathcal{O},N_g}(M_h, \chi) \right] ,$$
$$\tag{2.53}$$

where $R_{\mathcal{O},N_g}(M_h, \chi)$ is called the response function of the expected number of HOD galaxies (hosted in a halo of mass M_h) in the presence of a large scale perturbation $\mathcal{O}$. The response term is often not considered in literature but it has been shown by Voivodic & Barreira (2021) that it can indeed be substantial for specific perturbations.

2.4   Fields in projection

The discussion so far has been restricted to 3D quantities such as the matter density fluctuations $\delta(\mathbf{x}, \tau)$ where $\mathbf{x}$ is the comoving spatial coordinate and τ the conformal time coordinate. The 3D 2-point function $\xi_{L,\delta}(|\Delta\mathbf{x}|, \tau)$ and the corresponding power spectrum $P_{L,\delta}(k, \tau)$ have been defined accordingly. However, due to many reasons, these 3D quantities are not directly observable (Friedrich, 2018):

1. We can only access those values of $\delta(\mathbf{x}, \tau)$ for which the position $\mathbf{x}$ and conformal time τ lie on our past light-cone. In other words, the observed objects which are located at different *comoving* distances (see equation (2.54)) from us do not possess the same conformal time coordinate τ.

2. We cannot get exact position information of the observed tracer objects from current wide-area photometric surveys (e.g. DES) and can only infer their distribution along our line-of-sight direction $\hat{\mathbf{n}}$.

Hence, what we actually observe in the night sky are 2D projections of the 3D density fields (through galaxies or other tracers) onto the celestial unit sphere. To formalise this we can write any cosmic field $f^{\mathrm{3D}}[\boldsymbol{\chi}, \tau]$ that we observe on our past light-cone at 3D comoving position $\boldsymbol{\chi}$ and corresponding conformal lookback time[7] $\tau = \tau_0 - \chi$ (where τ_0 is the conformal time today and χ the radial comoving distance), can be projected onto the 2D celestial sphere to obtain the weighted line-of-sight 2D quantity $f(\hat{\mathbf{n}})$ towards a radial unit direction $\hat{\mathbf{n}}$ (Bartelmann & Schneider, 2001)

$$f(\hat{\mathbf{n}}) = \int \mathrm{d}\chi \, q_{\mathrm{f}}(\chi) f^{\mathrm{3D}}\left[\chi\hat{\mathbf{n}}, \tau_0 - \chi\right] , \tag{2.55}$$

where $q_{\mathrm{f}}(\chi)$ is a particular weighting kernel over which f^{3D} is projected. Examples are the projected galaxy number density or the weak lensing convergence field that we observe on the celestial sphere. Assuming that the angular extent of the field of view is small — spanning an area of a few square degrees — we can make the flat-sky approximation, where we denote the position on the sky as a 2D planar vector $\boldsymbol{\theta} = (\theta_x, \theta_y)$ and express f as

$$f(\boldsymbol{\theta}) = \int \mathrm{d}\chi \, q_{\mathrm{f}}(\chi) f^{\mathrm{3D}}\left[(\chi\boldsymbol{\theta}, \chi), \tau_0 - \chi\right] . \tag{2.56}$$

2.4.1 Weak gravitational lensing fields

One way to probe the δ field is through gravitational lensing. Gravitational lensing is the bending of light rays coming from background objects (sources e.g. galaxies) by the gravitational potential of foreground objects (lenses e.g. galaxy cluster) resulting in our observation of shifted, magnified and distorted images of the sources. When the gravitational potential is very strong (or the light ray of a single source object passes very close to a single massive lens), the bending of the light rays by the lens can result in multiple images of the source object. This is known as *strong gravitational lensing*. However, in most cases (farther away from the centers of galaxy clusters where the gravitational potential is weak), a light ray from a background source only suffers slight deflections in its trajectory by many

---

[7]The conformal lookback time is related to the comoving distance via (using speed of light $c = 1$) (Hogg, 1999):

$$\chi = \int_t^{t_0} \frac{\mathrm{d}t}{a(t)} = \int_\tau^{\tau_0} \mathrm{d}\tau = \tau_0 - \tau . \tag{2.54}$$

foreground lenses while on its way to us. Hence, we do not observe multiple images but only a miniscule distortion in the image of the source. This is usually the case when light from sources pass through the foreground matter distribution of the LSS. The distortion is so small that one can only see this effect on a statistical basis through correlations of the alignment of the weakly but coherently distorted images of many background source galaxies. This is known as *weak gravitational lensing* and serves as a probe for investigating the distribution of matter in the LSS which causes the distortion of the images. This field can be interpreted as the *shear* caused by a weighted line-of-sight projection of the 3D matter density field — known as the weak lensing convergence field. Following equation (2.56), the weak lensing convergence field $\kappa(\boldsymbol{\theta})$ acting on source galaxies situated at the radial comoving distance χ_s can be written as a line-of-sight projection of the 3D matter density contrast field δ^{3D}

$$\kappa(\boldsymbol{\theta}) = \int \mathrm{d}\chi \, q(\chi) \delta^{3D} \big[(\chi\boldsymbol{\theta}, \chi), \tau_0 - \chi \big] \;, \tag{2.57}$$

with projection kernel $q(\chi)$ (also known as lensing efficiency) for the case when all source galaxies are located in a Dirac-δ function like tomographic bin at χ_s:

$$q(\chi) = \frac{3 H_0^2 \Omega_{\mathrm{m},0}}{2c^2} \frac{\chi}{a(\chi)} \frac{\chi_s - \chi}{\chi_s} \;; \qquad \text{with } \chi \leq \chi_s \;. \tag{2.58}$$

However, it is straight forward to write $q(\chi)$ for an ensemble of sources (instead of a single source) distribution of source galaxies in a tomographic redshift bin following a normalized distribution $p(\chi')$ (e.g. see Schneider, 2006; Kilbinger, 2015:

$$q_\kappa(\chi) = \frac{3 H_0^2 \Omega_{\mathrm{m},0}}{2c^2} \frac{\chi}{a(\chi)} \int_\chi^{\chi_{\mathrm{lim}}} \mathrm{d}\chi' p(\chi') \frac{\chi' - \chi}{\chi'} \;. \tag{2.59}$$

Note that q_κ has units of $[\mathrm{Mpc}^{-1}]$. In the equations above $\Omega_{\mathrm{m},0}$ is the total matter density parameter of the Universe today, H_0 the Hubble parameter today, a the scale factor and c the speed of light.

The convergence and the associated complex shear field are related to each other through second-order derivatives of the lensing potential $\psi(\boldsymbol{\theta})$ in the 2D sky-plane (Schneider, 2006):

$$\kappa(\boldsymbol{\theta}) = \frac{1}{2} \left(\partial_x^2 + \partial_y^2 \right) \psi(\boldsymbol{\theta}), \qquad \gamma(\boldsymbol{\theta}) = \frac{1}{2} \left(\partial_x^2 - \partial_y^2 + 2i\partial_x\partial_y \right) \psi(\boldsymbol{\theta}) \tag{2.60}$$

where $\psi(\boldsymbol{\theta})$ is the line-of-sight projection of the 3D Newtonian gravitational potential $\Phi \big[(\chi\boldsymbol{\theta}, \chi), \tau_0 - \chi \big]$ of the total matter distribution:

$$\psi(\boldsymbol{\theta}) = \frac{2}{c^2} \int \mathrm{d}\chi \, \frac{\chi_s - \chi}{\chi_s \, \chi} \, \Phi \big[(\chi\boldsymbol{\theta}, \chi), \tau_0 - \chi \big] \;; \qquad \text{with } \chi_s > \chi \;. \tag{2.61}$$

The shear $\gamma(\boldsymbol{\theta}) = \gamma_1(\boldsymbol{\theta}) + i\gamma_2(\boldsymbol{\theta})$ at a given location $\boldsymbol{\theta}$ is a complex quantity where the shear components γ_1 and γ_2 are specified in a chosen Cartesian frame (in 2D flat-sky). However, one is free to rotate the coordinates by any arbitrary angle β. With respect to this reference rotation angle β, one defines the rotated shear components of the shear at position $\boldsymbol{\theta}$ as (Schneider, 2006) With respect to this reference rotation angle β, one defines the rotated shear

$$\gamma_\beta(\boldsymbol{\theta}) \equiv -e^{-2i\beta}\gamma(\boldsymbol{\theta})$$
$$= -e^{-2i\beta}\big[\gamma_1(\boldsymbol{\theta}) + i\gamma_2(\boldsymbol{\theta})\big] \ .$$

In particular, for given a pair of points $\boldsymbol{\theta}$ and $\boldsymbol{\vartheta}$ on the field which are separated by the 2D vector $\boldsymbol{\alpha} \equiv \boldsymbol{\vartheta} - \boldsymbol{\theta}$, one can therefore write the *tangential* (+) and *cross* (×) components of the shear for this particular pair of points along the separation direction $\beta = \phi_{\boldsymbol{\alpha}}$ (polar angle of $\boldsymbol{\alpha}$). For example, the tangential and cross components of the shear at location $\boldsymbol{\vartheta}$ (and similarly for $\boldsymbol{\theta}$) with respect to direction $\phi_{\boldsymbol{\alpha}}$ are then expressed as:

$$\gamma_{\phi_{\boldsymbol{\alpha}}}(\boldsymbol{\vartheta}) \equiv \gamma_+(\boldsymbol{\vartheta}, \phi_{\boldsymbol{\alpha}}) + i\gamma_\times(\boldsymbol{\vartheta}, \phi_{\boldsymbol{\alpha}}) \equiv -e^{-2i\phi_{\boldsymbol{\alpha}}}\big[\gamma_1(\boldsymbol{\vartheta}) + i\gamma_2(\boldsymbol{\vartheta})\big] \ . \tag{2.62}$$

In the 2D Fourier plane, the shear $\gamma(\boldsymbol{\ell})$ is related to $\kappa(\boldsymbol{\ell})$ as (Schneider, 2006; Kilbinger, 2015)

$$\gamma(\boldsymbol{\ell}) \ = \ \frac{(\ell_x + i\,\ell_y)^2}{\ell^2}\,\kappa(\boldsymbol{\ell}) \ = \ e^{2i\phi_{\boldsymbol{\ell}}}\kappa(\boldsymbol{\ell}) \ ; \qquad \text{for } \ell \neq 0 \tag{2.63}$$

where $\ell = \sqrt{\ell_x^2 + \ell_y^2}$ and $\phi_{\boldsymbol{\ell}} = \arctan\left(\frac{\ell_y}{\ell_x}\right)$ is the polar angle of $\boldsymbol{\ell}$.

In the weak lensing regime, a source object is a distant galaxy whose image is distorted by the foreground matter density fluctuations. What we observe directly through galaxy imaging photometric surveys such as Dark Energy Survey (DES) is the shape of each galaxy, characterised by its ellipticity. As mentioned previously, the shear field distorts the images (ellipticities) of the background source galaxy and therefore studying the observed ellipticities gives us a gateway to estimate the gravitational shear field. To estimate the ellipticity from the image of a galaxy one often uses the moments of the brightness distribution of the galaxy (Seitz & Schneider, 1996; Bartelmann & Schneider, 2001). Besides the induced ellipticity on the shape of the source galaxy by the shear field, the unlensed galaxy also has its own *intrinsic ellipticity* ie.. its original shape projected in the sky. Assuming there are no additional distortion from observational systematic errors, the relation between the observed ellipticity e^{obs} and the intrinsic ellipticity e^{int} of a galaxy is given by (Schneider, 2006)

$$e^{\text{obs}} = \frac{e^{\text{int}} + \gamma}{1 + \gamma^* e^{\text{int}}} \ , \tag{2.64}$$

where γ^* represents the complex conjugate of the shear field. The ellipticity itself depends on how observations measure the light distribution of a galaxy and there can be various definitions (see Schneider (2006)). It is infeasible in the weak lensing regime to probe the shear

value at a point in the field by measuring the ellipticity of just one neighbouring galaxy since both observed. However, based on one crucial assumption that unlensed galaxies are oriented randomly (Bartelmann & Schneider, 2001), measuring a large number of galaxies around that field point and taking the ensemble average of the observed ellipticities, the ensemble average of intrinsic ellipticities would vanish due to their random orientations and we are left with $\langle e^{\text{obs}} \rangle \approx \gamma$. This is only an approximation and what we actually measure from ellipticities is a quantity called *reduced shear* (Bartelmann & Schneider, 2001):

$$g = \frac{\gamma}{1 - \kappa} \ , \tag{2.65}$$

which in the limit of weak lensing regime would become $g \approx \gamma$ since $\kappa \ll 1$. One should also note here that the shear measurement procedure from galaxy images may suffer from calibration biases (Troxel et al., 2018) which need to be carefully accounted for in the form of nuisance systematic parameters so as to obtain unbiased cosmological results.

Furthermore, the assumption that the intrinsic ellipticities of the source galaxies is random is of course not true since neighbouring galaxies which live inside a common gravitational potential well would have their orientations aligned with the gradient of the local tidal field of the potential which would in turn lead to correlations already between the intrinsic shapes and orientations of the galaxies before their shapes are lensed. Accounting for this *intrinsic alignment* of the galaxy shapes is an area of active research in weak lensing research as it is of paramount importance to isolate this effect of intrinsic alignment from the weak lensing signal (Troxel & Ishak, 2015).

2.4.2 Projected galaxy density field

Another projected 2D field is the observed average number counts of galaxies $\bar{N}$ projected across a redshift range, characterised by an interval in comoving distance $[\chi_{lower}, \chi_{upper}]$ (we suppress the limits in the following expressions for clarity)

$$\bar{N} \equiv \int \mathrm{d}\chi \ \frac{\mathrm{d}V}{\mathrm{d}\chi} \ \bar{n}_g^{\text{3D}}(\tau) \ , \tag{2.66}$$

where $\frac{\mathrm{d}V}{\mathrm{d}\chi}(\chi)$ is the cosmological volume element, which for the whole spherical sky reads $\frac{\mathrm{d}V}{\mathrm{d}\chi}(\chi) = 4\pi\chi^2$. Following equation (2.56), the projected 2D number counts field of galaxies

$N(\boldsymbol{\theta})$ can be written as a line-of-sight projection along comoving radial coordinate χ:

$$
\begin{aligned}
N(\boldsymbol{\theta}) &= \int \mathrm{d}\chi \, \frac{\mathrm{d}V}{\mathrm{d}\chi} \, n_g^{3\mathrm{D}}(\boldsymbol{x}, \tau) \\
&= \int \mathrm{d}\chi \, \frac{\mathrm{d}V}{\mathrm{d}\chi} \, \bar{n}_g^{3\mathrm{D}}(\tau)[1 + \delta_g^{3\mathrm{D}}(\boldsymbol{x}, \tau)] \\
&= \int \mathrm{d}\chi \, \frac{\mathrm{d}V}{\mathrm{d}\chi} \, \bar{n}_g^{3\mathrm{D}}(\tau) + \int \mathrm{d}\chi \, \frac{\mathrm{d}V}{\mathrm{d}\chi} \, \bar{n}_g^{3\mathrm{D}}(\tau) \, \delta_g^{3\mathrm{D}}(\boldsymbol{x}, \tau) \\
&= \bar{N} \left[1 + \frac{1}{\bar{N}} \int \mathrm{d}\chi \, \frac{\mathrm{d}V}{\mathrm{d}\chi} \, \bar{n}_g^{3\mathrm{D}}(\tau) \, \delta_g^{3\mathrm{D}}(\boldsymbol{x}, \tau) \right] \\
&\equiv \bar{N} \left[1 + \delta_g^{2\mathrm{D}}(\boldsymbol{\theta}) \right] ,
\end{aligned}
\tag{2.67}
$$

where in the last line we have identified the projected galaxy number density contrast field $\delta_g^{2\mathrm{D}}(\boldsymbol{\theta})$ as:

$$
\begin{aligned}
\delta_g^{2\mathrm{D}}(\boldsymbol{\theta}) &= \frac{1}{\bar{N}} \int \mathrm{d}\chi \, \frac{\mathrm{d}V}{\mathrm{d}\chi} \, \bar{n}_g^{3\mathrm{D}}(\tau) \, \delta_g^{3\mathrm{D}}(\boldsymbol{x}, \tau) \\
&= \int \mathrm{d}\chi \, q_g(\chi) \, \delta_g^{3\mathrm{D}}(\boldsymbol{x}, \tau) .
\end{aligned}
\tag{2.68}
$$

with the galaxy projection kernel (units $[\mathrm{Mpc}^{-1}]$):

$$
q_g(\chi) = \frac{1}{\bar{N}} \frac{\mathrm{d}V}{\mathrm{d}\chi} \bar{n}_g^{3\mathrm{D}}(\tau) .
\tag{2.69}
$$

From (2.66) it implies that

$$
\int \mathrm{d}\chi \, q_g(\chi) = \frac{1}{\bar{N}} \int \mathrm{d}\chi \frac{\mathrm{d}V}{\mathrm{d}\chi} \bar{n}_g^{3\mathrm{D}}(\tau) = 1 ,
\tag{2.70}
$$

i.e., the $q_g(\chi)$ is a normalized projection kernel of galaxies and one may identify it as the observed distribution of galaxies $p(\chi)$ from a given LSS survey.

Explicit relation for the projected galaxy number density contrast field

As an example, we can write down a few terms of this $\delta_g^{2\mathrm{D}}(\boldsymbol{\theta})$ field for the following bias expansion[8] of the 3D galaxy density contrast field:

$$
\delta_g^{3\mathrm{D}}(\boldsymbol{x}, \tau) = b_\delta(\tau) \, \delta(\boldsymbol{x}, \tau) + b_{\delta^2}(\tau) \, \delta^2(\boldsymbol{x}, \tau) + b_{s^2}(\tau) \, \boldsymbol{K}^2(\boldsymbol{x}, \tau) + \left[\epsilon(\boldsymbol{x}, \tau) + \epsilon_\delta(\boldsymbol{x}, \tau) \, \delta(\boldsymbol{x}, \tau) \right] ,
\tag{2.71}
$$

where $\delta(\boldsymbol{x}, \tau) \equiv \delta_m^{3\mathrm{D}}(\boldsymbol{x}, \tau)$ is the 3D matter density contrast and $\boldsymbol{K}^2(\boldsymbol{x}, \tau) = K_{ij}(\boldsymbol{x}, \tau)K^{ij}(\boldsymbol{x}, \tau)$ is the tidal field squared. Note that we have kept the terms which would contribute if we

---

[8]in literature, one usually denotes the bias parameters as b_n which are related to b_{δ^n} via $b_{\delta^n} = \frac{b_n}{n!}$ such

perform an SPT like expansion up to second order in perturbations of the matter field i.e. $\delta_m^{3D} \equiv \delta = \delta^{(1)} + \delta^{(2)}$. Plugging this into the equation above for $\delta_g^{2D}(\boldsymbol{\theta})$ we get (this is analogous to equation (2.56)):

$$
\begin{aligned}
\delta_g^{2D}(\boldsymbol{\theta}) &= \int \mathrm{d}\chi \, q_g(\chi) \, \delta_g^{3D}(\boldsymbol{x}, \tau) \\
&= \int \mathrm{d}\chi \, q_g(\chi) \, b_\delta(\tau) \, \delta(\boldsymbol{x}, \tau) + \int \mathrm{d}\chi \, q_g(\chi) \, b_{\delta^2}(\tau) \, \delta^2(\boldsymbol{x}, \tau) + \int \mathrm{d}\chi \, q_g(\chi) \, b_{s^2}(\tau) \, \boldsymbol{K}^2(\boldsymbol{x}, \tau) + \\
&\quad + \left[\int \mathrm{d}\chi \, q_g(\chi) \, \epsilon(\boldsymbol{x}, \tau) + \int \mathrm{d}\chi \, q_g(\chi) \, \epsilon_\delta(\boldsymbol{x}, \tau) \, \delta(\boldsymbol{x}, \tau) \right] \\
&= \int \mathrm{d}\chi \, q_{b_\delta}(\chi) \, \delta(\boldsymbol{x}, \tau) + \int \mathrm{d}\chi \, q_{b_{\delta^2}}(\chi) \, \delta^2(\boldsymbol{x}, \tau) + \int \mathrm{d}\chi \, q_{b_{s^2}}(\chi) \, \boldsymbol{K}^2(\boldsymbol{x}, \tau) \\
&\quad + \left[\int \mathrm{d}\chi \, q_g(\chi) \, \epsilon(\boldsymbol{x}, \tau) + \int \mathrm{d}\chi \, q_g(\chi) \, \epsilon_\delta(\boldsymbol{x}, \tau) \, \delta(\boldsymbol{x}, \tau) \right] ,
\end{aligned}
$$
$$(2.72)$$

where in the last equality we have identified particular bias projection kernels $q_{b_{\mathcal{O}}}(\chi)$ (e.g. $q_{b_\delta}(\chi)$, $q_{b_{\delta^2}}(\chi)$, $q_{b_{s^2}}(\chi)$) for a given operator $\mathcal{O}(\boldsymbol{x}, \tau)$ as:

$$ q_{b_{\mathcal{O}}}(\chi) = q_g(\chi) \, b_{\mathcal{O}}(\chi) , \tag{2.73} $$

where we have expressed the conformal time coordinate τ in terms of comoving distance χ. Note that the projection kernels of the stochastic bias terms are simply $q_g(\chi)$.

All these bias projection kernels have units $[\mathrm{Mpc}^{-1}]$. The idea of the above equation was to make explicit that whenever one encounters in some projected correlation function of galaxies (e.g. projected power spectrum) an instance of a relevant bias term one should simply use the corresponding $q_{b_{\mathcal{O}}}(\chi)$ kernel as defined above. Often, one also makes the assumption that a given bias term $q_{b_{\mathcal{O}}}^i$ inside a particular redshift bin i is a constant. In such a case the relation becomes even simpler

$$ q_{b_{\mathcal{O}}}^i(\chi) = b_{\mathcal{O}}^i \, q_g^i(\chi) , \tag{2.74} $$

where $q_g^i(\chi) = p^i(\chi)$ is the distribution of galaxies inside the redshift bin i.

2.5   Statistics on lensing and projected galaxy fields

Now that we have introduced the two projected fields that are routinely observed in galaxy imaging LSS surveys, we will discuss what the statistical quantities are that can be measured and analysed from these fields. We will first start with the discussion of the conventional 2-point methods employed by most LSS surveys, the so-called 3×2-point correlation

---

that $b_\delta = b_1$ and $b_{\delta^2} = \frac{b_2}{2}$.

functions (see Sec. 2.5.1). However, the late-time matter and galaxy density fields are non-Gaussian distributed (Bernardeau et al., 2002) and thus have information contained in higher-order moments that are not captured by 2-point correlation functions alone. Hence, going beyond 2-point statistics and investigating higher-order correlation functions is of great interest as they can enable even tighter constraints on cosmological parameters. Efforts on this front using cosmic shear or galaxy data include the 3-point cosmic shear correlation functions (3PCF) and third-order aperture mass moments (Takada & Jain, 2004; Schneider et al., 2005; Semboloni et al., 2013; Fu et al., 2014; Secco & DES Collaboration, 2022; Heydenreich et al., 2022), galaxy-galaxy-galaxy lensing (Schneider, P. & Watts, P., 2005; Linke et al., 2022), density-split statistics Friedrich et al. (2018); Gruen & DES Collaboration (2018); Burger et al. (2020); Burger, Pierre A. et al. (2023), the lensing aperture mass and convergence PDF (Barthelemy et al., 2021; Boyle et al., 2021; Giblin et al., 2023), third-order convergence moments (Jain & Seljak, 1997; Gatti & DES Collaboration, 2022) and weak lensing peak statistics (Harnois-Déraps et al., 2021; Zürcher & DES Collaboration, 2022; Davies et al., 2022; Lanzieri et al., 2023). These works focus mostly on cosmic shear data, with only a few analysing the galaxy and shear fields together. In particular, a robust framework for joint galaxy and shear cross-correlation analyses for a higher-order equivalent to the 3×2PCFs that can be directly applied to galaxy imaging data to obtain improved cosmological constraints has not been developed so far. Hence, in Sec. 2.5.2 I will introduce the concept of a practical 3PCF statistic that we have developed for application to galaxy imaging datasets to extract higher-order information. In Sec. 2.5.3 I will discuss another beyond 2PCF statistic, the PDF of the galaxy-matter density field.

2.5.1 3×2-point correlation functions

Using the cosmic shear and galaxy density contrast fields, one can construct three 2PCFs consisting of auto and cross-correlations (see chapter 7):

- Cosmic shear 2PCFs $\xi_\pm$ (shear-shear) defined by correlating the *rotated shear* $\gamma_{\phi_{\boldsymbol{\alpha}}}$ at two angular positions $\boldsymbol{\theta}$ and $\boldsymbol{\theta} + \boldsymbol{\alpha}$ on the shear field, where $\gamma_{\phi_{\boldsymbol{\alpha}}}$ at each point is computed along the direction $\phi_{\boldsymbol{\alpha}}$ of the separation vector $\boldsymbol{\alpha}$ between the two points (see Schneider et al. (2002); Jarvis et al. (2004)),

$$
\begin{aligned}
\xi_+^{ij}(\alpha) &\equiv \langle \gamma_{\phi_{\boldsymbol{\alpha}}}^i(\boldsymbol{\theta}) \gamma_{\phi_{\boldsymbol{\alpha}}}^{j*}(\boldsymbol{\theta} + \boldsymbol{\alpha}) \rangle = \int \frac{\mathrm{d}\ell\,\ell}{2\pi} \mathcal{P}_\kappa^{ij}(\ell) J_0(\ell\alpha), \\
\xi_-^{ij}(\alpha) &\equiv \langle \gamma_{\phi_{\boldsymbol{\alpha}}}^i(\boldsymbol{\theta}) \gamma_{\phi_{\boldsymbol{\alpha}}}^j(\boldsymbol{\theta} + \boldsymbol{\alpha}) \rangle = \int \frac{\mathrm{d}\ell\,\ell}{2\pi} \mathcal{P}_\kappa^{ij}(\ell) J_4(\ell\alpha);
\end{aligned}
\tag{2.75}
$$

- Angular galaxy clustering 2PCF ξ_g (galaxy-galaxy) measured by correlating two points separated by α on the galaxy density contrast field δ_g^{2D} Krause et al. (2021):

$$
\xi_g^{ij}(\alpha) \equiv \langle \delta_g^{\mathrm{2D},i}(\boldsymbol{\theta}) \delta_g^{\mathrm{2D},j}(\boldsymbol{\theta} + \boldsymbol{\alpha}) \rangle = \int \frac{\mathrm{d}\ell\,\ell}{2\pi} \mathcal{P}_g^{ij}(\ell) J_0(\ell\alpha);
\tag{2.76}
$$

- Tangential shear 2PCF ξ_t (galaxy-shear), which is the cross-correlation of the foreground galaxy density field with the rotated shear of a background source galaxy along the direction of the separation vector $\boldsymbol{\alpha}$ joining the foreground lens and the background source galaxy.[9] It can be written as (see Krause et al. (2021)):

$$\xi_t^{ij}(\alpha) \equiv \langle \delta_g^{\text{2D},i}(\boldsymbol{\theta}) \gamma_{\phi_{\boldsymbol{\alpha}}}^j(\boldsymbol{\theta} + \boldsymbol{\alpha}) \rangle = \int \frac{\mathrm{d}\ell\,\ell}{2\pi} \mathcal{P}_t^{ij}(\ell) J_2(\ell\alpha) \ . \tag{2.77}$$

In the equations above, the superscripts i, j denote tomographic bins of the background shear source galaxies or the foreground lens galaxies. We consider only the so-called E-mode shear fields, for which the imaginary parts of $\xi_\pm$ and ξ_t vanish. In the last equalities in Eqs. (2.75), (2.76) and (2.77), we have related the real space 2PCFs to the corresponding lensing/galaxy (cross-) power spectra through inverse harmonic transforms (with J_n being the n-th order ordinary Bessel function of the first kind). These spectra can in turn be expressed as line-of-sight projections of the 3D matter/galaxy (cross-) power spectra using the Limber approximation Limber (1954); Kaiser (1992); Krause et al. (2021). In this approximation, one assumes that the correlation between the two shells at comoving distances χ_1 and χ_2 approaches zero faster as $\Delta\chi = \chi_2 - \chi_1 \to \pm\infty$ than due to the average variations in the distribution kernels $q(\chi_1)$ and $q(\chi_2)$ (Bartelmann & Schneider, 2001; Friedrich, 2018).

$$\mathcal{P}_\kappa^{ij}(\ell) = \int \mathrm{d}\chi \frac{q_\kappa^i(\chi) q_\kappa^j(\chi)}{\chi^2} P_{mm}^{\text{3D}}\left(\frac{\ell}{\chi}, \chi\right), \tag{2.78a}$$

$$\mathcal{P}_g^{ij}(\ell) = \int \mathrm{d}\chi \frac{q_g^i(\chi) q_g^j(\chi)}{\chi^2} P_{gg}^{\text{3D}}\left(\frac{\ell}{\chi}, \chi\right), \tag{2.78b}$$

$$\mathcal{P}_t^{ij}(\ell) = \int \mathrm{d}\chi \frac{q_g^i(\chi) q_\kappa^j(\chi)}{\chi^2} P_{gm}^{\text{3D}}\left(\frac{\ell}{\chi}, \chi\right). \tag{2.78c}$$

Here, we have defined the convergence power spectrum in Fourier space as

$$(2\pi)^2 \mathcal{P}_\kappa^{ij}(\ell) \delta_D(\boldsymbol{\ell} + \boldsymbol{\ell}') = \langle \kappa^i(\boldsymbol{\ell}) \kappa^j(\boldsymbol{\ell}') \rangle,$$

the 2D galaxy number density contrast power spectrum as

$$(2\pi)^2 \mathcal{P}_g^{ij}(\ell) \delta_D(\boldsymbol{\ell} + \boldsymbol{\ell}') = \langle \delta_g^{\text{2D},i}(\boldsymbol{\ell}) \delta_g^{\text{2D},j}(\boldsymbol{\ell}') \rangle,$$

and the convergence-galaxy cross-power spectrum as

$$(2\pi)^2 \mathcal{P}_t^{ij}(\ell) \delta_D(\boldsymbol{\ell} + \boldsymbol{\ell}') = \langle \delta_g^{\text{2D},i}(\boldsymbol{\ell}) \kappa^j(\boldsymbol{\ell}') \rangle.$$

One can use the nonlinear matter power spectrum e.g. from `Halofit` (Takahashi et al., 2012) or `HMCODE` (Mead et al., 2015) to evaluate P_{mm}^{3D} that enters the calculation of $\xi_\pm$.

---

[9]This 2PCF is also known as *galaxy-galaxy lensing* in literature, but we refrain from calling it so to

To evaluate P_{gg}^{3D} and P_{gm}^{3D} (i.e. galaxy-galaxy and galaxy-matter power spectra), which enter the calculation of ξ_g and ξ_t, we will rely on standard perturbation theory (SPT) (see Appendix C of chapter 7). Working at leading order (tree-level) and evaluating these 3D spectra as $P_{gg}^{3D} = b_1^2 P_{mm}^{3D} + P_{\epsilon\epsilon}^{3D}$ and $P_{gm}^{3D} = b_1 P_{mm}^{3D}$, where $P_{\epsilon\epsilon}^{3D}$ is the power spectrum of the stochastic field $\epsilon(\boldsymbol{x})$ (see chapter 7).

These three 2-point statistics employed in weak gravitational lensing surveys are together known as the 3×2-point correlation functions (3×2PCFs), jointly probing projections of the late-time power spectrum of matter and galaxy density perturbations. Analysing them simultaneously allow one to constrain both the b_1 and the A_s (or σ_8) parameters (this is a so-called *degeneracy*), where A_s is the amplitude of the primordial scalar power spectrum; at leading-order, we have

$$\xi_\pm \propto A_s;$$

$$\xi_t \propto b_1 A_s;$$

$$\xi_g \propto b_1^2 A_s.$$

These scaling arguments are useful to understand the behaviour of the different 2PCFs with respect to parameters.

2.5.2 Integrated 3-point correlation functions

Calculating an estimate for the 2-point correlation function at a given separation (or in Fourier space, the power spectrum) of the density field in a volume having N galaxies requires taking all pairs of galaxies from the sample of N. On the other hand, calculating the 3-point correlation function (or in Fourier space, the bispectrum) requires taking all triples from N which is computationally much more expensive (Chiang et al., 2014). However, imposing some restrictions, one can derive information from specific limits of the bispectrum (Chiang et al., 2014) without the need for calculating the full bispectrum (or the full 3-point function). As described in Chiang et al. (2014, 2015), the bispectrum (analogously, the 3PCF) is generated by non-linear gravitational evolution (and possibly inflationary physics) and in the so called 'squeezed-limit' ($k_3 \ll k_1 \approx k_2$) describes the influence of large-scale density fluctuations (mode k_3) on small scale structure formation (modes k_1 and k_2). In Chiang et al. (2014), the authors showed that this 'squeezed-limit' bispectrum n 3D can be measured by an *integral* of the full bispectrum — hence called the *integrated bispectrum* — that is dominated by the squeezed-configurations. The authors further showed that if one divides a volume V into many 'subvolumes' V_L each located at distinct positions $\mathbf{r}_L$ (with dimension of length L of the subvolume); the **integrated bispectrum** $\mathcal{B}^{3D}(\mathbf{k})$ is given by the ensemble average of the product of the *mean density contrast* $\bar{\delta}(\mathbf{r}_L)$ — relates to the large scale mode k_3 — measured in each subvolume (e.g. consider the density field and volume in 3D space) and the so called *position-dependent*

---

avoid confusions with the galaxy-galaxy clustering 2PCF.

power spectrum $P(\mathbf{k}; \mathbf{r}_L)$ measured in the corresponding subvolumes — this relates to the small scale modes k_1 and k_2. Therefore, the integrated bispectrum is written as

$$\mathcal{B}^{3D}(\mathbf{k}) \equiv \left\langle \bar{\delta}(\mathbf{r}_L) P(\mathbf{k}; \mathbf{r}_L) \right\rangle . \tag{2.79}$$

where, the *position-dependent power spectrum* $P(\mathbf{k}; \mathbf{r}_L)$ is related to the 2-point correlation of the Fourier modes $\delta(\mathbf{k}; \mathbf{r}_L)$, the local Fourier transform of the real space density fluctuation field $\delta(\mathbf{r}; \mathbf{r}_L)$, through the following formula (Chiang et al., 2014):

$$P(\mathbf{k}; \mathbf{r}_L) \equiv \frac{1}{V_L} \left\langle \delta(\mathbf{k}; \mathbf{r}_L) \delta^*(\mathbf{k}; \mathbf{r}_L) \right\rangle = \frac{1}{V_L} |\delta(\mathbf{k}; \mathbf{r}_L)|^2 . \tag{2.80}$$

As one can already guess, the real space counterpart of the integrated bispectrum is given by the **integrated 3-point correlation function** (see Chiang et al. (2015) for details) and for the 3D density field is written as (using statistical homogeneity and isotropy of the field)

$$\zeta^{3D}(r) \equiv \left\langle \bar{\delta}(\mathbf{r}_L) \xi(r; \mathbf{r}_L) \right\rangle . \tag{2.81}$$

In real space, the squeezed-limit criterion for this statistic holds when the separation r of $\xi(r; \mathbf{r}_L)$ is much smaller than the dimension of the subvolume i.e. $r \ll L$ (Chiang et al., 2015). In this case, $\zeta^{3D}(r)$ can be physically interpreted as the response of the small scale density fluctuations (encoded in $\xi(r; \mathbf{r}_L)$) to the mean density $\bar{\delta}(\mathbf{r}_L)$ which acts as a large-scale 'background density'. This is called the response approach and the interested reader is referred to Chiang et al. (2015) for more details. However, this does not mean that this statistic given in equation (2.81) can only be evaluated for the squeezed-limit. One can evaluate it for both squeezed and non-squeezed configurations. In the latter case, there might not be some straightforward physical interpretation as that for the squeezed-limit, nevertheless, being a higher order statistic it is still a probe for investigating non-linearities (in the form of non-Gaussianities). As demonstrated in Chiang et al. (2015), the $\zeta^{3D}(r)$ is an integral of the full 3-point correlation function $\zeta(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$. These works Chiang et al. (2014, 2015) studied the integrated 3-point correlation function for the galaxy density field in 3D. In this thesis we have taken this concept and developed a full framework of the integrated 3-point correlation functions for projected 2D weak lensing and galaxy density fields along with their cross-correlations. This enables us to extend the conventional 3×2-point cross-correlation functions for 2D fields to the integrated 3-point correlation function framework and extract higher-order information from projected fields.

In essence, an integrated 3PCF in 2D is simply the correlation between (i) a position dependent 1-point weighted 2D mean density of a projected field within a patch of the galaxy imaging survey with (ii) the position-dependent 2PCF measured at the same patch location (see Fig. 1 of chapter 7). For the shear and galaxy density contrast fields, one can therefore correlate the 1-point mean shear signal (called the lensing aperture mass M_a) or the mean galaxy density contrast M_g with the standard 3×2-point correlation functions measured within the same patches. This enables a total of 6 integrated 3PCFs. The derivation steps are given in chapter 7 and we simply quote the results here:

$$\zeta_{a\pm}^{ijk}(\alpha) = \left\langle M_a^i(\boldsymbol{\theta}_C) \, \xi_{\pm}^{jk}(\alpha; \boldsymbol{\theta}_C) \right\rangle = \frac{1}{A_{2pt}(\alpha)} \int \frac{d\ell\, \ell}{2\pi} \, \mathcal{B}_{a\pm}^{ijk}(\ell) J_{0/4}(\ell\alpha), \tag{2.82}$$

$$\zeta_{g\pm}^{ijk}(\alpha) = \left\langle M_g^i(\boldsymbol{\theta}_C)\, \xi_\pm^{jk}(\alpha; \boldsymbol{\theta}_C) \right\rangle = \frac{1}{A_W A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi}\, \mathcal{B}_{g\pm}^{ijk}(\ell) J_{0/4}(\ell\alpha), \tag{2.83}$$

$$\zeta_{ag}^{ijk}(\alpha) = \left\langle M_a^i(\boldsymbol{\theta}_C)\, \xi_g^{jk}(\alpha; \boldsymbol{\theta}_C) \right\rangle = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi}\, \mathcal{B}_{ag}^{ijk}(\ell) J_0(\ell\alpha), \tag{2.84}$$

$$\zeta_{gg}^{ijk}(\alpha) = \left\langle M_g^i(\boldsymbol{\theta}_C)\, \xi_g^{jk}(\alpha; \boldsymbol{\theta}_C) \right\rangle = \frac{1}{A_W A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi}\, \mathcal{B}_{gg}^{ijk}(\ell) J_0(\ell\alpha), \tag{2.85}$$

$$\zeta_{at}^{ijk}(\alpha) = \left\langle M_a^i(\boldsymbol{\theta}_C)\, \xi_t^{jk}(\alpha; \boldsymbol{\theta}_C) \right\rangle = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi}\, \mathcal{B}_{at}^{ijk}(\ell) J_2(\ell\alpha) \,, \tag{2.86}$$

$$\zeta_{gt}^{ijk}(\alpha) = \left\langle M_g^i(\boldsymbol{\theta}_C)\, \xi_t^{jk}(\alpha; \boldsymbol{\theta}_C) \right\rangle = \frac{1}{A_W A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi}\, \mathcal{B}_{gt}^{ijk}(\ell) J_2(\ell\alpha) \,, \tag{2.87}$$

where i labels the source or lens tomographic bin inside which we measure either M_a or M_g, and j and k denote the tomographic bins used to compute the three position-dependent 2PCFs $\xi_\pm, \xi_g, \xi_t$. The angle brackets denote ensemble average (or in practice, averaging over all patch positions $\boldsymbol{\theta}_C$). The $A_{2\mathrm{pt}}$ and A_W are area normalisation factors in 2D, similar to the volume normalisation term V_L for the integrated 3-point correlation function in 3D. The equations above show the real-space ζ in terms of their corresponding Fourier-space counterparts called the *integrated bispectra* $\mathcal{B}(\ell)$. These integrated bispectra can be expressed in terms of line-of-sight projections of 3D matter and galaxy (cross-) bispectra (i.e. $B_{mmm}^{3\mathrm{D}}, B_{gmm}^{3\mathrm{D}}, B_{ggm}^{3\mathrm{D}}, B_{ggg}^{3\mathrm{D}}$) using the Limber approximation (Buchalter et al., 2000) as

$$\mathcal{B}_{a\pm}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_\kappa^i(\chi) q_\kappa^j(\chi) q_\kappa^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{mmm}^{3\mathrm{D}}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi\right) e^{2i\left(\phi_{\boldsymbol{\ell}_2} \mp \phi_{-\boldsymbol{\ell}_{12}}\right)}$$
$$\times U(\boldsymbol{\ell}_1) W(\boldsymbol{\ell}_2 + \boldsymbol{\ell}) W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}), \tag{2.88}$$

$$\mathcal{B}_{g\pm}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_g^i(\chi) q_\kappa^j(\chi) q_\kappa^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{gmm}^{3\mathrm{D}}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi\right) e^{2i\left(\phi_{\boldsymbol{\ell}_2} \mp \phi_{-\boldsymbol{\ell}_{12}}\right)}$$
$$\times W(\boldsymbol{\ell}_1) W(\boldsymbol{\ell}_2 + \boldsymbol{\ell}) W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}), \tag{2.89}$$

$$\mathcal{B}_{ag}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_\kappa^i(\chi) q_g^j(\chi) q_g^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{mgg}^{3\mathrm{D}}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi\right) U(\boldsymbol{\ell}_1) W(\boldsymbol{\ell}_2 + \boldsymbol{\ell}) W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}), \tag{2.90}$$

$$\mathcal{B}_{gg}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_g^i(\chi) q_g^j(\chi) q_g^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{ggg}^{3\mathrm{D}}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi\right) W(\boldsymbol{\ell}_1) W(\boldsymbol{\ell}_2 + \boldsymbol{\ell}) W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}), \tag{2.91}$$

$$\mathcal{B}_{at}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_\kappa^i(\chi) q_g^j(\chi) q_\kappa^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{mgm}^{3\mathrm{D}} \left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi \right) e^{2i\phi_{-\boldsymbol{\ell}_{12}}}$$
$$\times U(\boldsymbol{\ell}_1) W(\boldsymbol{\ell}_2 + \boldsymbol{\ell}) W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}), \tag{2.92}$$

$$\mathcal{B}_{gt}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_g^i(\chi) q_g^j(\chi) q_\kappa^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{ggm}^{3\mathrm{D}} \left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi \right) e^{2i\phi_{-\boldsymbol{\ell}_{12}}}$$
$$\times W(\boldsymbol{\ell}_1) W(\boldsymbol{\ell}_2 + \boldsymbol{\ell}) W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}), \tag{2.93}$$

where $U(\boldsymbol{\ell})$ and $W(\boldsymbol{\ell})$ are Fourier space representations of 2D compensated and tophat filters, respectively, within which the the position-dependent 1-point mean density and the 2PCFs are measured. Modelling the bispectra at leading-order in perturbation theory, one can find that the various integrated 3PCFs display different scalings of the galaxy bias terms. Concretely,

$$\zeta_{a\pm} \propto A_s^2;$$
$$\zeta_{g\pm}, \zeta_{at} \propto \{b_1 A_s^2, b_2 A_s^2, b_{s^2} A_s^2\};$$
$$\zeta_{ag}, \zeta_{gt} \propto \{b_1^2 A_s^2, b_1 b_2 A_s^2, b_1 b_{s^2} A_s^2, b_1 A_s/\bar{n}\};$$
$$\zeta_{gg} \propto \{b_1^3 A_s^2, b_1^2 b_2 A_s^2, b_1^2 b_{s^2} A_s^2, b_1^2 A_s/\bar{n}\};$$

here, $\bar{n}$ denotes the mean number density of galaxies in a given tomographic bin. The different sensitivity of the 6 integrated 3PCFs to the galaxy bias and cosmological parameters relative to the 3 global 2PCFs discussed in Sec. 2.5.1 indicates that joint analyses of these statistics can help to lift parameter degeneracies, leading to tighter parameter constraints overall.

2.5.3   1-point PDF of cosmic density fields

So far we have discussed N-point correlation functions, specifically the integrated 3-point correlation function, as a practical way to probe the non-Gaussianity in the late-time density fields (in 3D or in projection). However, besides multi-point correlations there are other ways of analysing the cosmic density fields which can extract information that is not captured by 2PCFs. One such way is by studying the full-shape of the joint 1-point probability density function (PDF) $p(\delta_m, \delta_g)$ of the local matter δ_m and galaxy density δ_g fluctuation fields. Compared to conventional statistical techniques measured locally on these two fields which only investigate the variance of δ_m, the variance of δ_g and their cross-covariance, the *full shape of the joint PDF* on the other hand is sensitive to information in *all* local higher-order 1-point joint-moments of these two fields. Similar to the galaxy-shear integrated 3PCFs, this can enable tighter constraints on cosmological parameters as well as

allow for the extraction of more detailed information about the galaxy-matter connection between the invisible dark matter and the observed galaxy density field that traces it (Uhlemann et al., 2020; Friedrich et al., 2020; Boyle et al., 2021; Friedrich et al., 2022). The latter includes not only the galaxy bias parameters but also the stochasticity parameters i.e. the scatter between galaxy density and matter density fluctuations (Friedrich et al., 2018; Gruen & DES Collaboration, 2018).

Practically, the joint PDF of the field is measured first by taking two projected fields, δ_m (e.g. this can be probed by the weak lensing convergence which is a line-of-sight projected matter density contrast as described in equation (2.57)) and foreground projected galaxy number density field δ_g. At a given location, one then measures within a tophat cylindrical 2D filter (of a given radius) the smoothed projected matter density contrast (or convergence) as well as the foreground galaxy number counts. Performing this measurement at several locations one can then simply create the normalised histogram of these two quantities which is a measurement of the joint 1-point PDF of the two fields. In order to theoretically model the joint PDF one uses Bayes theorem to write:

$$p(\delta_g, \delta_m) = p(\delta_m)p(\delta_g|\delta_m) \ . \tag{2.94}$$

The modelling of $p(\delta_m)$ in quasi-linear scales can be performed analytically using the theory of large deviations as applied in cosmology (see Friedrich et al. (2022); Barthelemy et al. (2021) for details). On the other hand, using the galaxy biasing formalism discussed in Sec. 2.3 along with a stochasticity model for the galaxy sample one is interested in, one can model $p(\delta_g|\delta_m)$. This has been done in Friedrich et al. (2022) and is presented in Chapter 8. This framework thus constitutes a robust way of studying the cosmos besides N-point correlation functions.

Chapter 3

The integrated 3-point correlation function of cosmic shear

Author's contribution

In this work I laid down the framework for measuring and modelling the integrated 3-point correlation function (3PCF) of the directly observable weak lensing cosmic shear field. In this publication as first author I performed: the theoretical modelling, numerical calculations, processing of T17 lensing maps from Takahashi et al. (2017), creation of lognormal simulations using the `FLASK` software (Xavier et al., 2016), measurements of the integrated 3PCF from these simulations, validation of the modelling along with Fisher parameter constraints forecast and also wrote the paper. All the other authors contributed through valuable discussions, ideas and thorough reviewing of the paper draft.

The integrated three-point correlation function of cosmic shear

Anik Halder [1,2]* Oliver Friedrich,[3,4] Stella Seitz[1,2] and Tamas N. Varga[1,2]

[1]*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstraße 1, D-81679 München, Germany*
[2]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse 1, D-85748 Garching, Germany*
[3]*Kavli Institute for Cosmology, University of Cambridge, Cambridge CB3 0HA, UK*
[4]*Churchill College, University of Cambridge, Cambridge CB3 0DS, UK*

ABSTRACT

We present the integrated three-point shear correlation function $i\zeta_\pm$ – a higher order statistic of the cosmic shear field – which can be directly estimated in wide-area weak lensing surveys without measuring the full three-point shear correlation function, making this a practical and complementary tool to two-point statistics for weak lensing cosmology. We define it as the one-point aperture mass statistic $M_{\rm ap}$ measured at different locations on the shear field correlated with the corresponding *local* two-point shear correlation function $\xi_\pm$. Building upon existing work on the integrated bispectrum of the weak lensing convergence field, we present a theoretical framework for computing the integrated three-point function in real space for any projected field within the flat-sky approximation and apply it to cosmic shear. Using analytical formulae for the non-linear matter power spectrum and bispectrum, we model $i\zeta_\pm$ and validate it on N-body simulations within the uncertainties expected from the sixth year cosmic shear data of the Dark Energy Survey. We also explore the Fisher information content of $i\zeta_\pm$ and perform a joint analysis with $\xi_\pm$ for two tomographic source redshift bins with realistic shape noise to analyse its power in constraining cosmological parameters. We find that the joint analysis of $\xi_\pm$ and $i\zeta_\pm$ has the potential to considerably improve parameter constraints from $\xi_\pm$ alone, and can be particularly useful in improving the figure of merit of the dynamical dark energy equation of state parameters from cosmic shear data.

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

1 INTRODUCTION

Weak gravitational lensing involves the study of the cosmic shear field γ – coherent distortions imprinted in the shapes of background source galaxies by the gravitational lensing effect of foreground matter distribution in the Universe (Bartelmann & Schneider 2001; Schneider 2006; Kilbinger 2015). Statistical analysis of the shear field facilitates the inference of various cosmological model parameters describing the foreground (late-time) matter field. The spatial distribution of these late-time matter density fluctuations consists of several, moderately underdense regions (e.g. voids) and relatively fewer, but highly overdense regions (e.g. galaxies, galaxy clusters) that have emerged through the interplay of gravitational and baryonic processes over billions of years. As a consequence, the late-time density fluctuations follow a positively skewed non-Gaussian distribution. However, most of the statistical analyses currently performed on cosmic shear data are focused on the evaluation of two-point shear correlation functions $\xi_\pm$ (Troxel et al. 2018; Hamana et al. 2020; Asgari et al. 2021), which are insensitive to the information contained in the higher order moments of the distribution. Therefore, the need for exploring methods beyond two-point statistics in the vast amounts of observed shear data is of paramount importance. These higher order statistics may not constrain cosmological parameters better than two-point correlation functions. However, due

to different dependence on the parameters, they hold the potential to break parameter degeneracies that appear in two-point analyses.

The three-point correlation function of cosmic shear (γ-3PCF), generalized third-order aperture mass statistics (Schneider & Lombardi 2003; Kilbinger & Schneider 2005; Schneider, Kilbinger & Lombardi 2005), and weak lensing convergence bispectrum (Takada & Jain 2004; Kayo & Takada 2013; Sato & Nishimichi 2013) are examples of third-order statistics that can probe the full three-point information of the observed weak lensing field. Cosmological constraints using the γ-3PCF were first reported by Semboloni et al. (2011) in the COSMOS survey and by Fu et al. (2014) in the CFHTLS survey.[1] However, in current weak lensing surveys (such as DES, KiDs, HSC),[2] which span thousand square degrees and larger areas on the sky (much larger than COSMOS and CFHTLS), measuring and analysing the full γ-3PCF remains unexplored due to both theoretical and observational challenges.

Hence, in recent years, many alternate methods to probe parts of the higher order information in the cosmic shear field have been

---

[1]COSMOS – Cosmic Evolution Survey, https://cosmos.astro.caltech.edu; CFHTLS – Canada–France–Hawaii Telescope Legacy Survey, https://www.cfht.hawaii.edu/Science/CFHTLS/.

[2]DES – Dark Energy Survey, https://www.darkenergysurvey.org; KiDS – Kilo Degree Survey, http://kids.strw.leidenuniv.nl/index.php; HSC – Hyper Suprime-Cam Survey, https://hsc.mtk.nao.ac.jp/ssp/.

* E-mail: ahalder@usm.lmu.de

proposed and some even measured in data, which although do not capture the full three-point information, are easier to measure and model than γ-3PCF. Examples are shear peak statistics (Kacprzak et al. 2016), shear peak counts and minima (Zürcher et al. 2021), density split statistics (Friedrich et al. 2018; Gruen et al. 2018; Burger et al. 2020), lensing mass-map moments (Chang et al. 2018; Gatti et al. 2020), and joint analyses of shear peaks with $\xi_\pm$ (Harnois-Déraps et al. 2020; Martinet et al. 2021) to name a few. Most of them show potential in putting tighter constraints on cosmological parameters obtained from $\xi_\pm$ alone.

In this paper, we propose another such statistic that can be measured directly from cosmic shear data, namely,[3] *the integrated three-point shear correlation function* $i\zeta_\pm$. We define the statistic as the *aperture mass* measured using a compensated filter at several locations, and correlate them with the *position-dependent shear two-point correlation function* measured within top-hat patches at the corresponding locations. Some key aspects that we explore in this paper are the following.

(i) This statistic is the real space counterpart of the recently introduced *integrated bispectrum* of the weak lensing convergence field κ as studied by Munshi et al. (2020b) and Jung et al. (2021). In this paper, we build upon the existing work and formulate a theoretical model for our real-space statistic on the shear field γ and validate it on simulated cosmic shear maps.

(ii) The most desirable feature of $i\zeta_\pm$ is that it can be easily measured from the observed shear field, a direct observable. This is possible because we define $i\zeta_\pm$ using an aperture mass – a weighted measurement of the shear field at a given location using a compensated window that filters out a constant convergence mass sheet – and the position-dependent two-point shear correlation function that is intuitively the $\xi_\pm$ measured within top-hat patches (with area of a few square degrees). Our definition is different from Munshi et al. (2020b) who work with the convergence field in Fourier space and accordingly define the integrated convergence bispectrum iB_κ using the local mean convergence measured within a top-hat patch instead of using a compensated filter. If one would want to measure iB_κ, then it would first be necessary to construct a convergence map from the observed shear field. This map-making process is not at all straightforward in the presence of complicated survey geometry and masks.

(iii) We investigate the information content of $i\zeta_\pm$ for a DES-sized tomographic survey in terms of Fisher constraints on cosmological parameters. This is the first work to perform such an analysis in the context of the integrated weak lensing bispectrum.

We organize the paper in the following manner. In Section 2, we formulate the integrated three-point function statistic for any projected field within the flat-sky approximation and then apply it to the case for the cosmic shear field in Section 3. In Section 4, we describe the simulations and numerical methods we use in order to measure and theoretically model the statistic. Finally, in Section 5, we validate our theoretical model on the simulations and present Fisher constraints on cosmological parameters. Throughout this paper we assume flat cosmology i.e. $\Omega_K = 0$. As we mainly work with projected 2D quantities, we differentiate them from 3D quantities by explicitly specifying the subscript or superscript '3D' for the latter.

2 THEORY I: GENERAL FORMALISM

In this section, we formulate the general framework of equations required for describing the integrated three-point function (in real space) and the integrated bispectrum (in Fourier space) of any projected 2D field within the flat-sky approximation. For this section and the next, we provide a summary of this technical part of the paper at the end of Section 3. Readers may feel free to skip these theoretical details and directly refer to the summary in Section 3.4.

2.1 Projected fields

Any cosmic field $f^{3D}[\boldsymbol{\chi}, \eta]$ that we observe on our past light-cone at 3D comoving position $\boldsymbol{\chi}$ and corresponding conformal lookback time $\eta = \eta_0 - \chi$ (where η_0 is the conformal time today and χ the radial comoving distance) can be projected on to the 2D celestial sphere to obtain the weighted line-of-sight 2D quantity $f(\hat{\boldsymbol{n}})$ towards a radial unit direction $\hat{\boldsymbol{n}}$ (Bartelmann & Schneider 2001),

$$f(\hat{\boldsymbol{n}}) = \int \mathrm{d}\chi \, q_\mathrm{f}(\chi) f^{3D}\big[\chi\hat{\boldsymbol{n}}, \eta_0 - \chi\big], \tag{1}$$

where $q_\mathrm{f}(\chi)$ is a particular weighting kernel over which f^{3D} is projected. Examples are the projected galaxy number density or the weak lensing convergence field that we observe on the celestial sphere. Assuming that the angular extent of the field of view is small – spanning an area of a few square degrees – we can make the flat-sky approximation, where we denote the position on the sky as a 2D planar vector $\boldsymbol{\theta} = (\theta_x, \theta_y)$ and express f as

$$f(\boldsymbol{\theta}) = \int \mathrm{d}\chi \, q_\mathrm{f}(\chi) f^{3D}\big[(\chi\boldsymbol{\theta}, \chi), \eta_0 - \chi\big]. \tag{2}$$

2.2 The projected power spectrum and bispectrum

The 2D power spectrum (P_gh) and bispectrum (B_fgh) of projected fields f, g, h are defined as (Bartelmann & Schneider 2001)

$$\big\langle g(\boldsymbol{l}_1)h(\boldsymbol{l}_2)\big\rangle \equiv (2\pi)^2 \delta_\mathrm{D}(\boldsymbol{l}_1 + \boldsymbol{l}_2) \, P_\mathrm{gh}(\boldsymbol{l}_1), \tag{3}$$

$$\big\langle f(\boldsymbol{l}_1)g(\boldsymbol{l}_2)h(\boldsymbol{l}_3)\big\rangle \equiv (2\pi)^2 \delta_\mathrm{D}(\boldsymbol{l}_1 + \boldsymbol{l}_2 + \boldsymbol{l}_3) \, B_\mathrm{fgh}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3), \tag{4}$$

where $\langle \cdots \rangle$ denotes an ensemble average over different realizations of the Universe and δ_D denotes the Dirac delta function. $f(\boldsymbol{l})$ corresponds to the Fourier space representation[4] of field $f(\boldsymbol{\theta})$ (see Appendix A); and similarly for fields g, h. The bispectrum is defined for closed triangle configurations $\boldsymbol{l}_1 + \boldsymbol{l}_2 + \boldsymbol{l}_3 = 0$.

These 2D spectra can be expressed as line-of-sight projections of the power spectrum $P_\mathrm{gh}^{3D}(\boldsymbol{k}, \eta)$, and bispectrum $B_\mathrm{fgh}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \eta)$ of the 3D fields f^{3D}, g^{3D}, h^{3D} with $\boldsymbol{k}_i$ corresponding to 3D Fourier wave vectors. This can be computed using the Limber approximation (Limber 1954; Kaiser 1992; Buchalter, Kamionkowski & Jaffe 2000):

$$P_\mathrm{gh}(\boldsymbol{l}) = \int \mathrm{d}\chi \, \frac{q_\mathrm{g}(\chi)q_\mathrm{h}(\chi)}{\chi^2} P_\mathrm{gh}^{3D}\left(\boldsymbol{k} = \frac{\boldsymbol{l}}{\chi}, \eta_0 - \chi\right), \tag{5}$$

---

[3]The i in $i\zeta_\pm$ stands for 'integrated' and should not be confused with the complex imaginary unit $\sqrt{-1}$.

[4]In this paper, we do not use any distinguishing symbol (e.g. the commonly used tilde) for separately denoting the Fourier space representation of the field f. The Fourier representation is left understood when f appears with argument $\boldsymbol{l}$ or $\boldsymbol{q}$ (2D Fourier wave vectors).

$$B_{\text{fgh}}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3) = \int \mathrm{d}\chi \, \frac{q_{\text{f}}(\chi) q_{\text{g}}(\chi) q_{\text{h}}(\chi)}{\chi^4} B_{\text{fgh}}^{\text{3D}}\left(\frac{\boldsymbol{l}_1}{\chi}, \frac{\boldsymbol{l}_2}{\chi}, \frac{\boldsymbol{l}_3}{\chi}, \eta_0 - \chi\right),$$
(6)

where $q_{\text{f}}(\chi)$, $q_{\text{g}}(\chi)$, and $q_{\text{h}}(\chi)$ are the weighting kernels with which f^{3D}, g^{3D}, and h^{3D} are projected, respectively. Under the assumptions of an isotropic Universe, the power spectrum is independent of the direction of the wave vector and the bispectrum does not depend on the orientation of the closed triangle of its wave vectors. It should be noted again that these expressions are written assuming that the Universe is flat. However, it is straightforward to generalize these equations to a universe with non-zero spatial curvature (see Bartelmann & Schneider 2001; Schneider 2006).

2.3 The integrated three-point function and integrated bispectrum of projected fields

The integrated bispectrum $iB^{\text{3D}}(k)$ of the 3D matter density contrast field was first studied by Chiang et al. (2014) who defined it as the correlation of the local mean density perturbation and the position-dependent power spectrum evaluated within 3D subvolumes. They showed that this correlation can be expressed as integrals over different $\boldsymbol{k}$-modes of the full 3D matter density contrast bispectrum B_{δ}^{3D}. Chiang et al. (2015) studied the real space counterpart of $iB^{\text{3D}}(k)$, namely the integrated three-point function $i\zeta^{\text{3D}}(r)$ that they showed to be the correlation of the local mean density perturbation and the position-dependent two-point correlation function within 3D subvolumes and presented the first detection of $i\zeta^{\text{3D}}(r)$ in the Baryon Oscillation Spectroscopic Survey (BOSS) Data Release 10 (DR10) constant-mass (CMASS) galaxy sample. The integrated bispectrum has also found other applications, for example in studying the Lyman α forest, quasars (Doux et al. 2016; Chiang et al. 2017; Chiang & Slosar 2018), and also the 21-cm line in the epoch of reionization (Giri et al. 2019). Recently, Munshi & Coles (2017), Munshi et al. (2020a,b), and Jung et al. (2020, 2021) have extended the formalism to the integrated bispectrum $iB(l)$ of projected 2D fields. In particular, Munshi et al. (2020b) studied this in the context of the weak lensing convergence field and developed various theoretical models for the same. In this section, we build upon the mathematical formalism of the integrated bispectrum developed in these previous works and introduce its real space counterpart the integrated three-point function $i\zeta(\theta)$ for any projected 2D field.

2.3.1 Projected field within 2D window

The central quantity to our discussion will be the projected field $f(\boldsymbol{\theta}; \boldsymbol{\theta}_C)$ at a given location $\boldsymbol{\theta}$ on the flat-sky, weighted by an azimuthally symmetric 2D window function W (of a given size or characteristic scale) centred at $\boldsymbol{\theta}_C$,

$$f(\boldsymbol{\theta}; \boldsymbol{\theta}_C) \equiv f(\boldsymbol{\theta}) W(\boldsymbol{\theta}_C - \boldsymbol{\theta}),$$
(7)

where $W(\boldsymbol{\theta}_C - \boldsymbol{\theta}) = W(\boldsymbol{\theta} - \boldsymbol{\theta}_C) = W(|\boldsymbol{\theta}_C - \boldsymbol{\theta}|)$. For example, if the window function centred at $\boldsymbol{\theta}_C$ is a top-hat of size θ_{T}, then $f(\boldsymbol{\theta}; \boldsymbol{\theta}_C) = f(\boldsymbol{\theta})$ only when $|\boldsymbol{\theta}_C - \boldsymbol{\theta}| \leq \theta_{\text{T}}$, otherwise $f(\boldsymbol{\theta}; \boldsymbol{\theta}_C) = 0$. Its local Fourier transform (see Appendix A) is given by

$$f(\boldsymbol{l}; \boldsymbol{\theta}_C) \equiv \mathcal{F}_{\text{2D}}[f(\boldsymbol{\theta}; \boldsymbol{\theta}_C)] = \int \mathrm{d}^2\boldsymbol{\theta} \, f(\boldsymbol{\theta}) W(\boldsymbol{\theta}_C - \boldsymbol{\theta}) \mathrm{e}^{-i\boldsymbol{l}\cdot\boldsymbol{\theta}}$$
$$= \int \frac{\mathrm{d}^2\boldsymbol{l}_1}{(2\pi)^2} f(\boldsymbol{l}_1) W(\boldsymbol{l}_1 - \boldsymbol{l}) \mathrm{e}^{i(\boldsymbol{l}_1 - \boldsymbol{l})\cdot\boldsymbol{\theta}_C},$$
(8)

where f can be any complex/real 2D field defined in a tomographic bin with projection kernel q_{f}, e.g. projected galaxy density contrast, weak lensing convergence κ, and weak lensing shear γ. $f(\boldsymbol{l})$ and $W(\boldsymbol{l})$ are the Fourier space representations of $f(\boldsymbol{\theta})$ and $W(\boldsymbol{\theta})$, respectively. If f is a real field, i.e. $f^*(-\boldsymbol{l}) = f(\boldsymbol{l})$, then we can easily see that $f^*(\boldsymbol{l}; \boldsymbol{\theta}_C) = f(-\boldsymbol{l}; \boldsymbol{\theta}_C)$.

2.3.2 Position-dependent weighted mean of projected field

We can now find the weighted mean of $f(\boldsymbol{\theta}; \boldsymbol{\theta}_C)$ defined within the 2D window[5] W_{1pt} at $\boldsymbol{\theta}_C$ as

$$\bar{f}(\boldsymbol{\theta}_C) \equiv \frac{1}{A_{\text{1pt}}} \int \mathrm{d}^2\boldsymbol{\theta} \, f(\boldsymbol{\theta}; \boldsymbol{\theta}_C) = \frac{1}{A_{\text{1pt}}} \int \mathrm{d}^2\boldsymbol{\theta} \, f(\boldsymbol{\theta}) W_{\text{1pt}}(\boldsymbol{\theta}_C - \boldsymbol{\theta})$$
$$= \frac{1}{A_{\text{1pt}}} \int \frac{\mathrm{d}^2\boldsymbol{l}}{(2\pi)^2} f(\boldsymbol{l}) W_{\text{1pt}}(\boldsymbol{l}) \mathrm{e}^{i\boldsymbol{l}\cdot\boldsymbol{\theta}_C},$$
(9)

where for the final equality we have used the convolution theorem and have defined the one-point area normalization term as

$$A_{\text{1pt}} \equiv \int \mathrm{d}^2\boldsymbol{\theta} \, W_{\text{1pt}}(\boldsymbol{\theta}_C - \boldsymbol{\theta}).$$
(10)

Note that this normalization term is a purely geometric factor independent of the location $\boldsymbol{\theta}_C$ of the window (evaluating it at any $\boldsymbol{\theta}_C$ gives the same result and for simplicity we evaluate it at $\boldsymbol{\theta}_C = \boldsymbol{0}$; see also Footnote 6). If we use a normalized window function, i.e. $A_{\text{1pt}} = 1$, or a compensated filter (Schneider 2006) – which shall be important when we consider aperture masses (see Section 3.2) – then we do not need to consider this normalization factor. From equations (8) and (9) we can see that

$$\bar{f}(\boldsymbol{\theta}_C) = \frac{1}{A_{\text{1pt}}} f(\boldsymbol{l} = \boldsymbol{0}; \boldsymbol{\theta}_C).$$
(11)

2.3.3 Position-dependent two-point function of projected fields

The two-point correlation (as a function of the separation 2D vector $\boldsymbol{\alpha}$) of projected fields g and h is defined as

$$\xi_{\text{gh}}(\boldsymbol{\alpha}) \equiv \langle g(\boldsymbol{\theta}) h(\boldsymbol{\theta} + \boldsymbol{\alpha}) \rangle.$$
(12)

This is the real space counterpart of the projected power spectrum $P_{\text{gh}}(\boldsymbol{l})$:

$$\xi_{\text{gh}}(\boldsymbol{\alpha}) = \mathcal{F}_{\text{2D}}^{-1}[P_{\text{gh}}(\boldsymbol{l})] = \int \frac{\mathrm{d}^2\boldsymbol{l}}{(2\pi)^2} P_{\text{gh}}(\boldsymbol{l}) \mathrm{e}^{i\boldsymbol{l}\cdot\boldsymbol{\alpha}}.$$
(13)

Considering isotropic fields, this inverse 2D Fourier transformation becomes an inverse Hankel transform (see Appendix A): $\xi_{\text{gh}}(\alpha) = \mathcal{F}_{\text{2D}}^{-1}[P_{\text{gh}}(l)]$ where the correlation function (power spectrum) is independent of the direction of the separation vector $\boldsymbol{\alpha}$ (Fourier mode $\boldsymbol{l}$).

For ergodic fields, we can write the expression for this two-point correlation function evaluated within a finite region of area A as

$$\hat{\xi}_{\text{gh}}(\boldsymbol{\alpha}) \equiv \frac{1}{A} \int \mathrm{d}^2\boldsymbol{\theta} \, g(\boldsymbol{\theta}) h(\boldsymbol{\theta} + \boldsymbol{\alpha}),$$
(14)

where the integrand for a given separation $\boldsymbol{\alpha}$ is defined only for those points $\boldsymbol{\theta}$ for which both $\boldsymbol{\theta}$ and $\boldsymbol{\theta} + \boldsymbol{\alpha}$ lie within the boundary

---

[5]We use the subscript '1pt' for the window function in the equation for the weighted mean of a field inside the window W_{1pt} at a given location to distinguish it from the case when we compute the position-dependent two-point function within a different window W at the same location (see equation 15).

of the region under consideration. As $A \rightarrow \infty$, $\hat{\xi}_{gh}(\boldsymbol{\alpha}) \rightarrow \xi_{gh}(\boldsymbol{\alpha})$. However, if the region spans only a small area (e.g. a small 2D aperture on the sky), then this limit does not hold and instead the expression $\hat{\xi}_{gh}(\boldsymbol{\alpha})$ evaluates to a value that depends on the location of the aperture. Hence, we now formally define the expression for the *position-dependent two-point correlation function* $\hat{\xi}_{gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)$ of the projected fields g and h both defined within a 2D aperture W centred at $\boldsymbol{\theta}_C$ as

$$
\begin{aligned}
\hat{\xi}_{gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) &\equiv \frac{1}{A_{2pt}(\boldsymbol{\alpha})} \int d^2\boldsymbol{\theta} \ g(\boldsymbol{\theta}; \boldsymbol{\theta}_C) h(\boldsymbol{\theta} + \boldsymbol{\alpha}; \boldsymbol{\theta}_C) \\
&= \frac{1}{A_{2pt}(\boldsymbol{\alpha})} \int d^2\boldsymbol{\theta} \ g(\boldsymbol{\theta}) W(\boldsymbol{\theta}_C - \boldsymbol{\theta}) \\
&\quad \times h(\boldsymbol{\theta} + \boldsymbol{\alpha}) W(\boldsymbol{\theta}_C - \boldsymbol{\theta} - \boldsymbol{\alpha}) \\
&= \frac{1}{A_{2pt}(\boldsymbol{\alpha})} \int \frac{d^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_2}{(2\pi)^2} \int \frac{d^2\boldsymbol{q}}{(2\pi)^2} \ g(\boldsymbol{l}_1) h(\boldsymbol{l}_2) \\
&\quad \times W(\boldsymbol{q}) W(\boldsymbol{l}_1 + \boldsymbol{l}_2 - \boldsymbol{q}) \ e^{i(\boldsymbol{l}_1 + \boldsymbol{l}_2) \cdot \boldsymbol{\theta}_C} e^{i(\boldsymbol{q} - \boldsymbol{l}_1) \cdot \boldsymbol{\alpha}},
\end{aligned}
\tag{15}
$$

where $\boldsymbol{l}_i, \boldsymbol{q}$ are 2D Fourier wave vectors. In the above equation $A_{2pt}(\boldsymbol{\alpha})$ is the area normalization for this projected position-dependent two-point function and is given by

$$
\begin{aligned}
A_{2pt}(\boldsymbol{\alpha}) &\equiv \int d^2\boldsymbol{\theta} \ W(\boldsymbol{\theta}_C - \boldsymbol{\theta}) W(\boldsymbol{\theta}_C - \boldsymbol{\theta} - \boldsymbol{\alpha}) \\
&= \int \frac{d^2\boldsymbol{q}}{(2\pi)^2} \ W(\boldsymbol{q}) W(-\boldsymbol{q}) \ e^{i\boldsymbol{q} \cdot \boldsymbol{\alpha}},
\end{aligned}
\tag{16}
$$

which for simplicity we evaluate (using the first equality) at $\boldsymbol{\theta}_C = \boldsymbol{0}$ as this term is independent of the window's location $\boldsymbol{\theta}_C$.[6] However, it is important to note that this area normalization depends on the separation vector $\boldsymbol{\alpha}$ under consideration, unlike A_{1pt} defined in equation (10). For azimuthally symmetric window functions that we are interested in, it follows from isotropy considerations that this normalization term only depends on the magnitude α of the separation vector, i.e. $A_{2pt}(\boldsymbol{\alpha}) = A_{2pt}(\alpha)$. Hence, one can evaluate this term for any polar angle $\phi_{\boldsymbol{\alpha}}$ (e.g. defined with respect to the *x*-axis of the flat-sky coordinate system). For simplicity, we shall consider $\phi_{\boldsymbol{\alpha}} = 0$.

On the other hand, for the position-dependent two-point correlation function of field g with the complex-conjugated field h^* we have

$$
\begin{aligned}
\hat{\xi}_{gh^*}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) &\equiv \frac{1}{A_{2pt}(\boldsymbol{\alpha})} \int d^2\boldsymbol{\theta} \ g(\boldsymbol{\theta}; \boldsymbol{\theta}_C) h^*(\boldsymbol{\theta} + \boldsymbol{\alpha}; \boldsymbol{\theta}_C) \\
&= \frac{1}{A_{2pt}(\boldsymbol{\alpha})} \int \frac{d^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_2}{(2\pi)^2} \int \frac{d^2\boldsymbol{q}}{(2\pi)^2} \ g(\boldsymbol{l}_1) h^*(-\boldsymbol{l}_2) \\
&\quad \times W(\boldsymbol{q}) W(\boldsymbol{l}_1 + \boldsymbol{l}_2 - \boldsymbol{q}) \ e^{i(\boldsymbol{l}_1 + \boldsymbol{l}_2) \cdot \boldsymbol{\theta}_C} e^{i(\boldsymbol{q} - \boldsymbol{l}_1) \cdot \boldsymbol{\alpha}}.
\end{aligned}
\tag{17}
$$

In case the field h is real, i.e. $h^*(-\boldsymbol{l}) = h(\boldsymbol{l})$, it follows from equation (15) that $\hat{\xi}_{gh^*}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) = \hat{\xi}_{gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)$.

The position-dependent correlation function gives an unbiased estimate of the two-point correlation function, i.e. $\langle \hat{\xi}_{gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) \rangle = \xi_{gh}(\boldsymbol{\alpha})$. Also, when we consider the fields and the window functions to be isotropic, then the above expressions only depend on the magnitude α of the separation vector, i.e. $\hat{\xi}_{gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) = \hat{\xi}_{gh}(\alpha; \boldsymbol{\theta}_C)$.

---

[6]Of course, this is only true when we do not consider holes and masks in the data. To account for this, one may randomly throw away some points inside a window centred at $\boldsymbol{\theta}_C$ so as to have only those pairs of points $\{\boldsymbol{\theta}, \boldsymbol{\theta} + \boldsymbol{\alpha}\}$ yielding the same effective area of another window at $\boldsymbol{\theta}'_C$ but which has masks and holes within its aperture.

2.3.4 Position-dependent power spectrum of projected fields

The power spectrum is the forward Fourier transform of the two-point correlation function. Hence, we define the Fourier space counterpart of $\hat{\xi}_{gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)$ as

$$
\begin{aligned}
\hat{P}_{gh}(\boldsymbol{l}; \boldsymbol{\theta}_C) &\equiv \mathcal{F}_{2D}[A_{2pt}(\boldsymbol{\alpha}) \hat{\xi}_{gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)] \\
&= \int d^2\boldsymbol{\alpha} \ A_{2pt}(\boldsymbol{\alpha}) \hat{\xi}_{gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) e^{-i\boldsymbol{l} \cdot \boldsymbol{\alpha}} \\
&= \int \frac{d^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_2}{(2\pi)^2} \ g(\boldsymbol{l}_1) h(\boldsymbol{l}_2) \\
&\quad \times W(\boldsymbol{l}_1 + \boldsymbol{l}) W(\boldsymbol{l}_2 - \boldsymbol{l}) e^{i(\boldsymbol{l}_1 + \boldsymbol{l}_2) \cdot \boldsymbol{\theta}_C} \\
&= g(-\boldsymbol{l}; \boldsymbol{\theta}_C) h(\boldsymbol{l}; \boldsymbol{\theta}_C).
\end{aligned}
\tag{18}
$$

This is slightly different from the *position-dependent power spectrum* definition of Chiang et al. (2014) who define it for the 3D matter density contrast field in their equation (2.3) with a constant volume normalization term. On the other hand, we factor out the scale-dependent area normalization term $A_{2pt}(\boldsymbol{\alpha})$ in our definition of $\hat{P}_{gh}$.

Similarly, the Fourier space counterpart of $\hat{\xi}_{gh^*}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)$ can be written as

$$
\begin{aligned}
\hat{P}_{gh^*}(\boldsymbol{l}; \boldsymbol{\theta}_C) &\equiv \mathcal{F}_{2D}[A_{2pt}(\boldsymbol{\alpha}) \hat{\xi}_{gh^*}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)] \\
&= \int \frac{d^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_2}{(2\pi)^2} \ g(\boldsymbol{l}_1) h^*(-\boldsymbol{l}_2) \\
&\quad \times W(\boldsymbol{l}_1 + \boldsymbol{l}) W(\boldsymbol{l}_2 - \boldsymbol{l}) e^{i(\boldsymbol{l}_1 + \boldsymbol{l}_2) \cdot \boldsymbol{\theta}_C} \\
&= g(-\boldsymbol{l}; \boldsymbol{\theta}_C) h^*(-\boldsymbol{l}; \boldsymbol{\theta}_C).
\end{aligned}
\tag{19}
$$

When the field h is real, $\hat{P}_{gh^*}(\boldsymbol{l}; \boldsymbol{\theta}_C) = \hat{P}_{gh}(\boldsymbol{l}; \boldsymbol{\theta}_C)$.

2.3.5 Integrated three-point function of projected fields

We now define the integrated three-point function of projected fields analogous to the 3D case (Chiang et al. 2015) – the ensemble average (over different locations $\boldsymbol{\theta}_C$) of the product of the position-dependent weighted mean and the position-dependent two-point function of projected fields:

$$
\begin{aligned}
i\zeta(\boldsymbol{\alpha}) &\equiv \left\langle \bar{f}(\boldsymbol{\theta}_C) \hat{\xi}_{gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) \right\rangle \\
&= \frac{1}{A_{1pt} A_{2pt}(\boldsymbol{\alpha})} \int d^2\boldsymbol{\theta}_1 \int d^2\boldsymbol{\theta}_2 \left\langle f(\boldsymbol{\theta}_1) g(\boldsymbol{\theta}_2) h(\boldsymbol{\theta}_2 + \boldsymbol{\alpha}) \right\rangle \\
&\quad \times W_{1pt}(\boldsymbol{\theta}_C - \boldsymbol{\theta}_1) W(\boldsymbol{\theta}_C - \boldsymbol{\theta}_2) W(\boldsymbol{\theta}_C - \boldsymbol{\theta}_2 - \boldsymbol{\alpha}) \\
&= \frac{1}{A_{1pt} A_{2pt}(\boldsymbol{\alpha})} \int \frac{d^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_2}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_3}{(2\pi)^2} \int \frac{d^2\boldsymbol{q}}{(2\pi)^2} \\
&\quad \times \left\langle f(\boldsymbol{l}_1) g(\boldsymbol{l}_2) h(\boldsymbol{l}_3) \right\rangle e^{i(\boldsymbol{l}_1 + \boldsymbol{l}_2 + \boldsymbol{l}_3) \cdot \boldsymbol{\theta}_C} \\
&\quad \times W_{1pt}(\boldsymbol{l}_1) W(\boldsymbol{q}) W(\boldsymbol{l}_2 + \boldsymbol{l}_3 - \boldsymbol{q}) e^{i(\boldsymbol{q} - \boldsymbol{l}_2) \cdot \boldsymbol{\alpha}},
\end{aligned}
\tag{20}
$$

and for the case with complex-conjugated field h^*:

$$
\begin{aligned}
i\zeta_*(\boldsymbol{\alpha}) &\equiv \left\langle \bar{f}(\boldsymbol{\theta}_C) \hat{\xi}_{gh^*}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) \right\rangle \\
&= \frac{1}{A_{1pt} A_{2pt}(\boldsymbol{\alpha})} \int \frac{d^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_2}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_3}{(2\pi)^2} \int \frac{d^2\boldsymbol{q}}{(2\pi)^2} \\
&\quad \times \left\langle f(\boldsymbol{l}_1) g(\boldsymbol{l}_2) h^*(-\boldsymbol{l}_3) \right\rangle e^{i(\boldsymbol{l}_1 + \boldsymbol{l}_2 + \boldsymbol{l}_3) \cdot \boldsymbol{\theta}_C} \\
&\quad \times W_{1pt}(\boldsymbol{l}_1) W(\boldsymbol{q}) W(\boldsymbol{l}_2 + \boldsymbol{l}_3 - \boldsymbol{q}) e^{i(\boldsymbol{q} - \boldsymbol{l}_2) \cdot \boldsymbol{\alpha}}.
\end{aligned}
\tag{21}
$$

For a real field h, it follows that $i\zeta_*(\boldsymbol{\alpha}) = i\zeta(\boldsymbol{\alpha})$.

2.3.6 Integrated bispectrum of projected fields

The Fourier space counterparts of the above equations can be written as

$$i B(\boldsymbol{l}) \equiv \mathcal{F}_{2D} \left[A_{2pt}(\boldsymbol{\alpha}) i \zeta(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) \right]$$

$$= \frac{1}{A_{1pt}} \int \frac{d^2 \boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2 \boldsymbol{l}_2}{(2\pi)^2} \int \frac{d^2 \boldsymbol{l}_3}{(2\pi)^2} \left\langle f(\boldsymbol{l}_1) g(\boldsymbol{l}_2) h(\boldsymbol{l}_3) \right\rangle$$

$$\times e^{i(\boldsymbol{l}_1 + \boldsymbol{l}_2 + \boldsymbol{l}_3) \cdot \boldsymbol{\theta}_C} W_{1pt}(\boldsymbol{l}_1) W(\boldsymbol{l}_2 + \boldsymbol{l}) W(\boldsymbol{l}_3 - \boldsymbol{l})$$

$$= \left\langle \bar{f}(\boldsymbol{\theta}_C) \, \hat{P}_{gh}(\boldsymbol{l}; \boldsymbol{\theta}_C) \right\rangle, \tag{22}$$

$$i B_*(\boldsymbol{l}) \equiv \mathcal{F}_{2D} \left[A_{2pt}(\boldsymbol{\alpha}) i \zeta_*(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) \right]$$

$$= \frac{1}{A_{1pt}} \int \frac{d^2 \boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2 \boldsymbol{l}_2}{(2\pi)^2} \int \frac{d^2 \boldsymbol{l}_3}{(2\pi)^2} \left\langle f(\boldsymbol{l}_1) g(\boldsymbol{l}_2) h^* (-\boldsymbol{l}_3) \right\rangle$$

$$\times e^{i(\boldsymbol{l}_1 + \boldsymbol{l}_2 + \boldsymbol{l}_3) \cdot \boldsymbol{\theta}_C} W_{1pt}(\boldsymbol{l}_1) W(\boldsymbol{l}_2 + \boldsymbol{l}) W(\boldsymbol{l}_3 - \boldsymbol{l})$$

$$= \left\langle \bar{f}(\boldsymbol{\theta}_C) \, \hat{P}_{gh^*} (\boldsymbol{l}; \boldsymbol{\theta}_C) \right\rangle, \tag{23}$$

where the last lines of both these equations show that the integrated bispectrum is the ensemble average of the position-dependent weighted mean and the position-dependent power spectrum of the projected fields.

From isotropy considerations (of the fields and of the symmetric window functions) we have $i B(\boldsymbol{l}) = i B(l)$ and $i \zeta(\boldsymbol{\alpha}) = i \zeta(\alpha)$. We can thereby relate the integrated three-point function to the integrated bispectrum through an inverse Hankel transform:

$$i \zeta(\alpha) = \frac{1}{A_{2pt}(\alpha)} \mathcal{F}_{2D}^{-1}[i B(l)]. \tag{24}$$

The formalism for the integrated bispectrum and integrated three-point function we have developed so far is very general and applicable to any projected field within the flat-sky approximation. For the curved-sky formulation of the projected integrated bispectrum the reader is referred to the work by Jung et al. (2020).

In this paper, we shall look into only one application of our formalism for the integrated three-point function – on the cosmic shear field.

3 THEORY II: APPLICATION

Having developed the general framework of equations for computing the integrated three-point function for any projected field, we now apply it to the weak lensing shear field and formulate the equations for the *integrated three-point shear correlation function*.

3.1 Weak lensing basics

The light from background (source) galaxies is weakly deflected by the foreground (lens) intervening total matter distribution. This causes a coherent distortion pattern in the observed shapes of these background galaxies and is known as the cosmic shear field. This field can be interpreted as the shear caused by a weighted line-of-sight projection of the 3D matter density field – known as the weak lensing convergence field. Statistical analysis of this shear field (directly observable) through the widely used two-point shear correlation function allows one to infer about the projected power spectrum of the total matter distribution (theoretically predictable) and thereby constrain cosmological parameters.

Following equation (2), the weak lensing convergence field $\kappa(\boldsymbol{\theta})$ acting on source galaxies situated at the radial comoving distance χ_s

can be written as a line-of-sight projection of the 3D matter density contrast field δ^{3D}:

$$\kappa(\boldsymbol{\theta}) = \int d\chi \; q(\chi) \delta^{3D} \left[(\chi \boldsymbol{\theta}, \chi), \eta_0 - \chi \right], \tag{25}$$

with projection kernel $q(\chi)$ (also known as lensing efficiency) written as[7](Kilbinger 2015)

$$q(\chi) = \frac{3 H_0^2 \Omega_m}{2 c^2} \frac{\chi}{a(\chi)} \frac{\chi_s - \chi}{\chi_s}; \qquad \text{with } \chi \leq \chi_s, \tag{26}$$

where Ω_m is the total matter density parameter of the Universe today, H_0 the Hubble parameter today, a the scale factor, and c the speed of light. The convergence and the associated complex shear field are related to each other through second-order derivatives of the lensing potential $\psi(\boldsymbol{\theta})$ in the 2D sky-plane (Schneider 2006):

$$\kappa(\boldsymbol{\theta}) = \frac{1}{2} \left(\partial_x^2 + \partial_y^2 \right) \psi(\boldsymbol{\theta}), \quad \gamma(\boldsymbol{\theta}) = \frac{1}{2} \left(\partial_x^2 - \partial_y^2 + 2i \partial_x \partial_y \right) \psi(\boldsymbol{\theta}), \tag{27}$$

where $\psi(\boldsymbol{\theta})$ is the line-of-sight projection of the 3D Newtonian gravitational potential $\Phi \left[(\chi \boldsymbol{\theta}, \chi), \eta_0 - \chi \right]$ of the total matter distribution:

$$\psi(\boldsymbol{\theta}) = \frac{2}{c^2} \int d\chi \; \frac{\chi_s - \chi}{\chi_s \chi} \; \Phi \left[(\chi \boldsymbol{\theta}, \chi), \eta_0 - \chi \right]; \qquad \text{with } \chi_s > \chi. \tag{28}$$

The shear $\gamma(\boldsymbol{\theta}) = \gamma_1(\boldsymbol{\theta}) + i \gamma_2(\boldsymbol{\theta})$ at a given location $\boldsymbol{\theta}$ is a complex quantity where the shear components γ_1 and γ_2 are specified in a chosen Cartesian frame (in 2D flat-sky). However, one is free to rotate the coordinates by any arbitrary angle β. With respect to this reference rotation angle β, one defines the *tangential* and *cross* components of the shear at position $\boldsymbol{\theta}$ as (Schneider 2006)

$$\gamma_t(\boldsymbol{\theta}, \beta) + i \gamma_\times(\boldsymbol{\theta}, \beta) \equiv -e^{-2i\beta} \left[\gamma_1(\boldsymbol{\theta}) + i \gamma_2(\boldsymbol{\theta}) \right]. \tag{29}$$

Now, given a pair of points $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ on the field that are separated by the 2D vector $\boldsymbol{\alpha} \equiv \boldsymbol{\theta}_2 - \boldsymbol{\theta}_1$, one can write the tangential and cross components of the shear for this particular pair of points along the separation direction $\beta = \phi_{\boldsymbol{\alpha}}$ (polar angle of $\boldsymbol{\alpha}$) as

$$\gamma_t(\boldsymbol{\theta}_j, \phi_{\boldsymbol{\alpha}}) + i \gamma_\times(\boldsymbol{\theta}_j, \phi_{\boldsymbol{\alpha}}) \equiv -e^{-2i\phi_{\boldsymbol{\alpha}}} \left[\gamma_1(\boldsymbol{\theta}_j) + i \gamma_2(\boldsymbol{\theta}_j) \right], \tag{30}$$

where $j = 1, 2$.

In the 2D Fourier plane, the shear $\gamma(\boldsymbol{l})$ is related to $\kappa(\boldsymbol{l})$ as (Schneider 2006; Kilbinger 2015)

$$\gamma(\boldsymbol{l}) = \frac{(l_x + i \, l_y)^2}{l^2} \kappa(\boldsymbol{l}) = e^{2i\phi_l} \kappa(\boldsymbol{l}); \qquad \text{for } l \neq 0, \tag{31}$$

where $l = \sqrt{l_x^2 + l_y^2}$ and $\phi_l = \arctan\left(\frac{l_y}{l_x}\right)$ is the polar angle of $\boldsymbol{l}$.

The weak lensing convergence power spectrum $P_{\kappa, gh}$ can be defined through equation (3) – $\left\langle \kappa_g(\boldsymbol{l}_1) \kappa_h(\boldsymbol{l}_2) \right\rangle \equiv (2\pi)^2 \delta_D(\boldsymbol{l}_1 + \boldsymbol{l}_2) P_{\kappa, gh}(\boldsymbol{l}_1)$ for the convergence fields κ_g and κ_h, each defined with projection kernels $q_g(\chi)$ and $q_h(\chi)$ for two different redshift bins (see equation 25) with sources located at $\chi_{s, g}$ and $\chi_{s, h}$, respectively. It can be further expressed through equation (5) as

$$P_{\kappa, gh}(\boldsymbol{l}) = \int d\chi \frac{q_g(\chi) q_h(\chi)}{\chi^2} P_\delta^{3D} \left(\boldsymbol{k} = \frac{\boldsymbol{l}}{\chi}, \eta_0 - \chi \right), \tag{32}$$

[7] In this paper, we only consider the case when all source galaxies are located in a Dirac-δ function like bin at χ_s. However, it is straightforward to write $q(\chi)$ for a general distribution of source galaxies in a tomographic redshift bin (e.g. see Schneider 2006).

where $P_\delta^{3D}(\boldsymbol{k}, \eta)$ is the 3D matter density contrast power spectrum.

Similarly, the weak lensing convergence bispectrum defined through equation (4) – $\langle \kappa_f(\boldsymbol{l}_1)\kappa_g(\boldsymbol{l}_2)\kappa_h(\boldsymbol{l}_3) \rangle \equiv (2\pi)^2 \delta_D(\boldsymbol{l}_1 + \boldsymbol{l}_2 + \boldsymbol{l}_3)B_{\kappa,fgh}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3)$ of the convergence fields κ_f, κ_g, and κ_h with projection kernels $q_f(\chi)$, $q_g(\chi)$, and $q_h(\chi)$, respectively, can be expressed through equation (6) as

$$B_{\kappa,fgh}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3)$$
$$= \int d\chi \frac{q_f(\chi)q_g(\chi)q_h(\chi)}{\chi^4} B_\delta^{3D}\left(\frac{\boldsymbol{l}_1}{\chi}, \frac{\boldsymbol{l}_2}{\chi}, \frac{\boldsymbol{l}_3}{\chi}, \eta_0 - \chi\right), \quad (33)$$

where $B_\delta^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \eta)$ is the 3D bispectrum of the matter density contrast field and $\boldsymbol{k}_i = \frac{\boldsymbol{l}_i}{\chi}$. From the statistical isotropy of the density contrast field, both P_δ^{3D} and B_δ^{3D} are independent of the direction of the $\boldsymbol{k}_i$ wave vectors.

3.2 Shear two-point correlation function and aperture mass

A widely used statistic to investigate the shear field $\gamma(\boldsymbol{\theta})$ is the two-point shear correlation function. Using the notation $\gamma_{t,j} \equiv \gamma_t(\theta_j, \phi_\alpha)$ and $\gamma_{\times,j} \equiv \gamma_\times(\theta_j, \phi_\alpha)$, the two-point shear correlations (as a function of separation vector $\boldsymbol{\alpha}$) are defined as (Schneider & Lombardi 2003; Jarvis, Bernstein & Jain 2004)

$$\xi_+(\boldsymbol{\alpha}) \equiv \langle \gamma_{t,1}\, \gamma_{t,2} \rangle + \langle \gamma_{\times,1}\, \gamma_{\times,2} \rangle = \langle \gamma(\boldsymbol{\theta}_1)\gamma^*(\boldsymbol{\theta}_2) \rangle,$$
$$\xi_-(\boldsymbol{\alpha}) \equiv \langle \gamma_{t,1}\, \gamma_{t,2} \rangle - \langle \gamma_{\times,1}\, \gamma_{\times,2} \rangle = \langle \gamma(\boldsymbol{\theta}_1)\gamma(\boldsymbol{\theta}_2)\, e^{-4i\phi_\alpha} \rangle, \quad (34)$$

where the ensemble averages are over all pairs of points $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2\}$ with $\boldsymbol{\theta}_2 = \boldsymbol{\theta}_1 + \boldsymbol{\alpha}$.

Considering a pair of shear fields γ_g and γ_h with projection kernels $q_g(\chi)$ and $q_h(\chi)$, respectively, the shear two-point cross-correlations $\xi_{\pm, gh}$ between the two fields can then be written as

$$\xi_{+,gh}(\boldsymbol{\alpha}) \equiv \langle \gamma_g(\boldsymbol{\theta})\gamma_h^*(\boldsymbol{\theta} + \boldsymbol{\alpha}) \rangle,$$
$$\xi_{-,gh}(\boldsymbol{\alpha}) \equiv \langle \gamma_g(\boldsymbol{\theta})\gamma_h(\boldsymbol{\theta} + \boldsymbol{\alpha})\, e^{-4i\phi_\alpha} \rangle. \quad (35)$$

In general, both the correlations are complex quantities but have vanishing imaginary parts only for the so-called E-mode shear fields (which we consider in this paper; Schneider, van Waerbeke & Mellier 2002; Kilbinger 2015). Moreover, from statistical isotropy of the fields it follows that $\xi_{\pm,gh}(\boldsymbol{\alpha}) = \xi_{\pm,gh}(\alpha)$. These shear correlations are related to the convergence power spectrum (equation 32) through inverse Hankel transforms (see Appendix A; Schneider 2006; Kilbinger 2015):

$$\xi_{+,gh}(\alpha) = \mathcal{F}_{2D}^{-1}[P_{\kappa,gh}(l)] = \int \frac{dl\, l}{2\pi} P_{\kappa,gh}(l)\, J_0(l\alpha),$$
$$\xi_{-,gh}(\alpha) = \mathcal{F}_{2D}^{-1}[P_{\kappa,gh}(l)\, e^{-4i\phi_l}] = \int \frac{dl\, l}{2\pi} P_{\kappa,gh}(l)\, J_4(l\alpha), \quad (36)$$

where $J_0(x)$ and $J_4(x)$ are the zeroth- and fourth-order Bessel functions of the first kind, respectively.

We can now write the position-dependent two-point correlation functions $\hat{\xi}_{\pm,gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)$ of the shear field within a 2D window W centred at position $\boldsymbol{\theta}_C$. Using equations (17) and (31) and the first line of equation (35), we can write the $\hat{\xi}_{+,gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)$ correlation as

$$\hat{\xi}_{+,gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) \equiv \frac{1}{A_{2pt}(\boldsymbol{\alpha})} \int d^2\boldsymbol{\theta}\, \gamma_g(\boldsymbol{\theta}; \boldsymbol{\theta}_C)\gamma_h^*(\boldsymbol{\theta} + \boldsymbol{\alpha}; \boldsymbol{\theta}_C)$$
$$= \frac{1}{A_{2pt}(\boldsymbol{\alpha})} \int \frac{d^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_2}{(2\pi)^2} \int \frac{d^2\boldsymbol{q}}{(2\pi)^2}$$

$$\times \kappa_g(\boldsymbol{l}_1)\kappa_h(\boldsymbol{l}_2) \times e^{2i(\phi_1-\phi_2)} W(\boldsymbol{q})W(\boldsymbol{l}_1 + \boldsymbol{l}_2 - \boldsymbol{q})$$
$$\times e^{i(\boldsymbol{l}_1+\boldsymbol{l}_2)\cdot\boldsymbol{\theta}_C}\, e^{i(\boldsymbol{q}-\boldsymbol{l}_1)\cdot\boldsymbol{\alpha}}, \quad (37)$$

where ϕ_1 and ϕ_2 are the polar angles of the Fourier modes $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$, respectively.

Taking into account the phase factor $e^{-4i\phi_\alpha}$ present in the second line of equation (35), we can write the $\hat{\xi}_{-,gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)$ using equations (15) and (31) as

$$\hat{\xi}_{-,gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) \equiv \frac{1}{A_{2pt}(\boldsymbol{\alpha})} \int d^2\boldsymbol{\theta}\, \gamma_g(\boldsymbol{\theta}; \boldsymbol{\theta}_C)\gamma_h(\boldsymbol{\theta} + \boldsymbol{\alpha}; \boldsymbol{\theta}_C)\, e^{-4i\phi_\alpha}$$
$$= \frac{1}{A_{2pt}(\boldsymbol{\alpha})} \int \frac{d^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_2}{(2\pi)^2} \int \frac{d^2\boldsymbol{q}}{(2\pi)^2} \kappa_g(\boldsymbol{l}_1)\kappa_h(\boldsymbol{l}_2)$$
$$\times e^{2i(\phi_1+\phi_2)} W(\boldsymbol{q})W(\boldsymbol{l}_1 + \boldsymbol{l}_2 - \boldsymbol{q})\, e^{i(\boldsymbol{l}_1+\boldsymbol{l}_2)\cdot\boldsymbol{\theta}_C}$$
$$\times e^{i(\boldsymbol{q}-\boldsymbol{l}_1)\cdot\boldsymbol{\alpha}}\, e^{-4i\phi_\alpha}. \quad (38)$$

For isotropic window functions W, both the estimators are independent of the direction of $\boldsymbol{\alpha}$, i.e. $\hat{\xi}_{\pm,gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) = \hat{\xi}_{\pm,gh}(\alpha; \boldsymbol{\theta}_C)$. Moreover, taking the ensemble average of the above equations we can see that $\langle \hat{\xi}_{\pm,gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) \rangle = \xi_{\pm,gh}(\boldsymbol{\alpha})$.

Along these lines we can also define the position-dependent shear power spectra expressions as the Fourier space counterparts of the above equations. Using equations (19) and (18), respectively (with an extra phase factor $e^{4i\phi_\alpha}$ in the latter), we get

$$\hat{P}_{+,gh}(\boldsymbol{l}; \boldsymbol{\theta}_C) \equiv \mathcal{F}_{2D}\left[A_{2pt}(\boldsymbol{\alpha})\hat{\xi}_{+,gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)\right]$$
$$= \gamma_g(-\boldsymbol{l}; \boldsymbol{\theta}_C)\gamma_h^*(-\boldsymbol{l}; \boldsymbol{\theta}_C)$$
$$= \int \frac{d^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_2}{(2\pi)^2} \kappa_g(\boldsymbol{l}_1)\kappa_h(\boldsymbol{l}_2)\, e^{2i(\phi_1-\phi_2)}$$
$$\times W(\boldsymbol{l}_1 + \boldsymbol{l})W(\boldsymbol{l}_2 - \boldsymbol{l})\, e^{i(\boldsymbol{l}_1+\boldsymbol{l}_2)\cdot\boldsymbol{\theta}_C} \quad (39)$$

and

$$\hat{P}_{-,gh}(\boldsymbol{l}; \boldsymbol{\theta}_C) \equiv \mathcal{F}_{2D}\left[A_{2pt}(\boldsymbol{\alpha})\hat{\xi}_{-,gh}(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)\, e^{4i\phi_\alpha}\right]$$
$$= \gamma_g(-\boldsymbol{l}; \boldsymbol{\theta}_C)\gamma_h(\boldsymbol{l}; \boldsymbol{\theta}_C)$$
$$= \int \frac{d^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{d^2\boldsymbol{l}_2}{(2\pi)^2} \kappa_g(\boldsymbol{l}_1)\kappa_h(\boldsymbol{l}_2)\, e^{2i(\phi_1+\phi_2)}$$
$$\times W(\boldsymbol{l}_1 + \boldsymbol{l})W(\boldsymbol{l}_2 - \boldsymbol{l})\, e^{i(\boldsymbol{l}_1+\boldsymbol{l}_2)\cdot\boldsymbol{\theta}_C}. \quad (40)$$

In this paper, we shall use a top-hat (disc) window function W of radius θ_T inside which we shall evaluate the two-point shear correlations:

$$W(\boldsymbol{\theta}) = W(\theta) = \begin{cases} 1 & \theta \leq \theta_T, \\ 0 & \theta > \theta_T, \end{cases} \quad (41)$$

and the Fourier transform of this window function reads

$$W(\boldsymbol{l}) = W(l) = \int d^2\boldsymbol{\theta}\, W(\theta)\, e^{-i\boldsymbol{l}\cdot\boldsymbol{\theta}} = 2\pi\theta_T^2 \frac{J_1(l\theta_T)}{l\theta_T}, \quad (42)$$

where J_1 is the first-order ordinary Bessel function of the first kind. One should note that this form of the top-hat window function is not normalized since $\int d^2\boldsymbol{\theta}\, W(\theta) = \pi\theta_T^2$.

Another statistic used for investigating the convergence/shear field is the aperture mass $M_{ap}(\boldsymbol{\theta}_C)$ that measures the weighted κ – a projected surface mass – inside an aperture U located at a given

point $\boldsymbol{\theta}_C$ (Kaiser 1995; Schneider 1996, 2006):

$$M_{\mathrm{ap}}(\boldsymbol{\theta}_C) = \int \mathrm{d}^2\boldsymbol{\theta}\, \kappa(\boldsymbol{\theta})\, U(\boldsymbol{\theta}_C - \boldsymbol{\theta})$$

$$= \int \frac{\mathrm{d}^2\boldsymbol{l}}{(2\pi)^2}\, \kappa(\boldsymbol{l})U(\boldsymbol{l})\, \mathrm{e}^{i\boldsymbol{l}\cdot\boldsymbol{\theta}_C}, \tag{43}$$

where the azimuthally symmetric aperture $U(\boldsymbol{\theta}) = U(\theta)$ has a characteristic size scale θ_{ap} and in the second line we have expanded the equation with Fourier space expressions (see equation 9). Furthermore, if U is a compensated window function, i.e. its integral over its support vanishes $\int \mathrm{d}^2\boldsymbol{\theta}\, U(\boldsymbol{\theta}_C - \boldsymbol{\theta}) = 0$, then a very interesting property of the aperture mass is that it can be directly evaluated from the shear field as a weighted tangential shear within an azimuthally symmetric aperture Q (of size θ_{ap}) located at $\boldsymbol{\theta}_C$ (Kaiser 1995; Schneider 1996, 2006):

$$M_{\mathrm{ap}}(\boldsymbol{\theta}_C) = \int \mathrm{d}^2\boldsymbol{\theta}\, \gamma_{\mathrm{t}}(\boldsymbol{\theta}, \phi_{\boldsymbol{\theta}_C - \boldsymbol{\theta}})\, Q(\boldsymbol{\theta}_C - \boldsymbol{\theta}), \tag{44}$$

where the tangential shear $\gamma_{\mathrm{t}}(\boldsymbol{\theta}, \phi_{\boldsymbol{\theta}_C - \boldsymbol{\theta}})$ at any given location $\boldsymbol{\theta}$ is defined with respect to $\phi_{\boldsymbol{\theta}_C - \boldsymbol{\theta}}$ that is the polar angle of the separation vector between $\boldsymbol{\theta}$ and the centre of the aperture $\boldsymbol{\theta}_C$. The azimuthally symmetric aperture Q has the form (Schneider 2006)

$$Q(\boldsymbol{\theta}) = Q(\theta) = -U(\theta) + \frac{2}{\theta^2} \int_0^\theta \mathrm{d}\theta'\, \theta' U(\theta'). \tag{45}$$

The aperture mass statistic can be interpreted as a position-dependent weighted mean of the shear/convergence field (see equation 9) with $W_{\mathrm{1pt}} = U$. However, as we define it using a compensated filter, an area normalization term for this statistic is irrelevant (see equation 10).

For the filter functions Q and U, several choices have been investigated. In this paper, we use the forms proposed by Crittenden et al. (2002; see also Kilbinger & Schneider 2005; Schneider et al. 2005):

$$U(\theta) = \frac{1}{2\pi\theta_{\mathrm{ap}}^2} \left(1 - \frac{\theta^2}{2\theta_{\mathrm{ap}}^2}\right) \exp\left(-\frac{\theta^2}{2\theta_{\mathrm{ap}}^2}\right),$$

$$Q(\theta) = \frac{\theta^2}{4\pi\theta_{\mathrm{ap}}^4} \exp\left(-\frac{\theta^2}{2\theta_{\mathrm{ap}}^2}\right). \tag{46}$$

We shall also work closely with the Fourier space representation of U for our theoretical modelling:

$$U(\boldsymbol{l}) = U(l) = \int \mathrm{d}^2\boldsymbol{\theta}\, U(\theta)\, \mathrm{e}^{-i\boldsymbol{l}\cdot\boldsymbol{\theta}} = \frac{l^2\theta_{\mathrm{ap}}^2}{2} \exp\left(-\frac{l^2\theta_{\mathrm{ap}}^2}{2}\right). \tag{47}$$

3.3 Integrated three-point shear correlation function

We now have all the necessary ingredients to define the integrated three-point function (see Section 2.3.5) of the cosmic shear field as follows:

$$i\zeta_{\pm,\mathrm{fgh}}(\boldsymbol{\alpha}) \equiv \left\langle M_{\mathrm{ap,f}}(\boldsymbol{\theta}_C)\, \hat{\xi}_{\pm,\mathrm{gh}}(\boldsymbol{\alpha};\boldsymbol{\theta}_C) \right\rangle, \tag{48}$$

where $M_{\mathrm{ap,f}}(\boldsymbol{\theta}_C)$ is the aperture mass at location $\boldsymbol{\theta}_C$ (see equations 43 and 44) evaluated from the shear field γ_{f} with projection kernel $q_{\mathrm{f}}(\chi)$, and $\hat{\xi}_{\pm,\mathrm{gh}}(\boldsymbol{\alpha};\boldsymbol{\theta}_C)$ are the position-dependent shear two-point correlation functions (see equations 37 and 38) computed inside a top-hat patch centred at $\boldsymbol{\theta}_C$ from fields γ_{g} and γ_{h} with projection kernels $q_{\mathrm{g}}(\chi)$ and $q_{\mathrm{h}}(\chi)$, respectively. Note again that each of these projection kernels indicates source redshifts corresponding to different comoving distances $\chi_{s,\mathrm{f}}$, $\chi_{s,\mathrm{g}}$, and $\chi_{s,\mathrm{h}}$, respectively.

Using equations (21), (43), and (37), we can write the expression for the $i\zeta_+$ correlation function as

$$i\zeta_{+,\mathrm{fgh}}(\boldsymbol{\alpha}) \equiv \left\langle M_{\mathrm{ap,f}}(\boldsymbol{\theta}_C)\, \hat{\xi}_{+,\mathrm{gh}}(\boldsymbol{\alpha};\boldsymbol{\theta}_C) \right\rangle$$

$$= \frac{1}{A_{\mathrm{2pt}}(\boldsymbol{\alpha})} \int \mathrm{d}^2\boldsymbol{\theta}_1 \int \mathrm{d}^2\boldsymbol{\theta}_2 \left\langle \kappa_{\mathrm{f}}(\boldsymbol{\theta}_1)\gamma_{\mathrm{g}}(\boldsymbol{\theta}_2)\gamma_{\mathrm{h}}^*(\boldsymbol{\theta}_2 + \boldsymbol{\alpha}) \right\rangle$$

$$\times U(\boldsymbol{\theta}_C - \boldsymbol{\theta}_1)W(\boldsymbol{\theta}_C - \boldsymbol{\theta}_2)W(\boldsymbol{\theta}_C - \boldsymbol{\theta}_2 - \boldsymbol{\alpha})$$

$$= \frac{1}{A_{\mathrm{2pt}}(\boldsymbol{\alpha})} \int \frac{\mathrm{d}^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{\mathrm{d}^2\boldsymbol{l}_2}{(2\pi)^2} \int \frac{\mathrm{d}^2\boldsymbol{q}}{(2\pi)^2}$$

$$\times B_{\kappa,\mathrm{fgh}}(\boldsymbol{l}_1, \boldsymbol{l}_2, -\boldsymbol{l}_1 - \boldsymbol{l}_2)\, \mathrm{e}^{2i(\phi_2 - \phi_{-1-2})}$$

$$\times U(\boldsymbol{l}_1)W(\boldsymbol{q})W(-\boldsymbol{l}_1 - \boldsymbol{q})\, \mathrm{e}^{i(\boldsymbol{q} - \boldsymbol{l}_2)\cdot\boldsymbol{\alpha}}, \tag{49}$$

where ϕ_{-1-2} is the polar angle of the $-\boldsymbol{l}_1 - \boldsymbol{l}_2$ 2D Fourier mode and in the last equality we have used the definition of the convergence bispectrum B_κ that can be further expressed in terms of a line-of-sight projection of the 3D matter density bispectrum using equation (33) to obtain

$$i\zeta_{+,\mathrm{fgh}}(\boldsymbol{\alpha}) = \frac{1}{A_{\mathrm{2pt}}(\boldsymbol{\alpha})} \int \mathrm{d}\chi \frac{q_{\mathrm{f}}(\chi)q_{\mathrm{g}}(\chi)q_{\mathrm{h}}(\chi)}{\chi^4} \int \frac{\mathrm{d}^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{\mathrm{d}^2\boldsymbol{l}_2}{(2\pi)^2}$$

$$\times \int \frac{\mathrm{d}^2\boldsymbol{q}}{(2\pi)^2}\, B_\delta^{\mathrm{3D}}\left(\frac{\boldsymbol{l}_1}{\chi}, \frac{\boldsymbol{l}_2}{\chi}, \frac{-\boldsymbol{l}_1 - \boldsymbol{l}_2}{\chi}, \eta_0 - \chi\right)$$

$$\times \mathrm{e}^{2i(\phi_2 - \phi_{-1-2})}$$

$$\times U(\boldsymbol{l}_1)W(\boldsymbol{q})W(-\boldsymbol{l}_1 - \boldsymbol{q})\, \mathrm{e}^{i(\boldsymbol{q} - \boldsymbol{l}_2)\cdot\boldsymbol{\alpha}}. \tag{50}$$

Similarly, using equation (38) the $i\zeta_-$ correlation reads

$$i\zeta_{-,\mathrm{fgh}}(\boldsymbol{\alpha}) \equiv \left\langle M_{\mathrm{ap,f}}(\boldsymbol{\theta}_C)\, \hat{\xi}_{-,\mathrm{gh}}(\boldsymbol{\alpha};\boldsymbol{\theta}_C) \right\rangle$$

$$= \frac{1}{A_{\mathrm{2pt}}(\boldsymbol{\alpha})} \int \mathrm{d}^2\boldsymbol{\theta}_1 \int \mathrm{d}^2\boldsymbol{\theta}_2 \left\langle \kappa_{\mathrm{f}}(\boldsymbol{\theta}_1)\gamma_{\mathrm{g}}(\boldsymbol{\theta}_2)\gamma_{\mathrm{h}}(\boldsymbol{\theta}_2 + \boldsymbol{\alpha}) \right\rangle$$

$$\times \mathrm{e}^{-4i\phi_\alpha} U(\boldsymbol{\theta}_C - \boldsymbol{\theta}_1)W(\boldsymbol{\theta}_C - \boldsymbol{\theta}_2)W(\boldsymbol{\theta}_C - \boldsymbol{\theta}_2 - \boldsymbol{\alpha})$$

$$= \frac{1}{A_{\mathrm{2pt}}(\boldsymbol{\alpha})} \int \mathrm{d}\chi \frac{q_{\mathrm{f}}(\chi)q_{\mathrm{g}}(\chi)q_{\mathrm{h}}(\chi)}{\chi^4} \int \frac{\mathrm{d}^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{\mathrm{d}^2\boldsymbol{l}_2}{(2\pi)^2}$$

$$\times \int \frac{\mathrm{d}^2\boldsymbol{q}}{(2\pi)^2}\, B_\delta^{\mathrm{3D}}\left(\frac{\boldsymbol{l}_1}{\chi}, \frac{\boldsymbol{l}_2}{\chi}, \frac{-\boldsymbol{l}_1 - \boldsymbol{l}_2}{\chi}, \eta_0 - \chi\right)$$

$$\times \mathrm{e}^{2i(\phi_2 + \phi_{-1-2})}$$

$$\times U(\boldsymbol{l}_1)W(\boldsymbol{q})W(-\boldsymbol{l}_1 - \boldsymbol{q})\, \mathrm{e}^{i(\boldsymbol{q} - \boldsymbol{l}_2)\cdot\boldsymbol{\alpha}}\, \mathrm{e}^{-4i\phi_\alpha}. \tag{51}$$

As stated before, for isotropic window functions, these correlations are independent of the direction of $\boldsymbol{\alpha}$, i.e. $i\zeta_{\pm,\mathrm{fgh}}(\boldsymbol{\alpha}) = i\zeta_{\pm,\mathrm{fgh}}(\alpha)$. One thing to note is the similarity between the expressions of $i\zeta_{\pm,\mathrm{fgh}}$ and the generalized third-order aperture mass statistics with different compensated filter radii as proposed by of Schneider et al. (2005, see their section 6). Our expressions can be interpreted as a special case of these generalized aperture mass statistics where we use two top-hat filters of same radii and one compensated filter with a different size instead of using three compensated filters.

Computationally, it is more convenient to arrive at these expressions for the integrated three-point shear correlation functions from the inverse Fourier transforms of the integrated shear bispectra that we define as

$$i B_{+,\mathrm{fgh}}(\boldsymbol{l}) \equiv \mathcal{F}_{\mathrm{2D}}\left[A_{\mathrm{2pt}}(\boldsymbol{\alpha})i\zeta_{+,\mathrm{fgh}}(\boldsymbol{\alpha})\right],$$

$$i B_{-,\mathrm{fgh}}(\boldsymbol{l}) \equiv \mathcal{F}_{\mathrm{2D}}\left[A_{\mathrm{2pt}}(\boldsymbol{\alpha})i\zeta_{-,\mathrm{fgh}}(\boldsymbol{\alpha})\, \mathrm{e}^{4i\phi_\alpha}\right]. \tag{52}$$

Upon simplification, the expressions for these integrated bispectra read

$$
\begin{aligned}
i\, B_{\pm,\mathrm{fgh}}(\boldsymbol{l}) = {} & \int \mathrm{d}\chi \, \frac{q_{\mathrm{f}}(\chi) q_{\mathrm{g}}(\chi) q_{\mathrm{h}}(\chi)}{\chi^4} \int \frac{\mathrm{d}^2 l_1}{(2\pi)^2} \int \frac{\mathrm{d}^2 l_2}{(2\pi)^2} \\
& \times B_\delta^{\mathrm{3D}}\left(\frac{\boldsymbol{l}_1}{\chi}, \frac{\boldsymbol{l}_2}{\chi}, \frac{-\boldsymbol{l}_1 - \boldsymbol{l}_2}{\chi}, \eta_0 - \chi\right) \mathrm{e}^{2\mathrm{i}(\phi_2 \mp \phi_{-1-2})} \\
& \times U(\boldsymbol{l}_1) W(\boldsymbol{l}_2 + \boldsymbol{l}) W(-\boldsymbol{l}_1 - \boldsymbol{l}_2 - \boldsymbol{l}) \\
= {} & \left\langle M_{\mathrm{ap,f}}(\boldsymbol{\theta}_C)\, \hat{P}_{\pm,\mathrm{gh}}(\boldsymbol{l}; \boldsymbol{\theta}_C)\right\rangle,
\end{aligned} \tag{53}
$$

where the last equality confirms our expectation (see equations 22 and 23) that the integrated bispectrum of the shear field is the correlation of the aperture mass and the position-dependent shear power spectrum.

Because of the isotropy argument, we have $i\,B_{\pm,\mathrm{fgh}}(\boldsymbol{l}) = i\,B_{\pm,\mathrm{fgh}}(l)$, i.e. the above equation is true for any polar angle ϕ_l and the integrated three-point functions are then inverse Hankel transforms of these integrated bispectra:

$$
\begin{aligned}
i\zeta_{+,\mathrm{fgh}}(\alpha) &= \frac{1}{A_{\mathrm{2pt}}(\alpha)} \mathcal{F}_{\mathrm{2D}}^{-1}\left[i\,B_{+,\mathrm{fgh}}(l)\right] \\
&= \frac{1}{A_{\mathrm{2pt}}(\alpha)} \int \frac{\mathrm{d}l\, l}{2\pi} \, i\,B_{+,\mathrm{fgh}}(l)\, J_0(l\alpha), \\
i\zeta_{-,\mathrm{fgh}}(\alpha) &= \frac{1}{A_{\mathrm{2pt}}(\alpha)} \mathcal{F}_{\mathrm{2D}}^{-1}\left[i\,B_{-,\mathrm{fgh}}(l)\, \mathrm{e}^{-4\mathrm{i}\phi_l}\right] \\
&= \frac{1}{A_{\mathrm{2pt}}(\alpha)} \int \frac{\mathrm{d}l\, l}{2\pi} \, i\,B_{-,\mathrm{fgh}}(l)\, J_4(l\alpha).
\end{aligned} \tag{54}
$$

The $J_0(l\alpha)$ filter puts more weight on low-l values of the integrated bispectrum than the $J_4(l\alpha)$ filter at a given angular separation α. Hence, $i\zeta_+(\alpha)$ is more sensitive to large-scale fluctuations (lower l) than $i\zeta_-(\alpha)$ at the same angular separation α.

3.4 Summary

So far, we have developed the following.

(i) The integrated three-point shear correlation function $i\zeta_\pm$ can be estimated from the cosmic shear field by measuring the aperture mass statistic (with a compensated filter) at different locations and then correlating it with the position-dependent two-point shear correlation function (evaluated inside top-hat apertures) located at the corresponding locations (see equation 48).

(ii) Given a prescription of the 3D matter density bispectrum $B_\delta^{\mathrm{3D}}(k_1, k_2, k_3, \eta)$ for a set of cosmological parameters, we can theoretically predict the $i\zeta_\pm$ through an inverse Hankel transform of the integrated shear bispectrum $iB_\pm$ – an integral of the convergence bispectrum (see equations 33, 53, and 54). This is analogous to the way in which one obtains the shear two-point correlation function $\xi_\pm$ from the convergence power spectrum that is in turn related to the 3D matter density power spectrum $P_\delta^{\mathrm{3D}}(k, \eta)$ through a line-of-sight projection (see equations 32 and 36).

(iii) In Section 2, we provide a general framework of equations for the integrated three-point function (equations 20 and 21) and the integrated bispectrum (equations 22 and 23) that can be extended to the analysis of any projected field within the flat-sky approximation.

We shall now proceed to measure the $\xi_\pm$ and $i\zeta_\pm$ statistics on simulated cosmic shear data and also perform theoretical calculations for the same using the equations mentioned above. We will test the accuracy of our models on the simulations and then investigate their constraining power on cosmological parameters.

4 SIMULATIONS, MEASUREMENTS, AND NUMERICAL METHODS FOR THEORETICAL MODELLING

In this section, we describe the simulations (Sections 4.1 and 4.2) we use in order to measure our data vector and the data-covariance matrix (Section 4.3). We will then discuss the methods we use in order to theoretically model the data vector in Section 4.4.

4.1 Takahashi et al. (2017) N-body simulations

We use the publicly available simulated data sets from Takahashi et al. (2017, hereafter T17)[8] cosmological simulations. The simulations were generated primarily for the gravitational lensing studies for the HSC Survey. In this paper, we use the full-sky light-cone weak lensing shear and convergence maps of the simulation suite.

These data sets were obtained from a cold dark matter (CDM) only cosmological N-body simulation in periodic cubic boxes. The simulation setting consisted of 14 boxes of increasing side lengths L, $2L$, $3L$, ..., $14L$ (with $L = 450$ Mpc h^{-1}), nested around a common vertex (see fig. 1 of T17). Each box contained 2048^3 particles (smaller boxes hence have better spatial and mass resolution) and their initial conditions were set with second-order Lagrangian perturbation theory (Crocce, Pueblas & Scoccimarro 2006) with an initial power spectrum computed for a flat ΛCDM cosmology with the following parameters:[9] $\Omega_{\mathrm{cdm}} = 0.233$, $\Omega_{\mathrm{b}} = 0.046$, $\Omega_{\mathrm{m}} = \Omega_{\mathrm{cdm}} + \Omega_{\mathrm{b}} = 0.279$, $\Omega_{\mathrm{de}} = \Omega_\Lambda = 0.721$, $h = 0.7$, $\sigma_8 = 0.82$, and $n_{\mathrm{s}} = 0.97$. We adopt this set of parameters as our fiducial cosmology. The particles in each box were then made to evolve from the initial conditions using the N-body gravity solver code GADGET2 (Springel, Yoshida & White 2001; Springel 2005). The evolved particle distribution of the different nested boxes was combined in layers of shells, each 150 Mpc h^{-1} thick, to obtain full-sky light-cone matter density contrast inside the shells. The simulation boxes were also ray traced using the multiple-lens plane ray-tracing algorithm GRAYTRIX (Hamana et al. 2015; Shirasaki, Hamana & Yoshida 2015) to obtain full-sky weak lensing convergence/shear maps (in HEALPIX format; Górski et al. 2005; Zonca et al. 2019) for several Dirac-δ-like source redshift bins. Multiple simulations were run to produce 108 realizations for each of their data products. The authors report that the average matter power spectra from their several realizations of the simulations agreed with the theoretical revised HALOFIT power spectrum (Smith et al. 2003, later revised by Takahashi et al. 2012) to within 5 (10) per cent for $k < 5(6)$ h^{-1} Mpc at $z < 1$. They also provide correction formulae for their 3D and angular power spectra in order to account for the discrepancies stemming from the finite shell thickness, angular resolution, and finite simulation box size effect in their simulations. We refer the reader to our Appendix B for a summary of those corrections.

In this paper, for validating the two-point and integrated three-point shear correlation functions (see Section 4.3) we use the 108 full-sky weak lensing convergence and shear maps from the simulation suite. These maps come in the HEALPIX format (Górski et al. 2005; Zonca et al. 2019) for various angular resolutions. We only use the maps with NSIDE = 4096 (angular pixel scale of 0.82 arcmin) at source redshifts $z_1 = 0.5739$ and $z_2 = 1.0334$. For reference, these two redshifts correspond closely to the mean redshifts of the second

---

[8]The data products of the simulation are available at http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/.
[9]The density parameter for species X is defined at $\eta = \eta_0$, i.e. $\Omega_X \equiv \Omega_{X, 0}$.

and fourth photometric source redshift bins that have been used in the cosmic shear two-point analyses of the DES (Troxel et al. 2018).

4.2 FLASK lognormal simulations

A crucial part of any cosmological analysis involves the calculation of the covariance matrix of a data vector – which for us shall consist of two-point and integrated three-point correlations of the shear field (see Section 4.3). The estimation of the inverse of this data-covariance, namely the precision matrix, is particularly important for forecasting cosmological parameter constraints (e.g. see Section 5.2). Although an analytically modelled data-covariance matrix can be inverted easily as it is inherently noise-free, it needs to be modelled sufficiently accurately. An easier approach is to estimate the covariance for a desired data vector from an ensemble of realistic N-body simulations. However, this comes at a cost that the sample covariance suffers from noise when estimated from a finite number of mock simulations. The inversion of such a noisy matrix comes with its own challenges. In order to beat down this noise in the precision matrix one therefore needs a large ensemble of independent simulations – with the number of simulations required to be much larger than the size of the data vector (see Taylor, Joachimi & Kitching 2013). Unfortunately, for our purpose, we have only 108 independent T17 simulations to estimate the data-covariance of our quite high-dimensional data vector that will result in a noisy covariance matrix estimate (see Appendix F). Hence we need another way to estimate the covariance. Many possible methods to circumvent the problem have been suggested in literature such as resampling techniques for estimating the covariance matrix using a few mocks (Escoffier et al. 2016), shrinkage estimators (Joachimi 2017), or to use lognormal simulations to name a few.

We choose the option of simulating a large ensemble of full-sky lognormal random fields for estimating the data-covariance matrix. Lognormal random fields have been extensively studied in the cosmological context (Coles & Jones 1991) and have been shown to be a very good approximation for the one-point probability density function (PDF) of the weak lensing convergence/shear (Hilbert, Hartlap & Schneider 2011; Xavier, Abdalla & Joachimi 2016) or the distribution of the late time matter density contrast fields (Friedrich et al. 2018; Gruen et al. 2018). This assumption has been confirmed from the DES Science Verification data for the convergence field (Clerkin et al. 2017) and most recently been used to compute covariances for the two-point shear correlations for the third year data analysis of the DES (Friedrich et al. 2020). We further discuss and test the validity of modelling the data-covariance matrix with lognormal simulations in Appendix F. We show that our lognormal data-covariance and its inverse are indeed a good model as the Fisher parameter constraints shown in Section 5.2 are hardly affected when we correct the lognormal model with residual terms measured from the T17 simulations.

We use the publicly available Full-sky Lognormal Astro-fields Simulation Kit (FLASK) tool[10] (Xavier et al. 2016) that can be used to create realizations of correlated lognormal fields on the celestial sphere at different redshifts. Concisely, FLASK draws from a lognormal variable κ with the PDF (Xavier et al. 2016):

$$p(\kappa) = \begin{cases} \frac{\exp\left(-\frac{1}{2\sigma^2}[\ln(\kappa+\lambda)-\mu]^2\right)}{\sqrt{2\pi}\sigma(\kappa+\lambda)} & \kappa > -\lambda, \\ 0 & \text{otherwise,} \end{cases} \quad (55)$$

---

[10]Currently hosted at http://www.astro.iag.usp.br/~flask/.

where μ and σ^2 are the mean and variance of the associated normal variable and λ is the lognormal shift parameter marking the lower limit for possible values that κ can realize. Using FLASK we create lognormal mocks of the T17 convergence/shear fields that approximately follow the one-point PDFs of the T17 maps at redshifts z_1 and z_2, respectively. As input to FLASK, one needs to provide the convergence power spectra $P_{\kappa, gh}(l)$ and the lognormal shift parameters λ_i for the two redshifts (with $g, h, i = 1, 2$). We obtain the power spectra by projecting the 3D matter density contrast power spectrum $P_\delta^{3D}(k, \eta)$ along the line of sight as described in equation (32). We use the open-source Boltzmann solver code CLASS[11] (Blas, Lesgourgues & Tram 2011; Lesgourgues 2011) for computing the non-linear $P_\delta^{3D}(k, \eta)$ in the fiducial T17 cosmology for which we use the revised HALOFIT prescription for the non-linear matter power spectrum (Smith et al. 2003; Bird, Viel & Haehnelt 2012; Takahashi et al. 2012) that is included in CLASS. For obtaining the lognormal shift parameters we follow the strategy of Hilbert et al. (2011) and fit the above form of the lognormal PDF to the one-point PDF of the T17 maps at both redshifts and get the following values[12]: $\lambda_1 = 0.012$ and $\lambda_2 = 0.031$. Using these settings we generate 1000 correlated FLASK pairs (each pair consists of two maps at source redshifts z_1 and z_2, respectively) of full-sky shear maps in HEALPIX format with NSIDE = 4096.

The T17 simulations are pure gravitational lensing shear/convergence maps without any noise. In real data, the shear is obtained from the measured ellipticities of background galaxies that – besides the gravitational shear effect – are subject to different sources of noise such as non-circular intrinsic ellipticities of the galaxies, measurement noise, noise from point spread function correction, etc. In our covariance matrix we want to include the effect of this *shape noise*. This is important when we want to forecast realistic constraints on cosmological parameters. In principle, this can be modelled by adding a complex noise term $N(\theta) = N_1(\theta) + iN_2(\theta)$ to the shear field $\gamma\theta) = \gamma_1(\theta) + i\gamma_2(\theta)$ (Pires et al. 2020), where θ represents a pixel on the HEALPIX shear map. The noise components N_1, N_2 can both be modelled as uncorrelated Gaussian variables with zero-mean and variance:

$$\sigma_N^2 = \frac{\sigma_\epsilon^2}{n_g A_{pix}}, \quad (56)$$

where A_{pix} is the area of the pixel at the given NSIDE, σ_ϵ is the dispersion of intrinsic galaxy ellipticities that we set to be 0.3 as found for weak lensing surveys (Leauthaud et al. 2007; Schrabback et al. 2018), and n_g is the number of observed galaxies per square arcminute for which we assume a value of 5 at each redshift bin. Note that for the two Dirac-δ source redshift bins we consider, this adds up to give 10 galaxies per square arcminutes that is in accordance with the expected number density of galaxies for the full DES Year 6 cosmic shear data. To every pixel in a FLASK generated shear map we add an independent draw of each Gaussian noise term. We then convert these noisy shear maps into noisy convergence maps on the curved sky using a Kaiser–Squires (KS; Kaiser & Squires 1993) mass map reconstruction method as described in section 2.1 of Gatti et al. (2020; see also Chang et al. 2018). This process of first adding

---

[11]Currently hosted at http://class-code.net. We use version v2.9.4 of the code.
[12]Precisely, we only fit the PDF to the first of the 108 T17 maps at both redshifts to obtain the quoted λ_i values. We have also tested the fits on other maps at each redshift and the values for the logshift parameters differ in only the third decimal place whose effect on the summary statistics evaluated from the corresponding FLASK maps is insignificant. The values for the other fit parameters are: $\mu_1 = -4.578$, $\mu_2 = -3.565$ and $\sigma_1^2 = 0.351$, $\sigma_2^2 = 0.205$.

noise to the shear field and then converting it to a convergence map is more accurate than the usually prevalent way of adding independent Gaussian noise to the pixels of the noiseless convergence field. This is because convergence at a given pixel is a convolution of the shear in several pixels around the desired location. This makes the noise in the convergence at a given pixel be correlated with the noise in neighbouring pixels. Although the KS method ensures this, the approach where uncorrelated Gaussian noise is added to the pixels of a noiseless convergence map directly does not account for it and is therefore not entirely accurate.

4.3 Measurements: data vector and data-covariance matrix

We carry out measurements of the position-dependent two-point shear correlations $\hat{\xi}_{\pm,\mathrm{gh}}(\alpha; \boldsymbol{\theta}_C)$ on the T17 and FLASK shear maps at source redshifts z_1 and z_2 (i.e. $g, h = 1, 2$) within top-hat windows W with radius $\theta_\mathrm{T} = 75$ arcmin. Approximately, this results in a circular patch of area 5 deg^2 (which is small enough for the flat-sky approximation to hold). We use the publicly available code TREECORR[13] (Jarvis et al. 2004) to measure these correlations in 20 log-spaced bins with angular separations $5 < \alpha < 140$ arcmin. To be precise, we execute TREECORR on those pixels of the map that lie within a disc of radius θ_T centred at a given location $\boldsymbol{\theta}_C$ in order to obtain $\hat{\xi}_{\pm,\mathrm{gh}}(\alpha; \boldsymbol{\theta}_C)$.

For computing the aperture mass $M_{\mathrm{ap,f}}(\boldsymbol{\theta}_C)$ (with $f = 1, 2$) we use a compensated window U with an aperture scale[14] $\theta_{\mathrm{ap}} = 70$ arcmin. From a convergence map at a given source redshift z_f, we measure the aperture mass at location $\boldsymbol{\theta}_C$ through a convolution of the U filter with pixels in the neighbourhood of $\boldsymbol{\theta}_C$ (see equation 43). Note that it is completely equivalent to compute the aperture mass from the corresponding shear map at z_f by convolving shear pixels with the Q filter (with the same aperture scale size as that of U, see equation 44) and completely skip the KS convergence map making procedure (see Harnois-Déraps et al. 2020). Hence, the way in which we compute the aperture mass using the convergence field is redundant. As we are working in a simulated setting and do not consider holes and masks in our data, the map making procedure is straightforward. However, this is not the case in real data and it is then practical to evaluate the aperture mass from the shear map directly.

In the 108 T17 noiseless simulation maps, we evaluate the above statistics at locations distributed over the full sky. We do not do this for every pixel in the HEALPIX map but rather choose well-separated pixels (about $2\theta_\mathrm{T}$ apart – the diameter of W) for which the top-hat patches at those chosen pixels only slightly overlap with the patches centred at neighbouring chosen pixels. The overlap is not a problem and allows us to maximize the area over which we evaluate the statistics. For computing the two-point shear correlations $\xi_{\pm,\mathrm{gh}}(\alpha)$ in a given map we take the average of the position-dependent shear correlations evaluated at all chosen patches on the map (see the discussion after equation 38). The integrated three-point shear correlations $i\zeta_{\pm,\mathrm{fgh}}(\alpha)$ are evaluated by first taking the product of the aperture mass and the position-dependent shear correlation at a chosen location and then performing an average of this product

evaluated at all other locations (see equation 48) for a specific realization.

We perform the same measurements on the FLASK maps (with shape noise). Unlike the T17 maps, we do not distribute patches over the whole sky but rather cut out two big circular footprints of 5000 deg^2 (approximately the size of the DES footprint) in each hemisphere of a FLASK map and restrict the distribution of patches to within the extent of each footprint. In each FLASK map, the two footprints are widely separated that allows us to treat each region as an independent survey realization. This helps to maximize the use of our FLASK simulations and allows us to have a total of 2000 DES-like realizations (from 1000 FLASK maps) that we consider sufficient for the estimation of the covariance matrix of our data vector for a DES-sized survey as the number of realizations is much larger than the maximum size of our data vector that we discuss next.

For the two source redshifts z_1 and z_2, our data vector D_i evaluated from the ith simulation realization (T17 or FLASK) consists of the two-point shear cross-correlations and the integrated three-point shear cross-correlations (each correlation function evaluated at 20 angular separations α) as depicted below:

$$D_i \equiv (\xi_{\pm,11}, \; \xi_{\pm,22}, \; \xi_{\pm,12}, \; i\zeta_{\pm,111}, \; i\zeta_{\pm,222}, \; i\zeta_{\pm,122}, \; i\zeta_{\pm,211})^\mathrm{T}, \quad (57)$$

where T stands for transpose. This gives a data vector of size $N_d = 7 \times 2 \times 20 = 280$ elements. The mean data vector is obtained by taking an average of the individual data vectors obtained from each of the N_r realizations:

$$\overline{D} = \frac{1}{N_r} \sum_{i=1}^{N_r} D_i. \quad (58)$$

On the other hand, we evaluate our covariance matrix of the data vector as

$$\hat{\mathbf{C}} = \frac{1}{N_r - 1} \sum_{j=1}^{N_r} \left(D_j - \overline{D}\right) \left(D_j - \overline{D}\right)^\mathrm{T}, \quad (59)$$

thus resulting in an $N_d \times N_d = 280 \times 280$ matrix. For validating our theoretical model for the data vector we compare it with the mean data vector from the 108 T17 noiseless maps. For obtaining our DES-like data-covariance matrix (with impact of shape noise) we evaluate it from the $N_r = 2000$ footprints cut out from the FLASK simulations.

4.4 Methods for theoretical modelling

In this section, we detail the numerical recipes that go into the theoretical computation of the constituents of the model vector M which we evaluate for the fiducial T17 cosmology (see Section 4.1). For modelling the two-point shear correlations $\xi_{\pm,\mathrm{gh}}(\alpha)$ (see equation 36) we need to compute the convergence power spectrum $P_{\kappa,\mathrm{gh}}(l)$. As already stated before, we use the public Boltzmann solver code CLASS[15] to compute the non-linear revised HALOFIT 3D matter power spectrum $P_\delta^{\mathrm{3D}}(k, \eta)$ that we integrate along the line of sight to obtain $P_{\kappa,\mathrm{gh}}(l)$. We use the one-dimensional adaptive quadrature integration routine from the GNU Scientific Library (GSL)[16] (Gough 2009) to perform the integration. To partly correct for the flat-sky and the Limber approximation that goes into the derivation for the

---

[13]Currently hosted at: https://rmjarvis.github.io/TreeCorr/_build/html/index.html#.

[14]We found that for $\theta_{\mathrm{ap}} = 70$ arcmin the amplitude of the iB_+ signal was larger than other aperture scales when measured in combination with the top-hat patch of $\theta_\mathrm{T} = 75$ arcmin. Optimization of the filter sizes remains an interesting avenue to explore.

[15]To be precise, we use the C++ wrapper of the code (version v2.9.4) that can be obtained from the official repository, currently hosted at: https://github.com/lesgourg/class_public.

[16]Currently hosted at: http://www.gnu.org/software/gsl/.

expressions of the shear correlations, we multiply the convergence power spectrum by an l-dependent correction factor proposed by Kitching et al. (2017):

$$C_{\kappa,\mathrm{gh}}(l) \equiv \frac{(l+2)(l+1)l(l-1)}{\left(l+\frac{1}{2}\right)^4} P_{\kappa,\mathrm{gh}}\left(l+\frac{1}{2}\right). \quad (60)$$

Moreover, instead of performing the inverse Hankel transform l-integrals (i.e. the $\mathcal{F}_{2D}^{-1}[\cdots]$ operations in equation 36) for converting the Fourier space power spectra to shear correlations, we use expressions with summation over l as given in Friedrich et al. (2020; see also Stebbins 1996):

$$\xi_{\pm,\mathrm{gh}}(\alpha) = \sum_{l>2} \frac{2l+1}{4\pi} \frac{2\left(G_{l,2}^+(\cos\alpha) \pm G_{l,2}^-(\cos\alpha)\right)}{l^2(l+1)^2} C_{\kappa,\mathrm{gh}}(l), \quad (61)$$

where the functions $G_{l,2}^\pm(x)$ can be expressed in terms of second-order associated Legendre polynomials $\mathcal{P}_{l,2}(x)$ (Stebbins 1996; Friedrich et al. 2020):

$$G_{l,2}^+(x) \pm G_{l,2}^-(x) = \mathcal{P}_{l,2}(x)\left(\frac{4 - l \pm 2x(l-1)}{1-x^2} - \frac{l(l-1)}{2}\right) + \mathcal{P}_{l-1,2}(x)\frac{(l+2)(x \mp 2)}{1-x^2}. \quad (62)$$

These equations are exact for a curved-sky treatment and more accurate than the inverse Hankel transforms; the latter resulting in increasing errors for larger angular separations (Kitching et al. 2017). The expressions can be easily evaluated using the GSL library.

For computing the integrated three-point functions $i\zeta_{\pm,\mathrm{fgh}}(\alpha)$, we first need to evaluate the integrated shear bispectra $iB_{\pm,\mathrm{fgh}}(l)$ (see equation 53). We use the fitting formula for the 3D dark matter bispectrum $B_\delta^{3D}(k_1, k_2, k_3, \eta)$ by Gil-Marín et al. (2012, hereafter GM, see more in Appendix C) with the revised HALOFIT non-linear power spectrum implementation in CLASS that we then integrate over the l_i-modes and along the line of sight to obtain $iB_{\pm,\mathrm{fgh}}(l)$. For numerically computing the five-dimensional integration in equation (53) we use the publicly available adaptive multidimensional integration package CUBATURE[17] and evaluate each integrated bispectrum for 157 l-modes log-spaced in the range $1 \le l \le 20\,000$. In converting to real space, we again replace the required inverse Hankel integrals of $iB_{\pm,\mathrm{fgh}}(l)$ in equation (54) by the summation over l expressions in equation (61) to obtain $i\zeta_{\pm,\mathrm{fgh}}(\alpha)$. In order to do so, we first linearly interpolate the $iB_{\pm,\mathrm{fgh}}(l)$ between the 157 log-spaced l-modes to get the $iB_{\pm,\mathrm{fgh}}(l)$ for every integer-l multipole within the range specified above. We then use the interpolated value at every multipole to perform the summation.

In order to validate the theoretical model for the twp-point and integrated three-point shear correlations on the T17 simulations, we also need to account for the effects in the simulations due to limited angular resolution of the maps, finite simulation box size, and finite thickness of the lens shells as reported by T17. We include these corrections in our theory power spectra as summarized in Appendix B.

5 RESULTS AND DISCUSSION

We now present the results of our measurements and theory calculations. In Section 5.1, we test the accuracy of our model in describing the T17 data vector within the uncertainties expected from the sixth year cosmic shear data of the DES using the FLASK covariance

---

[17]Currently hosted at: https://github.com/stevengj/cubature.
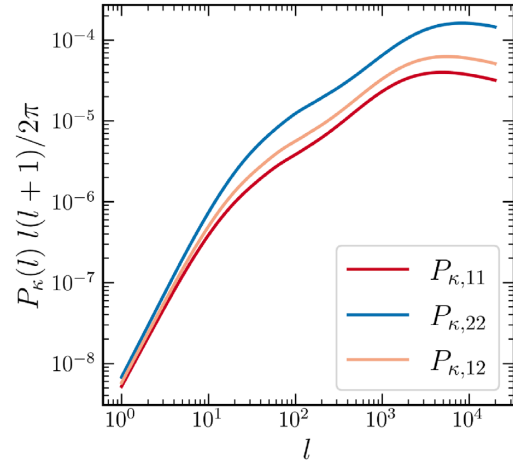
Figure 1. The scaled convergence auto- and cross-power spectra $P_\kappa(l)$ for two tomographic source redshift bins $z_1 = 0.5739$ and $z_2 = 1.0334$.
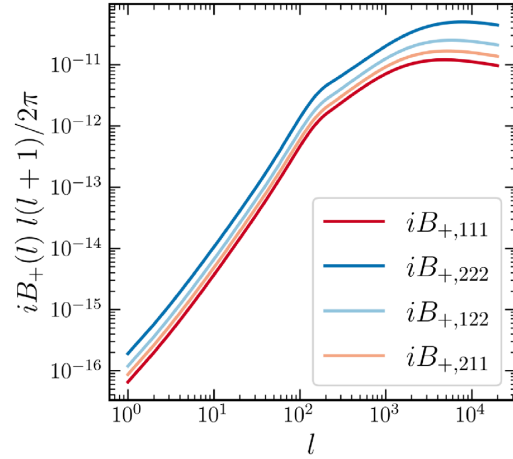
Figure 2. The scaled integrated bispectra $iB_+(l)$ for two tomographic source redshift bins $z_1 = 0.5739$ and $z_2 = 1.0334$. These spectra have been computed using a compensated filter of size $\theta_\mathrm{ap} = 70$ arcmin and top-hat window of radius $\theta_\mathrm{T} = 75$ arcmin.

matrix. In Section 5.2, we explore the Fisher constraining power on cosmological parameters that can be obtained on performing a joint analysis of $\xi_\pm$ and $i\zeta_\pm$.

The results of the theory computation of the convergence power spectra $P_{\kappa,\mathrm{gh}}(l)$ for source redshifts $z_1 = 0.5739$ and $z_2 = 1.0334$ (where g, $h = 1$, 2) are shown in Fig. 1. It is clear from the figure that the convergence power spectrum for sources at higher redshift, i.e. $P_{\kappa,22}$, is larger than the lower redshift power spectrum, $P_{\kappa,11}$, indicating the presence of more amount of deflecting material between the observer and the source at larger redshifts; in other words, a larger lensing efficiency for sources situated at a higher redshift (see equation 26). Also, the spectra are smooth as features like the baryonic acoustic oscillations that are prominent in the 3D matter power spectrum are smeared out due to the mixing of 3D k-modes into 2D l-modes through the line-of-sight projection (see equation 32). In Fig. 2, we show the integrated bispectra $iB_{+,\mathrm{fgh}}(l)$ for the two source redshifts z_1 and z_2 (where $f, g, h = 1, 2$). As mentioned before, the integrated bispectra are evaluated using a compensated filter of size $\theta_\mathrm{ap} = 70$ arcmin and two top-hat windows of radii $\theta_\mathrm{T} = 75$ arcmin. Other cross-combinations besides the four cross-spectra shown in the figure, e.g. $iB_{+,112}(l)$ and $iB_{+,212}(l)$ are the same as
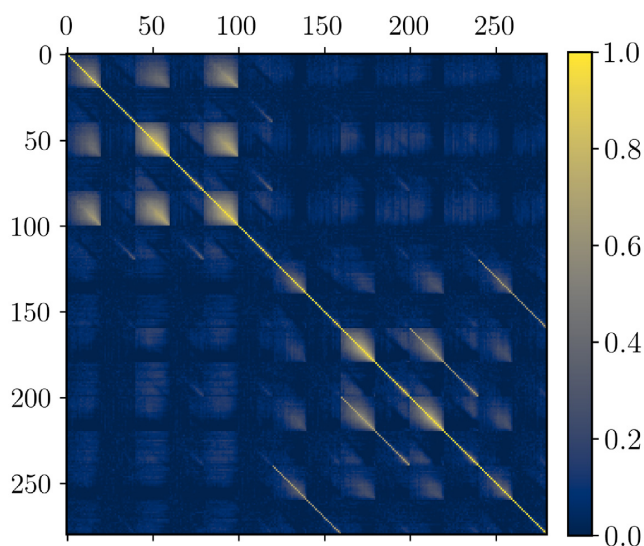
Figure 3. The 280×280 data-correlation matrix (normalized version of $\hat{\mathbf{C}}$, see equation 59) estimated from 2000 DES Year 6 sized footprints in FLASK lognormal sky maps that include realistic shape noise for two tomographic source redshift bins $z_1 = 0.5739$ and $z_2 = 1.0334$. Each 20×20 box around the diagonal indicates the correlation matrix for the 20 separation bins α of each of the 14 components of the data vector $D = (\xi_{+, 11}(\alpha), \xi_{-, 11}(\alpha), \xi_{+, 22}(\alpha), \xi_{-, 22}(\alpha), ..., i\zeta_{+, 211}(\alpha), i\zeta_{-, 211}(\alpha))^{\mathrm{T}}$ (see equation 57). The off-diagonal boxes indicate the cross-correlations between the angular bins of different correlation functions.

$iB_{+, 211}(l)$ and $iB_{+, 122}(l)$, respectively (e.g. this can be easily verified from equation 53). Hence, they add no extra information and we only consider these four. The $iB_{-, \mathrm{fgh}}$ spectra look similar to $iB_{+, \mathrm{fgh}}$ and are not shown separately.

It should be noted here that the high-l end of the integrated shear bispectra pick up significant contributions from squeezed configurations of the convergence bispectrum B_κ since the high-l values correspond to computing the position-dependent correlation function in real space on angular scales much smaller than the size of the patch ($l \gg 2\pi/2\theta_{\mathrm{T}} \approx 145$). As shown before for the 3D integrated bispectrum by Chiang et al. (2014) and for the 2D convergence bispectrum by Barreira et al. (2019) and Munshi et al. (2020b), this in turn corresponds to picking up the squeezed bispectrum configurations. However, it should be noted that the low-l end of iB picks up contribution from triangle configurations other than squeezed as the angular scales that the low-l correspond to are close to the diameter of the patch where the squeezed limit does not hold (see Fig. D1 and discussion in Appendix D for more details).

5.1 Validation on T17 simulations

In Fig. 4, we show each component of the data vector $\overline{D}$ (black dots) evaluated from the mean of 108 T17 simulated maps for the two source redshifts. The error bars on the data points indicate the standard deviation over the 108 maps (note that these are noiseless simulations). The grey shaded region is the 1σ standard deviation computed from the data-covariance matrix $\hat{\mathbf{C}}$ estimated from 2000 DES Year 6 sized footprints in FLASK lognormal sky maps that include realistic shape noise (see Fig. 3). The model vector for each statistic is also shown in the plots (in blue) where we also include the corrections proposed by T17 to account for the various resolution effects of the T17 simulation (see Appendix B). The $\xi_\pm$ models are in good agreement with the T17 measurements within both the scatter of

the simulations and the DES error bars. This is another confirmation of the result already reported by T17 that the convergence power spectrum (that we obtain using the revised HALOFIT 3D matter power spectrum) matches with the T17 simulations after taking into account the resolution corrections (see Appendix B). Our model predictions for the $i\zeta_+$ statistic also agrees well on all angular scales with the T17 simulations not only within the grey DES error bars but also within the scatter of the T17 simulations (black error bars). However, this is not the case for $i\zeta_-$ models as they are seen to be in agreement with the T17 simulations only on larger angular scales but overpredict the simulations on smaller scales. This stems from an inaccuracy of the GM bispectrum fitting formula. At the small angular scales, the $i\zeta_-$ with its fourth-order Bessel function J_4 (see equation 54 and discussion after the equation) is most sensitive to the very high-l values of the integrated bispectrum. At these very high l values, the integrated bispectrum signal is mostly due to the contributions from the highly squeezed configurations of the convergence bispectrum (see discussion in Appendix D). The GM formula, on the other hand, is known to overestimate these highly squeezed bispectrum configurations (Sato & Nishimichi 2013; Namikawa et al. 2019; Takahashi et al. 2020) and hence causes the overestimation of the $i\zeta_-$ signal on the small angular scales. On the other hand, $i\zeta_+$ has a zeroth-order Bessel function J_0 weighting that is more sensitive to lower l values (for a given angular scale) of the integrated bispectrum compared to $i\zeta_-$. At low to moderate l, the integrated bispectrum receives contribution from not so highly squeezed and other bispectrum triangle configurations where the GM fitting function works reasonably well. In Appendix D, we show the results of using a more accurate bispectrum fitting function BIHALOFIT (Takahashi et al. 2020) that correctly estimates the squeezed configurations and allows for an improved modelling of the $i\zeta_-$ correlations down to smaller angular scales (see Fig. D3). However, BIHALOFIT is currently only applicable to wCDM cosmologies (i.e. $w_0 = $ constant and $w_a = 0$), whereas one of the major goals of our analysis is to investigate the constraining power of $i\zeta_\pm$ for cosmologies with dynamical dark energy $w_a \neq 0$ (see Section 5.2 and the discussion in Appendix D). On the other hand, the GM fitting function is applicable to cosmologies with dynamical dark energy (as previously shown by Sato & Nishimichi 2013) and hence we choose it as our fiducial bispectrum model instead of BIHALOFIT.

To compare how well the model vector M describes the data vector $\overline{D}$ of a given statistic quantitatively, we compute the χ^2 value as

$$\chi^2 = (\overline{D} - M)^{\mathrm{T}} \mathbf{C}^{-1} (\overline{D} - M), \tag{63}$$

where $\mathbf{C}^{-1}$ is an unbiased estimate of the inverse data-covariance matrix $\hat{\mathbf{C}}^{-1}$ measured from N_r realizations for a data vector containing N_d elements (Hartlap, Simon & Schneider 2007):

$$\mathbf{C}^{-1} = \frac{N_r - N_d - 2}{N_r - 1} \hat{\mathbf{C}}^{-1}. \tag{64}$$

Note that the estimation of this unbiased inverse data-covariance matrix requires $N_r > N_d + 2$. Moreover, a relatively small number of N_r compared to N_d results in a highly noisy inverse covariance estimate (see Appendix F). Hence, one usually needs $N_r \gg N_d$.

Using the χ^2 value, computed using the FLASK covariance matrix (see Fig. 3), we make angular scale-cuts for every individual statistic (see the red-dashed vertical lines in Fig. 4). For making a scale-cut we impose two conditions. First, the χ^2 value of a given statistic using all angular bins larger than the scale-cut must be lower than a threshold value of 0.15. Secondly, the fractional change in the χ^2 value when ignoring the smallest bin right after the scale-cut should be less than 15 per cent. For further analyses, this enables us to
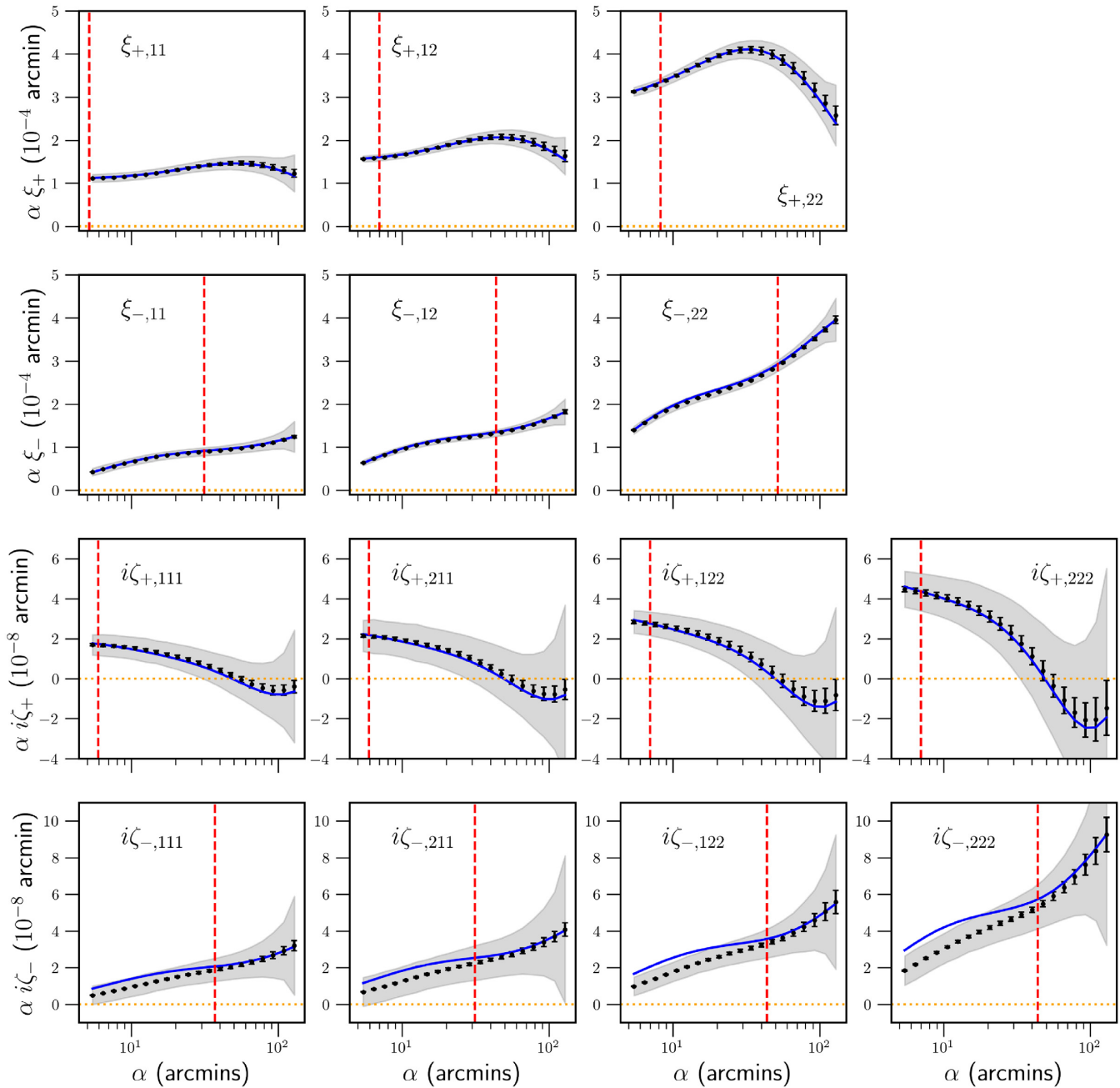
Figure 4. The two-point shear correlation functions $\xi_\pm(\alpha)$ and the integrated three-point shear correlation functions $i\zeta_\pm(\alpha)$ for two tomographic source redshift bins $z_1 = 0.5739$ and $z_2 = 1.0334$. The black dots with the error bars show the mean and the 1σ standard deviation of the measurements from the 108 T17 simulation maps, respectively. The grey shaded regions show the 1σ standard deviation for these statistics in DES Year 6 sized footprints obtained from the data-covariance matrix estimated using FLASK lognormal simulations with realistic shape noise. The blue curves show the theoretical model predictions for the statistics. The theory curves include the corrections needed to account for finite angular resolution, simulation box size, and shell thickness effects in the T17 simulations (see Appendix B). The integrated three-point functions have been computed using a compensated filter of size $\theta_{ap} = 70$ arcmin and top-hat window of radius $\theta_T = 75$ arcmin. The red-dashed lines denote the angular scale-cuts imposed on the data/model vectors using a χ^2 criterion (see text). The angular bins smaller than the scale-cuts are not included in further analyses.

include only those parts of the model vectors that agree very well with the simulations with respect to the DES-like uncertainties.

In Table 1, we report the signal-to-noise ratio (S/N) of the various statistics after imposing the angular scale-cuts. The S/N is computed as (Chang et al. 2019)

$$\text{S/N} = \sqrt{\overline{D}^{\text{T}} \mathbf{C}^{-1} \overline{D}},$$ (65)

and it indicates the statistical significance of the data vector. We also report the corresponding χ^2 values for the data vectors. Although we require the χ^2 for each individual statistic e.g. $\xi_{+,11}$ etc. to be below 0.15, there is no such restriction for the joint data vectors. The low χ^2 value of 1.08 for the entire data vector (after the scale-cuts) confirms that the model is in good agreement with the simulations within the DES uncertainties. We also check in Appendix E (see Fig. E1) whether any remaining

Table 1. Signal-to-noise ratio (S/N) for the T17 simulation data vectors computed with the FLASK covariance matrix. The χ^2 values for the theory model with respect to the data vector are also reported along with the length of the data vector. All reported quantities are evaluated after imposing angular scale-cuts (see Fig. 4). The data vector for each statistic includes all auto- and cross-correlations for both tomographic bins $z_1 = 0.5739$ and $z_2 = 1.0334$, e.g. $\xi_+ = (\xi_{+,11}, \xi_{+,12}, \xi_{+,22})^T, i\zeta_+ = (i\zeta_{+,111}, i\zeta_{+,222}, i\zeta_{+,122}, i\zeta_{+,211})^T$, etc.

Data vector	Length of data vector	S/N	χ^2
ξ_+	55	43.65	0.26
ξ_-	22	36.77	0.30
$\xi_\pm$	77	47.26	0.57
$i\zeta_+$	74	8.06	0.16
$i\zeta_-$	31	7.91	0.26
$i\zeta_\pm$	105	9.41	0.51
$\xi_\pm$ and $i\zeta_\pm$	182	48.40	1.08

systematic offset between the T17 data vector and our model vector after imposing the scale-cuts can cause any large parameter biases in our Fisher forecasts (see next section). We verify that the systematic offset for each parameter from the corresponding fiducial parameter value is smaller than one-third of the 1σ constraints expected from the Fisher analysis of the entire data vector.

Although the S/N of the $i\zeta_\pm$ is not as high as $\xi_\pm$ for a DES-like survey, the non-zero signals measured from the simulations without having had to compute the full three-point correlation function shows the ease of measurement and also the potential of the integrated three-point shear correlation function to probe higher order information of the highly non-Gaussian late-time matter density field.

5.2 Fisher forecast on cosmological parameter constraints

Having validated our theory model for the integrated three-point shear correlations – $i\zeta_+$ on all angular scales that we are interested in and $i\zeta_-$ on large angular scales – we shall now address the Fisher information content of this statistic on cosmological parameters when analysed jointly with the two-point shear correlation function. The Fisher information matrix **F** for a model vector M that depends on a set of parameters $\boldsymbol{\pi}$ reads (Huterer 2002; Dodelson & Schmidt 2020)

$$F_{ij} = \left(\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_i}\right)^T \mathbf{C}^{-1} \left(\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_j}\right), \tag{66}$$

where F_{ij} corresponds to an element of **F** for the model parameters π_i and π_j. The partial derivative of the model vector with respect to a model parameter π_i can be computed using a four-point central difference quotient[18] (also known as five-point stencil derivative; Abramowitz & Stegun 1964):

$$\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_i}$$
$$= \frac{-M(\pi_i + 2\delta_i) + 8M(\pi_i + \delta_i) - 8M(\pi_i - \delta_i) + M(\pi_i - 2\delta_i)}{12\delta_i}, \tag{67}$$

where δ_i is a small change of the parameter π_i about its fiducial value, and $M(\pi_i \pm \delta_i)$ means evaluating the model vector at the changed parameters $\pi_i \pm \delta_i$ while keeping all other parameters fixed. For our purpose we shall be interested in the cosmological

parameters $\boldsymbol{\pi} = \{\Omega_{\mathrm{cdm}}, \sigma_8, n_{\mathrm{s}}, w_0, w_a\}$ where w_0 and w_a indicate the dynamical dark energy equation of state parameters in the Chevallier–Polarski–Linder (CPL) parametrization (Chevallier & Polarski 2001; Linder 2003) adopted by the Dark Energy Task Force (Albrecht et al. 2006) to compare different dark energy probes. The fiducial values for our cosmological parameters are the same as that of the T17 simulations i.e. $\boldsymbol{\pi} = \{0.233, 0.82, 0.97, -1, 0\}$. For the first four parameters we choose the step sizes δ_i to be 4, 2, 10, and 8 per cent of the corresponding fiducial values. For the w_a parameter we adopt $\delta_{w_a} = 0.16$. We keep other parameters such as Ω_{b}, h fixed to their fiducial (T17) values and keep the flatness of the Universe unchanged. This means when varying Ω_{cdm}, the amount of dark energy in the Universe is adjusted accordingly. These step sizes were motivated from Yahia-Cherif et al. (2021) who proposed optimal steps for the five-point stencil derivative for Fisher analysis with the galaxy power spectrum. For our analysis we use slightly larger steps than them but within the proposed range of steps for the parameters. Our steps were found as a trade-off between neither being too big[19] (in order to obtain accurate derivatives, i.e. have low truncation errors) nor being too small (such that the derivatives are not dominated by numerical noise, i.e. have low rounding-off errors). We do not impose any priors on these five cosmological parameters.

The inverse of the Fisher matrix gives the parameter covariance matrix $\mathbf{C}_{\boldsymbol{\pi}}$ under the assumptions that the measured data vector is drawn from a multivariate Gaussian distribution[20] and that the dependence of M on the parameters $\boldsymbol{\pi}$ is close to linear (Trotta 2017; Uhlemann et al. 2020):

$$\mathbf{C}_{\boldsymbol{\pi}} = \mathbf{F}^{-1}. \tag{68}$$

Hence, using the derivatives and the expected data-covariance matrix (for a DES-sized survey) we can compute this parameter covariance matrix and forecast error contours on the cosmological parameters that we are interested in. In Figs 5 and 6, we show the derivatives of some components of our model vector with respect to the five cosmological parameters, normalized by the standard deviation for each component obtained from the FLASK covariance matrix (in other words, dividing the derivative of a statistic by the corresponding grey shaded error in Fig. 4 for a given separation bin α). This gives a visual estimate of the shape and amplitude of the ingredients of the Fisher matrix. It is clear that the way in which the amplitudes and shapes of $i\zeta_+$ derivatives (as a function of α) differ from one parameter to another is different compared to ξ_+ that results in slightly altered orientations of the error contours of each statistic in the parameter planes. This can be seen in Fig. 7. The error contours from $\xi_\pm$ are shown in blue, the contours from $i\zeta_\pm$ are shown in green dashed ellipses, and the joint contours of the two together in orange. For clarity, we also remove the integrated three-point function contours and show only the $\xi_\pm$ and the joint contours on the right-hand panel of the figure. Although the $i\zeta_\pm$ alone has larger contours compared to $\xi_\pm$ – due to the lower amplitudes of the derivatives (see Figs 5 and 6) that partly stems from the low S/N of $i\zeta_\pm$ (see Table 1) – the degeneracy directions are slightly different. A joint analysis of $\xi_\pm$ along with $i\zeta_\pm$

---

[18]We prefer to use four-point to two-point central difference quotient for obtaining more accurate first derivatives (see also Yahia-Cherif et al. 2021).

[19]To ensure that the steps were not too large, we verified that the δ_i were smaller than one-third of the 1σ marginalized Fisher constraints on the parameters (see Table 2) for the joint model in our final analysis.

[20]To go beyond the assumption that the data vector is drawn from a multivariate Gaussian distribution, one can also perform cosmological parameter inference with the integrated three-point shear correlation function in a likelihood-free inference set-up (Alsing, Wandelt & Feeney 2018) as advocated recently for other weak lensing summary statistics by Jeffrey, Alsing & Lanusse 2021).
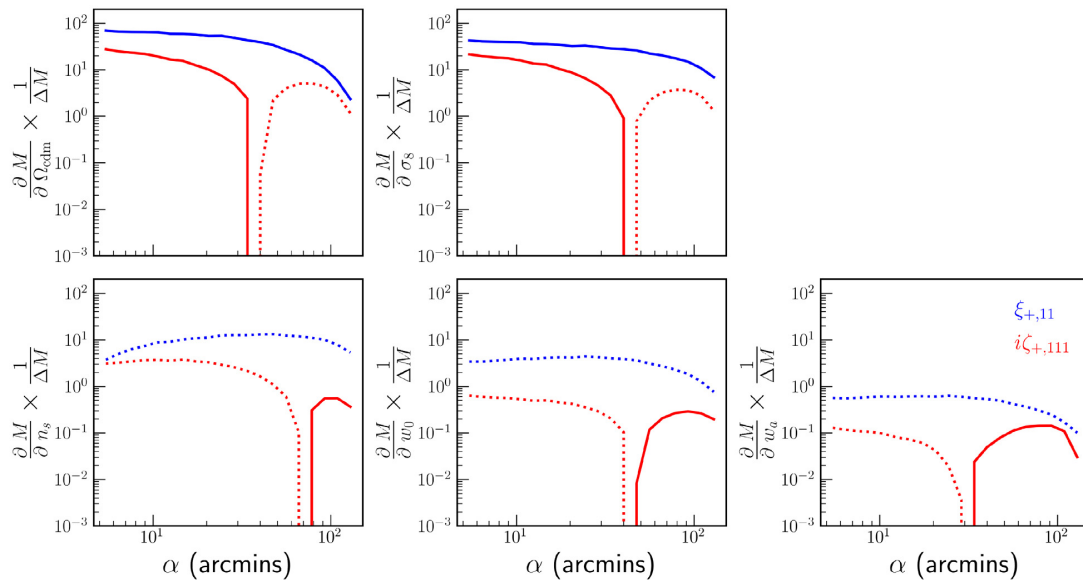
Figure 5. The derivatives $\frac{\partial M}{\partial \pi_i}$ of two components of the model vector $M - \xi_{+,11}$ (blue) and $i\zeta_{+,111}$ (red) shear correlation functions for source redshift bin $z_1 = 0.5739$ – with respect to the five cosmological parameters $\boldsymbol{\pi} = \{\Omega_{\mathrm{cdm}}, \sigma_8, n_{\mathrm{s}}, w_0, w_a\}$ and normalized by the corresponding 1σ standard deviation $\Delta M(\alpha)$ (from the FLASK covariance matrix) at a given α. The derivatives are negated (indicated with dotted lines) where $\frac{\partial M}{\partial \pi_i} < 0$.
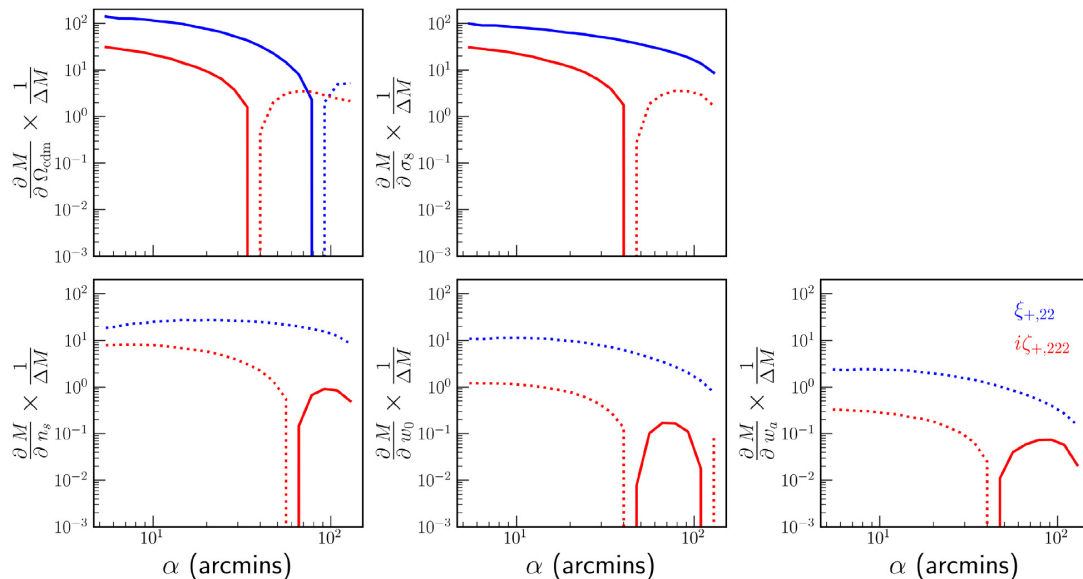
Figure 6. Same as Fig. 5, but for $\xi_{+,22}$ (blue) and $i\zeta_{+,222}$ (red) shear correlation functions for source redshift bin $z_2 = 1.0334$.

thus helps to alleviate some of the parameter degeneracies present in $\xi_\pm$ alone and result in a significant decrease in the contour sizes. The contribution from $i\zeta_\pm$ to the joint contours is significant with respect to the w_0, w_a parameters. This can be reasoned by investigating the derivatives (see Figs 5 and 6) of the statistics with respect to the dark energy equation of state parameters. The derivatives change more significantly for the different source redshifts with respect to w_0, w_a for $i\zeta_+$ compared to ξ_+. This can be attributed to the fact that the two-point shear correlation is a projection of the 3D power spectrum along the line of sight with a weighting of $\frac{q^2(\chi)}{\chi^2}$ (see equations 36 and 32), whereas the integrated three-point shear correlation function has a factor of $\frac{q^3(\chi)}{\chi^4}$ (see equation 50) implying that the latter is weighted more heavily at lower redshifts (or smaller

χ), especially in the dark-energy-dominated era. This sensitivity of the projected integrated three-point function to w_0 and w_a shows potential in probing the dynamical dark energy equation of state from cosmic shear data.[21] Quantitatively, this can also be seen from the marginalized 1σ constraints $\sigma(\pi_i) = \sqrt{C_{\boldsymbol{\pi},ii}}$ of the w_0 and w_a parameters for our analysis reported in the second column of Table 2.

---

[21] Interestingly, Byun et al. (2017) found that the 3D integrated bispectrum is relatively insensitive in constraining the dynamical dark energy parameters compared to the 3D power spectrum. However, as we find, the sensitivity to w_0 and w_a is different for the projected 2D integrated bispectrum compared to the projected 2D power spectrum – arising due to the different geometric projection kernel weighting terms in their respective line-of-sight projections.
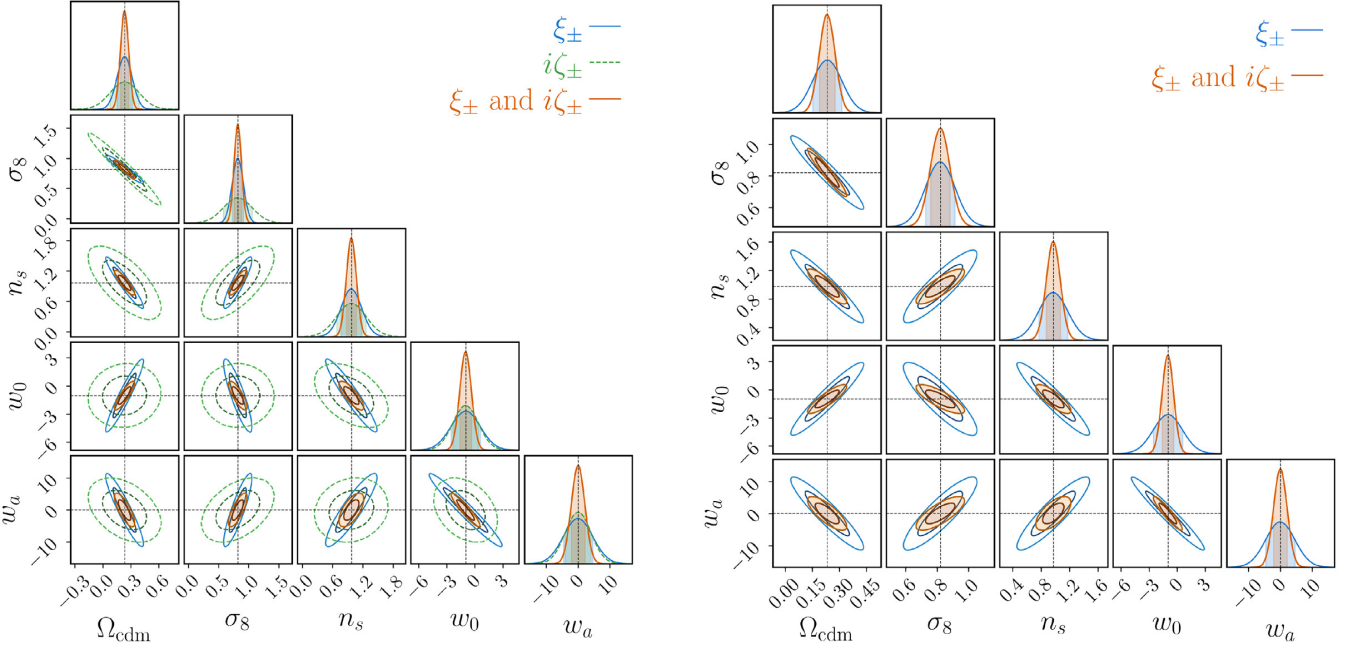
Figure 7. Left-hand panel: Fisher contours for the five cosmological parameters $\pi = \{\Omega_{\text{cdm}}, \sigma_8, n_s, w_0, w_a\}$ for the model vectors $- \xi_\pm$ (blue), $i\zeta_\pm$ (green dashed), and their joint model vector (orange) using the FLASK DES-like covariance matrix with realistic shape noise in a two tomographic source redshift bin setting with $z_1 = 0.5739$ and $z_2 = 1.0334$. The contours are centred around the fiducial parameter values (black dotted lines) and are computed after imposing the angular scale-cuts on the model vectors (see Fig. 4). Right-hand panel: same as left-hand panel but zoomed in and only showing the $\xi_\pm$ (blue) and the joint contours (orange).

Table 2. Comparison of our work in real space using shear two-point and *integrated* three-point correlations $\xi_\pm$, $i\zeta_\pm$ against previous works in Fourier space by Takada & Jain (2004), Kayo & Takada (2013), and Sato & Nishimichi (2013) who used the convergence power spectrum P_κ and the *full* convergence bispectrum B_κ. Some of the symbols used in the table that have not been defined in the text earlier are: the total number of source galaxies over all tomographic bins $n_g = \int dz\, p(z)$, where $p(z)$ is the entire source galaxy distribution, α_s is the spectral running index parameter, and $A_s = \delta_\zeta^2$ is the normalization parameter of the primordial power spectrum. We present the marginalized 1σ constraints on the cosmological parameters $\sigma(\pi_i)$ along with the dark energy FoM for our work using $\xi_\pm$, $i\zeta_\pm$ and the combined (joint) data vector of the two, respectively (with scale-cuts on the data vector). The step sizes δ_i (for computing the derivatives of the model with respect to the parameters) when specified in per cent are relative to the fiducial parameter values. We also show corresponding values reported in the other works. The '–' indicates values that are not explicitly reported or inapplicable to the other works.

	Our work	Takada & Jain (2004)	Kayo & Takada (2013)	Sato & Nishimichi (2013)
Total # source galaxies n_g (per arcmin^2)	10	100; $p(z)$ following Huterer (2002)	20	25
σ_ϵ	0.3	0.4	–	0.22
Area coverage (deg^2)	5000	4000	1500	1100
# source redshift bins for tomography	2	2	2	3
Source redshifts z_i	$z_1 = 0.5739, z_2 = 1.0334$	$0 \leq z_1 \leq 1.3, z_2 > 1.3$	–	$z_1 = 0.6, z_2 = 1.0, z_3 = 1.5$
Type of source redshift bin $p_i(z)$	Dirac-δ function	Equal $n_{g,i}$ in each bin from $n(z)$	Top-hat function	Dirac-δ function
Field	Cosmic shear γ	Convergence κ	Convergence κ	Convergence κ
Analysis in real or Fourier space	Real	Fourier	Fourier	Fourier
Data vectors (DVs)	$\xi_\pm(\alpha), i\zeta_\pm(\alpha)$, joint	$P_\kappa(l), B_\kappa(l_1, l_2, l_3)$, joint	$P_\kappa(l), B_\kappa(l_1, l_2, l_3)$, joint	$P_\kappa(l), B_\kappa(l_1, l_2, l_3)$, joint
Minimum and maximum scales	$5 < \alpha < 140$ arcmin	$50 \leq l_i \leq 3000$	$10 \leq l_i \leq 2000$	$72 \leq l_i \leq 2000$
Data-covariance	Lognormal simulations	Theoretical (only Gaussian covariance)	Theoretical	Theoretical
Cross-covariance for joint DV	Yes	No	Yes	Yes
Fisher analysis parameters	$\Omega_{\text{cdm}}, \sigma_8, n_s, w_0, w_a$	$\Omega_{\text{de}}, \Omega_b, h, n_s, \sigma_8, w_0, w_a$	$\Omega_{\text{de}}, \Omega_m h^2, \Omega_b h^2, n_s, \alpha_s, \delta_\zeta, w_0, w_a$	$\Omega_{\text{de}}, \Omega_{\text{cdm}} h^2, n_s, A_s, w_0, w_a$
Derivative step sizes δ_i	4, 2, 10, 8 per cent, 0.16	$\delta_{w_a} = 0.1$; 5 per cent for other parameters	–	$\delta_{w_a} = 0.5$; 10 per cent for other parameters
Analysis with flat or non-flat Universe	Flat	Flat	Non-flat	Flat
Priors in analysis	None	Planck priors on Ω_b, h, n_s	Planck priors on all parameters	None
Marginalized $\sigma(\Omega_{\text{cdm}})$	0.08, 0.16, 0.04	–	–	–
Marginalized $\sigma(\sigma_8)$	0.09, 0.24, 0.06	–	–	–
Marginalized $\sigma(n_s)$	0.20, 0.29, 0.10	–	–	–
Marginalized $\sigma(w_0)$	1.55, 1.36, 0.62	0.34, 0.32, 0.11	0.51, 0.62, 0.38	–
Marginalized $\sigma(w_a)$	4.56, 4.01, 2.11	0.93, 0.91, 0.36	1.30, 1.60, 0.94	–
Dark energy FoM	0.78, 0.19, 2.28	–	11, 7.2, 20	5, 15, 25

The constraints obtained from the joint analysis of $\xi_\pm$ and $i\zeta_\pm$, i.e. $\sigma(w_0) = 0.62$, $\sigma(w_a) = 2.11$, are significantly smaller than those present in the individual analysis of $\xi_\pm$, i.e. $\sigma(w_0) = 1.55$, $\sigma(w_a) = 4.56$, or in $i\zeta_\pm$, i.e. $\sigma(w_0) = 1.36$, $\sigma(w_a) = 4.01$. The same is true for the other cosmological parameters. Alternatively, one often quotes the dark energy figure of merit (FoM) defined as (Albrecht et al. 2006; Sato & Nishimichi 2013)

$$\text{FoM} \equiv \frac{1}{\sqrt{\det\left(\mathbf{C}_\pi[w_0, w_a]\right)}} \tag{69}$$

to characterize the power of a survey to constrain these two parameters. The higher the FoM, the stronger are the constraints in the $w_0 - w_a$ plane. For a DES-like survey, our joint analysis has a FoM $= 2.28$ that is almost three times larger than the FoM $= 0.78$ that we get from $\xi_\pm$ shear correlations alone; visually, this is reflected from the smaller size of the orange contours in the right-hand panel of Fig. 7 compared to the blue contours. The above quoted numbers are with the scale-cuts assumed in our analysis. We expect that including smaller angular scales will show more improvement on the marginalized constraints and also on the FoM.[22] However this needs the development of more accurate models down to small angular scales. We also show for comparison, the marginalized constraints and FoM from previous works by Takada & Jain (2004), Kayo & Takada (2013), and Sato & Nishimichi (2013) who investigated the convergence power spectrum and the full convergence bispectrum. Their reported constraints are significantly better than ours that we associate to several differences in their analysis settings to ours, e.g. higher n_g, no assumed scale-cuts, and for Kayo & Takada (2013) they assumed priors on the parameters of their Fisher analysis (see their fig. 1), whereas we do not impose any priors. Most importantly, these works investigate the constraining power of the full convergence bispectrum, whereas we study only an integrated quantity of the bispectrum. The full bispectrum can be targeted to probe general bispectrum configurations thereby probing more information than integrated quantities of the bispectrum. Of course, this is also true in real space for the full three-point shear correlation function γ-3PCF or the generalized third-order aperture mass statistics (Schneider et al. 2005). All these statistics should ideally be able to constrain the dark energy equation of state parameters better than the integrated three-point shear correlation function. However, all of them rely on the accurate measurement of the full γ-3PCF (or the bispectrum) from data that are still unexplored in current wide-area weak lensing surveys. The integrated three-point shear correlation function is much easier to measure and holds potential to improve upon the parameter constraints obtained from two-point shear analyses alone. On the theory side, we expect that including other effects such as galaxy intrinsic alignments, baryonic feedback, impact of massive neutrinos, etc. should be easier to tune into the $i\zeta_\pm$ model compared to including them for the full shear three-point correlation function. From both observational and theoretical aspects, this makes the integrated three-point shear correlation function a promising statistic to explore in current and future cosmic shear data.

6 CONCLUSIONS

In this paper, we propose a higher order statistic – *the integrated three-point shear correlation function* – that can be measured directly

from the cosmic shear field observed in current wide-area weak-lensing surveys such as DES, KiDS, HSC, and future surveys like Rubin Observatory Legacy Survey of Space and Time (LSST) and *Euclid*.[23] The following are the key results of this work.

(i) The integrated three-point shear correlation function $i\zeta_\pm$ can be measured by dividing a large survey area into several top-hat patches (each having an area of a few square degrees) and correlating the *position-dependent (local) two-point shear correlation function* inside each patch with the *aperture mass statistic* evaluated at the centre of the corresponding patch using a compensated filter. For fixed filter sizes, the $i\zeta_\pm(\alpha)$ is a function of a single variable – the separation scale α at which the local two-point shear correlation function is measured. This makes it analogous to the full shear two-point correlation function $\xi_\pm(\alpha)$ that is widely measured in weak lensing surveys (see Fig. 4).

(ii) We develop a theoretical model for $i\zeta_\pm$ that is the real space counterpart of *the integrated convergence bispectrum* as introduced by Munshi et al. (2020b) in Fourier space. The authors, however, formulated the integrated bispectrum using equal-sized top-hat patches on the convergence field. Working in real space with cosmic shear, we instead propose the usage of a combination of compensated (for the aperture mass statistic) and top-hat filters (for the local two-point shear correlation) of different sizes allowing for the evaluation of the statistic directly from cosmic shear data without any need for constructing a convergence map. We compute our theoretical models using the GM bispectrum fitting formula with the revised HALOFIT non-linear matter power spectrum (Takahashi et al. 2012) implementation in the CLASS software (Blas et al. 2011; Lesgourgues 2011).

(iii) We validate our model for the integrated three-point function using the weak lensing shear simulations from T17. We find that our theoretical predictions are in excellent agreement for the measured '+' integrated three-point functions $i\zeta_+$ (analogous to the ξ_+ shear two-point correlation function) within the scatter of the simulations for multiple source redshifts and the cross-correlations thereof (see Fig. 4). However, our model for the '−' integrated three-point shear correlation functions $i\zeta_-$ (analogous to the ξ_- shear correlation function) agrees with the simulations on large angular scales but overpredicts the simulation results on small scales. We associate this with the overestimation of the bispectrum by the GM fitting formula for the highly squeezed configurations of the bispectrum that the $i\zeta_-$ correlation function is mainly sensitive to. A more theoretically motivated formalism, e.g. using the response function approach to modelling the squeezed lensing bispectrum as recently studied by Barreira et al. (2019) – who also formulated the effect of baryons on the squeezed bispectrum – may help to accurately model the $i\zeta_-$ correlation functions down to smaller angular scales. This also shows the potential in encoding effects of non-linear processes (e.g. baryonic feedback) in the integrated three-point function that we expect to be easier compared to modelling them for the full three-point shear correlation function. This is left as a direction for future work. In wCDM cosmologies, one can use a more accurate 3D matter bispectrum fitting function such as BIHALOFIT (Takahashi et al. 2020) to achieve improved modelling of the $i\zeta_-$ correlations (see Appendix D).

(iv) Making appropriate scale-cuts on the model vectors, we use the Fisher matrix formalism to forecast constraints on cosmological parameters for a DES Year 6 sized survey with realistic shape

---

[22]For example, assuming that our model is correct on all angular scales and without imposing any scale-cuts, we find that the FoM for the joint data vector improves by over a factor of 2.

[23]See https://www.lsst.org and https://www.euclid-ec.org.

noise in a two-redshift bin tomographic setting (see Fig. 7). For the data-covariance matrix we use a set of lognormal simulations using the FLASK tool (Xavier et al. 2016). We find that the joint analysis of the integrated three-point function and the two-point shear correlation functions can allow for a significant improvement in the parameter constraints compared to those obtained from two-point shear correlation functions alone (see Table 2). This is because the responses of the integrated three-point shear correlations to the cosmological parameters are different from that of two-point shear correlations thereby resulting in slightly different degeneracy directions in the parameter planes (see Figs 5 and 6). In particular, we find that the integrated three-point function has the potential to significantly improve the dark energy FoM on a combined analysis with two-point shear correlation functions. This arises due to the derivatives of the integrated three-point function (with respect to the dark energy equation of state parameters) varying considerably in shape and amplitude compared to the derivatives of the two-point shear correlation. This can be attributed to the fact that the line-of-sight projection kernel in the expression for the convergence bispectrum is weighted considerably more heavily down to low redshifts (in the late-time dark-energy-dominated era) compared to the convergence power spectrum (see equations 33 and 32). This can be very useful for probing the dark energy equation of state parameters from cosmic shear data alone and makes the integrated three-point shear correlation function a promising method to probe higher order information content of the shear field and thereby complement two-point shear analysis.

Theoretically, the integrated three-point function (or the integrated bispectrum) of the lensing convergence field should be easier to work with than the $i\zeta_\pm$ shear correlation function that we investigate in this paper. However, observationally, the former requires one to go through the convergence map making process from the cosmic shear field. This process becomes challenging when the observed shear field has complicated masks and survey geometry. Although our analysis involves a simulated set-up with simplifying assumptions such as a circular survey footprint without masks and holes, accounting for the masking effects is straightforward as our statistic is designed to be measured directly from the cosmic shear data (where the masking effects are inherent) without the need for any map making. The integrated three-point shear correlation function with its ease of measurement through the two-point position-dependent shear correlation function and the one-point aperture mass statistic is tailor-made for application to real data.

Although we have concentrated on the integrated three-point function of the cosmic shear field, we provide a general framework of equations in Section 2 that can be used for computing the integrated three-point function for any projected field, e.g. galaxy counts field and its cross-correlations with the shear field. This will be explored in future works.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

The data for the *N*-body simulations used in this paper were accessed from the public domain: http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/. The lognormal simulations used in this work were generated using the publicly available FLASK software: http://www.astro.iag.usp.br/flask/.

REFERENCES

Abramowitz M., Stegun I. A., 1964, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover, New York
Albrecht A. et al., 2006, preprint (astro-ph/0609591)
Alsing J., Wandelt B., Feeney S., 2018, MNRAS, 477, 2874
Asgari M. et al., 2021, A&A, 645, A104
Barreira A., Nelson D., Pillepich A., Springel V., Schmidt F., Pakmor R., Hernquist L., Vogelsberger M., 2019, MNRAS, 488, 2079
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Bird S., Viel M., Haehnelt M. G., 2012, MNRAS, 420, 2551
Blas D., Lesgourgues J., Tram T., 2011, J. Cosmol. Astropart. Phys., 07, 034
Buchalter A., Kamionkowski M., Jaffe A. H., 2000, ApJ, 530, 36
Burger P., Schneider P., Demchenko V., Harnois-Deraps J., Heymans C., Hildebrandt H., Unruh S., 2020, A&A, 642, A161
Byun J., Eggemeier A., Regan D., Seery D., Smith R. E., 2017, MNRAS, 471, 1581
Chang C. et al., 2018, MNRAS, 475, 3165
Chang C. et al., 2019, MNRAS, 482, 3696
Chevallier M., Polarski D., 2001, Int. J. Mod. Phys. D, 10, 213
Chiang C.-T., Slosar A., 2018, J. Cosmol. Astropart. Phys., 01, 012
Chiang C.-T., Wagner C., Schmidt F., Komatsu E., 2014, J. Cosmol. Astropart. Phys., 05, 048
Chiang C.-T., Wagner C., Sánchez A. G., Schmidt F., Komatsu E., 2015, J. Cosmol. Astropart. Phys., 09, 028
Chiang C.-T., Cieplak A. M., Schmidt F., Slosar A., 2017, J. Cosmol. Astropart. Phys., 06, 022
Clerkin L. et al., 2017, MNRAS, 466, 1444
Coles P., Jones B., 1991, MNRAS, 248, 1
Crittenden R. G., Natarajan P., Pen U.-L., Theuns T., 2002, ApJ, 568, 20
Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369
Dodelson S., Schmidt F., 2020, Modern Cosmology, 2nd edn. Academic Press, New York
Dodelson S., Schneider M. D., 2013, Phys. Rev. D, 88, 063537
Doux C., Schaan E., Aubourg E., Ganga K., Lee K.-G., Spergel D. N., Tréguer J., 2016, Phys. Rev. D, 94, 103506
Escoffier S. et al., 2016, preprint (arXiv:1606.00233)
Friedrich O., Eifler T., 2018, MNRAS, 473, 4150
Friedrich O. et al., 2018, Phys. Rev. D, 98, 023508
Friedrich O. et al., 2020, preprint (arXiv:2012.08568)
Fu L. et al., 2014, MNRAS, 441, 2725
Gatti M. et al., 2020, MNRAS, 498, 4060

Gil-Marín H., Wagner C., Fragkoudi F., Jimenez R., Verde L., 2012, J. Cosmol. Astropart. Phys., 02, 047 (GM)

Giri S. K., D'Aloisio A., Mellema G., Komatsu E., Ghara R., Majumdar S., 2019, J. Cosmol. Astropart. Phys., 02, 058

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759

Gough B., 2009, GNU Scientific Library Reference Manual, 3rd edn. Network Theory Ltd., Godalming, Surrey

Gruen D. et al., 2018, Phys. Rev. D, 98, 023507

Hamana T., Sakurai J., Koike M., Miller L., 2015, PASJ, 67, 34

Hamana T. et al., 2020, PASJ, 72, 16

Harnois-Déraps J., Martinet N., Castro T., Dolag K., Giblin B., Heymans C., Hildebrandt H., Xia Q., 2020, preprint (arXiv:2012.02777)

Harris C. R. et al., 2020, Nature, 585, 357

Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399

Hilbert S., Hartlap J., Schneider P., 2011, A&A, 536, A85

Hinton S. R., 2016, J. Open Source Softw., 1, 00045

Hunter J. D., 2007, Comput. Sci. Eng., 9, 90

Huterer D., 2002, Phys. Rev. D, 65, 063001

Jarvis M., Bernstein G., Jain B., 2004, MNRAS, 352, 338

Jeffrey N., Alsing J., Lanusse F., 2021, MNRAS, 501, 954

Joachimi B., 2017, MNRAS, 466, L83

Jung G., Oppizzi F., Ravenni A., Liguori M., 2020, J. Cosmol. Astropart. Phys., 06, 035

Jung G., Namikawa T., Liguori M., Munshi D., Heavens A., 2021, J. Cosmol. Astropart. Phys., 06, 055

Kacprzak T. et al., 2016, MNRAS, 463, 3653

Kaiser N., 1992, ApJ, 388, 272

Kaiser N., 1995, ApJ, 439, L1

Kaiser N., Squires G., 1993, ApJ, 404, 441

Kayo I., Takada M., 2013, preprint (arXiv:1306.4684)

Kilbinger M., 2015, Rep. Progress Phys., 78, 086901

Kilbinger M., Schneider P., 2005, A&A, 442, 69

Kitching T. D., Alsing J., Heavens A. F., Jimenez R., McEwen J. D., Verde L., 2017, MNRAS, 469, 2737

Leauthaud A. et al., 2007, ApJS, 172, 219

Lesgourgues J., 2011, preprint (arXiv:1104.2932)

Limber D. N., 1954, ApJ, 119, 655

Linder E. V., 2003, Phys. Rev. Lett., 90, 091301

Martinet N., Harnois-Déraps J., Jullo E., Schneider P., 2021, A&A, 646, A62

Munshi D., Coles P., 2017, J. Cosmol. Astropart. Phys., 02, 010

Munshi D., Namikawa T., Kitching T., McEwen J., Takahashi R., Bouchet F., Taruya A., Bose B., 2020a, MNRAS, 493, 3985

Munshi D., McEwen J. D., Kitching T., Fosalba P., Teyssier R., Stadel J., 2020b, J. Cosmol. Astropart. Phys., 05, 043

Namikawa T., Bose B., Bouchet F. R., Takahashi R., Taruya A., 2019, Phys. Rev. D, 99, 063511

Pires S. et al., 2020, A&A, 638, A141

Sato M., Nishimichi T., 2013, Phys. Rev. D, 87, 123538

Schneider P., 1996, MNRAS, 283, 837

Schneider P., 2006, in Meylan G., Jetzer P., North P., eds, Saas-Fee Advanced Course 33, Gravitational Lensing: Strong, Weak and Micro. Springer-Verlag, Berlin, p. 269

Schneider P., Lombardi M., 2003, A&A, 397, 809

Schneider P., van Waerbeke L., Mellier Y., 2002, A&A, 389, 729

Schneider P., Kilbinger M., Lombardi M., 2005, A&A, 431, 9

Schrabback T. et al., 2018, A&A, 610, A85

Scoccimarro R., Couchman H. M. P., 2001, MNRAS, 325, 1312

Scoccimarro R., Frieman J. A., 1999, ApJ, 520, 35

Semboloni E., Schrabback T., van Waerbeke L., Vafaei S., Hartlap J., Hilbert S., 2011, MNRAS, 410, 143

Shirasaki M., Hamana T., Yoshida N., 2015, MNRAS, 453, 3043

Smith R. E. et al., 2003, MNRAS, 341, 1311

Springel V., 2005, MNRAS, 364, 1105

Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79

Stebbins A., 1996, preprint (astro-ph/9609149)

Takada M., Jain B., 2004, MNRAS, 348, 897

Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152

Takahashi R., Hamana T., Shirasaki M., Namikawa T., Nishimichi T., Osato K., Shiroyama K., 2017, ApJ, 850, 24(T17)

Takahashi R., Nishimichi T., Namikawa T., Taruya A., Kayo I., Osato K., Kobayashi Y., Shirasaki M., 2020, ApJ, 895, 113

Taylor A., Joachimi B., Kitching T., 2013, MNRAS, 432, 1928

Trotta R., 2017, preprint (arXiv:1701.01467)

Troxel M. et al., 2018, Phys. Rev. D, 98, 043528

Uhlemann C., Friedrich O., Villaescusa-Navarro F., Banerjee A., Codis S., 2020, MNRAS, 495, 4006

Xavier H. S., Abdalla F. B., Joachimi B., 2016, MNRAS, 459, 3693

Yahia-Cherif S. et al., 2021, A&A, 649, A52

Zonca A., Singer L., Lenz D., Reinecke M., Rosset C., Hivon E., Gorski K., 2019, J. Open Source Softw., 4, 1298

Zürcher D., Fluri J., Sgier R., Kacprzak T., Refregier A., 2021, J. Cosmol. Astropart. Phys., 01, 028

APPENDIX A: FOURIER AND HANKEL TRANSFORMS

The forward and inverse Fourier transforms of a field f in the 2D sky-plane can be written as

$$f(\boldsymbol{l}) = \mathcal{F}_{2D}[f(\boldsymbol{\theta})] \equiv \int d^2\boldsymbol{\theta}\ f(\boldsymbol{\theta})\,e^{-i\boldsymbol{l}\cdot\boldsymbol{\theta}} \qquad \text{(forward FT),}$$

$$f(\boldsymbol{\theta}) = \mathcal{F}_{2D}^{-1}[f(\boldsymbol{l})] \equiv \int \frac{d^2\boldsymbol{l}}{(2\pi)^2}\ f(\boldsymbol{l})\,e^{i\boldsymbol{l}\cdot\boldsymbol{\theta}} \qquad \text{(inverse FT),} \qquad \text{(A1)}$$

where $\boldsymbol{l} = (l_x, l_y)$ is the 2D Fourier wave vector. If the field f is real, i.e. $f^*(\boldsymbol{\theta}) = f(\boldsymbol{\theta})$, then it follows from the above equation that $f^*(\boldsymbol{l}) = f(-\boldsymbol{l})$.

If a function (e.g. correlation function) $\xi(\boldsymbol{\alpha})$ defined in the 2D sky plane is independent of the direction of the vector $\boldsymbol{\alpha}$, i.e. $\xi(\boldsymbol{\alpha}) = \xi(\alpha)$, then it follows from the Fourier transformation equation (A1) and from the properties of ordinary Bessel functions that (Schneider 2006; Dodelson & Schmidt 2020)

$$P(l) \equiv \mathcal{F}_{2D}[\xi(\alpha)] = \int d^2\boldsymbol{\alpha}\ \xi(\alpha)\,e^{-i\boldsymbol{l}\cdot\boldsymbol{\alpha}}$$

$$= 2\pi \int d\alpha\ \alpha\ \xi(\alpha)\ J_0(l\alpha),$$

$$\xi(\alpha) \equiv \mathcal{F}_{2D}^{-1}[P(l)] = \int \frac{d^2\boldsymbol{l}}{(2\pi)^2}\ P(l)\,e^{i\boldsymbol{l}\cdot\boldsymbol{\alpha}}$$

$$= \int \frac{dl\ l}{2\pi}\ P(l)\ J_0(l\alpha), \qquad \text{(A2)}$$

where $J_0(x)$ is the zeroth-order Bessel function of the first kind.

On the other hand, the 2D Fourier transform of $\xi(\alpha)$ with a complex phase factor $e^{4i\phi_\alpha}$ and its inverse transform reads

$$P(l) \equiv \mathcal{F}_{2D}\left[\xi(\alpha)\,e^{4i\phi_\alpha}\right] = \int d^2\boldsymbol{\alpha}\ \xi(\alpha)\,e^{-i\boldsymbol{l}\cdot\boldsymbol{\alpha}}\,e^{4i\phi_\alpha}$$

$$= 2\pi \int d\alpha\ \alpha\ \xi(\alpha)\ J_4(l\alpha),$$

$$\xi(\alpha) \equiv \mathcal{F}_{2D}^{-1}\left[P(l)\,e^{-4i\phi_l}\right] = \int \frac{d^2\boldsymbol{l}}{(2\pi)^2}\ P(l)\,e^{i\boldsymbol{l}\cdot\boldsymbol{\alpha}}\,e^{-4i\phi_l}$$

$$= \int \frac{dl\ l}{2\pi}\ P(l)\ J_4(l\alpha), \qquad \text{(A3)}$$

where ϕ_α is the polar angle of $\boldsymbol{\alpha}$, ϕ_l is the polar angle of $\boldsymbol{l}$, and $J_4(x)$ is the fourth-order Bessel function of the first kind. These equations are Hankel transformations.

APPENDIX B: T17 SIMULATIONS POWER SPECTRA CORRECTION FORMULAE

T17 found that the convergence power spectra that were measured from the mean of the 108 simulated sky maps in their simulation suite, slightly underestimated the theoretical power spectrum calculated with revised HALOFIT (which we are also using in this paper). They associated three effects that caused the underestimation and provided correction factors to the theory formulae to take them into account.

(i) *Finite simulation-box-size effect*. In appendix B of T17, the authors report that in order to consider the effect of density fluctuations larger than the simulation-box-size L on the angular power spectrum, one needs to impose the condition that for $k < 2\pi/L$, the matter power spectrum $P_\delta^{3D}(k, \eta) \overset{!}{=} 0$, as the box does not include fluctuations larger than L.

(ii) *Finite lens-shell effect*. The T17 lensing maps were produced by ray tracing through lens shells of finite thickness in the simulation boxes (see Section 4.1). The finite thickness affects the angular power spectrum of surface density fluctuations on a shell. To account for this, T17 suggest to convolve the matter power spectrum with the window function of the shell (see their appendix B). They provide a fitting formula for the convolved power spectrum:

$$P_\delta^{3D}(k, \eta) \longrightarrow \frac{(1 + c_1 k^{-\alpha_1})^{\alpha_1}}{(1 + c_2 k^{-\alpha_2})^{\alpha_3}} P_\delta^{3D}(k, \eta), \tag{B1}$$

with $c_1 = 9.5171 \times 10^{-4}$, $c_2 = 5.1543 \times 10^{-3}$, $\alpha_1 = 1.3063$, $\alpha_2 = 1.1475$, and $\alpha_3 = 0.62793$ that they find to be in good agreement with the analytically computed convolved power spectrum up to redshift $z < 7.1$ that is well within the range considered in this paper.

(iii) *Finite angular resolution of sky maps*. For a given NSIDE of a HEALPIX map, the angular power spectrum $C(l)$ measured from the sky map is underestimated compared to the theoretical power spectrum at large l due to lack of angular resolution. To account for this in the theory spectrum, T17 suggest a damping factor at small scales (high-l) given by

$$C(l) \longrightarrow \frac{C(l)}{1 + (l/l_{\text{res}})^2}, \tag{B2}$$

where $l_{\text{res}} = 1.6 \times \text{NSIDE}$.

APPENDIX C: 3D MATTER BISPECTRUM

The 3D matter bispectrum at leading order (tree-level) in density perturbations as computed with standard Eulerian perturbation theory (PT) for Gaussian initial conditions is written as (Bernardeau et al. 2002; Dodelson & Schmidt 2020)

$$B_{\delta,\text{tree}}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \eta) = 2 F_2(\boldsymbol{k}_1, \boldsymbol{k}_2, \eta) P_{\delta,L}^{3D}(k_1, \eta) P_{\delta,L}^{3D}(k_2, \eta)$$
$$+ \text{ cyclic permutations,} \tag{C1}$$

where

$$P_{\delta,L}^{3D}(k, \eta) = D_+^2(\eta) P_{\delta,L}^{3D}(k, \eta_0) \tag{C2}$$

is the 3D linear matter power spectrum today evolved to time η using the linear growth factor $D_+(\eta)$ that is normalized to unity today, i.e. $D_+(\eta_0) = 1$. $F_2(\boldsymbol{k}_i, \boldsymbol{k}_j, \eta)$ is a symmetrized two-point mode coupling kernel that in a general ΛCDM universe takes the form (Friedrich et al. 2018)

$$F_2(\boldsymbol{k}_i, \boldsymbol{k}_j, \eta) = \mu(\eta) + \frac{1}{2} \cos(\phi_{ij}) \left(\frac{k_i}{k_j} + \frac{k_j}{k_i} \right)$$
$$+ [1 - \mu(\eta)] \cos^2(\phi_{ij}), \tag{C3}$$

where ϕ_{ij} is the angle between the two wave vectors $\boldsymbol{k}_i$ and $\boldsymbol{k}_j$. In an Einstein–de Sitter (EdS) universe, the function $\mu(\eta)$ is a constant and takes the value $\mu(\eta) = \frac{5}{7}$. However, this form of the bispectrum only works in the linear regime (large physical scales) and fails in the non-linear regime. To improve upon this, one can go on to include higher order PT corrections but calculating the higher order terms is cumbersome. Another way of predicting the non-linear matter bispectrum is to propose a fitting formula for the bispectrum and calibrate the function's parameters using the bispectra measured from CDM N-body simulations. This approach was first taken by Scoccimarro & Frieman (1999) and later improved by Scoccimarro & Couchman (2001) and GM. In this paper, we use the bispectrum fitting formula of GM:

$$B_\delta^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \eta) = 2 F_2^{\text{eff}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \eta) P_\delta^{3D}(k_1, \eta) P_\delta^{3D}(k_2, \eta)$$
$$+ \text{ cyclic permutations,} \tag{C4}$$

where $P_\delta^{3D}(k, \eta)$ is the 3D non-linear matter power spectrum (e.g. obtained using revised HALOFIT; Takahashi et al. 2012) and the effective mode coupling kernel $F_2^{\text{eff}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \eta)$ is a modified version of the EdS F_2 kernel and reads

$$F_2^{\text{eff}}(\boldsymbol{k}_i, \boldsymbol{k}_j, \eta) = \frac{5}{7} a(k_i, \eta) a(k_j, \eta)$$
$$+ \frac{1}{2} \cos(\phi_{ij}) \left(\frac{k_i}{k_j} + \frac{k_j}{k_i} \right) b(k_i, \eta) b(k_j, \eta)$$
$$+ \frac{2}{7} \cos^2(\phi_{ij}) c(k_i, \eta) c(k_j, \eta). \tag{C5}$$

The functions $a(k, \eta)$, $b(k, \eta)$, and $c(k, \eta)$ are fitting formulae calibrated with N-body simulations to interpolate the results between the linear (tree-level bispectrum) and the non-linear regime bispectrum measured from the simulations:

$$a(k, \eta) = \frac{1 + \sigma_8^{a_6}(\eta) [0.7 Q_3(n_{\text{eff}})]^{1/2} (q a_1)^{n_{\text{eff}} + a_2}}{1 + (q a_1)^{n_{\text{eff}} + a_2}},$$

$$b(k, \eta) = \frac{1 + 0.2 a_3 (n_{\text{eff}} + 3)(q a_7)^{n_{\text{eff}} + 3 + a_8}}{1 + (q a_7)^{n_{\text{eff}} + 3.5 + a_8}},$$

$$c(k, \eta) = \frac{1 + 4.5 a_4 / [1.5 + (n_{\text{eff}} + 3)^4] (q a_5)^{n_{\text{eff}} + 3 + a_9}}{1 + (q a_5)^{n_{\text{eff}} + 3.5 + a_9}}. \tag{C6}$$

Although these functions have been expressed in terms of conformal time, it is completely equivalent to replace η with the corresponding redshift z as the time argument in the above expressions. $\sigma_8(\eta)$ is the standard deviation of matter density fluctuations today linearly evolved to time η, i.e. $\sigma_8(\eta) = D_+(\eta) \sigma_8(\eta_0)$. The effective logarithmic slope of the linear matter power spectrum today $n_{\text{eff}}(k)$ reads

$$n_{\text{eff}}(k) = \frac{d \log P_{\delta,L}^{3D}(k, \eta_0)}{d \log k}. \tag{C7}$$

$q \equiv k/k_{\text{nl}}$ is defined with the scale $k_{\text{nl}}(\eta)$ at which non-linearities start to become important and is defined as

$$\frac{k_{\text{nl}}^3 P_{\delta,L}^{3D}(k_{\text{nl}}, \eta)}{2\pi^2} \equiv 1, \tag{C8}$$

and the function $Q_3(n_{\text{eff}})$ is defined as

$$Q_3(n_{\text{eff}}) \equiv \frac{4 - 2^{n_{\text{eff}}}}{1 + 2^{n_{\text{eff}} + 1}}. \tag{C9}$$

The values for the parameters calibrated using simulations as found by GM are

$a_1 = 0.484$, $a_2 = 3.740$, $a_3 = -0.849$, $a_4 = 0.392$, $a_5 = 1.013$,
$a_6 = -0.575$, $a_7 = 0.128$, $a_8 = -0.722$, $a_9 = -0.926$.

As reported by GM, the fitting formula with these parameter values works reasonably well for $z < 1.5$ and for $k < 0.4$ Mpc^{-1} h in ΛCDM cosmologies. However, in this paper we use this fitting function for non-ΛCDM cosmologies, in particular to predict the bispectrum for cosmologies with varying dark energy equation of state parameters by encoding the information of the latter into the fitting formula through the linear and non-linear (revised HALOFIT) power spectra and $\sigma_8(\eta)$ obtained using CLASS. Our approach is similar to what has previously been done by Sato & Nishimichi (2013) who verified that the GM formula reasonably described the lensing bispectrum measured in N-body simulations with dynamical dark energy. Another approach, for wCDM cosmologies (i.e. w_0 = constant and $w_a = 0$) can be taken by using the recently introduced BIHALOFIT fitting function for the matter bispectrum by Takahashi et al. (2020) that is more accurate than the GM fitting function especially in predicting the highly squeezed configurations of the matter bispectrum that the GM formula overestimates. We show results of modelling $i\zeta_\pm$ using BIHALOFIT in Appendix D.

APPENDIX D: INTEGRATED SHEAR BISPECTRUM USING DIFFERENT APPROXIMATIONS

Here we compare the results of computing the integrated bispectrum $iB_{+,222}(l)$ and correspondingly the integrated three-point shear correlation function $i\zeta_{+,222}(\alpha)$ for the source redshift $z_2 = 1.0334$ when using different 3D matter bispectrum approximations in equations (53) and (50). In Fig. D1, we plot the integrated bispectrum when computed with the GM bispectrum fitting formula (as already shown in Fig. 2) along with the prediction when using the tree-level bispectrum (see equation C1). We also plot the integrated bispectrum with the GM bispectrum but only when considering elongated/squeezed configurations, i.e. when two modes of the bispectrum are at least larger than two times the smallest mode.

From the figure, it is clear that the tree-level bispectrum and the GM formula match on low-l ($l \ll 100$) corresponding to large angular
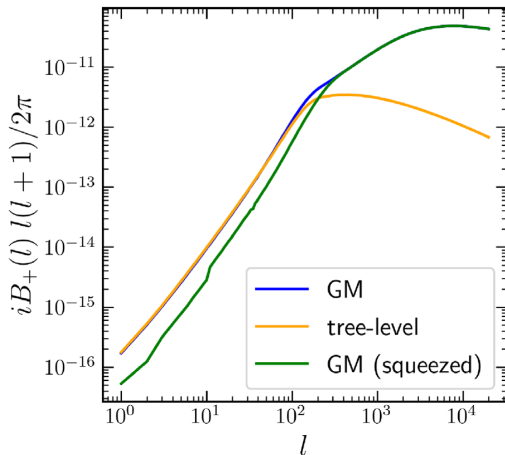
Figure D1. The scaled integrated shear bispectrum $iB_{+,222}(l)$ for source redshift bin $z_2 = 1.0334$ computed with equation (53) using the Gil-Marín et al. (2012 – GM) fitting formula (blue), the tree-level bispectrum (orange), and the GM formula but only when considering squeezed configurations (green). The non-smoothness in the green curve on low-l is numerical artefacts arising from the integration routine being forced to exclude sampled points in the integration volume for non-squeezed configurations. The computations for iB_+ were performed using a compensated filter of size $\theta_{ap} = 70$ arcmin and top-hat window of radius $\theta_T = 75$ arcmin.
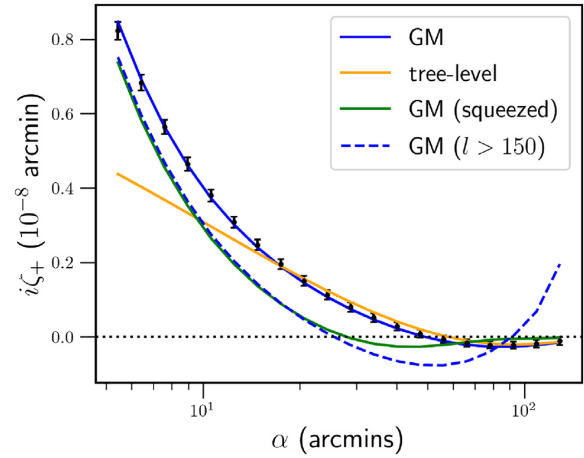
Figure D2. The integrated three-point function $i\zeta_{+,222}(\alpha)$ for source redshift bin $z_2 = 1.0334$. The black dots with the error bars show the mean and the 1σ standard deviation of the measurements from the 108 T17 maps, respectively. The blue-solid curve shows the model prediction using the GM bispectrum and in blue-dashed the prediction using only $l > 150$. The orange curve shows the predicted signal using the tree-level bispectrum and in green the signal using only the squeezed configurations of the GM formula. The theory curves include the corrections needed to account for finite angular resolution, simulation box size, and shell thickness effects in the T17 simulations. The computations use a compensated filter of size $\theta_{ap} = 70$ arcmin and top-hat window of radius $\theta_T = 75$ arcmin.

scales, but differ significantly on small scales that correspond to the non-linear regime (high-l). On the other hand, when computing the iB_+ with only the squeezed configurations of the GM bispectrum, we find that the result matches with the full GM result only in the high-l end, indicating that most of the iB_+ signal is dominated by squeezed configurations for l much larger than the characteristic mode corresponding to the diameter of the patch within which the position-dependent shear correlation is measured, i.e. $l \gg 2\pi/(2\theta_T) \approx 145$. However, for low-$l$ modes corresponding to scales approximately the size of the patch or larger, the squeezed configuration result underestimates the full GM result. This shows that our statistic probes not only the squeezed but partly also other bispectrum configurations. This also explains the non-smoothed behaviour of the squeezed iB_+ result on low-l as the integration routine is forced to exclude sampled points in the integration volume for the non-squeezed configurations that contribute mostly at low-l. The non-smoothness is insignificant and does not affect the computation of the $i\zeta_+$ signal that we discuss next.

In Fig. D2, we show the corresponding real space $i\zeta_{+,222}(\alpha)$ predictions by Hankel transforming (actually using equation 61) the integrated shear bispectra computed above and compare them with the result of the T17 simulations. The GM bispectrum computed prediction matches well with the simulations as already seen in Fig. 4. The tree-level bispectrum computed $i\zeta_+(\alpha)$ signal only captures the result on the largest angular scales but heavily deviates in the non-linear regime. The squeezed configuration calculation of the GM bispectrum follows the trend of the simulation on small scales while slightly underestimating the measured signal. This can be attributed to the fact that at a given small angular separation α, $i\zeta_+(\alpha)$ receives contributions not only from the high-l end of $iB_+(l)$ but also from the low-l end that correspond to scales larger than the separation scale (see equation 61). As seen in Fig. D1, the squeezed bispectrum $iB_+(l)$ underestimates the full GM bispectrum result in the low-l end thereby explaining the slight deficit. On larger scales, the squeezed
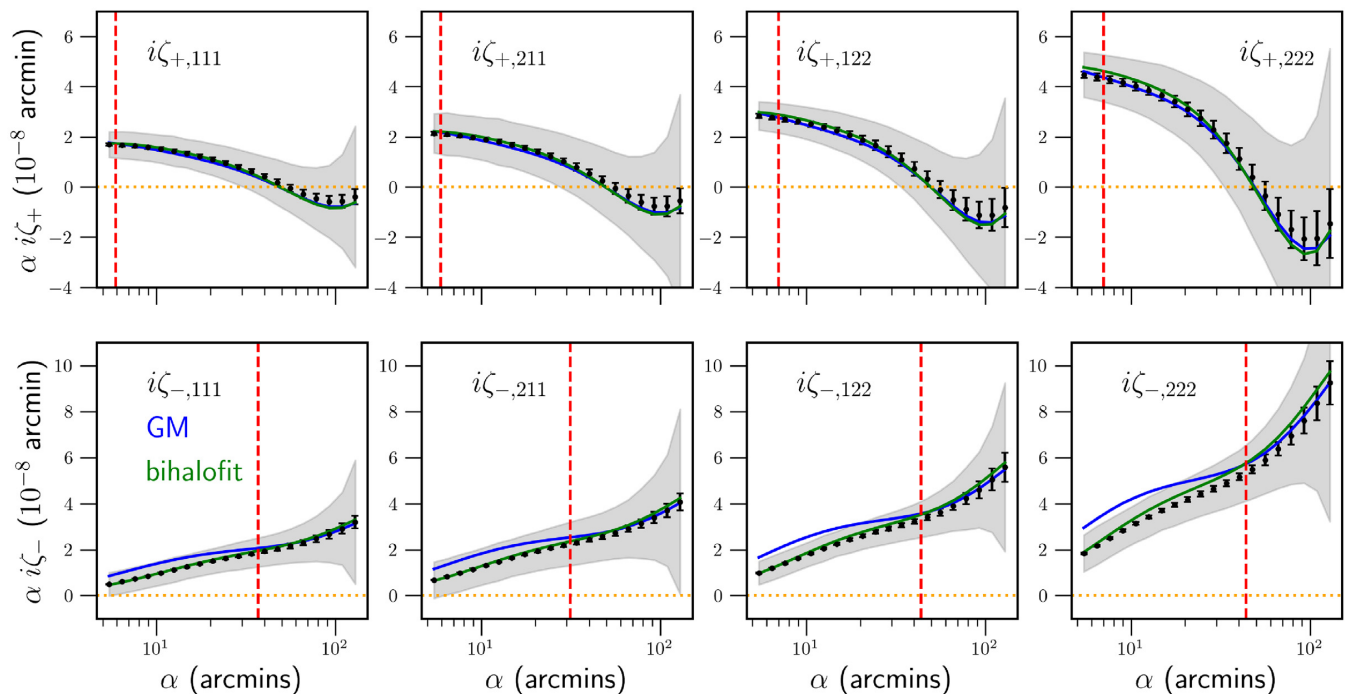
Figure D3. Same as the lower two panels of Fig. 4 depicting the integrated three-point shear correlation functions $i\zeta_\pm(\alpha)$ for two tomographic source redshift bins $z_1 = 0.5739$ and $z_2 = 1.0334$. The blue curves show the theoretical model predictions for the statistics using the 3D matter bispectrum fitting formula by Gil-Marín et al. (2012) (GM – our fiducial model for Fisher analysis). The green curves show the model predictions using a more recent 3D bispectrum fitting formula by Takahashi et al. (2020) (BIHALOFIT). As described in the text, although the BIHALOFIT fitting formula enables more accurate modelling of the $i\zeta_-$ correlations than the GM formula on small angular scales, we do not use it for our Fisher analysis as it is currently not applicable to cosmologies with dynamical dark energy.

bispectrum fails to describe the simulation results showing that the squeezed-limit approximation does not hold as α approaches the size of the patch. We also show the result of the $i\zeta_+(\alpha)$ signal computation using the GM bispectrum but restricting the Hankel summation of $iB_+(l)$ to include only $l > 150$, i.e. modes corresponding to scales much smaller than the size of the patch. Although the result does not describe the simulations, the signal matches the squeezed bispectrum results on the small scales confirming that the $i\zeta_+$ signal is indeed described by the squeezed-limit bispectrum on these scales. However, it is worth noting that on very small scales (smaller than 5 arcmin) the $i\zeta_+$ prediction with the full GM bispectrum will eventually fail to describe the T17 simulations as the integrated bispectrum result at extremely small scales (very high-l) receives most contribution from highly squeezed bispectrum configurations that are known to be overestimated by the GM formula (Namikawa et al. 2019; Takahashi et al. 2020). This was apparent for the $i\zeta_-$ signals (see Fig. 4) that are already sensitive to the very high-l values of the integrated bispectrum for angular separations around 30 arcmin (due to the J_4 Bessel function weighting).

To show that one can indeed improve the modelling at smaller angular scales for the $i\zeta_-$ correlations, we use the recently introduced BIHALOFIT fitting function by Takahashi et al. (2020) for the 3D matter bispectrum and compare it against the results obtained using the GM fitting function that we have adopted as our fiducial modelling choice. The results comparing them both to the T17 simulations are shown in Fig. D3. As already depicted in Fig. 4, the GM model predictions for $i\zeta_\pm$ are shown in blue and the corresponding angular scale-cuts in the red-dashed vertical lines. The green-solid curves show the theoretical predictions using the BIHALOFIT fitting formula and demonstrates the significant improvement achieved in modelling

the $i\zeta_-$ simulation results at the smaller angular scales. This is exactly due to the fact that the squeezed bispectrum configurations are more accurately predicted by BIHALOFIT than the GM fitting function (Takahashi et al. 2020). Using BIHALOFIT would thus allow to push the currently imposed angular scale-cuts down to even smaller scales and retain larger parts of the data vector. Nevertheless, we still use the $i\zeta_\pm$ model predictions with the GM formula for the Fisher analysis because currently BIHALOFIT is only applicable to wCDM cosmologies (i.e. $w_0 = $ constant and $w_a = 0$), whereas a major goal of our analysis (see Section 5.2) is to investigate the constraining power of $i\zeta_\pm$ for cosmologies with $w_a \neq 0$. The GM fitting function in combination with the revised HALOFIT non-linear power spectrum has no such restriction and has previously been validated for cosmologies with dynamical dark energy (Sato & Nishimichi 2013). Furthermore, as we only use those parts of the GM computed $i\zeta_\pm$ model vectors that have been validated on the T17 simulations (ensured with the imposed scale-cuts shown in Fig. 4), our parameter constraints are more on the conservative side (see Fig. 7). Including smaller angular scales of the $i\zeta_-$ data vector with improved modelling is expected to only improve the overall constraining power. However, further improved modelling in cosmologies with $w_a \neq 0$ is beyond the scope of this paper and is left as a direction for future work.

APPENDIX E: IMPACT OF SYSTEMATIC OFFSET BETWEEN MODEL AND DATA VECTORS ON PARAMETER CONSTRAINTS

Here we discuss the impact of the remaining systematic offset between the model M and data vector $\overline{D}$ (see Fig. 4) after imposing the angular scale-cuts in our analysis. A systematic offset would

Best-fit parameters
$\xi_{\pm}$ and $i\zeta_{\pm}$
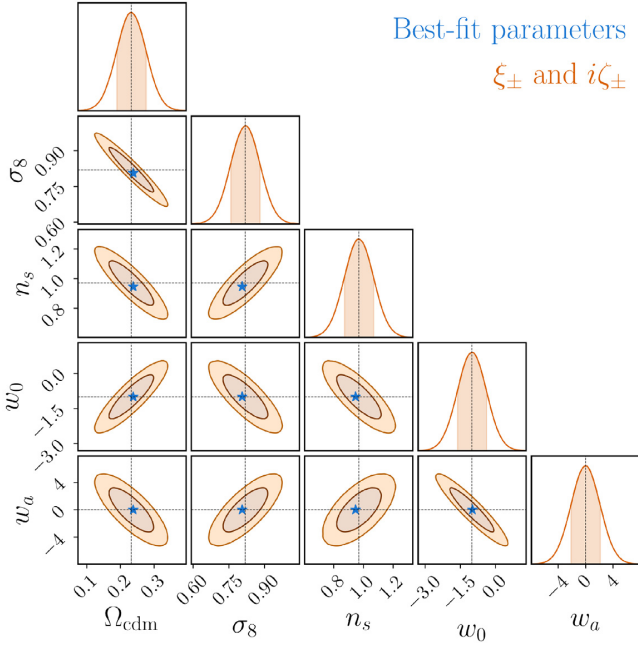
Figure E1. Offsets between fiducial parameters (black dotted lines) and best-fitting parameters (blue stars) in the parameter planes. The Fisher contours expected from the analysis of the entire model vector (after imposing scale-cuts) are shown in orange centred around the fiducial parameters.

amount to a bias in our parameter constraints that would cause the Fisher contours in Fig. 7 to be centred around the wrong cosmological values $\boldsymbol{\pi}^0$ – in our case the fiducial parameters. In other words, we want to explore how much the best-fitting[24] parameters $\boldsymbol{\pi}^{\mathrm{MP}}$ of the model describing the data vector are off from $\boldsymbol{\pi}^0$. In order to do so, we need to minimize the $\chi^2(\boldsymbol{\pi})$ as a function of the parameters (see equation 63) between the data and model. We already saw in Section 5.1 that the $\chi^2(\boldsymbol{\pi} = \boldsymbol{\pi}^0)$ between $\overline{D}$ and $M(\boldsymbol{\pi}^0)$ has a value of 1.08. We now want to find the parameters $\boldsymbol{\pi}^{\mathrm{MP}}$ that describe the data vector with the lowest χ^2_{MP}. We adopt the approach of Friedrich et al. (2020, see their section 5.1) and study a linearized approximation of the model vector as a function of the parameters $M(\boldsymbol{\pi})$ around the fiducial parameters $\boldsymbol{\pi}^0$. This allows us to write the best-fitting parameters as (see equation 32 of Friedrich et al. 2020)

$$\boldsymbol{\pi}^{\mathrm{MP}} = \boldsymbol{\pi}^0 + \mathbf{F}^{-1}\boldsymbol{x}, \tag{E1}$$

where we have assumed no priors on the parameters. $\mathbf{F}$ is the Fisher matrix (see equation 66) of the model vector and $\boldsymbol{x}$ is another vector with components:

$$x_i = \left(\overline{D} - M(\boldsymbol{\pi^0})\right)^{\mathrm{T}} \mathbf{C}^{-1} \left(\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_i}\right), \tag{E2}$$

where $\mathbf{C}$ is the data-covariance matrix (see equation 59) and $\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_i}$ are the derivatives of the model with respect to the parameters, evaluated at the fiducial values $\boldsymbol{\pi} = \boldsymbol{\pi}^0$. We show our best-fitting parameters for the model describing the entire T17 data vector (after imposing the scale-cuts) in Fig. E1 that can be seen to scatter very closely around the fiducial parameters. We also plot the orange contours (see Fig. 7) of the parameters from the Fisher analysis for the entire data vector (with $\xi_{\pm}$ and $i\zeta_{\pm}$ including the assumed scale-cuts). The absolute offsets of the best-fitting parameters from the fiducial values in units

[24]MP stands for maximum posterior in the notation of Friedrich et al. (2020).

of the marginalized 1σ Fisher constraints for the five parameters $\Omega_{\mathrm{cdm}}, \sigma_8, n_{\mathrm{s}}, w_0, w_a$ are 0.12, 0.22, 0.25, 0.02, and 0.02, respectively. As these offsets are smaller than one-third the marginalized 1σ constraints in the parameter planes, we conclude that our fiducial model after imposing the angular scale-cuts describes the T17 data vector very well and there is no significant bias in our results. Ideally, one should include these offsets as a systematic error but as they are not significant we deem it safe to ignore for our analysis.

APPENDIX F: VALIDATING THE USE OF LOGNORMAL DATA-COVARIANCE ON PARAMETER CONSTRAINTS USING THE PRECISION MATRIX EXPANSION

In this appendix, we test whether the use of lognormal FLASK simulations (see Section 4.2) as a model for computing the data-covariance matrix $\hat{\mathbf{C}}$ and its inverse, the precision matrix, causes any significant over/underestimation of the Fisher constraints presented in Section 5.2. For our purpose we use the precision matrix expansion (PME) formalism developed in section 3 of Friedrich & Eifler (2018) that we explain below.

Let us assume that on one hand we know the true data-covariance $\mathbf{C}_{\mathrm{true}}$ (e.g. from N-body simulations) and on the other hand we have a model $\mathbf{C}$ for the covariance (e.g. lognormal model). We can then write

$$\begin{aligned}
\mathbf{C}_{\mathrm{true}} &= \mathbf{C} + \mathbf{C}_{\mathrm{true}} - \mathbf{C} \\
&= \left(\mathbf{1} + (\mathbf{C}_{\mathrm{true}} - \mathbf{C})\mathbf{C}^{-1}\right)\mathbf{C} \\
&= (\mathbf{1} + \mathbf{X})\,\mathbf{C},
\end{aligned} \tag{F1}$$

where $\mathbf{1}$ is the identity matrix and $\mathbf{X} \equiv (\mathbf{C}_{\mathrm{true}} - \mathbf{C})\mathbf{C}^{-1}$. The true precision matrix, i.e. $\mathbf{C}_{\mathrm{true}}^{-1}$, can then be expressed as

$$\begin{aligned}
\mathbf{C}_{\mathrm{true}}^{-1} &= \left((\mathbf{1} + \mathbf{X})\,\mathbf{C}\right)^{-1} \\
&= \mathbf{C}^{-1}\,(\mathbf{1} + \mathbf{X})^{-1} \\
&= \mathbf{C}^{-1}\left(\mathbf{1} - \mathbf{X} + \mathbf{X}^2 + \mathcal{O}[\mathbf{X}^3]\right),
\end{aligned} \tag{F2}$$

where in the last line we have used the geometric series expansion of $(\mathbf{1} + \mathbf{X})^{-1}$. We can now write estimates for the true precision matrix for different orders in $\mathbf{X}$. Up to the zeroth-order we have

$$\mathbf{C}_{\mathrm{true,0th}}^{-1} \equiv \mathbf{C}^{-1}, \tag{F3}$$

which is the inverse of the model covariance matrix. This is exactly what we have used as our fiducial precision matrix throughout the main text of the paper for computing the χ^2 values and for our Fisher analysis. To remind ourselves, we first estimate the model covariance matrix $\hat{\mathbf{C}}$ from FLASK lognormal simulations and we then write (see equation 64) an unbiased estimate for the inverse model covariance $\mathbf{C}^{-1}$ using the Hartlap et al. (2007) correction factor that assumes that the estimated covariance matrix $\hat{\mathbf{C}}$ is distributed according to a Wishart distribution. Now, going up to first order in PME we have

$$\begin{aligned}
\mathbf{C}_{\mathrm{true,1st}}^{-1} &\equiv \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{X} \\
&= \mathbf{C}^{-1} - \mathbf{C}^{-1}(\hat{\mathbf{C}}_{\mathrm{true}} - \mathbf{C})\mathbf{C}^{-1} \\
&= 2\mathbf{C}^{-1} - \mathbf{C}^{-1}\hat{\mathbf{C}}_{\mathrm{true}}\mathbf{C}^{-1}.
\end{aligned} \tag{F4}$$

The second term in the last line is the leading order correction to our model precision matrix $\mathbf{C}^{-1}$ and as depicted is written using a direct estimate $\hat{\mathbf{C}}_{\mathrm{true}}$ of the true covariance matrix. This is a crucial point because even if one has only a few N-body mocks to estimate $\mathbf{C}_{\mathrm{true}}$, the correction term does not involve the inversion of this estimate. This is applicable for us as we only have a finite number of T17
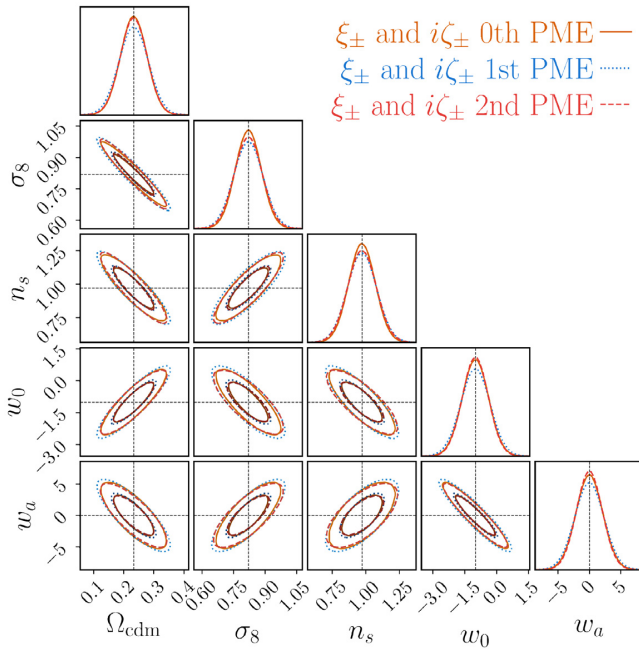
Figure F1. The Fisher contours expected from the analysis of the entire model vector (after imposing scale-cuts) for three different use cases of the inverse covariance matrix – zeroth-order precision matrix expansion (PME; orange solid – same as the constraints shown in right-hand panel of Fig. 7), first-order PME (blue dotted), and second-order PME (red dashed). The higher order PME terms hardly change the contours obtained when using the zeroth-order precision matrix that is estimated from lognormal simulations.

Table F1. The dark energy figure of merit (FoM) as measured from the marginalized Fisher constraints in the w_0–w_a plane (see equation 69) where the parameter covariance $\mathbf{C}_\pi$ is measured in the Fisher analysis with three different versions of the precision matrix: the zeroth-order PME (fiducial lognormal precision matrix), first-order PME, and second-order PME. The results are shown for $\xi_\pm$, $i\zeta_\pm$ and their combined data vector. The factor of improvement achieved in the FoM of the joint $\xi_\pm$ and $i\zeta_\pm$ compared to $\xi_\pm$ alone is also shown.

Data vector	Zeroth-order PME	First-order PME	Second-order PME
FoM $\xi_\pm$	0.78	0.77	0.76
FoM $i\zeta_\pm$	0.19	0.19	0.18
FoM $\xi_\pm$ and $i\zeta_\pm$	2.28	1.86	2.40
Factor improvement	2.9	2.4	3.2

mocks to estimate $\hat{\mathbf{C}}_{\text{true}}$. In order to do this, we first add shape noise to all the 108 T17 simulations (similar to what we did for the FLASK lognormal maps as described in Section 4.2) and then cut out two big circular footprints of 5000 deg^2 (approximately the size of the DES footprint) in each hemisphere of a given map. This gives us effectively 216 DES-sized true mocks from which we estimate $\hat{\mathbf{C}}_{\text{true}}$. To ensure that equation (F4) is an unbiased estimator, we evaluate the two $\mathbf{C}^{-1}$ terms appearing in the $\mathbf{C}^{-1}\hat{\mathbf{C}}_{\text{true}}\mathbf{C}^{-1}$ from two independent sets of FLASK lognormal simulations. We also symmetrize the final term after its evaluation.

Practically, we can use the $N_r = 216$ T17 footprints estimated true data-covariance $\hat{\mathbf{C}}_{\text{true}}$ and invert it to obtain the precision matrix with the Hartlap correction formula (see equation 64). This is

possible as N_r is larger than the size of our entire data vector after imposing scale-cuts $N_d = 182$ (see Table 1). However, as discussed in Taylor et al. (2013), the uncertainty that one encounters in the estimation of the precision matrix goes as $\sqrt{2/(N_r - N_d - 4)}$ that upon using the T17 covariance would result in a very large error of roughly 25 per cent on the Fisher matrix. Furthermore, even when the covariance estimate can be safely inverted (i.e. when $N_r \gg N_d$) the noise in the precision matrix estimate still leads to a significant additional scatter in maximum-likelihood parameters in actual likelihood analyses, unless $N_r - N_d \gg N_d$ (cf. Dodelson & Schneider 2013, as well as fig. 1 of Friedrich & Eifler 2018). Keeping this in mind and the fact that in future applications we may add further redshift bins and hence consider even bigger data vectors, we opt for the strategy of Friedrich & Eifler (2018) instead of standard inversion.

Following equation (12) of Friedrich & Eifler (2018), we can also write an unbiased estimator of the true precision matrix up to second order:

$$
\begin{aligned}
\mathbf{C}^{-1}_{\text{true, 2nd}} &\equiv \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{X} + \mathbf{C}^{-1}\mathbf{X}^2 \\
&= 3\mathbf{C}^{-1} - 3\mathbf{C}^{-1}\hat{\mathbf{C}}_{\text{true}}\mathbf{C}^{-1} \\
&\quad + \mathbf{C}^{-1}\frac{\nu^2\hat{\mathbf{C}}_{\text{true}}\mathbf{C}^{-1}\hat{\mathbf{C}}_{\text{true}} - \nu\hat{\mathbf{C}}_{\text{true}}\text{tr}\left(\mathbf{C}^{-1}\hat{\mathbf{C}}_{\text{true}}\right)}{\nu^2 + \nu - 2}\mathbf{C}^{-1}, \quad \text{(F5)}
\end{aligned}
$$

where $\nu = N_r - 1$ with N_r being the number of mock realizations used for estimating $\hat{\mathbf{C}}_{\text{true}}$ and $\text{tr}(\cdots)$ stands for evaluating the trace of laska matrix. Just as the correction factor that Hartlap et al. (2007) advocated for the unbiased estimation of the inverse of a matrix that is Wishart distributed, the final term in the above equation also stems from the same assumption, i.e. $\hat{\mathbf{C}}_{\text{true}}$ is Wishart distributed. For more details, the reader is referred to section 3 and appendix B of Friedrich & Eifler (2018). For evaluating the final term, as mentioned earlier, we again ensure that each of the involved $\mathbf{C}^{-1}$ is estimated from independent sets of FLASK lognormal simulations.

In Fig. F1, we show the effect on the Fisher parameter contours of the entire data vector (with $\xi_\pm$ and $i\zeta_\pm$) when using the precision matrix as evaluated with the zeroth-order PME (orange solid), first-order PME (blue dotted), and second-order PME (red dashed) expressions. The correction induced to the parameter constraints at first-order in PME hardly changes the results obtained using the zeroth-order PME (our fiducial lognormal model precision matrix). The result of adding even higher order terms up to second-order in PME shows remarkable agreement with the fiducial contours (compare the red dashed and orange solid ellipses).

We also evaluated the FoM in the w_0–w_a plane for $\xi_\pm$, $i\zeta_\pm$ and their combined data vector. The values are reported in Table F1. One sees that the relative factor of improvement of the FoM on performing a Fisher analysis of the joint data vector as compared to only $\xi_\pm$ is larger than 2 in all cases – irrespective of whether one uses the fiducial lognormal model precision matrix or after correcting the model precision matrix with leading order PME terms estimated from the T17 simulations. All these results lead us to conclude that our analysis with the FLASK lognormal covariance matrix is well justified and gives us robust qualitative and quantitative estimates for the significant improvement achieved in the parameter constraints upon complementing two-point shear analysis with the measurement of integrated three-point shear correlation functions.

This paper has been typeset from a TeX/LaTeX file prepared by the author.

Chapter 4

Response approach to the integrated shear 3-point correlation function: the impact of baryonic effects on small scales

Bibliographic and copyright information

Author's contribution

This work extends the previous work in chapter 3, where I introduced the integrated shear 3-point correlation function $\zeta_\pm$. It overcomes a major shortcoming in its modelling on the small angular scales (which in the previous paper were excluded by using conservative scale-cuts; e.g. see lower panel of Fig. 4 of chapter 3.). In summary, using the so called *response function approach to perturbation theory* framework developed by the second author A. Barreira (Barreira & Schmidt, 2017), we not only are able to achieve improved accuracy of the model of $\zeta_\pm$ (see for example Fig. 4 of this paper) but also can include baryonic feedback effect. The latter is an astrophysical effect which redistributes mat-

ter on small-scales and hence contaminates the true cosmological lensing signal seen in 2-point or in this case also in the integrated 3-point correlation function analysis. This is a major systematic effect that needs to be accounted for carefully in order to obtain unbiased cosmological results. I used the response approach method and implemented it in the modelling of $\zeta_\pm$. As first author I performed the theoretical modelling, numerical calculations, validation of the model against simulations, parameter constraint forecasts and wrote the full draft. The second author contributed significantly by guiding me in the modelling aspects, brainstorming ideas, and thorough reviewing of the draft.

Response approach to the integrated shear 3-point correlation function: the impact of baryonic effects on small scales

Anik Halder [1,2]★ and Alexandre Barreira[3,4]★

[1]*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstraße 1, 81679 München, Germany*
[2]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstraße 1, 85748 Garching, Germany*
[3]*Excellence Cluster ORIGINS, Boltzmannstraße 2, 85748 Garching, Germany*
[4]*Fakultät für Physik, Ludwig-Maximilians-Universität, Schellingstraße 4, 80799 München, Germany*

ABSTRACT

The integrated shear 3-point correlation function $\zeta_\pm$ is a higher-order statistic of the cosmic shear field that describes the modulation of the 2-point correlation function $\xi_\pm$ by long-wavelength features in the field. Here, we introduce a new theoretical model to calculate $\zeta_\pm$ that is accurate on small angular scales, and that allows to take baryonic feedback effects into account. Our model builds on the realization that the small-scale $\zeta_\pm$ is dominated by the non-linear matter bispectrum in the squeezed limit, which can be evaluated accurately using the non-linear matter power spectrum and its first-order response functions to density and tidal field perturbations. We demonstrate the accuracy of our model by showing that it reproduces the small-scale $\zeta_\pm$ measured in simulated cosmic shear maps. The impact of baryonic feedback enters effectively only through the corresponding impact on the non-linear matter power spectrum, thereby permitting to account for these astrophysical effects on $\zeta_\pm$ similarly to how they are currently accounted for on $\xi_\pm$. Using a simple idealized Fisher matrix forecast for a DES-like survey we find that, compared to $\xi_\pm$, a combined $\xi_\pm$ & $\zeta_\pm$ analysis can lead to improvements of order 20–40 per cent on the constraints of cosmological parameters such as σ_8 or the dark energy equation of state parameter w_0. We find similar levels of improvement on the constraints of the baryonic feedback parameters, which strengthens the prospects for cosmic shear data to obtain tight constraints not only on cosmology but also on astrophysical feedback models. These encouraging results motivate future works on the integrated shear 3-point correlation function towards applications to real survey data.

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

1 INTRODUCTION

The *cosmic shear field* is the name given to the coherent distortion pattern on the shapes of distant background galaxies, that is induced by the weak gravitational lensing effect caused by the intervening matter distribution. Statistical analyses of cosmic shear data thus let us directly probe the large-scale structure in our Universe, and consequently, enable us to place tight constraints on the parameters of our cosmological models and address key questions such as the nature of dark energy, dark matter, and gravity. Indeed, cosmic shear data has already had a marked impact in cosmology, most notably with the recent analyses of the data from surveys like DES (Abbott et al. 2021), KiDS (Heymans et al. 2021), and HSC (Hikage et al. 2019), and this progress is expected to be taken to a new level when the data from the larger Euclid (Laureijs et al. 2011), Vera Rubin (LSST Dark Energy Science Collaboration 2012), and Nancy Roman (Spergel et al. 2015) surveys are analysed in the future.

The majority of the cosmic shear analyses performed to date are based on the 2-point correlation function (2PCF), $\xi_\pm(\alpha)$, i.e. the correlation between the cosmic shear field at two points separated by an angle α on the sky (in harmonic/Fourier space, this is called the power spectrum). However, the cosmic shear field is non-Gaussian distributed, and as a result, there is additional, independent information beyond the 2PCF that is crucial to access in order to maximize the constraining power of the data. The most natural first step beyond 2PCF analyses is to study the 3-point correlation function (3PCF) of the cosmic shear field (the bispectrum in harmonic/Fourier space). The 2PCF and 3PCF depend differently on the cosmological parameters, and so combined analyses of these two statistics allow us to break degeneracies and obtain tighter constraints on the parameter values (Takada & Jain 2004; Kayo & Takada 2013; Sato & Nishimichi 2013). The 3PCF is however appreciably more complicated than the 2PCF, which is why these analyses are not yet routine in the cosmic shear literature. For example, on the measurement side, the shear 3PCF lives in a higher-dimensional parameter space (it is a function of the size and angles of the sides of triangles connecting three points on the sky), which makes its estimation from observational data more challenging. Further, on the theory side, predicting the 3PCF requires accurate models for the three-dimensional matter bispectrum on small scales (Takahashi et al. 2020; Aricò et al. 2021), which is still a challenging enterprise.

These complications get further exacerbated by the need to also account for photometric redshift uncertainties, shear calibration, and masking, as well as galaxy intrinsic alignments and baryonic feedback. For these reasons, the first attempts to incorporate higher-order information into cosmic shear analyses have focused on simpler summary statistics, including mass aperture moments (Semboloni et al. 2010; Fu et al. 2014; Gatti et al. 2021; Martinet et al. 2021), lensing peaks (Kacprzak et al. 2016; Harnois-Déraps et al. 2020; Zürcher et al. 2022), or density-split statistics (Friedrich et al. 2018; Gruen et al. 2018; Burger et al. 2020, 2022).

Here, we focus our attention on a particularly promising way of accessing 3PCF information in the cosmic shear field using a statistic called the *integrated shear 3-point correlation function $\zeta_\pm(\alpha)$*, which has been described recently[1] in Halder et al. (2021) [see also Munshi et al. 2021, and for studies of its harmonic counterpart we refer to Munshi et al. (2020) and Jung et al. (2021)]. Concretely, $\zeta_\pm(\alpha)$ describes the correlation between (i) the *shear 2PCF* measured locally inside well-defined patches on the sky, and (ii) the *1-point aperture shear mass* of the patches. This statistic admits a very well-defined physical interpretation as the modulation of the small-scale shear 2PCF by long-wavelength features of the cosmic shear field.[2] This can be shown to be sensitive to a certain integral of the three-dimensional matter bispectrum (hence the name *integrated*), which is how one can access 3-point function information. A key practical advantage of this statistic is that it requires only measurements of the shear aperture mass and 2PCF, which can both be obtained from cosmic shear catalogues using existing, well-tested numerical algorithms.

As we will see below, the key theoretical ingredient to predict $\zeta_\pm$ is the three-dimensional matter bispectrum $B_\delta^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$, where the $\boldsymbol{k}_i$ are wavevectors in Fourier space. In Halder et al. (2021), this was calculated using the fitting function of Gil-Marín et al. (2012), which was fitted only on scales $k \lesssim 0.4\,h\,\mathrm{Mpc}^{-1}$, and as a result, it could not be used to describe the parts of $\zeta_\pm$ that get contributions from the non-linear regime of structure formation on smaller scales. In this paper, one of our goals is to remedy this by putting forward an alternative calculation of $\zeta_\pm$ that is accurate on small scales, and that can therefore maximize the utility of this statistic to constrain cosmology. As we will discuss below, the key observation behind our calculation is that, on small scales, the integrated cosmic shear 3PCF is dominated by the *squeezed limit* of the matter bispectrum, i.e. the limit in which one of the wave vectors $\boldsymbol{k}_i$ is much smaller than the other two. This is fortunate since this particular limit of the matter bispectrum can be described very efficiently and accurately using the *response approach to perturbation theory* developed by Barreira & Schmidt (2017a). The response approach is a rigorous extension of standard perturbation theory (SPT) (Bernardeau et al. 2002) that allows the evaluation of squeezed *N*-point interactions in the non-linear regime of structure formation. This semi-analytical approach takes as inputs the non-linear matter power spectrum and its *response functions* to long-wavelength perturbations, which are much easier to predict and calibrate using *N*-body simulations compared to the full non-linear matter bispectrum. One of our main results in this paper is the demonstration of the accuracy of the response approach to describe the integrated shear 3PCF $\zeta_\pm(\alpha)$

deep in the non-linear, small-scale regime of structure formation, by comparing against results from direct simulation of cosmic shear maps.

Another advantage of the response approach that we highlight and focus on in this paper concerns the relative ease with which the impact of baryonic feedback effects can be taken into account. On small distance scales ($k \gtrsim 1\,h\,\mathrm{Mpc}^{-1}$), baryonic effects such as the energy released by active galactic nuclei (AGN) inside dark matter haloes are known to have a marked impact on the small-scale cosmic shear field (Chisari et al. 2019), and cannot be ignored at the risk of obtaining strongly biased cosmological constraints (Semboloni et al. 2011; Eifler et al. 2015; Huang et al. 2019; Schneider et al. 2020). The size and time-dependence of the baryonic effects are however currently very unknown, which makes this one of the most serious modelling challenges in cosmic shear data analyses. One way around this problem is to simply discard the parts of the data that are expected to be affected by baryonic effects, but this is manifestly suboptimal. A more interesting approach involves describing the impact of baryonic effects on the theory predictions with a set of extra parameters that can be fitted alongside the cosmological ones. At the 2-point function level, there is already a significant amount of work devoted to the modelling of baryonic effects, including through empirical fitting formulae (Mohammed & Seljak 2014; Harnois-Déraps et al. 2015), analyses based on principal components (Eifler et al. 2015; Huang et al. 2019, 2021), extensions of the halo model (Semboloni et al. 2011; Mead et al. 2015, 2021), and *baryonification* techniques of gravity-only simulations (Schneider & Teyssier 2015; Schneider et al. 2020; Aricò et al. 2020). Owing to its extra complexity, and despite interesting first steps (Semboloni, Hoekstra & Schaye 2013; Foreman et al. 2020; Aricò et al. 2021; Takahashi et al. 2020), the same progress at the 3-point function level has naturally lagged behind. Herein lies the other advantage of the response approach: we will see below that $\zeta_\pm$ depends on baryonic physics effectively only via the non-linear matter power spectrum, and thus, the incorporation of the baryonic effects on $\zeta_\pm$ can be made as straightforward as that on the 2PCF $\xi_\pm$.

In this paper, in particular, we investigate the impact of baryonic effects on $\zeta_\pm$ with the aid of the HMCODE developed by Mead et al. (2015), which accounts for baryonic effects through two parameters that describe their impact on the internal structure of haloes. Using a Fisher matrix forecast analysis for a DES-sized tomographic survey, we will see that the combination of $\xi_\pm$ and $\zeta_\pm$ information can lead to significant improvements in cosmological parameter constraints, importantly, even after marginalizing over the baryonic feedback parameters. The size of the improvements can depend on the details of the forecast analyses and varies from one parameter to another, but our results show that $\zeta_\pm$ data has the potential to improve the constraints on parameters like σ_8 or the dark energy equation of state parameter w_0 by ≈ 20–40 per cent. Interestingly, we will also see that the addition of $\zeta_\pm$ helps to tighten the constraints on the baryonic feedback parameters themselves, which strengthens the opportunity for cosmic shear data to constrain not only cosmology but also astrophysical models of AGN feedback (Huang et al. 2021). This provides further motivation to include $\zeta_\pm$ information in constraint analyses of cosmic shear data, reinforcing the promising potential of this statistic that Halder et al. (2021) had highlighted before.

The rest of this paper is organized as follows. In Section 2, we review the formalism behind the integrated shear 3PCF $\zeta_\pm$, and introduce our theoretical framework to model the matter bispectrum using the response function approach, including the incorporation of baryonic feedback effects. In Section 3, we describe the main numerical details of the simulation data that we adopt from

---

[1]We drop the letter *i* from the notation $i\zeta_\pm$ used in Halder et al. (2021) to avoid confusion with the imaginary unit $i = \sqrt{-1}$.

[2]This statistic was first introduced by Chiang et al. (2014), Chiang et al. (2015) in the context of three-dimensional galaxy clustering analyses, where it admits a similar physical interpretation.

Halder et al. (2021) to demonstrate the accuracy of our theoretical model for $\zeta_\pm$. Section 4 contains our main numerical results and discussion: we begin by validating our theoretical model against the numerical simulations, and then go through the results of our Fisher matrix constraints. Finally, we summarize and conclude in Section 5.

2 THEORY

In this section, we outline the key concepts of the integrated shear 3PCF (Section 2.1) and of our model for the matter bispectrum based on the response function approach (Section 2.2). These topics have been introduced before in Halder et al. (2021) and Barreira & Schmidt (2017a), respectively, to which we refer the reader for more details and derivations. We also discuss how we incorporate baryonic effects in our theoretical predictions using the HMCODE formalism of Mead et al. (2015) (Section 2.3).

2.1 Integrated shear 3-point correlation function

The integrated shear 3PCF is defined as

$$\zeta_{\pm,\text{fgh}}(\boldsymbol{\alpha}) \equiv \left\langle M_{\text{ap,f}}(\boldsymbol{\theta}_C)\, \hat{\xi}_{\pm,\text{gh}}(\boldsymbol{\alpha};\boldsymbol{\theta}_C) \right\rangle, \tag{1}$$

where $\langle \rangle$ denotes ensemble averaging, $M_{\text{ap}}(\boldsymbol{\theta}_C)$ is the 1-point aperture mass statistic in some patch of the sky centred at $\boldsymbol{\theta}_C$, and $\hat{\xi}_{\pm}(\boldsymbol{\alpha};\boldsymbol{\theta}_C)$ is the local shear 2PCF evaluated in the same patch. The subscripts $_{\text{f, g, and h}}$ denote the galaxy source redshift bin; for example $\hat{\xi}_{\pm,\text{gh}}$ is the 2-point cross-correlation function of the fields γ_{g}, γ_{h}, which are respectively, the cosmic shear fields estimated from galaxy shapes at redshifts g and h. It is in the sense of the correlation in this equation that we can identify $\zeta_\pm$ as describing the modulation of the small-scale shear 2-point correlation function $\xi_\pm$ by the local shear mass M_{ap}. We discuss next the two ingredients that enter the right-hand side of equation (1).

The 1-point aperture mass statistic $M_{\text{ap}}(\boldsymbol{\theta}_C)$ measures the weighted lensing convergence field $\kappa(\boldsymbol{\theta})$ inside an aperture U centred at $\boldsymbol{\theta}_C$ (Kaiser 1995; Schneider 1996, 2006):

$$M_{\text{ap}}(\boldsymbol{\theta}_C) = \int \mathrm{d}^2\boldsymbol{\theta}\, \kappa(\boldsymbol{\theta})\, U(\boldsymbol{\theta}_C - \boldsymbol{\theta}), \tag{2}$$

where $U(\boldsymbol{\theta}) = U(\theta)$ is an azimuthally symmetric filter with size θ_{ap}. The κ and γ fields are related to each other in Fourier space as

$$\gamma(\boldsymbol{l}) = e^{2i\phi_l}\kappa(\boldsymbol{l}), \tag{3}$$

where $\boldsymbol{l}$ is the 2D Fourier wave vector and ϕ_l is its polar angle (note we always work in the flat-sky limit). The convergence field is not directly observable, but interestingly, if U is a compensated filter, i.e. $\int \mathrm{d}^2\boldsymbol{\theta}\, U(\boldsymbol{\theta}_C - \boldsymbol{\theta}) = 0$, then the aperture mass can be directly evaluated from the observed shear field γ as (Kaiser 1995; Schneider 1996, 2006):

$$M_{\text{ap}}(\boldsymbol{\theta}_C) = \int \mathrm{d}^2\boldsymbol{\theta}\, \gamma_{\text{t}}(\boldsymbol{\theta}, \phi_{\boldsymbol{\theta}_C-\boldsymbol{\theta}})\, Q(\boldsymbol{\theta}_C - \boldsymbol{\theta}), \tag{4}$$

where $\gamma_{\text{t}}(\boldsymbol{\theta}, \phi_{\boldsymbol{\theta}_C-\boldsymbol{\theta}})$ is the tangential component of the shear at location $\boldsymbol{\theta}$ defined with respect to the polar angle $\phi_{\boldsymbol{\theta}_C-\boldsymbol{\theta}}$ of the separation vector between $\boldsymbol{\theta}$ and the centre of the aperture $\boldsymbol{\theta}_C$. This equation highlights the ease with which one can actually measure M_{ap} from the observed shear field without having to go through the process of creating a convergence mass map (cf. equation 2). In this paper, we consider the following forms of the filters U and Q

(Crittenden et al. 2002; Kilbinger & Schneider 2005):

$$U(\theta) = \frac{1}{2\pi\theta_{\text{ap}}^2} \left(1 - \frac{\theta^2}{2\theta_{\text{ap}}^2}\right) \exp\left(-\frac{\theta^2}{2\theta_{\text{ap}}^2}\right),$$

$$Q(\theta) \equiv -U(\theta) + \frac{2}{\theta^2} \int_0^\theta \mathrm{d}\theta'\, \theta' U(\theta')$$

$$= \frac{\theta^2}{4\pi\theta_{\text{ap}}^4}\, \exp\left(-\frac{\theta^2}{2\theta_{\text{ap}}^2}\right). \tag{5}$$

Below we will need U in Fourier space, where it is given by

$$U(\boldsymbol{l}) = U(l) = \int \mathrm{d}^2\boldsymbol{\theta}\, U(\theta)e^{-i\boldsymbol{l}\cdot\boldsymbol{\theta}} = \frac{l^2\theta_{\text{ap}}^2}{2}\, \exp\left(-\frac{l^2\theta_{\text{ap}}^2}{2}\right). \tag{6}$$

The second ingredient in equation (1) is $\hat{\xi}_\pm(\boldsymbol{\alpha};\boldsymbol{\theta}_C)$: the 2PCF of the *windowed* cosmic shear field $\gamma(\boldsymbol{\theta};\boldsymbol{\theta}_C) \equiv \gamma(\boldsymbol{\theta})W(\boldsymbol{\theta}_C - \boldsymbol{\theta})$, with the window function taken to be a top-hat of size θ_T centred at $\boldsymbol{\theta}_C$

$$W(\boldsymbol{\theta}) = W(\theta) = \begin{cases} 1 & \theta \leq \theta_T \\ 0 & \theta > \theta_T \end{cases}. \tag{7}$$

The Fourier transform of the windowed shear field is thus given by

$$\gamma(\boldsymbol{l};\boldsymbol{\theta}_C) = \int \frac{\mathrm{d}^2\boldsymbol{l}'}{(2\pi)^2}\gamma(\boldsymbol{l}')W(\boldsymbol{l}' - \boldsymbol{l})e^{i(\boldsymbol{l}'-\boldsymbol{l})\cdot\boldsymbol{\theta}_C}, \tag{8}$$

with

$$W(\boldsymbol{l}) = W(l) = \int \mathrm{d}^2\boldsymbol{\theta}\, W(\theta)e^{-i\boldsymbol{l}\cdot\boldsymbol{\theta}} = 2\pi\theta_T^2\, \frac{J_1(l\theta_T)}{l\theta_T}, \tag{9}$$

and where J_n is the nth-order ordinary Bessel function of the first kind. The two local *position-dependent* 2PCFs that appear in equation (1) are defined as

$$\hat{\xi}_+(\boldsymbol{\alpha};\boldsymbol{\theta}_C) \equiv \frac{1}{A_{\text{2pt}}(\boldsymbol{\alpha})} \int \mathrm{d}^2\boldsymbol{\theta}\, \gamma(\boldsymbol{\theta};\boldsymbol{\theta}_C)\gamma^*(\boldsymbol{\theta} + \boldsymbol{\alpha};\boldsymbol{\theta}_C)$$

$$\hat{\xi}_-(\boldsymbol{\alpha};\boldsymbol{\theta}_C) \equiv \frac{1}{A_{\text{2pt}}(\boldsymbol{\alpha})} \int \mathrm{d}^2\boldsymbol{\theta}\, \gamma(\boldsymbol{\theta};\boldsymbol{\theta}_C)\gamma(\boldsymbol{\theta} + \boldsymbol{\alpha};\boldsymbol{\theta}_C)e^{-4i\phi_\alpha}, \tag{10}$$

where ϕ_α is the polar angle of the spatial separation vector $\boldsymbol{\alpha}$, $*$ denotes complex conjugation, and $A_{\text{2pt}}(\boldsymbol{\alpha}) \equiv \int \mathrm{d}^2\boldsymbol{\theta}\, W(\boldsymbol{\theta}_C - \boldsymbol{\theta})W(\boldsymbol{\theta}_C - \boldsymbol{\theta} - \boldsymbol{\alpha})$ is the area normalization factor. For our isotropic window function $W(\boldsymbol{\theta}) = W(\theta)$, it follows that both the normalization term and $\hat{\xi}_\pm(\boldsymbol{\alpha};\boldsymbol{\theta}_C)$ depend only on the magnitude α of the separation vector, i.e. $A_{\text{2pt}}(\boldsymbol{\alpha}) = A_{\text{2pt}}(\alpha)$ and $\hat{\xi}_\pm(\boldsymbol{\alpha};\boldsymbol{\theta}_C) = \hat{\xi}_\pm(\alpha;\boldsymbol{\theta}_C)$.

In this paper, we are also interested in the global 2PCF of the whole cosmic shear field $\gamma(\boldsymbol{\theta})$ [i.e. not just the windowed one $\gamma(\boldsymbol{\theta};\boldsymbol{\theta}_C)$], which can be written in terms of the convergence power spectrum $P_\kappa(l)$ through inverse Hankel transforms [e.g. see appendix A of Halder et al. (2021)]

$$\xi_{+,\text{gh}}(\alpha) = \int \frac{\mathrm{d}l\, l}{2\pi}\, P_{\kappa,\text{gh}}(l)\, J_0(l\alpha),$$

$$\xi_{-,\text{gh}}(\alpha) = \int \frac{\mathrm{d}l\, l}{2\pi}\, P_{\kappa,\text{gh}}(l)\, J_4(l\alpha), \tag{11}$$

where

$$P_{\kappa,\text{gh}}(l) = \int \mathrm{d}\chi \frac{q_{\text{g}}(\chi)q_{\text{h}}(\chi)}{\chi^2} P_\delta^{\text{3D}}\left(k = \frac{l}{\chi}, \chi\right), \tag{12}$$

and P_δ^{3D} denotes the three-dimensional matter power spectrum (note that throughout this paper, we assume a flat cosmology). The lensing kernel functions are given by

$$q_{\text{f}}(\chi) = \frac{3H_0^2\Omega_{\text{m,0}}}{2c^2} \frac{\chi}{a(\chi)} \frac{\chi_s^{\text{f}} - \chi}{\chi_s^{\text{f}}}, \tag{13}$$

where χ is the comoving distance, $\chi_s^{\rm f}$ is the comoving distance out to the galaxies in source redshift bin f, $a(\chi)$ is the scale factor, $\Omega_{\rm m, 0}$ is the fractional cosmic matter density today, $H_0 = 100\, h\,{\rm km\,s^{-1}\,Mpc^{-1}}$ is the Hubble expansion rate today, and c is the speed of light; we assume for simplicity that the galaxies in each tomographic bin are all at a single source redshift, but it is straightforward to generalize beyond this by writing $q_{\rm f}(\chi)$ for a general distribution of source galaxies in a tomographic redshift bin (e.g. see Schneider 2006). When we evaluate the equations above, we apply the l-dependent correction of Kitching et al. (2017) to correct for flat-sky and Limber approximation effects [see, for example, equation (60) of Halder et al. (2021)]. Furthermore, rather than using the inverse Hankel transform integrals directly in equation (11), we use the expressions with summation over l as given in Friedrich et al. (2021) (see their equation 9), which are exact in the curved-sky case and more accurate in that they take into account the finite bin widths in which the correlations are measured in the data. We summarize these auxiliary equations in Appendix A.

Finally, putting all the ingredients together and following the derivation of Halder et al. (2021), the two integrated 3PCFs in equation (1) can be written as

$$\zeta_{+,\mathrm{fgh}}(\alpha) = \frac{1}{A_{\mathrm{2pt}}(\alpha)} \int \frac{\mathrm{d}l\, l}{2\pi}\, \mathcal{B}_{+,\mathrm{fgh}}(l)\, J_0(l\alpha),$$

$$\zeta_{-,\mathrm{fgh}}(\alpha) = \frac{1}{A_{\mathrm{2pt}}(\alpha)} \int \frac{\mathrm{d}l\, l}{2\pi}\, \mathcal{B}_{-,\mathrm{fgh}}(l)\, J_4(l\alpha), \qquad (14)$$

where the integrated shear bispectra read

$$\mathcal{B}_{\pm,\mathrm{fgh}}(\boldsymbol{l}) = \int \mathrm{d}\chi\, \frac{q_{\mathrm{f}}(\chi)q_{\mathrm{g}}(\chi)q_{\mathrm{h}}(\chi)}{\chi^4} \int \frac{\mathrm{d}^2\boldsymbol{l}_1}{(2\pi)^2} \int \frac{\mathrm{d}^2\boldsymbol{l}_2}{(2\pi)^2}$$
$$\times\, B_\delta^{\mathrm{3D}}\left(\frac{\boldsymbol{l}_1}{\chi}, \frac{\boldsymbol{l}_2}{\chi}, \frac{-\boldsymbol{l}_1 - \boldsymbol{l}_2}{\chi}, \chi\right) e^{2i(\phi_2 \mp \phi_{-1-2})}$$
$$\times\, U(\boldsymbol{l}_1) W(\boldsymbol{l}_2 + \boldsymbol{l}) W(-\boldsymbol{l}_1 - \boldsymbol{l}_2 - \boldsymbol{l}). \qquad (15)$$

In this equation, B_δ^{3D} denotes the three-dimensional matter bispectrum, ϕ_1 and ϕ_2 are the polar angles of the Fourier modes $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$, respectively, and ϕ_{-1-2} is the polar angle of $-\boldsymbol{l}_1 - \boldsymbol{l}_2$. For our isotropic window functions U and W, these integrated shear 3PCFs and bispectra are direction independent, i.e. $\zeta_{\pm,\mathrm{fgh}}(\boldsymbol{\alpha}) = \zeta_{\pm,\mathrm{fgh}}(\alpha)$ and $\mathcal{B}_{\pm,\mathrm{fgh}}(\boldsymbol{l}) = \mathcal{B}_{\pm,\mathrm{fgh}}(l)$, respectively. As with the 2PCF, we again use the l-summation strategy of Friedrich et al. (2021) instead of the direct inverse Hankel transforms in order to convert the integrated shear bispectra to the real space correlation functions.

2.2 Model for the matter bispectrum

The evaluation of the integrated shear bispectrum in equation (15) requires predicting the three-dimensional non-linear matter bispectrum B_δ^{3D}, which is defined as

$$\langle \delta_m(\boldsymbol{k}_1)\delta_m(\boldsymbol{k}_2)\delta_m(\boldsymbol{k}_3) \rangle = (2\pi)^3 \delta_D(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) B_\delta^{\mathrm{3D}}(k_1, k_2, k_3), \qquad (16)$$

where $\delta_m(\boldsymbol{k})$ is the Fourier transform of the three-dimensional matter density contrast. In standard perturbation theory (SPT), the tree-level matter bispectrum is given by Bernardeau et al. (2002):

$$B_{\delta,\mathrm{tree}}^{\mathrm{3D}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \tau) = 2\, F_2(\boldsymbol{k}_1, \boldsymbol{k}_2)\, P_{\delta,L}^{\mathrm{3D}}(k_1, \tau)\, P_{\delta,L}^{\mathrm{3D}}(k_2, \tau)$$
$$+ \text{cyclic permutations}, \qquad (17)$$

where $P_{\delta,L}^{\mathrm{3D}}(k, \tau)$ is the three-dimensional linear matter power spectrum and $F_2(\boldsymbol{k}_i, \boldsymbol{k}_j)$ is the symmetrized two-point mode coupling

kernel:

$$F_2(\boldsymbol{k}_i, \boldsymbol{k}_j) = \frac{5}{7} + \frac{1}{2}\mu_{\boldsymbol{k}_i, \boldsymbol{k}_j}\left(\frac{k_i}{k_j} + \frac{k_j}{k_i}\right) + \frac{2}{7}\mu_{\boldsymbol{k}_i, \boldsymbol{k}_j}^2, \qquad (18)$$

where $\mu_{\boldsymbol{k}_i, \boldsymbol{k}_j} \equiv \boldsymbol{k}_i \cdot \boldsymbol{k}_j / (k_i k_j)$ is the cosine of the angle between the two Fourier modes $\boldsymbol{k}_i$ and $\boldsymbol{k}_j$. This expression is only valid in the weakly non-linear regime of structure formation, and it is therefore insufficient to accurately model the integrated shear bispectrum, as we will see below.

In order to model B_δ^{3D} in the non-linear regime, one needs to either go beyond tree-level in perturbation theory and consider the one-loop or two-loop bispectrum (Lazanu & Liguori 2018; Baldauf et al. 2021), or rely on fitting formulae calibrated using N-body simulations (Scoccimarro & Couchman 2001; Gil-Marín et al. 2012; Takahashi et al. 2020). In their previous work on the integrated shear 3PCF, Halder et al. (2021) used the bispectrum fitting formula from Gil-Marín et al. (2012) (hereafter referred to as GM), which can be written as

$$B_{\delta,\mathrm{GM}}^{\mathrm{3D}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \tau) = 2\, F_2^{\mathrm{eff}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \tau)\, P_\delta^{\mathrm{3D}}(k_1, \tau)\, P_\delta^{\mathrm{3D}}(k_2, \tau)$$
$$+ \text{cyclic permutations}, \qquad (19)$$

where $P_\delta^{\mathrm{3D}}(k, \tau)$ is the three-dimensional non-linear matter power spectrum and $F_2^{\mathrm{eff}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \tau)$ is the following modified version of the F_2 kernel:

$$F_2^{\mathrm{eff}}(\boldsymbol{k}_i, \boldsymbol{k}_j, \tau) = \frac{5}{7}a(k_i, \tau)a(k_j, \tau)$$
$$+ \frac{1}{2}\mu_{\boldsymbol{k}_i, \boldsymbol{k}_j}\left(\frac{k_i}{k_j} + \frac{k_j}{k_i}\right) b(k_i, \tau)b(k_j, \tau)$$
$$+ \frac{2}{7}\mu_{\boldsymbol{k}_i, \boldsymbol{k}_j}^2 c(k_i, \tau)c(k_j, \tau), \qquad (20)$$

where $a(k, \tau)$, $b(k, \tau)$, and $c(k, \tau)$ are fitting functions calibrated using measurements of the matter bispectrum from gravity-only N-body simulations up to wave numbers $k < 0.4\, h\,\mathrm{Mpc^{-1}}$ (see Gil-Marín et al. 2012 for the form and parameters of these functions).

The work of Halder et al. (2021) showed that although the GM fitting function $B_{\delta,\mathrm{GM}}^{\mathrm{3D}}$ is able to describe the $\zeta_+(\alpha)$ correlation measured from simulations very well down to angular scales of $\alpha \approx 5$ arcmin, the same is not true for the $\zeta_-(\alpha)$ case, for which the GM function begins to breakdown on scales of a few tenths of arcmin (we reproduce this result below in Fig. 4). This is because ζ_- is more sensitive to the non-linear regime of structure formation and the GM fitting formula was calibrated only on quasi-linear scales ($k < 0.4\, h\,\mathrm{Mpc^{-1}}$) with applications to galaxy clustering observations in mind. One way to improve upon this is to use the recent `bihalofit` formula for the matter bispectrum from Takahashi et al. (2020), which was calibrated using the matter bispectrum from simulations in the non-linear regime, and which Halder et al. (2021) showed does describe well both the $\zeta_+(\alpha)$ and $\zeta_-(\alpha)$ statistics measured from gravity-only simulations (see fig. D3 there, but we also reproduce this result in Fig. 4).

The development of `bihalofit` is an important step forward in our ability to predict the matter bispectrum in the non-linear regime of structure formation, but in its current form it cannot still be readily used to account for the impact of baryonic effects on small scales. Note that Takahashi et al. (2020) do provide a baryon-ratio formula that accounts for the specific impact of baryonic effects in the IllustrisTNG galaxy formation model (Weinberger et al. 2017; Pillepich et al. 2018), but for applications to cosmic shear data we need to be able to make predictions as a function of the baryonic feedback parameters (and not just a single set such as IllustrisTNG) that we can then marginalize over. We will see below how this is

something that can be straightforwardly achieved with the response approach to perturbation theory.

2.2.1 Response approach to the squeezed matter bispectrum

The response approach to perturbation theory developed by Barreira & Schmidt (2017a) is a formalism that allows to evaluate certain mode-coupling terms in SPT in the non-linear regime. The first step of the response approach involves noting that the small-scale matter power spectrum can be regarded as a biased tracer of large-scale structure (see Desjacques, Jeong & Schmidt 2018 for a review on biasing), i.e. it can be expanded as

$$P_\delta^{3D}(\boldsymbol{k}, \tau; \boldsymbol{x}) = P_\delta^{3D}(k, \tau)\big[1 + R_1(k, \tau)\delta_m^L(\boldsymbol{x}, \tau)$$
$$+ R_K(k, \tau)\hat{k}_i\hat{k}_j K_{ij}^L(\boldsymbol{x}, \tau)\big], \quad (21)$$

where $P_\delta^{3D}(\boldsymbol{k}, \tau; \boldsymbol{x})$ is the *local* power spectrum measured in some volume around position $\boldsymbol{x}$, $P_\delta^{3D}(k, \tau)$ is its *global* cosmic average, δ_m^L is a large-scale isotropic matter density perturbation, and $K_{ij}^L = \big[\partial_i\partial_j/\nabla^2 - \delta_{ij}/3\big]\delta_m^L$ is a large-scale tidal field; the superscript ^{L} indicates that these are large-scale perturbations that are in the linear/quasi-linear regime of structure formation.[3] The coefficients R_1 and R_K are called first-order *power spectrum response functions*, and describe physically the response of the small-scale matter power spectrum to the presence of large-scale overdensities and tidal fields, respectively. These response functions can be written as (Li, Hu & Takada 2014; Wagner et al. 2015; Barreira & Schmidt 2017a)

$$R_1(k, \tau) = 1 - \frac{1}{3}\frac{d\ln P_\delta^{3D}(k, \tau)}{d\ln k} + G_1(k, \tau), \quad (22)$$

$$R_K(k, \tau) = G_K(k, \tau) - \frac{d\ln P_\delta^{3D}(k, \tau)}{d\ln k}, \quad (23)$$

where G_1 and G_K are so-called *growth-only* response functions, which can be measured very efficiently in the non-linear regime of structure formation using separate universe simulations. In this paper, we use the measurements of G_1 from Wagner et al. (2015) and the measurements of G_K from Schmidt et al. (2018) (see also Stücker et al. 2021); the time and scale-dependence of the resulting $R_1(k, \tau)$, $R_K(k, \tau)$ functions can be seen alongside one another in fig. 1 of Barreira, Krause & Schmidt (2018).

The second step of the response approach is the realization that certain combinations of power spectrum response functions can be identified as *resummed* perturbation theory kernels in the squeezed limit. Concretely, for the case of the matter bispectrum that we are interested in here, we can write (see Barreira & Schmidt 2017a for the derivation)

$$B_{\delta, RF}^{3D}(\boldsymbol{k}_s, \boldsymbol{k}_h, -\boldsymbol{k}_{sh}, \tau) = \left[R_1(k_h, \tau) + \left(\mu_{\boldsymbol{k}_h, \boldsymbol{k}_s}^2 - \frac{1}{3}\right)R_K(k_h, \tau)\right]$$
$$\times P_\delta^{3D}(k_h, \tau)P_{\delta, L}^{3D}(k_s, \tau) + \mathcal{O}\left[\frac{k_s^2}{k_h^2}\right], \quad (24)$$

where the Fourier mode $\boldsymbol{k}_s$ is called a *soft* mode (it describes large scales), $\boldsymbol{k}_h$ is called a *hard* mode (it describes small scales), and $-\boldsymbol{k}_{sh} = -\boldsymbol{k}_s - \boldsymbol{k}_h$; the subscript $_{RF}$ stands for *response function*. This equation is valid strictly in the squeezed limit, i.e. $k_s \ll k_h \approx |-\boldsymbol{k}_{sh}|$, with the corrections scaling as k_s^2/k_h^2. By comparing this equation to equation (17), we note that the term in squared brackets R_1

---

[3]This expansion implicitly assumes that the *local* power spectrum is measured within a volume V_{loc} that is sufficiently inside the large-scale perturbations, i.e. $V_{loc}^{1/3} \ll S$, where S is the wavelength of the δ_m^L, K_{ij}^L perturbations.

$+ (\mu^2 - 1/3)R_K$ can be identified as a generalized F_2 SPT kernel, the power spectrum of the hard mode is now the non-linear matter power spectrum, but the power spectrum of the soft mode must remain in the linear regime. With equation (24), we can thus evaluate the squeezed matter bispectrum for k_s in the linear regime, but importantly since we use results for R_1, R_K, and P_δ^{3D} obtained using *N*-body simulations, the result is valid for non-linear values of the hard mode k_h; it is in this sense that the response approach extends the validity of SPT to the non-linear regime.

For the case of the bispectrum, the response functions R_1 and R_K are the only two that are needed, but we note for completeness that the response approach can be also used to evaluate terms that contribute to correlation functions beyond 3-point by letting the expansion of equation (21) to continue to higher-order (i.e. including terms like $R_2[\delta_m^L]^2$): for example, Barreira & Schmidt (2017b) and Barreira et al. (2018) used the response approach to calculate the covariance of the matter power spectrum (which is a 4-point function), and Barreira (2019) used it to calculate the covariance of the squeezed matter bispectrum (which contains terms up to the 6-point function).

2.2.2 The joint model for the non-linear matter bispectrum

We have not yet discussed what is especial about the ability of the response approach to predict the squeezed-limit bispectrum in the non-linear regime, i.e. why is it sufficient to evaluate this particular limit in the non-linear regime, but not the remainder of the bispectrum configurations? The answer to this question rests on the observation that, *on small scales (high-l), the integrated lensing bispectrum $\mathcal{B}_\pm$ in equation (15) is dominated by the squeezed limit of the three-dimensional matter bispectrum B_δ^{3D}*. We will verify this explicitly numerically below (see also appendix D of Halder et al. 2021), but the form of equation (15) can be used already to understand the reason why. The key point to note is that the product of the window functions $U(\boldsymbol{l}_a)$ and $W(\boldsymbol{l}_a)$ works as a low-pass filter, i.e. it becomes small whenever $\boldsymbol{l}_a$ describes scales smaller than the scale of the patches. Concretely, the $U(\boldsymbol{l}_1)$ term ensures the integral is sizeable only if $l_1 \sim \sqrt{2}/\theta_{ap} < 2\pi/\theta_{ap}$, where recall θ_{ap} is the size of the aperture mass compensated filter. Moreover, if $l \gg 2\pi/\theta_T$, i.e. we are interested in evaluating $\mathcal{B}_\pm(l)$ for modes well within the top-hat patch W, then the window function term $W(\boldsymbol{l}_2 + \boldsymbol{l})$ effectively constrains $\boldsymbol{l} \approx -\boldsymbol{l}_2$, which is in turn much larger in amplitude than $l_1 \lesssim 2\pi/\theta_{ap}$ since $\theta_{ap} \approx \theta_T$ (we consider both to have similar sizes of order 70 arcmin; cf. Section 3.3). This enforces the hierarchy $|-\boldsymbol{l}_1 - \boldsymbol{l}_2| \approx |\boldsymbol{l}_2| \gg |\boldsymbol{l}_1|$ between the bispectrum modes in equation (15), i.e. the bispectrum is pushed to the squeezed limit, which is why the response approach calculation is sufficient for these high-*l* modes. In other words, if $l \gg 2\pi/\theta_T$, the contributions from non-squeezed configurations that arise as one integrates over $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ (and which cannot be described by the response approach) are negligible as they are suppressed by the window function terms.

On the other hand, for *l* values comparable to the scale of the patches, $l \lesssim 2\pi/\theta_T$, the term $W(\boldsymbol{l}_2 + \boldsymbol{l})$ no longer enforces $l_2 \gg l_1$, the bispectrum thus contributes through non-squeezed configurations, and the response approach is not applicable. Importantly, however, if the typical size of the patches θ_{ap}, θ_T is large enough, then for $l \sim 2\pi/\theta_T$, we are in a regime where a fitting formula like the GM is able to provide accurate results.

Following these observations, the model of the non-linear matter bispectrum that we utilize in this paper corresponds to the following *stitching* of the GM and RF approach expressions:

(i) Given the lengths of 3 Fourier modes k_1, k_2, and k_3 at which we want to evaluate the bispectrum $B_\delta^{3D}(k_1, k_2, k_3, \tau)$, we arrange the modes in descending order and name them k_h, k_m, and k_s such that $k_h \geq k_m \geq k_s$.

(ii) We quantify the *squeezeness* of a given configuration by the parameter $f_{sq} \equiv k_m/k_s$, where the larger the value, the more squeezed the triangle is.

(iii) We evaluate the matter bispectrum as

$$B_\delta^{3D}(k_1, k_2, k_3, \tau) = \begin{cases} B_{\delta,RF}^{3D} & , \quad f_{sq} \geq f_{sq}^{thr} \Rightarrow \text{squeezed} \\ \\ B_{\delta,GM}^{3D} & , \quad \text{otherwise} \end{cases}, \qquad (25)$$

where f_{sq}^{thr} is a parameter that sets the threshold above which we dub a given triangle as squeezed and evaluate the bispectrum using the response approach.

The optimal choice for f_{sq}^{thr} is determined by a balance between the accuracy of the response approach and GM results at the edge of their regimes of validity. On the one hand, if f_{sq}^{thr} is chosen too low, then the response approach result will not be as accurate since the triangle is not very squeezed; in equation (24), the corrections to the response result scale as $(k_s/k_h)^2 \approx (k_s/k_m)^2 = 1/f_{sq}^2$, which become larger as $f_{sq}^{thr} \to 1$. On the other hand, for large values of f_{sq}^{thr}, the GM branch will be switched on and contribute sizeably for squeezed configurations in the non-linear regime, for which the GM fitting formula becomes less accurate; for instance, in the limit of $f_{sq}^{thr} \to \infty$, the response calculation is never used and we are left in the situation where we always use the GM formula [equivalent to the modelling setup studied in Halder et al. (2021)]. We will return to these considerations below when we examine the impact of different choices for f_{sq}^{thr}.

As we will see below, the sharp transition between the two branches in equation (25) does not translate into any visible discontinuous artefacts in the numerical predictions for $\zeta_\pm(\alpha)$ as they are effectively smoothed out by the integrals in equation (15). A smoother and continuous transition between the two branches could none the less be devised, but we leave this for future work. Note also that the GM formula serves in our calculation as a representative of any matter bispectrum calculation that is accurate in the quasi-linear regime; for instance, our main conclusions in this paper hold equally if instead of GM we had used the `bihalofit` formula.

2.3 Baryonic effects

The incorporation of baryonic feedback effects in our theoretical predictions for the shear 2PCF and integrated shear 3PCF can be done through their impact on the three-dimensional matter power spectrum P_δ^{3D} and bispectrum B_δ^{3D} in equations (12) and (15), respectively. In this paper, we use the `HMCODE` framework of Mead et al. 2015 to model the impact of baryons on P_δ^{3D}. The `HMCODE` is a modified version of the halo model (see Cooray & Sheth 2002 for a review) that introduces two parameters η_0 and c_{min} that can be varied to mimic the typical impact from baryonic physics effects; primarily the impact of adiabatic contraction by radiative cooling in the inner parts of haloes, and the strength of AGN feedback that can expel gas to large radii and suppress the amplitude of the power spectrum on scales $k \gtrsim 1\,h\,\mathrm{Mpc}^{-1}$. In Mead et al. (2015), these parameters were shown to provide a reasonable description of the power spectrum measured from the OWLS (Schaye et al. 2010) suite of hydrodynamical simulations that include these baryonic physics. We use the `HMCODE` implementation in the publicly available

Boltzmann solver code `CLASS`[4] (Blas, Lesgourgues & Tram 2011). The `HMCODE` code has been recently upgraded in Mead et al. (2021) and it was used by the KiDS collaboration to model baryonic effects in their cosmic shear data analysis (Heymans et al. 2021).

At the matter bispectrum level, Foreman et al. (2020) have recently studied the impact of baryonic physics in a series of different hydrodynamical simulations, and the `bihalofit` fitting formulae of Takahashi et al. (2020) also admits the impact of the baryonic physics in the IllustrisTNG galaxy formation model; as we noted already above, this does not yet allow to make predictions as a function of different baryonic physics parameters, which is what is needed to marginalize over their uncertain impact in cosmic shear data analyses. An interesting step in this direction, however, has been taken recently by Aricò et al. (2021) who showed that the *baryonification* approach (Schneider & Teyssier 2015; Schneider et al. 2020) is able to reproduce well the bispectrum measured in a series of different hydrodynamical simulations. This is an interesting way forward that allows to predict the matter bispectrum as a function of cosmological and baryonic feedback parameters, but which has not yet been realized in the form of a concrete code for fast numerical predictions [e.g. an emulator like the one developed in Aricò et al. (2021) for the matter power spectrum].

Fortunately, for the case of the integrated shear 3PCF that we focus on here, we can build on the work of Barreira et al. (2019), who showed that the impact of baryonic effects on the squeezed matter bispectrum can be effectively predicted from that on the matter power spectrum alone. Concretely, from equations (22), (23), and (24), we observe that baryonic effects impact the squeezed matter bispectrum only through P_δ^{3D}, and the growth-only responses G_1 and G_K. The functions G_1 and G_K, however, are expected to depend only very weakly on the baryonic physics effects. This is because they measure the dependence of the power spectrum on the large-scale environment, which is affected by baryonic physics to a much smaller degree compared to the impact of baryonic effects on the power spectrum itself. Indeed, using separate universe simulations with the IllustrisTNG galaxy formation model, Barreira et al. (2019) showed that the measured G_1 was virtually identical to the same measurements on gravity-only simulations.[5] The same has not been explicitly checked yet for G_K, but this reasoning suggests that it is also affected negligibly by baryonic effects. This allows us to straightforwardly account for baryonic effects on the squeezed matter bispectrum using also the `HMCODE` through its predictions for P_δ^{3D}; the ease with which baryonic effects can be propagated to the squeezed-limit matter bispectrum is one of the key advantages of using the response approach to predict the integrated shear 3PCF.

This addresses how we account for baryonic physics effects in the RF branch of equation (25) for squeezed configurations, but not in the GM branch for non-squeezed configurations. Here, we follow a strategy similar to Semboloni et al. (2013) who studied the

---

[4]Precisely, we use the c++ wrapper of the code (version v2.9.4) which can be obtained from the official repository, currently hosted at: https://github.com/lesgourg/class_public.

[5]Furthermore, Foreman et al. (2020) found that neglecting the impact of baryonic effects on G_1 leads also to very good agreement with the matter squeezed bispectrum measured in the Illustris (Vogelsberger et al. 2014) and EAGLE (Schaye et al. 2015) hydrodynamical simulations (see their Fig. 19). Interestingly, some differences were observed in the case of the BAHAMAS simulations (McCarthy et al. 2017), which the authors speculated could be due to a per cent-level impact of baryonic effects on G_1 in the BAHAMAS model. This is a small effect compared to the larger impact on the power spectrum itself, but would be interesting to investigate in the future.

impact of baryonic feedback on the third-order aperture mass statistic $\langle M_{ap}^3 \rangle$ using the matter bispectrum fitting formula of Scoccimarro & Couchman (2001) (a predecessor of GM): concretely, we account for baryonic effects through P_δ^{3D} in the GM formula of equation (19) using the HMCODE, but keep F_2^{eff} unchanged. This *ad hoc* solution does not affect our results significantly because the GM formula contributes primarily only on large scales that are not affected by the baryonic physics processes. Furthermore, Foreman et al. (2020) have also shown that this ad hoc strategy is actually able to reproduce well the impact of baryonic effects for equilateral configurations down to scales $k \sim 0.5 \, h \, \mathrm{Mpc}^{-1}$. That is, if there is a residual impact of baryonic physics on these configurations/scales, this may be actually well captured by this modification of the GM formula.

3 SIMULATIONS AND SURVEY SETUP

In this section, we briefly present the simulated cosmic shear data and the measurements that we use to test our theoretical predictions; this is essentially the same as in Halder et al. (2021) (see their Section 4 for more details).

3.1 T17 *N*-body simulations

We use the publicly available weak lensing data products from the cosmological simulations run by Takahashi et al. (2017).[6] In the following, we refer to these as the T17 simulations. These are a set of gravity-only cosmological *N*-body simulations run in periodic cubic boxes for a flat ΛCDM cosmology with the following parameters (we assume this as our fiducial cosmology): $\Omega_{cdm} = 0.233$, $\Omega_b = 0.046$, $\Omega_m = \Omega_{cdm} + \Omega_b = 0.279$, $\Omega_\Lambda = 0.721$, $h = 0.7$, $\sigma_8 = 0.82$, and $n_s = 0.97$. The particles in each simulation box were evolved from initial conditions using the *N*-body gravity solver GADGET2 (Springel, Yoshida & White 2001; Springel 2005) and they were then ray traced using the multiple-lens plane ray-tracing algorithm GRAYTRIX (Hamana et al. 2015; Shirasaki, Hamana & Yoshida 2015) to obtain 108 independent all-sky convergence/shear realizations for several source redshifts in HEALPIX format (Górski et al. 2005; Zonca et al. 2019) (see Takahashi et al. 2017 for more details). In this paper, we use the 108 full-sky weak lensing convergence and shear maps with NSIDE $= 4096$ (angular pixel scale of 0.86 arcmin) at source redshifts $z_1 = 0.5739$ and $z_2 = 1.0334$. In our results, here we make use of the correction formulae that Takahashi et al. (2017) put forward to account for numerical artefacts associated with the thickness of the lens planes, angular resolution, and finite simulation box size. We refer the reader to Takahashi et al. (2017) for more details or to appendix B of Halder et al. (2021) for a summary of these corrections.

3.2 FLASK Lognormal simulations

In order to estimate the covariance matrix of our data vector, we use the 1000 full-sky lognormal mock shear maps generated by Halder et al. (2021) using the publicly available FLASK tool.[7] Each mock consists of two lognormal shear fields simulated at source redshifts z_1 and z_2 in HEALPIX format with NSIDE $= 4096$. The mocks were created by fitting a lognormal PDF to the 1-point PDFs of the T17

convergence maps. The interested reader is referred to Section 4.2 of Halder et al. (2021) for the details about the creation of the FLASK mocks. The lognormal mocks from FLASK are noiseless and so in order to mimic realistic noise in weak lensing surveys, a complex *shape-noise* term $N(\theta) = N_1(\theta) + i N_2(\theta)$ is added to the shear field $\gamma(\theta)$ (Pires et al. 2020), where θ represents a pixel on the HEALPIX shear map. The noise components N_1, N_2 can both be modelled as uncorrelated Gaussian variables with zero mean and variance

$$\sigma_N^2 = \frac{\sigma_\epsilon^2}{n_g \cdot A_{pix}}, \tag{26}$$

where A_{pix} is the area of the pixel, σ_ϵ is the dispersion of intrinsic galaxy ellipticities which is set to 0.3, and n_g is the number of observed galaxies per squared arcminute, which we take to be 5 for both redshift bins. This is comparable to the expected number density of $n_g = 10 \, \mathrm{arcmin}^{-2}$ for the full DES Year 6 cosmic shear data.

3.3 Data vector and covariance matrix

We use the same data vector and covariance matrix measurements of the shear 2PCFs and the integrated shear 3PCFs obtained by Halder et al. (2021) using these simulation data products and the publicly available code TREECORR[8] (Jarvis, Bernstein & Jain 2004). We briefly review these measurements next.

The position-dependent shear 2PCFs $\hat{\xi}_{\pm, gh}(\alpha; \theta_C)$ were measured on the shear maps at source redshifts z_1 and z_2 (i.e. $g, h = 1, 2$) within top-hat windows W with radius $\theta_T = 75$ arcmin in 20 log-spaced angular bins within the range $\alpha \in [5, 140]$ arcmin. The aperture mass $M_{ap,f}(\theta_C)$ (with $f = 1, 2$) was measured using a compensated window U with an aperture scale $\theta_{ap} = 70$ arcmin. These *window* patches were distributed to cover the whole sky with only slight overlap between adjacent patches.

The global shear 2PCFs $\xi_{\pm, gh}(\alpha)$ in a given map were computed by averaging over the local position-dependent shear correlations evaluated inside all patches. On the other hand, the integrated shear 3PCFs $\zeta_{\pm, fgh}(\alpha)$ were evaluated by taking the average of the product of the aperture mass and the position-dependent shear 2PCFs over all the patches as in equation (1). The same measurements were performed on the various FLASK maps (with shape-noise), however, with patches distributed inside two big circular footprints of 5000 square degrees (approximately the size of the DES footprint) in each hemisphere of each FLASK map. This results in 2000 DES-like realizations which were used to estimate the covariance matrix.

The data vector evaluated from a single simulation realization (T17 or FLASK) consists of the shear 2PCFs and the integrated shear 3PCFs at the two source redshifts z_1 and z_2 (including the cross-correlation between redshift bins). The mean data vector was obtained by taking the average of the individual data vectors obtained from each of the 108 T17 realizations:

$$D \equiv \left(\xi_{\pm, 11}, \, \xi_{\pm, 12}, \, \xi_{\pm, 22}, \, \zeta_{\pm, 111}, \, \zeta_{\pm, 112}, \, \zeta_{\pm, 122}, \, \zeta_{\pm, 222} \right). \tag{27}$$

This data vector has $N_d = 7 \times 2 \times 20 = 280$ elements. On the other hand, the covariance of the data vector $\hat{\mathbf{C}}$ was estimated from the 2000 FLASK footprints [see fig. 3 of Halder et al. (2021) for the corresponding correlation matrix]. In our Fisher matrix analysis below, we will require the inverse covariance matrix, which we

---

[6]The data products of the simulation are available at http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/.

[7]Full-sky Lognormal Astro-fields Simulation Kit (FLASK) - currently hosted at http://www.astro.iag.usp.br/~flask/.

[8]Currently hosted at: https://rmjarvis.github.io/TreeCorr/_build/html/index.html#.
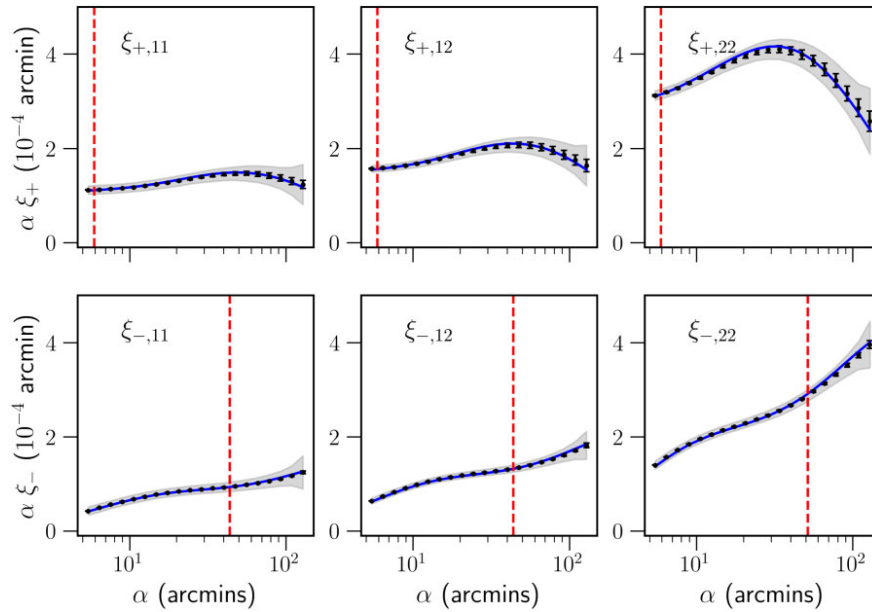
Figure 1. The shear 2PCFs $\xi_\pm(\alpha)$ for two tomographic source redshift bins $z_1 \approx 0.57$ and $z_2 \approx 1.03$. The black dots with the error bars show the mean and the standard deviation of the measurements from the 108 T17 all-sky simulation maps, respectively. The grey shaded regions indicate the standard deviation expected for these statistics in DES-sized footprints (obtained from the diagonal of our covariance matrix). The blue curves show the theoretical predictions of equation (11) obtained using the HMCODE. The result shown includes the numerical resolution corrections described in Takahashi et al. (2017). The red-dashed vertical lines mark the angular scale cuts that we apply in our Fisher matrix analysis to remove the parts of the data vector (on the left of the line) that are affected by baryonic feedback (see Section 4.2.1 for details).

evaluate as

$$\mathbf{C}^{-1} = \frac{N_r - N_d - 2}{N_r - 1} \, \hat{\mathbf{C}}^{-1}, \tag{28}$$

where $\hat{\mathbf{C}}^{-1}$ is the inverse of $\hat{\mathbf{C}}$ and the numerical prefactor corrects for the numerical bias of inverting a noisy covariance matrix estimated from a finite number of realizations $N_r = 2000$ (Hartlap, Simon & Schneider 2007). We refer the reader to appendix F of Halder et al. (2021) for validation checks of this covariance matrix calculation.

4 RESULTS

In this section, we discuss the performance of our model of the integrated shear 3PCF by comparing against the measurements from the T17 simulations. We also present the results of our Fisher matrix forecast analysis for a DES-sized survey, where we look into constraints on both cosmological and baryonic feedback parameters.

As shown in Section 2.1, in order to theoretically predict the 2PCFs $\xi_{\pm, \mathrm{gh}}(\alpha)$, and integrated shear 3PCFs $\zeta_{\pm, \mathrm{fgh}}(\alpha)$, we require the evaluation of the convergence power spectrum $P_{\kappa, \mathrm{gh}}(l)$ in equation (12), and the integrated shear bispectra $\mathcal{B}_{\pm, \mathrm{fgh}}(l)$ in equation (15). We numerically evaluate these spectra for 120 l-modes log-spaced in the range $1 \leq l \leq 20000$ and linearly interpolate between these 120 values to obtain the spectra at every other multipole. To evaluate $\mathcal{B}_{\pm, \mathrm{fgh}}(l)$, we use the Monte Carlo Vegas algorithm (Lepage 1980) from the GNU Scientific Library gsl[9] (Gough 2009) to perform the multidimensional integration in equation (15); we integrate l_1 and l_2 from 0 to 25000, and ϕ_1 and ϕ_2 from 0 to 2π.

---

[9]Currently hosted at: http://www.gnu.org/software/gsl/.

4.1 Comparison to simulations

Fig. 1 shows the $\xi_\pm$ components of the data vector. The black dots with error bars show the mean and standard deviation of the measurements from the 108 all-sky T17 maps. The grey shaded region indicates the standard deviation computed from the diagonal of the DES-like covariance matrix $\mathbf{C}$. The model vectors for $\xi_\pm$ computed with equations (11) and (12) are shown in blue; note these include the corrections proposed by Takahashi et al. 2017 to account for the various resolution effects of the T17 simulation. The $\xi_\pm$ predictions are in excellent agreement with the T17 measurements, and well within both the simulation error bars and DES uncertainty.

Fig. 2 shows the $\zeta_\pm$ components of the data vector. The black dots with error bars are again the measurements from the 108 T17 maps, and the grey shaded area shows the standard deviation obtained from the DES-like covariance matrix. The solid red lines show the analytical predictions for $\zeta_\pm$ obtained using equations (14) and (15), together with our GM + RF model of the non-linear matter bispectrum in equation (25) for $f_{\mathrm{sq}}^{\mathrm{thr}} = 7$. This value of $f_{\mathrm{sq}}^{\mathrm{thr}}$ was determined with a simple minimum χ^2 diagnostic using all of the $\zeta_\pm$ components of the data vector (we discuss the impact of $f_{\mathrm{sq}}^{\mathrm{thr}}$ below). The dashed lines show the same but with the GM branch in equation (25) artificially set to zero (i.e. only squeezed configurations contribute), and the dot–dashed line shows the outcome from setting the RF branch to zero instead (i.e. only non-squeezed configurations contribute). Indeed, as anticipated from our discussion in Section 2.2.2, the contributions from the squeezed configurations become more important on small angular scales, deep in the non-linear regime of structure formation. For the case of ζ_-, the squeezed configurations become dominant for $\alpha \lesssim 20$ arcmin, and the figure shows that if evaluated with the response approach, then they are able to describe the simulation measurements very well. At a fixed angular scale α, the ζ_+ statistic is less sensitive to
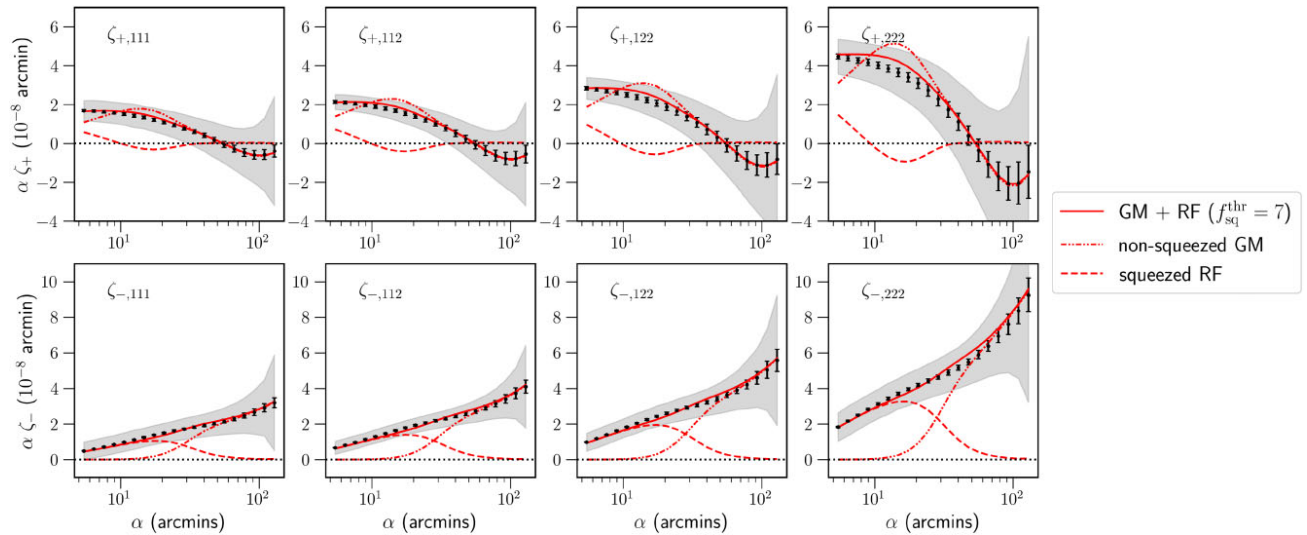
Figure 2. The integrated shear 3PCFs $\zeta_\pm(\alpha)$ for two tomographic source redshift bins $z_1 \approx 0.57$ and $z_2 \approx 1.03$. The black dots with the error bars show the mean and the standard deviation of the measurements from the 108 T17 all-sky simulation maps, respectively. The grey shaded regions indicate the standard deviation expected for these statistics in DES-sized footprints (obtained from the diagonal of our covariance matrix). The solid red lines show the theoretical predictions of equations (14) using the joint GM + RF bispectrum model (cf. equation 25) for $f_{sq}^{thr} = 7$. The dot-dashed line shows the contribution to the total result from only non-squeezed configurations [i.e. setting the RF branch in equation (25) to zero], and the dashed line shows the contribution from only squeezed configurations [i.e. setting the GM branch in equation (25) to zero]. The result includes the numerical resolution corrections described in Takahashi et al. (2017).

higher-l modes compared to ζ_- in equation (14) (this is because of the different shapes of the Bessel functions J_0 and J_4), and as a result, the contribution from squeezed configurations is not as significant.

Fig. 3 illustrates the impact of different choices for the threshold parameter f_{sq}^{thr} on the $\zeta_\pm$ predictions. Let us discuss first the result for ζ_- in the lower panels. On angular scales $\alpha \lesssim 15$ arcmin, the result is dominated by squeezed configurations and the figure shows that it is independent of the choice of the threshold parameter in the range $f_{sq}^{thr} \in [3, 11]$. This indicates that, on these scales, ζ_- is determined by very squeezed triangles with at least $f_{sq} \geq 11$, which the response approach can evaluate very accurately. On the other hand, the fact that for $\alpha \gtrsim 80$ arcmin, the result is also independent of $f_{sq}^{thr} \in [3, 11]$ indicates that ζ_- is determined by triangle configurations that are closer to equilateral with at least $f_{sq} \leq 3$. These large-scale configurations are in turn well captured by the GM fitting formula. On scales in between these two limits, the result is seen to depend on f_{sq}^{thr}, i.e. it is sensitive to the fraction of triangles with $f_{sq} \in [3, 11]$ that are evaluated with the response approach or with the GM formula. Concretely, lowering f_{sq}^{thr} gives more emphasis to the RF branch in equation (25), but since the error of the RF result scales as $1/f_{sq}^2$, the calculation is also less accurate. Conversely, increasing f_{sq}^{thr} gives more emphasis to the GM branch, but since these are scales where non-linear contributions are already important, the GM formula becomes also less accurate. It is the competition between the accuracy of the RF and GM branches in these transition regimes that determines the optimal choice of f_{sq}^{thr}. The discussion for the ζ_+ results shown in the upper panels of Fig. 3 follows along similar lines, with the main difference being the lack of a regime on small angular scales where the result is independent of the threshold parameter in the range $f_{sq}^{thr} \in [3, 11]$. This is again associated with the shapes of the Bessel functions J_0 and J_4, which make ζ_+ less sensitive to high-l values, and consequently, to the contribution from very squeezed triangles.

It is important to note that the impact of f_{sq}^{thr} shown in Fig. 3 is peculiar to our choice of using the GM formula to evaluate the non-squeezed branch of equation (25). If instead of GM we would

have used the `bihalofit` formula, then there would be a certain value of f_{sq}^{thr} above which the goodness-of-fit would always be reasonable since `bihalofit` is relatively accurate for all triangle configurations and in the non-linear regime. In other words, we would expect to find a value of f_{sq}^{thr} below which the goodness-of-fit becomes bad as the error of the RF expression becomes large. In situations like these, in which the RF branch is used together with a calculation that is accurate on all scales, then the criteria to determine f_{sq}^{thr} should be that (i) it is just large enough to ensure the RF result is used accurately for sufficiently squeezed triangles, but (ii) not too large to still let the RF branch provide the dominant contribution on scales where baryonic effects are important. We have explicitly checked that $f_{sq}^{thr} = 7$ satisfies also these criteria by replacing the GM formula with `bihalofit` in equation (25) with $f_{sq}^{thr} = 7$, and noting that the goodness-of-fit is effectively the same, and that the RF branch dominates the contribution on scales where baryonic effects are important (as determined using the strategy described in the next section).

Fig. 4 compares the outcome of equation (14) to predict $\zeta_\pm$ using four different methods to evaluate the three-dimensional matter bispectrum in equation (15): tree-level SPT (orange), the GM formula (blue), the `bihalofit` formula (Takahashi et al. 2020) (green), and our joint GM + RF bispectrum model with $f_{sq}^{thr} = 7$ (red). The tree-level bispectrum calculation gives only a poor fit to the simulation results, which is as expected since it is only a decent approximation on very large scales. As found previously in Halder et al. (2021), the GM result provides a good description of ζ_+ on all angular scales shown, as well as of ζ_- for $\alpha \gtrsim 30$ arcmin. As discussed above in Fig. 2, on smaller scales for ζ_-, the result begins to be dominated by squeezed configurations in the non-linear regime, whose contribution the GM fitting function manifestly overestimates (cf. Fig. 13 of Takahashi et al. 2020). Instead, with its ability to accurately describe the squeezed matter bispectrum in the non-linear regime, the response approach is able to fix these shortcomings of the GM fitting function, as seen by the excellent agreement between the red solid line and the simulation data points for the small-scale
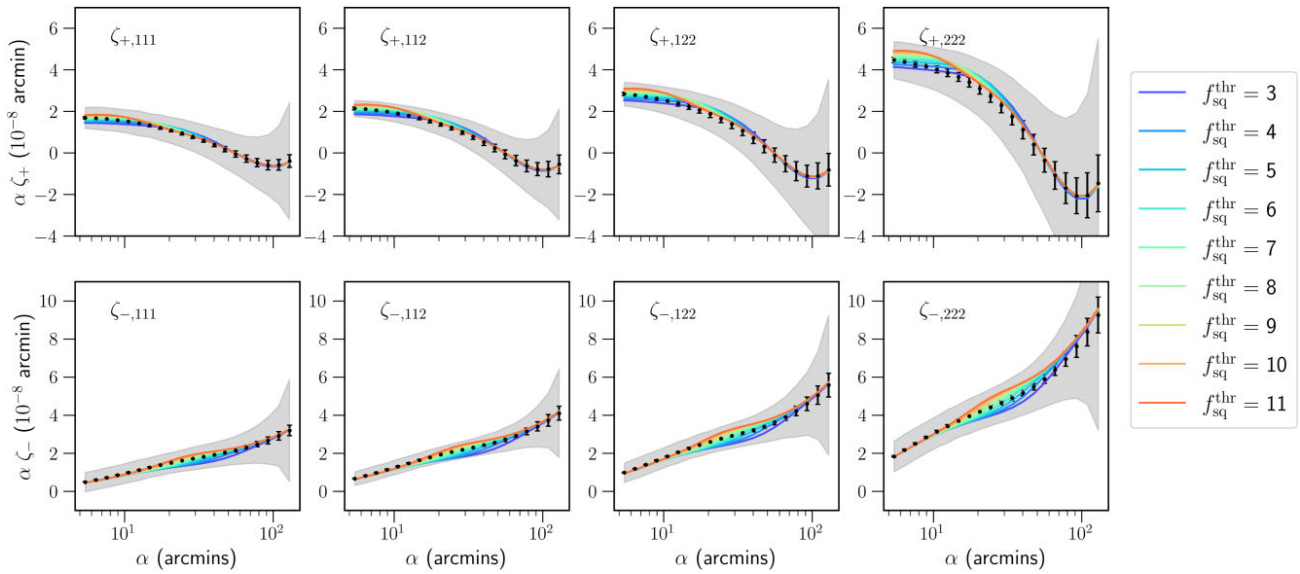
Figure 3. Same as Fig. 2, but with the theoretical predictions shown for different values of the threshold parameter f_{sq}^{thr}, as labelled.
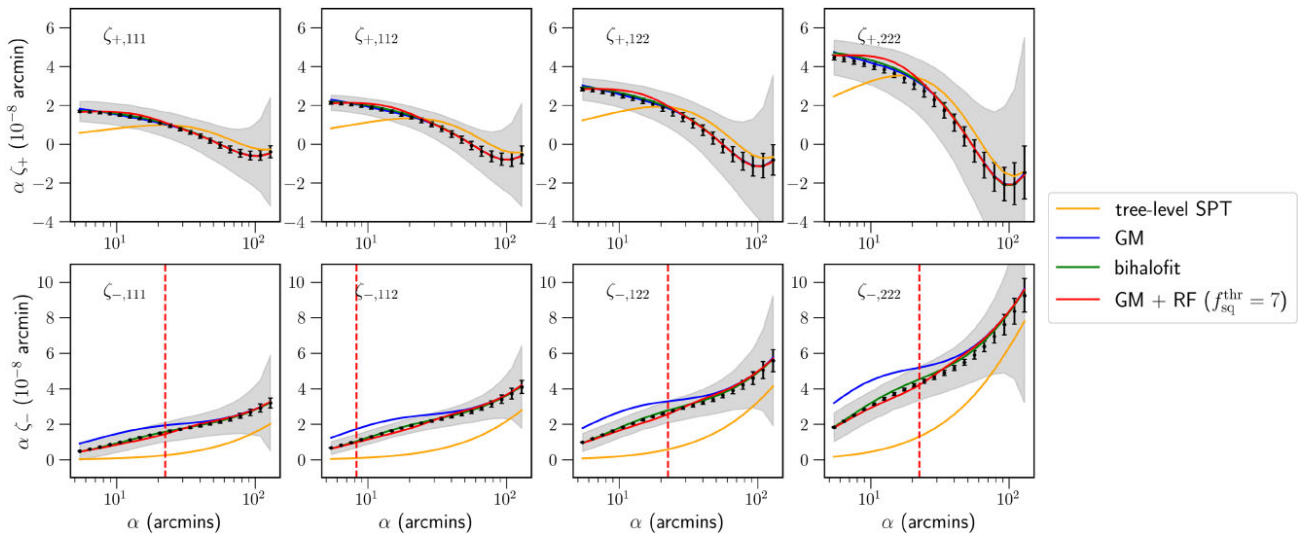
Figure 4. Same as Fig. 2, but for different methods to calculate the three-dimensional matter bispectrum: tree-level (orange), the GM formula (blue), the bihalofit formula (green), and our GM + RF model with $f_{sq}^{thr} = 7$ (red). The red-dashed vertical lines mark the angular scale cuts that we apply in our Fisher matrix analysis to remove the parts of the data vector (on the left of the line) that are affected by baryonic feedback (see Section 4.2.1 for details). Note that the ζ_+ statistics do not have any imposed scale cuts on the scales shown.

ζ_-. We have explicitly checked that in the limit of very large f_{sq}^{thr}, the GM + RF result eventually becomes indistinguishable from the GM-only result, as expected.

For the ζ_- statistic, Fig. 4 shows that the GM + RF (red) and `bihalofit` (green) approaches display effectively the same goodness-of-fit to the simulation results, but for ζ_+ the GM + RF approach is seen to slightly overestimate the result at $\alpha \sim 10$–15 arcmin (this is best seen in the $\zeta_{+,122}$ and $\zeta_{+,222}$ panels). This has to do with our choice of $f_{sq}^{thr} = 7$, which we determined by inspecting the global χ^2 goodness-of-fit using both ζ_- and ζ_+. An improved strategy would have been to choose different values of f_{sq}^{thr} for ζ_+ and ζ_- (or even for different tomographic bins), which is in fact the most reasonable thing to do given that these two statistics get manifestly different contributions from squeezed and non-squeezed configurations, as shown in Fig. 2. Here, we proceed with our global

choice of $f_{sq}^{thr} = 7$ for simplicity, but also because the impact of this choice is still well within the expected DES uncertainty.

4.2 Fisher forecasts and baryonic effects

We now investigate the Fisher information content of the combination of the $\xi_\pm$ and $\zeta_\pm$ correlation functions on both the cosmological and baryonic feedback parameters. We begin with the description of our forecast setup in Section 4.2.1, and discuss the results in Section 4.2.2.

4.2.1 Fisher forecast setup

The Fisher information matrix **F** for a model vector *M* that depends on a set of parameters *π* is given by (Tegmark, Taylor & Heavens
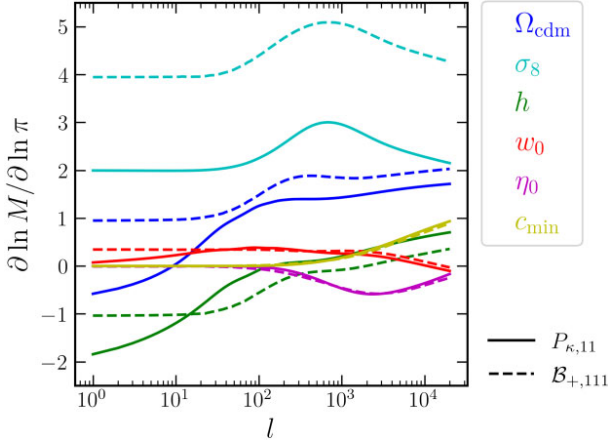
Figure 5. Logarithmic derivatives of the convergence power spectrum (solid) and integrated shear bispectrum (dashed) with respect to the six parameters $\boldsymbol{\pi} = \{\Omega_{\mathrm{cdm}}, \sigma_8, h, w_0, \eta_0, c_{\min}\}$ for source redshift $z_1 \approx 0.57$.

1997)

$$F_{ij} = \left(\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_i}\right)^{\mathrm{T}} \mathbf{C}^{-1} \left(\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_j}\right), \qquad (29)$$

where F_{ij} is the element of $\mathbf{F}$ associated with the parameters π_i and π_j, and $\mathbf{C}^{-1}$ is the inverse data covariance matrix. The partial derivative of the model vector with respect to the parameter π_i can be computed using a 2-point central difference:

$$\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_i} = \frac{M(\pi_i + \delta_i) - M(\pi_i - \delta_i)}{2\delta_i}, \qquad (30)$$

where δ_i is a small change of the parameter π_i around its fiducial value, and $M(\pi_i \pm \delta_i)$ is the model vector evaluated at the changed parameter $\pi_i \pm \delta_i$ with all other parameters fixed to their fiducial values. We consider four cosmological and two baryonic feedback parameters $\boldsymbol{\pi} = \{\Omega_{\mathrm{cdm}}, \sigma_8, h, w_0, \eta_0, c_{\min}\}$, where w_0 is the dark energy equation of state parameter (assumed time-independent) and recall η_0, $c_{\min}$ are the two baryonic feedback parameters of the HMCODE. The fiducial values are $\boldsymbol{\pi}_0 = \{0.233, 0.82, 0.7, -1, 0.603, 3.13\}$. For the cosmological parameters, this is the same as in the T17 simulations, and for the two baryonic parameters we consider the default gravity-only values as determined by Mead et al. (2015) by fitting against the COSMIC EMU power spectrum of Heitmann et al. (2014). When we differentiate w.r.t. Ω_{cdm} we keep the baryon density Ω_b fixed, but adjust the dark energy density to keep the universe spatially flat. The parameter covariance matrix $\mathbf{C}_{\boldsymbol{\pi}}$ is given by the inverse of the Fisher matrix

$$\mathbf{C}_{\boldsymbol{\pi}} = \mathbf{F}^{-1}, \qquad (31)$$

and is what can be used to forecast constraints on the parameters $\boldsymbol{\pi}$.

Before discussing the forecast results, it is interesting to inspect first the derivatives $\partial M/\partial \boldsymbol{\pi}$ of the model vector. Instead of looking at the derivatives in real space, e.g. $\partial \xi_\pm(\alpha)/\partial \pi_i$ and $\partial \zeta_\pm(\alpha)/\partial \pi_i$, we find it more intuitive to show the derivatives in Fourier space, i.e. $\partial P_\kappa(l)/\partial \pi_i$ and $\partial \mathcal{B}_\pm(l)/\partial \pi_i$; note the result in real space at fixed α is a mixture of contributions from several l modes. Fig. 5 shows the logarithmic derivatives of $P_{\kappa,\,11}$ (solid) and $\mathcal{B}_{+,111}$ (dashed) at source redshift $z_1 \approx 0.57$ with respect to our six parameters. We do not show all tomographic combinations of $P_{\kappa,\mathrm{fg}}$ and $\mathcal{B}_{\pm,\mathrm{fgh}}$ for brevity and because they share the same following takeaway points:

(i) The derivatives w.r.t. the baryonic parameters η_0, $c_{\min}$ become sizeable only for $l \gtrsim 200$, and their shape and size are similar for both

P_κ and $\mathcal{B}_+$ (cf. solid and dashed magenta and yellow curves). This is expected since at high l these two statistics have similar dependencies on the *non-linear* three-dimensional matter power spectrum. Concretely, from equation (12), we can write $P_\kappa \sim \int \mathrm{d}\chi (q^2/\chi^2) P_\delta^{3D}$, and from equation (15), we can write the high-l integrated shear bispectrum as[10] $\mathcal{B}_+ \sim \int \mathrm{d}\chi (\cdots)(q^3/\chi^4) P_\delta^{3D}$. That is, P_κ and $\mathcal{B}_+$ depend differently on the baryonic effects only in that the different dependence of the integrands on q and χ weights differently the time evolution of the impact of baryons on the matter power spectrum P_δ^{3D}. This time evolution is not very strong in the HMCODE of Mead et al. 2015 since η_0, $c_{\min}$ are constant in time, but we note that other parametrizations of baryonic effects with more complicated time evolutions may lead to more distinctive impacts on P_κ and $\mathcal{B}_+$.

(ii) For the cosmological parameters Ω_{cdm}, σ_8, h, and w_0, the size and scale-dependence of the derivatives of P_κ are now visibly more different than those of $\mathcal{B}_+$. For σ_8, this is due to the impact on the amplitude of the matter power spectrum, and as expected, $\mathcal{B}_+$ is more sensitive as it depends on two powers of the matter power spectrum. For Ω_{cdm}, h, and w_0, this is instead due to a combination of the direct impact that these parameters have on both the lensing kernel factors q^2/χ^2 and q^3/χ^4, and the matter power spectrum.

In the results that we discuss next, we consider cases in which all angular scales are used, as well as cases in which we discard the parts of the data vectors that are expected to be significantly affected by baryonic effects. We determine these scale cuts as follows. We compute the theoretical predictions for two scenarios using the HMCODE: (i) one which was fitted to the matter power spectrum from the AGN run of the OWLS simulations, $\eta_0 = 0.76$ and $c_{\min} = 2.32$; and (ii) another which was fitted to the gravity-only counterpart of the same simulation, $\eta_0 = 0.64$ and $c_{\min} = 3.43$ (see Mead et al. 2015 for more details). We then evaluate the χ^2 quantity

$$\chi^2(\alpha_{\min}) = \sum_{\alpha_i, \alpha_j > \alpha_{\min}} \left(M_{\mathrm{AGN}}(\alpha_i) - M_{\mathrm{Grav}}(\alpha_i)\right)(\mathbf{C}^{-1})_{\alpha_i \alpha_j}$$
$$\times \left(M_{\mathrm{AGN}}(\alpha_j) - M_{\mathrm{Grav}}(\alpha_j)\right), \qquad (32)$$

where M_{AGN} and M_{Grav} denote the two model vectors, $\mathbf{C}$ is our covariance matrix, and $\alpha_{\min}$ is a minimum angular scale cut. Our *baryon scale cuts* are then given by the smallest value of $\alpha_{\min}$ for which χ^2 is less than 0.3. The angular scale cuts determined in this way for $\xi_\pm$ and $\zeta_\pm$ are marked by the vertical red dashed lines in Figs 1 and 4, respectively; that is, in our discussion below, we deem all scales α larger than the scale cut to be relatively unaffected by the impact of baryonic effects. Note how these scale cuts are less aggressive (i.e. more scales are kept) on $\zeta_\pm$ compared to $\xi_\pm$; they are also less stringent on ξ_+, ζ_+ than on ξ_-, ζ_-.

As a check of this χ^2-strategy, we further estimate the expected bias on the cosmological parameters from ignoring baryonic effects on the scales that we consider baryon-free. In order to do this, we adopt the Gaussian linear model (see Seehars et al. 2016; Dodelson, Shapiro & White 2006) and make a linearized approximation of the model vector around the 4 fiducial cosmological parameters $\boldsymbol{\pi}_0^{\mathrm{cosmo}}$. The bias, quantified as the difference between the best-fitting $\boldsymbol{\pi}_{\mathrm{BF}}^{\mathrm{cosmo}}$

---

[10]We stress these are schematic equations to facilitate making the point. Note also that although the squeezed matter bispectrum in the response approach depends on two powers of the matter power spectrum in equation (24), one of them is effectively always in the linear regime and is not affected by baryonic effects. That is why we only write one power of P_δ^{3D} to roughly explain the origin of the baryonic impact (we also ignore the contribution from response functions in equation (24)).
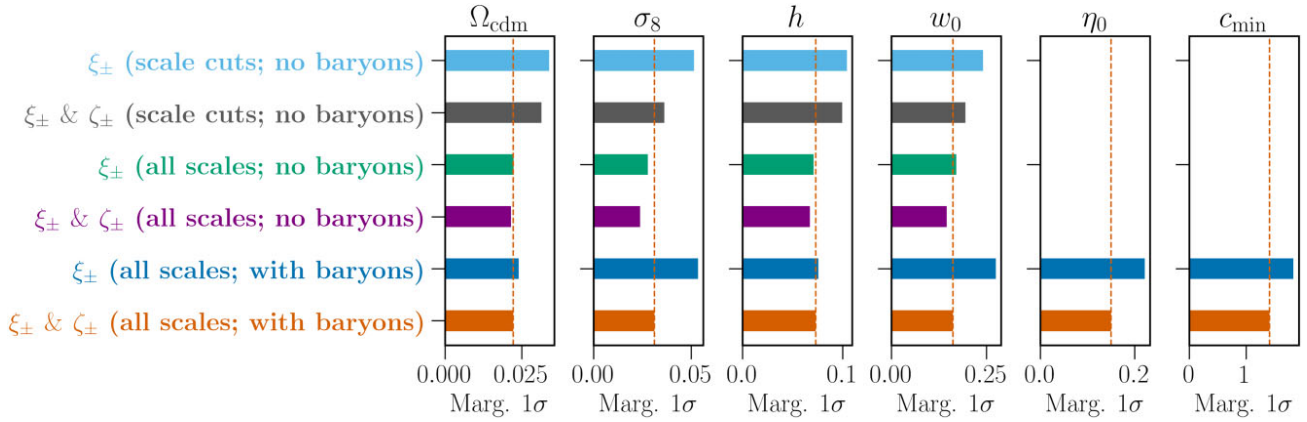
Figure 6. Marginalized 1σ constraints for the parameters $\boldsymbol{\pi} = \{\Omega_{cdm}, \sigma_8, h, w_0, \eta_0, c_{min}\}$. The columns are for the different parameters and the rows indicate the constraints for different combinations of $\xi_\pm$ and $\zeta_\pm$ for the three different ways to deal with baryonic effects (see the text in Section 4.2.2), as labelled. These are the marginalized 1σ constraints for the same cases shown in Figs 7 and 8.

and fiducial $\boldsymbol{\pi}_0^{cosmo}$ cosmological parameters, is then estimated as

$$\boldsymbol{\pi}_{BF}^{cosmo} - \boldsymbol{\pi}_0^{cosmo} = \mathbf{F}^{-1}\boldsymbol{x}, \tag{33}$$

where $\boldsymbol{x}$ is a vector with components

$$x_i = (M_{AGN} - M_{Grav})^T \mathbf{C}^{-1} \left(\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_i} \right), \tag{34}$$

and all the relevant quantities are evaluated using only the baryon-free scales. In units of the marginalized 1σ uncertainties (grey bars in Fig. 6), the biases for the 4 cosmological parameters Ω_{cdm}, σ_8, h, and w_0 are 0.17, 0.69, 0.22, and 0.51, respectively. These are contained well within the corresponding 1σ error bars, and so we deem our scale cut strategy to be sufficiently adequate to our purpose here to roughly estimate the scales where baryonic effects begin to play a role, and investigate the gains from extending the analyses on to smaller scales. We note here that in all our results, we ignore the impact of baryonic effects on the covariance matrix. This has been shown to be negligible by Barreira et al. (2019) and Schneider et al. (2020) for 2-point correlation function analyses; to the best of our knowledge, the same has never been checked for analyses using 3-point correlation function information.

4.2.2 Fisher forecast results

We show and discuss results for the following three different ways to deal with the baryonic feedback effects:

(i) Case A: *scale cuts; no baryons*. In this case, we consider only the parts of the data vector that have been deemed as baryon-free (cf. vertical lines in Figs 1 and 4). We constrain the parameters Ω_{cdm}, σ_8, h, and w_0, and keep the baryonic parameters η_0, c_{min} fixed to their fiducial gravity-only values.

(ii) Case B: *all scales; no baryons*. Here we consider all angular scales, including those affected by baryonic feedback, but continue to keep the baryonic feedback parameters fixed to their fiducial values.

(iii) Case C: *all scales; with baryons*. This is the same as case B, but varying also the baryonic feedback parameters, i.e. we constrain the six parameters Ω_{cdm}, σ_8, h, w_0, η_0, and c_{min}.

Fig. 6 shows the one-dimensional marginalized 1σ constraints obtained with $\xi_\pm$ alone and with the combination $\xi_\pm$ & $\zeta_\pm$, for each of the three cases above, as labelled. The left-hand and right-hand panels of Fig. 7 show the corresponding corner plots with two-

dimensional marginalized 1σ and 2σ constraints for cases A and B, respectively; Fig. 8 shows the same for case C. The main overall takeaway message is that, for all of the cases shown, the combination $\xi_\pm$ & $\zeta_\pm$ always leads to improved constraints compared to $\xi_\pm$ alone. This is as expected from the different dependence of the $\xi_\pm$ and $\zeta_\pm$ statistics on the different parameters (cf. Fig. 5), which works to break degeneracies and leads to tighter constraints. Specifically, relative to $\xi_\pm$ alone, the constraints obtained with the combination $\xi_\pm$ & $\zeta_\pm$ on the four cosmological parameters $\{\Omega_{cdm}, \sigma_8, h, \text{and } w_0\}$ are tighter by $\{7, 30, 4, 19\}$ per cent for case A, $\{4, 14, 5, 15\}$ per cent for case B, and $\{8, 42, 3, 41\}$ per cent for case C, respectively.

As one would expect, the inclusion of information from all scales in case B compared to the scale cuts imposed in case A results in tighter constraints on all four cosmological parameters. Concretely, for the combination $\xi_\pm$ & $\zeta_\pm$, the constraints on $\{\Omega_{cdm}, \sigma_8, h, \text{and } w_0\}$ in case B are better than in case A by $\{32, 35, 32, 25\}$ per cent, respectively (cf. 2nd versus 4th lines in Fig. 6). We also note that the impact of $\zeta_\pm$ is less pronounced in case B compared to case A because of the less aggressive scale cuts that are imposed on $\zeta_\pm$ compared to $\xi_\pm$ (cf. red vertical lines in Figs 1 and 4), i.e. relative to the size of the corresponding covariances, baryonic feedback effects impact $\xi_\pm$ more strongly than $\zeta_\pm$. In other words, when all scales are included in the analysis from case A to case B, the constraining power coming from $\xi_\pm$ increases more compared to $\zeta_\pm$. We stress however that case B is a highly idealized scenario that assumes perfect knowledge of how baryonic physics impact the small-scale cosmic shear signal, but which is helpful to analyse anyway to help appreciate the additional amount of information encoded on those small scales.

A more realistic approach to the analysis of cosmic shear data on small scales is therefore that of case C, which accounts also for uncertainties on the baryonic physics parameters. In this case, Fig. 6 shows that, for the combination $\xi_\pm$ & $\zeta_\pm$, the inclusion of small-scale information results also in improved cosmological constraints, despite the added uncertainty coming from marginalizing over the two baryonic feedback parameters. Specifically, the constraints on $\{\Omega_{cdm}, \sigma_8, h, \text{and } w_0\}$ in case C are better than those on case A by $\{29, 14, 26, 17\}$ per cent, respectively (cf. 2nd versus 6th lines in Fig. 6).

Interestingly, the inclusion of information from $\zeta_\pm$ in case C results also in visibly tighter constraints on the two baryonic feedback parameters: the constraints on η_0 and c_{min} obtained with the combination $\xi_\pm$ & $\zeta_\pm$ improve by 32 per cent and 23 per cent,
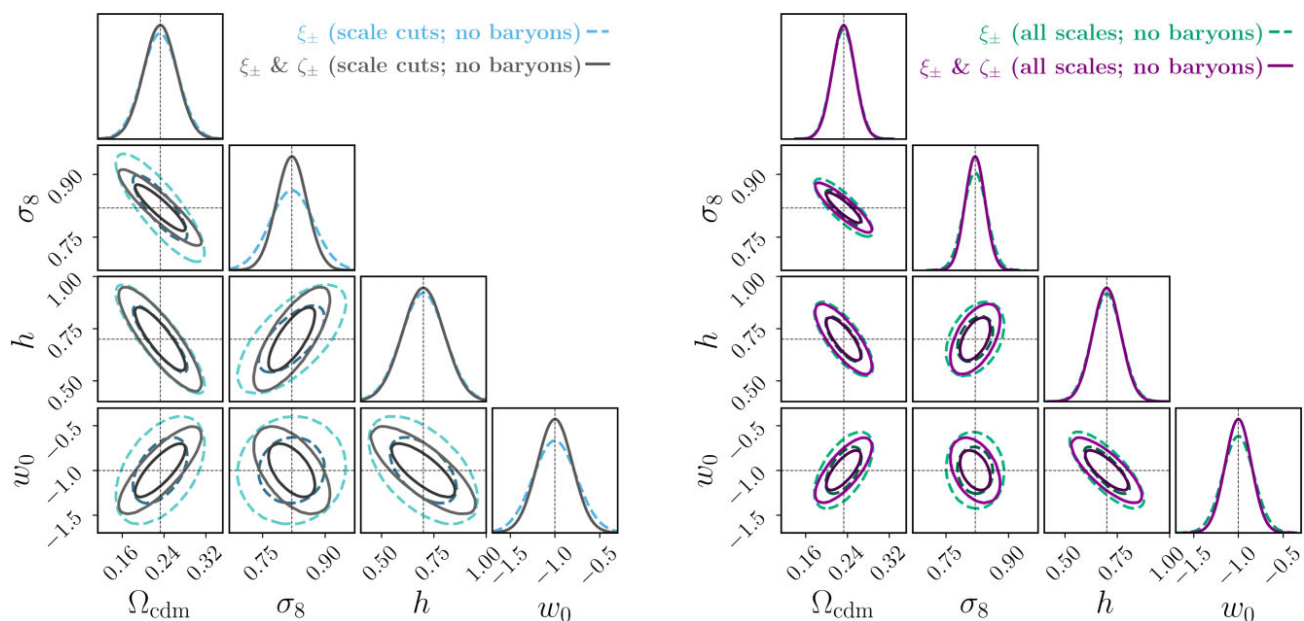
Figure 7. Marginalized 1- and 2-dimensional constraints for the cosmological parameters $\boldsymbol{\pi} = \{\Omega_{cdm}, \sigma_8, h, w_0\}$. The results are shown in dashed and solid for the $\xi_\pm$ and combined $\xi_\pm$ & $\zeta_\pm$ data vectors, respectively; the black dotted lines mark the fiducial values and the two sets of contours mark 1σ and 2σ confidence limits. The left-hand panel is for case A (*scale cuts; no baryons*) and the right-hand panel is for case B (*all scales; no baryons*). Recall from Section 4.2.2 that for these two cases the baryonic feedback parameters η_0, c_{min} are kept fixed to their fiducial values.
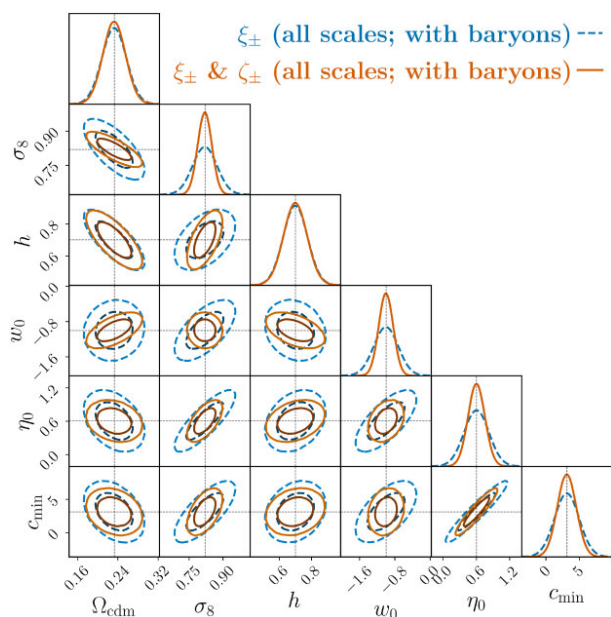
Figure 8. Same as the right part of Fig. 7, but for case C (*all scales; with baryons*), i.e. also varying the baryonic feedback parameters η_0 and c_{min}.

respectively, compared to $\xi_\pm$ alone. This may appear as a surprising result since, as we discussed in Fig. 5, the $\xi_\pm$ and $\zeta_\pm$ statistics respond very similarly to changes to the parameters η_0 and c_{min}. The reason behind these improved constraints can be explained by the correlations that these parameters display with the rest of the cosmological parameters. For example, Fig. 8 shows that η_0 and c_{min} are correlated with the cosmological parameters σ_8 and w_0, and thus, a tightening of the constraints on the latter by $\zeta_\pm$ information also results in a tightening of the constraints on the former. We stress however that these considerations should be interpreted in

light of the parametrization of baryonic feedback effects that is implemented in the HMCODE of Mead et al. (2015), and that more involved parametrizations (e.g. including time-dependent baryonic feedback parameters) may result in different constraints. We defer these investigations to future work.

Finally, Fig. 9 shows the marginalized 1σ constraints in the same format as in Fig. 6, but with the dark energy equation of state held fixed at its fiducial value $w_0 = -1$. In this case as well, the parameter constraints improve when $\zeta_\pm$ is added to the analysis: relative to $\xi_\pm$ alone, the constraints obtained with the combination $\xi_\pm$ & $\zeta_\pm$ on the three cosmological parameters $\{\Omega_{cdm}, \sigma_8, \text{ and } h\}$ are tighter by $\{38, 41, 29\}$ per cent for case A, $\{20, 21, 8\}$ per cent for case B, and $\{26, 25, 14\}$ per cent for case C, respectively. Two noteworthy differences from these constraints, compared to those in which w_0 is also varied are: (i) the improvements brought by $\zeta_\pm$ on the Ω_{cdm} and h constraints become more pronounced; but (ii) the same improvements on the baryonic feedback parameter constraints become less pronounced. This last point indicates that the improvement on the η_0 and c_{min} constraints observed in Fig. 6 largely followed from the breaking of degeneracies with w_0 by $\zeta_\pm$.

5 SUMMARY AND CONCLUSIONS

The *integrated shear 3-point correlation function* (3PCF) $\zeta_\pm$ introduced in Halder et al. (2021) is a higher-order weak lensing statistic that can be directly measured from cosmic shear data by correlating *1-point mass statistics* with *2-point correlation functions* measured within well-defined patches (or apertures) distributed across a survey footprint (cf. equation 1). The main theoretical ingredient to make predictions for this statistic is the three-dimensional non-linear matter bispectrum (cf. equations 14 and 15). In their previous work, Halder et al. (2021) made predictions for $\zeta_\pm$ using the GM (Gil-Marín et al. 2012) and bihalofit (Takahashi et al. 2020) fitting formulae to evaluate the matter bispectrum, but both these approaches have drawbacks when applied to predictions for $\zeta_\pm$ in the non-linear
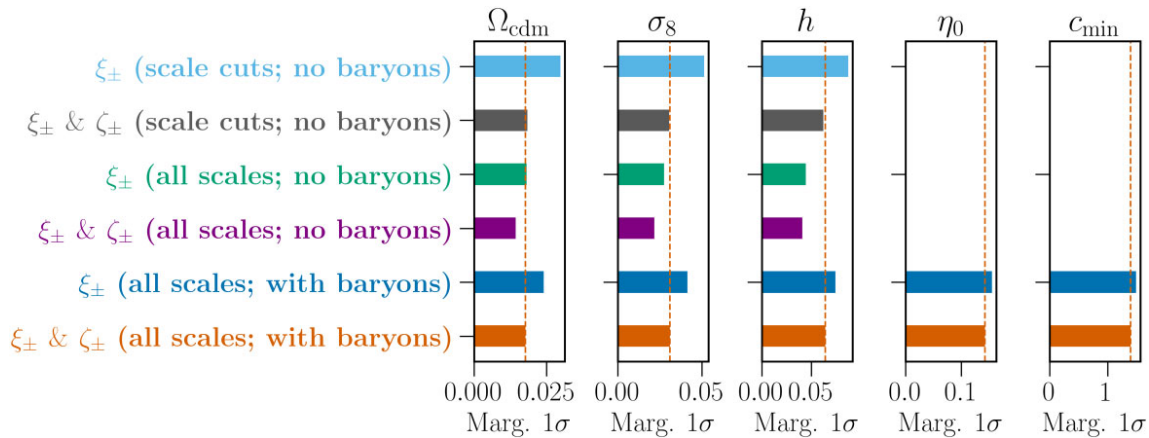
Figure 9. Same as Fig. 6 but dropping the dark energy equation of state parameter from the constraints by fixing it to its fiducial value $w_0 = -1$.

regime on small angular scales: the GM formula becomes inaccurate as it was not developed to describe the matter bispectrum in the non-linear regime, and while `bihalofit` works well on small scales, it cannot currently make predictions as a function of different baryonic feedback parameters.

In this paper, we used the response approach to perturbation theory developed by Barreira & Schmidt (2017a) to develop a new and improved method to calculate the non-linear matter bispectrum to predict the small-scale $\zeta_{\pm}$ statistic. The key observation behind our calculation is that, on small angular scales, the $\zeta_{\pm}$ statistic is dominated by squeezed bispectrum configurations, which can be evaluated accurately in the non-linear regime using the response approach (cf. discussion in Section 2.2.2). The result is given in terms of the non-linear matter power spectrum and its first-order response functions to mass overdensities and tidal fields (cf. equation 24). Importantly, however, the impact of baryonic effects enters the calculation of $\zeta_{\pm}$ effectively only through their impact on the non-linear matter power spectrum, for which several strategies already exist in the literature; here, we adopted the formalism of the HMCODE of Mead et al. (2015) to illustrate our calculation.

Our main objective was two fold: (i) illustrate and validate the application of the response approach to the small-scale $\zeta_{\pm}$, and (ii) study the improvements on cosmological constraints from combining measurements of $\zeta_{\pm}$ with the shear 2-point correlation function (2PCF) $\xi_{\pm}$, while taking baryonic effects into account. Our main results can be summarized as follows:

(i) A model for the matter bispectrum that uses the GM formula[11] for non-squeezed, and the response approach for squeezed bispec-trum configurations (cf. Section 2.2.2) is able to describe very well the $\zeta_{\pm}$ statistic measured from simulated cosmic shear maps down to scales $\alpha = 5$ arcmin (cf. Fig. 4). This joint model is characterized by a parameter f_{sq}^{thr} that defines whether a bispectrum configuration is dubbed as squeezed (cf. equation 25), and which can be calibrated using gravity-only simulations.

(ii) Using a simple Fisher matrix forecast analysis for a tomo-graphic DES-like survey, we found that the data combination $\xi_{\pm}$ & $\zeta_{\pm}$ can lead to substantial improvements in cosmological constraints, compared to standard analyses with $\xi_{\pm}$ alone. Concretely, our nu-merical results showed that $\zeta_{\pm}$ information can tighten the constraints of parameters like σ_8 or w_0 by about 20–40 per cent. We note that

these exact figures can depend on many analysis details (cf. Fig. 6 versus Fig. 9), but they do provide motivation to include $\zeta_{\pm}$ in future real-data analyses.

(iii) We also found that the inclusion of small angular scales in combined $\xi_{\pm}$ & $\zeta_{\pm}$ analyses, even after marginalizing over baryonic uncertainties, generically leads to improved constraints on the cosmological parameters (e.g. 15–20 per cent for w_0), compared to analyses that apply scale cuts to remove the parts of the data vector that are expected to be affected by baryonic effects. When considering these small scales, our numerical results also showed that the addition of $\zeta_{\pm}$ data could lead to improvements of order 20–30 per cent on the constraints of the two HMCODE feedback pa-rameters η_0 and c_{min} (cf. Fig. 6). This illustrates that the ability to incorporate baryonic physics effects on $\zeta_{\pm}$ is important not only to constrain cosmology, but it can also help constrain astrophysics models of baryonic feedback in hydrodynamical simulations.

The steps we took in this paper help bring the integrated shear 3PCF one step closer to applications to real data. This is not only because of the ability of our theoretical model to make accurate predictions for $\zeta_{\pm}$ on small scales in itself, but also because it equips us with a strategy to identify the scales that are expected to be most affected by baryonic physics effects, should these still be conservatively chosen to be removed from real data analyses. We should note therefore that many of our considerations in this paper are actually not peculiar to the response approach, and hold generically to any calculation of the matter bispectrum that is able to take baryonic effects into account. The response approach is indeed an elegant and efficient way to do so, but future approaches based on direct emulation of the matter bispectrum with baryonic effects and/or generalized versions of the `bihalofit` formula will be interesting to consider as well. Further, as discussed in Halder et al. (2021), the mathematical formalism behind the integrated 3PCF can be straightforwardly generalized beyond cosmic shear data to include also correlations with the foreground galaxy distribution. These and other developments will be explored in future works.

In conclusion, our discussion underlines the importance of cosmic shear studies beyond the 2PCF, and the benefits from developing theoretical models to describe the signal on small scales where baryonic effects are important. In particular, future works on $\zeta_{\pm}$ are especially well motivated since these statistics can be readily measured from existing and forthcoming cosmic shear data using well-tested techniques, and this may well result in interesting new

---

[11]Or any other formula accurate in the quasi-linear regime.

constraints on not only cosmology, but also the complex astrophysics of baryonic feedback.

DATA AVAILABILITY

The data for the *N*-body simulations used in this article were accessed from the public domain: http://cosmo.phys.hirosaki-u.ac.jp/takahasi /allsky_raytracing/. The other data underlying this article will be shared on reasonable request to the authors.

REFERENCES

Abbott T. M. C. et al., 2021, Phys. Rev. D, 105, 023520
Aricò G., Angulo R. E., Hernández-Monteagudo C., Contreras S., Zennaro M., Pellejero-Ibañez M., Rosas-Guevara Y., 2020, MNRAS, 495, 4800
Aricò G., Angulo R. E., Hernández-Monteagudo C., Contreras S., Zennaro M., 2021, MNRAS, 503, 3596
Aricò G., Angulo R. E., Contreras S., Ondaro-Mallea L., Pellejero-Ibañez M., Zennaro M., 2021, MNRAS, 506, 4070
Baldauf T., Garny M., Taule P., Steele T., 2021, Phys. Rev. D, 104, 123551
Barreira A., 2019, J. Cosmology Astropart. Phys., 2019, 008
Barreira A., Schmidt F., 2017a, J. Cosmology Astropart. Phys., 2017, 053
Barreira A., Schmidt F., 2017b, J. Cosmology Astropart. Phys., 2017, 051
Barreira A., Krause E., Schmidt F., 2018, J. Cosmology Astropart. Phys., 2018, 015
Barreira A., Nelson D., Pillepich A., Springel V., Schmidt F., Pakmor R., Hernquist L., Vogelsberger M., 2019, MNRAS, 488, 2079
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Blas D., Lesgourgues J., Tram T., 2011, J. Cosmology Astropart. Phys., 2011, 034
Burger P., Schneider P., Demchenko V., Harnois-Deraps J., Heymans C., Hildebrandt H., Unruh S., 2020, A&A, 642, A161
Burger P., Friedrich O., Harnois-Déraps J., Schneider P., 2022, A&A, 661, A137
Chiang C.-T., Wagner C., Schmidt F., Komatsu E., 2014, J. Cosmology Astropart. Phys., 2014, 048
Chiang C.-T., Wagner C., Sánchez A. G., Schmidt F., Komatsu E., 2015, J. Cosmology Astropart. Phys., 09, 028
Chisari N. E. et al., 2019, Open J. Astrophys., 2, 4
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Crittenden R. G., Natarajan P., Pen U.-L., Theuns T., 2002, ApJ, 568, 20
Desjacques V., Jeong D., Schmidt F., 2018, Phys. Rep., 733, 1
Dodelson S., Shapiro C., White M., 2006, Phys. Rev. D, 73
Eifler T., Krause E., Dodelson S., Zentner A. R., Hearin A. P., Gnedin N. Y., 2015, MNRAS, 454, 2451

Foreman S., Coulton W., Villaescusa-Navarro F., Barreira A., 2020, MNRAS, 498, 2887
Friedrich O. et al., 2018, Phys. Rev. D, 98
Friedrich O. et al., 2021, MNRAS, 508, 3125
Fu L. et al., 2014, MNRAS, 441, 2725
Gatti M. et al., 2021, Dark Energy Survey Year 3 Results: Cosmology with Moments of Weak Lensing Mass Maps. preprint (arXiv:2110.10141)
Gil-Marín H., Wagner C., Fragkoudi F., Jimenez R., Verde L., 2012, J. Cosmology Astropart. Phys., 2012, 047
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Gough B., 2009, GNU Scientific Library Reference Manual. Network Theory Ltd.
Gruen D. et al., 2018, Phys. Rev. D, 98, 023507
Halder A., Friedrich O., Seitz S., Varga T. N., 2021, MNRAS, 506, 2780
Hamana T., Sakurai J., Koike M., Miller L., 2015, Publ. Astron. Soc. Japan, 67, 34
Harnois-Déraps J., van Waerbeke L., Viola M., Heymans C., 2015, MNRAS, 450, 1212
Harnois-Déraps J., Martinet N., Castro T., Dolag K., Giblin B., Heymans C., Hildebrandt H., Xia Q., MNRAS, 2021, 506
Harris C. R. et al., 2020, Nature, 585, 357
Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399
Heitmann K., Lawrence E., Kwan J., Habib S., Higdon D., 2014, ApJ, 780, 111
Heymans C. et al., 2021, A&A, 646, A140
Hikage C. et al., 2019, Publ. Astron. Soc. Japan, 71, 43
Hinton S. R., 2016, J. Open Source Softw., 1, 00045
Huang H.-J., Eifler T., Mandelbaum R., Dodelson S., 2019, MNRAS, 488, 1652
Huang H.-J., Eifler T., Mandelbaum R., Dodelson S., 2019, MNRAS, 488, 1652
Huang H.-J. et al., 2021, MNRAS, 502, 6010
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jarvis M., Bernstein G., Jain B., 2004, MNRAS, 352, 338
Jung G., Namikawa T., Liguori M., Munshi D., Heavens A., 2021, J. Cosmology Astropart. Phys., 2021, 055
Kacprzak T. et al., 2016, MNRAS, 463, 3653
Kaiser N., 1995, ApJ, 439, L1
Kayo I., Takada M., 2013, Cosmological Parameters from Weak Lensing Power Spectrum and Bispectrum Tomography: Including the Non-Gaussian Errors. preprint (arXiv:1306.4684)
Kilbinger M., Schneider P., 2005, A&A, 442, 69
Kitching T. D., Alsing J., Heavens A. F., Jimenez R., McEwen J. D., Verde L., 2017, MNRAS, 469, 2737
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lazanu A., Liguori M., 2018, J. Cosmology Astropart. Phys., 04, 055
Lepage G. P., 1980, Technical report, VEGAS - An Adaptive Multi-Dimensional Integration Program. Cornell Univ. Lab. Nucl. Stud., Ithaca, NY
Lesgourgues J., 2011, The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview. preprint (arXiv:1104.2932)
Li Y., Hu W., Takada M., 2014, Phys. Rev. D, 89, 083519
LSST Dark Energy Science Collaboration, 2012, preprint (arXiv:1211.0310)
Martinet N., Harnois-Déraps J., Jullo E., Schneider P., 2021, A&A, 646, A62
McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936
Mead A. J., Peacock J. A., Heymans C., Joudaki S., Heavens A. F., 2015, MNRAS, 454, 1958
Mead A. J., Brieden S., Tr'oster T., Heymans C., 2021, MNRAS, 502, 1401
Mohammed I., Seljak U., 2014, MNRAS, 445, 3382
Munshi D., McEwen J., Kitching T., Fosalba P., Teyssier R., Stadel J., 2020, J. Cosmology Astropart. Phys., 2020, 043
Munshi D., Jung G., Kitching T. D., McEwen J., Liguori M., Namikawa T., Heavens A., 2021, preprint (arXiv:2104.01185)
Pillepich A., et al., 2018, MNRAS, 473, 4077
Pires S. et al., 2020, A&A, 638, A141
Sato M., Nishimichi T., 2013, Phys. Rev. D, 87

Schaye J. et al., 2010, MNRAS, 402, 1536

Schaye J. et al., 2015, MNRAS, 446, 521

Schmidt A. S., White S. D. M., Schmidt F., Stücker J., 2018, MNRAS, 479, 162

Schneider P., 1996, MNRAS, 283, 837

Schneider P., Kochanek C., Wambsganss J., 2006, Gravitational Lensing: Strong, Weak and Micro, Springer Berlin, Heidelberg

Schneider A., Teyssier R., 2015, J. Cosmology Astropart. Phys., 2015, 049

Schneider A., Stoira N., Refregier A., Weiss A. J., Knabenhans M., Stadel J., Teyssier R., 2020, J. Cosmology Astropart. Phys., 2020, 019

Scoccimarro R., Couchman H. M. P., 2001, MNRAS, 325, 1312

Seehars S., Grandis S., Amara A., Refregier A., 2016, Phys. Rev. D, 93

Semboloni E., Schrabback T., van Waerbeke L., Vafaei S., Hartlap J., Hilbert S., 2010, MNRAS, 410, 143

Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, MNRAS, 417, 2020

Semboloni E., Hoekstra H., Schaye J., 2013, MNRAS, 434, 148

Shirasaki M., Hamana T., Yoshida N., 2015, MNRAS, 453, 3043

Spergel D. et al., 2015, preprint (arXiv:1503.03757)

Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79

Springel V., 2005, MNRAS, 364, 1105

Stücker J., Schmidt A. S., White S. D. M., Schmidt F., Hahn O., 2021, MNRAS, 503, 1473

Takada M., Jain B., 2004, MNRAS, 348, 897

Takahashi R., Hamana T., Shirasaki M., Namikawa T., Nishimichi T., Osato K., Shiroyama K., 2017, ApJ, 850, 24

Takahashi R., Nishimichi T., Namikawa T., Taruya A., Kayo I., Osato K., Kobayashi Y., Shirasaki M., 2020, ApJ, 895, 113

Tegmark M., Taylor A. N., Heavens A. F., 1997, ApJ, 480, 22

Vogelsberger M. et al., 2014, Nature, 509, 177

Wagner C., Schmidt F., Chiang C. T., Komatsu E., 2015, MNRAS, 448, L11

Weinberger R. et al., 2017, MNRAS, 465, 3291

Xavier H. S., Abdalla F. B., Joachimi B., 2016, MNRAS, 459, 3693

Zonca A., Singer L., Lenz D., Reinecke M., Rosset C., Hivon E., Gorski K., 2019, J. Open Source Softw., 4, 1298

Zürcher D. et al., 2022, MNRAS, 511, 2075

APPENDIX A: AUXILIARY EQUATIONS FOR THEORETICAL MODELLING

As mentioned in Section 2.1, to partly correct for the flat-sky and the Limber approximations that go into the derivation e.g. of the convergence power spectrum in equation (12), we apply an l-dependent correction factor proposed by Kitching et al. (2017):

$$C_{\kappa,\mathrm{gh}}(l) \equiv \frac{(l+2)(l+1)l(l-1)}{\left(l+\frac{1}{2}\right)^4} P_{\kappa,\mathrm{gh}}\left(l+\frac{1}{2}\right). \tag{A1}$$

Also, while converting the Fourier space power to shear correlation functions using the inverse Hankel transform l-integrals [e.g. integrals involving the $J_{0,4}$ Bessel functions in equation (11)], we use expressions with summation over l as given in Friedrich et al. (2021):

$$\xi_{\pm,\mathrm{gh}}(\alpha) = \sum_{l\geq 2} \frac{2l+1}{4\pi} \frac{2\left(\overline{G^+_{l,2}}(\cos\alpha) \pm \overline{G^-_{l,2}}(\cos\alpha)\right)}{l^2(l+1)^2} C_{\kappa,\mathrm{gh}}(l), \tag{A2}$$

where the functions $\overline{G^\pm_{l,2}}(x)$ account for the finite bin width in which the correlation function is actually measured at a given angular separation α, i.e. $\alpha \in [\alpha_{\min}, \alpha_{\max}]$. These can be expressed in terms of associated Legendre polynomials $\mathcal{P}_l(x)$ and their analytic forms can be found in appendix B of Friedrich et al. (2021). These equations are exact for a curved-sky treatment and more accurate than the inverse Hankel transforms which are strictly only valid in the flat-sky approximation and do not take into account the finite bin widths in which the correlations are measured in data.

This paper has been typeset from a TEX/LATEX file prepared by the author.

Chapter 5

Cosmology from the integrated shear 3-point correlation function: simulated likelihood analyses with machine-learning emulators

Bibliographic and copyright information

Author's contribution

In this work we further extended the model of the integrated shear 3PCF $\zeta_\pm$ to account for more weak lensing systematic effects and also developed dedicated emulators to produce rapid model predictions of $\zeta_\pm$ for MCMC analyses. As second author in this paper, together with the first author Z. Gong and third author A. Barreira, I formulated analytical models to account for systematic effects such as shear calibration bias, photometric redshift uncertainty and intrinsic alignment in $\zeta_\pm$. To assess the validity of our systematic infused theoretical model, I processed the T17 (Takahashi et al., 2017) lensing simulations

to mimic realistic redshift distributions and shape noise for DESY3-like survey (also created `FLASK` lognormal mocks (Xavier et al., 2016)), performed the measurements of $\zeta_\pm$ from the simulations as well as provide all the training and test datasets for building the the machine-learning emulator. I extensively supported and guided the first author who (i) developed the neural network emulator for $\zeta_\pm$ using the `COSMOPOWER` framework (Spurio Mancini et al., 2022), (ii) carried out the numerical calculations along with all the simulated likelihood analyses of $\zeta_\pm$ for various analysis setups and (iii) wrote the paper. I also contributed to the drafting process by assisting in thorough reviewing of the manuscript. The other authors contributed through valuable discussions, ideas and in reviewing of the paper draft.

Cosmology from the integrated shear 3-point correlation function: simulated likelihood analyses with machine-learning emulators

Zhengyangguang Gong,[a,b] **Anik Halder,**[a,b] **Alexandre Barreira,**[c,d] **Stella Seitz**[a,b] **and Oliver Friedrich**[a]

[a]Universitäts-Sternwarte, Fakultät für Physik,
 Ludwig-Maximilians Universität München,
 Scheinerstraße 1, 81679 München, Germany
[b]Max Planck Institute for Extraterrestrial Physics,
 Giessenbachstraße 1, 85748 Garching, Germany
[c]Excellence Cluster ORIGINS,
 Boltzmannstraße 2, 85748 Garching, Germany
[d]Ludwig-Maximilians-Universität,
 Schellingstraße 4, 80799 München, Germany

 E-mail: lgong@usm.lmu.de, ahalder@usm.lmu.de, alex.barreira@origins-cluster.de, stella@usm.lmu.de, oliver.friedrich@physik.uni-muenchen.de

Abstract. The integrated shear 3-point correlation function $\zeta_\pm$ measures the correlation between the local shear 2-point function $\xi_\pm$ and the 1-point shear aperture mass in patches of the sky. Unlike other higher-order statistics, $\zeta_\pm$ can be efficiently measured from cosmic shear data, and it admits accurate theory predictions on a wide range of scales as a function of cosmological and baryonic feedback parameters. Here, we develop and test a likelihood analysis pipeline for cosmological constraints using $\zeta_\pm$. We incorporate treatment of systematic effects from photometric redshift uncertainties, shear calibration bias and galaxy intrinsic alignments. We also develop an accurate neural-network emulator for fast theory predictions in MCMC parameter inference analyses. We test our pipeline using realistic cosmic shear maps based on N-body simulations with a DES Y3-like footprint, mask and source tomographic bins, finding unbiased parameter constraints. Relative to $\xi_\pm$-only, adding $\zeta_\pm$ can lead to $\approx 10 - 25\%$ improvements on the constraints of parameters like A_s (or σ_8) and w_0. We find no evidence in $\xi_\pm + \zeta_\pm$ constraints of a significant mitigation of the impact of systematics.

We also investigate the impact of the size of the apertures where $\zeta_\pm$ is measured, and of the strategy to estimate the covariance matrix (N-body vs. lognormal). Our analysis solidifies the strong potential of the $\zeta_\pm$ statistic and puts forward a pipeline that can be readily used to improve cosmological constraints using real cosmic shear data.

Contents

1 Introduction

The weak gravitational lensing effect is the bending of the light of background source galaxies by foreground gravitational potentials [1, 2]. This induces a coherent distortion pattern in the observed shape of the background galaxies that is called the *cosmic shear field*. The statistics of this field depend on the three-dimensional large-scale structure, hence cosmic shear studies offer a powerful way to address key questions in cosmology such as the structure formation history, the nature of dark energy and dark matter, and the laws of gravity on large scales. Indeed, cosmic shear is one of the most active research areas in large-scale structure today: the DES [3], KiDS [4] and HSC-SSP [5] surveys have recently presented cosmological constraints from their cosmic shear data, and more accurate and bigger data sets will be available soon with missions like Euclid [6], Vera Rubin's LSST [7] and Nancy Roman [8].

The majority of cosmic shear analyses are based on the shear 2-point correlation function (2PCF), or its Fourier counterpart the lensing power spectrum. These statistics completely characterize the information content of Gaussian random fields, which our Universe was close to at the earliest stages of its evolution, as well as today on sufficiently large-scales. At late times, however, the evolution of matter density fluctuations becomes nonlinear on small scales, inducing non-Gaussian features in the cosmic shear field that cannot be described by 2PCF alone. Higher-order statistics are thus needed to access the non-Gaussian information.

The shear 3-point correlation function (3PCF; or its Fourier counterpart the lensing bispectrum) is the natural first step beyond the 2PCF [9–14]. However, being a more complicated statistic, it is more challenging to measure observationally, as well as to predict

theoretically. Concretely, compared to the 2PCF which depends only on the distance between two points in the survey footprint, the 3PCF is a function of the size and shape of triangles connecting three points, which requires more demanding estimators. Additionally, theoretical predictions require accurate prescriptions for the nonlinear matter bispectrum, which despite recent progress [15, 16], are still not as developed as the matter power spectrum that enters the shear 2PCF. Further complications arise by the need to account for baryonic feedback effects, as well as systematics effects such as photometric redshift uncertainties, shear multiplicative bias and galaxy intrinsic alignments (IA). This helps explain why existing real-data constraints using higher-order shear information are based not on the full 3-point correlation function, but on other statistics including aperture moments [17–21], lensing peaks [22–24], density-split statistics [25–29] and persistent homology of cosmic shear [30, 31]. The shear 3PCF was recently measured using DES Year 3 (Y3) data [20], although only in patches over the survey and not over the whole footprint as that would be too computationally demanding.

In this paper, we focus on a particular kind of shear 3PCF called the *integrated shear 3-point correlation function* [32]. This statistic corresponds to the correlation between the shear 2PCF measured in patches of the sky with the 1-point shear aperture mass in those patches.[1] Physically, this statistic describes the modulation of the local shear 2PCF by long-wavelength features in the cosmic shear field. The integrated shear 3PCF enjoys two key advantages relative to other higher-order shear statistics. The first is that it is straightforward to measure from the data as it requires only conventional and well-tested shear 2PCF estimators. The second is that, as shown in ref. [38], this statistic is sensitive to the squeezed matter bispectrum that can be evaluated accurately in the nonlinear regime using the response approach to perturbation theory [39]. Importantly, the response approach allows to account for the impact of baryonic feedback on small scales, which is crucial to design scale cuts and/or marginalize over these uncertainties in real data analyses.

Our goal here is to develop and test a likelihood analysis pipeline to reliably extract cosmology from real cosmic shear data using the integrated shear 3PCF. Concretely, we incorporate the impact of baryonic feedback (as in ref. [38]), as well as of photometric redshift uncertainties, shear multiplicative bias and galaxy IA. We also develop a neural-network (NN) emulator for the theory model to enable fast theory predictions in Monte-Carlo Markov Chain (MCMC) parameter inference analyses. We test our analysis pipeline on simulated cosmic shear maps with DES Y3-like survey footprints and source galaxy redshift distributions. We study in particular (i) the ability of the theory model to return unbiased parameter constraints,[2] (ii) the impact of the size of the aperture where the integrated shear 3PCF is measured, (iii) the ability of combined 2PCF and 3PCF analyses to mitigate the impact of systematic uncertainties, and (iv) the impact of different data vector covariance estimates.

In terms of constraining power, we find that the integrated shear 3PCF leads to improvements of $\approx 10 - 25\%$ on the constraints of parameters like the amplitude of primordial density fluctuations A_s (or equivalently σ_8) or the dark energy equation of state parameter w_0. This is consistent with the previous findings of refs. [32, 38] based on idealized Fisher matrix forecasts, but now in the context of realistically simulated MCMC likelihood analyses. Our results thus strongly motivate as next steps exploring the power of this statistic to improve cosmological constraints using real cosmic shear data.

---

[1]See also refs. [33, 34] for earlier applications of the same idea in the context of the three-dimensional galaxy distribution, and refs. [35–37] for studies of the Fourier counterpart of the integrated shear 3PCF.

[2]Throughout the paper we loosely use the term "unbiased constraints" to mean that the 68% posterior credible intervals encompass the true model parameter values.

This paper is structured as follows: In section 2 we review the theoretical formalism behind the integrated shear 3PCF and describe how we incorporate lensing systematic effects. In section 3 we describe the construction of our DES Y3-like cosmic shear maps, as well as the measurements of the shear 2PCF, integrated 3PCF and their (cross) covariance matrices. We describe and discuss the performance of our NN emulator of the theory predictions for fast MCMC likelihood analyses in section 4. Our main numerical results are shown in section 5. We summarize and conclude in section 6. Appendix A describes our modelling of the galaxy IA.

2 Theoretical formalism

In this section we describe the theory behind the integrated shear 3PCF. We begin with a recap of the model of refs. [32, 38], and then discuss how we incorporate lensing systematics.

2.1 Integrated shear 3-point correlation function

The integrated shear 3PCF, $\zeta_{\pm,ijk}(\boldsymbol{\alpha})$, is defined as

$$\zeta_{\pm,ijk}(\boldsymbol{\alpha}) \equiv \left\langle M_{\mathrm{ap},i}(\boldsymbol{\theta}_C)\hat{\xi}_{\pm,jk}(\boldsymbol{\alpha};\boldsymbol{\theta}_C)\right\rangle, \tag{2.1}$$

where $M_{\mathrm{ap},i}(\boldsymbol{\theta}_C)$ is the 1-point aperture mass statistic measured on a patch of the survey centered at angular position $\boldsymbol{\theta}_C$, and $\hat{\xi}_{\pm,jk}(\boldsymbol{\alpha};\boldsymbol{\theta}_C)$ is the shear 2PCF measured on the same patch of the sky; $\boldsymbol{\alpha}$ describes angular separations. The angle brackets denote ensemble average (or in practice, averaging over all positions $\boldsymbol{\theta}_C$) and the subscripts i, j, k denote tomographic source bins, i.e. $\hat{\xi}_{\pm,jk}$ is the 2PCF of the shear fields from galaxy shape measurements at the redshift bins j and k. This equation makes apparent the interpretation of the shear 3PCF as describing the spatial modulation of the local 2PCF by the local shear mass aperture, which describes larger-scale features in the shear field.

The aperture mass $M_{\mathrm{ap}}(\boldsymbol{\theta}_C)$ is defined as [2, 40]

$$M_{\mathrm{ap}}(\boldsymbol{\theta}_C) = \int \mathrm{d}^2\boldsymbol{\theta}\ \kappa(\boldsymbol{\theta})U(\boldsymbol{\theta}_C - \boldsymbol{\theta}), \tag{2.2}$$

where $\kappa(\boldsymbol{\theta})$ is the lensing convergence field, and U is an azimuthally symmetric filter function with angular size θ_{ap}. The convergence field is not directly observable, but if U is a compensated filter satisfying $\int \mathrm{d}^2\boldsymbol{\theta}\ U(\boldsymbol{\theta}_C - \boldsymbol{\theta}) = 0$, then $M_{\mathrm{ap}}(\boldsymbol{\theta}_C)$ can be expressed as

$$M_{\mathrm{ap}}(\boldsymbol{\theta}_C) = \int \mathrm{d}^2\boldsymbol{\theta}\ \gamma_{\mathrm{t}}(\boldsymbol{\theta}, \phi_{\boldsymbol{\theta}_C-\boldsymbol{\theta}})Q(\boldsymbol{\theta}_C - \boldsymbol{\theta}), \tag{2.3}$$

where γ_{t} is the tangential component of the shear field (which is directly observable), $\phi_{\boldsymbol{\theta}_C-\boldsymbol{\theta}}$ is the polar angle of the angular separation between $\boldsymbol{\theta}_C$ and $\boldsymbol{\theta}$, and Q is a filter function related to U. As in previous works, we adopt the following form for U and Q [41]

$$U(\theta) = \frac{1}{2\pi\theta_{\mathrm{ap}}^2}\left(1 - \frac{\theta^2}{2\theta_{\mathrm{ap}}^2}\right)\exp\left(-\frac{\theta^2}{2\theta_{\mathrm{ap}}^2}\right), \tag{2.4}$$

$$Q(\theta) = \frac{\theta^2}{4\pi\theta_{\mathrm{ap}}^2}\exp\left(-\frac{\theta^2}{2\theta_{\mathrm{ap}}^2}\right); \tag{2.5}$$

note the filters depend only on the magnitude of the arguments because of the azimuthal symmetry. The Fourier transform of U, which appears in equations below, is given by

$$U(\ell) = \int \mathrm{d}^2\boldsymbol{\theta} \, U(\theta)e^{-i\boldsymbol{\ell}\cdot\boldsymbol{\theta}} = \frac{\ell^2\theta_{\mathrm{ap}}^2}{2}\exp\left(-\frac{\ell^2\theta_{\mathrm{ap}}^2}{2}\right), \tag{2.6}$$

where $\boldsymbol{\ell}$ is a two-dimensional wavevector on the sky (we assume the flat-sky approximation).

The other term in eq. (2.1), $\hat{\xi}_\pm(\boldsymbol{\alpha}; \boldsymbol{\theta}_C)$, is the 2PCF of the *windowed* shear field $\gamma(\boldsymbol{\theta}; \boldsymbol{\theta}_C) \equiv \gamma(\boldsymbol{\theta})W(\boldsymbol{\theta}_C - \boldsymbol{\theta})$, where the window function W is a top-hat of size θ_{T} at position $\boldsymbol{\theta}_C$. The two 2PCFs are defined as

$$\hat{\xi}_+(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) \equiv \frac{1}{A_{2\mathrm{pt}}(\boldsymbol{\alpha})} \int \mathrm{d}^2\boldsymbol{\theta} \, \gamma(\boldsymbol{\theta}; \boldsymbol{\theta}_C)\gamma^*(\boldsymbol{\theta} + \boldsymbol{\alpha}; \boldsymbol{\theta}_C)$$

$$\hat{\xi}_-(\boldsymbol{\alpha}; \boldsymbol{\theta}_C) \equiv \frac{1}{A_{2\mathrm{pt}}(\boldsymbol{\alpha})} \int \mathrm{d}^2\boldsymbol{\theta} \, \gamma(\boldsymbol{\theta}; \boldsymbol{\theta}_C)\gamma(\boldsymbol{\theta} + \boldsymbol{\alpha}; \boldsymbol{\theta}_C)e^{-4i\phi_\alpha}, \tag{2.7}$$

where * denotes complex conjugation, $\phi_{\boldsymbol{\alpha}}$ is the polar angle of $\boldsymbol{\alpha}$, and $A_{2\mathrm{pt}}(\boldsymbol{\alpha}) \equiv \int \mathrm{d}^2\boldsymbol{\theta} \, W(\boldsymbol{\theta}_C - \boldsymbol{\theta})W(\boldsymbol{\theta}_C - \boldsymbol{\theta} - \boldsymbol{\alpha})$. The Fourier transform of W appears in equations below, and is given by

$$W(\boldsymbol{l}) = W(l) = 2\pi\theta_{\mathrm{T}}^2 \, \frac{J_1(l\theta_{\mathrm{T}})}{l\theta_{\mathrm{T}}}, \tag{2.8}$$

where J_n is the nth-order Bessel function of the first kind.

Skipping the details of the derivation [32], the two 3PCF in eq. (2.1) can be written as

$$\zeta_{+,ijk}(\boldsymbol{\alpha}) = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\ell}{2\pi}\mathcal{B}_{+,ijk}^{2\mathrm{D}}(\boldsymbol{\ell})J_0(\ell\alpha), \tag{2.9}$$

$$\zeta_{-,ijk}(\boldsymbol{\alpha}) = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\ell}{2\pi}\mathcal{B}_{-,ijk}^{2\mathrm{D}}(\boldsymbol{\ell})J_4(\ell\alpha), \tag{2.10}$$

where $\mathcal{B}_\pm^{2\mathrm{D}}$ is called the integrated lensing bispectrum, and it is given by (in the Limber approximation)

$$\mathcal{B}_{\pm,ijk}^{2\mathrm{D}}(\boldsymbol{\ell}) = \int \mathrm{d}\chi\frac{q^i(\chi)q^j(\chi)q^k(\chi)}{\chi^4} \int \frac{\mathrm{d}^2\boldsymbol{\ell}_1}{(2\pi)^2} \int \frac{\mathrm{d}^2\boldsymbol{\ell}_2}{(2\pi)^2}B_\delta^{3\mathrm{D}}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2}{\chi}, \chi\right)$$

$$\times \, e^{2i(\phi_2 \mp \phi_{-1-2})}U(\boldsymbol{\ell}_1)W(\boldsymbol{\ell}_2 + \boldsymbol{\ell})W(-\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2 - \boldsymbol{\ell}). \tag{2.11}$$

In this equation, B_δ^{3D} is the 3-dimensional matter bispectrum (discussed below), ϕ_2 is the polar angle of $\boldsymbol{\ell}_2$, ϕ_{-1-2} is the polar angle of $-\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2$, and $q(\chi)$ is the lensing kernel

$$q^i(\chi) = \frac{3H_0^2\Omega_m}{2c^2}\frac{\chi}{a(\chi)} \int_\chi \mathrm{d}\chi'n_s^i(\chi')\frac{\chi' - \chi}{\chi'}, \tag{2.12}$$

where $n_s^i(\chi)$ is the galaxy source number density distribution for the redshift tomographic bin i, χ denotes comoving distances, H_0 is the Hubble parameter, Ω_m is the cosmic matter density parameter today, c is the speed of light and $a(\chi)$ is the scale factor; note that throughout the paper we always assume spatially flat cosmologies.

In our results, we will consider also the *global* shear 2PCF, which can be evaluated as

$$\xi_{+,ij}(\alpha) = \int \frac{\mathrm{d}\ell\ell}{2\pi}P_{\kappa,ij}(\ell)J_0(\ell\alpha), \tag{2.13}$$

$$\xi_{-,ij}(\alpha) = \int \frac{\mathrm{d}\ell\ell}{2\pi}P_{\kappa,ij}(\ell)J_4(\ell\alpha), \tag{2.14}$$

where $P_{\kappa,ij}$ is the convergence power spectrum given by (in the Limber approximation)

$$P_{\kappa,ij}(\ell) = \int \mathrm{d}\chi \frac{q^i(\chi)q^j(\chi)}{\chi^2} P_\delta^{3D}\left(k = \frac{\ell}{\chi}, \chi\right), \qquad (2.15)$$

with P_δ^{3D} the three-dimensional matter power spectrum.

2.2 The three-dimensional matter bispectrum model

A key ingredient to evaluate $\zeta_\pm$ is the three-dimensional matter bispectrum B_δ^{3D} in eq. (2.11), which we evaluate following ref. [38] as

$$B_\delta^{3D}(\boldsymbol{k_1}, \boldsymbol{k_2}, \boldsymbol{k_3}, \chi) = \begin{cases} B_{\delta,\mathrm{RF}}^{3D}, & f_{\mathrm{sq}} \geq f_{\mathrm{sq}}^{\mathrm{thr}} \Longrightarrow \text{squeezed} \\ B_{\delta,\mathrm{GM}}^{3D}, & \text{otherwise} \end{cases}, \qquad (2.16)$$

where $B_{\delta,\mathrm{RF}}^{3D}$ is the bispectrum expression of the response function approach valid for squeezed configurations, and $B_{\delta,\mathrm{GM}}^{3D}$ is the bispectrum fitting formula of ref. [42]. The parameter f_{sq} is defined as $f_{\mathrm{sq}} = k_m/k_s$, with k_s (k_m) the smallest (intermediate) of the amplitudes of the three modes $\boldsymbol{k_i}$. As explained in ref. [38], this equation guarantees that the response function branch correctly evaluates the squeezed matter bispectrum configurations in the nonlinear regime, which determine the value of $\zeta_\pm$ on small angular scales. The value of $f_{\mathrm{sq}}^{\mathrm{thr}}$ is the threshold that defines whether a given bispectrum configuration is dubbed as squeezed or not. Ref. [38] found that a range of values around $f_{\mathrm{sq}}^{\mathrm{thr}} \approx 7$ yield good fits to simulation measurements; in this paper we adopt $f_{\mathrm{sq}}^{\mathrm{thr}} = 7$.

The response function branch in eq. (2.16) is given by

$$B_{\delta,\mathrm{RF}}^{3D}(\boldsymbol{k_1}, \boldsymbol{k_2}, \boldsymbol{k_3}, z) = \left[R_1(k_h, z) + \left(\mu_{\boldsymbol{k_h}, \boldsymbol{k_s}}^2 - \frac{1}{3}\right) R_K(k_h, z)\right] P_\delta^{3D}(k_h, z) P_{\delta,L}^{3D}(k_s, z), \quad (2.17)$$

where $\boldsymbol{k_h}$ denotes the mode $\boldsymbol{k_i}$ with the highest magnitude, $\mu_{\boldsymbol{k_i}, \boldsymbol{k_j}}$ is the cosine of the angle between $\boldsymbol{k_i}$ and $\boldsymbol{k_j}$, $P_{\delta,L}^{3D}$ is the three-dimensional linear matter power spectrum and $R_1(k, z)$ and $R_K(k, z)$ are the first-order response functions of the matter power spectrum to large-scale density and tidal fields:

$$R_1(k, z) = 1 - \frac{1}{3}\frac{\mathrm{d}\ln P_\delta^{3D}(k, z)}{\mathrm{d}\ln k} + G_1(k, z), \qquad (2.18)$$

$$R_K(k, z) = G_K(k, z) - \frac{\mathrm{d}\ln P_\delta^{3D}(k, z)}{\mathrm{d}\ln k}. \qquad (2.19)$$

In these expressions, G_1 and G_K are the so-called *growth-only* response functions, which can be measured in the nonlinear regime of structure formation using separate universe simulations. Just as in ref. [38], we use the results of ref. [43] for G_1 and ref. [44] for G_K.

The GM branch is in turn given by

$$B_{\delta,\mathrm{GM}}^{3D}(\boldsymbol{k_1}, \boldsymbol{k_2}, \boldsymbol{k_3}, z) = 2F_2^{\mathrm{eff}}(\boldsymbol{k_1}, \boldsymbol{k_2}, z) P_\delta^{3D}(k_1, z) P_\delta^{3D}(k_2, z) + \text{cyclic permutations}, \quad (2.20)$$

where $F_2^{\mathrm{eff}}(\boldsymbol{k_1}, \boldsymbol{k_2}, z)$ is a modified version of the 2-point mode coupling kernel with free functions calibrated against N-body simulations [42].

In this paper, we evaluate the nonlinear matter power spectrum using the `HMcode` [45] implementation inside the publicly available Boltzmann code `CLASS` [46]; to model the impact

of baryonic feedback effects, we adopt the single parameter $c_{\min}$ parametrization, where $c_{\min}$ roughly describes the strength of feedback by active galactic nuclei (AGN). As discussed in refs. [47, 48], mode-coupling terms like F_2^{eff}, G_1 and G_K are expected to be very weakly dependent on baryonic physics. This way, the impact of baryonic effects on the bispectrum is trivially propagated by that on the power spectrum; note that in practice the baryonic effects impact only the response function branch in eq. (2.16), since the GM branch contributes only on large scales [38] where baryonic effects have a negligible role.

2.3 Systematic error effects

Reference [38] has shown how to include the impact of baryonic feedback effects on $\zeta_\pm$, which are one of the main non-cosmological contaminants in cosmic shear analyses. In this subsection we describe how we take into account a series of other important systematic effects, namely photometric redshift uncertainties, multiplicative shear bias and galaxy IA.

Photometric redshift (photo-z) uncertainties have a direct impact on the galaxy source redshift distribution. Here, we follow a strategy commonly adopted in real-data analyses and parametrize their effect through a single shift parameter Δz defined as

$$n_s^i(z) = \hat{n}_s^i(z + \Delta z^i),\tag{2.21}$$

where $\hat{n}_s^i$ is the default estimate for the galaxy source redshift bin i. This simple way to account for photo-z uncertainties was found sufficient at the statistical power of DES-Y3 analyses (see figure 10 in ref. [49]).

Again, as common in the literature, we model biases from the shear measurement pipeline with multiplicative factors $1 + m_i$ for each tomographic bin i. In practice, this implies the following transformations of $\xi_\pm$ and $\zeta_\pm$,

$$\xi_{\pm,ij}(\alpha) \longrightarrow (1 + m_i)(1 + m_j)\xi_{\pm,ij}(\alpha),\tag{2.22}$$

$$\zeta_{\pm,ijk}(\alpha) \longrightarrow (1 + m_i)(1 + m_j)(1 + m_k)\zeta_{\pm,ijk}(\alpha).\tag{2.23}$$

We assume that any additive bias component is well calibrated by lensing image simulations and removed from the measurement pipeline [50].

Finally, we consider the effect of galaxy IA that describe intrinsic correlations between the shapes of source galaxies and their local tidal fields, i.e. correlations that are not induced by the gravitational lensing effect. We adopt the *nonlinear linear alignment* (NLA) model [51, 52] for both $\xi_\pm$ and $\zeta_\pm$. In practice, the incorporation of IA in our theory predictions is equivalent to transforming the lensing kernels as (see appendix A for more details)

$$q^i(\chi) \longrightarrow q^i(\chi) + f_{\text{IA}}(z(\chi))\frac{n_s^i(\chi)}{\bar{n}_s^i}\frac{\mathrm{d}z}{\mathrm{d}\chi},\tag{2.24}$$

with $\bar{n}_s^i$ the mean source galaxy density in tomographic bin i and [53, 54]

$$f_{\text{IA}}(z) = -A_{\text{IA},0}\left(\frac{1+z}{1+z_0}\right)^{\alpha_{\text{IA}}}\frac{c_1\rho_{\text{crit}}\Omega_{\text{m},0}}{D(z)},\tag{2.25}$$

where $A_{\text{IA},0}$ is the IA amplitude, α_{IA} is a power index and $D(z)$ is the linear growth factor. We adopt $z_0 = 0.62$, $c_1\rho_{\text{crit}} = 0.0134$ [52, 55]. In our results below we keep the power index fixed to $\alpha_{\text{IA}} = 0$ for simplicity; note that simultaneously varying $A_{\text{IA},0}$ and α_{IA} in MCMC

constraints can lead to posterior projection effects that could artificially bias the marginalized constraints of $A_{\mathrm{IA},0}$ towards zero.

In our modelling of IA, $\xi_\pm$ acquires terms $\propto f_{\mathrm{IA}}, f_{\mathrm{IA}}^2$, and $\zeta_\pm$ terms $\propto f_{\mathrm{IA}}, f_{\mathrm{IA}}^2, f_{\mathrm{IA}}^3$. These are different from the terms displayed in ref. [56]; further notice that our eq. (2.25) differs from the corresponding eq. (27) in ref. [56] by a multiplicative factor $1/(1+z)$. We shall return to the impact of different IA treatments when we discuss our numerical results.

3 Data vector and covariance from simulations

In this section we describe the DES Y3-like simulated cosmic shear maps that we use to measure the $\xi_\pm$ and $\zeta_\pm$ data vectors and to estimate their covariance matrices.

3.1 Shear maps from N-body simulations

Our main cosmic shear maps are obtained using the publicly available N-body simulation data developed by Takahashi et al. [57] (hereafter referred to as T17). In particular, we make use of the 108 independent full-sky cosmic shear maps for several Dirac-delta source distributions at redshifts between $z = 0.05$ and $z = 5.3$. The cosmology of the simulations is flat ΛCDM with parameters: $\Omega_m = 0.279$, $\Omega_b = 0.046$, $h = 0.7$, $\sigma_8 = 0.82$, $n_s = 0.97$.

We consider DES Y3-like galaxy source redshift distributions to construct our cosmic shear maps. For simplicity, rather than considering the four source bins utilized in the DES Y3 analysis, we merge them into two as follows. Let N_1 and N_2 be the total number of galaxies in the first two DES source distributions $n_{s,\mathrm{DES1}}(z)$ and $n_{s,\mathrm{DES2}}(z)$, respectively (see figure 6 and 11 in ref. [58]). Then, our first source redshift bin is obtained as $\hat{n}_s^1 = (N_1 n_{s,\mathrm{DES1}} + N_2 n_{s,\mathrm{DES2}})/(N_1 + N_2)$; and similarly for our second source redshift, using the third and fourth DES Y3 source distributions. The source redshift distributions that we consider in this paper are shown on the left of figure 1. For each of the 108 T17 realizations, we build two full-sky shear maps by summing the T17 shear maps weighted by each of the two source redshift distributions. The vertical lines on the left of figure 1 mark the source redshift of the T17 maps we use.

We then apply the DES Y3 footprint to each of the full-sky shear maps. In order to maximize the utility of each full-sky map, we place 5 footprints in each with minimal overlap, as illustrated on the right of figure 1. For each of our two source bins, this provides us with $108 \times 5 = 540$ DES Y3-like shear maps on which we can measure $\xi_\pm$, $\zeta_\pm$ and their covariance.

Finally, we add DES Y3 levels of shape noise to our maps as follows. Using the angular positions of the source galaxies in the DES Y3 shape catalogue [59], we assign to each of our pixels the galaxy ellipticities and measurement weights that are also present in those catalogues. We then randomly rotate the ellipticities of the galaxies assigned to each pixel. The shape noise γ_{noise} is the average of these randomly rotated ellipticities weighted by the corresponding measurement weights. This is added to the shear values of the T17 maps γ_{sim} to generate the shear measurement in each pixel γ_{pix}. Concretely,

$$\gamma_{\mathrm{pix}} = \gamma_{\mathrm{noise}} + \gamma_{\mathrm{sim}} = \frac{\sum_{j=1}^{N} \omega_j \gamma_{j,\mathrm{DES}} \exp(i\phi_j)}{\sum_{j=1}^{N} \omega_j} + \gamma_{\mathrm{sim}}\,, \tag{3.1}$$

where N is the number of galaxies in a given pixel, $\gamma_{j,\mathrm{DES}}$ and ω_j are the measured ellipticity and weight of the jth galaxy and each angle ϕ_j is drawn uniformly from $[0, 2\pi]$; note that the average value of γ_{noise} across all pixels is zero, but each pixel has in general nonzero values.
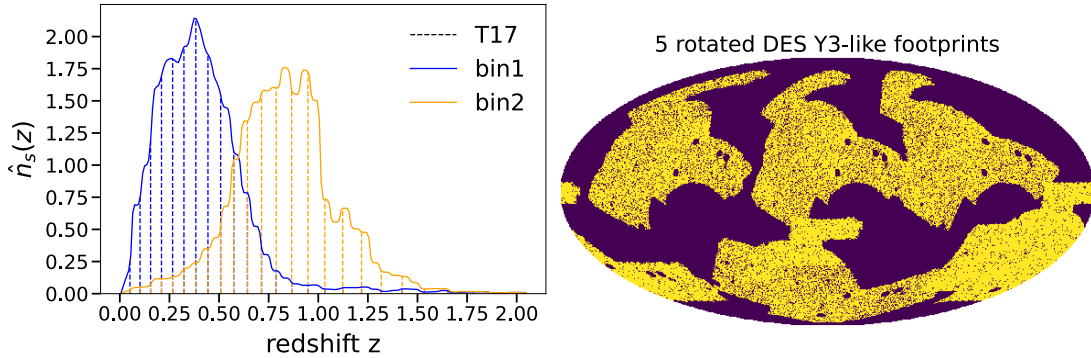
Figure 1. *Left panel*: The two galaxy source redshift distributions that we consider in this paper. Each is a combination of two of the four DES Y3 source distributions. The vertical dashed lines mark the source redshifts of the T17 shear maps, which are weighted by the galaxy source distributions to produce our shear maps. *Right panel*: The Mollweide projection map with the placement of 5 DES Y3 survey footprints after the selection with Q filters of 90 arcmin in a full-sky map; dark blue pixels indicate masked/unobserved regions. This allows us to measure 5 DES Y3-like realizations of $\xi_\pm$ and $\zeta_\pm$ from each full-sky map.

3.2 Shear maps from lognormal realizations

In addition to the T17-based shear maps, we also consider DES Y3-like maps from lognormal lensing realizations generated with the Full-sky Lognormal Astro-fields Simulation Kit [60] (hereafter referred to as FLASK). FLASK takes as input the lensing convergence power spectrum, which we compute theoretically for the T17 cosmology and our two galaxy source redshift distributions. FLASK requires also the value of a logshift parameter, which we obtain by fitting a lognormal probability distribution function (PDF) to the PDF of the T17 maps (see section 4.2 of [32] for more details about the generation of our FLASK shear maps). For each of our two source bins, we generate a total of 300 independent FLASK full-sky cosmic shear maps, on which we place 5 DES Y3-like footprints analogously to the T17 full-sky maps (cf. right panel of figure 1). We add shape noise following the strategy described above for the T17 maps. For each of the two source bins then, we have a total of $5 \times 300 = 1500$ lognormal realizations of a DES Y3-like footprint on which we can measure $\xi_\pm$ and $\zeta_\pm$.

3.3 Data vector and covariance measurements

We use the `Treecorr` code [61] to measure $\xi_{\pm,ij}(\alpha)$ on 15 log-spaced angular bins between 5 and 250 arcmin; these are scales comparable to those adopted in the DES Y3 analysis [62]. We measure the auto- and cross-correlation of the two source redshift bins, yielding a total of 6 shear 2PCFs. The measurements from the T17 maps are shown by the black dots in figure 2.

In order to measure $\zeta_{\pm,ijk}(\alpha)$, we use the `Treecorr` code to compute the position-dependent shear 2PCF and 1-point aperture masses within patches of the footprint; we assume the same size $\theta_{\rm ap}$ and $\theta_{\rm T}$ for the aperture mass and position-dependent 2PCF. The 2PCF in each patch is measured in 15 log-spaced angular bins between 5 and $2\theta_{\rm T} - 10$ arcmin, and the 1-point aperture mass is evaluated using eq. (2.3) with the integral up to $5\theta_{\rm ap}$. The $\zeta_\pm$ is obtained by averaging the product of the shear 2PCF and 1-point aperture mass across
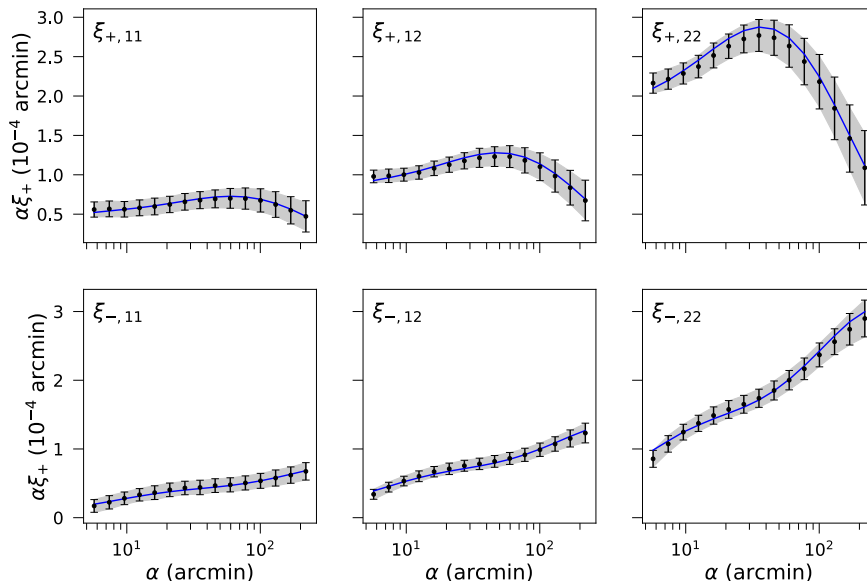
Figure 2. The shear 2PCF $\xi_\pm(\alpha)$ measured from our DES Y3-like footprints for two galaxy source redshift bins. The black dots with the error bars show the mean and the standard deviation of the measurements from the 540 T17 shear maps. For comparison, the grey shaded bands show the standard deviation computed using the 1500 FLASK shear maps. The blue curves show the theoretical result obtained using eqs. (2.13) and (2.14).

all patches selected in the footprint. For our two source redshift bins, we have 8 integrated auto- and cross-3PCF $\zeta_{\pm,ijk}(\alpha)$. The measurements from the T17 maps are shown by the black dots in figure 3 for an aperture size of $\theta_{\mathrm{ap}} = \theta_{\mathrm{T}} = 90$ arcmin.[3]

We estimate the covariance matrix of our data vectors as

$$\hat{C} = \frac{1}{N_s - 1} \sum_{i=1}^{N_s} (\hat{d}_i - \hat{d})(\hat{d}_i - \hat{d})^{\mathrm{T}}, \tag{3.2}$$

where N_s is the number of footprint realizations (540 for T17 and 1500 for FLASK), $\hat{d}_i$ is the data vector of the i-th realization and $\hat{d}$ is the mean data vector across all realizations. When evaluating the inverse covariance matrix, we correct it as

$$\hat{C}^{-1} = \left[\frac{N_{\mathrm{s}} - N_{\mathrm{d}} - 2}{N_{\mathrm{s}} - 1} \right] [1 + A + B(N_{\mathrm{p}} + 1)] C^{-1}, \tag{3.3}$$

where

$$A = \frac{2}{(N_{\mathrm{s}} - N_{\mathrm{d}} - 1)(N_{\mathrm{s}} - N_{\mathrm{d}} - 4)}, \tag{3.4}$$

$$B = \frac{N_{\mathrm{s}} - N_{\mathrm{d}} - 2}{(N_{\mathrm{s}} - N_{\mathrm{d}} - 1)(N_{\mathrm{s}} - N_{\mathrm{d}} - 4)}, \tag{3.5}$$

---

[3]As a technical point, in our measurements of $\zeta_\pm$ we consider only survey patches where the fraction of unmasked pixels is larger than 80% for the top-hat filter W and larger than 70% for the Q filter up to $5\theta_{\mathrm{ap}}$ of aperture radius. Holes and masked pixels inside the footprint contribute to the counting of these fractions, in addition to pixels outside the survey footprint. This ensures our measurements are not affected by too many unmasked pixels in the patches, as confirmed by their excellent agreement with the theory predictions for both $\xi_\pm$ and $\zeta_\pm$ in figures 2 and 3, respectively.
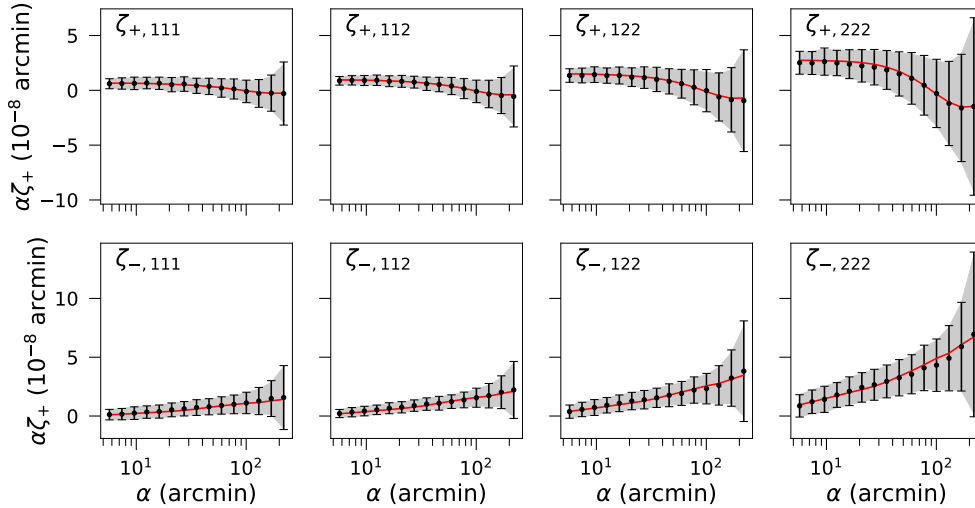
Figure 3. The integrated shear 3PCF $\zeta_\pm(\alpha)$ measured from our DES Y3-like footprints for two galaxy source redshift bins. The black dots with the error bars show the mean and the standard deviation of the measurements from the 540 T17 shear maps for a filter size of 90 arcmin. The grey shaded bands show the standard deviation computed using the 1500 FLASK shear maps. The red curves show the theoretical result from eqs. (2.9) and (2.10).

and N_d is the size of the data vector ($N_d = 90$ for $\xi_\pm$, $N_d = 120$ for $\zeta_\pm$, and $N_d = 210$ for their combination), N_p is the number of inference parameters and C^{-1} is the directly inverted covariance. The first term in brackets is the bias correction on the inverse covariance from ref. [63], while the second term is a correction factor from ref. [64].

The FLASK covariance matrix has the advantage of having less numerical noise because of the larger N_s, but the disadvantage of corresponding to lognormal realizations of cosmic shear maps, which are not as realistic as the T17 ones from N-body simulations. The left panel of figure 4 compares the correlation matrix $r_{mn} = \hat{C}_{mn}/\sqrt{\hat{C}_{mm}\hat{C}_{nn}}$ from the FLASK (upper triangle) and T17 (lower triangle) maps; the indices m, n run over the data vector entries. Reassuringly, the two covariance matrices display broadly the same correlations. There are however some differences that are better seen in the right panel of figure 4 which shows the relative difference between the two covariances. We investigate the impact of these differences in the parameter constraints when we discuss our results below.

We note that both our covariance matrices do not appropriately account for super-sample covariance (SSC) [65, 66], i.e. the variance induced by the gravitational coupling between observed modes inside the survey and unobserved modes with wavelengths larger than the survey size. The SSC is the dominant off-diagonal contribution in 2-point function analyses [67], and it is expected to be a smaller contribution to the squeezed bispectrum configurations that dominate the small-scale $\zeta_\pm$ [68]. Our quoted error bars for $\xi_\pm$-only analyses are thus expected to be underestimated, and consequently, our quoted improvements from $\zeta_\pm$ are conservative; i.e. the relative improvement from $\zeta_\pm$ is expected to be larger in analyses that appropriately account for SSC. We defer the inclusion of SSC to future work.
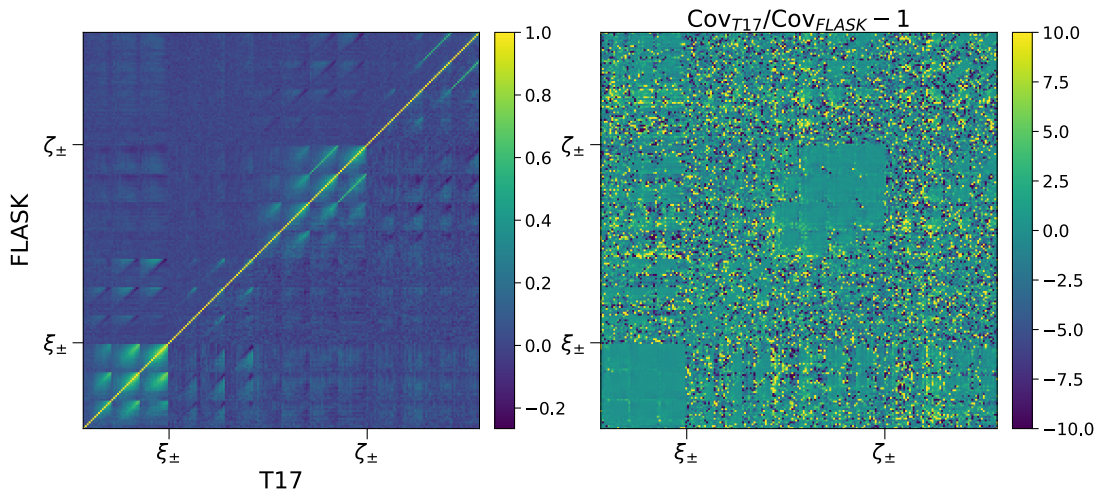
Figure 4. Comparison between the T17 and FLASK covariance matrices. The left panel shows the correlation coefficient from FLASK in the upper triangle and T17 in the lower triangle part of the matrix. The right panel shows the relative difference between the two covariance estimates (the color coding is limited to ± 10 to exclude a few extreme values for visibility). The ordering of the matrix entries is according to: $\{\xi_{+,11}, \xi_{+,12}, \xi_{+,22}, \xi_{-,11}, \xi_{-,12}, \xi_{-,22}, \zeta_{+,111}, \zeta_{+,112}, \zeta_{+,122}, \zeta_{+,222}, \zeta_{-,111}, \zeta_{-,112}, \zeta_{-,122}, \zeta_{-,222}\}$. The result shown for $\zeta_{\pm}$ is estimated using 90 arcmin apertures.

4   Emulators for $\boldsymbol{\xi_{\pm}}$ and $\boldsymbol{\zeta_{\pm}}$

The evaluation of the integrated lensing bispectrum $\mathcal{B}^{\mathrm{2D}}_{\pm,ijk}(\boldsymbol{\ell})$ is the key computational bottleneck when evaluating $\zeta_{\pm}$ using eqs. (2.9) and (2.10), and thus the quantity that we wish to emulate. However, rather than emulating $\mathcal{B}^{\mathrm{2D}}_{\pm,ijk}(\boldsymbol{\ell})$ directly, we emulate only the part of the integrand in eq. (2.11) given by

$$\int \frac{\mathrm{d}^2 \boldsymbol{\ell}_1}{(2\pi)^2} \int \frac{\mathrm{d}^2 \boldsymbol{\ell}_2}{(2\pi)^2} B^{\mathrm{3D}}_{\delta} \left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2}{\chi}, \chi\right) \times e^{2i(\phi_2 \mp \phi_{-1-2})} U(\boldsymbol{\ell}_1) W(\boldsymbol{\ell}_2 + \boldsymbol{\ell}) W(-\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2 - \boldsymbol{\ell}).$$
(4.1)

This leaves out the part involving the line-of-sight integration in eq. (2.11), but has the advantage of allowing for more flexibility to adjust the source redshift distributions, including bypassing the need to emulate any of the systematic parameters mentioned in section 2.3. The training of the emulator still needs to be redone for different sizes of the U and W filters. The direct evaluation of $\xi_{\pm}$ in an MCMC exploration of the parameter space would not impose a serious computational burden, but we emulate its calculation anyway for extra speed. In this case we emulate simply the three-dimensional matter power spectrum P^{3D}_{δ} in eq. (2.15).

We build our emulator by training a neural network (NN) on a Latin hypercube with 10^5 training nodes. The emulated parameters comprise the cosmological parameters $\{\Omega_{\mathrm{m}}, A_s, w_0\}$, the baryonic feedback parameter c_{min}, as well as the redshift z which we need to emulate to perform the line-of-sight integrations in eqs. (2.11) and (2.15). The ranges of the cosmological and baryonic parameters are listed in table 1 (note we rescale A_s to $\ln(10^{10}A_s)$), and for redshift we consider $z \in [0, 2.1]$. The NN architecture is that of the `Cosmopower` code [69],[4]

---

[4]https://alessiospuriomancini.github.io/cosmopower/.

	Prior range
Cosmological parameters (emulated)	
Ω_{m}	$U\,[0.16, 0.45]$
$\ln(10^{10}A_s)$	$U\,[1.61, 4.20]$
w_0	$U\,[-3.33, -0.33]$
Baryonic feedback parameter (emulated)	
$c_{\min}$	$U\,[1.0, 5.5]$
Systematic parameters (not emulated)	
Δz_1	$\mathcal{N}(0.0, 0.023)$
Δz_2	$\mathcal{N}(0.0, 0.020)$
m_1	$\mathcal{N}(0.0261, 0.012)$
m_2	$\mathcal{N}(-0.061, 0.011)$
$A_{\mathrm{IA},0}$	$U\,[-5.0, 5.0]$
α_{IA}	0 (fixed)

Table 1. Model parameters considered in this paper. The parameters that enter our NN emulator are the cosmological parameters Ω_{m}, $\ln(10^{10}A_s)$, w_0, and the baryonic feedback parameter $c_{\min}$. The photo-z, shear calibration and IA systematic parameters do not need to be emulated because the predictions for different values are fast to obtain. In our MCMC analyses we vary these parameters within the listed uniform prior ranges (U) or assuming Gaussian priors $\mathcal{N}(\mu, \sigma)$ with mean μ and standard deviation σ. The listed priors for the systematic parameters are inspired by those assumed in the DES Y3 analyses [54, 62].

which was originally developed to emulate 2-point statistics, but which can be straightforwardly applied to emulate eq. (4.1). The input layers of the NN are the cosmological, baryonic and redshift parameters. For $\zeta_\pm$, the output of the NN is the quantity in eq. (4.1) in 100 log-spaced ℓ bins between $\ell = 2$ and $\ell = 15000$. In the training set, the supervised learning labels are the same quantity obtained by directly evaluating eq. (4.1) using Monte-Carlo integration. For $\xi_\pm$ the output is the three-dimensional matter power spectrum in 100 log-spaced ℓ bins as in the right-hand side of eq. (2.15) between $\ell = 2$ and $\ell = 15000$.

We test the emulators using another Latin hypercube with 10^3 test nodes with the same prior ranges of the training set. We quantify the performance of the emulator with the expression

$$\epsilon \equiv \left| \frac{\chi^2_{\mathrm{emu},i}}{\chi^2_{\mathrm{test},i}} - 1 \right|, \tag{4.2}$$

where $\chi^2_{\mathrm{emu},i}$ is the χ^2 value associated with the ith test node, defined w.r.t. the data vector $\hat{d}_{\mathrm{T}17}$ generated by the theory model at T17 cosmological parameters. Concretely, $\chi^2_{\mathrm{emu},i} = \left(\hat{d}_{\mathrm{emu},i} - \hat{d}_{\mathrm{T}17}\right)^t \hat{C}^{-1} \left(\hat{d}_{\mathrm{emu},i} - \hat{d}_{\mathrm{T}17}\right)$, with $\hat{d}_{\mathrm{emu},i}$ the emulator prediction and $\hat{C}^{-1}$ the T17 inverse covariance matrix. The quantity $\chi^2_{\mathrm{test},i}$ is defined analogously, but replacing the emulator result at each test node with the test label prediction. The ϵ metric describes how similar the emulator would behave to the theory model in likelihood analyses. The smaller the value of ϵ, the better the accuracy of the emulator.

Figures 5 and 6 show the outcome of this test for $\xi_\pm$ and $\zeta_\pm$, respectively. We show ϵ projected only on the Ω_{m} — w_0 plane, but the takeaways are common to other projections. For $\xi_\pm$, effectively all of the test nodes have χ^2 relative differences $\epsilon < 0.05$. The performance
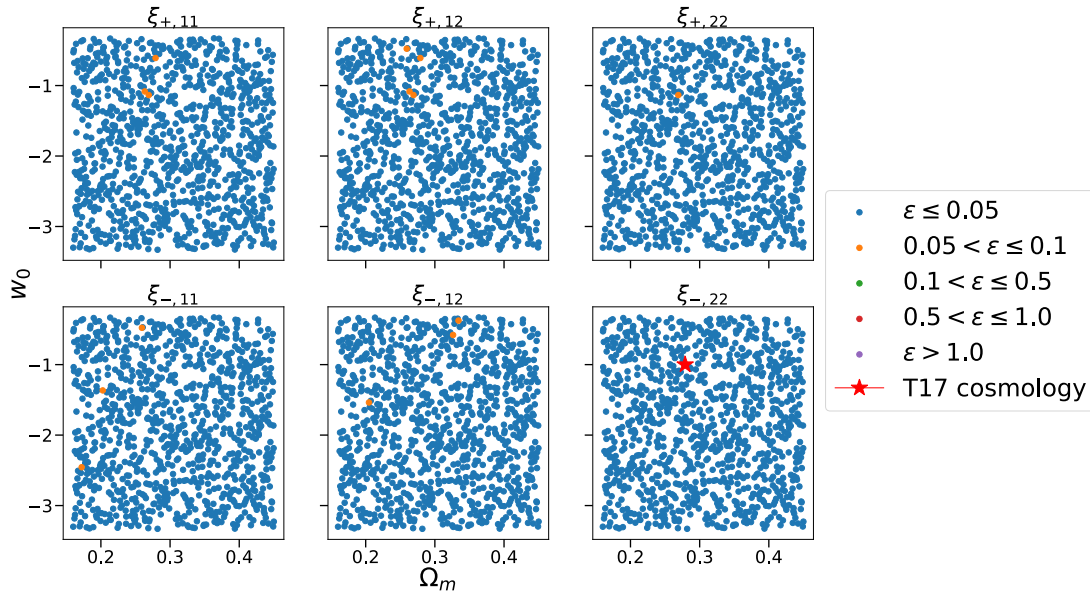
Figure 5. Performance of the $\xi_\pm$ emulator on 10^3 test nodes projected on the Ω_m — w_0 plane. The colors show the absolute value of the χ^2 relative difference ϵ defined in eq. (4.2); if $\epsilon < 0.05$, this means the emulator describes the χ^2 w.r.t. the T17 cosmology (marked by the red star) to better than 5%.
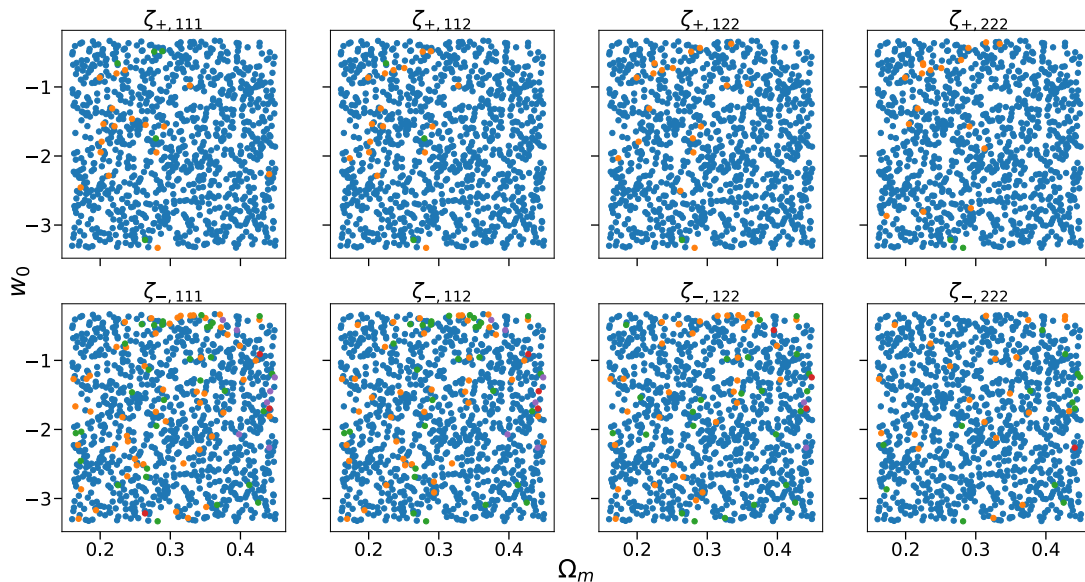
Figure 6. Same as figure 5, but for the performance of the $\zeta_\pm$ emulator, instead of $\xi_\pm$. The result is for 90 arcmin apertures. The color coding is the same as in figure 5.

gets reduced slightly for $\zeta_\pm$ with 92% (95%) of the test nodes having $\epsilon < 0.05$ ($\epsilon < 0.1$); the result in figure 6 is for apertures with 90 arcmin, but we have checked the performance is equivalent for other apertures as well. If the true χ^2 value of some point in parameter space is $\chi^2_\mathrm{test} = 1$, then $\epsilon < 0.1$ implies $\chi^2_\mathrm{emu} \in [0.9, 1.1]$. Effectively all of the test nodes for both $\xi_\pm$ and $\zeta_\pm$ satisfy this satisfactory criterion.

5 Results: simulated likelihood analyses with MCMC

In this section we present our main numerical results from simulated likelihood analyses with MCMC. Unless otherwise specified, we consider the parameter priors listed in table 1, and sample the parameter space assuming a Gaussian likelihood function,

$$\mathcal{L}(\boldsymbol{\theta}) \propto \exp\left[-\frac{1}{2}\left(\mu(\boldsymbol{\theta}) - \hat{d}\right)^t C^{-1}\left(\mu(\boldsymbol{\theta}) - \hat{d}\right)\right], \tag{5.1}$$

where $\hat{d}$ is the assumed data vector, C the covariance matrix and $\mu(\boldsymbol{\theta})$ the theory prediction for model parameters $\boldsymbol{\theta}$. We utilize the sampler code `affine`[5] based on `tensorflow`. With the available NVIDIA A100 GPU (Graphics Processing Unit) hardware, emulator and sampler, we are able to sample an order of 10^6 points in an hour's timescale.

Next, we validate our model using the T17 $\xi_\pm$ and $\zeta_\pm$ data vectors in section 5.1, investigate the impact of the aperture size in $\zeta_\pm$ constraints in section 5.2, discuss the impact of the systematic parameters in section 5.3, and check the impact from using the T17 or FLASK covariance matrices in section 5.4. All of the marginalized two-dimensional constraints shown throughout display contours with the 1σ and 2σ confidence regions.

5.1 Validation on the T17 cosmic shear maps

Figure 7 shows the constraints on the cosmological and baryonic feedback parameters for the data vector from the T17 shear maps (cf. black points in figures 2 and 3) and the FLASK covariance matrix. The result is for $\zeta_\pm$ measured using 90 arcmin apertures. We keep the systematic parameters fixed to zero in these constraints, which is the case for our T17 maps. In addition to the correction factors in eq. (3.3), in this section we consider also the factor $[1 + B(N_{\rm d} - N_{\rm p})]^{-1}$ from ref. [70] due to statistical noise in our covariance matrix estimate.

The key takeaway from figure 7 is that our theory model and emulator recover unbiased constraints: the T17 parameters (dashed black lines) are contained well within the 1σ confidence levels for both the $\xi_\pm$-only (green) and $\xi_\pm + \zeta_\pm$ constraints (red). The ability of our theory model to recover unbiased cosmological constraints could have already been anticipated from the good agreement between theory and simulations in figures 2 and 3.

As a test, we have repeated the analysis in figure 7 but adopting the t-distribution likelihood function from ref. [71], instead of a Gaussian likelihood. The result (not shown) is practically indistinguishable from that in figure 7 for both $\xi_\pm$ and $\xi_\pm + \zeta_{\pm,\{90'\}}$, suggesting the exact choice of the likelihood function does not critically affect our results.

5.2 The impact of the aperture size

When measuring $\zeta_\pm$ one of the decisions concerns the choice of the apertures on which to measure the 1-point shear aperture mass and local $\xi_\pm$. To investigate the impact of this, we perform likelihood analyses with a noiseless data vector generated with the theory model using the T17 parameters. In these tests, we use the FLASK covariance, and vary also the systematic parameters with the priors listed in table 1. The main result is shown in table 2, which lists the relative improvement of the combined $\xi_\pm + \zeta_\pm$ constraints relative to $\xi_\pm$-only, for different aperture sizes and combinations. Figure 8 shows the actual parameter constraints for two aperture choices: a single aperture with 90 arcmin (blue) and the combination of five apertures with sizes $\{50, 70, 90, 110, 130\}$ arcmin (red).
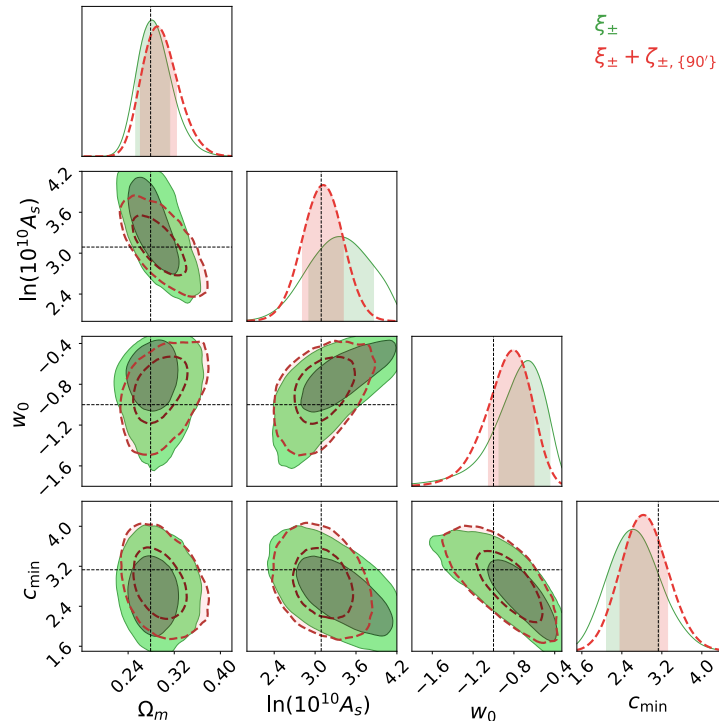

Figure 7. Parameter constraints obtained with the T17 data vector and FLASK covariance matrix. The constraints in green and red are for the $\xi_\pm$-only and $\xi_\pm + \zeta_\pm$ data vectors, respectively. The black dashed lines mark the T17 parameters.

Aperture sizes (arcmin)	Ω_{m}	$\ln\left(10^{10}A_s\right)$	w_0	$c_{\min}$
50	1.2%	9.0%	18.1%	4.8%
70	1.2%	16.9%	31.9%	11.6%
90	**3.7**%	**20.2**%	**38.4**%	**15.1**%
110	1.2%	19.1%	34.1%	11.0%
130	1.2%	16.9%	32.6%	12.3%
$\{50, 70, 90\}$	2.5%	24.7%	39.1%	15.8%
$\{50, 90, 130\}$	3.7%	23.6%	41.3%	16.4%
$\{70, 90, 110\}$	6.2%	25.8%	39.1%	15.1%
$\{90, 110, 130\}$	8.6%	25.9%	42.8%	15.8%
$\{50, 70, 90, 110, 130\}$	**12.4**%	**28.1**%	**44.9**%	**19.9**%

Table 2. Relative improvement of combined $\xi_\pm + \zeta_\pm$ constraints relative to $\xi_\pm$-only for different values and combinations of the aperture sizes. The best single- and combined-filter cases are highlighted in bold. The result is for a noiseless data vector from the theory model, the FLASK covariance, and marginalizing over the systematic parameters.
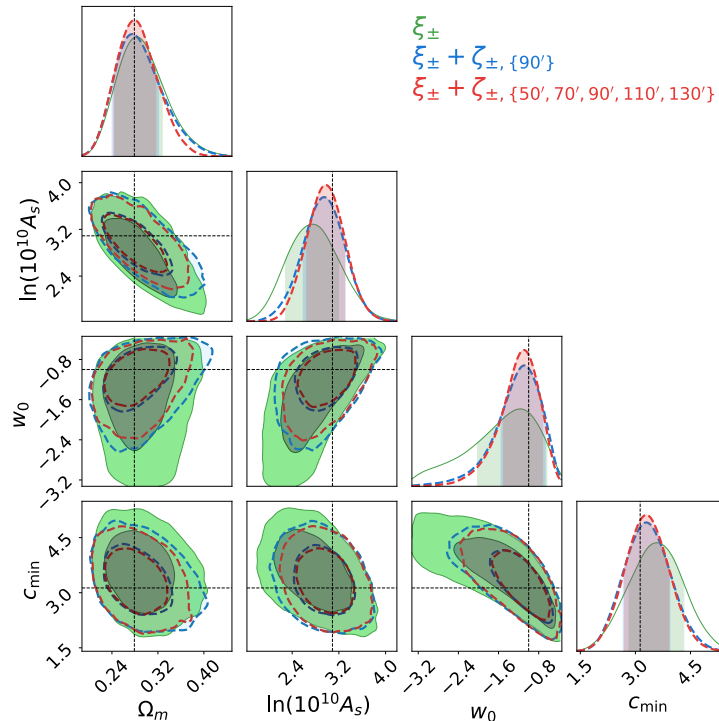
Figure 8. Impact of the aperture size in $\xi_\pm + \zeta_\pm$ constraints. The contours in green are for $\xi_\pm$-only constraints. The contours in blue are for $\xi_\pm + \zeta_\pm$ constraints using a single aperture with size 90 arcmin, and in red for the combination of five filter sizes $\{50, 70, 90, 110, 130\}$ arcmin. The result is for a noiseless data vector from the theory model with the T17 parameters (dashed lines), the FLASK covariance, and marginalizing over the systematic parameters.

Regarding the single aperture cases, table 2 shows that the constraints improve first from 50 to 90 arcmin, but then degrade from 90 to 130 arcmin. This follows from the combination of the following effects. Smaller apertures have the advantage of providing $\zeta_\pm$ with higher signal-to-noise ratio since there are more apertures over which the average of eq. (2.1) can be taken. They have, however, the disadvantage that the local $\xi_\pm$ is measured over a more reduced range of angular scales inside each patch. Conversely, bigger apertures allow to probe the local $\xi_\pm$ on larger scales, but at the price of less signal-to-noise as one averages over a smaller number of patches on the sky.[6] In general, different aperture sizes are sensitive to different configurations of the small-scale squeezed-limit bispectrum [38], which can contain varying cosmological information and impact the final parameter constraints.

For the aperture sizes shown, the balance between these effects is optimal for apertures with 90 arcmin, which gives the best constraints. Concretely, the addition of $\zeta_\pm$ to the constraints leads to improvements of 4% for Ω_m, 20% for $\ln\left(10^{10}A_s\right)$, 38% for w_0 and 15% for c_min. These figures are in line with the previous findings of refs. [32, 38] based on idealized Fisher-matrix forecasts, but extended here to more realistic MCMC-based analyses.

The $\zeta_\pm$ measured over slightly different aperture sizes are expected to be substantially correlated due to the large overlap of the regions where the local $\xi_\pm$ is measured. However, the

---

[6]In particular, in the limit of very large apertures, the $\xi_\pm$ measured in the patches become almost perfectly correlated with the $\xi_\pm$ of the whole survey, effectively contributing with no independent information.
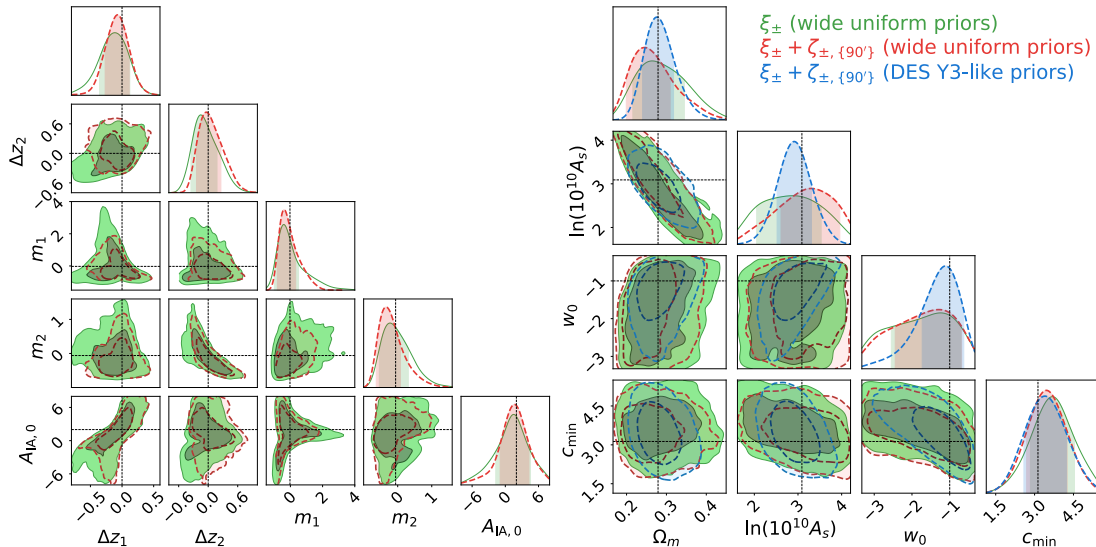
Figure 9. Parameter constraints for different priors on the systematic parameters. The green and red contours are for $\xi_\pm$ and $\xi_\pm + \zeta_\pm$ assuming wide uniform priors on the systematic parameters. The blue contours are for $\xi_\pm + \zeta_\pm$ with DES Y3-like Gaussian priors on the systematic parameters. The left panel shows the systematic parameter constraints, and the right panel shows the constraints on the cosmological and baryonic feedback parameters. The left panels do not show the contours with DES Y3-like priors as they are too small to be clearly seen. The result is for a noiseless data vector drawn from the theory model with the T17 parameters (except we set $A_{\mathrm{IA},0} = 2$; cf. dashed lines), and using the FLASK covariance.

lower part of table 2 shows that there is still enough independent information to improve the constraints further by combining different apertures. For the cases shown, the best constraints are obtained when combining all apertures $\{50, 70, 90, 110, 130\}$ arcmin: the improvements become 12% for Ω_m, 28% for $\ln\left(10^{10} A_s\right)$, 45% for w_0 and 20% for c_min. These improvements need however to be contrasted with the complications that they add to the analyses. For example, this comes with the price of a much larger data vector, which puts pressure on the numerical requirements for reliable covariance estimates from simulations. In this paper, this pressure was still manageable for a DES Y3-like survey with two tomographic bins, but future survey analysis settings will have larger areas and more source redshift bins as well. The decision of how many filters to combine should thus be made case by case.

5.3 The impact of systematics and their modelling

We turn our attention now to the impact of systematics (photo-z, shear calibration and IA) in $\zeta_\pm$ constraints. This is interesting as $\xi_\pm$ and $\zeta_\pm$ depend differently on systematics, and so combined analyses can potentially mitigate the degradation caused by these additional free parameters, leading to better cosmological constraints [73–75]. Indeed, this has been studied recently in ref. [56], where it was shown that combining lensing 2- and 3-point correlation function information in a survey like Euclid could lead even to the self-calibration of the systematic parameters to levels that reduce the need for external calibration data sets.

The green and red contours in figure 9 show the constraints for $\xi_\pm$ and $\xi_\pm + \zeta_\pm$, but instead of the tight DES Y3-like priors that we have assumed so far for the systematic parameters
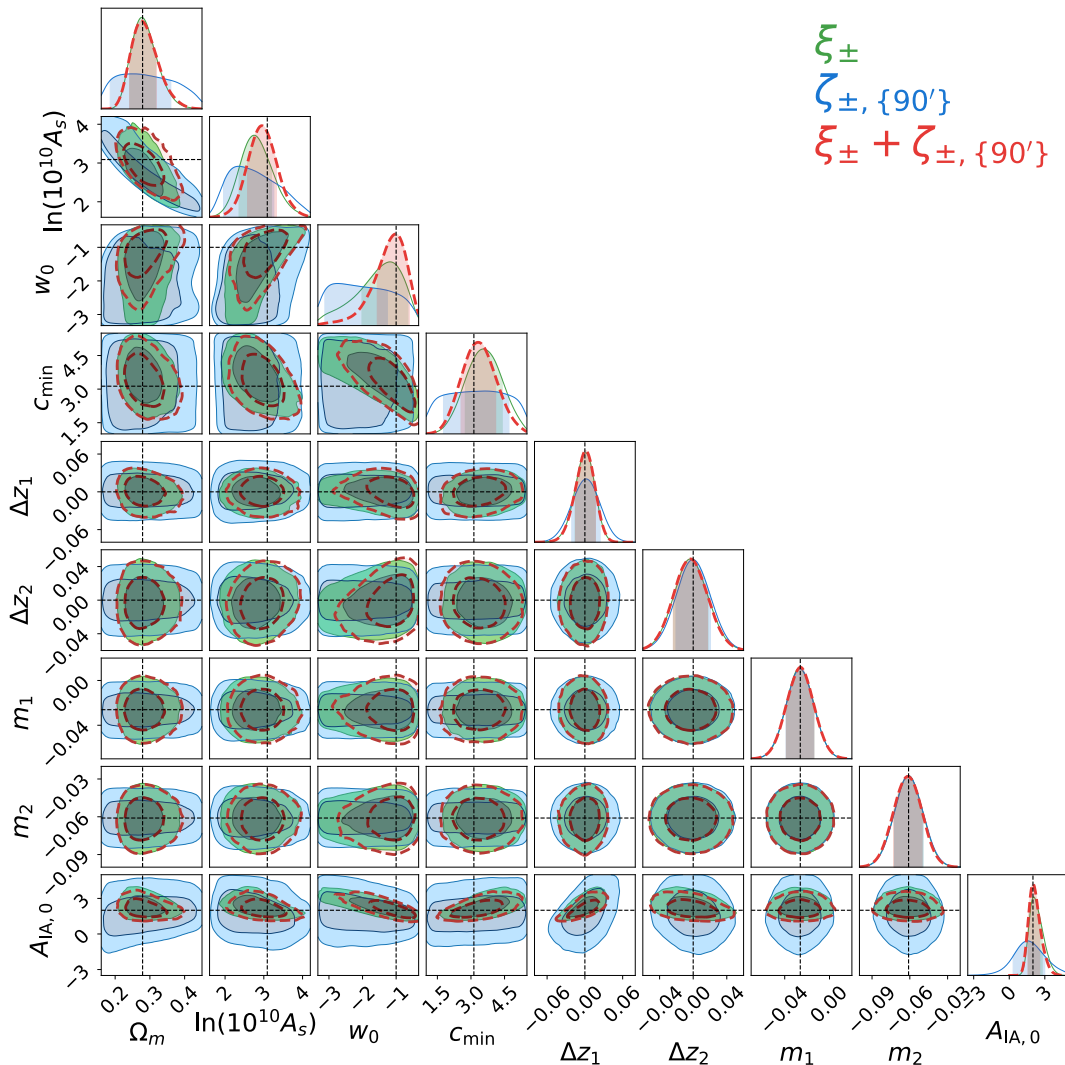
Figure 10. Parameter constraints obtained with our default NLA IA modelling, but on a data vector generated with the NLA IA model used in refs. [54, 56, 72]. The result is for the T17 parameters (except we set $A_{\rm IA,0} = 2$; cf. dashed lines) and the FLASK covariance. The green, blue and red contours are for $\xi_\pm$, $\zeta_\pm$ and $\xi_\pm + \zeta_\pm$ constraints. We use DES Y3-like Gaussian priors for the systematic parameters.

(cf. table 1), we assume now wide priors for them. The result is for a noiseless realization of the data vector for the T17 parameters, with the exception that we set $A_{\rm IA,0} = 2$ in this subsection. The improvements on the cosmological and baryonic parameters from adding $\zeta_\pm$ are 15.4% for $\Omega_{\rm m}$, 8.8% for $\ln(10^{10}A_s)$, 4.9% for w_0 and 8.8% for $c_{\rm min}$. Compared to the case where we marginalize over tight DES Y3-like Gaussian priors, varying the systematic parameters over wide priors degrades the improvement by factors of 2.3, 7.8 and 1.7 for $\ln(10^{10}A_s)$, w_0 and $c_{\rm min}$ respectively. Furthermore, contrary to the case in ref. [56], the improvements that still exist do not appear to be associated with a significant self-calibration of the systematic parameters. This can be seen also on the left of figure 9, where the constraints on the systematic parameters in the combined $\xi_\pm + \zeta_\pm$ case (red) show improvements of 21% for

Δz_1, 8% for Δz_2, 24% for m_1, 17% for m_2 and 18% for $A_{\mathrm{IA},0}$. There is indeed a visible level of systematics self-calibration from combining $\xi_\pm$ with $\zeta_\pm$, but which still yields constraints that are substantially larger than using the externally calibrated DES Y3-like priors (blue).

The quantitative differences to the analysis of ref. [56] could be at least partly due to some of the following reasons. First, ref. [56] considers 3-point correlation function information by taking the equilateral lensing bispectrum as the data, whereas we consider $\zeta_\pm$ that probes predominantly the squeezed lensing bispectrum [32, 38]. Second, ref. [56] considers a treatment of the NLA IA model that is not the same as ours (cf. appendix A). Further, the results of ref. [56] are based on Fisher matrix analyses, whereas ours are for simulated likelihood analyses with MCMC sampling. This can be especially important given how strongly non-Gaussian the marginalized posteriors of the systematic parameters are on the left of figure 9. Finally, our analysis is for a DES Y3-like survey assuming two tomographic bins, whereas ref. [56] considers a larger Euclid-like survey with five tomographic bins, and thus a higher-dimensional subspace of systematic parameters. A deep investigation of the origin of the differences between the results of the two works would be interesting to pursue, but that is beyond the scope of the present paper.

We investigate also potential biases in the constraints of the $A_{\mathrm{IA},0}$ parameter from assuming different IA models in shear 3-point correlation function analyses. In particular, we wish to contrast the NLA model used in this paper (cf. section 2.3 and appendix A) with that in ref. [54] which comes from refs. [56, 72]. To do so we generate a noiseless data vector with the T17 parameters and $A_{\mathrm{IA},0} = 2$ *assuming the IA parametrization of ref. [54]*, which we subsequently analyse by running MCMC constraints *assuming our IA modelling strategy*. At the $\xi_\pm$ level, the two IA treatments are equivalent, but there are differences at the level of the 3-point correlation functions (cf. appendix A).[7] Figure 10 shows the corresponding constraints for $\xi_\pm$ (green), $\zeta_\pm$ (blue) and $\xi_\pm + \zeta_\pm$ (red), with all yielding unbiased constraints, including $A_{\mathrm{IA},0}$. That is, at the level of the constraining power of our DES Y3-like setup, the differences between the two NLA IA models do not have any significant impact. We note, however, that whether the same conclusion holds for other survey setups should be checked on a case-by-case basis.

5.4 The impact of different covariance estimates

We compare in figure 11 the parameter constraints obtained with the T17 covariance matrix (left) with those obtained using FLASK (right). In order to make a fair comparison, in this subsection we constructed a new FLASK covariance with the same number of footprint realizations as T17 ($N_s = 540$), and with a source redshift distribution matching the discretized one of the T17 simulations described in section 3.1. The result in figure 11 is for a noiseless realization of the data vector from the theory model at the T17 parameters, and with the systematic parameters marginalized with the DES Y3-like Gaussian priors. Table 3 lists the corresponding improvements from adding $\zeta_\pm$ information to the constraints.

The two covariance matrices yield effectively the same parameter posteriors for $\xi_\pm$-only constraints; cf. similarity between the green contours on the left and right of figure 11. There are however some differences in the combined $\xi_\pm + \zeta_\pm$ constraints shown in red, with the FLASK covariance yielding smaller parameter error bars for most parameters. In particular, the improvements from $\zeta_\pm$ can be factors of $\approx 1.2 - 1.9$ larger with the FLASK covariance compared to T17.

---

[7]Among other, the model of ref. [54] includes terms $\propto A_{\mathrm{IA},0}^4$, whereas ours stops at third order $\propto A_{\mathrm{IA},0}^3$, as expected for a three-point correlation function.
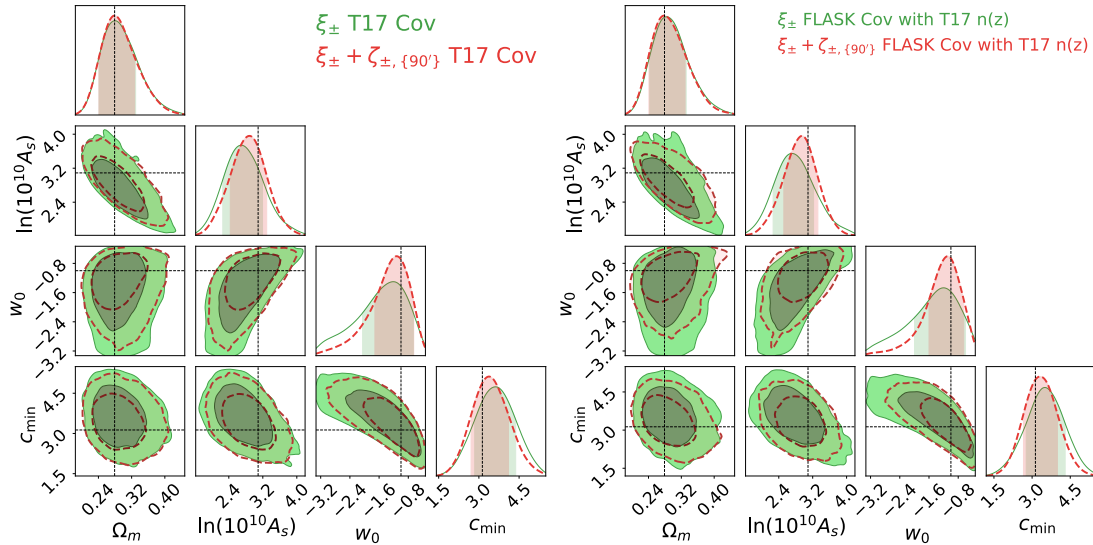
Figure 11. Impact of the covariance matrix estimate on the parameter constraints. The left and right panels show the result for the covariance estimated from the T17 and FLASK shear maps, respectively. In both panels, the result is for a noiseless realization of the data vector from the theory model at the T17 parameters (dashed lines); the green and red contours are for $\xi_\pm$-only and $\xi_\pm + \zeta_\pm$, respectively. The result is for apertures with sizes 90 arcmin and systematics marginalized with the DES Y3-like Gaussian priors.

Covariance type	$\Omega_{\rm m}$	$\ln\left(10^{10}A_s\right)$	w_0	$c_{\rm min}$
FLASK (lognormal)	1.1%	16.7%	32.1%	12.4%
T17 (N-body simulations)	3.5%	8.8%	26.1%	8.7%

Table 3. Impact of the covariance matrix estimate on the improvement of $\xi_\pm + \zeta_\pm$ constraints, relative to $\xi_\pm$-only. The result is for the aperture with size 90 arcmin.

The T17 and FLASK covariances in this subsection are estimated from ensembles of 540 shear maps and so have the same noise level. These differences may indicate they are intrinsic to the different ability of N-body simulations and lognormal realizations to capture the covariance of $\zeta_\pm$,[8] or due to residual statistical fluctuations for $N_s = 540$. We leave a more detailed investigation of the impact of the covariance matrix, including covariances calculated analytically [13, 68], to future work.

6 Summary & conclusion

The integrated shear 3PCF $\zeta_\pm$ [32, 38] is a higher-order cosmic shear statistic that measures the correlation between the shear 2PCF measured in patches of the sky and the shear aperture mass in the same patches (cf. eq. (2.1)). On small scales, $\zeta_\pm$ probes primarily the cosmological

---

[8]The covariance of a 3-point function contains terms up to the 6-point function, which are not as faithfully captured in lognormal realizations, compared to N-body simulations.

information encoded in the squeezed-limit lensing bispectrum. Two of the key advantages of $\zeta_\pm$ compared to other higher-order cosmic shear statistics are that (i) it can be straightforwardly evaluated from the data using efficient and well-tested 2-point correlation function estimators (i.e. it does not explicitly require dedicated and more expensive 3-point estimators) and (ii) it admits a theoretical model based on the response approach to perturbation theory [38] that is accurate in the nonlinear regime of structure formation, allowing to reliably account for the impact of baryonic physics.

In this paper, we developed an analysis pipeline that can be directly applied to real cosmic shear data to obtain cosmological constraints from $\zeta_\pm$ and its combination with $\xi_\pm$. Compared to previous works on $\zeta_\pm$, the main significant advances in this paper are (i) the incorporation of lensing systematics associated with photo-z uncertainties, shear calibration biases and galaxy IA (cf. section 2.3), and (ii) the development of a NN-based emulator for fast theory predictions to enable MCMC parameter inference. We tested our pipeline on a set of realistic cosmic shear maps based on N-body simulations, with DES Y3-like survey footprint, mask and source redshift distributions (cf. section 3).

In our tests of the analysis pipeline we have investigated in particular (i) the accuracy of the theory model (cf. section 5.1), (ii) the impact of the size of the apertures used to measure $\zeta_\pm$ (cf. section 5.2), (iii) the impact of lensing systematics (cf. section 5.3) and (iv) the impact of N-body simulation vs. lognormal estimates of the data vector covariance matrix (cf. section 5.4). Our main findings can be summarized as follows:

- Our analysis pipeline is accurate (cf. figures 2 and 3) and able to yield unbiased parameter constraints from our N-body simulation DES Y3-like data vectors (cf. figure 7).

- For the range of aperture sizes $\{50, 70, 90, 110, 130\}$ arcmin, 90 arcmin is what results in the largest information gain from $\zeta_\pm$. The combination of several filter sizes can improve the constraints further (cf. table 2), but at the cost of dealing with a larger data vector and covariance matrix.

- Although $\xi_\pm$ and $\zeta_\pm$ depend differently on the systematic parameters, we do not find significant improvements in their constraints in combined $\xi_\pm + \zeta_\pm$ analyses; i.e. the mitigation of systematic effects still requires prior calibration from external data (cf. figure 9). This is in contrast with the findings in ref. [56], although this may be due to differences in the 3-point correlation function studied, survey setup and other analysis details. At the level of the DES Y3 constraining power, different modelling strategies for IA lead also to no significant biases in parameter constraints (cf. figure 10).

- Relative to $\xi_\pm$-only constraints with the N-body covariance matrix, adding $\zeta_\pm$ leads to improvements of 4% for $\Omega_{\rm m}$, 9% for $\ln\left(10^{10}A_s\right)$, 26% for w_0 and 9% for $c_{\rm min}$. Except for $\Omega_{\rm m}$, these are factors of $\approx 1.2 - 1.9$ smaller compared to the FLASK covariance. This may be due to residual statistical fluctuations at the level of our number of simulation realizations ($N_s = 540$), or simply that lognormal realizations do not provide reliable estimates of the $\zeta_\pm$ covariance matrix.

Overall, our results corroborate with a realistic MCMC-based simulated likelihood analysis the encouraging findings from previous idealized Fisher matrix forecasts [32, 38]. The analysis pipeline developed and tested here can be readily applied to real survey data, enabling the exploration of the potential of the integrated shear 3PCF $\zeta_\pm$ to improve cosmological parameter constraints using cosmic shear observations.

Acknowledgments

Data availability. The numerical data underlying the analysis of this paper may be shared upon reasonable request to the authors.

A The modelling of intrinsic alignments

In this appendix we describe our modelling of galaxy intrinsic alignments in $\xi_\pm$ and $\zeta_\pm$.

General considerations

The observed galaxy ellipticity in cosmic shear observations $\epsilon_{\rm obs}$ is a combination of the gravitational (G) lensing shear component γ and the intrinsic (I) ellipticity of the galaxies $\epsilon_{\rm I}$ induced by correlations with local gravitational tidal fields at the source (in this appendix, we ignore the random stochastic component that would contribute as shape noise):

$$\epsilon_{\rm obs}^i(\boldsymbol{\theta}) = \gamma^i(\boldsymbol{\theta}) + \epsilon_{\rm I}^i(\boldsymbol{\theta}), \tag{A.1}$$

where i denotes a specific source galaxy redshift bin. The lensing shear is related to the lensing convergence κ as [2]

$$\gamma(\boldsymbol{\ell}) = e^{2i\phi_\ell}\kappa(\boldsymbol{\ell})\,;\quad \kappa(\boldsymbol{\ell}) = \int {\rm d}^2\boldsymbol{\theta}\ \kappa(\boldsymbol{\theta})e^{-i\boldsymbol{\ell\theta}}\,;\quad \kappa(\boldsymbol{\theta}) = \int {\rm d}\chi\ q(\chi)\delta_m(\boldsymbol{\theta}\chi,\chi), \tag{A.2}$$

where δ_m is the three-dimensional matter density contrast.[9] In analogy, we can write for the intrinsic component $\epsilon_{\rm I}$

$$\epsilon_{\rm I}(\boldsymbol{\ell}) = e^{2i\phi_\ell}\kappa_{\rm I}(\boldsymbol{\ell})\,;\quad \kappa_{\rm I}(\boldsymbol{\ell}) = \int {\rm d}^2\boldsymbol{\theta}\ \kappa_{\rm I}(\boldsymbol{\theta})e^{-i\boldsymbol{\ell\theta}}\,;\quad \kappa_{\rm I}(\boldsymbol{\theta}) = \int {\rm d}\chi\ n(\chi)\delta_{\rm I}(\boldsymbol{\theta}\chi,\chi), \tag{A.3}$$

where $\delta_{\rm I}$ is a three-dimensional field that determines *effectively* the intrinsic alignment (IA) of the galaxies with their local gravitational tidal fields; note also that the line-of-sight kernel is now just the source galaxy distribution $n(\chi)$, and not the lensing kernel $q(\chi)$.

In the popular *nonlinear linear alignment* (NLA) model [51, 52], one writes

$$\delta_{\rm I}(\boldsymbol{x}, z) = f_{\rm IA}(z)\delta_m(\boldsymbol{x}, z), \tag{A.4}$$

---

[9]To ease the notation, we distinguish between real- and harmonic-space variables by their arguments. For example, $\kappa(\boldsymbol{\theta})$ and $\kappa(\boldsymbol{\ell})$ are the lensing convergence in real and harmonic space, respectively.

treating δ_m as the nonlinear matter density contrast. The amplitude $f_{IA}(z)$ is

$$f_{IA}(z) = -A_{IA,0} \left(\frac{1+z}{1+z_0}\right)^{\alpha_{IA}} \frac{c_1 \rho_{crit} \Omega_{m,0}}{D(z)}, \tag{A.5}$$

where $A_{IA,0}$, α_{IA} are free redshift-independent parameters, $c_1 = 5 \times 10^{-14} \, (h^2 M_\odot / \mathrm{Mpc}^3)^{-1}$ [52], ρ_{crit} is the critical cosmic energy density, $D(z)$ is the growth factor normalized to unity today, and z_0 is some reasonable pivot redshift value.

Note that this is only an *effective* parametrization of the impact of IA in cosmic shear observations. A more rigorous approach would involve a description of the relation of galaxy shapes and tidal fields in 3D, subsequently projected to the sky plane. This is the approach described in refs. [82, 83] based on bias expansions in effective field theory, which is however valid only in the quasi-linear, large-scale regime of structure formation. Extensions of the NLA model to include nonlinear corrections to eq. (A.4) also exist [84].

Contributions to $\xi_\pm$

The two shear 2PCF $\xi_\pm$ are given by

$$\xi_{+,\mathrm{obs}}^{ij}(\alpha) = \langle \epsilon_{obs}^i(\boldsymbol{\theta}) \epsilon_{obs}^{j*}(\boldsymbol{\theta} + \boldsymbol{\alpha}) \rangle \tag{A.6}$$

$$\xi_{-,\mathrm{obs}}^{ij}(\alpha) = \langle \epsilon_{obs}^i(\boldsymbol{\theta}) \epsilon_{obs}^{j}(\boldsymbol{\theta} + \boldsymbol{\alpha}) e^{-4i\phi_\alpha} \rangle, \tag{A.7}$$

and each can be decomposed into GG, GI, IG and II terms as

$$\xi_{\pm,\mathrm{obs}}^{ij} = \xi_{\pm,\mathrm{GG}}^{ij} + \xi_{\pm,\mathrm{GI}}^{ij} + \xi_{\pm,\mathrm{IG}}^{ij} + \xi_{\pm,\mathrm{II}}^{ij}. \tag{A.8}$$

The GI case of $\xi_{+,\mathrm{obs}}^{ij}$, for example, is given by (the derivations are analogous for all terms):

$$\begin{aligned} \xi_{+,\mathrm{GI}}^{ij}(\alpha) &= \langle \gamma^i(\boldsymbol{\theta}) \epsilon_I^{j*}(\boldsymbol{\theta} + \boldsymbol{\alpha}) \rangle \\ &= \int \frac{d\ell \ell}{2\pi} P_{\kappa \kappa_I}^{ij}(\ell) J_0(\ell \alpha), \end{aligned} \tag{A.9}$$

where $P_{\kappa \kappa_I}^{ij}(\ell)$ is defined as $(2\pi)^2 P_{\kappa \kappa_I}^{ij}(\ell) \delta_D(\boldsymbol{\ell} + \boldsymbol{\ell}') = \langle \kappa^i(\boldsymbol{\ell}) \kappa_I^j(\boldsymbol{\ell}') \rangle$ and given by

$$P_{\kappa \kappa_I}^{ij}(\ell) = \int d\chi \frac{q^i(\chi) n^j(\chi)}{\chi^2} P_{\delta_m \delta_I}^{3D}(\ell/\chi, \chi). \tag{A.10}$$

The $P_{\delta_m \delta_I}^{3D}$ is defined as $(2\pi)^3 P_{\delta_m \delta_I}^{3D}(\boldsymbol{k}_1 + \boldsymbol{k}_2) = \langle \delta_m(\boldsymbol{k}_1) \delta_I(\boldsymbol{k}_2) \rangle$, and in the NLA model it is

$$P_{\delta_m \delta_I}^{3D}(k, z) = f_{IA}(z) P_{\delta_m \delta_m}^{3D}(k, z). \tag{A.11}$$

That is, the GI contribution to $\xi_{+,\mathrm{obs}}^{ij}$ can be obtained by replacing the jth lensing kernel $q^j(\chi)$ in the expression of the GG term with $n^j(\chi) f_{IA}$. It follows as a result that all contributions from GG, GI, IG and II can be obtained by replacing all lensing kernels $q(\chi)$ with $q(\chi) + n(\chi) f_{IA}$, as in eq. (2.24). This yields terms $\propto f_{IA}^0$ (GG), $\propto f_{IA}$ (GI, IG) and $\propto f_{IA}^2$ (II).

Contributions to $\zeta_\pm$

The observed integrated shear 3PCF $\zeta_\pm$ is defined as

$$\zeta_{\pm,\text{obs}}^{ijk}(\alpha) = \left\langle \hat{M}_{\text{ap,obs}}^{i}(\boldsymbol{\theta}_C)\hat{\xi}_{\pm,\text{obs}}^{jk}(\alpha;\boldsymbol{\theta}_C)\right\rangle. \tag{A.12}$$

The position-dependent shear 2PCF $\hat{\xi}_{\pm,\text{obs}}^{jk}(\alpha;\boldsymbol{\theta}_C)$ also contains GG, GI, IG and II terms. Further, the IA terms also contribute to the 1-point aperture mass $\hat{M}_{\text{ap,obs}}^{i}(\boldsymbol{\theta}_C)$, which contains G and I terms as

$$\hat{M}_{\text{ap,obs}}^{i}(\boldsymbol{\theta}_C) = \int \mathrm{d}^2\boldsymbol{\theta}\Big[\kappa^i(\boldsymbol{\theta}) + \kappa_{\text{I}}^i(\boldsymbol{\theta})\Big]U(\boldsymbol{\theta}_C - \boldsymbol{\theta}). \tag{A.13}$$

This thus generates the following 8 contributions to $\zeta_{\pm,\text{obs}}^{ijk}(\alpha)$:

$$\zeta_{\pm,\text{obs}}^{ijk} = \zeta_{\pm,\text{GGG}}^{ijk} + \zeta_{\pm,\text{GGI}}^{ijk} + \zeta_{\pm,\text{GIG}}^{ijk} + \zeta_{\pm,\text{GII}}^{ijk} + \zeta_{\pm,\text{IGG}}^{ijk} + \zeta_{\pm,\text{IGI}}^{ijk} + \zeta_{\pm,\text{IIG}}^{ijk} + \zeta_{\pm,\text{III}}^{ijk}. \tag{A.14}$$

Again, just as a single example, the IIG case for $\zeta_{+,\text{obs}}$ can be written as

$$\zeta_{+,\text{IIG}}^{ijk}(\alpha) = \frac{1}{A(\alpha)} \int \frac{\mathrm{d}\ell\ell}{2\pi} \mathcal{B}_{+,\text{IIG}}^{ijk}(\ell)J_0(\ell\alpha), \tag{A.15}$$

where

$$\mathcal{B}_{+,\text{IIG}}^{ijk}(\ell) = \frac{\mathrm{d}^2\boldsymbol{\ell}_1}{(2\pi)^2} \int \frac{\mathrm{d}^2\boldsymbol{\ell}_2}{(2\pi)^2} B_{\kappa_{\text{I}}\kappa_{\text{I}}\kappa}^{ijk}\left(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,-\boldsymbol{\ell}_{12}\right)e^{2i(\phi_2-\phi_{-1-2})}U(\boldsymbol{\ell}_1)W(\boldsymbol{\ell}+\boldsymbol{\ell}_2)W(-\boldsymbol{\ell}-\boldsymbol{\ell}_{12}), \tag{A.16}$$

with $\boldsymbol{\ell}_{12} = \boldsymbol{\ell}_1 + \boldsymbol{\ell}_2$ and

$$B_{\kappa_{\text{I}}\kappa_{\text{I}}\kappa_{\text{G}}}^{ijk}\left(\boldsymbol{\ell}_a,\boldsymbol{\ell}_b,\boldsymbol{\ell}_c\right) = \int \mathrm{d}\chi \frac{n^i(\chi)n^j(\chi)q^k(\chi)}{\chi^4}B_{\delta_{\text{I}}\delta_{\text{I}}\delta_m}^{\text{3D}}\left(\frac{\boldsymbol{\ell}_a}{\chi},\frac{\boldsymbol{\ell}_b}{\chi},\frac{\boldsymbol{\ell}_c}{\chi};\chi\right). \tag{A.17}$$

The derivation of these expressions is the same as the usual gravitational lensing GGG expression, except one replaces the first two instances of κ by κ_{I}. In the NLA model, $B_{\delta_{\text{I}}\delta_{\text{I}}\delta_m}^{\text{3D}} = f_{\text{IA}}^2 B_{\delta_m\delta_m\delta_m}^{\text{3D}}$. That is, the IIG contribution to $\zeta_{+,\text{obs}}^{ijk}(\alpha)$ term can be obtained from GGG by simply replacing the ith and jth lensing kernels $q^i(\chi)$, $q^j(\chi)$ with $n^i(\chi)f_{\text{IA}}$ and $n^j(\chi)f_{\text{IA}}$. It follows as a result that all of the 8 contributions to $\zeta_{\pm,\text{obs}}^{ijk}(\alpha)$ can be obtained by replacing all lensing kernels $q(\chi)$ with $q(\chi) + n(\chi)f_{\text{IA}}$, as in eq. (2.24). This yields terms $\propto f_{\text{IA}}^0$ (GGG), $\propto f_{\text{IA}}$ (GGI, GIG, IGG) and $\propto f_{\text{IA}}^2$ (GII, IGI, IIG) and $\propto f_{\text{IA}}^3$ (III).

These 3-point contributions from galaxy IA are different than those derived in ref. [56] using also the NLA model. Among other differences, their III term is $\propto f_{\text{IA}}^4$ and their GII + IGI + IIG terms are $\propto f_{\text{IA}}^3$ (cf. their eqs. (30–32)). Reference [56] does not provide a detailed derivation of their expressions, which keeps us from inspecting this issue further. We emphasise, however, that the NLA model is in itself only an approximation of the effect of galaxy IA on small-scales, and so even our expressions should be interpreted in light of this.

References

[1] M. Bartelmann and P. Schneider, *Weak gravitational lensing*, *Phys. Rept.* **340** (2001) 291 [astro-ph/9912508] [InSPIRE].

[2] C.S. Kochanek, *The Saas Fee Lectures on strong gravitational lensing*, in the proceedings of the *33rd Advanced Saas Fee Course on Gravitational Lensing: Strong, Weak, and Micro*, Les Diablerets, Switzerland, April 7–12, 2003, [astro-ph/0407232] [InSPIRE].

[3] DES collaboration, *Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing*, *Phys. Rev. D* **105** (2022) 023520 [`arXiv:2105.13549`] [ɪɴSPIRE].

[4] KɪDS collaboration, *KiDS-1000 Cosmology: Cosmic shear constraints and comparison between two point statistics*, *Astron. Astrophys.* **645** (2021) A104 [`arXiv:2007.15633`] [ɪɴSPIRE].

[5] T. Hamana et al., *Cosmological constraints from cosmic shear two-point correlation functions with HSC survey first-year data*, *Publ. Astron. Soc. Jap.* **72** (2020) PublicationsoftheAstronomicalSocietyofJaan,Volume72,Issue1,February2020,16,htts://doiorg/101093/asj/sz138 [`arXiv:1906.06041`] [ɪɴSPIRE].

[6] EUCLID collaboration, *Euclid preparation — VI. Verifying the performance of cosmic shear experiments*, *Astron. Astrophys.* **635** (2020) A139 [*Erratum ibid.* **638** (2020) C2] [`arXiv:1910.10521`] [ɪɴSPIRE].

[7] LSST DARK ENERGY SCIENCE collaboration, *Large Synoptic Survey Telescope: Dark Energy Science Collaboration*, `arXiv:1211.0310` [ɪɴSPIRE].

[8] M. Yamamoto et al., *Weak gravitational lensing shear estimation with metacalibration for the Roman High-Latitude Imaging Survey*, *Mon. Not. Roy. Astron. Soc.* **519** (2023) 4241 [`arXiv:2203.08845`] [ɪɴSPIRE].

[9] P. Schneider and M. Lombardi, *The three-point correlation function of cosmic shear: I. The natural components*, *Astron. Astrophys.* **397** (2003) 809 [`astro-ph/0207454`] [ɪɴSPIRE].

[10] M. Takada and B. Jain, *Cosmological parameters from lensing power spectrum and bispectrum tomography*, *Mon. Not. Roy. Astron. Soc.* **348** (2004) 897 [`astro-ph/0310125`] [ɪɴSPIRE].

[11] P. Schneider, M. Kilbinger and M. Lombardi, *The three-point correlation function of cosmic shear. 2. Relation to the bispectrum of the projected mass density and generalized third-order aperture measures*, *Astron. Astrophys.* **431** (2005) 9 [`astro-ph/0308328`] [ɪɴSPIRE].

[12] S. Dodelson and P. Zhang, *The weak lensing bispectrum*, *Phys. Rev. D* **72** (2005) 083001 [`astro-ph/0501063`] [ɪɴSPIRE].

[13] I. Kayo, M. Takada and B. Jain, *Information content of weak lensing power spectrum and bispectrum: including the non-Gaussian error covariance matrix*, *Mon. Not. Roy. Astron. Soc.* **429** (2013) 344 [`arXiv:1207.6322`] [ɪɴSPIRE].

[14] M. Sato and T. Nishimichi, *Impact of the non-Gaussian covariance of the weak lensing power spectrum and bispectrum on cosmological parameter estimation*, *Phys. Rev. D* **87** (2013) 123538 [`arXiv:1301.3588`] [ɪɴSPIRE].

[15] N. McCullagh, D. Jeong and A.S. Szalay, *Toward accurate modelling of the non-linear matter bispectrum: standard perturbation theory and transients from initial conditions*, *Mon. Not. Roy. Astron. Soc.* **455** (2016) 2945 [`arXiv:1507.07824`] [ɪɴSPIRE].

[16] R. Takahashi et al., *Fitting the nonlinear matter bispectrum by the Halofit approach*, *Astrophys. J.* **895** (2020) 113 [`arXiv:1911.07886`] [ɪɴSPIRE].

[17] E. Semboloni et al., *Weak lensing from space: first cosmological constraints from three-point shear statistics*, *Mon. Not. Roy. Astron. Soc.* **410** (2011) 143 [`arXiv:1005.4941`] [ɪɴSPIRE].

[18] L. Fu et al., *CFHTLenS: Cosmological constraints from a combination of cosmic shear two-point and three-point correlations*, *Mon. Not. Roy. Astron. Soc.* **441** (2014) 2725 [`arXiv:1404.5469`] [ɪɴSPIRE].

[19] A. Barthelemy, S. Codis and F. Bernardeau, *Probability distribution function of the aperture mass field with large deviation theory*, *Mon. Not. Roy. Astron. Soc.* **503** (2021) 5204 [`arXiv:2012.03831`] [ɪɴSPIRE].

[20] DES collaboration, *Dark Energy Survey Year 3 Results: Three-point shear correlations and mass aperture moments*, *Phys. Rev. D* **105** (2022) 103537 [`arXiv:2201.05227`] [ɪɴSPIRE].

[21] S. Heydenreich, L. Linke, P. Burger and P. Schneider, *A roadmap to cosmological parameter analysis with third-order shear statistics — I. Modelling and validation*, *Astron. Astrophys.* **672** (2023) A44 [`arXiv:2208.11686`] [INSPIRE].

[22] DES collaboration, *Cosmology constraints from shear peak statistics in Dark Energy Survey Science Verification data*, *Mon. Not. Roy. Astron. Soc.* **463** (2016) 3653 [`arXiv:1603.05040`] [INSPIRE].

[23] J. Harnois-Déraps et al., *Cosmic shear cosmology beyond two-point statistics: a combined peak count and correlation function analysis of DES-Y1*, *Mon. Not. Roy. Astron. Soc.* **506** (2021) 1623 [`arXiv:2012.02777`] [INSPIRE].

[24] DES collaboration, *Dark energy survey year 3 results: Cosmology with peaks using an emulator approach*, *Mon. Not. Roy. Astron. Soc.* **511** (2022) 2075 [`arXiv:2110.10135`] [INSPIRE].

[25] DES collaboration, *Density Split Statistics: Joint Model of Counts and Lensing in Cells*, *Phys. Rev. D* **98** (2018) 023508 [`arXiv:1710.05162`] [INSPIRE].

[26] DES collaboration, *Density Split Statistics: Cosmological Constraints from Counts and Lensing in Cells in DES Y1 and SDSS Data*, *Phys. Rev. D* **98** (2018) 023507 [`arXiv:1710.05045`] [INSPIRE].

[27] P. Burger et al., *An adapted filter function for density split statistics in weak lensing*, *Astron. Astrophys.* **642** (2020) A161 [`arXiv:2006.10778`] [INSPIRE].

[28] P. Burger, O. Friedrich, J. Harnois-Déraps and P. Schneider, *A revised density split statistic model for general filters*, *Astron. Astrophys.* **661** (2022) A137 [`arXiv:2106.13214`] [INSPIRE].

[29] P.A. Burger et al., *KiDS-1000 cosmology: Constraints from density split statistics*, *Astron. Astrophys.* **669** (2023) A69 [`arXiv:2208.02171`] [INSPIRE].

[30] S. Heydenreich, B. Brück and J. Harnois-Déraps, *Persistent homology in cosmic shear: constraining parameters with topological data analysis*, *Astron. Astrophys.* **648** (2021) A74 [`arXiv:2007.13724`] [INSPIRE].

[31] S. Heydenreich et al., *Persistent homology in cosmic shear — II. A tomographic analysis of DES-Y1*, *Astron. Astrophys.* **667** (2022) A125 [`arXiv:2204.11831`] [INSPIRE].

[32] A. Halder, O. Friedrich, S. Seitz and T.N. Varga, *The integrated three-point correlation function of cosmic shear*, *Mon. Not. Roy. Astron. Soc.* **506** (2021) 2780 [`arXiv:2102.10177`] [INSPIRE].

[33] C.-T. Chiang, C. Wagner, F. Schmidt and E. Komatsu, *Position-dependent power spectrum of the large-scale structure: a novel method to measure the squeezed-limit bispectrum*, *JCAP* **05** (2014) 048 [`arXiv:1403.3411`] [INSPIRE].

[34] C.-T. Chiang et al., *Position-dependent correlation function from the SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 10 CMASS Sample*, *JCAP* **09** (2015) 028 [`arXiv:1504.03322`] [INSPIRE].

[35] D. Munshi and P. Coles, *The Integrated Bispectrum and Beyond*, *JCAP* **02** (2017) 010 [`arXiv:1608.04345`] [INSPIRE].

[36] G. Jung et al., *The integrated angular bispectrum of weak lensing*, *JCAP* **06** (2021) 055 [`arXiv:2102.05521`] [INSPIRE].

[37] D. Munshi et al., *Position-dependent correlation function of weak-lensing convergence*, *Phys. Rev. D* **107** (2023) 043516 [`arXiv:2104.01185`] [INSPIRE].

[38] A. Halder and A. Barreira, *Response approach to the integrated shear 3-point correlation function: the impact of baryonic effects on small scales*, *Mon. Not. Roy. Astron. Soc.* **515** (2022) 4639 [`arXiv:2201.05607`] [INSPIRE].

[39] A. Barreira and F. Schmidt, *Responses in Large-Scale Structure*, *JCAP* **06** (2017) 053 [`arXiv:1703.09212`] [INSPIRE].

[40] P. Schneider, *Detection of (dark) matter concentrations via weak gravitational lensing*, *Mon. Not. Roy. Astron. Soc.* **283** (1996) 837 [`astro-ph/9601039`] [InSPIRE].

[41] R.G. Crittenden, P. Natarajan, U.-L. Pen and T. Theuns, *Discriminating weak lensing from intrinsic spin correlations using the curl-gradient decomposition*, *Astrophys. J.* **568** (2002) 20 [`astro-ph/0012336`] [InSPIRE].

[42] H. Gil-Marín et al., *An improved fitting formula for the dark matter bispectrum*, *JCAP* **02** (2012) 047 [`arXiv:1111.4477`] [InSPIRE].

[43] C. Wagner, F. Schmidt, C.-T. Chiang and E. Komatsu, *Separate Universe Simulations*, *Mon. Not. Roy. Astron. Soc.* **448** (2015) L11 [`arXiv:1409.6294`] [InSPIRE].

[44] A.S. Schmidt, S.D.M. White, F. Schmidt and J. Stücker, *Cosmological N-Body Simulations with a Large-Scale Tidal Field*, *Mon. Not. Roy. Astron. Soc.* **479** (2018) 162 [`arXiv:1803.03274`] [InSPIRE].

[45] A. Mead et al., *An accurate halo model for fitting non-linear cosmological power spectra and baryonic feedback models*, *Mon. Not. Roy. Astron. Soc.* **454** (2015) 1958 [`arXiv:1505.07833`] [InSPIRE].

[46] D. Blas, J. Lesgourgues and T. Tram, *The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes*, *JCAP* **07** (2011) 034 [`arXiv:1104.2933`] [InSPIRE].

[47] A. Barreira et al., *Separate Universe Simulations with IllustrisTNG: baryonic effects on power spectrum responses and higher-order statistics*, *Mon. Not. Roy. Astron. Soc.* **488** (2019) 2079 [`arXiv:1904.02070`] [InSPIRE].

[48] S. Foreman, W. Coulton, F. Villaescusa-Navarro and A. Barreira, *Baryonic effects on the matter bispectrum*, *Mon. Not. Roy. Astron. Soc.* **498** (2020) 2887 [`arXiv:1910.03597`] [InSPIRE].

[49] DES collaboration, *Dark Energy Survey Year 3 results: Cosmology from cosmic shear and robustness to data calibration*, *Phys. Rev. D* **105** (2022) 023514 [`arXiv:2105.13543`] [InSPIRE].

[50] DES collaboration, *Dark Energy Survey Y3 results: blending shear and redshift biases in image simulations*, *Mon. Not. Roy. Astron. Soc.* **509** (2021) 3371 [`arXiv:2012.08567`] [InSPIRE].

[51] C.M. Hirata et al., *Intrinsic galaxy alignments from the 2SLAQ and SDSS surveys: Luminosity and redshift scalings and implications for weak lensing surveys*, *Mon. Not. Roy. Astron. Soc.* **381** (2007) 1197 [`astro-ph/0701671`] [InSPIRE].

[52] S. Bridle and L. King, *Dark energy constraints from cosmic shear power spectra: impact of intrinsic alignments on photometric redshift requirements*, *New J. Phys.* **9** (2007) 444 [`arXiv:0705.0166`] [InSPIRE].

[53] DES collaboration, *Dark Energy Survey Year 1 Results: Multi-Probe Methodology and Simulated Likelihood Analyses*, `arXiv:1706.09359` [InSPIRE].

[54] DES collaboration, *Dark Energy Survey Year 3 results: Cosmology with moments of weak lensing mass maps*, *Phys. Rev. D* **106** (2022) 083509 [`arXiv:2110.10141`] [InSPIRE].

[55] A.J.S. Hamilton, *Formulae for growth factors in expanding universes containing matter and a cosmological constant*, *Mon. Not. Roy. Astron. Soc.* **322** (2001) 419 [`astro-ph/0006089`] [InSPIRE].

[56] S. Pyne and B. Joachimi, *Self-calibration of weak lensing systematic effects using combined two- and three-point statistics*, *Mon. Not. Roy. Astron. Soc.* **503** (2021) 2300 [`arXiv:2010.00614`] [InSPIRE].

[57] R. Takahashi et al., *Full-sky Gravitational Lensing Simulation for Large-area Galaxy Surveys and Cosmic Microwave Background Experiments*, *Astrophys. J.* **850** (2017) 24 [`arXiv:1706.01472`] [InSPIRE].

[58] DES collaboration, *Dark Energy Survey Year 3 results: redshift calibration of the weak lensing source galaxies*, *Mon. Not. Roy. Astron. Soc.* **505** (2021) 4249 [arXiv:2012.08566] [INSPIRE].

[59] DES collaboration, *Dark energy survey year 3 results: weak lensing shape catalogue*, *Mon. Not. Roy. Astron. Soc.* **504** (2021) 4312 [arXiv:2011.03408] [INSPIRE].

[60] H.S. Xavier, F.B. Abdalla and B. Joachimi, *Improving lognormal models for cosmological fields*, *Mon. Not. Roy. Astron. Soc.* **459** (2016) 3693 [arXiv:1602.08503] [INSPIRE].

[61] M. Jarvis, G. Bernstein and B. Jain, *The skewness of the aperture mass statistic*, *Mon. Not. Roy. Astron. Soc.* **352** (2004) 338 [astro-ph/0307393] [INSPIRE].

[62] DES collaboration, *Dark Energy Survey Year 3 results: Cosmology from cosmic shear and robustness to modeling uncertainty*, *Phys. Rev. D* **105** (2022) 023515 [arXiv:2105.13544] [INSPIRE].

[63] J. Hartlap, P. Simon and P. Schneider, *Why your model parameter confidences might be too optimistic: Unbiased estimation of the inverse covariance matrix*, *Astron. Astrophys.* **464** (2007) 399 [astro-ph/0608064] [INSPIRE].

[64] W.J. Percival et al., *The Clustering of Galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Including covariance matrix errors*, *Mon. Not. Roy. Astron. Soc.* **439** (2014) 2531 [arXiv:1312.4841] [INSPIRE].

[65] M. Takada and W. Hu, *Power Spectrum Super-Sample Covariance*, *Phys. Rev. D* **87** (2013) 123504 [arXiv:1302.6994] [INSPIRE].

[66] A. Barreira, E. Krause and F. Schmidt, *Complete super-sample lensing covariance in the response approach*, *JCAP* **06** (2018) 015 [arXiv:1711.07467] [INSPIRE].

[67] A. Barreira, E. Krause and F. Schmidt, *Accurate cosmic shear errors: do we need ensembles of simulations?*, *JCAP* **10** (2018) 053 [arXiv:1807.04266] [INSPIRE].

[68] A. Barreira, *The squeezed matter bispectrum covariance with responses*, *JCAP* **03** (2019) 008 [arXiv:1901.01243] [INSPIRE].

[69] A. Spurio Mancini et al., *CosmoPower: emulating cosmological power spectra for accelerated Bayesian inference from next-generation surveys*, *Mon. Not. Roy. Astron. Soc.* **511** (2022) 1771 [arXiv:2106.03846] [INSPIRE].

[70] S. Dodelson and M.D. Schneider, *The Effect of Covariance Estimator Error on Cosmological Parameter Constraints*, *Phys. Rev. D* **88** (2013) 063537 [arXiv:1304.2593] [INSPIRE].

[71] W.J. Percival, O. Friedrich, E. Sellentin and A. Heavens, *Matching Bayesian and frequentist coverage probabilities when using an approximate data covariance matrix*, *Mon. Not. Roy. Astron. Soc.* **510** (2022) 3207 [arXiv:2108.10402] [INSPIRE].

[72] S. Pyne, A. Tenneti and B. Joachimi, *Three-point intrinsic alignments of dark matter haloes in the IllustrisTNG simulation*, *Mon. Not. Roy. Astron. Soc.* **516** (2022) 1829 [arXiv:2204.10342] [INSPIRE].

[73] D. Huterer, M. Takada, G. Bernstein and B. Jain, *Systematic errors in future weak lensing surveys: Requirements and prospects for self-calibration*, *Mon. Not. Roy. Astron. Soc.* **366** (2006) 101 [astro-ph/0506030] [INSPIRE].

[74] E. Semboloni, C. Heymans, L. van Waerbeke and P. Schneider, *Sources of contamination to weak lensing three-point statistics: constraints from N-body simulations*, *Mon. Not. Roy. Astron. Soc.* **388** (2008) 991 [arXiv:0802.3978] [INSPIRE].

[75] M.A. Troxel and M. Ishak, *Self-Calibration Technique for 3-point Intrinsic Alignment Correlations in Weak Lensing Surveys*, *Mon. Not. Roy. Astron. Soc.* **419** (2012) 1804 [arXiv:1109.4896] [INSPIRE].

[76] A. Zonca et al., *healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python*, *J. Open Source Softw.* **4** (2019) 1298 [ɪɴSPIRE].

[77] G.P. Lepage, *Adaptive multidimensional integration: VEGAS enhanced*, *J. Comput. Phys.* **439** (2021) 110386 [`arXiv:2009.05112`] [ɪɴSPIRE].

[78] A.G.G. Matthews et al., *GPflow: A Gaussian process library using TensorFlow*, `arXiv:1610.08733` [`DOI:10.48550/ARXIV.1610.08733`].

[79] C.R. Harris et al., *Array programming with NumPy*, *Nature* **585** (2020) 357 [`arXiv:2006.10256`] [ɪɴSPIRE].

[80] J.D. Hunter, *Matplotlib: A 2D Graphics Environment*, *Comput. Sci. Eng.* **9** (2007) 90 [ɪɴSPIRE].

[81] S. Hinton, *ChainConsumer*, *J. Open Source Softw.* **1** (2016) 45.

[82] Z. Vlah, N.E. Chisari and F. Schmidt, *An EFT description of galaxy intrinsic alignments*, *JCAP* **01** (2020) 025 [`arXiv:1910.08085`] [ɪɴSPIRE].

[83] Z. Vlah, N.E. Chisari and F. Schmidt, *Galaxy shape statistics in the effective field theory*, *JCAP* **05** (2021) 061 [`arXiv:2012.04114`] [ɪɴSPIRE].

[84] J. Blazek, N. MacCrann, M.A. Troxel and X. Fang, *Beyond linear galaxy alignments*, *Phys. Rev. D* **100** (2019) 103506 [`arXiv:1708.09247`] [ɪɴSPIRE].

Chapter 6

Blinded cosmological constraints from the integrated shear 3-point correlation function in DESY3 cosmic shear data

Building on the modelling of the integrated shear 3-point correlation function $\zeta_\pm$ presented in chapters 3, 4, 5, *we have also taken the major step and applied our framework for measuring and analysing $\zeta_\pm$ in real cosmic shear data from the Year 3 data release of the Dark Energy Survey (DESY3).* We do the analysis in a blinded manner (Muir et al., 2020). *Blinding* an analysis is to hide an experiment's critical results (e.g. removing tick and labels while plotting), in our case the cosmological parameter constraints from DESY3 with $\zeta_\pm$, until all decisions affecting its analysis have been finalised in order to prevent ourselves against confirmation bias. We use the combination of the first two and the third and fourth source distribution bins of DESY3 as shown in Fig. 1 of Chapter 5 to perform our measurements of $\zeta_\pm$ inside compensated and tophat filters both having scale radii of 90 arcminutes. Using the MCMC analysis pipeline developed in Gong et al. (2023) (see Chapter 5), we have jointly analysed $\zeta_\pm$ alongside the cosmic shear 2PCF $\xi_\pm$ in blinded cosmic shear DESY3 catalogs. The whole blinding procedure is discussed in detail in Muir et al. (2020). Only when the analysers of a statistic are sure that they have finalised all analysis choices, the unblinded (true catalogs) can be used. This process is called *unblinding.* As of writing of the thesis,[1] this work is currently within the DES collaboration undergoing more tests with all four tomographic source redshift bins (instead of the combined tomographic bins as used in this chapter) and therefore is yet to be unblinded (Halder and DES Collaboration (in prep.)). In this brief chapter I present the key cosmological contours obtained from the application of our framework to the blinded DESY3 data.

---

[1]We note that at the time of writing this thesis, the DESY3 data release has already been *unblinded* following the 2PCF analyses (DES Collaboration, 2022) and the dataset has been made publicly available `https://des.ncsa.illinois.edu/releases/y3a2`. However, for the initial analyses of the integrated shear 3PCF we nevertheless choose to work with the blinded dataset.

120

6. Blinded cosmological constraints from the integrated shear 3-point correlation function in DESY3 cosmic shear data

Firstly, in Figure 6.1 we show the blinded measurements of the integrated shear 3PCFs $\zeta_\pm$, including auto and cross-correlations between the two redshift bins, and the best-fitting theory models obtained on performing an MCMC analysis around the measured blinded data vector using the $\zeta_\pm$ emulator developed in Gong et al. (2023) (see chapter 5). It is reassuring to see that the measured trends of $\zeta_\pm$ are consistent with the measurements seen in simulated data in previous chapters 3, 4, 5.
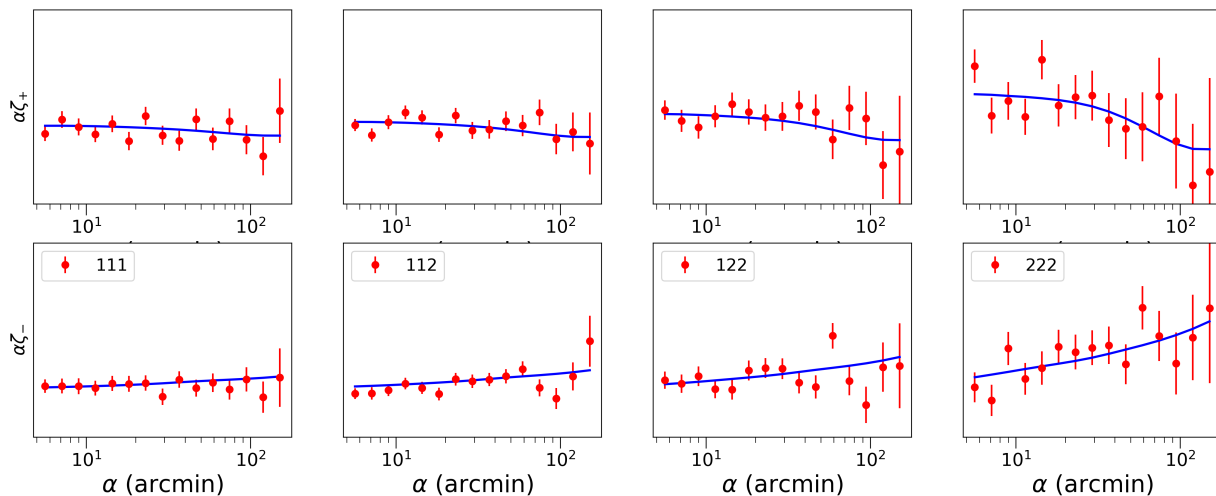
Figure 6.1: Integrated shear 3-point correlation functions $\zeta_\pm$ measured in the blinded Dark Energy Survey Year 3 (DESY3) cosmic shear data catalogs in two effective tomographic redshift bins data (red) within 90' filters. Due to blinding, we do not plot the y-axis tick marks. The blue curves show the best-fitting theory predictions after running MCMC chains around the blinded data vectors using the the $\zeta_\pm$ emulator developed in Gong et al. (2023) (see chapter 5).

In Figure 6.2 we present the MCMC constraints from analysing $\zeta_\pm$ alongside $\xi_\pm$ on three cosmological parameters: the amplitude of matter density fluctuations A_s, amount of total matter density in units of today's critical density Ω_m, and dark energy equation of state w_0, after the marginalisation of systematic parameters described in chapters 4, 5. The blinded contours already demonstrate two key results:

1. *The addition of $\zeta_\pm$ (red) to $\xi_\pm$ (blue) yields significant tightening of cosmological parameter constraints, especially on w_0 ($\sim$ **40% improvement** relative to 2PCF alone).* This is in line with the simulated likelihood analysis forecasts presented in chapter 5 and also the expectation that $\zeta_\pm$ is sensitive to the dark energy equation of state parameter due to the different line-of-sight projection weighting of the lensing bispectrum compared to $\xi_\pm$ (see chapters 3, 4).

2. Besides being tighter than the $\xi_\pm$ alone contours, the joint $\xi_\pm$ & $\zeta_\pm$ constraints in the data are at the same time consistent with $\xi_\pm$ (as already seen in the simulated analyses in chapter 5).
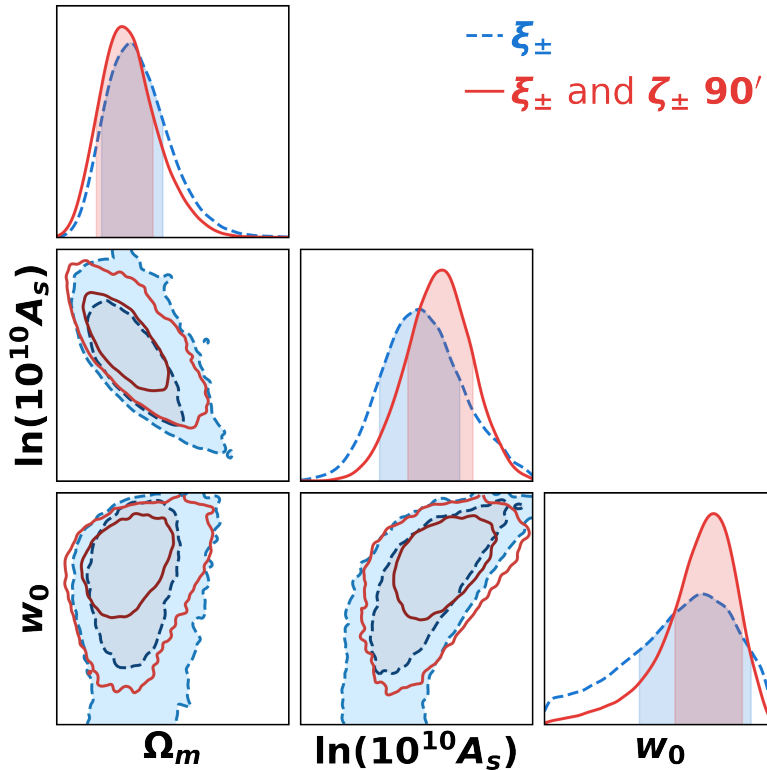
Figure 6.2: Cosmological parameter constraints with $\zeta_\pm$ from Dark Energy Survey Year 3 (DESY3) cosmic shear data (*blinded*; Halder and DES Collaboration (in prep.)) measured within 90' filters. Due to blinding, we do not plot the tick marks. *Adding $\zeta_\pm$ to the traditional shear 2PCF $\xi_\pm$* (cf. blue vs. red contours) *significantly improves the precision on cosmological parameters such as the dark energy equation of state w_0 ($\sim$ **40%** improvement*) (other parameters including lensing systematic effects are marginalised over). MCMC chains were run with the $\zeta_\pm$ emulator developed in Gong et al. (2023) (see chapter 5).

Although we are yet to unblind these results to obtain precise values on these parameters, the results already indicate the power of harvesting higher-order information in the form of the integrated 3PCF from cosmic shear datasets. In Figure 6.3, besides the constraints on the cosmological parameters we also show the constraints on two systematic parameters: baryonic feedback effects c_{min} and the nonlinear intrinsic alignment (IA) model amplitude of galaxies A_{IA}. As discussed in chapters 4, 5 these are two astrophysical systematic effects which significantly contaminate the true cosmological lensing signal and hence need to be carefully accounted for to obtain unbiased cosmological results. Although not as significant as the cosmological parameters, we find that adding $\zeta_\pm$ nevertheless brings non-negligible improvements on the constraints on these two parameters as well. This shows that besides cosmology, the integrated 3PCF can also break degeneracies between cosmological and systematic parameters appearing at the 2-point level and there-
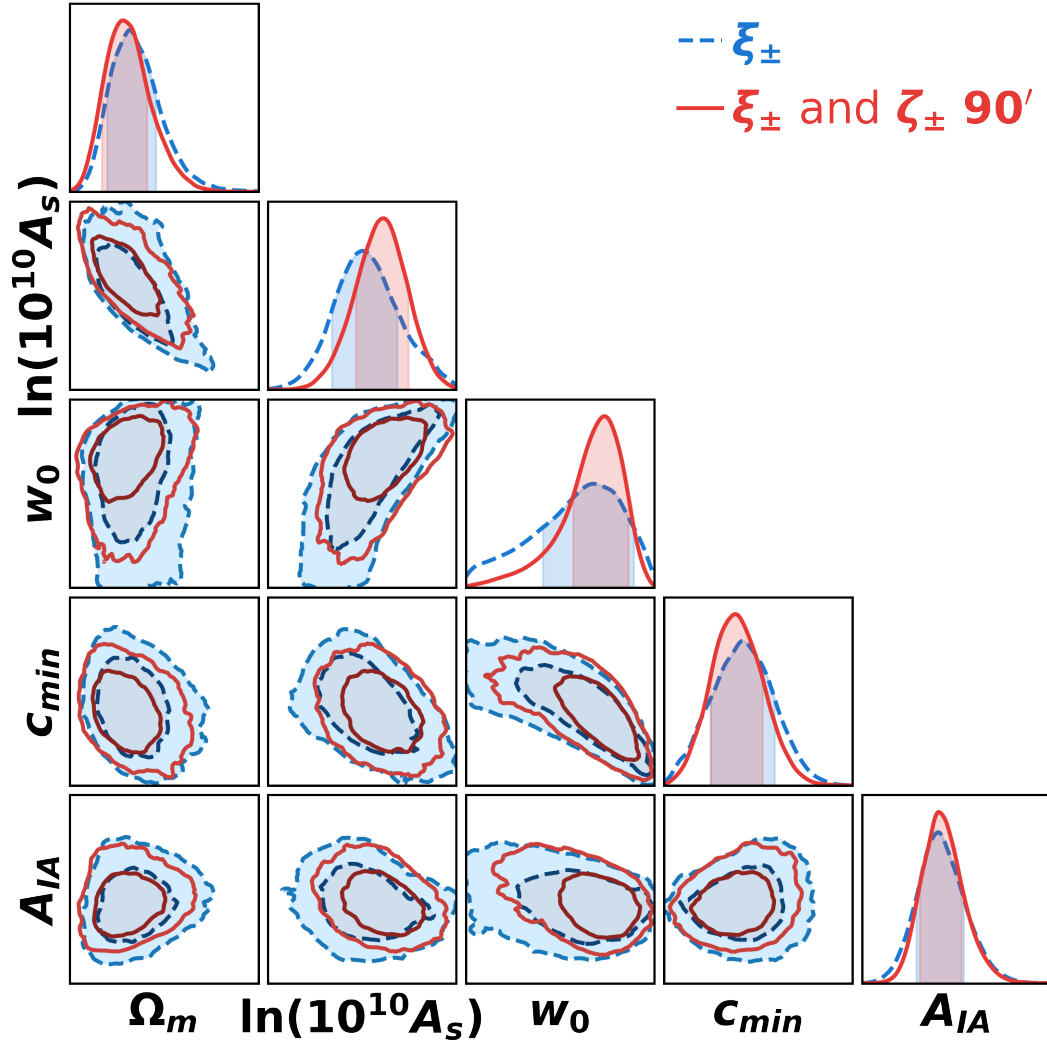
Figure 6.3:   Same as Fig. 6.2 but now also including two systematic parameters — baryonic feedback effect c_{min} and the nonlinear intrinsic alignment (IA) amplitude of galaxies. Adding $\zeta_\pm$ to $\xi_\pm$ (cf. blue vs. red contours) also brings non-negligible improvements on the precision of these systematic parameters.

fore help in constraining the systematic parameters better as well. This indicates that higher-order statistics such as integrated shear 3PCF can be used to even self-calibrate systematic parameters from data.

We note here that the setup used for this blinded analysis is not the final setup envisioned for the full analysis. We plan to analyse more filter sizes for $\zeta_\pm$ besides 90 arcminutes. As forecasted in chapter 5, this would enable even further improvements relative to $\xi_\pm$ on the parameters compared to a single filter. Moreover, instead of combining the 4 DESY3 source redshift bins into two as done for the blinded analysis (following our strategy of chapter 5) we plan to perform the analysis with all the 4 redshift bins. This can enable

even further breaking of degeneracies for the w_0 parameter with even further possibilities of cross-correlations between different tomographic bins for $\zeta_\pm$ compared to that for $\xi_\pm$. Constraining dark energy to the percent level is one of the key goals in upcoming Stage-IV surveys such as Vera Rubin Observatory LSST and Euclid Satellite mission, and our blinded study already shows that $\zeta_\pm$ will prove extremely useful to achieve this in the coming years.

Chapter 7

Beyond 3×2-point cosmology: the integrated shear and galaxy 3-point correlation functions

Bibliographic and copyright information

Author's contribution

Based on the promising results obtained using the integrated shear 3PCF in chapters 3, 4, 5, 6, in this paper I have proposed the integrated 3-point cross-correlations between cosmic shear and the foreground galaxy density fields, extending the 3×2PCFs methodology (see Sec. 2.5.1) to the practical higher-order integrated 3PCF framework. As first author: I conceptualised the framework, performed the theoretical modelling, numerical calculations, creation of mock HOD galaxy simulations from the T17 simulation suite (Takahashi et al., 2017), measurements of the integrated 3PCF from the simulations along with the MCMC as well as Fisher forecasts on parameter constraints and also wrote the paper. All the other authors contributed through valuable discussions, ideas and thorough reviewing of the paper draft.

Beyond 3×2-point cosmology: the integrated shear and galaxy 3-point correlation functions

Anik Halder,[a,b] **Zhengyangguang Gong,**[a,b] **Alexandre Barreira,**[c,d] **Oliver Friedrich,**[a,c] **Stella Seitz**[a,b] **and Daniel Gruen**[a,c]

[a]Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians-Universität München,
Scheinerstraße 1, 81679 München, Germany

[b]Max Planck Institute for Extraterrestrial Physics,
Giessenbachstraße 1, 85748 Garching, Germany

[c]Excellence Cluster ORIGINS,
Boltzmannstraße 2, 85748 Garching, Germany

[d]Ludwig-Maximilians-Universität München,
Schellingstraße 4, 80799 München, Germany

E-mail: ahalder@usm.lmu.de, lgong@usm.lmu.de, alex.barreira@origins-cluster.de,
oliver.friedrich@physik.uni-muenchen.de, stella@usm.lmu.de,
Daniel.Gruen@lmu.de

https://doi.org/10.1088/1475-7516/2023/10/028

Abstract. We present the integrated 3-point correlation functions (3PCF) involving both the cosmic shear and the galaxy density fields. These are a set of higher-order statistics that describe the modulation of local 2-point correlation functions (2PCF) by large-scale features in the fields, and which are easy to measure from galaxy imaging surveys. Based on previous works on the shear-only integrated 3PCF, we develop the theoretical framework for modelling 5 new statistics involving the galaxy field and its cross-correlations with cosmic shear. Using realistic galaxy and cosmic shear mocks from simulations, we determine the regime of validity of our models based on leading-order standard perturbation theory with an MCMC analysis that recovers unbiased constraints of the amplitude of fluctuations parameter A_s and the linear and quadratic galaxy bias parameters b_1 and b_2. Using Fisher matrix forecasts for a DES-Y3-like survey, relative to baseline analyses with conventional 3×2PCFs, we find that the addition of the shear-only integrated 3PCF can improve cosmological parameter constraints by 20–40%. The subsequent addition of the new statistics introduced in this paper can lead to further improvements of 10–20%, even when utilizing only conservatively large scales where the tree-level models are valid. Our results motivate future work on the galaxy and shear integrated 3PCFs, which offer a practical way to extend standard analyses based on 3×2PCFs to systematically probe the non-Gaussian information content of cosmic density fields.

JCAP10(2023)028

Contents

1 Introduction

Three popular 2-point statistics employed in weak gravitational lensing surveys are: (i) the 2-point cosmic shear correlation function $\xi_{\pm}$, (ii) the angular clustering of foreground lens galaxies ξ_g, and (iii) the average tangential shear signal of source galaxies around foreground lens galaxies ξ_t. Together they are known as the 3×2-point correlation functions (3×2PCFs), jointly probing projections of the late-time power spectrum of matter and galaxy density perturbations. These statistics form a key analysis tool in current surveys such as DES [1], KiDS [2], HSC-SSP [3], and are expected to continue to provide even tighter constraints on

cosmological parameters when upcoming missions like Euclid [4], Vera Rubin's LSST [5] and the Nancy Roman Space Telescope [6] go online. However, the late-time matter and galaxy density fields are non-Gaussian distributed [7] and thus have information contained in higher-order moments that are not captured by 2-point correlation functions alone. Hence, going beyond 2-point statistics and investigating higher-order correlation functions is of great interest as they can enable even tighter constraints on cosmological parameters. Efforts on this front using cosmic shear or galaxy data include the 3-point cosmic shear correlation functions (3PCF) and third-order aperture mass moments [8–13], galaxy-galaxy-galaxy lensing [14, 15], density-split statistics [16–19], the lensing aperture mass and convergence PDF [20–22], third-order convergence moments [23, 24] and weak lensing peak statistics [25–28]. These works focus mostly on cosmic shear data, with only a few analysing the galaxy and shear fields together. In particular, a robust framework for joint galaxy and shear cross-correlation analyses for a higher-order equivalent to the 3×2PCFs that can be directly applied to galaxy imaging data to obtain improved cosmological constraints has not been developed so far.

Previous works [29–31] have developed a practical higher-order cosmic shear 3-point statistic called the *integrated shear 3-point correlation function* $\zeta_{a\pm}$, which measures the correlation of the 1-point lensing aperture mass and position-dependent cosmic shear 2-point correlation function measured within sub-patches of the sky.[1] This statistic admits a well defined physical interpretation as the modulation of the small-scale shear 2PCF by long-wavelength features of the cosmic shear field. In this paper, we extend this framework to also include the projected foreground galaxy density field and its cross-correlations with the shear field at the integrated 3PCF level. The new integrated 3PCFs that we introduce in this paper are obtained by (i) measuring a position-dependent shear 2PCF, a position-dependent galaxy 2PCF, and a position-dependent tangential shear signal (i.e. a galaxy-shear cross-2PCF) within survey patches, and then (ii) correlating each of them with either a 1-point lensing aperture mass M_a or a 1-point average number density of lens galaxies M_g measured at the same patch locations. This yields a total of 6 galaxy and shear integrated 3-point correlation functions ζ, as illustrated in figure 1.

Concretely, in this paper we introduce these new statistics, derive analytical expressions for them, and evaluate them using either leading-order standard perturbation theory (SPT) or the response function (RF) approach to perturbation theory [36]. For the integrated 3PCFs that involve the galaxy density field (which are all except the $\zeta_{a\pm}$ statistic studied previously in refs. [29–31] that can be computed accurately in the nonlinear regime using RF), we investigate the regime of validity of our SPT models against measurements from realistic DES-like shear and galaxy mocks. Through Fisher forecasts for a DES-like survey, we also investigate the constraining power of these statistics relative to the conventional 3×2PCFs. We find that the addition of $\zeta_{a\pm}$ (the integrated 3PCF computed from only the shear field) already allows for significant improvements of $20-40\%$ on parameters such as the cold-dark matter density $\Omega_{\rm cdm}$, amplitude of primordial density fluctuations A_s, dark energy equation of state parameter w_0, reduced Hubble parameter h and even the linear galaxy bias parameter b_1. Adding the galaxy integrated 3PCFs and utilizing only the large scales where their leading-order SPT models are valid allows for further improvements of $\sim 10\%$, which can increase to $\sim 20\%$ when investigating extended cosmologies such as with a dynamical dark energy equation of state parameter. Our results strongly motivate further work on these

---

[1]Readers interested in the development history of the integrated 3PCF are referred to refs. [29–35] for details.
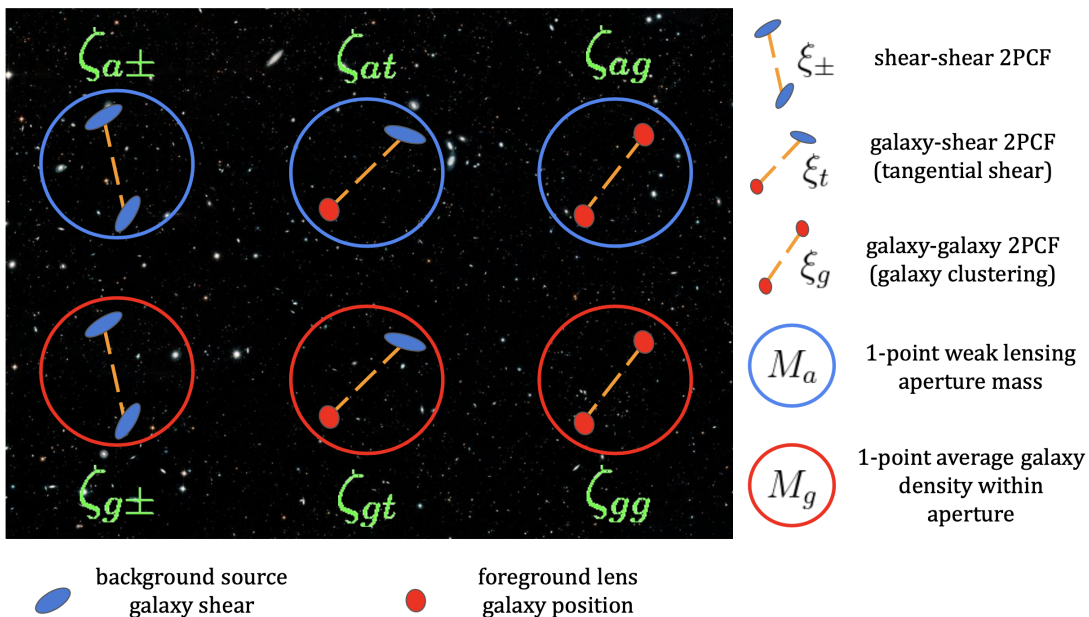
Figure 1. Illustration showing the ingredients needed for computing the integrated 3PCFs using both cosmic shear (filled blue ellipses) and foreground lens galaxy (filled red ellipses) fields. The ingredients are the position-dependent shear 2PCFs $\xi_{\pm}$, position-dependent tangential shear 2PCFs ξ_t and the position-dependent galaxy clustering 2PCFs ξ_g measured inside apertures on the sky (left, middle and right, respectively). Also needed are the 1-point weak lensing aperture mass M_a (blue circles) and the 1-point average galaxy density contrast M_g (red circles), both measured in the same apertures/patches. The 6 integrated 3PCFs $\zeta_{a\pm}, \zeta_{at}, \zeta_{ag}, \zeta_{g\pm}, \zeta_{gt}, \zeta_{gg}$ (left to right starting from top-left in the illustration) are measured by correlating the 3 position-dependent 2PCFs and the 2 aperture masses. Background image: Hubble Legacy Field. Credit: NASA, ESA and Hubble Legacy Field team.

practical higher-order galaxy and shear integrated 3-point statistics, which have the potential to improve cosmological, as well as galaxy bias parameter constraints using real survey data.

The rest of this paper is as follows. In sections 2, 3 and 4 we present the theory expressions for the 3×2PCFs, the position-dependent 1- and 2-point statistics, and the 6 integrated 3PCFs ζ. We present the lensing and galaxy mocks we use in this work in section 5, and our data vector measurements and covariance estimation in section 6. In section 7 we present the comparison of our theoretical models against simulations (section 7.1), the validation of our models using Monte Carlo Markov Chain (MCMC) analyses (section 7.2), and Fisher forecasts (section 7.3) to investigate the constraining power of the integrated 3PCFs on cosmological and galaxy bias parameters. We summarise and conclude in section 8. Appendix A describes more details about the 2D projected shear and galaxy density fields. In appendix B we present expressions for the galaxy bias terms calculated in the halo occupation distribution (HOD) formalism. Appendix C details the expressions of the galaxy-matter power- and bi-spectrum models at leading order in SPT. In appendix D we discuss details of the point mass contribution to the correlations involving the tangential shear 2PCF.

2 2-point correlations of shear and galaxy fields

We study the correlations of two projected cosmic density fields, namely the weak lensing shear field and the projected galaxy density contrast field.[2] We refer the reader to appendix A for further details about these two fields.

The weak lensing convergence $\kappa(\boldsymbol{\theta})$ can be expressed as a line-of-sight projection of the 3D matter density contrast field δ_m^{3D} [37–39]

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_{\text{lim}}} d\chi \, q_\kappa(\chi) \, \delta_m^{3D}(\boldsymbol{x}, \tau), \tag{2.1}$$

where $\boldsymbol{x} = (\chi\boldsymbol{\theta}, \chi)$ is the 3D comoving position with χ the comoving radial coordinate, $\tau = \tau_0 - \chi$ the conformal time coordinate with τ_0 the present day conformal time, and $q_\kappa(\chi)$ the lensing projection kernel for source galaxies which follow a normalized distribution $p(\chi')$:

$$q_\kappa(\chi) = \frac{3H_0^2\Omega_{\text{m}}}{2c^2} \frac{\chi}{a(\chi)} \int_\chi^{\chi_{\text{lim}}} d\chi' p(\chi') \frac{\chi' - \chi}{\chi'}. \tag{2.2}$$

Here, χ_{lim} is the upper integral limit of the comoving coordinate usually taken to be the size of the comoving horizon of the observable Universe and $a(\chi)$ is the scale factor of the Universe parametrized in terms of χ which we also utilise as a time coordinate instead of τ. H_0 is the Hubble parameter and Ω_{m} is the total matter density parameter today. The weak lensing shear $\gamma(\boldsymbol{\theta}) = \gamma_1(\boldsymbol{\theta}) + i\gamma_2(\boldsymbol{\theta})$ at a given angular position $\boldsymbol{\theta}$ on the sky, which can be directly estimated using source galaxy shapes, is a complex quantity where the shear components γ_1 and γ_2 are specified in a chosen Cartesian frame (we will work in 2D flat-sky). One can relate this complex shear field to the convergence field through second-order derivatives of the 2D lensing potential. In Fourier space, the shear $\gamma(\boldsymbol{\ell})$ is related to the convergence $\kappa(\boldsymbol{\ell})$ via[3]

$$\gamma(\boldsymbol{\ell}) = \frac{(\ell_x + i \, \ell_y)^2}{\ell^2} \kappa(\boldsymbol{\ell}) = e^{2i\phi_{\boldsymbol{\ell}}}\kappa(\boldsymbol{\ell}) \, ; \qquad \text{for } \ell \neq 0, \tag{2.3}$$

where $\ell = \sqrt{\ell_x^2 + \ell_y^2}$ and $\phi_{\boldsymbol{\ell}} = \arctan(\ell_y/\ell_x)$ is the polar angle of $\boldsymbol{\ell}$.

The local comoving number density of galaxies at comoving position $\boldsymbol{x}$ and time τ can be written in terms of the number density contrast of galaxies $\delta_g^{3D}(\boldsymbol{x}, \tau)$ as

$$n_g^{3D}(\boldsymbol{x}, \tau) = \bar{n}_g^{3D}(\tau)[1 + \delta_g^{3D}(\boldsymbol{x}, \tau)], \tag{2.4}$$

where $\bar{n}_g^{3D}(\tau)$ is the cosmic mean comoving number density of galaxies. The 2D projected galaxy density contrast field $\delta_g^{2D}(\boldsymbol{\theta})$ can then be defined as a line-of-sight projection of δ_g^{3D}

$$\delta_g^{2D}(\boldsymbol{\theta}) = \int d\chi \, q_g(\chi) \, \delta_g^{3D}(\boldsymbol{x}, \tau), \tag{2.5}$$

where $q_g(\chi)$ is a normalized projection kernel of galaxies, which in this case is identified as the observed distribution of foreground lens galaxies i.e. $q_g(\chi) = p(\chi)$. The δ_g^{3D} field is considered to be a *tracer* of the underlying matter field $\delta \equiv \delta_m^{3D}$, which can be expressed using

---

[2]To simplify the language, we will drop 'projected' in 'projected galaxy density field' from hereon.

[3]To ease the notation, we distinguish between real- and Fourier-space variables by their arguments. For example, $\kappa(\boldsymbol{\theta})$ and $\kappa(\boldsymbol{\ell})$ are the lensing convergence representations in real and Fourier space, respectively.

a series expansion of *operators* $\mathcal{O}$ with accompanying *bias* coefficients $b_\mathcal{O}(\tau)$ and *stochasticity* parameters $\epsilon_\mathcal{O}(\boldsymbol{x}, \tau)$ (see ref. [40] for a comprehensive review):

$$\delta_g^{\rm 3D}(\boldsymbol{x}, \tau) = \sum_\mathcal{O} b_\mathcal{O}(\tau)\mathcal{O}(\boldsymbol{x}, \tau) + \left[\epsilon(\boldsymbol{x}, \tau) + \sum_\mathcal{O} \epsilon_\mathcal{O}(\boldsymbol{x}, \tau)\,\mathcal{O}(\boldsymbol{x}, \tau)\right]. \tag{2.6}$$

Specifically, up to second order in perturbations this equation can be written as

$$\delta_g^{\rm 3D}(\boldsymbol{x}, \tau) = b_\delta(\tau)\,\delta(\boldsymbol{x}, \tau) + b_{\delta^2}(\tau)\,\delta^2(\boldsymbol{x}, \tau) + b_{K^2}(\tau)\,K^2(\boldsymbol{x}, \tau) + \left[\epsilon(\boldsymbol{x}, \tau) + \epsilon_\delta(\boldsymbol{x}, \tau)\,\delta(\boldsymbol{x}, \tau)\right], \tag{2.7}$$

where $K^2 = K_{ij}K^{ij}$ is the square of the 3D tidal field $K_{ij} = \left(\frac{\partial_i \partial_j}{\nabla^2} - \frac{\delta_{ij}}{3}\right)\delta_m^{\rm 3D}$. In this equation we have ignored higher-order spatial derivatives of $\delta_m^{\rm 3D}$. The bias parameters $b_\mathcal{O}(\tau)$ are interpreted as the response of the local number density of galaxies $n_g^{\rm 3D}(\boldsymbol{x}, \tau)$ to changes in the amplitude of the operators $\mathcal{O}(\boldsymbol{x}, \tau)$; they absorb the complicated details of small-scale galaxy formation and evolution [41]. These bias terms are often expressed using another notation $b_1 \equiv b_\delta, b_2 \equiv 2b_{\delta^2}, b_{s^2} \equiv 2b_{K^2}$ which we adopt throughout. The terms inside the square brackets in eqs. (2.6) and (2.7) denote the non-deterministic (stochastic) part of the galaxy-matter relation which arise due to perturbations on small scales in the underlying $\mathcal{O}$ fields. Similar to the deterministic bias relation, the stochastic relation comes with its own free parameters $\epsilon_\mathcal{O}$ and an offset term ϵ.[4] We shall consider the bias and stochastic terms to be time-independent inside a given galaxy redshift bin.

Using the (cross-) correlations of the cosmic shear and galaxy density contrast fields, we can construct three 2PCFs:

- Cosmic shear 2PCFs $\xi_\pm$ (shear-shear) defined by correlating the *rotated shear* γ_{ϕ_α} at two angular positions $\boldsymbol{\theta}$ and $\boldsymbol{\theta} + \boldsymbol{\alpha}$ on the shear field, where γ_{ϕ_α} at each point is computed along the direction ϕ_α of the separation vector $\boldsymbol{\alpha}$ between the two points (see appendix A or refs. [42, 43]),

$$\begin{aligned} \xi_+^{ij}(\alpha) &\equiv \langle \gamma_{\phi_\alpha}^i(\boldsymbol{\theta})\gamma_{\phi_\alpha}^{j*}(\boldsymbol{\theta} + \boldsymbol{\alpha})\rangle = \int \frac{\mathrm{d}\ell\,\ell}{2\pi}\mathcal{P}_\kappa^{ij}(\ell)J_0(\ell\alpha), \\ \xi_-^{ij}(\alpha) &\equiv \langle \gamma_{\phi_\alpha}^i(\boldsymbol{\theta})\gamma_{\phi_\alpha}^j(\boldsymbol{\theta} + \boldsymbol{\alpha})\rangle = \int \frac{\mathrm{d}\ell\,\ell}{2\pi}\mathcal{P}_\kappa^{ij}(\ell)J_4(\ell\alpha); \end{aligned} \tag{2.8}$$

- Angular galaxy clustering 2PCF ξ_g (galaxy-galaxy) measured by correlating two points separated by α on the galaxy density contrast field $\delta_g^{\rm 2D}$ [44]:

$$\xi_g^{ij}(\alpha) \equiv \langle \delta_g^{{\rm 2D},i}(\boldsymbol{\theta})\delta_g^{{\rm 2D},j}(\boldsymbol{\theta} + \boldsymbol{\alpha})\rangle = \int \frac{\mathrm{d}\ell\,\ell}{2\pi}\mathcal{P}_g^{ij}(\ell)J_0(\ell\alpha); \tag{2.9}$$

- Tangential shear 2PCF ξ_t (galaxy-shear), which is the cross-correlation of the foreground galaxy density field with the rotated shear of a background source galaxy along the direction of the separation vector $\boldsymbol{\alpha}$ joining the foreground lens and the background source galaxy.[5] It can be written as (see appendix D or ref. [44]):

$$\xi_t^{ij}(\alpha) \equiv \langle \delta_g^{{\rm 2D},i}(\boldsymbol{\theta})\gamma_{\phi_\alpha}^j(\boldsymbol{\theta} + \boldsymbol{\alpha})\rangle = \underbrace{\int \frac{\mathrm{d}\ell\,\ell}{2\pi}\mathcal{P}_t^{ij}(\ell)J_2(\ell\alpha)}_{\xi_t^{ij,{\rm PT}}(\alpha)} + \frac{\mathcal{M}_t^i}{\alpha^2}, \tag{2.10}$$

---

[4]We assume the stochastic terms to be Poisson random variables (see section C). Exploring non-Poisson stochasticity in the context of the integrated 3PCF is left to future work.

[5]This 2PCF is also known as *galaxy-galaxy lensing* in literature, but we refrain from calling it so to avoid confusions with the galaxy-galaxy clustering 2PCF.

where we include the contribution from the so-called *point-mass* term, whose amplitude $\mathcal{M}_t$ is a free parameter of the model. These terms are due to the fact that there are small nonlinear scales, which are not well captured by perturbation theory ξ_t^{PT}, but which can still contribute to the signal on large scales due to the nonlocal nature of the tangential shear signal (see appendix D or ref. [45] for more details).

In the equations above, the superscripts i, j denote tomographic bins of the background shear source galaxies or the foreground lens galaxies. We consider only the so-called E-mode shear fields, for which the imaginary parts of $\xi_\pm$ and ξ_t vanish. In the last equalities in eqs. (2.8), (2.9) and (2.10), we have related the real space 2PCFs to the corresponding lensing/galaxy (cross-) power spectra through inverse harmonic transforms (with J_n being the n-th order ordinary Bessel function of the first kind). These spectra can in turn be expressed as line-of-sight projections of the 3D matter/galaxy (cross-) power spectra using the Limber approximation [44, 46, 47] (see appendix C):

$$\mathcal{P}_\kappa^{ij}(\ell) = \int \mathrm{d}\chi \frac{q_\kappa^i(\chi) q_\kappa^j(\chi)}{\chi^2} P_{mm}^{\mathrm{3D}}\left(\frac{\ell}{\chi}, \chi\right), \tag{2.11a}$$

$$\mathcal{P}_g^{ij}(\ell) = \int \mathrm{d}\chi \frac{q_g^i(\chi) q_g^j(\chi)}{\chi^2} P_{gg}^{\mathrm{3D}}\left(\frac{\ell}{\chi}, \chi\right), \tag{2.11b}$$

$$\mathcal{P}_t^{ij}(\ell) = \int \mathrm{d}\chi \frac{q_g^i(\chi) q_\kappa^j(\chi)}{\chi^2} P_{gm}^{\mathrm{3D}}\left(\frac{\ell}{\chi}, \chi\right). \tag{2.11c}$$

Here, we have defined the convergence power spectrum in Fourier space as $(2\pi)^2 \mathcal{P}_\kappa^{ij}(\ell)\delta_D(\boldsymbol{\ell} + \boldsymbol{\ell}') = \langle \kappa^i(\boldsymbol{\ell})\kappa^j(\boldsymbol{\ell}')\rangle$, the 2D galaxy number density contrast power spectrum as $(2\pi)^2 \mathcal{P}_g^{ij}(\ell)\delta_D(\boldsymbol{\ell} + \boldsymbol{\ell}') = \langle \delta_g^{\mathrm{2D},i}(\boldsymbol{\ell})\delta_g^{\mathrm{2D},j}(\boldsymbol{\ell}')\rangle$, and the convergence-galaxy cross-power spectrum as $(2\pi)^2 \mathcal{P}_t^{ij}(\ell)\delta_D(\boldsymbol{\ell} + \boldsymbol{\ell}') = \langle \delta_g^{\mathrm{2D},i}(\boldsymbol{\ell})\kappa^j(\boldsymbol{\ell}')\rangle$. We evaluate only the auto-correlations of galaxies in ξ_g within the same foreground galaxy redshift bin, i.e. $i = j$, because the cross-correlation between galaxy density fields ξ_g^{ij} in different redshift bins $i \neq j$ is small.[6]

We use `HMCODE` [48] to evaluate the nonlinear matter power spectrum P_{mm}^{3D} that enters the calculation of $\xi_\pm$. To evaluate P_{gg}^{3D} and P_{gm}^{3D}, which enter the calculation of ξ_g and ξ_t, we will rely on standard perturbation theory (SPT). Concretely, we work to leading order (tree-level) and evaluate these 3D spectra as $P_{gg} = b_1^2 P_{mm} + P_{\epsilon\epsilon}$ and $P_{gm} = b_1 P_{mm}$, where $P_{\epsilon\epsilon}$ is the power spectrum of the stochastic field $\epsilon(\boldsymbol{x})$ (see appendix C). Note that owing to our inability to make predictions for galaxy clustering observations on small, nonlinear scales, the galaxy-related statistics will be limited to larger scales compared to $\xi_\pm$.

We note already here that 3×2PCFs analyses are able to simultaneously constrain both the b_1 and the A_s parameters, where A_s is the amplitude of the primordial scalar power spectrum; at leading-order, $\xi_\pm \propto A_s$, $\xi_t \propto b_1 A_s$ and $\xi_g \propto b_1^2 A_s$. We will see later that when we also consider the integrated 3PCFs involving the galaxy density field, which display different scalings with b_1 and A_s, they will allow for further breaking of degeneracies and help put tighter constraints.

3 Position-dependent statistics of shear and galaxy fields

Having looked at the global 2PCFs, we turn our attention now to position-dependent quantities, i.e. statistics of the fields within sub-patches of the survey (cf. figure 1).

---

[6]Effects such as lensing magnification can induce non-zero correlations between galaxies in different redshift bins [44], but we defer the modelling of these effects to future work.

3.1 Position-dependent 2-point correlation functions

First, we consider the case of the *position-dependent 2PCFs* which we define as 2PCFs measured inside finite patches. Following the mathematical formalism of ref. [29], we can write the angle-averaged cosmic shear, galaxy clustering and the tangential shear position-dependent 2PCFs as

$$\xi_+^{ij}(\alpha; \boldsymbol{\theta}_C) = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\phi_{\boldsymbol{\alpha}}}{2\pi} \int \mathrm{d}^2\boldsymbol{\theta} \; \gamma_{\phi_{\boldsymbol{\alpha}}}^i(\boldsymbol{\theta}; \boldsymbol{\theta}_C) \gamma_{\phi_{\boldsymbol{\alpha}}}^{j*}(\boldsymbol{\theta} + \boldsymbol{\alpha}; \boldsymbol{\theta}_C), \tag{3.1a}$$

$$\xi_-^{ij}(\alpha; \boldsymbol{\theta}_C) = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\phi_{\boldsymbol{\alpha}}}{2\pi} \int \mathrm{d}^2\boldsymbol{\theta} \; \gamma_{\phi_{\boldsymbol{\alpha}}}^i(\boldsymbol{\theta}; \boldsymbol{\theta}_C) \gamma_{\phi_{\boldsymbol{\alpha}}}^j(\boldsymbol{\theta} + \boldsymbol{\alpha}; \boldsymbol{\theta}_C), \tag{3.1b}$$

$$\xi_g^{ij}(\alpha; \boldsymbol{\theta}_C) = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\phi_{\boldsymbol{\alpha}}}{2\pi} \int \mathrm{d}^2\boldsymbol{\theta} \; \delta_g^{2\mathrm{D},i}(\boldsymbol{\theta}; \boldsymbol{\theta}_C) \delta_g^{2\mathrm{D},j}(\boldsymbol{\theta} + \boldsymbol{\alpha}; \boldsymbol{\theta}_C), \tag{3.1c}$$

$$\xi_t^{ij}(\alpha; \boldsymbol{\theta}_C) = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\phi_{\boldsymbol{\alpha}}}{2\pi} \int \mathrm{d}^2\boldsymbol{\theta} \; \delta_g^{2\mathrm{D},i}(\boldsymbol{\theta}; \boldsymbol{\theta}_C) \gamma_{\phi_{\boldsymbol{\alpha}}}^j(\boldsymbol{\theta} + \boldsymbol{\alpha}; \boldsymbol{\theta}_C), \tag{3.1d}$$

where the *windowed rotated shear* with respect to direction $\phi_{\boldsymbol{\alpha}}$ at location $\boldsymbol{\theta}$ inside a top-hat aperture W centred at $\boldsymbol{\theta}_C$ reads

$$\gamma_{\phi_{\boldsymbol{\alpha}}}(\boldsymbol{\theta}; \boldsymbol{\theta}_C) = \gamma_{\phi_{\boldsymbol{\alpha}}}(\boldsymbol{\theta}) W(\boldsymbol{\theta} - \boldsymbol{\theta}_C), \tag{3.2}$$

and

$$\delta_g^{2\mathrm{D}}(\boldsymbol{\theta}; \boldsymbol{\theta}_C) = \delta_g^{2\mathrm{D}}(\boldsymbol{\theta}) W(\boldsymbol{\theta} - \boldsymbol{\theta}_C) \tag{3.3}$$

is the *windowed 2D galaxy density contrast* at location $\boldsymbol{\theta}$ inside the same top-hat. The area normalization factor is given by

$$A_{2\mathrm{pt}}(\alpha) = \int \frac{\mathrm{d}\phi_{\boldsymbol{\alpha}}}{2\pi} \int \mathrm{d}^2\boldsymbol{\theta} \; W(\boldsymbol{\theta} - \boldsymbol{\theta}_C) W(\boldsymbol{\theta} + \boldsymbol{\alpha} - \boldsymbol{\theta}_C) = \int \frac{\mathrm{d}\ell \, \ell}{2\pi} \; W(\ell)^2 J_0(\ell\alpha), \tag{3.4}$$

where in the last equality we have used the Fourier space representation of the window function and used the fact that we adopt only azimuthally symmetric apertures W. Specifically, for a top-hat filter of angular radius θ_T, we have:

$$W(\boldsymbol{\ell}) = W(\ell) = 2\pi\theta_T^2 \frac{J_1(\ell\theta_T)}{\ell\theta_T}. \tag{3.5}$$

In Fourier space, these statistics can be expressed as (using eq. (2.3) and following similar derivation steps as in refs. [29, 30])

$$\xi_+^{ij}(\alpha; \boldsymbol{\theta}_C) = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int_{\phi_{\boldsymbol{\alpha}}} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} \int_{\boldsymbol{q}_1} \kappa^i(\boldsymbol{\ell}_1) \kappa^j(\boldsymbol{\ell}_2) e^{2i(\phi_{\boldsymbol{\ell}_1} - \phi_{\boldsymbol{\ell}_2})} W(\boldsymbol{q}_1) W(\boldsymbol{\ell}_{12} - \boldsymbol{q}_1) e^{i(\boldsymbol{q}_1 - \boldsymbol{\ell}_1)\cdot\boldsymbol{\alpha}} e^{i\boldsymbol{\ell}_{12}\cdot\boldsymbol{\theta}_C},$$
$$\tag{3.6a}$$

$$\xi_-^{ij}(\alpha; \boldsymbol{\theta}_C) = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int_{\phi_{\boldsymbol{\alpha}}} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} \int_{\boldsymbol{q}_1} \kappa^i(\boldsymbol{\ell}_1) \kappa^j(\boldsymbol{\ell}_2) e^{2i(\phi_{\boldsymbol{\ell}_1} + \phi_{\boldsymbol{\ell}_2} - 4i\phi_{\boldsymbol{\alpha}})} W(\boldsymbol{q}_1) W(\boldsymbol{\ell}_{12} - \boldsymbol{q}_1)$$
$$\times \; e^{i(\boldsymbol{q}_1 - \boldsymbol{\ell}_1)\cdot\boldsymbol{\alpha}} e^{i\boldsymbol{\ell}_{12}\cdot\boldsymbol{\theta}_C}, \tag{3.6b}$$

$$\xi_g^{ij}(\alpha; \boldsymbol{\theta}_C) = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int_{\phi_{\boldsymbol{\alpha}}} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} \int_{\boldsymbol{q}_1} \delta_g^{2\mathrm{D},i}(\boldsymbol{\ell}_1) \delta_g^{2\mathrm{D},j}(\boldsymbol{\ell}_2) W(\boldsymbol{q}_1) W(\boldsymbol{\ell}_{12} - \boldsymbol{q}_1) e^{i(\boldsymbol{q}_1 - \boldsymbol{\ell}_1)\cdot\boldsymbol{\alpha}} e^{i\boldsymbol{\ell}_{12}\cdot\boldsymbol{\theta}_C},$$
$$\tag{3.6c}$$

$$\xi_t^{ij}(\alpha; \boldsymbol{\theta}_C) = \frac{-1}{A_{2\mathrm{pt}}(\alpha)} \int_{\phi_{\boldsymbol{\alpha}}} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} \int_{\boldsymbol{q}_1} \delta_g^{2\mathrm{D},i}(\boldsymbol{\ell}_1) \kappa^j(\boldsymbol{\ell}_2) e^{2i(\phi_{\boldsymbol{\ell}_2} - \phi_{\boldsymbol{\alpha}})} W(\boldsymbol{q}_1) W(\boldsymbol{\ell}_{12} - \boldsymbol{q}_1)$$
$$\times \; e^{i(\boldsymbol{q}_1 - \boldsymbol{\ell}_1)\cdot\boldsymbol{\alpha}} e^{i\boldsymbol{\ell}_{12}\cdot\boldsymbol{\theta}_C}. \tag{3.6d}$$

The $\phi_{\boldsymbol{\ell}_1}$ and $\phi_{\boldsymbol{\ell}_2}$ are polar angles of the wavevectors $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2$, respectively. We also defined the shorthand notations $\int_{\phi_{\boldsymbol{\alpha}}} \equiv \int \mathrm{d}\phi_{\boldsymbol{\alpha}}/(2\pi)$, $\int_{\boldsymbol{\ell}} \equiv \int \mathrm{d}^2\boldsymbol{\ell}/(2\pi)^2$ and $\boldsymbol{\ell}_{12\ldots n} \equiv \boldsymbol{\ell}_1 + \boldsymbol{\ell}_2 + \ldots + \boldsymbol{\ell}_n$. It can be shown that, as expected, the ensemble averages of these position-dependent 2PCFs give the corresponding global 2PCFs.

3.2 Position-dependent 1-point statistics

In order to predict the integrated 3PCFs we also need *position-dependent 1-point statistics*. We consider in particular the 1-point lensing aperture mass statistic and the 1-point average galaxy density contrast.

The *1-point lensing aperture mass* in shear tomographic bin i is defined as [38]

$$
M_a^i(\boldsymbol{\theta}_C) = \int \mathrm{d}^2\boldsymbol{\theta}\; \gamma_t^i(\boldsymbol{\theta}, \phi_{\boldsymbol{\theta}-\boldsymbol{\theta}_C}) Q(\boldsymbol{\theta} - \boldsymbol{\theta}_C) = \int \mathrm{d}^2\boldsymbol{\theta}\; \kappa^i(\boldsymbol{\theta}) U(\boldsymbol{\theta} - \boldsymbol{\theta}_C) = \int_{\boldsymbol{\ell}} \kappa^i(\boldsymbol{\ell}) U(\boldsymbol{\ell}) e^{i\boldsymbol{\ell}\cdot\boldsymbol{\theta}_C},
\tag{3.7}
$$

where in the last equality we have used the Fourier space representation of the aperture mass. In practice, this can be measured as a weighted mean of the tangential shear field $\gamma_t \equiv \Re[\gamma_{\phi_{\boldsymbol{\theta}-\boldsymbol{\theta}_C}}]$ inside a compensated filter Q centred at location $\boldsymbol{\theta}_C$; the Fourier representation in terms of the lensing convergence is useful from a theoretical modelling perspective. These compensated filters by definition satisfy:

$$
\int \mathrm{d}^2\boldsymbol{\theta}\; U(\boldsymbol{\theta} - \boldsymbol{\theta}_C) = \int \mathrm{d}^2\boldsymbol{\theta}\; Q(\boldsymbol{\theta} - \boldsymbol{\theta}_C) = 0,
\tag{3.8}
$$

and hence an area normalisation term for this lensing aperture mass is irrelevant. We adopt the following azimuthally-symmetric form for the U and Q filters [49]:

$$
U(\theta) = \frac{1}{2\pi\theta_{ap}^2}\left(1 - \frac{\theta^2}{2\theta_{ap}^2}\right)\exp\left(-\frac{\theta^2}{2\theta_{ap}^2}\right)\;; \qquad Q(\theta) = \frac{\theta^2}{4\pi\theta_{ap}^2}\exp\left(-\frac{\theta^2}{2\theta_{ap}^2}\right),
\tag{3.9}
$$

where the aperture scale of the compensated filter is denoted by θ_{ap}. The Fourier space expression for U (which we use for theoretical predictions) is given by

$$
U(\ell) = \int \mathrm{d}^2\boldsymbol{\theta}\; U(\theta) e^{-i\boldsymbol{\ell}\cdot\boldsymbol{\theta}} = \frac{\ell^2\theta_{ap}^2}{2}\exp\left(-\frac{\ell^2\theta_{ap}^2}{2}\right).
\tag{3.10}
$$

Additionally, we also define the *1-point projected average galaxy density* within a top-hat filter W centred at location $\boldsymbol{\theta}_C$ and measured in foreground lens tomographic bin i as

$$
M_g^i(\boldsymbol{\theta}_C) \equiv \frac{1}{A_W} \int \mathrm{d}^2\boldsymbol{\theta}\; \delta_g^{2\mathrm{D},i}(\boldsymbol{\theta}) W(\boldsymbol{\theta} - \boldsymbol{\theta}_C) = \frac{1}{A_W} \int_{\boldsymbol{\ell}} \delta_g^{2\mathrm{D},i}(\boldsymbol{\ell}) W(\boldsymbol{\ell}) e^{i\boldsymbol{\ell}\cdot\boldsymbol{\theta}_C}.
\tag{3.11}
$$

The area normalisation term in this case is given by

$$
A_W = \int \mathrm{d}^2\boldsymbol{\theta}\; W(\boldsymbol{\theta} - \boldsymbol{\theta}_C),
\tag{3.12}
$$

which is simply the area enclosed by the top-hat filter.

4 Integrated 3-point correlations of shear and galaxy fields

We now have all of the ingredients needed to compute the integrated 3-point correlation functions involving the cosmic shear and the galaxy density contrast fields. In essence, an integrated 3PCF is simply the correlation between (i) a position dependent 1-point weighted mean within a patch of the survey with (ii) the position-dependent 2PCF measured at the same patch location (see figure 1). In refs. [29–31], the authors studied the case of the integrated shear 3PCF $\zeta_{a\pm}(\alpha)$,[7] which corresponds to the correlation between the 1-point lensing aperture mass $M_a(\boldsymbol{\theta}_C)$ with the position-dependent shear 2PCFs $\xi_\pm(\alpha;\boldsymbol{\theta}_C)$. With the galaxy density contrast field, we can construct 5 additional such cross-correlations, enabling a total of 6 integrated 3PCFs. The derivation steps are similar for all 6 statistics [29], which can be written as

$$\zeta_{a\pm}^{ijk}(\alpha) = \left\langle M_a^i(\boldsymbol{\theta}_C)\, \xi_\pm^{jk}(\alpha;\boldsymbol{\theta}_C)\right\rangle = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi}\, \mathcal{B}_{a\pm}^{ijk}(\ell) J_{0/4}(\ell\alpha), \tag{4.1a}$$

$$\zeta_{g\pm}^{ijk}(\alpha) = \left\langle M_g^i(\boldsymbol{\theta}_C)\, \xi_\pm^{jk}(\alpha;\boldsymbol{\theta}_C)\right\rangle = \frac{1}{A_W A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi}\, \mathcal{B}_{g\pm}^{ijk}(\ell) J_{0/4}(\ell\alpha), \tag{4.1b}$$

$$\zeta_{ag}^{ijk}(\alpha) = \left\langle M_a^i(\boldsymbol{\theta}_C)\, \xi_g^{jk}(\alpha;\boldsymbol{\theta}_C)\right\rangle = \frac{1}{A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi}\, \mathcal{B}_{ag}^{ijk}(\ell) J_0(\ell\alpha), \tag{4.1c}$$

$$\zeta_{gg}^{ijk}(\alpha) = \left\langle M_g^i(\boldsymbol{\theta}_C)\, \xi_g^{jk}(\alpha;\boldsymbol{\theta}_C)\right\rangle = \frac{1}{A_W A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi}\, \mathcal{B}_{gg}^{ijk}(\ell) J_0(\ell\alpha), \tag{4.1d}$$

$$\zeta_{at}^{ijk}(\alpha) = \left\langle M_a^i(\boldsymbol{\theta}_C)\, \xi_t^{jk}(\alpha;\boldsymbol{\theta}_C)\right\rangle = \underbrace{\frac{1}{A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi}\, \mathcal{B}_{at}^{ijk}(\ell) J_2(\ell\alpha)}_{\zeta_{at}^{ijk,\mathrm{PT}}(\alpha)} + \frac{\mathcal{M}_{at}^j}{\alpha^2}, \tag{4.1e}$$

$$\zeta_{gt}^{ijk}(\alpha) = \left\langle M_g^i(\boldsymbol{\theta}_C)\, \xi_t^{jk}(\alpha;\boldsymbol{\theta}_C)\right\rangle = \underbrace{\frac{1}{A_W A_{2\mathrm{pt}}(\alpha)} \int \frac{\mathrm{d}\ell\, \ell}{2\pi}\, \mathcal{B}_{gt}^{ijk}(\ell) J_2(\ell\alpha)}_{\zeta_{gt}^{ijk,\mathrm{PT}}(\alpha)} + \frac{\mathcal{M}_{gt}^j}{\alpha^2}, \tag{4.1f}$$

where i labels the source or lens tomographic bin inside which we measure either M_a or M_g, and j and k denote the tomographic bins used to compute the three position-dependent 2PCFs $\xi_\pm, \xi_g, \xi_t$. The angle brackets denote ensemble average (or in practice, averaging over all patch positions $\boldsymbol{\theta}_C$).

The equations above write the real-space ζ in terms of their corresponding Fourier-space counterparts called the *integrated bispectra* $\mathcal{B}(\ell)$. These integrated bispectra can be expressed in terms of line-of-sight projections of 3D matter and galaxy (cross-) bispectra

---

[7]In ref. [30] the authors denoted the integrated shear 3PCF as $\zeta_\pm$. To be consistent with the notations of the other integrated 3PCFs involving the galaxy field that we introduce in this work, we denote the integrated shear 3PCF as $\zeta_{a\pm}$ with the added subscript 'a' to specify the involvement of the lensing aperture mass M_a in this statistic.

using the Limber approximation [50] as (see appendix C)[8]

$$\mathcal{B}_{a\pm}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_\kappa^i(\chi)q_\kappa^j(\chi)q_\kappa^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{mmm}^{\mathrm{3D}}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi\right) e^{2i\left(\phi_{\boldsymbol{\ell}_2} \mp \phi_{-\boldsymbol{\ell}_{12}}\right)}$$
$$\times U(\boldsymbol{\ell}_1)W(\boldsymbol{\ell}_2 + \boldsymbol{\ell})W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}), \tag{4.2a}$$

$$\mathcal{B}_{g\pm}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_g^i(\chi)q_\kappa^j(\chi)q_\kappa^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{gmm}^{\mathrm{3D}}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi\right) e^{2i\left(\phi_{\boldsymbol{\ell}_2} \mp \phi_{-\boldsymbol{\ell}_{12}}\right)}$$
$$\times W(\boldsymbol{\ell}_1)W(\boldsymbol{\ell}_2 + \boldsymbol{\ell})W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}), \tag{4.2b}$$

$$\mathcal{B}_{ag}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_\kappa^i(\chi)q_g^j(\chi)q_g^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{mgg}^{\mathrm{3D}}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi\right) U(\boldsymbol{\ell}_1)W(\boldsymbol{\ell}_2 + \boldsymbol{\ell})W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}),$$
$$\tag{4.2c}$$

$$\mathcal{B}_{gg}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_g^i(\chi)q_g^j(\chi)q_g^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{ggg}^{\mathrm{3D}}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi\right) W(\boldsymbol{\ell}_1)W(\boldsymbol{\ell}_2 + \boldsymbol{\ell})W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}),$$
$$\tag{4.2d}$$

$$\mathcal{B}_{at}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_\kappa^i(\chi)q_g^j(\chi)q_\kappa^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{mgm}^{\mathrm{3D}}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi\right) e^{2i\phi_{-\boldsymbol{\ell}_{12}}}$$
$$\times U(\boldsymbol{\ell}_1)W(\boldsymbol{\ell}_2 + \boldsymbol{\ell})W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}), \tag{4.2e}$$

$$\mathcal{B}_{gt}^{ijk}(\ell) = \int \mathrm{d}\chi \frac{q_g^i(\chi)q_g^j(\chi)q_\kappa^k(\chi)}{\chi^4} \int_{\boldsymbol{\ell}_1} \int_{\boldsymbol{\ell}_2} B_{ggm}^{\mathrm{3D}}\left(\frac{\boldsymbol{\ell}_1}{\chi}, \frac{\boldsymbol{\ell}_2}{\chi}, \frac{-\boldsymbol{\ell}_{12}}{\chi}; \chi\right) e^{2i\phi_{-\boldsymbol{\ell}_{12}}}$$
$$\times W(\boldsymbol{\ell}_1)W(\boldsymbol{\ell}_2 + \boldsymbol{\ell})W(-\boldsymbol{\ell}_{12} - \boldsymbol{\ell}). \tag{4.2f}$$

For the shear-only $\mathcal{B}_{a\pm}$, ref. [30] showed that the 3D nonlinear matter bispectrum B_{mmm}^{3D} can be modelled accurately using the response approach to perturbation theory [36]. This is the calculation we adopt here, which allows to evaluate $\zeta_{a\pm}$ down to nonlinear scales as a function of cosmological and baryonic physics parameters. For the remainder of the integrated bispectra that involve the galaxy density field, we model the corresponding bispectra at leading-order in perturbation theory; we do not display all of these expressions here, but the interested reader can find them in appendix C. In particular, the various integrated 3PCFs display different scalings of the galaxy bias terms. Concretely, $\zeta_{a\pm} \propto A_s^2$; $\zeta_{g\pm}, \zeta_{at} \propto \{b_1 A_s^2, b_2 A_s^2, b_{s^2} A_s^2\}$; $\zeta_{ag}, \zeta_{gt} \propto \{b_1^2 A_s^2, b_1 b_2 A_s^2, b_1 b_{s^2} A_s^2, b_1 A_s/\bar{n}\}$ and $\zeta_{gg} \propto \{b_1^3 A_s^2, b_1^2 b_2 A_s^2, b_1^2 b_{s^2} A_s^2, b_1^2 A_s/\bar{n}\}$; here, $\bar{n}$ denotes the mean number density of galaxies in a given tomographic bin. The different sensitivity of the 6 integrated 3PCFs to the galaxy bias and cosmological parameters relative to the 3 global 2PCFs[9] discussed in section 2 indicates that joint analyses of these statistics can help to lift parameter degeneracies, leading to tighter parameter constraints overall.

In eqs. (4.1e) and (4.1f) for ζ_{at} and ζ_{gt}, we note again the presence of point-mass contributions with amplitude $\mathcal{M}_{at}$ and $\mathcal{M}_{gt}$. These statistics involve the position-dependent

---

[8]In order to perform the numerical integrations in the predictions for $\mathcal{B}$ we use the Monte-Carlo Vegas algorithm [51]. Moreover, instead of using the inverse Hankel transform integrals directly to convert the $\mathcal{B}(\ell)$ to real space integrated 3PCFs $\zeta(\alpha)$ (and also the $\mathcal{P}(\ell)$ to real space 2PCFs $\xi(\alpha)$), we use expressions with summation over ℓ as given in ref. [52] (see their eq. (9)), which are exact in the curved-sky case and more accurate in that they take into account the finite bin widths in which the correlation functions are measured in the data.

[9]See refs. [29, 30] for detailed discussions on the different dependence of $\xi_\pm$ and $\zeta_{a\pm}$ on cosmological parameters.

tangential shear 2PCF, which is why these parameters are introduced due to the nonlocality of the tangential shear signal. The derivation of these point-mass term contributions is shown in appendix D.

We also note that the correlations involving the 1-point average galaxy density $M_g(\boldsymbol{\theta}_C)$ are susceptible to imaging systematics as they directly probe the number density of galaxies within apertures. The presence of a position-dependent systematic effect affecting the observed foreground lens galaxy number count at different locations $\boldsymbol{\theta}_C$ on a survey footprint may therefore impact the measurements of the $\zeta_{g\pm}, \zeta_{gt}, \zeta_{gg}$ statistics. We leave the investigation of the impact of such observational systematic effects in the integrated 3PCFs to future work.

5 Simulations

In this section we present the simulated data we use in order to test our theoretical models of the integrated 3PCFs discussed in the previous section. We use the publicly available cosmological simulation data from ref. [53] (hereafter referred to as the T17 simulations).[10] In our work, we use the full-sky lightcone halo catalogues and cosmic shear lensing maps of the simulation suite. The simulation data products were obtained from a gravity-only N-body simulation in a ΛCDM cosmology with the following parameters: $\Omega_{\rm cdm} = 0.233$, $\Omega_{\rm b} = 0.046$, $\Omega_{\rm m} = \Omega_{\rm cdm} + \Omega_{\rm b} = 0.279$, $\Omega_\Lambda = 0.721$, $h = 0.7$, $\sigma_8 = 0.82$ and $n_s = 0.97$. The amplitude of the primordial scalar perturbations A_s corresponding to the T17 value of σ_8 is $A_s = 2.197 \times 10^{-9}$ (we work with A_s in our paper instead of σ_8). Halos and sub-halos in the simulation were identified using the six-dimensional phase-space friends-of-friends algorithm ROCKSTAR [54]. These halo catalogues were combined in layers of shells to obtain full-sky lightcone halo catalogues. The simulation boxes were also ray traced using the multiple-lens plane ray-tracing algorithm GRAYTRIX [55, 56] to obtain weak lensing shear maps for several source redshifts between $z = 0.05$ and $z = 5.3$. We utilize 108 realizations of these data products, obtained from multiple realizations of the T17 simulations.

As discussed in refs. [53, 57], the T17 maps suffer from systematic effects associated with, for example, the thickness of the lens shells that may need to be incorporated into the theory predictions for fairer comparisons. In this work we do not perform these corrections explicitly since, as we will see below, we find good agreement between simulations and our theory predictions at DES Y3-like survey precision.

5.1 Simulated weak lensing shear maps

To create a mock shear map from the T17 simulations we use a realistic source galaxy distribution. In order to do that we combine the simulated T17 cosmic shear data products at individual source redshifts according to a source distribution $p(z)$ inspired from the DES Year 3 analysis. We use the same scheme as that used in ref. [31], in particular we consider their second tomographic bin (cf. blue dashed distribution in figure 2 of ref. [31]; this corresponds to a combination of the third and fourth DES Year 3 tomographic bins). The shear map is in Healpix [58] pixel format to which we add 5 galaxies per arcmin² to mimic the shape noise level expected for DES Year 6 (we refer the reader to ref. [29] for more details about the addition of shape noise).

---

[10]The simulation data products are at http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/.
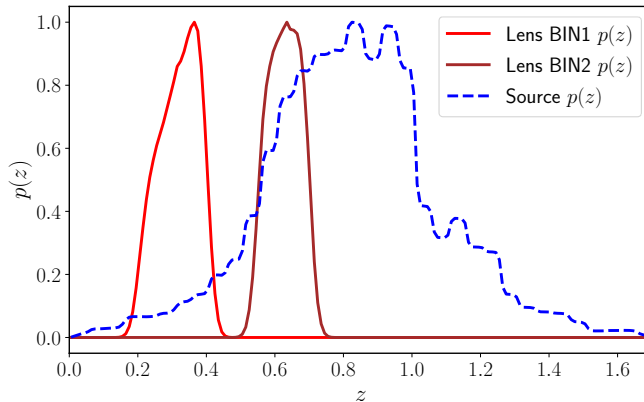
Figure 2. The distribution of lens and source galaxies used for creating our mock galaxy and shear sky maps. The red and brown curves indicate the $p(z)$ of the mock HOD lens galaxies in redshift BIN1 and BIN2, respectively. The source galaxy sample $p(z)$ is shown in blue dashed. For visualization only, the distributions are scaled to have the same maximum at unity.

5.2 Mock HOD galaxy catalogues

For the purpose of measuring the correlations on the galaxy density field we require mock galaxy catalogues. Being a gravity-only N-body simulation, the T17 suite does not come with galaxy catalogues. We hence create our own full-sky galaxy mocks by populating the T17 halo catalogues using an empirical Halo Occupation Distribution (HOD) method [59] based on the widely used halo model of large-scale structure (see ref. [60] for a review). Briefly, an HOD model describes a probability distribution $P(N_g|M_h)$ for a halo mass M_h to host N_g galaxies. We follow ref. [61] who used a 4-parameter HOD model to investigate the lens galaxy samples used in the DES data.

Concretely, the HOD model separates the contribution from central and satellite galaxies, and has the following functional forms for the mean values of the central and satellite galaxies, respectively:

$$\langle N_{\rm cen}|M_h\rangle = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\log M_h - \log M_{\min}}{\sigma_{\log M_h}}\right)\right],\tag{5.1}$$

$$\langle N_{\rm sat}|M_h\rangle = \langle N_{\rm cen}|M_h\rangle\left(\frac{M_h}{M_1}\right)^{\gamma}.\tag{5.2}$$

The first equation describes the mean number of central galaxies hosted by halos of mass M_h; $M_{\min}$ and $\sigma_{\log M_h}$ are the parameters of a smooth step-function. One can understand $M_{\min}$ as the mass at which half of the halos with this mass host a central galaxy and $\sigma_{\log M_h}$ gives the scatter around the halo mass M_h. The second equation gives the mean number of satellite galaxies within halos of mass M_h, and is parametrized by γ, a power-law index for the mass dependence of the number of satellites and M_1, the threshold mass for halos to start hosting at least one satellite. The total mean number of galaxies hosted by halos of mass M_h is

$$\langle N_g|M_h\rangle = \langle N_{\rm cen}|M_h\rangle + \langle N_{\rm sat}|M_h\rangle.\tag{5.3}$$

Table 1 lists our HOD parameters for two lens tomographic redshift bins that we use in our analyses. They are close to those in ref. [61] for their first and third MagLim sample

Lens bin	$\log_{10} M_{\min}$	$\sigma_{\log_{10} M_h}$	$\log_{10} M_1$	γ	$\bar{n}_g^{\mathrm{3D}}$ [Mpc^{-3}]	b_1	b_2	b_{s^2}
BIN1 (z=0.2–0.4)	12.40	0.2	13.40	0.65	1.2×10^{-3}	1.32	-0.52	-0.18
BIN2 (z=0.55–0.7)	12.03	0.014	13.37	0.52	0.4×10^{-3}	1.93	0.19	-0.53

Table 1. HOD model parameters used in this work for populating mock lens galaxies in the T17 halo catalogues. Halo masses are expressed in units of $M_\odot$. Also listed are the galaxy number density and bias parameter values obtained for each redshift bin.

bins (see table D2 of ref. [61]), and as a result, they result in similar galaxy number densities to the MagLim sample. For our HOD parameters, using eq. (B.2) in appendix B, we have also computed the galaxy bias parameters $b_\mathcal{O}(z)$ averaged over the redshift distributions of the lens bins. We use the halo mass function formula of ref. [62] and (i) the halo b_1 fitting function from ref. [63] to compute b_1 for the galaxies, (ii) the halo $b_2(b_1)$ fitting function from ref. [64] to compute the galaxy b_2, and (iii) the co-evolution relation $b_{s^2} = -\frac{4}{7}(b_1 - 1)$ [40, 65] to obtain the galaxy b_{s^2}. We take these values to be the fiducial bias parameters of our HOD samples; the bias values are listed in table 1.

To create the actual mock galaxy catalogues from the T17 simulation, we first combine the halo shells to obtain the halos in our two lens redshift intervals. Identifying M_h with M_{200b} (the mass enclosed inside a radius where the mean density is 200 times the background matter density), we use the HOD model described above to populate each halo in the catalogue with galaxies. For a given halo we perform a Bernoulli draw with expectation given by eq. (5.1) to get N_{cen} and a Poisson random draw with expectation given by eq. (5.2) to obtain N_{sat}. The central galaxies are placed at the halo centres, whereas the satellite galaxies are placed randomly around the halo centre following a Navarro-Frenk-White distribution [66] (this is as in ref. [67]). We note further that we restrict ourselves to using halos with masses $M_{200b} > 1.1 \times 10^{12} \, M_\odot/h$ for BIN1 and $M_{200b} > 5.1 \times 10^{12} \, M_\odot/h$ for BIN2 to remain largely unaffected by the mass resolution limit of the simulation in the respective redshift ranges (see table 1 of ref. [53] for mass-cut details).

In order to obtain smoothed looking distribution of lens galaxies in a tomographic bin as expected from photometric galaxy imaging surveys like DES, we first populate HOD galaxies in halos within and beyond the boundaries of the desired redshift range of the tomographic bin. To every true redshift z_{true} of these simulated HOD galaxies, we associate a mock 'observed' galaxy redshift $z_{\mathrm{obs}} = z_{\mathrm{true}} + \delta z$ where δz is drawn from a Gaussian with mean 0 and standard deviation $\sigma_z = 0.02$ to mimic a photometric uncertainty in the galaxy's redshift [68, 69]. We then select those galaxies whose z_{obs} fall within the tomographic bin's redshift range and use their corresponding z_{true} values to obtain the distribution. The shapes of the mock galaxy distributions in the two redshift bins are shown in figure 2. Finally, we project the galaxy catalogues to 2D grids in `Healpix` format with `NSIDE` = 2048 to obtain galaxy number counts maps, which we use to measure the auto- and cross-correlations of the galaxy and cosmic shear fields.

6 Measurements and data covariance

We measure the galaxy and shear correlations on 6 non-overlapping 5000 deg^2 circular footprints carved from each all-sky T17 galaxy/shear map; the footprint area is chosen to be representative of a DES-like survey. Over the 108 T17 realizations, this results in a total of 648 DES-sized galaxy and shear maps that we use to obtain our mean data vector and to estimate its covariance.

We follow the same measurement strategy (using the public code `TreeCorr` [43]) as ref. [29] for the shear-only $\zeta_{a\pm}$ statistic. The position-dependent 2PCFs[11] $\hat{\xi}(\alpha; \boldsymbol{\theta}_C)$ (cf. figure 1) are measured on the shear and galaxy density maps within top-hat windows W with radius $\theta_T = 130$ arcmin in 15 log-spaced angular bins within the range $\alpha \in [5, 250]$ arcmin. The 1-point lensing aperture mass $\hat{M}_a(\boldsymbol{\theta}_C)$ at location $\boldsymbol{\theta}_C$ is estimated using shear measurements within a compensated window Q with an aperture scale $\theta_{ap} = 130$ arcmin (visualized as blue apertures in figure 1). The 1-point average galaxy density contrast $\hat{M}_g(\boldsymbol{\theta}_C)$ on the other hand is measured by taking the mean of the pixel values of the foreground lens galaxy density map within the same top-hat window W where the $\hat{\xi}(\alpha; \boldsymbol{\theta}_C)$ are measured (visualized as red apertures in figure 1). In each 5000 deg^2 footprint, these apertures are distributed to cover the whole area with only slight overlap between adjacent patches resulting in order 1000 patches across the footprint. The integrated 3PCFs are then estimated as

$$\hat{\zeta}_{xy}(\alpha) = \frac{1}{N_p} \sum_{i=1}^{N_p} \hat{M}_x(\boldsymbol{\theta}_{C,i}) \, \hat{\xi}_y(\alpha, \boldsymbol{\theta}_{C,i}), \tag{6.1}$$

where $x \in \{a, g\}$, $y \in \{\pm, t, g\}$ and the sum runs over the N_p patches centered at $\boldsymbol{\theta}_{C,i}$. On the other hand, the global 2PCFs $\hat{\xi}(\alpha)$ in a given footprint are estimated with `TreeCorr` by using all the pixel values for shear or the galaxy density contrast within the entire footprint.

The full mean data vector is then obtained as the average over the estimates from the 648 mock footprints. For the case of a single lens and single source redshift bin, the data vector consists of the following correlations:

$$d \equiv \{ \underbrace{\xi_+, \xi_-, \zeta_{a+}, \zeta_{a-}}_{\text{shear-only}}, \underbrace{\xi_g, \zeta_{gg}}_{\text{galaxy-only}}, \underbrace{\xi_t, \zeta_{at}, \zeta_{ag}, \zeta_{g+}, \zeta_{g-}, \zeta_{gt}}_{\text{galaxy-shear cross-correlations}} \}. \tag{6.2}$$

$$\underbrace{\phantom{\xi_g, \zeta_{gg}, \xi_t, \zeta_{at}, \zeta_{ag}, \zeta_{g+}, \zeta_{g-}, \zeta_{gt}}}_{\text{galaxy correlations}}$$

To aid in our discussions below, we organize the data vector into different types of contributions. The first 4 components are the *cosmic shear-only* $\xi_\pm$ and $\zeta_{a\pm}$ correlations; the 5th and 6th components are the *galaxy-only* correlations ξ_g, ζ_{gg}; the remaining components correspond to the *galaxy-shear cross-correlations*. We denote the galaxy-only and the galaxy-shear cross-correlations together as *galaxy correlations* indicating that they involve the galaxy field. Accounting for our other lens redshift bin results in additional galaxy correlation terms.

The data covariance matrix is estimated from the mocks as

$$\hat{\mathbf{C}} = \frac{1}{N_s - 1} \sum_{i=1}^{N_s} (\hat{d}_i - \hat{d})(\hat{d}_i - \hat{d})^{\mathrm{T}}, \tag{6.3}$$

where $N_s = 648$ is the number of mock footprints, $\hat{d}_i$ is the data vector measured in the i-th footprint, and $\hat{d}$ the sample mean over the N_s realizations. To get an unbiased estimate of the inverse covariance matrix we apply the correction [70]

$$\mathbf{C}^{-1} = \frac{N_s - N_d - 2}{N_s - 1} \hat{\mathbf{C}}^{-1}, \tag{6.4}$$

where N_d is the length of the data vector. For our two lens bins and single source bin we have 20 components in d, each with 15 data points making a total of 300 data points before the application of any scale cuts. All of the components of our data vector are shown by the black points with error bars in figure 3; the error bars shown are the square root of the diagonal of the covariance matrix.

---

[11] The hat in $\hat{\xi}(\alpha; \boldsymbol{\theta}_C)$ indicates that this is an estimate of the corresponding statistic from data.

7 Results

In this section we present our main numerical results for the modelling and cosmological constraining power of the integrated 3PCFs. First, we discuss the regime of validity of our perturbation theory model for the galaxy correlations, i.e. $\xi_t, \xi_g, \zeta_{at}, \zeta_{ag}, \zeta_{g\pm}, \zeta_{gt}$. Then, we demonstrate the ability of our galaxy correlation models to recover unbiased constraints on A_s, as well as on the bias parameters b_1 and b_2 of the mock galaxy samples through an MCMC likelihood analysis. Finally, we present Fisher forecast results where we investigate the constraining power of a joint 3×2PCF and integrated 3PCF analysis using all the correlations i.e., both cosmic shear-only ($\xi_\pm$, $\zeta_{a\pm}$) as well as galaxy correlations.

7.1 Comparison of theoretical models to measurements from simulations

Figure 3 compares the components of our data vector in eq. (6.2) (black dots) with our theory model predictions from sections 2 and 4 evaluated at the fiducial cosmology and bias parameters (red curves).

The first row shows the cosmic shear-only correlations, namely the global shear 2PCFs $\xi_\pm$ and the integrated shear-only 3PCFs $\zeta_{a\pm}$; recall these are evaluated using the nonlinear matter power spectrum from the `HMCODE` and the *response approach* bispectrum model from ref. [30], respectively. As previously discussed in ref. [30], the theoretical models for these shear-only correlations are in excellent agreement with the simulation measurements on all angular scales probed. Further, our model allows to readily incorporate the impact of baryonic feedback on small scales. We follow the strategy of ref. [30] (see their section 4.2) to determine scale cuts to remove scales that are expected to be severely affected by baryonic physics. In figure 3, these are shown by the purple dashed vertical lines in the first row of panels; concretely, the data to the right of these lines are deemed to be unaffected by baryonic feedback.

The second and third row of panels in figure 3 show the galaxy correlations for Lens BIN1, i.e. the ξ_t, ξ_g 2PCFs and the $\zeta_{at}, \zeta_{ag}, \zeta_{g\pm}, \zeta_{gt}, \zeta_{gg}$ integrated 3PCFs. The first and second panels in the second row show $\xi_t(\alpha)$ and $\xi_g(\alpha)$. For ξ_t, our tree-level perturbation theory model is in good agreement with the simulation results on only large angular scales α, whereas for ξ_g, the tree-level model displays a good fit down to comparatively smaller angular scales. This is as expected since at a given α, ξ_t is more sensitive to larger multipoles ℓ (smaller nonlinear scales) compared to ξ_g; this is because the former's J_2 Bessel function in the Fourier- to real-space conversion (cf. eqs. (2.10), (2.9)) weights the nonlinear scales more than the J_0 function of the latter. Thus, ξ_t gets more contributions from scales where our perturbation theory model breaks down, hence the poorer agreement between theory and simulations in the figure.

The third and fourth panels in the second row show the integrated 3PCFs $\zeta_{at}(\alpha)$ and $\zeta_{ag}(\alpha)$, whereas the panels in the third row are for the integrated 3PCFs $\zeta_{g+}, \zeta_{g-}, \zeta_{gt}, \zeta_{gg}$. The galaxy correlations ζ show similar trends as ξ_t, ξ_g: the models agree with the simulations on large angular scales, but become discrepant on smaller scales where the tree-level models break down (this can be similarly understood in terms of the J_n weightings of each statistic). We note that in our fiducial predictions for $\xi_t, \zeta_{at}, \zeta_{gt}$ we set the corresponding point-mass terms $\mathcal{M}$ to zero. The values of these parameters cannot be predicted from first principles (they capture a complicated interplay of higher-order galaxy bias, stochastic terms and nonlinear matter fluctuations), and so in our MCMC validation analysis below we will treat them as free parameters.
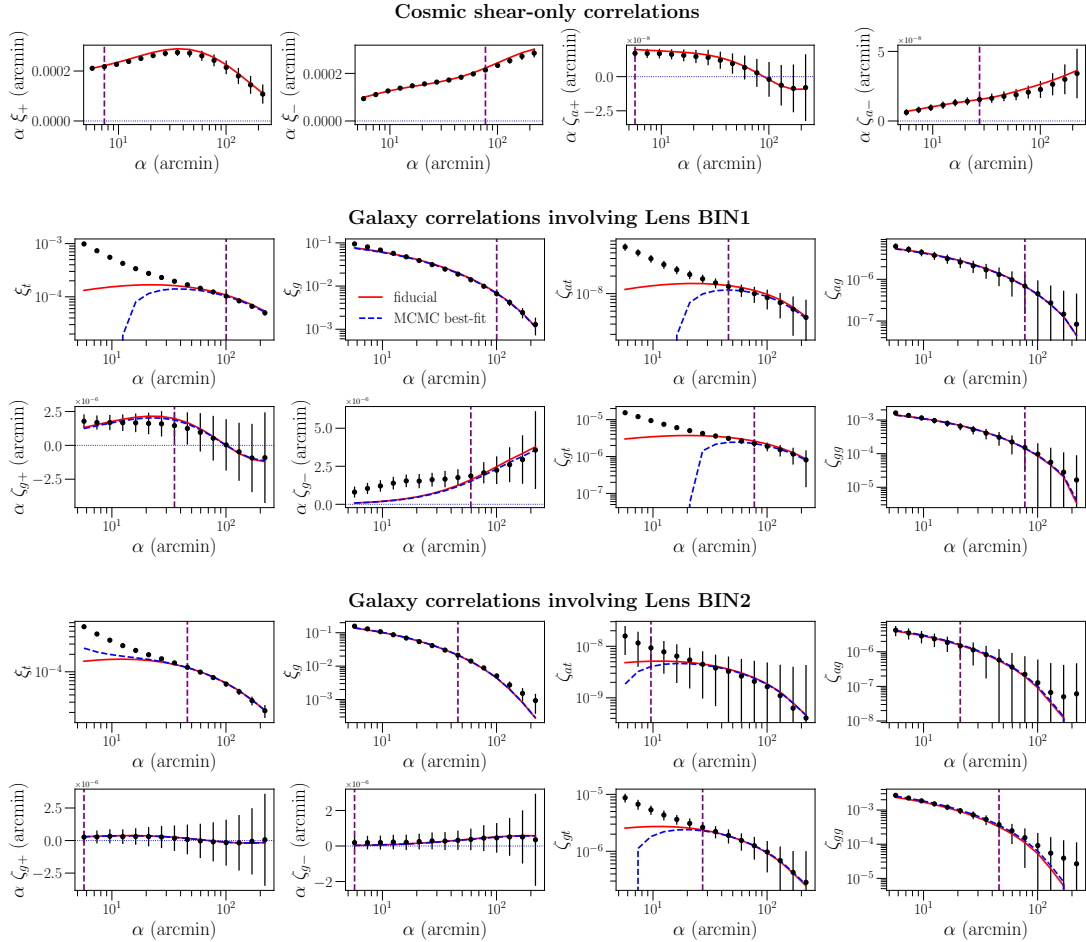
Figure 3. *First row*: the shear-only 2PCF $\xi_\pm$ and the integrated 3PCF $\zeta_{a\pm}$ computed from the T17 mock source tomographic bin. The black dots with the error bars show the mean and the standard deviation of the measurements from our 648 T17 mocks, respectively. The red curves show the model predictions using the fully nonlinear theoretical recipes for these shear correlations computed at the fiducial cosmological parameters of the simulations. The purple dashed vertical lines show our scale cuts to remove scales affected by baryonic feedback effects. *Second and third rows*: the galaxy 2PCFs ξ_t, ξ_g and the integrated galaxy 3PCFs $\zeta_{at}, \zeta_{ag}, \zeta_{g\pm}, \zeta_{gt}, \zeta_{gg}$ involving galaxy correlations and cross-shear correlations computed with galaxies in Lens BIN1. The red curves show the tree-level perturbation theory models computed using the fiducial cosmological parameters of the simulations and the fiducial galaxy bias parameter values evaluated using the HOD approach for the lens galaxy sample. The purple dashed vertical lines show our conservative scale cuts to remove scales where our tree-level perturbation theory model breaks down; scales below these cuts are not included in the MCMC and Fisher forecast analyses. The blue dashed curves show the theory predictions computed at the maximum posterior of the MCMC analysis performed in section 7.2. *Fourth and fifth rows*: same as the second and third rows but for Lens BIN2 instead of BIN1.

We conservatively estimate the regime of validity of our perturbation theory models for the galaxy correlations as follows. We compute the theoretical predictions for two scenarios: (i) using the fiducial tree-level model $d_{\rm tree}$ and (ii) another model $d_{k_{\rm NL}}$ where we

artificially set all Fourier modes k larger than the non-linear scale $k_{\rm NL}$ to zero.[12] Using these predictions, we progressively discard small angular scales from our data vector until $\chi^2 \equiv (d_{\rm tree} - d_{k_{\rm NL}})^{\rm T} \mathbf{C}^{-1} (d_{\rm tree} - d_{k_{\rm NL}}) < 0.3$ is satisfied. This roughly identifies a minimum scale $\alpha_{\rm min}$ below which our perturbation theory model begins to fail significantly; the resulting scale cuts are marked by the purple dashed vertical lines in the galaxy correlations panels.[13] We note that our criteria for determining these scale cuts are conservative. In particular, allowing the point-mass terms to vary, which we currently set to zero, could allow for greater reach down to smaller angular scales in ξ_t, ζ_{at}, and ζ_{gt}. This could in turn enable the use of higher signal-to-noise data points. As our primary aim is to explore the first-order information gain that the galaxy correlations can already bring from scales where leading-order PT models are valid, we choose to adopt these conservative scale cuts and defer the investigation of more accurate models on smaller angular scales to future works.

Finally, similar considerations hold for the fourth and fifth rows of panels in figure 3, which show the same as the second and third rows, but for the Lens BIN2. We note only that the error bars in some of the panels are larger than those for BIN1, which is as expected by the fact that the number density of galaxies in BIN2 is approximately 3 times smaller than that of BIN 1 (cf. table 1).

## 7.2 MCMC validation of the galaxy correlations modelling

Using the measured galaxy correlations with the scale cuts, we now test whether our tree-level perturbation theory models can correctly recover the fiducial parameters in an MCMC constraint analysis; we do not consider the shear-only correlations in this section to focus on the galaxy correlations modelling. We investigate in particular the constraints on the galaxy bias parameters b_1 and b_2, as well as the cosmological parameter A_s.[14] We assume a Gaussian likelihood function and wide uniform priors for the parameters varied: in addition to A_s and the b_1, b_2 of the two samples, we also vary the point-mass $\mathcal{M}$ amplitude parameters that contribute to the $\xi_t, \zeta_{at}, \zeta_{gt}$ statistics. We do not sample b_{s^2} but let it vary according to the co-evolution relation for $b_{s^2}(b_1)$. We use the publicly available affine sampler emcee [71] to perform the MCMC analysis. The results are shown in the contour plot in figure 4. The figure is for the result marginalized over the point-mass terms, but see figure 8 in appendix D for the constraints on all varied parameters.

The figure shows that our tree-level models for ξ_t, ξ_g, which are only sensitive to A_s and b_1, correctly recover the fiducial values of the parameters within the 68% credible intervals (blue dashed contours). When considering the galaxy integrated 3PCFs (red contours), the constraints remain unbiased, but they become visibly tighter. This shows that our conservative scale cuts are adequate to return unbiased results, while still letting our tree-level models explore the non-Gaussian information in the galaxy correlations to improve the parameter constraints. Relative to the galaxy 2PCFs constraints (blue), the addition of the galaxy integrated 3PCFs (red) improves the constraints on $\ln(10^{10}A_s)$, $b_1^{\rm BIN1}$, $b_1^{\rm BIN2}$ by approximately 30%, 20%, 25%, respectively. These improvements are associated to the breaking of degeneracies between A_s and b_1 in the galaxy 2PCFs by the galaxy integrated

---

[12]The nonlinear scale is defined implicitly as $k_{\rm NL}^3 P_{\rm lin}^{\rm 3D}(k_{\rm NL}, z)/(2\pi)^3 = 1$.

[13]Note these scale cuts are not the same as those applied on the shear-only statistics, which ensure instead that the scales are not affected by baryonic feedback effects. We assume also that perturbation theory breaks down on scales larger than the scales where baryonic effects are important, as is reasonable.

[14]We take A_s as the only cosmological parameter in the MCMC results as predictions for it are rapid to obtain. The extension to other cosmological parameters would require the construction of dedicated emulators for fast theory predictions (e.g. see ref. [31] for an emulator for $\zeta_{a\pm}$), which is beyond the scope of this work.
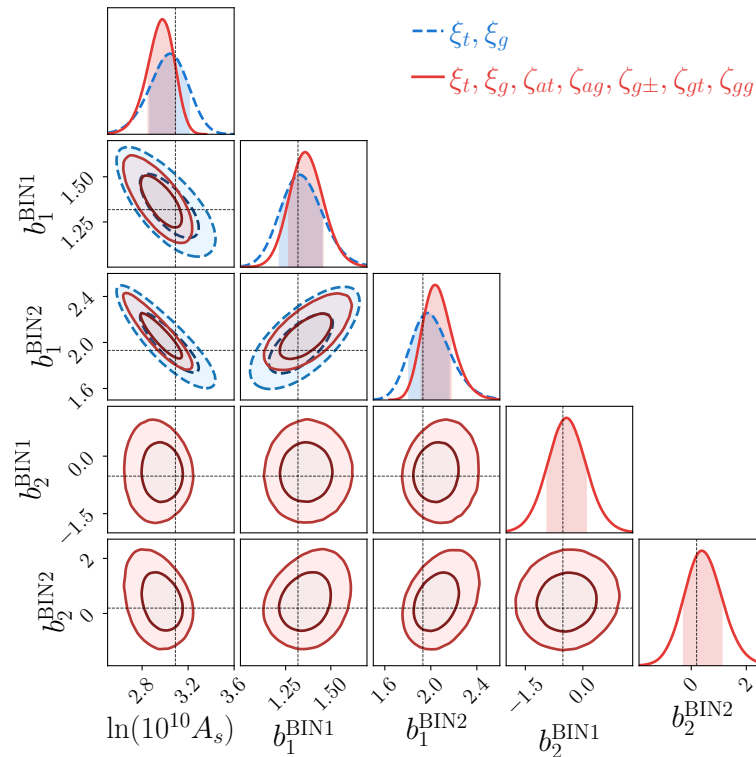
Figure 4. Marginalized one- and two-dimensional MCMC constraints on the parameters $\ln(10^{10}A_s)$, b_1^{BIN1}, b_1^{BIN2}, b_2^{BIN1}, b_2^{BIN2} obtained from the galaxy 2PCFs ξ_t, ξ_g (blue dashed) and considering also the galaxy integrated 3PCFs, $\xi_t, \xi_g, \zeta_{at}, \zeta_{ag}, \zeta_{g\pm}, \zeta_{gt}, \zeta_{gg}$ (red solid). We consider the average T17 measurements as the data vector with the scale cuts discussed in section 7.1. The fiducial values of the parameters are marked by the black dashed lines. Note that at tree-level, the galaxy 2PCFs cannot constrain the second-order b_2 parameters. The point-mass terms associated with tangential shear correlations are marginalized over; see figure 8 in appendix D for their constraints. All other parameters are fixed to the fiducial values of the simulation and the tidal bias terms are varied according to the co-evolution $b_{s^2}(b_1)$ relation.

3PCFs. The galaxy integrated 3PCFs can also constrain the b_2 parameter, which is not possible with ξ_t, ξ_g at tree-level.

The predictions of the tree-level models using the best-fitting parameters from this MCMC analysis are shown by the blue dashed curves in figure 3. As expected, they agree with the predictions for the fiducial parameters (red curves) on scales larger than our assumed scale cuts.

7.3 Fisher forecasts for a DES-Y3-like survey

We now investigate in the context of Fisher matrix forecasts the ability of combined 2PCFs and integrated 3PCFs to constrain cosmological parameters; from hereon we consider also the cosmic shear-only 2PCFs $\xi_\pm$ and integrated 3PCFs $\zeta_{a\pm}$. The Fisher information matrix $\mathbf{F}$ for a model vector M depending on parameters π, assuming a constant data-covariance

C, is given by [72]

$$F_{ij} = \left(\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_i}\right)^{\mathrm{T}} \mathbf{C}^{-1} \left(\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_j}\right), \tag{7.1}$$

where F_{ij} is an element of the matrix $\mathbf{F}$ associated with the parameters π_i and π_j, and $\mathbf{C}^{-1}$ is the inverse data covariance matrix in eq. (6.4). The partial derivative of the model vector with respect to the parameter π_i can be computed using a 2-point central difference:

$$\frac{\partial M(\boldsymbol{\pi})}{\partial \pi_i} = \frac{M(\pi_i + \delta_i) - M(\pi_i - \delta_i)}{2\delta_i}, \tag{7.2}$$

where δ_i is a small change in the parameter π_i around its fiducial value, and $M(\pi_i \pm \delta_i)$ is the model vector computed at the changed parameter $\pi_i \pm \delta_i$ with all the other parameters fixed. We consider the following cosmological[15] and baryonic feedback parameters $\boldsymbol{\pi}_{\mathrm{cosmo}} = \{\Omega_{\mathrm{cdm}}, \ln(10^{10}A_s), w_0, w_a, h, c_{\min}\}$, where w_0, w_a are the dynamical dark energy equation of state parameters (in the CPL parametrization $w(z) = w_0 + w_a z/(1 + z)$ [74]) and $c_{\min}$ is a baryonic feedback parameter of the HMCODE [48] nonlinear matter power spectrum which enters in our modelling of the cosmic shear-only statistics $\xi_\pm$ and $\zeta_{a\pm}$. The fiducial values are $\boldsymbol{\pi}_0 = \{0.233, 3.089, -1, 0.0, 0.7, 3.13\}$; the cosmological parameters are the same as the T17 simulations, and the baryonic parameter is the gravity-only value as determined by ref. [48]. When varying the Ω_{cdm} parameter we keep the baryon density Ω_{b} fixed, but adjust the dark energy density to keep the universe spatially flat. In addition, we also vary the galaxy bias b_1, b_2 and point-mass $\mathcal{M}$ parameters for both lens samples. The fiducial values of the galaxy bias terms are given in table 1; the point-mass term fiducial values are assumed to be zero in our analysis. When varying b_1 we evaluate the tidal bias terms b_{s^2} according to the co-evolution relation $b_{s^2}(b_1)$.

The parameter covariance matrix $\mathbf{C}_{\boldsymbol{\pi}}$ is given by the inverse of the Fisher matrix

$$\mathbf{C}_{\boldsymbol{\pi}} = \mathbf{F}^{-1}, \tag{7.3}$$

which we use to forecast constraints on the model parameters. In our results below, we report the Fisher constraints on the cosmological parameters, baryonic feedback parameter and linear bias parameter b_1, marginalizing over the second-order bias b_2 and point-mass terms. We present results for three different combinations of 2- and 3-point statistics: (i) a 3×2PCF-only analysis, labelled as $\{\xi_\pm, \xi_t, \xi_g\}$ and shown in blue colour; (ii) the same, but adding the shear-only integrated 3PCF, labelled as $\{\xi_\pm, \xi_t, \xi_g, \zeta_{a\pm}\}$ and shown in green; and (iii) using all of the statistics discussed in this paper combined, labelled as *all correlations* and shown in red. For each of these, we discuss three analysis setup cases:

- *Case A*: constraints on the parameters $\{\Omega_{\mathrm{cdm}}, \ln(10^{10}A_s), w_0, h, b_1^{\mathrm{BIN1}}, b_1^{\mathrm{BIN2}}\}$, assuming scale cuts on all statistics (cf. vertical lines in figure 3). Recall that the scale cuts on the cosmic shear-only statistics $\xi_\pm, \zeta_{a\pm}$ are the ones deemed as baryon-free, and so we keep $c_{\min}$ fixed to the fiducial value (we also fix w_a). The 1σ marginalized Fisher constraints are shown in figure 5. Relative to the 3×2PCF constraints (blue), the

---

[15] Instead of A_s, we could have alternatively opted to quantify the amplitude of matter fluctuations with the parameter σ_8 (or $S_8 = \sigma_8\sqrt{\Omega_{\mathrm{m}}/0.3}$), as is most commonly done in the weak lensing literature. We opted for A_s since, unlike σ_8, it is not affected by h in its definition. By virtue of this, A_s allows to more unambiguously read out constraints on the amplitude of fluctuations (see ref. [73] for a discussion). We have explicitly checked nonetheless in a separate Fisher analysis that our main qualitative and quantitative conclusions remain the same if the parameter $\ln(10^{10}A_s)$ is replaced by S_8.
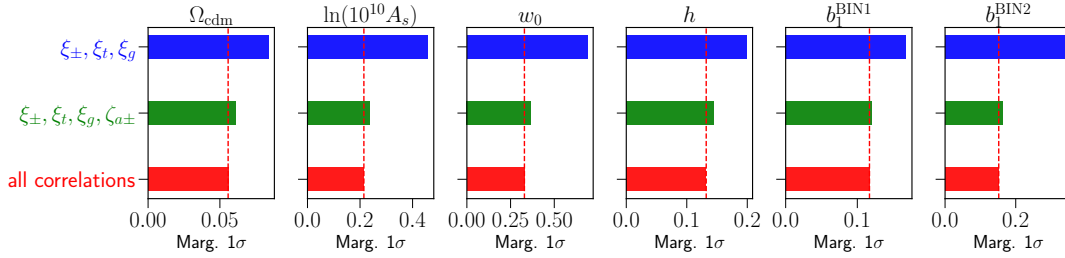
Figure 5. Marginalized 1σ Fisher constraints for analysis setup case A: constraints on the parameters $\{\Omega_{\mathrm{cdm}}, \ln(10^{10}A_s), w_0, h, b_1^{\mathrm{BIN1}}, b_1^{\mathrm{BIN2}}\}$ with scale cuts on all of the statistics. The columns are for the different parameters and the rows for different combinations of the 2PCFs and integrated 3PCFs. The second-order bias parameters b_2 and point-mass terms are marginalized over. The red dashed vertical lines serve as a guide to the eye for comparing the constraints from *all correlations* with those from other combinations of ξ and ζ.
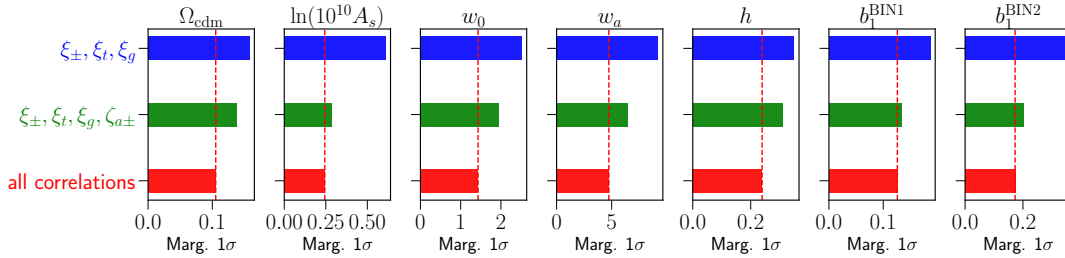
Figure 6. Same as figure 5 but for analysis setup case B, in which w_a is also a free parameter.

addition of the shear-only integrated 3PCF $\zeta_{a\pm}$ (green) improves the constraints on $\{\Omega_{\mathrm{cdm}}, \ln(10^{10}A_s), w_0, h\}$ by $\{27, 48, 47, 27\}\%$, respectively. The addition of all other integrated 3PCFs involving the galaxy density field leads to further $\{9, 10, 10, 9\}\%$ improvements (red vs. green). The constraints on the bias parameters $\{b_1^{\mathrm{BIN1}}, b_1^{\mathrm{BIN2}}\}$ are also improved by the integrated 3PCFs: $\{28, 52\}\%$ from the blue to the green and $\{3, 7\}\%$ from the green to the red constraints.

- *Case B* is the same as case A, but with w_a treated as a free parameter; the constraints are shown in figure 6. Relative to the $3\times$2PCFs, $\zeta_{a\pm}$ leads to improvements of $\{13, 53, 22, 29, 11\}\%$ on $\{\Omega_{\mathrm{cdm}}, \ln(10^{10}A_s), w_0, w_a, h\}$ (green vs. blue). The addition of all other galaxy integrated 3PCF correlations leads to further improvements of $\{24, 15, 27, 27, 23\}\%$. The same improvements for $\{b_1^{\mathrm{BIN1}}, b_1^{\mathrm{BIN2}}\}$ are $\{28, 42\}\%$ (green vs. blue) and $\{6, 14\}\%$ (red vs. green).

- *Case C* is the same as case A, but without any scale cuts on the shear-only statistics $\xi_\pm, \zeta_{a\pm}$ (we still apply the scale cuts on the galaxy correlations). Hence, in this case we also vary the baryonic feedback parameter c_{min}. We fix w_a again to its fiducial value. The results are shown in figure 7. The $\zeta_{a\pm}$ statistic improves the constraints on $\{\Omega_{\mathrm{cdm}}, \ln(10^{10}A_s), w_0, h, c_{\mathrm{min}}\}$ by $\{34, 47, 47, 34, 27\}\%$, relative to $3\times$2PCFs (green vs. blue). The rest of the galaxy correlations can further improve these by $\{8, 10, 10, 9, 2\}\%$ (red vs. green). The same improvements for $\{b_1^{\mathrm{BIN1}}, b_1^{\mathrm{BIN2}}\}$ are $\{25, 48\}\%$ (green vs. blue) and $\{3, 6\}\%$ (red vs. green).
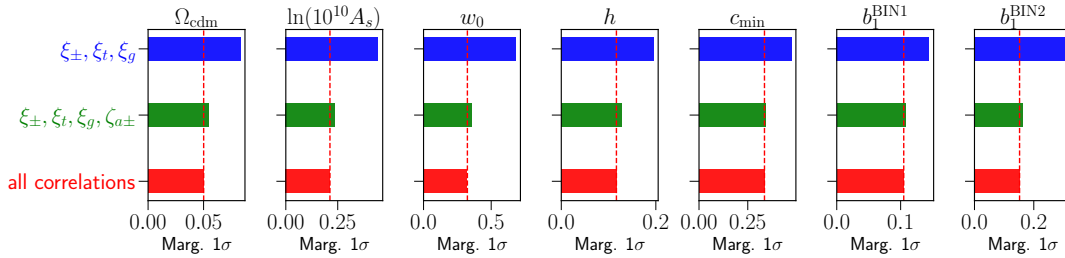
Figure 7. Same as figure 5 but for analysis setup case C: no scale cuts on the shear-only statistics $\xi_\pm, \zeta_{a\pm}$ (conservative scale cuts are still imposed on the galaxy correlations) and $c_{\min}$ is a free parameter.

The main takeaway from figures 5, 6, 7 is that the bulk of the improvements from adding integrated 3PCF information to 3×2PCF-only analyses comes from the addition of the cosmic shear-only integrated 3PCF $\zeta_{a\pm}$ ($20-40\%$; green bars), with the remainder of the integrated 3PCFs involving the galaxy density field adding smaller, but still significant improvements of approximately 10% (red bars). The smaller improvements by these galaxy correlations must be interpreted however in light of the conservative range of scales we assumed for them; higher-order perturbation theory calculations of the galaxy-matter bispectrum valid on smaller scales may result in further information gains in cosmological parameter constraints even when marginalizing over larger number of bias parameters (see ref. [75]). In addition, even for our conservative scale cuts, these galaxy correlations show the potential to already lead to improvements of up to 20% in parameter constraints when w_a is included as a free parameter. Besides cosmology, another advantage of the galaxy integrated 3PCFs is that they can tighten constraints not only on linear bias, but also higher-order galaxy bias parameters.

We note further that the improvements reported here are also tied to other analysis setup choices such as the parameter space, number of lens and source tomographic bins and galaxy number density. For example, the precise numbers may change when increasing the number of tomographic bins in the analysis. In any case, we expect that the addition of the integrated 3PCFs will always help to lift parameter degeneracies present at the 2PCF level and lead generically to improved constraints.

8 Summary and conclusion

The integrated 3-point function is a practical statistic that measures the correlation between the local 2-point correlation function and 1-point averages in patches of the survey, and which probes the squeezed-limit of the full 3-point function. In refs. [29–31] this formalism has been developed for the case of the cosmic shear field, where the statistic is known as the integrated shear 3-point correlation function $\zeta_{a\pm}$. In this paper, we have extended the formalism to include the foreground galaxy distribution and its cross-correlations with the shear field, which results in 5 new integrated 3PCFs, $\{\zeta_{at}, \zeta_{ag}, \zeta_{g\pm}, \zeta_{gt}, \zeta_{gg}\}$. In total, these 6 statistics describe the correlation between (i) three position-dependent 2PCFs, namely cosmic shear 2PCF $\xi_\pm$, tangential shear 2PCF ξ_t and galaxy clustering 2PCF ξ_g, and (ii) two average 1-point statistics, namely the lensing aperture mass M_a and the average foreground galaxy density M_g (see figure 1 for an illustration). This forms a set of higher-order galaxy and shear statistics that can be readily measured from survey data, and thus be used to improve cosmological parameter constraints relative to standard analyses based on 2PCFs alone.

The main objectives of our work were to:

1. Introduce the integrated 3PCFs involving the galaxy and cosmic shear fields, and put forward theory model predictions to evaluate them (section 4 and appendix C). For the shear-only $\zeta_{a\pm}$ statistic we use the response approach to perturbation theory which is accurate in the nonlinear regime. For all other statistics involving the galaxy density field we use tree-level perturbation theory.

2. Identify the regime of validity of the theory predictions against measurements from realistic mock galaxy and lensing simulated data (see sections 5, 6 and 7.2). We considered a DES Y3-like setup with two foreground lens galaxy bins and a single source lensing bin.

3. Investigate the improvement in cosmological, baryonic and galaxy bias parameter constraints from adding the integrated 3PCFs to standard analyses based on global 3×2PCFs.

Our main results can be summarized as follows:

- Concerning the integrated 3PCFs involving the galaxy density field that are new to this work, we find that even when restricting to large angular scales (with conservative scale cuts) where our tree-level perturbation theory models are valid (cf. figure 3), these higher-order statistics can still lead to improvements in parameter constraints. In an MCMC constraint analysis on the cosmological parameter A_s and the galaxy bias parameters b_1 and b_2, we found that the corresponding fiducial parameter values can be recovered within 68% credible intervals and the addition of the integrated 3PCFs $\zeta_{at}, \zeta_{ag}, \zeta_{g\pm}, \zeta_{gt}, \zeta_{gg}$, could lead to $20 - 30\%$ improvements over the constraints from the galaxy 2PCFs ξ_t, ξ_g.

- Using Fisher matrix forecasts for a DES-Y3-like survey, we find that the addition of the shear-only integrated 3PCF $\zeta_{a\pm}$ can lead to $20 - 40\%$ improvements on the constraints of parameters like $\Omega_{\rm cdm}, \ln(10^{10} A_s), w_0, h, b_1$, relative to the standard analysis with the 3×2PCFs alone (cf. green vs. blue in figures 5 and 7).

- The addition of the remainder integrated 3PCFs involving the galaxy density field, even when restricted to conservatively large scales, can further improve the constraints by $\sim 10\%$ (cf. red vs. green in figures 5 and 7). These improvements depend however on the specific analysis setup: for example, in constraints where the dynamical dark energy parameter w_a is free, these improvements can become $\sim 15 - 25\%$ (cf. figure 6).

These results are encouraging and motivate further developments on the theory modelling front. This includes more accurate modelling of the galaxy-matter bispectrum on smaller scales (e.g. one-loop bispectrum [75]) to utilize the higher signal-to-noise ζ data points currently excluded in our analysis due to conservative scale cuts, modelling of redshift space distortions [76] and wide-angle effects [77], as well as observational systematic effects such as galaxy intrinsic alignments [78], photometric redshift uncertainty and shear calibration bias. In particular, it would be interesting to generalize the work of ref. [31] on the integrated shear 3PCF, who investigated the optimal size of apertures for measuring $\zeta_{a\pm}$ and developed an end-to-end $\zeta_{a\pm}$ MCMC analysis pipeline, to the case of the integrated 3PCFs involving the galaxy density field. Jointly analysing the 3×2PCFs and the integrated 3PCFs

in cases beyond the standard cosmological parameters such as in searches for primordial non-Gaussianity and massive neutrinos using galaxy imaging and CMB lensing surveys would also be interesting avenues to explore in future works.

Overall, our results indicate that there is important cosmological information in integrated 3-point correlation functions involving the galaxy field and its cross-correlations with cosmic shear. Crucially, these statistics can be straightforwardly measured using existing and well-tested estimators for 1- and 2-point statistics, enabling the exploration of 3-point function information in current galaxy imaging surveys such as DES, and in the future using Euclid and Vera Rubin LSST data.

Acknowledgments

Data availability. The numerical data underlying the analysis of this paper may be shared upon reasonable request to the authors.

A Projected galaxy and weak lensing fields

In this appendix we outline the background behind the cosmic shear and the projected galaxy density contrast fields we consider in this work.

A.1 Tangential and cross components of the shear field

The complex weak lensing shear $\gamma(\boldsymbol{\theta}) = \gamma_1(\boldsymbol{\theta}) + i\gamma_2(\boldsymbol{\theta})$ has components γ_1 and γ_2 specified in a given Cartesian frame. However, one is free to rotate the coordinates by any arbitrary angle β. With respect to this reference rotation angle β, one defines the *rotated shear*

$$\gamma_\beta(\boldsymbol{\theta}) \equiv -e^{-2i\beta}\gamma(\boldsymbol{\theta}) = -e^{-2i\beta}[\gamma_1(\boldsymbol{\theta}) + i\gamma_2(\boldsymbol{\theta})], \tag{A.1}$$

and correspondingly the *tangential* (t) and *cross* ($\times$) components of the shear at position $\boldsymbol{\theta}$ w.r.t. the reference rotation angle β as

$$\gamma_\beta(\boldsymbol{\theta}) \equiv \gamma_t(\boldsymbol{\theta}, \beta) + i\gamma_\times(\boldsymbol{\theta}, \beta) = -e^{-2i\beta}[\gamma_1(\boldsymbol{\theta}) + i\gamma_2(\boldsymbol{\theta})]. \tag{A.2}$$

In particular, given a pair of points $\boldsymbol{\theta}$ and $\boldsymbol{\vartheta}$ on the sky separated by $\boldsymbol{\alpha} \equiv \boldsymbol{\vartheta} - \boldsymbol{\theta}$, one can write the tangential and cross components of the shear along $\beta = \phi_{\boldsymbol{\alpha}}$ (where $\phi_{\boldsymbol{\alpha}}$ is polar angle of $\boldsymbol{\alpha}$).

A.2 Projected galaxy number density contrast field

The projected number of galaxies $N(\boldsymbol{\theta})$ at position $\boldsymbol{\theta}$ can be written as a line-of-sight projection of the three-dimensional galaxy number density $n_g^{3D}(\boldsymbol{x}, \tau)$ along the comoving radial coordinate χ:

$$N(\boldsymbol{\theta}) = \int d\chi \, \frac{dV}{d\chi} \, n_g^{3D}(\boldsymbol{x}, \tau) \equiv \bar{N} \left[1 + \delta_g^{2D}(\boldsymbol{\theta}) \right], \tag{A.3}$$

where $\boldsymbol{x} = (\chi\boldsymbol{\theta}, \chi)$, $\frac{dV}{d\chi}(\chi)$ is the cosmological volume element (which is survey specific and for example for the whole spherical sky is $\frac{dV}{d\chi}(\chi) = 4\pi\chi^2$) and $\bar{N} \equiv \int d\chi \, \frac{dV}{d\chi} \, \bar{n}_g^{3D}(\tau(\chi))$ is the average number count of galaxies. The projected galaxy number density contrast field $\delta_g^{2D}(\boldsymbol{\theta})$ is defined as

$$\delta_g^{2D}(\boldsymbol{\theta}) \equiv \frac{1}{\bar{N}} \int d\chi \, \frac{dV}{d\chi} \, \bar{n}_g^{3D}(\tau) \, \delta_g^{3D}(\boldsymbol{x}, \tau) = \int d\chi \, q_g(\chi) \, \delta_g^{3D}(\boldsymbol{x}, \tau), \tag{A.4}$$

where in the second equality we have identified the galaxy projection kernel as $q_g(\chi) = \frac{1}{\bar{N}} \frac{dV}{d\chi} \bar{n}_g^{3D}(\tau(\chi))$.

B HOD expressions for galaxy bias

The halo model and halo occupation distribution (HOD) approach offer a useful framework to make predictions for the galaxy bias parameters defined in eq. (2.7). The halo model assumes that galaxies reside inside dark matter halos with some mass. In this HOD approach the effective global number density of galaxies hosted within halos of mass $M_h \in [M_{h,\min}, M_{h,\max}]$ is given by

$$\bar{n}_g^{3D}(\tau) = \int_{M_{h,\min}}^{M_{h,\max}} dM_h \, \frac{d\bar{n}_h^{3D}}{dM_h}(M_h, \tau) \, \bar{N}_g(M_h, \tau), \tag{B.1}$$

where $d\bar{n}_h^{3D}/dM_h$ is the global halo mass function (number density of dark matter halos in an infinitesimal mass bin dM_h around halos of mass M_h) and $\bar{N}_g(M_h, \tau)$ is the expected number of galaxies residing inside dark matter halos of mass M_h at time τ. The galaxy bias parameters $b_{\mathcal{O}}$ are in turn expressed as [83]

$$b_{\mathcal{O}}(\tau) = \frac{1}{\bar{n}_g^{3D}(\tau)} \int_{M_{h,\min}}^{M_{h,\max}} dM_h \, \frac{d\bar{n}_h^{3D}}{dM_h}(M_h, \tau) \, \bar{N}_g(M_h, \tau) \left[b_{\mathcal{O},h}(M_h, \tau) + R_{\mathcal{O}, \bar{N}_g}(M_h, \tau) \right], \tag{B.2}$$

where $b_{\mathcal{O},h}$ is the bias parameter of dark matter halos of mass M_h, and $R_{\mathcal{O}, \bar{N}_g}(M_h, \tau)$ is called the *response function* of $\bar{N}_g$ for large-scale perturbations $\mathcal{O}$. This response function describes the modulation of $\bar{N}_g$ by the $\mathcal{O}$ perturbations, in the same way that the bias parameters $b_{\mathcal{O},h}$ describe the modulation of the halo mass function. In our HOD catalogues, we have assumed $\bar{N}_g$ to be the same everywhere inside the simulation box irrespective of the local density and tidal field values, which corresponds to assuming $R_{\mathcal{O}, \bar{N}_g} = 0$. In order to get an effective bias $b_{\mathcal{O}}^i$ parameter of a galaxy sample in a tomographic bin i we take the expectation value of the galaxy bias parameter $b_{\mathcal{O}}(z)$ over the redshift distribution $p^i(z)$ of the bin:

$$b_{\mathcal{O}}^i = \int dz \, p^i(z) b_{\mathcal{O}}(z). \tag{B.3}$$

C Power spectra and bispectra of galaxy and matter density fields

In this appendix we discuss the leading-order (tree-level) standard perturbation theory (PT) models of the 3D galaxy-matter power spectra and bispectra used in our work. We can write the Fourier transform of eq. (2.7) as

$$
\delta_g^{\rm 3D}(\boldsymbol{k},\tau) = b_1(\tau)\delta(\boldsymbol{k},\tau) + \frac{1}{2}\int \frac{{\rm d}^3\boldsymbol{q}}{(2\pi)^3}\,\delta(\boldsymbol{q},\tau)\delta(\boldsymbol{k}-\boldsymbol{q},\tau)\Big(b_2(\tau) + b_{s^2}(\tau)S_2(\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q})\Big)
$$
$$
+ \left[\epsilon(\boldsymbol{k},\tau) + \int \frac{{\rm d}^3\boldsymbol{q}}{(2\pi)^3}\,\epsilon_\delta(\boldsymbol{q},\tau)\delta(\boldsymbol{k}-\boldsymbol{q},\tau)\right],
\tag{C.1}
$$

where $\epsilon_\mathcal{O}$ are random Poisson variables with vanishing expectation values that are uncorrelated with the density fields. The term $S_2(\boldsymbol{k},\boldsymbol{q})$ is the operator which generates the Fourier representation of the square of the tidal tensor K^2 [65]

$$
S_2(\boldsymbol{k},\boldsymbol{q}) = \frac{(\boldsymbol{k}\cdot\boldsymbol{q})^2}{(kq)^2} - \frac{1}{3}\,.
\tag{C.2}
$$

In the equations that follow we drop the time τ from the arguments to ease the notation.

C.1 3D power spectra

The 3D matter power spectrum $P_{mm}^{\rm 3D}(k)$, the galaxy-matter cross-power spectrum $P_{gm}^{\rm 3D}(k)$, and the galaxy power spectrum $P_{gg}^{\rm 3D}(k)$ are defined as

$$
\langle\delta_m^{\rm 3D}(\boldsymbol{k})\delta_m^{\rm 3D}(\boldsymbol{k}')\rangle = (2\pi)^3\delta_D(\boldsymbol{k}+\boldsymbol{k}')P_{mm}^{\rm 3D}(k),
\tag{C.3a}
$$
$$
\langle\delta_g^{\rm 3D}(\boldsymbol{k})\delta_m^{\rm 3D}(\boldsymbol{k}')\rangle = (2\pi)^3\delta_D(\boldsymbol{k}+\boldsymbol{k}')P_{gm}^{\rm 3D}(k),
\tag{C.3b}
$$
$$
\langle\delta_g^{\rm 3D}(\boldsymbol{k})\delta_g^{\rm 3D}(\boldsymbol{k}')\rangle = (2\pi)^3\delta_D(\boldsymbol{k}+\boldsymbol{k}')P_{gg}^{\rm 3D}(k).
\tag{C.3c}
$$

At tree-level perturbation theory these are given by

$$
P_{mm}^{\rm 3D}(k) = P_{\rm lin}^{\rm 3D}(k),
\tag{C.4a}
$$
$$
P_{gm}^{\rm 3D}(k) = b_1 P_{\rm lin}^{\rm 3D}(k),
\tag{C.4b}
$$
$$
P_{gg}^{\rm 3D}(k) = b_1^2 P_{\rm lin}^{\rm 3D}(k) + P_{\epsilon\epsilon}^{\rm 3D}(k),
\tag{C.4c}
$$

where $P_{\epsilon\epsilon}^{\rm 3D}$ is the power spectrum of the stochastic field ϵ and $P_{\rm lin}^{\rm 3D}$ is the linear matter power spectrum which scales with the amplitude of the primordial scalar perturbations A_s, i.e. $P_{\rm lin}^{\rm 3D} \propto A_s$.

C.2 3D bispectra

The 3D matter bispectrum $B_{mmm}^{\rm 3D}$, the galaxy-matter-matter bispectrum $B_{gmm}^{\rm 3D}$, the galaxy-galaxy-matter bispectrum $B_{ggm}^{\rm 3D}$, and the galaxy bispectrum $B_{ggg}^{\rm 3D}$ are defined as

$$
\langle\delta_m^{\rm 3D}(\boldsymbol{k}_1)\delta_m^{\rm 3D}(\boldsymbol{k}_2)\delta_m^{\rm 3D}(\boldsymbol{k}_3)\rangle = (2\pi)^3\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_{mmm}^{\rm 3D}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3),
\tag{C.5a}
$$
$$
\langle\delta_g^{\rm 3D}(\boldsymbol{k}_1)\delta_m^{\rm 3D}(\boldsymbol{k}_2)\delta_m^{\rm 3D}(\boldsymbol{k}_3)\rangle = (2\pi)^3\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_{gmm}^{\rm 3D}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3),
\tag{C.5b}
$$
$$
\langle\delta_g^{\rm 3D}(\boldsymbol{k}_1)\delta_g^{\rm 3D}(\boldsymbol{k}_2)\delta_m^{\rm 3D}(\boldsymbol{k}_3)\rangle = (2\pi)^3\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_{ggm}^{\rm 3D}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3),
\tag{C.5c}
$$
$$
\langle\delta_g^{\rm 3D}(\boldsymbol{k}_1)\delta_g^{\rm 3D}(\boldsymbol{k}_2)\delta_g^{\rm 3D}(\boldsymbol{k}_3)\rangle = (2\pi)^3\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_{ggg}^{\rm 3D}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3).
\tag{C.5d}
$$

At tree-level perturbation theory these are expressed as [40, 76]:

$$B_{mmm}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = 2F_2(\boldsymbol{k}_1, \boldsymbol{k}_2)P_{\text{lin}}^{3D}(k_1)P_{\text{lin}}^{3D}(k_2) + 2F_2(\boldsymbol{k}_3, \boldsymbol{k}_1)P_{\text{lin}}^{3D}(k_3,)P_{\text{lin}}^{3D}(k_1)$$
$$+ 2F_2(\boldsymbol{k}_2, \boldsymbol{k}_3)P_{\text{lin}}^{3D}(k_2)P_{\text{lin}}^{3D}(k_3) \equiv B_{\text{tree}}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3), \tag{C.6a}$$

$$B_{gmm}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = b_1 B_{\text{tree}}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) + \left(b_2 + b_{s^2}S_2(\boldsymbol{k}_2, \boldsymbol{k}_3)\right)P_{\text{lin}}^{3D}(k_2)P_{\text{lin}}^{3D}(k_3), \tag{C.6b}$$

$$B_{ggm}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = b_1^2 B_{\text{tree}}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) + b_1\left(b_2 + b_{s^2}S_2(\boldsymbol{k}_3, \boldsymbol{k}_1)\right)P_{\text{lin}}^{3D}(k_3)P_{\text{lin}}^{3D}(k_1)$$
$$+ b_1\left(b_2 + b_{s^2}S_2(\boldsymbol{k}_2, \boldsymbol{k}_3)\right)P_{\text{lin}}^{3D}(k_2)P_{\text{lin}}^{3D}(k_3) + 2P_{\epsilon\epsilon_\delta}^{3D}(k_3)P_{\text{lin}}^{3D}(k_3), \tag{C.6c}$$

$$B_{ggg}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = b_1^3 B_{\text{tree}}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) + b_1^2\left(b_2 + b_{s^2}S_2(\boldsymbol{k}_1, \boldsymbol{k}_2)\right)P_{\text{lin}}^{3D}(k_1)P_{\text{lin}}^{3D}(k_2)$$
$$+ b_1^2\left(b_2 + b_{s^2}S_2(\boldsymbol{k}_3, \boldsymbol{k}_1)\right)P_{\text{lin}}^{3D}(k_3)P_{\text{lin}}^{3D}(k_1)$$
$$+ b_1^2\left(b_2 + b_{s^2}S_2(\boldsymbol{k}_2, \boldsymbol{k}_3)\right)P_{\text{lin}}^{3D}(k_2)P_{\text{lin}}^{3D}(k_3)$$
$$+ 2b_1\left(P_{\epsilon\epsilon_\delta}^{3D}(k_1)P_{\text{lin}}^{3D}(k_1) + P_{\epsilon\epsilon_\delta}^{3D}(k_2)P_{\text{lin}}^{3D}(k_2) + P_{\epsilon\epsilon_\delta}^{3D}(k_3)P_{\text{lin}}^{3D}(k_3)\right)$$
$$+ B_{\epsilon\epsilon\epsilon}^{3D}(k_1, k_2, k_3), \tag{C.6d}$$

where F_2 is the second-order gravitational mode-coupling kernel. The tree-level matter bispectrum B_{tree}^{3D} scales differently with A_s compared to the linear matter power spectrum, i.e. $B_{\text{tree}}^{3D} \propto (P_{\text{lin}}^{3D})^2 \propto A_s^2$. Further, the galaxy-matter bispectra scale differently with galaxy bias compared to the galaxy-matter power spectra. The $P_{\epsilon\epsilon}^{3D}$, $P_{\epsilon\epsilon_\delta}^{3D}$ and $B_{\epsilon\epsilon\epsilon}^{3D}$ are the power- and bi-spectra of the stochastic fields. Under the assumption of Poisson statistics for them, it follows that [40, 76]:

$$P_{\epsilon\epsilon}^{3D} = \frac{1}{\bar{n}_g^{3D}}, \qquad P_{\epsilon\epsilon_\delta}^{3D} = \frac{b_1}{2\bar{n}_g^{3D}}, \qquad B_{\epsilon\epsilon\epsilon}^{3D} = \frac{1}{(\bar{n}_g^{3D})^2}. \tag{C.7}$$

We note that to evaluate $B_{gmm}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ or $B_{mgm}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ it is important to associate the correct ordering of wavevectors $\boldsymbol{k}_i$ to the respective galaxy density 'g' parts of the correlations. For example in $B_{gmm}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$, the $\boldsymbol{k}_1$ mode is associated to δ_g, whereas in $B_{mgm}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ it is instead the mode $\boldsymbol{k}_2$. One thus needs to alter the wavevector arguments in eq. (C.6b) accordingly when calculating B_{mgm}^{3D}. Similar considerations hold when there are two instances of δ_g, e.g. in $B_{ggm}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ and $B_{mgg}^{3D}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$.

D Point-mass terms in tangential shear 2PCFs and integrated 3PCFs

The *mean* 3D matter density at position $\boldsymbol{x}+\boldsymbol{r}$ in the presence of a galaxy at $\boldsymbol{x}$ can be written as [84]

$$\rho_m^{3D}\left(\boldsymbol{x} + \boldsymbol{r} \mid n_g^{3D}(\boldsymbol{x})\right) \equiv \bar{\rho}_m^{3D}[1 + \xi_{gm}(r)], \tag{D.1}$$

where $\xi_{gm}(r) \equiv \langle \delta_g^{3D}(\boldsymbol{x})\delta_m^{3D}(\boldsymbol{x} + \boldsymbol{r})\rangle$ is the 3D galaxy-matter 2-point correlation function (we have assumed statistical homogeneity and isotropy of the Universe). The projected surface mass density on a 2D plane at position $\boldsymbol{R}$ around a galaxy located at the origin is given by

$$\Sigma_{gm}^{2D}(\boldsymbol{R}) \equiv \int_0^{\chi_{\text{lim}}} d\chi \, \rho_m^{3D}\left(\boldsymbol{r} \mid n_g^{3D}(\boldsymbol{x} = \boldsymbol{0})\right), \tag{D.2}$$

where $\boldsymbol{r} = [\boldsymbol{R} = (R, \phi), \chi]$ is expressed in cylindrical coordinates. The angle-averaged projected surface mass density at a distance R from the galaxy density then reads

$$\Sigma_{gm}(R) = \int_0^{\chi_{\rm lim}} {\rm d}\chi \, \bar{\rho}_m^{\rm 3D} \left[1 + \xi_{gm}(\sqrt{R^2 + \chi^2})\right], \tag{D.3}$$

where $r = \sqrt{R^2 + \chi^2}$. The average surface mass density within the disc of radius R around the galaxy position is given by

$$\bar{\Sigma}_{gm}(R) = \frac{2}{R^2} \int_0^R {\rm d}R' \, R' \, \Sigma_{gm}(R'). \tag{D.4}$$

The tangential shear signal is effectively a measure of the *excess surface mass density* $\Delta\Sigma_{gm}$, which is defined as

$$\begin{aligned}
\Delta\Sigma_{gm}(R) &= \bar{\Sigma}_{gm}(R) - \Sigma_{gm}(R) \\
&= \frac{2\bar{\rho}_m^{\rm 3D}}{R^2} \int_0^{\chi_{\rm lim}} {\rm d}\chi \, \int_0^R {\rm d}R' \, R' \, \xi_{gm}(\sqrt{R'^2 + \chi^2}) - \bar{\rho}_m^{\rm 3D} \int_0^{\chi_{\rm lim}} {\rm d}\chi \, \xi_{gm}(\sqrt{R^2 + \chi^2}).
\end{aligned} \tag{D.5}$$

This equation shows that the result at a given R gets contributions from ξ_{gm} on all distance scales below it through the integral $\int_0^R {\rm d}R'$, including scales where perturbation theory breaks down. To circumvent the problem and be able to make predictions for the tangential shear signal we write the galaxy-matter correlation function as

$$\xi_{gm}(r) = \xi_{gm}^{\rm PT}(r) + \xi_{gm}^{\rm res}(r), \tag{D.6}$$

where $\xi_{gm}^{\rm res}(r)$ is a residual term that is only nonzero for $r < r_{\rm min}$, with $r_{\rm min}$ denoting the scale below which perturbation theory breaks down. This way eq. (D.5) can be written as

$$\Delta\Sigma_{gm}(R > R_{\rm min}) = \Delta\Sigma_{gm}^{\rm PT}(R > R_{\rm min}) + \Delta\Sigma_{gm}^{\rm res}(R > R_{\rm min}), \tag{D.7}$$

where

$$\Delta\Sigma_{gm}^{\rm PT}(R > R_{\rm min}) = \frac{2\bar{\rho}_m^{\rm 3D}}{R^2} \int_0^{\chi_{\rm lim}} {\rm d}\chi \, \int_0^R {\rm d}R' \, R' \, \xi_{gm}^{\rm PT}(\sqrt{R'^2 + \chi^2}) - \bar{\rho}_m^{\rm 3D} \int_0^{\chi_{\rm lim}} {\rm d}\chi \, \xi_{gm}^{\rm PT}(\sqrt{R^2 + \chi^2}), \tag{D.8}$$

and

$$\Delta\Sigma_{gm}^{\rm res}(R > R_{\rm min}) = \frac{2\bar{\rho}_m^{\rm 3D}}{R^2} \int_0^{\chi_{\rm lim}} {\rm d}\chi \, \int_0^{R_{\rm min}} {\rm d}R' \, R' \, \xi_{gm}^{\rm res}(\sqrt{R'^2 + \chi^2}), \tag{D.9}$$

where we have used $\xi_{gm}^{\rm res}(R > R_{\rm min}) = 0$. Note that although eq. (D.8) integrates a perturbation theory model down to scales where it is not valid, the corrections to this when studying scales $R > R_{\rm min}$ are automatically absorbed by the residual contribution of eq. (D.9); in the thin-lens approximation (where the lens mass is sharply concentrated around χ_l, the comoving distance of the lens galaxies), the upper integration limit is given by $R_{\rm min} = \sqrt{r_{\rm min}^2 - \chi_l^2}$.

The observed tangential shear signal is related to the surface mass density via:

$$\xi_t(\alpha) \equiv \frac{\Delta\Sigma_{gm}(R = \chi_l \alpha)}{\Sigma_{\rm min}}, \tag{D.10}$$

where
$$\Sigma_{\min} = \frac{4\pi G}{c^2} \frac{(\chi_s - \chi_l)\chi_l}{\chi_s}, \tag{D.11}$$

and χ_s is the comoving distance of the source galaxies. This signal can thus be decomposed as

$$\xi_t(\alpha > \alpha_{\min}) = \xi_t^{\mathrm{PT}}(\alpha > \alpha_{\min}) + \xi_t^{\mathrm{res}}(\alpha > \alpha_{\min}), \tag{D.12}$$

where

$$\xi_t^{\mathrm{PT}}(\alpha > \alpha_{\min}) = \frac{\Delta\Sigma_{gm}^{\mathrm{PT}}(R = \chi_l\alpha)}{\Sigma_{\min}}, \tag{D.13}$$

and

$$\xi_t^{\mathrm{res}}(\alpha > \alpha_{\min}) = \frac{\Delta\Sigma_{gm}^{\mathrm{res}}(R = \chi_l\alpha)}{\Sigma_{\min}} = \frac{2\bar{\rho}_m^{\mathrm{3D}}}{\alpha^2\chi_l^2\,\Sigma_{\min}} \int_0^{\chi_{\mathrm{lim}}} \mathrm{d}\chi \int_0^{R_{\min}} \mathrm{d}R'\, R'\, \xi_{gm}^{\mathrm{res}}(\sqrt{R'^2 + \chi^2})$$
$$\equiv \frac{\mathcal{M}_t}{\alpha^2}. \tag{D.14}$$

The last equality defines the so-called *point-mass term* $\mathcal{M}_t$. Its value cannot be worked out analytically with perturbation theory, and so we treat it as a free model parameter. The final result is thus

$$\xi_t(\alpha) = \xi_t^{\mathrm{PT}}(\alpha) + \frac{\mathcal{M}_t}{\alpha^2}, \tag{D.15}$$

which matches eq. (2.10) in the main body of the paper. We refer the interested reader to refs. [45, 85–88] for more details about point-mass term contributions to the tangential shear 2PCF.

Similar point mass terms contribute to the integrated 3PCFs as well. Concretely, we can write the *position-dependent tangential shear 2PCF* with a position-dependent point-mass term $\mathcal{M}_t(\boldsymbol{\theta}_C)$ as

$$\xi_t(\alpha; \boldsymbol{\theta}_C) = \xi_t^{\mathrm{PT}}(\alpha; \boldsymbol{\theta}_C) + \frac{\mathcal{M}_t(\boldsymbol{\theta}_C)}{\alpha^2}. \tag{D.16}$$

The correlation of this statistic with the lensing aperture mass $M_a(\boldsymbol{\theta}_C)$ yields eq. (4.1e) for $\zeta_{at}(\alpha)$,

$$\zeta_{at}(\alpha) \equiv \langle M_a(\boldsymbol{\theta}_C)\xi_t(\alpha; \boldsymbol{\theta}_C)\rangle = \langle M_a(\boldsymbol{\theta}_C)\xi_t^{\mathrm{PT}}(\alpha; \boldsymbol{\theta}_C)\rangle + \frac{\langle M_a(\boldsymbol{\theta}_C)\mathcal{M}_t(\boldsymbol{\theta}_C)\rangle}{\alpha^2}$$
$$= \zeta_{at}^{\mathrm{PT}}(\alpha) + \frac{\mathcal{M}_{at}}{\alpha^2}, \tag{D.17}$$

where the last equality defines the *point-mass term* for ζ_{at}, $\mathcal{M}_{at}$. Similarly, the correlation with the mean number of galaxies within apertures yields eq. (4.1f) for $\zeta_{gt}(\alpha)$,

$$\zeta_{gt}(\alpha) \equiv \langle M_g(\boldsymbol{\theta}_C)\xi_t(\alpha; \boldsymbol{\theta}_C)\rangle = \langle M_g(\boldsymbol{\theta}_C)\xi_t^{\mathrm{PT}}(\alpha; \boldsymbol{\theta}_C)\rangle + \frac{\langle M_g(\boldsymbol{\theta}_C)\mathcal{M}_t(\boldsymbol{\theta}_C)\rangle}{\alpha^2}$$
$$= \zeta_{gt}^{\mathrm{PT}}(\alpha) + \frac{\mathcal{M}_{gt}}{\alpha^2}, \tag{D.18}$$

with $\mathcal{M}_{gt}$ a new point-mass term. Again, these new point-mass terms cannot be evaluated with perturbation theory, and so we treat them as free model parameters.
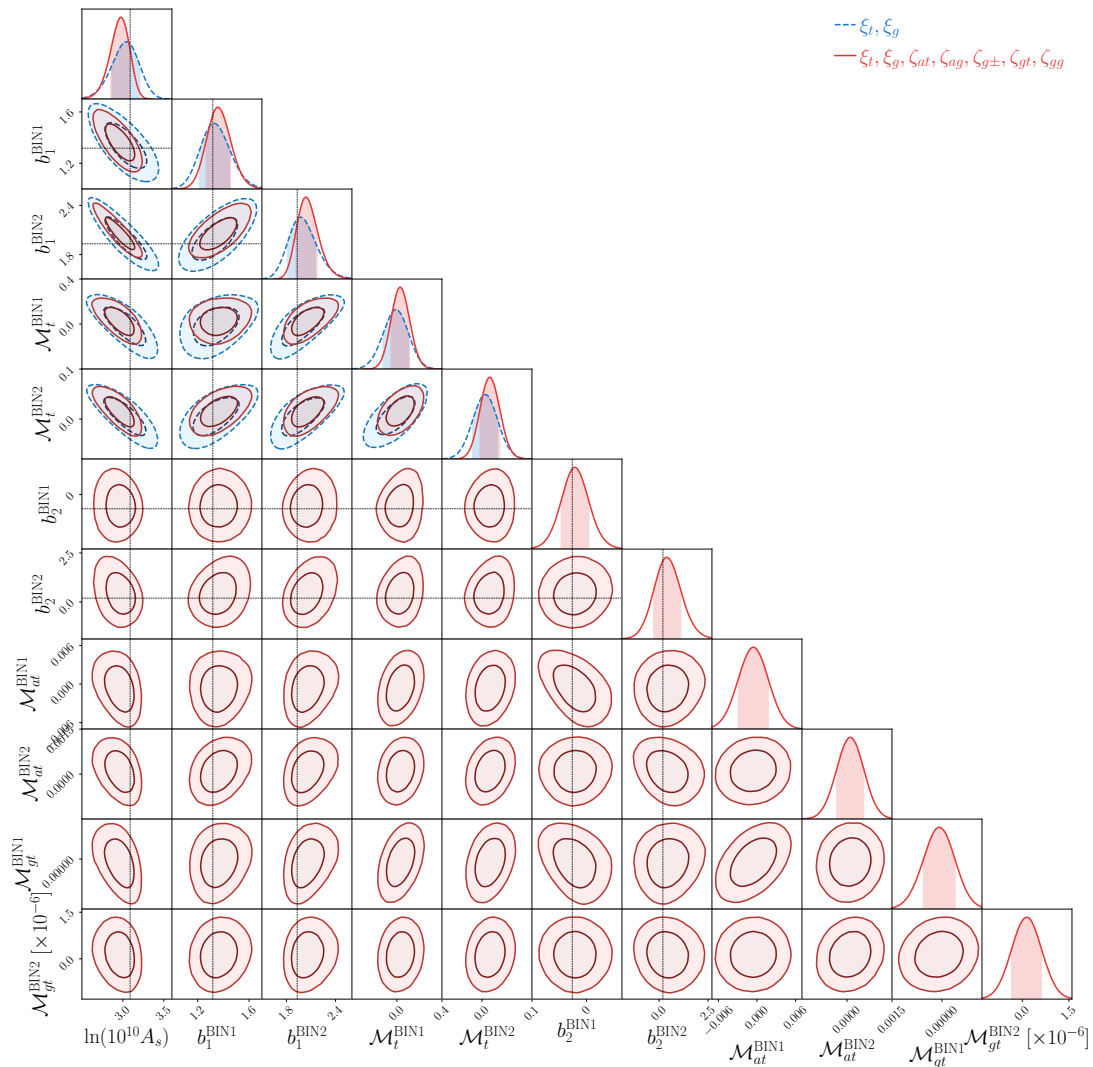
Figure 8. Same as figure 4 in the main body of the paper, but showing also the constraints on the point-mass terms. As there is no analytical way to straightforwardly calculate the expected value of the point-mass terms, these have no dotted lines marking their fiducial values.

Figure 8 shows the constraints from an MCMC analysis with the galaxy 2PCFs (blue) and their combination with the galaxy integrated 3PCFs (red); this is the same as figure 4 in the main body of the paper in section 7.2, except it also shows the constraints on the point-mass terms. Note each of the two lens galaxy samples has their associated point-mass terms. In the figure, it is interesting to note that the addition of the integrated 3PCFs leads to tighter constraints also on the $\mathcal{M}_t$ point-mass terms that contribute to ξ_t. We leave a more in depth study of the constraints on point-mass terms, including eventual insights on the density profile of the lens galaxies, to future work.

References

[1] DES collaboration, *Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing*, *Phys. Rev. D* **105** (2022) 023520 [`arXiv:2105.13549`] [ɪɴSPIRE].

[2] C. Heymans et al., *KiDS-1000 Cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints*, *Astron. Astrophys.* **646** (2021) A140 [`arXiv:2007.15632`] [ɪɴSPIRE].

[3] T. Hamana et al., *Cosmological constraints from cosmic shear two-point correlation functions with HSC survey first-year data*, *Publ. Astron. Soc. Jpn.* **72** (2020) 16 [`arXiv:1906.06041`] [*Erratum ibid.* **74** (2022) 488-491-491] [ɪɴSPIRE].

[4] EUCLID collaboration, *Euclid Definition Study Report*, `arXiv:1110.3193` [ɪɴSPIRE].

[5] LSST Dark Energy Science collaboration, *Large Synoptic Survey Telescope: Dark Energy Science Collaboration*, `arXiv:1211.0310` [ɪɴSPIRE].

[6] D. Spergel et al., *Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report*, `arXiv:1503.03757` [ɪɴSPIRE].

[7] F. Bernardeau, S. Colombi, E. Gaztanaga and R. Scoccimarro, *Large scale structure of the universe and cosmological perturbation theory*, *Phys. Rept.* **367** (2002) 1 [`astro-ph/0112551`] [ɪɴSPIRE].

[8] M. Takada and B. Jain, *Cosmological parameters from lensing power spectrum and bispectrum tomography*, *Mon. Not. Roy. Astron. Soc.* **348** (2004) 897 [`astro-ph/0310125`] [ɪɴSPIRE].

[9] P. Schneider, M. Kilbinger and M. Lombardi, *The three-point correlation function of cosmic shear. 2. Relation to the bispectrum of the projected mass density and generalized third-order aperture measures*, *Astron. Astrophys.* **431** (2005) 9 [`astro-ph/0308328`] [ɪɴSPIRE].

[10] E. Semboloni, H. Hoekstra and J. Schaye, *Effect of baryonic feedback on two- and three-point shear statistics: prospects for detection and improved modelling*, *Mon. Not. Roy. Astron. Soc.* **434** (2013) 148 [`arXiv:1210.7303`] [ɪɴSPIRE].

[11] L. Fu et al., *CFHTLenS: Cosmological constraints from a combination of cosmic shear two-point and three-point correlations*, *Mon. Not. Roy. Astron. Soc.* **441** (2014) 2725 [`arXiv:1404.5469`] [ɪɴSPIRE].

[12] DES collaboration, *Dark Energy Survey Year 3 Results: Three-point shear correlations and mass aperture moments*, *Phys. Rev. D* **105** (2022) 103537 [`arXiv:2201.05227`] [ɪɴSPIRE].

[13] S. Heydenreich, L. Linke, P. Burger and P. Schneider, *A roadmap to cosmological parameter analysis with third-order shear statistics — I. Modelling and validation*, *Astron. Astrophys.* **672** (2023) A44 [`arXiv:2208.11686`] [ɪɴSPIRE].

[14] P. Schneider and P. Watts, *Galaxy-galaxy-galaxy lensing: Third-order correlations between the galaxy and mass distributions in the Universe*, *Astron. Astrophys.* **432** (2005) 783 [`astro-ph/0502552`] [ɪɴSPIRE].

[15] L. Linke, P. Simon, P. Schneider, D.J. Farrow, J. Rödiger and A.H. Wright, *KiDS+VIKING+GAMA: Halo occupation distributions and correlations of satellite numbers with a new halo model of the galaxy-matter bispectrum for galaxy-galaxy-galaxy lensing*, *Astron. Astrophys.* **665** (2022) A38 [`arXiv:2204.02418`] [ɪɴSPIRE].

[16] DES collaboration, *Density Split Statistics: Joint Model of Counts and Lensing in Cells*, *Phys. Rev. D* **98** (2018) 023508 [`arXiv:1710.05162`] [ɪɴSPIRE].

[17] DES collaboration, *Density Split Statistics: Cosmological Constraints from Counts and Lensing in Cells in DES Y1 and SDSS Data*, *Phys. Rev. D* **98** (2018) 023507 [`arXiv:1710.05045`] [ɪɴSPIRE].

[18] P. Burger et al., *An adapted filter function for density split statistics in weak lensing*, *Astron. Astrophys.* **642** (2020) A161 [arXiv:2006.10778] [INSPIRE].

[19] P.A. Burger et al., *KiDS-1000 cosmology: Constraints from density split statistics*, *Astron. Astrophys.* **669** (2023) A69 [arXiv:2208.02171] [INSPIRE].

[20] A. Barthelemy, S. Codis and F. Bernardeau, *Probability distribution function of the aperture mass field with large deviation theory*, *Mon. Not. Roy. Astron. Soc.* **503** (2021) 5204 [arXiv:2012.03831] [INSPIRE].

[21] A. Boyle et al., *Nuw CDM cosmology from the weak-lensing convergence PDF*, *Mon. Not. Roy. Astron. Soc.* **505** (2021) 2886 [arXiv:2012.07771] [INSPIRE].

[22] B. Giblin, Y.-C. Cai and J. Harnois-Déraps, *Enhancing cosmic shear with the multiscale lensing probability density function*, *Mon. Not. Roy. Astron. Soc.* **520** (2023) 1721 [arXiv:2211.05708] [INSPIRE].

[23] B. Jain and U. Seljak, *Cosmological model predictions for weak lensing: Linear and nonlinear regimes*, *Astrophys. J.* **484** (1997) 560 [astro-ph/9611077] [INSPIRE].

[24] DES collaboration, *Dark Energy Survey Year 3 results: Cosmology with moments of weak lensing mass maps*, *Phys. Rev. D* **106** (2022) 083509 [arXiv:2110.10141] [INSPIRE].

[25] J. Harnois-Déraps et al., *Cosmic shear cosmology beyond two-point statistics: a combined peak count and correlation function analysis of DES-Y1*, *Mon. Not. Roy. Astron. Soc.* **506** (2021) 1623 [arXiv:2012.02777] [INSPIRE].

[26] DES collaboration, *Dark energy survey year 3 results: Cosmology with peaks using an emulator approach*, *Mon. Not. Roy. Astron. Soc.* **511** (2022) 2075 [arXiv:2110.10135] [INSPIRE].

[27] C.T. Davies, M. Cautun, B. Giblin, B. Li, J. Harnois-Déraps and Y.-C. Cai, *Cosmological forecasts with the clustering of weak lensing peaks*, *Mon. Not. Roy. Astron. Soc.* **513** (2022) 4729 [arXiv:2110.10164] [INSPIRE].

[28] LSST DARK ENERGY SCIENCE collaboration, *Forecasting the power of Higher Order Weak Lensing Statistics with automatically differentiable simulations*, arXiv:2305.07531 [INSPIRE].

[29] A. Halder, O. Friedrich, S. Seitz and T.N. Varga, *The integrated three-point correlation function of cosmic shear*, *Mon. Not. Roy. Astron. Soc.* **506** (2021) 2780 [arXiv:2102.10177] [INSPIRE].

[30] A. Halder and A. Barreira, *Response approach to the integrated shear 3-point correlation function: the impact of baryonic effects on small scales*, *Mon. Not. Roy. Astron. Soc.* **515** (2022) 4639 [arXiv:2201.05607] [INSPIRE].

[31] Z. Gong, A. Halder, A. Barreira, S. Seitz and O. Friedrich, *Cosmology from the integrated shear 3-point correlation function: simulated likelihood analyses with machine-learning emulators*, *JCAP* **07** (2023) 040 [arXiv:2304.01187] [INSPIRE].

[32] C.-T. Chiang, C. Wagner, F. Schmidt and E. Komatsu, *Position-dependent power spectrum of the large-scale structure: a novel method to measure the squeezed-limit bispectrum*, *JCAP* **05** (2014) 048 [arXiv:1403.3411] [INSPIRE].

[33] C.-T. Chiang, C. Wagner, A.G. Sánchez, F. Schmidt and E. Komatsu, *Position-dependent correlation function from the SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 10 CMASS Sample*, *JCAP* **09** (2015) 028 [arXiv:1504.03322] [INSPIRE].

[34] D. Munshi and P. Coles, *The Integrated Bispectrum and Beyond*, *JCAP* **02** (2017) 010 [arXiv:1608.04345] [INSPIRE].

[35] G. Jung, T. Namikawa, M. Liguori, D. Munshi and A. Heavens, *The integrated angular bispectrum of weak lensing*, *JCAP* **06** (2021) 055 [arXiv:2102.05521] [INSPIRE].

[36] A. Barreira and F. Schmidt, *Responses in Large-Scale Structure*, *JCAP* **06** (2017) 053 [arXiv:1703.09212] [INSPIRE].

[37] M. Bartelmann and P. Schneider, *Weak gravitational lensing*, *Phys. Rept.* **340** (2001) 291 [astro-ph/9912508] [INSPIRE].

[38] P. Schneider, *Weak gravitational lensing*, in *33rd Advanced Saas Fee Course on Gravitational Lensing: Strong, Weak, and Micro*, Les Diablerets, Switzerland (2006), pg. 269, https://doi.org/10.1007/978-3-540-30310-7_3 [astro-ph/0509252] [INSPIRE].

[39] M. Kilbinger, *Cosmology with cosmic shear observations: a review*, *Rept. Prog. Phys.* **78** (2015) 086901 [arXiv:1411.0115] [INSPIRE].

[40] V. Desjacques, D. Jeong and F. Schmidt, *Large-Scale Galaxy Bias*, *Phys. Rept.* **733** (2018) 1 [arXiv:1611.09787] [INSPIRE].

[41] A. Barreira, T. Lazeyras and F. Schmidt, *Galaxy bias from forward models: linear and second-order bias of IllustrisTNG galaxies*, *JCAP* **08** (2021) 029 [arXiv:2105.02876] [INSPIRE].

[42] P. Schneider, L. Van Waerbeke and Y. Mellier, *B-modes in cosmic shear from source redshift clustering*, *Astron. Astrophys.* **389** (2002) 729 [astro-ph/0112441] [INSPIRE].

[43] M. Jarvis, G. Bernstein and B. Jain, *The skewness of the aperture mass statistic*, *Mon. Not. Roy. Astron. Soc.* **352** (2004) 338 [astro-ph/0307393] [INSPIRE].

[44] DES collaboration, *Dark Energy Survey Year 3 Results: Multi-Probe Modeling Strategy and Validation*, arXiv:2105.13548 [INSPIRE].

[45] DES collaboration, *Inference from the Small Scales of Cosmic Shear with Current and Future Dark Energy Survey Data*, *Mon. Not. Roy. Astron. Soc.* **465** (2017) 2567 [arXiv:1608.01838] [INSPIRE].

[46] D.N. Limber, *The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. II*, *Astrophys. J.* **119** (1954) 655 [INSPIRE].

[47] N. Kaiser, *Weak gravitational lensing of distant galaxies*, *Astrophys. J.* **388** (1992) 272 [INSPIRE].

[48] A. Mead, J. Peacock, C. Heymans, S. Joudaki and A. Heavens, *An accurate halo model for fitting non-linear cosmological power spectra and baryonic feedback models*, *Mon. Not. Roy. Astron. Soc.* **454** (2015) 1958 [arXiv:1505.07833] [INSPIRE].

[49] R.G. Crittenden, P. Natarajan, U.-L. Pen and T. Theuns, *Discriminating weak lensing from intrinsic spin correlations using the curl-gradient decomposition*, *Astrophys. J.* **568** (2002) 20 [astro-ph/0012336] [INSPIRE].

[50] A. Buchalter, M. Kamionkowski and A.H. Jaffe, *The angular three-point correlation function in the quasilinear regime*, *Astrophys. J.* **530** (2000) 36 [astro-ph/9903486] [INSPIRE].

[51] G.P. Lepage, *VEGAS — an adaptive multi-dimensional integration program*, Technical Report Cornell University Ithaca, U.S.A. (1980), CLNS-447.

[52] DES collaboration, *Dark Energy Survey year 3 results: covariance modelling and its impact on parameter estimation and quality of fit*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 3125 [arXiv:2012.08568] [INSPIRE].

[53] R. Takahashi et al., *Full-sky Gravitational Lensing Simulation for Large-area Galaxy Surveys and Cosmic Microwave Background Experiments*, *Astrophys. J.* **850** (2017) 24 [arXiv:1706.01472] [INSPIRE].

[54] P.S. Behroozi, R.H. Wechsler and H.-Y. Wu, *The Rockstar Phase-Space Temporal Halo Finder and the Velocity Offsets of Cluster Cores*, *Astrophys. J.* **762** (2013) 109 [arXiv:1110.4372] [INSPIRE].

[55] T. Hamana, J. Sakurai, M. Koike and L. Miller, *Cosmological constraints from Subaru weak lensing cluster counts*, *Publ. Astron. Soc. Jpn.* **67** (2015) 34 [arXiv:1503.01851] [INSPIRE].

[56] M. Shirasaki, T. Hamana and N. Yoshida, *Probing cosmology with weak lensing selected clusters — I. Halo approach and all-sky simulations*, *Mon. Not. Roy. Astron. Soc.* **453** (2015) 3043 [arXiv:1504.05672] [INSPIRE].

[57] M. Shirasaki, T. Hamana, M. Takada, R. Takahashi and H. Miyatake, *Mock galaxy shape catalogues in the Subaru Hyper Suprime-Cam Survey*, *Mon. Not. Roy. Astron. Soc.* **486** (2019) 52 [arXiv:1901.09488] [INSPIRE].

[58] A. Zonca et al., *healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python*, *J. Open Source Softw.* **4** (2019) 1298 [INSPIRE].

[59] A.A. Berlind and D.H. Weinberg, *The Halo occupation distribution: Towards an empirical determination of the relation between galaxies and mass*, *Astrophys. J.* **575** (2002) 587 [astro-ph/0109001] [INSPIRE].

[60] A. Cooray and R.K. Sheth, *Halo Models of Large Scale Structure*, *Phys. Rept.* **372** (2002) 1 [astro-ph/0206508] [INSPIRE].

[61] DES collaboration, *Dark Energy Survey Year 3 results: galaxy-halo connection from galaxy-galaxy lensing*, *Mon. Not. Roy. Astron. Soc.* **509** (2022) 3119 [arXiv:2106.08438] [INSPIRE].

[62] J.L. Tinker et al., *Toward a halo mass function for precision cosmology: The Limits of universality*, *Astrophys. J.* **688** (2008) 709 [arXiv:0803.2706] [INSPIRE].

[63] J.L. Tinker et al., *The Large Scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests*, *Astrophys. J.* **724** (2010) 878 [arXiv:1001.3162] [INSPIRE].

[64] T. Lazeyras, C. Wagner, T. Baldauf and F. Schmidt, *Precision measurement of the local bias of dark matter halos*, *JCAP* **02** (2016) 018 [arXiv:1511.01096] [INSPIRE].

[65] T. Baldauf, U. Seljak, V. Desjacques and P. McDonald, *Evidence for Quadratic Tidal Tensor Bias from the Halo Bispectrum*, *Phys. Rev. D* **86** (2012) 083540 [arXiv:1201.4827] [INSPIRE].

[66] J.F. Navarro, C.S. Frenk and S.D.M. White, *The Structure of cold dark matter halos*, *Astrophys. J.* **462** (1996) 563 [astro-ph/9508025] [INSPIRE].

[67] O. Friedrich et al., *The PDF perspective on the tracer-matter connection: Lagrangian bias and non-Poissonian shot noise*, *Mon. Not. Roy. Astron. Soc.* **510** (2022) 5069 [arXiv:2107.02300] [INSPIRE].

[68] DES collaboration, *redMaGiC: Selecting Luminous Red Galaxies from the DES Science Verification Data*, *Mon. Not. Roy. Astron. Soc.* **461** (2016) 1431 [arXiv:1507.05460] [INSPIRE].

[69] DES collaboration, *Dark Energy Survey Year 3 results: Optimizing the lens sample in a combined galaxy clustering and galaxy-galaxy lensing analysis*, *Phys. Rev. D* **103** (2021) 043503 [arXiv:2011.03411] [INSPIRE].

[70] J. Hartlap, P. Simon and P. Schneider, *Why your model parameter confidences might be too optimistic: Unbiased estimation of the inverse covariance matrix*, *Astron. Astrophys.* **464** (2007) 399 [astro-ph/0608064] [INSPIRE].

[71] D. Foreman-Mackey, D.W. Hogg, D. Lang and J. Goodman, *emcee: The MCMC Hammer*, *Publ. Astron. Soc. Pac.* **125** (2013) 306 [arXiv:1202.3665] [INSPIRE].

[72] M. Tegmark, A. Taylor and A. Heavens, *Karhunen-Loeve eigenvalue problems in cosmology: How should we tackle large data sets?*, *Astrophys. J.* **480** (1997) 22 [astro-ph/9603021] [INSPIRE].

[73] A.G. Sanchez, *Arguments against using h^{-1} Mpc units in observational cosmology*, *Phys. Rev. D* **102** (2020) 123511 [arXiv:2002.07829] [INSPIRE].

[74] M. Chevallier and D. Polarski, *Accelerating universes with scaling dark matter*, *Int. J. Mod. Phys. D* **10** (2001) 213 [gr-qc/0009008] [INSPIRE].

[75] A. Eggemeier, R. Scoccimarro, R.E. Smith, M. Crocce, A. Pezzotta and A.G. Sánchez, *Testing one-loop galaxy bias: Joint analysis of power spectrum and bispectrum*, *Phys. Rev. D* **103** (2021) 123550 [arXiv:2102.06902] [INSPIRE].

[76] O. Leicht, T. Baldauf, J. Fergusson and P. Shellard, *Projected two- and three-point statistics: Forecasts and mitigation of non-linear RSDs*, *Mon. Not. Roy. Astron. Soc.* **503** (2021) 2137 [arXiv:2008.10199] [INSPIRE].

[77] K. Pardede, E. Di Dio and E. Castorina, *Wide-angle effects in the galaxy bispectrum*, *JCAP* **09** (2023) 030 [arXiv:2302.12789] [INSPIRE].

[78] D.M. Schmitz, C.M. Hirata, J. Blazek and E. Krause, *Time evolution of intrinsic alignments of galaxies*, *JCAP* **07** (2018) 030 [arXiv:1805.02649] [INSPIRE].

[79] J. Lesgourgues, *The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview*, arXiv:1104.2932 [INSPIRE].

[80] C.R. Harris et al., *Array programming with NumPy*, *Nature* **585** (2020) 357 [arXiv:2006.10256] [INSPIRE].

[81] J.D. Hunter, *Matplotlib: A 2D Graphics Environment*, *Comput. Sci. Eng.* **9** (2007) 90 [INSPIRE].

[82] S. Hinton, *ChainConsumer*, *J. Open Source Softw.* **1** (2016) 45.

[83] R. Voivodic and A. Barreira, *Responses of Halo Occupation Distributions: a new ingredient in the halo model & the impact on galaxy bias*, *JCAP* **05** (2021) 069 [arXiv:2012.04637] [INSPIRE].

[84] P.J.E. Peebles, *The large-scale structure of the universe*, Princeton University Press (1980).

[85] T. Baldauf, R.E. Smith, U. Seljak and R. Mandelbaum, *An algorithm for the direct reconstruction of the dark matter correlation function from weak lensing and galaxy clustering*, *Phys. Rev. D* **81** (2010) 063531 [arXiv:0911.4973] [INSPIRE].

[86] R. Mandelbaum, U. Seljak, T. Baldauf and R.E. Smith, *Precision cluster mass determination from weak lensing*, *Mon. Not. Roy. Astron. Soc.* **405** (2010) 2078 [arXiv:0911.4972] [INSPIRE].

[87] DES collaboration, *Dark Energy Survey year 3 results: Constraints on cosmological parameters and galaxy-bias models from galaxy clustering and galaxy-galaxy lensing using the redMaGiC sample*, *Phys. Rev. D* **106** (2022) 043520 [arXiv:2105.13545] [INSPIRE].

[88] DES collaboration, *Non-local contribution from small scales in galaxy-galaxy lensing: comparison of mitigation schemes*, *Mon. Not. Roy. Astron. Soc.* **522** (2023) 412 [arXiv:2212.03734] [INSPIRE].

Chapter 8

The PDF perspective on the tracer-matter connection: Lagrangian bias and non-Poissonian shot noise

Bibliographic and copyright information

Author's contribution

In this paper we introduced an accurate theoretical model of the joint 1-point PDF of galaxy and matter density fluctuations which was conceptualised and developed by the first author O. Friedrich. As second author in this publication I processed the matter density maps and halo catalogs from the T17 simulation suite (Takahashi et al., 2017) and also created and provided the mock HOD galaxy simulations on which we measured the joint PDF to validate the theoretical model. I wrote Sec. 3 of the paper and reviewed the rest of the draft. All the other authors contributed through valuable discussions, ideas and reviewing of the paper draft.

The PDF perspective on the tracer-matter connection: Lagrangian bias and non-Poissonian shot noise

Oliver Friedrich,[1,2]★ Anik Halder,[3,4] Aoife Boyle [ORCID],[5] Cora Uhlemann [ORCID],[6] Dylan Britt,[7,8] Sandrine Codis,[9] Daniel Gruen[3,8,10] and ChangHoon Hahn[11]

[1]*Kavli Institute for Cosmology, University of Cambridge, CB3 0HA Cambridge, UK*
[2]*Churchill College, University of Cambridge, CB3 0DS Cambridge, UK*
[3]*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, D-81679 München, Germany*
[4]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse 1, D-85748 Garching, Germany*
[5]*CNRS & Sorbonne Université, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[6]*School of Mathematics, Statistics and Physics, Newcastle University, Herschel Building, NE1 7RU Newcastle-upon-Tyne, UK*
[7]*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*
[8]*Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA*
[9]*AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France*
[10]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[11]*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton NJ 08544, USA*

ABSTRACT

We study the connection of matter density and its tracers from the probability density function (PDF) perspective. One aspect of this connection is the conditional expectation value $\langle \delta_{\mathrm{tracer}} | \delta_{\mathrm{m}} \rangle$ when averaging both tracer and matter density over some scale. We present a new way to incorporate a Lagrangian bias expansion of this expectation value into standard frameworks for modelling the PDF of density fluctuations and counts-in-cells statistics. Using N-body simulations and mock galaxy catalogues we confirm the accuracy of this expansion and compare it to the more commonly used Eulerian parametrization. For haloes hosting typical luminous red galaxies, the Lagrangian model provides a significantly better description of $\langle \delta_{\mathrm{tracer}} | \delta_{\mathrm{m}} \rangle$ at second order in perturbations. A second aspect of the matter-tracer connection is shot-noise, i.e. the scatter of tracer density around $\langle \delta_{\mathrm{tracer}} | \delta_{\mathrm{m}} \rangle$. It is well known that this noise can be significantly non-Poissonian and we validate the performance of a more general, two-parameter shot-noise model for different tracers and simulations. Both parts of our analysis are meant to pave the way for forthcoming applications to survey data.

Key words: theory – large-scale structure of Universe.

1 INTRODUCTION

Studying the evolution of the cosmic density field with the help of galaxy positions is like trying to understand a mountain range from knowing the location of (some of) its mountain peaks. One can hardly hope to infer the full profile of the density field from (a subset of) its luminous tracers. But one can hope that statistical properties of the galaxy density field can be expressed as functions of corresponding statistical properties of the total matter density field. For example, in the case of two-point statistics, one may assume that the galaxy clustering correlation function is just a multiple of the matter density correlation function (linear galaxy bias). In that case, any cosmological information contained in the shape of the matter density two-point function can still be retrieved from the galaxy density two-point function.

For such a program to be successful, one would optimally like to know the precise functional form that relates statistics of the matter density and galaxy density fields. And if there are unknown features in that functional form, then one would at least like to break

down these features into a well defined set of unknown numbers that parametrize our ignorance. The earliest attempt at finding such a parametrization was made by Kaiser (1984, Kaiser bias), who found that at sufficiently large-scales the two-point function of collapsed objects (clusters as modeled by overdense regions) is indeed proportional to the two-point function of the density field. At small scales, this picture of linear bias must be corrected due to halo exclusion and non-linear biasing effects (see e.g. Baldauf, Schaan & Zaldarriaga 2016; Desjacques, Jeong & Schmidt 2018; Ivanov, Simonović & Zaldarriaga 2020; Pandey et al. 2020; Baldauf et al. 2021). Even the simple linear bias model renders the amplitude of the galaxy clustering correlation function useless for inferring cosmological information. This degeneracy between galaxy bias and the variance of matter density fluctuations is broken when studying the full shape of the probability density function (PDF) of galaxy density fluctuations (Friedrich et al. 2018; Uhlemann et al. 2018a; Repp & Szapudi 2020). However, analysing the full PDF shape comes with the additional complication that one also has to understand the scatter between galaxy density and matter density fluctuations [shot-noise or stochasticity, see e.g. Friedrich et al. (2018) and Gruen et al. (2018) for a PDF context or Hamaus et al. (2010) and Desjacques et al. (2018) for stochasticity in two-point

★ E-mail: oliver.friedrich@physik.lmu.de

statistics]. Both for two-point and PDF statistics recent analyses had to employ quite complex models of the stochastic relation between matter density and galaxy density [e.g. Friedrich et al. (2018) and Gruen et al. (2018) using one parameter for galaxy bias and two parameters for density dependent shot-noise, Uhlemann et al. (2018a) using three parameters to describe a function relating the cumulative distribution function of matter and galaxy density fluctuations, and Ivanov et al. (2020) using three parameters for galaxy bias and one shot-noise amplitude].

In the PDF context, the bias of haloes (or galaxies) wrt. the matter density field is typically incorporated through an Eulerian expansion of the conditional expectation value $\langle \delta_{\mathrm{halo}} | \delta_{\mathrm{m}} \rangle$ (see e.g. Efstathiou 1995; Manera & Gaztañaga 2011; Clerkin et al. 2017; Friedrich et al. 2018; Gruen et al. 2018; Salvador et al. 2019; Repp & Szapudi 2020, with an exception found in Uhlemann et al. 2018c). This is somewhat unnatural, because standard methods to model the matter density PDF are typically built around the symmetric collapse of a leading order (saddle-point) configuration of the density field (e.g. Bernardeau 1994, 1995; Valageas 2002a; Bernardeau, Codis & Pichon 2015; Uhlemann et al. 2016, 2018b; Friedrich et al. 2020) which would seem to suggest a Lagrangian point-of-view.

We implement such a Lagrangian model in Section 2, where we also give a general overview of PDF modelling and also review the non-Poissonian shot-noise model of Friedrich et al. (2018) and Gruen et al. (2018) (hereafter F18 and G18). Section 3 presents details of the simulated data used in this study and in Section 4 we assess the importance of different aspects of our theory, by comparing our model of the joint PDF $p(\delta_{\mathrm{m}}, \delta_{\mathrm{g}})$ to the corresponding measured distribution of matter density and galaxy density fluctuations in those simulations. In particular, we are fitting both the Lagrangian and Eulerian bias models to measurements of $\langle \delta_{\mathrm{halo}} | \delta_{\mathrm{m}} \rangle$ in simulated data at different redshifts, for different smoothing scales and using haloes in different mass bins. We check whether the Lagrangian and Eulerian best-fitting parameters conform to consistency relations that should hold between them, and we compare them to corresponding values obtained from two-point function measurements and from analytical predictions of bias as a function of halo mass. Section 4.3 investigates details concerning shot-noise of tracer density fields. We discuss our results, summarize open questions, and give an outlook on future work in Section 5.

Throughout this paper, we consider the matter density and galaxy density fields averaged over cylindrical apertures (as opposed to e.g. spherical ones). This makes our results more directly applicable to line-of-sight projections of the cosmic density fields, since PDF-related statistics of such projected fields are most efficiently expressed in terms of line-of-sight integrals of corresponding cylindrical quantities (cf. Bernardeau & Valageas 2000; Friedrich et al. 2018; Barthelemy et al. 2020; Boyle et al. 2021; this is the equivalent of the Limber approximation – Limber 1953 – for two-point statistics). But our results can be easily transferred to the 3D case and to spherical apertures.

2 BIAS IN THE LANGUAGE OF PDF COSMOLOGY

2.1 Galaxy bias from the joint cumulant generating function of matter and galaxy density

In the following let $\delta_{m,R,L}(\boldsymbol{x}, z)$ and $\delta_{g,R,L}(\boldsymbol{x}, z)$ respectively be the matter and galaxy density contrast at redshift z and location $\boldsymbol{x}$ when averaging over a cylindrical aperture of radius R and length L (the orientation of the cylinder does not play a role in the following due

to statistical isotropy). In a statistically homogeneous and isotropic Universe, local moments of the form

$$\langle \delta_{m,R,L}(\boldsymbol{x}, z)^k \, \delta_{g,R,L}(\boldsymbol{x}, z)^l \rangle \tag{1}$$

do not depend on the spatial location $\boldsymbol{x}$ and we can define the joint moment generating function of matter and galaxy density contrast as

$$\psi_{R,L}(\lambda_m, \lambda_g, z) \equiv \sum_{k,l \geq 0} \langle \delta_{m,R,L}(\boldsymbol{x}, z)^k \, \delta_{g,R,L}(\boldsymbol{x}, z)^l \rangle \frac{\lambda_m^k \lambda_g^l}{k! \, l!}. \tag{2}$$

As evident from this definition, moments are obtained as derivatives of that function evaluated at $\lambda_{\mathrm{m}} = 0 = \lambda_{\mathrm{g}}$. For the rest of this subsection, we will suppress any dependencies of our notation on $\boldsymbol{x}$, z, R, and L. From the moment generating function ψ we define the cumulant generating function (CGF) as

$$\varphi(\lambda_{\mathrm{m}}, \lambda_{\mathrm{g}}) \equiv \log(\psi(\lambda_{\mathrm{m}}, \lambda_{\mathrm{g}}))$$

$$\equiv \sum_{k,l \geq 1} \langle \delta_{\mathrm{m}}^k \, \delta_{\mathrm{g}}^l \rangle_c \frac{\lambda_{\mathrm{m}}^k \lambda_{\mathrm{g}}^l}{k! \, l!}, \tag{3}$$

where the last line serves as a definition of the connected moments (or cumulants) $\langle \delta_{\mathrm{m}}^k \, \delta_{\mathrm{g}}^l \rangle_c$.

One quantity of interest for our study is the bias between galaxy density and matter density contrast as encoded by the conditional expectation value

$$\langle \delta_{\mathrm{g}} | \delta_{\mathrm{m}} \rangle = \frac{1}{p(\delta_{\mathrm{m}})} \int \mathrm{d}\delta_{\mathrm{g}} \, \delta_{\mathrm{g}} \, p(\delta_{\mathrm{g}}, \delta_{\mathrm{m}}). \tag{4}$$

Here $p(\delta_{\mathrm{m}})$ is the PDF of matter density contrast δ_{m} and $p(\delta_{\mathrm{g}}, \delta_{\mathrm{m}})$ is the joint PDF of both δ_{g} and δ_{m} (at the same location and redshift and averaged over the same cylindrical aperture). This joint PDF is related to the CGF via an inverse Laplace transformation (see e.g. Bernardeau & Valageas 2000; Valageas 2002a; Bernardeau et al. 2015; Friedrich et al. 2018). Hence, the above expectation value can be computed as

$$\langle \delta_{\mathrm{g}} | \delta_{\mathrm{m}} \rangle = \frac{1}{p(\delta_{\mathrm{m}})} \int \frac{\mathrm{d}\lambda_{\mathrm{g}} \mathrm{d}\lambda_{\mathrm{m}}}{(2\pi)^2} \, e^{-i\lambda_{\mathrm{m}}\delta_{\mathrm{m}} + \varphi(i\lambda_{\mathrm{m}}, i\lambda_{\mathrm{g}})} \int \mathrm{d}\delta_{\mathrm{g}} \, \delta_{\mathrm{g}} \, e^{-i\lambda_{\mathrm{g}}\delta_g}$$

$$= \frac{1}{p(\delta_{\mathrm{m}})} \int \frac{\mathrm{d}\lambda_{\mathrm{g}} \mathrm{d}\lambda_{\mathrm{m}}}{2\pi} \, e^{-i\lambda_{\mathrm{m}}\delta_{\mathrm{m}} + \varphi(i\lambda_{\mathrm{m}}, i\lambda_{\mathrm{g}})} \, i \, \frac{\mathrm{d}\delta_{\mathrm{Dirac}}(\lambda_{\mathrm{g}})}{\mathrm{d}\lambda_{\mathrm{g}}}$$

$$= \frac{\int \frac{\mathrm{d}\lambda_{\mathrm{m}}}{2\pi} \, e^{-i\lambda_{\mathrm{m}}\delta_{\mathrm{m}} + \varphi(i\lambda_{\mathrm{m}})} \, \partial_{l_g}\varphi(l_m, l_g)\big|_{l_m = i\lambda_{\mathrm{m}}, \, l_g = 0}}{\int \frac{\mathrm{d}\lambda_{\mathrm{m}}}{2\pi} \, e^{-i\lambda_{\mathrm{m}}\delta_m + \varphi(i\lambda_{\mathrm{m}})}}, \tag{5}$$

where $\varphi(\lambda_{\mathrm{m}})$ is the CGF of δ_{m} alone.

2.2 The joint CGF from functional integration

To calculate $\langle \delta_{\mathrm{g}} | \delta_{\mathrm{m}} \rangle$ according to equation (5) we need to know the joint CGF $\varphi_R(\lambda_{\mathrm{g}}, \lambda_{\mathrm{m}})$, where we have re-introduced the dependence on the radius R of our smoothing aperture, since we will occasionally vary R. The CGF can be calculated from the joint PDF as (Bernardeau et al. 2015)

$$e^{\varphi_R(\lambda_m, \lambda_g)} = \langle e^{\lambda_g \delta_{g,R} + \lambda_m \delta_{m,R}} \rangle$$

$$= \int \mathrm{d}\delta_{\mathrm{g,R}} \, \mathrm{d}\delta_{\mathrm{m,R}} \, p(\delta_{\mathrm{g,R}}, \delta_{\mathrm{m,R}}) \, e^{\lambda_g \delta_{\mathrm{g,R}} + \lambda_m \delta_{\mathrm{m,R}}}. \tag{6}$$

We want to stress again that our smoothing apertures are cylindrical, i.e. R is the radius of these cylinders. The only reason for our use of cylindrical filtering is that we prepare for an analysis in line-of-sight projected data (Friedrich et al. in prep). And the CGF of a line-of-sight projected density field can be calculated in a Limber-type approximation (Limber 1953; Bernardeau & Valageas 2000;

Friedrich et al. 2018; Barthelemy et al. 2020) from the CGF of the 3D density field in cylindrical apertures. But the following derivations apply in an almost identical manner to spherical filters as well.

Let us assume that both the galaxy density and matter density field are completely determined by the initial density field, or equivalently: today's linear density field which is related to the initial density field through linear growth. Then the expectation value in equation (6) can also be expressed through a functional integral over all possible configurations of the linear density contrast (Valageas 2002a). This yields

$$e^{\varphi_R(\lambda_m, \lambda_g)} = \int \mathcal{D}\delta_{\rm lin}\, e^{\lambda_g \delta_{g,R}[\delta_{\rm lin}] + \lambda_m \delta_{m,R}[\delta_{\rm lin}]}\, \mathcal{P}[\delta_{\rm lin}], \qquad (7)$$

where $\delta_{g,R}[\,\cdot\,]$ and $\delta_{m,R}[\,\cdot\,]$ are now functionals and $\mathcal{P}[\cdot]$ is the probability density functional of the random field $\delta_{\rm lin}(\boldsymbol{x})$. For Gaussian initial conditions $\mathcal{P}[\cdot]$ is a Gaussian functional and determined completely by the linear power spectrum (Valageas 2002a). By re-expressing the probability density functional of $\delta_{\rm lin}$ in terms of its cumulant generating functional, equation (7) can be brought into a more general – and for our purposes more convenient – form. We thus follow Friedrich et al. (2020) who derived that

$$e^{\varphi_R(\lambda_m, \lambda_g)} = \frac{1}{\mathcal{N}} \int \mathcal{D}\delta_{\rm lin}\, \mathcal{D}J_{\rm lin}\, e^{-S_{\lambda_m, \lambda_g}[\delta_{\rm lin}, J_{\rm lin}]}, \qquad (8)$$

where $J_{\rm lin}$ is an auxiliary source associated with the initial conditions and the action $S_{\lambda_m, \lambda_g}[\delta_{\rm lin}, J_{\rm lin}]$ is defined as

$$S_{\lambda_m, \lambda_g}[\delta_{\rm lin}, J_{\rm lin}] \equiv -\lambda_m \delta_{m,R}[\delta_{\rm lin}] - \lambda_g \delta_{g,R}[\delta_{\rm lin}] + i J_{\rm lin} \cdot \delta_{\rm lin} - \Phi[i J_{\rm lin}]. \qquad (9)$$

Here $\Phi[\,\cdot\,]$ is now the cumulant generating functional of the random field $\delta_{\rm lin}$, and $\mathcal{N}$ is an irrelevant normalization constant that drops in our final result (cf. Friedrich et al. 2020).

For detailed analyses of these and related functional integrals we e.g. refer the reader to Valageas (2002a, b), Ivanov, Kaurov & Sibiryakov (2019), and Friedrich et al. (2020). For the purpose of our study, we only state that the saddle point approximation to equation (8) yields

$$\varphi_R(\lambda_m, \lambda_g) \approx -S_{\lambda_m, \lambda_g}[\delta_{\rm lin}^*, J_{\rm lin}^*], \qquad (10)$$

where $\delta_{\rm lin}^*$ and $J_{\rm lin}^*$ are the saddle point configurations of the fields $\delta_{\rm lin}(\boldsymbol{x})$ and $J_{\rm lin}(\boldsymbol{x})$ which minimize the action $S_{\lambda_m, \lambda_g}[\cdot, \cdot]$ and hence give the largest contribution to the functional integral. These saddle point configurations can be shown to exhibit the same symmetry as the aperture used to define the functionals $\delta_{g,R}[\,\cdot\,]$ and $\delta_{m,R}[\,\cdot\,]$ (Friedrich et al. 2018, 2020; Valageas 2002a). In the case of long cylindrical apertures ($L \gg R$) this means that $\delta_{\rm lin}^*$ and $J_{\rm lin}^*$ will be cylindrically symmetric functions. They can even be explicitly calculated (Friedrich et al. 2020; Valageas 2002a), which is, however, not needed for our purposes. What is more important is the fact that the functional $\delta_{m,R}[\,\cdot\,]$ can be easily determined in the cylindrically symmetric situation. If we are observing the density field at redshift z then $\delta_{m,R}[\delta_{\rm lin}^*]$ is given by

$$\delta_{m,R}[\delta_{\rm lin}^*] = \mathcal{F}(\delta_{{\rm lin}, R_{\rm lin}}^*, z). \qquad (11)$$

Here $R_{\rm lin}$ is the initial (Lagrangian) radius of all the matter that is enclosed within R at redshift z, and $\delta_{{\rm lin}, R_{\rm lin}}^*$ is the average value of the saddle point configuration $\delta_{\rm lin}^*$ within this radius. Because of mass conservation $R_{\rm lin}$ is given by the (implicit) equation

$$R_{\rm lin} = R\sqrt{1 + \mathcal{F}(\delta_{{\rm lin}, R_{\rm lin}}^*, z)}, \qquad (12)$$

and the function $\mathcal{F}(\delta_{{\rm lin}, R_{\rm lin}}^*, z)$ describes how a cylindrically symmetric perturbation evolves when today's linear density contrast within

its initial radius is $\delta_{{\rm lin}, R_{\rm lin}}^*$. We detail the equations of motion needed to calculate $\mathcal{F}$ in Appendix A.

So far we have reviewed existing results on calculating the CGF and extended the notation of Friedrich et al. (2020) to the joint CGF of both galaxy density and matter density fluctuations as well as to cylindrical apertures instead of spherical ones. We will now see how the saddle point approximation of equation (10) allows for a practical implementation of a Lagrangian bias model within PDF theory.

2.3 Lagrangian bias along the saddle point configuration

To implement a parametric model for halo bias, let us have a closer look at the functional $\delta_{g,R}[\delta_{\rm lin}]$. Since the saddle point configuration is cylindrically symmetric, we will only consider cylindrically symmetric configurations and effectively consider 2D density fields. If $\delta_g(\boldsymbol{r})$ is the (smooth, shot-noise free) galaxy density contrast at (the 2D) location $\boldsymbol{r}$, then $\delta_{g,R}$ is given by

$$\delta_{g,R} = \frac{1}{\pi R^2} \int_{|\boldsymbol{r}| \leq R} {\rm d}^2 r\, \delta_g(\boldsymbol{r}). \qquad (13)$$

Tracing back the cylindrically collapsing evolution of the saddle point, a mass element at location $\boldsymbol{r}$ will originate from some initial (Lagrangian) location $\boldsymbol{q}$. Following standard Lagrangian bias parametrizations (see e.g. Lazeyras et al. 2016; Desjacques et al. 2018) we assume $\delta_g(\boldsymbol{r})$ can be expressed in terms of both the linear and non-linear matter density contrast field as

$$1 + \delta_g(\boldsymbol{r}) = (1 + \delta_{\rm m}(\boldsymbol{r})) \left(1 + b_1^L \delta_{\rm lin}(\boldsymbol{q}) + \frac{b_2^L}{2} \delta_{\rm lin}(\boldsymbol{q})^2\right), \qquad (14)$$

where we have stopped the bias expansion at quadratic order in today's linear density contrast. The cylindrical average $1 + \delta_{g,R}$ is then given by

$$\frac{1}{\pi R^2} \int_{|\boldsymbol{r}| \leq R} {\rm d}^2 r\, (1 + \delta_{\rm m}(\boldsymbol{r})) \left(1 + b_1^L \delta_{\rm lin}(\boldsymbol{q}) + \frac{b_2^L}{2} \delta_{\rm lin}(\boldsymbol{q})^2\right)$$

$$= \left(\frac{R_{\rm lin}}{R}\right)^2 \frac{1}{\pi R_{\rm lin}^2} \int_{|\boldsymbol{q}| \leq R_{\rm lin}} {\rm d}^2 q\, \left(1 + b_1^L \delta_{\rm lin}(\boldsymbol{q}) + \frac{b_2^L}{2} \delta_{\rm lin}(\boldsymbol{q})^2\right)$$

$$= \left(\frac{R_{\rm lin}}{R}\right)^2 \left(1 + b_1^L \delta_{{\rm lin}, R_{\rm lin}} + \frac{b_2^L}{2} [\delta_{\rm lin}^2]_{R_{\rm lin}}\right). \qquad (15)$$

Here $R_{\rm lin}$ is again the initial, Lagrangian (or *linear*) radius of the cylindrical perturbation now enclosed within R, $\delta_{{\rm lin}, R_{\rm lin}}$ is the average of today's linear density contrast within $R_{\rm lin}$, and $[\delta_{\rm lin}^2]_{R_{\rm lin}}$ is the average of the squared linear density contrast within $R_{\rm lin}$. Since we are considering cylindrically collapsing perturbations, the Lagrangian radius $R_{\rm lin}$ is related to R through
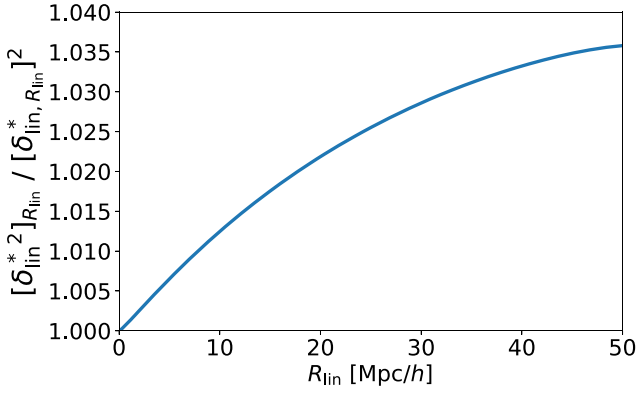
$$R_{\rm lin} = R\sqrt{1 + \delta_{\rm m,R}}. \qquad (16)$$

Hence, $\delta_{g,R}$ within our quadratic Lagrangian bias model is given by

$$1 + \delta_{g,R} = (1 + \delta_{m,R}) \left(1 + b_1^L \delta_{{\rm lin}, R_{\rm lin}} + \frac{b_2^L}{2} [\delta_{\rm lin}^2]_{R_{\rm lin}}\right). \qquad (17)$$

In Fig. 1, which is based on calculations presented in our Appendix B, we show that for the saddle point configuration $\delta_{\rm lin}^*$ the operations of squaring and cylindrically averaging approximately commute, i.e.

$$[\delta_{\rm lin}^{*\,2}]_{R_{\rm lin}} \approx (\delta_{{\rm lin}, R_{\rm lin}}^*)^2. \qquad (18)$$

Figure 1. At the saddle point configuration that dominates the path integral of equation (8) the operations of squaring and filtering the linear density contrast field commute approximately.

This allows us to express $\delta_{g,R}[\,\cdot\,]$ along the saddle point as

$$1 + \delta_{g,R}[\delta^*_{\text{lin}}, z] \approx (1 + \delta_{m,R}[\delta^*_{\text{lin}}, z])$$
$$\times \left(1 + b_1^L\, \delta^*_{\text{lin},R_{\text{lin}}} + \frac{b_2^L}{2}(\delta^*_{\text{lin},R_{\text{lin}}})^2\right). \quad (19)$$

We now have all the ingredients to formulate our main technical result. In complete analogy to the derivations of Friedrich et al. 2020 (but for cylindrical apertures and using the modified action of equation (9)) the task of determining the saddle point value of the action, $S_{\lambda_m,\lambda_g}[\delta^*_{\text{lin}}, J^*_{\text{lin}}]$, is equivalent to minimizing the 2D function

$$s_{\lambda_m,\lambda_g}(\delta, j) = -\lambda_g(1 + \mathcal{F}(\delta, z))\left(b_1^L\delta + \frac{b_2^L}{2}\delta^2\right) - (\lambda_m + \lambda_g)\,\mathcal{F}(\delta, z)$$
$$+ j\delta - \varphi_{\text{lin},R(1+\mathcal{F}(\delta,z))^{1/2}}(j). \quad (20)$$

Here $\varphi_{\text{lin},R}$ is the CGF of the linear density contrast (which is a quadratic function for Gaussian initial conditions) and δ and j should be understood as scalar variables. Minimizing $s_{\lambda_m,\lambda_g}(\delta, j)$ wrt. these variables yields an approximation of the joint CGF of matter density and galaxy density fluctuations via equation (10). This is the main result of our paper. Our formalism based on functional integration would be equivalent to a derivation within large deviation theory (LDT; see Bernardeau & Reimberg 2016, who introduced LDT for the matter density PDF), so we will refer to our calculation as the LDT model.

In practice we enhance the accuracy of this approximation with a linear-to-non-linear variance re-scaling of the CGF that leaves the reduced cumulants $S_n \equiv \langle\delta_m^n\rangle_c/\langle\delta_m^2\rangle_c^{n-1}$ unchanged (see e.g. section IV.A.2 of Friedrich et al. 2018). This, however, does not affect first derivatives of the CGF and has hence little impact on our calculation of $\langle\delta_g|\delta_m\rangle$ via equation (5). Numerical implementation of the minimization of $s_{\lambda_m,\lambda_g}(\delta, j)$ can be achieved in a manner similar to the one detailed step-by-step in section 4.6 of Friedrich et al. (2020). Equipped with the above approximation for the CGF we are now in a position to evaluate equation (5) and hence calculate the expectation value $\langle\delta_g|\delta_m\rangle$. In the following we will compare this Lagrangian bias model to an Eulerian model, which we directly define as a Taylor expansion of $\langle\delta_g|\delta_m\rangle$, i.e.

$$\langle\delta_g|\delta_m\rangle = b_1^E\,\delta_m + \frac{b_2^E}{2}\left(\delta_m^2 - \langle\delta_m^2\rangle\right). \quad (21)$$

This parametrization ignores tidal bias terms that can also contribute at second order in δ_m (e.g. Baldauf et al. 2012; Desjacques et al. 2018). Since we are averaging over cylindrical apertures we expect

these contributions to partially average out for the filtered density contrast (cf. fig. 3 of Baldauf et al. 2012) but our best-fitting values for b_2^E may absorb residual tidal contributions and hence may be slightly biased. We do not investigate this here. Subtracting the constant term $b_2^E/2 \cdot \langle\delta_m^2\rangle$ in equation (21) ensures that $\langle\delta_g\rangle = 0$. Note that this is not necessary in our Lagrangian model because of the Lagrangian-to-Eulerian mapping that is built into our path integral formulation.

2.4 Non-Poissonian shot-noise

The joint PDF of δ_m and δ_g can be expressed as

$$p(\delta_m, \delta_g) = p(\delta_m)\, p(\delta_g|\delta_m). \quad (22)$$

The matter density PDF $p(\delta_m)$ appearing on the right-hand side of this equation can be computed as the inverse Laplace transform of the CGF of δ_m (cf. the denominator in the last line of equation (5) as well as Bernardeau et al. 2015; Friedrich et al. 2018, 2020; Valageas 2002a for practical implementations of that transform). The second factor of the above equation, $p(\delta_g|\delta_m)$, is the conditional PDF of tracer density fluctuations given a fixed value of δ_m. In the previous subsections we have focused on computing the expectation value of that distribution, $\langle\delta_g|\delta_m\rangle$.

To model the full distribution $p(\delta_g|\delta_m)$ we have to consider stochasticity (resp. shot-noise) around the expectation value $\langle\delta_g|\delta_m\rangle$. This noise is often assumed to be Poissonian (see e.g. Efstathiou 1995; Clerkin et al. 2017; Salvador et al. 2019; Repp & Szapudi 2020). However, the results of F18 and G18 indicate that for certain types of tracers (in their case luminous red galaxies; cf. Rozo et al. 2016) this assumption can be in inaccurate (see also Hamaus, Seljak & Desjacques 2011; Dvornik et al. 2018 for non-Poissonian shot-noise in different contexts). To account for deviations from Poisson noise, F18 and G18 have modelled the distribution of a discrete random variable N with expectation value $\bar{N}$ as

$$P_\alpha(N) = \mathcal{N}\exp\left\{\frac{N}{\alpha}\ln\left[\frac{\bar{N}}{\alpha}\right] - \ln\Gamma\left[\frac{N}{\alpha} + 1\right] - \frac{\bar{N}}{\alpha}\right\}. \quad (23)$$

Here α parametrizes deviations from Poisson noise (with $\alpha = 1$ leading to a Poisson distribution), Γ is the gamma-function, and $\mathcal{N}$ is a normalization factor. That normalization is to a good approximation given by $1/\alpha$, though we do not rely on this here.

The above ansatz for $P_\alpha(N)$ can be used to model the distribution of tracer counts N_g in an aperture filled with a matter density contrast δ_m if we perform the identifications

$$N \to N_g$$
$$\bar{N} \to \bar{N}_g\,(1 + \langle\delta_g|\delta_m\rangle)$$
$$P_\alpha(N) \to P_\alpha(N_g|\delta_m),$$

where $\bar{N}_g$ is the mean tracer count across all apertures in a given survey volume. F18 and G18 then allow α to be a function of δ_m as well, hence making deviations from Poisson noise a function of the underlying matter density. They found that a linear ansatz,

$$\alpha(\delta_m) = \alpha_0 + \alpha_1\delta_m \quad (24)$$

describes the redMaGiC galaxy sample of the Buzzard N-body simulations (DeRose et al. 2019) well. We will test this linearity assumption here for a different set of simulations and different tracer samples of the large-scale structure.

3 SIMULATED DATA

The following section presents details of the different simulated data sets we use to test the theoretical ansatzes of Section 2.

3.1 T17 N-body simulations

We use publicly available data from cosmological simulations run by Takahashi et al. (2017).[1] In the following we refer to these as the T17 simulations. The simulations were generated primarily for the gravitational lensing studies for the Hyper Suprime Cam Survey. In this paper, we use the full-sky light-cone halo catalogues and matter density contrast shells of the simulation suite.

These data sets were obtained from a cold dark matter (CDM) only cosmological N-body simulation in periodic cubic boxes. The simulations consist of 14 boxes of increasing side lengths L, $2L$, $3L$,..., $14L$ (with $L = 450$ Mpc/h), nested around a common vertex (see fig. 1 of Takahashi et al. 2017). Each box contains 2048^3 particles (smaller boxes hence have better spatial and mass resolution) and their initial conditions were set with second-order Lagrangian perturbation theory (Crocce, Pueblas & Scoccimarro 2006) with an initial power spectrum computed for a flat ΛCDM cosmology with the following parameters: $\Omega_{cdm} = 0.233$, $\Omega_b = 0.046$, $\Omega_m = \Omega_{cdm} + \Omega_b = 0.279$, $\Omega_\Lambda = 0.721$, $h = 0.7$, $\sigma_8 = 0.82$, and $n_s = 0.97$. The particles in each box were then made to evolve from the initial conditions using the the N-body gravity solver GADGET2 (Springel, Yoshida & White 2001; Springel 2005). Dark matter haloes and sub-haloes in each simulation box were identified using the 6D phase-space friends-of-friends algorithm ROCKSTAR (Behroozi, Wechsler & Wu 2013). These ROCKSTAR halo catalogues and the evolved particle distribution of the different nested boxes are combined in layers of shells, each 150 Mpc/h thick, to obtain full-sky light cone halo catalogues and matter density contrast inside the shells, respectively. The simulation boxes were also ray traced using the multiple-lens plane ray-tracing algorithm GRAYTRIX (Hamana et al. 2015; Shirasaki, Hamana & Yoshida 2015) to obtain weak lensing convergence/shear maps for several source redshifts. Multiple simulations were run to produce 108 realizations (with labels r000 to r107) for each of these data products (see Takahashi et al. 2017 for more details). The authors report that the average matter power spectra from their several realizations of the simulations agreed with the theoretical revised Halofit power spectrum (Smith et al. 2003; Takahashi et al. 2012) to within 5 (10) per cent for $k < 5(6)$ h/Mpc at $z < 1$.

In this paper, for studying the bias as a function of halo properties we use the matter density contrast and the identified haloes in three 150 Mpc/h thick shells centred at $z = 0.476$, 0.751, 0.990 of realization r000 of the simulation suite. The all-sky halo catalogues come with a variety of halo properties such as halo mass, positions etc. of which we make use of the halo positions (right ascension, declination and redshift), halo mass M_{200b} (i.e. the mass contained in a radius within which the overdensity equals 200 times the background density), the virial radius of the halo R_{vir}, and the scale radius R_s, obtained by fitting an NFW profile (Navarro, Frenk & White 1996) to a given halo. The concentration parameter of the halo can then be calculated as $c \equiv R_{vir}/R_s$. Technically, our halo catalogues do contain sub-haloes. But the sub-halo fraction is negligible (< 0.1 per cent of the total halo population for the shell at $z = 0.476$ and even smaller for the other shells) such that for all practical purposes all haloes can be considered to be parent haloes.

3.2 Populating galaxies within T17 haloes using an halo occupation distribution approach

The T17 simulation suite does not come with galaxy catalogues. We would, however, like to validate our methods for typical luminous red galaxies (LRGs) similar to those observed by eBOSS (e.g. Zhai et al. 2017; Ross et al. 2020; $z \approx 0.7$). We hence create our own full-sky mock galaxy catalogue by populating the T17 halo catalogue at $z = 0.75$ using an empirical halo occupation distribution (HOD) method (Berlind & Weinberg 2002) based on the widely used halo model of large-scale structure (see Cooray & Sheth 2002 for a review). Briefly, an HOD describes a probability distribution $P(N_g|M_h)$, i.e. the probability that a given halo of mass M_h hosts N_g galaxies of a specific type (e.g. eBOSS LRG-like galaxies). We assume that the HOD does not depend on environment or formation history of the haloes (also known as assembly bias). We follow the work of Zhai et al. (2017) who empirically studied the clustering of more than 97 000 LRGs in the eBOSS survey within $z = 0.6$–0.9 (which contains the redshift range of the shell centred at $z = 0.75$) using a five-parameter HOD (we refer to this as the Zhai HOD). Zhai et al. (2017) parametrize their HOD by separating the contribution of a central galaxy from that of the satellite galaxies in a given halo of mass M_h. They characterized these contributions using the following functional forms for the mean values of the central and satellite galaxies,

$$\langle N_{cen}|M_h \rangle = \frac{1}{2} \left[1 + \mathrm{erf} \left(\frac{\log M_h - \log M_{min}}{\sigma_{\log M_h}} \right) \right], \tag{25}$$

$$\langle N_{sat}|M_h \rangle = \left(\frac{M_h}{M_{sat}} \right)^\gamma \exp \left(-\frac{M_{cut}}{M_h} \right) \langle N_{cen}|M_h \rangle. \tag{26}$$

The first of the above equations describes a smooth transition between having either 0 or 1 central galaxy with M_{min} being the mass at which half the haloes (in a given sample) host a central galaxy and $\sigma_{\log M_h}$ gives the scatter of the halo mass M_h at a fixed galaxy luminosity. The second equation gives the mean occupancy of satellite galaxies within the halo and is further parametrized by γ – a power-law index for the mass dependence of the number of satellites, M_{sat} – threshold mass for haloes to contain one satellite, and M_{cut} which allows for a halo-mass dependent cutoff. Together, the mean number of galaxies hosted within a halo of mass M_h is given by

$$\langle N_g|M_h \rangle = \langle N_{cen}|M_h \rangle + \langle N_{sat}|M_h \rangle. \tag{27}$$

Zhai et al. (2017) provide their best-fitting values for the five parameters by fitting analytical correlation functions[2] written in terms of their HOD to the observed galaxy clustering two-point correlation functions of the eBOSS LRGs sample. We report their best-fitting values here (see table 2 of Zhai et al. 2017): $\log M_{min} = 13.67$, $\log M_{sat} = 14.93$, $\gamma = 0.43$, $\log M_{cut} = 11.62$, $\sigma_{\log M_h} = 0.81$, where it is assumed that all the masses are expressed in units of $M_\odot/h$. In order to obtain these values Zhai et al. (2017) have adopted M_{200b} as their halo mass definition and we do so as well throughout our paper.

In order to create our mock galaxy catalogue from the T17 simulation, we use the HALOTOOLS software (Hearin et al. 2017) to first combine the T17 r000 halo shells which span the redshift range $z = 0.6$–0.9 to obtain a halo catalogue. Using M_{200b} as the mass proxy for the halo mass M_h, we use the Zhai HOD that we have described

---

[1]The data products of the simulation are available at http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/

[2]the one-halo and two-halo correlation functions, see e.g. appendix A of Coupon et al. (2012).

above along with their best-fitting parameters to populate each halo in the catalogue with galaxies. Note, however, that we restrict ourselves to haloes with masses $M_{200b} > 7.4 \times 10^{12}$ M$_\odot/h$ for the generation of our galaxy catalogue. This is to ensure that we have a similar number density of mock galaxies (per arcmin^2) as reported by Zhai et al. (2017) in their table 1 for the total BOSS + eBOSS LRG sample.[3] To this end, for a given halo we perform a Bernoulli draw with expectation given by equation (25) to get N_{cen} and a Poisson random draw with expectation given by equation (26) to obtain N_{sat}. The halo is then assigned to have a count of $N_{cen} + N_{sat}$ galaxies, where the central galaxy is placed at the same location as that of the parent halo's coordinates whereas a given satellite galaxy is placed at a distance r Mpc from the centre of the halo where r is a random realization[4] of a point drawn from an NFW profile. Besides the radial distance from the centre of the given halo, each satellite galaxy is assigned a uniformly distributed random angular direction on the sphere of radius r, from the centre of the halo. In this way, we create a mock full-sky eBOSS LRG like galaxy catalogue which we use for our analysis.

3.3 Quijote N-body simulations

The Quijote suite of N-body simulations (Villaescusa-Navarro et al. 2020) have been developed for quantifying the cosmological information content of large-scale structure observables. The suite consists of 43 100 simulations evaluated for more than 7000 cosmological models, varying the standard ΛCDM parameters, M_ν, and w. For our study we made use of the high-resolution runs of Quijote, which follow the evolution of 1024^3 particles over a comoving volume of 1 (Gpc/h)^3 starting from $z = 127$ for a fixed fiducial cosmology. Snapshots and halo catalogues (generated using a friends-of-friends algorithm) are publicly available for redshifts $z = 0, 0.5, 1, 2, 3$. Matter density PDFs are already included with the associated data products, and we extracted the joint tracer–matter PDFs. We refer the reader to Villaescusa-Navarro et al. (2020) for further details.

The Molino suite of mock galaxy catalogues has been created from the Quijote *N*-body simulations in order to extend cosmological forecasts to galaxy observables. The suite contains 75 000 mock galaxy catalogues that are constructed by applying the Zhai et al. (2017) HOD model (Section 3.2) to the Quijote halo catalogues. The galaxy catalogues are available at multiple cosmologies necessary for Fisher matrix forecasts (though here we only use catalogues at the Quijote fiducial cosmology of $(\Omega_m, \Omega_b, \sigma_8, n_s, h) = (0.3175, 0.049, 0.834, 0.9624, 0.6711)$).

4 COMPARISON OF THEORY AND SIMULATED DATA

We now compare the theoretical ansatzes developed in Section 2 to the simulated data described in Section 3. We start in Section 4.1 by looking at the joint PDF of matter density and our T17 synthetic galaxy sample. In Section 4.2 we then investigate the performance of our bias models as a function of mass, scale, and redshift. And in Section 4.3 we have a more detailed look at the shot-noise of different kinds of tracer samples.

4.1 The joint PDF of matter and galaxy density

In Fig. 2 we compare different models for the joint distribution of galaxy density and matter density fluctuations to a corresponding measurement of that distribution in the T17 simulations (cf. Section 3). The total matter density contrast of T17 is available in concentric shells of thickness 150 Mpc/h. For Fig. 2 we choose the shell centred around $z \approx 0.75$, which is e.g. similar to the average redshift of galaxy samples recently used in analyses of eBOSS (Zhai et al. 2017; de Mattia et al. 2021; Bautista et al. 2020; Gil-Marín et al. 2020; Tamone et al. 2020). The redshifts of the eBOSS LRG and ELG (emission line galaxy) samples span ranges that are significantly wider than 150 Mpc/h. Hence, the Limber-type approximation that one would employ when studying the line-of-sight projected PDF of these samples will not significantly deteriorate the accuracy we find here for the T17 shell width. To generate our mock galaxy sample we populates T17 haloes with the HOD described by Zhai et al. (2017, cf. our Section 3.2). To both the matter density and galaxy density map we then apply a circular top-hat filter with radius $R = 20$ Mpc/h perpendicular to the line-of-sight, i.e. we are averaging both fields in approximately cylindrical apertures of length $L = 150$ Mpc/h and radius $R = 20$ Mpc/h.

The blue contours in the two panels of Fig. 2 represent $1\sigma, 2\sigma$, and 3σ quantiles of the joint distribution $p(\delta_g, \delta_m)$ in our T17 + Zhai et al. mock data. The black contours represent the same quantiles for the theoretical model of $p(\delta_g, \delta_m)$ presented in Section 2. To obtain the Lagrangian bias parameters of that model, we have fit our theoretical prediction of $\langle \delta_g | \delta_m \rangle$ to measurements of that conditional expectation value in the simulated density fields. We performed these measurements in 25 equidistant bins of δ_m within a range that cuts 2 per cent of the probability from each tail of the PDF $p(\delta_m)$.

The red contours in the left-hand panel of Fig. 2 show the theoretical distribution $p(\delta_g, \delta_m)$ that one would obtain when assuming that $p(\delta_m)$ is a Gaussian PDF (and hence solely determined by its variance). Clearly, such a description is not sufficient for the matter density field at the smoothing scales and redshift considered here. The red contours in the right-hand panel of Fig. 2 show the distribution $p(\delta_g, \delta_m)$ that would be predicted when fitting only a linear Lagrangian bias model. Clearly, such a model does not sufficiently capture the curvature of $\langle \delta_g | \delta_m \rangle$ wrt. δ_m.

The difference between our best-fitting Lagrangian and Eulerian models for $p(\delta_g, \delta_m)$ is significantly smaller than the differences displayed in Fig. 2. Hence, we do not visualize them on the level of the full PDF, but for the conditional expectation values $\langle \delta_g | \delta_m \rangle$. Fig. 3 shows the residuals of $\langle \delta_g | \delta_m \rangle$ measured in our simulated data wrt. our best-fitting, quadratic Lagrangian, and Eulerian model (blue squares and red circles; within the range used to fit both models which cuts 2 per cent of probability from the tails of $p(\delta_m)$). We normalize these residuals by the 1σ standard deviations estimated with a jackknife scheme (cf. Section 4.2 for more details). The Lagrangian model manages to achieve a significantly better fit to our simulated data than the Eulerian one. The figure also shows the residuals of a best-fitting cubic Eulerian model which adds a term $b_3^E/6 \cdot (\delta_m^3 - \langle \delta_m^3 \rangle)$ to equation (21). This model performs very similar to the second order Lagrangian fit. Note, however, that the errorbars of Fig. 3 represent all-sky data, i.e. they might overestimate the accuracy required for realistic analyses and the second order Eulerian model may still perform well enough for

---

[3] A more accurate approach would be to re-fit our HOD parameters by matching a sufficiently constraining set of statistics of our mock galaxies to a target observed galaxy sample. This is, however, beyond the scope of this work.

[4] Precisely, we use the `mc_generate_nfw_radial_positions` method from `halotools` to draw a satellite galaxy's radial location r inside a given halo of mass M_{200b}, concentration parameter c, and redshift z.
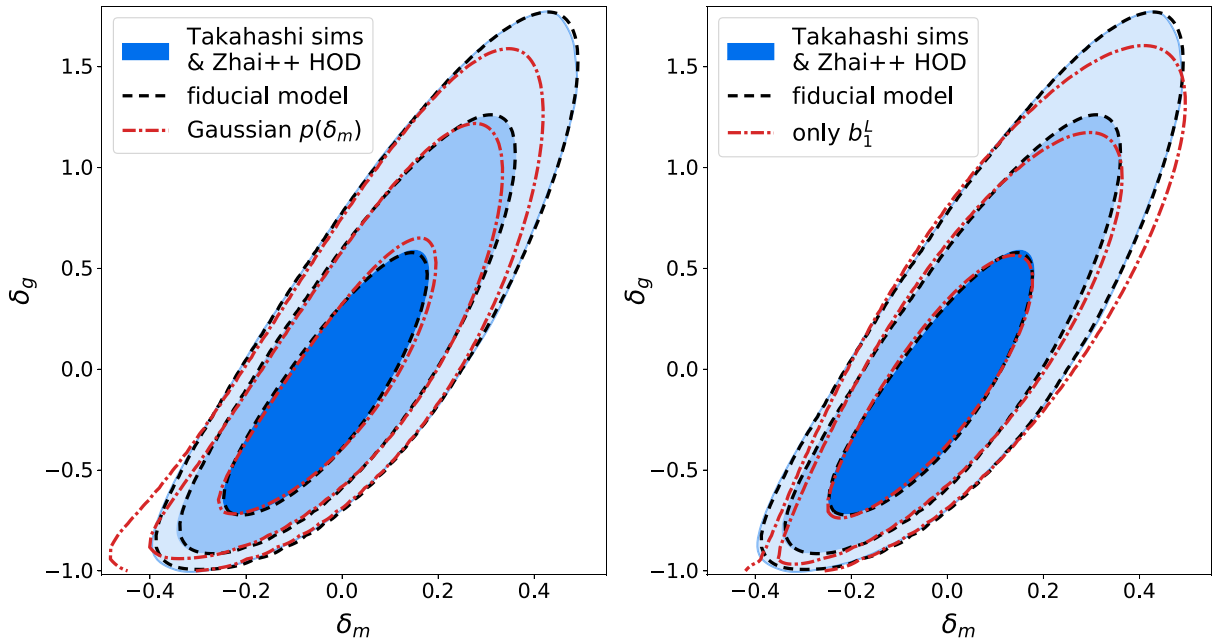
Figure 2. Comparing different models of the joint PDF $p(\delta_g, \delta_m)$ of galaxy and matter density fluctuations in cylindrical apertures of length $L = 150$ Mpc/h and radius $R = 20$ Mpc/h at redshift $z \approx 0.75$ to the distribution measured in simulated data. In both panels the blue contours represent the PDF measured in T17 mock data and using the mock galaxy catalogue described in Section 3.2. The black, dashed contours represents our fiducial model, which consists of three parts: an LDT model for the matter density PDF $p(\delta_m)$, a second order Lagrangian bias expansion within LDT, and a shot-noise model that allows for deviations from Poisson shot-noise (cf. Section 2.4 for the shot-noise model, and Section 4.3 for a detailed analysis of shot-noise in our simulations). The red, dash–dotted contours in the left-hand panel show what happens to the joint PDF model, if one assumes that $p(\delta_m)$ is Gaussian. The red, dash–dotted contours in the right-hand panel show a model that only fits a linear Lagrangian bias expansion.
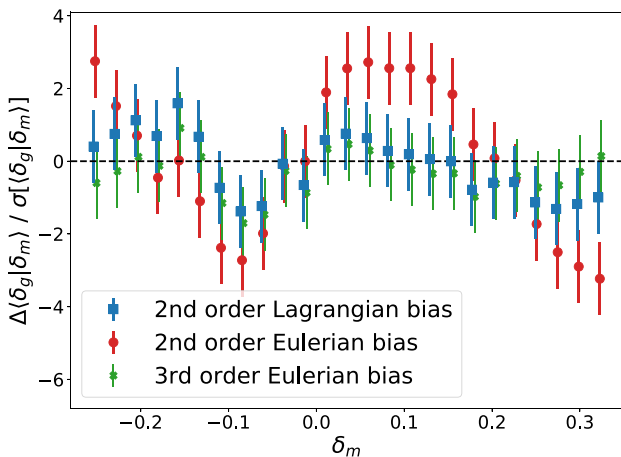

Figure 3. Residuals of $\langle \delta_g | \delta_m \rangle$ measured in simulated data (T17 haloes populated with the HOD description of Zhai et al. 2017) wrt. our best-fitting, quadratic Lagrangian (blue squares) and Eulerian (red circles) models. The figure uses the same scales and redshift as Fig. 2 and the residuals have been normalized by an estimate of the standard deviation of our measurements of $\langle \delta_g | \delta_m \rangle$. We also show that a third order Eulerian model (green crosses) performs similar to the second order Lagrangian one.

those. Also, we find in Section 4.2 that this comparison is somewhat mass dependent: the Lagrangian model tends to perform better for intermediate mass haloes, while the Eulerian one achieves better fits of $\langle \delta_g | \delta_m \rangle$ for very massive tracers of the density field.

4.2 Halo bias as a function of mass and consistency among bias measures

In Fig. 4 we show measurements of the conditional expectation value $\langle \delta_{\mathrm{halo}} | \delta_m \rangle$ in three different shells of the T17 simulations (with $z = 0.476, 0.751, 0.990$) and when averaging halo and matter densities in cylindrical apertures of radius $R = 20$ Mpc/h and length $L = 150$ Mpc/h. The different symbols in the figure represent measurements for different bins of halo mass. We choose bins of ± 10 per cent around the central masses $M_{\mathrm{200b}}/(10^{13} \, \mathrm{M}_\odot/h) = 0.5$ (squares), $= 1.0$ (circles), $= 2.0$ (crosses), $= 4.0$ (pluses), $= 8.0$ (hexagons), and $= 16.0$ (diamonds). The $z = 0.99$ shell of T17 does not resolve the lowest of these mass bins and the $z = 0.751$ shell only resolves haloes down to exactly $M_{\mathrm{200b}}/(10^{13} \, \mathrm{M}_\odot/h) = 0.5$, i.e. for that shell only the upper half of that bin enters our measurement. For each of the mass bins we measure $\langle \delta_{\mathrm{halo}} | \delta_m \rangle$ in 25 equidistant bins of δ_m and the lowest and upper most bound of these bins were chosen such as to cut away exactly 2 per cent of the probability from each tail of the underlying matter density PDF $p(\delta_m)$. We estimate the errorbars of each measurement from a jackknife approach (Norberg et al. 2009; Friedrich et al. 2016), splitting the all-sky maps of T17 into 196 sub-patches. The solid blue and dashed orange lines in the figure are best-fitting models from second order Lagrangian and Eulerian bias expansions, respectively (cf. Section 2).

We summarize the best-fitting values of our bias parameters as well as the χ^2 values between best-fitting models and measurements of $\langle \delta_{\mathrm{halo}} | \delta_m \rangle$ in Table 1. Taking into account that the noise in our covariance matrices adds a relative variance of about $\sqrt{2/(196 - 25 - 2)}$ to our best-fitting χ^2 (see e.g. Taylor, Joachimi & Kitching 2013, we add this noise in quadrature to the expected statistical scatter of
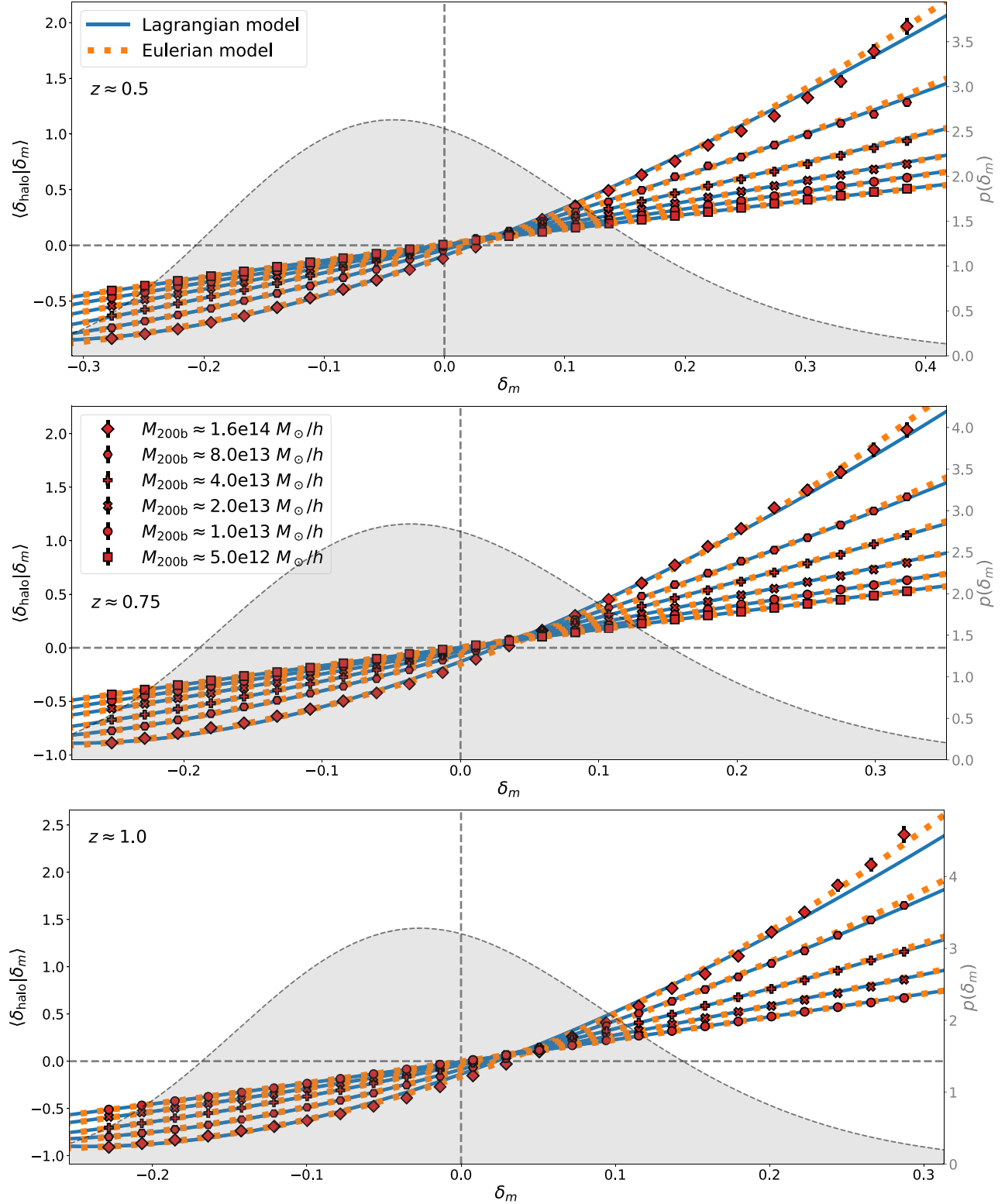
Figure 4. Conditional expectation value $\langle \delta_{\text{halo}} | \delta_m \rangle$ in cylindrical aperture of $R = 20$ Mpc/h and $L = 150$ Mpc/h at different redshifts and for different halo masses. Mass bins are $M_{200b}/(10^{13}\,\mathrm{M}_\odot/h) \approx 0.5$ (squares), ≈ 1.0 (circles), ≈ 2.0 (crosses), ≈ 4.0 (pluses), ≈ 8.0 (hexagons), and ≈ 16.0 (diamonds). The $z \approx 1.0$ shell of the T17 sims does not resolve the lowest mass bin. The solid blue and dashed orange lines are best-fitting models from second order Lagrangian and Eulerian bias expansions, respectively. Errorbars of the symbols are for an all-sky shell and are estimated from a jackknife procedure. The grey shaded area displays the PDF $p(\delta_{\mathrm{m}})$ (cf. right y-axis) and our binning of $\langle \delta_{\text{halo}} | \delta_m \rangle$ cuts away 2 per cent of the probability from the tails of that PDF.

Table 1. Best-fitting parameters and χ^2 values obtained from the fits shown in Figs 3 and 4.

$M_{200b}/(10^{13} M_\odot/h) \in$	b_1^L	b_2^L	χ_L^2	b_1^E	b_2^E	χ_E^2	$\chi_L^2 - \chi_E^2$
Optimally:							
Var$(\chi^2) = 6.78^2$ (statistical)							
$+ 2.49^2$ (cov. noise)							
$\Rightarrow \chi^2 \sim 23 \pm 7.22$							
$z \approx 0.5$:							
[0.45, 0.55]	0.39 ± 0.01	-0.72 ± 0.06	21.14	1.41 ± 0.01	-0.62 ± 0.06	20.17	0.97
[0.9, 1.1]	0.64 ± 0.01	-0.59 ± 0.08	20.28	1.66 ± 0.01	-0.42 ± 0.08	18.5	1.78
[1.8, 2.2]	0.96 ± 0.02	-0.43 ± 0.13	14.77	1.96 ± 0.01	-0.19 ± 0.14	14.13	0.64
[3.6, 4.4]	1.42 ± 0.02	0.35 ± 0.17	21.94	2.39 ± 0.02	0.73 ± 0.18	23.95	-1.99
[7.2, 8.8]	2.11 ± 0.04	2.6 ± 0.28	31.19	2.99 ± 0.03	3.36 ± 0.31	34.6	-3.4
[14.4, 17.6]	3.07 ± 0.06	6.72 ± 0.45	26.35	3.79 ± 0.05	8.32 ± 0.51	24.78	1.58
$z \approx 0.75$:							
[0.5, 0.55]	0.68 ± 0.01	-0.55 ± 0.12	16.53	1.7 ± 0.01	-0.39 ± 0.13	15.78	0.75
[0.9, 1.1]	0.97 ± 0.01	-0.26 ± 0.09	22.72	1.97 ± 0.01	-0.02 ± 0.09	22.05	0.67
[1.8, 2.2]	1.41 ± 0.01	0.75 ± 0.14	28.79	2.37 ± 0.01	1.11 ± 0.15	37.68	-8.88
[3.6, 4.4]	2.03 ± 0.02	2.2 ± 0.19	21.84	2.95 ± 0.02	2.87 ± 0.2	27.49	-5.64
[7.2, 8.8]	2.81 ± 0.04	5.16 ± 0.31	21.12	3.61 ± 0.03	6.31 ± 0.33	24.2	-3.07
[14.4, 17.6]	4.08 ± 0.08	11.35 ± 0.58	39.16	4.7 ± 0.06	14.25 ± 0.67	22.36	16.8
$z \approx 1.0$:							
[0.9, 1.1]	1.32 ± 0.01	0.29 ± 0.09	35.52	2.3 ± 0.01	0.62 ± 0.1	43.46	-7.93
[1.8, 2.2]	1.86 ± 0.01	1.74 ± 0.14	25.88	2.8 ± 0.01	2.28 ± 0.14	41.71	-15.81
[3.6, 4.4]	2.64 ± 0.02	4.23 ± 0.26	22.58	3.52 ± 0.02	5.25 ± 0.28	30.36	-7.78
[7.2, 8.8]	3.75 ± 0.05	10.19 ± 0.43	29.31	4.5 ± 0.04	12.31 ± 0.48	19.76	9.54
[14.4, 17.6]	4.96 ± 0.09	16.46 ± 0.76	70.05	5.6 ± 0.07	20.6 ± 0.89	45.73	24.32
Fitting $\langle\delta_g|\delta_m\rangle$ at $z \approx 0.75$	1.77 ± 0.008	2.11 ± 0.068	34.62	2.69 ± 0.007	2.69 ± 0.073	59.41	-24.79
(with the synthetic galaxies described							
in Section 3.2 and used for Fig. 3)							

χ^2), most of the fits in Fig. 4 agree with the measurements within either 1σ or 2σ. For the Lagrangian parametrization, only two of the overall 17 fits lie outside of 2σ. On average one would expect one such outlier. However, at least one of these outliers is at a very high χ^2 (6.5σ) and both of them are at the highest mass of their respective redshifts. Hence, there seems to be a systematic shortcoming of the Lagrangian model for very high halo masses.

The Eulerian parametrization performs somewhat better in these two instances (though it is still $>3\sigma$ off for the highest mass bin in the highest redshift shell). But in total, four of the Eulerian fits lie outside of 2σ. When ignoring the highest mass bins in the $z = 0.75$ and $z = 1.0$ shells, the Lagrangian model performs either similarly well or significantly better than the Eulerian one. This is the reason why the Lagrangian model was a significantly better fit to $\langle\delta_g|\delta_m\rangle$ for our synthetic galaxy sample discussed in Section 4.1 (cf. Fig. 3). The best-fitting parameters and χ^2 of that comparison are also displayed in Table 1.

Let us now investigate whether the best-fitting parameters of our bias models conform to basic theoretical expectations. In Fig. 5 we show the values obtained for the linear Lagrangian bias b_1^L at $z = 0.75$ as a function of the radius of our cylindrical aperture and for all of our different mass bins. The different symbols in the figure show measurements of b_1^L obtained from fitting our Lagrangian model for $\langle\delta_{halo}|\delta_m\rangle$ to T17 data. Above radii of $R_{cylinder} \approx 20-30$ Mpc/h there is only a mild scale dependence of these best-fitting values. Different horizontal lines in the figure display different theoretical predictions for the large-scale limit of b_1^L. Dashed lines show predictions obtained from the peak-background split (PBS) approach together with Press–Schechter halo mass function (Press & Schechter 1974), solid lines show predictions based on Sheth, Mo & Tormen (2001, i.e. including

their *moving barrier* correction), and dotted lines show predictions from the fitting formula of Tinker et al. (2010). All three sets of theoretical predictions match the large-scale limit of the bias values we fit with our Lagrangian parametrization to within 10 per cent accuracy. Surprisingly, the Press–Schechter predictions seem to match our measurements of b_1^L best (but with the Tinker et al. predictions performing very similarly). At very high masses ($\sim 8 \times 10^{13}$ M$_\odot$/h and $\sim 16 \times 10^{13}$ M$_\odot$/h) we find that the predictions of Sheth et al. (2001) are significantly lower than the other two models the biases measured from $\langle\delta_{halo}|\delta_m\rangle$ (and the other sets of predictions).

Lazeyras et al. (2016) have found a tight relationship between linear and quadratic Lagrangian bias, as measured from the response of halo density to changes in the overall matter density in a set of separate universe simulations. We expect our finding to closely match their results, because the expectation values $\langle\delta_{halo}|\delta_m\rangle$ resemble exactly that kind of response approach, with each of our apertures representing a (miniature) separate universe. In Fig. 6 we show our measurements of b_1^L and b_2^L in the three different redshift shells of the T17 data and for different mass bins. The colour coding of the mass bins is identical to that of Fig. 5 (higher bias values correspond to higher masses) and the different symbols represent fits to $\langle\delta_{halo}|\delta_m\rangle$ for different radii of our cylindrical aperture ($R = 10$, 20, 50 Mpc/h). The solid line in the figure displays the empirical relation found by Lazeyras et al. (2016). Despite directly measuring the Lagrangian bias parameters, they present their fit in terms of transformed, Eulerian biases. For reference, we translate that fit to Lagrangian space, which yields

$$b_2^L \approx -0.794 - 0.642\, b_1^L + 0.953\, (b_1^L)^2 + 0.008\, (b_1^L)^2 . \qquad (28)$$
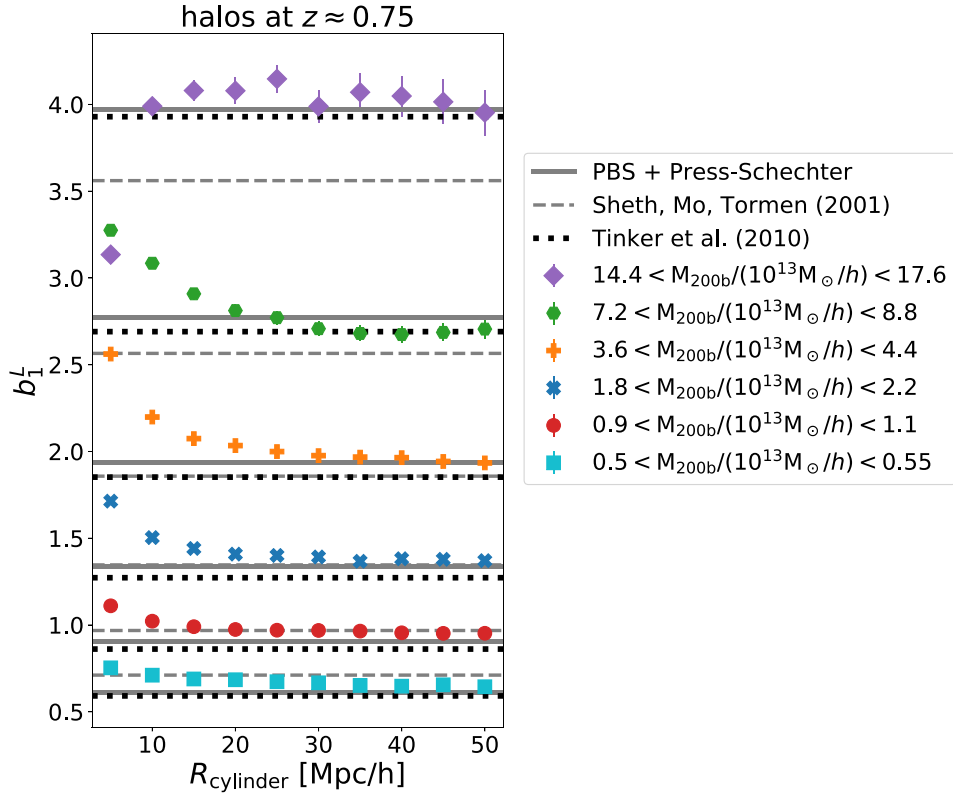
Figure 5. Measurement of linear Lagrangian bias through fits to $\langle \delta_{\mathrm{halo}} | \delta_{\mathrm{m}} \rangle$ in T17 simulated data. Different symbols (and colours) correspond to different halo mass bins and the *x*-axis represent the radius of the smoothing aperture used to measure $\langle \delta_{\mathrm{halo}} | \delta_{\mathrm{m}} \rangle$. Different horizontal lines correspond to different predictions of $b_1^L(M_{\mathrm{halo}})$ (see main text for details). We chose to leave those lines uncoloured for aesthetic reasons.

This relation indeed closely describes our measurements of b_1^L and b_2^L. This is encouraging and confirms that the bias parameters one would measure from our Lagrangian formalism in a PDF-type analysis indeed correspond to the bias parameters that have been investigated in other contexts. This is particularly important when considering combined analysis of the joint PDF $p(\delta_{\mathrm{tracer}}, \delta_{\mathrm{m}})$ and other summary statistics of the cosmic density field. But the agreement seen in Fig. 6 does unfortunately not mean that one can hope to eliminate one free parameter from our bias model. The tracers of the cosmic density field available in real analyses are galaxies, and in order to make use of the relation observed in Fig. 6 for such analyses one would have to model the HOD of these galaxies, which in itself would introduce a plethora of free parameters (see e.g. Dvornik et al. 2018). Hence, the strategy we aim for in future data analyses is to fit effective bias parameters for the tracer samples at hand, as we have e.g. done in Section 4.1. This is also why we do not further pursue accurate modelling of b_1^L as a function of halo mass.

In a next step, we want to check for consistency between the bias parameters measured from our Lagrangian and Eulerian models for $\langle \delta_{\mathrm{halo}} | \delta_{\mathrm{m}} \rangle$. In the large-scale limit b_1^L, b_2^L and b_1^E, b_2^E should be related by

$$b_1^E \approx 1 + b_1^L, \quad b_2^E \approx \frac{\nu - 1}{\nu} b_1^L + b_2^L, \tag{29}$$

where typically one assumes $\nu = 21/13$ (Wagner et al. 2015; Lazeyras et al. 2016). The derivation of that value uses a spherical collapse approximation which may not be appropriate for our cylindrical apertures. Substituting spherical with cylindrical collapse one arrives at $\nu = 7/5$ (Uhlemann et al. 2018c). In our situation we find both

values for ν to give very similar values of b_2^E (as calculated from b_1^L and b_2^L) and for cylinders of finite length, the truth is anyway expected to lie between both choices (see again Uhlemann et al. 2018c). So in the following we will stick with the spherical value such that $(\nu - 1)/\nu = 8/21$. In Fig. 7 we plot our measurements of b_1^E as a function of $1 + b_1^L$ and our measurements of b_2^E as function of $(\frac{8}{21} b_1^L + b_2^L)$. Different colours again represent different halo mass bins and different symbols represent different aperture radii. One can see that for our largest aperture (50 Mpc/*h*) the agreement with the relation (29) is indeed excellent (note that the measurement uncertainties of b_2^E and b_2^L are highly correlated, which is the reason why the measurements in the bottom panel are suspiciously spot on). This demonstrates that the machinery we have developed in Section 2 indeed represents a sensible Lagrangian bias model for PDF statistics.

Finally, we want to compare our linear bias values obtained from $\langle \delta_{\mathrm{halo}} | \delta_{\mathrm{m}} \rangle$ to the halo biases that would be inferred from measurements of large-scale two-point statistics (cf. Manera & Gaztañaga 2011, who have performed an analogous study for spherical apertures in simulation snapshots). Since we are using the T17 data in radial shells, we will consider the angular power spectra of the matter density and halo density fields in these shells projected on to the sky. Let C_ℓ^{mm} be the auto power spectrum of the matter density field in a particular shell, and let C_ℓ^{hm} be the cross power spectrum of matter and halo density. Following a similar procedure to that of Lazeyras et al. (2016) we assert that those are related by

$$C_\ell^{hm} \approx (b_1^{\mathrm{2pt}} + b_{\mathrm{NL}}^{\mathrm{2pt}} \ell^2) \, C_\ell^{mm}. \tag{30}$$
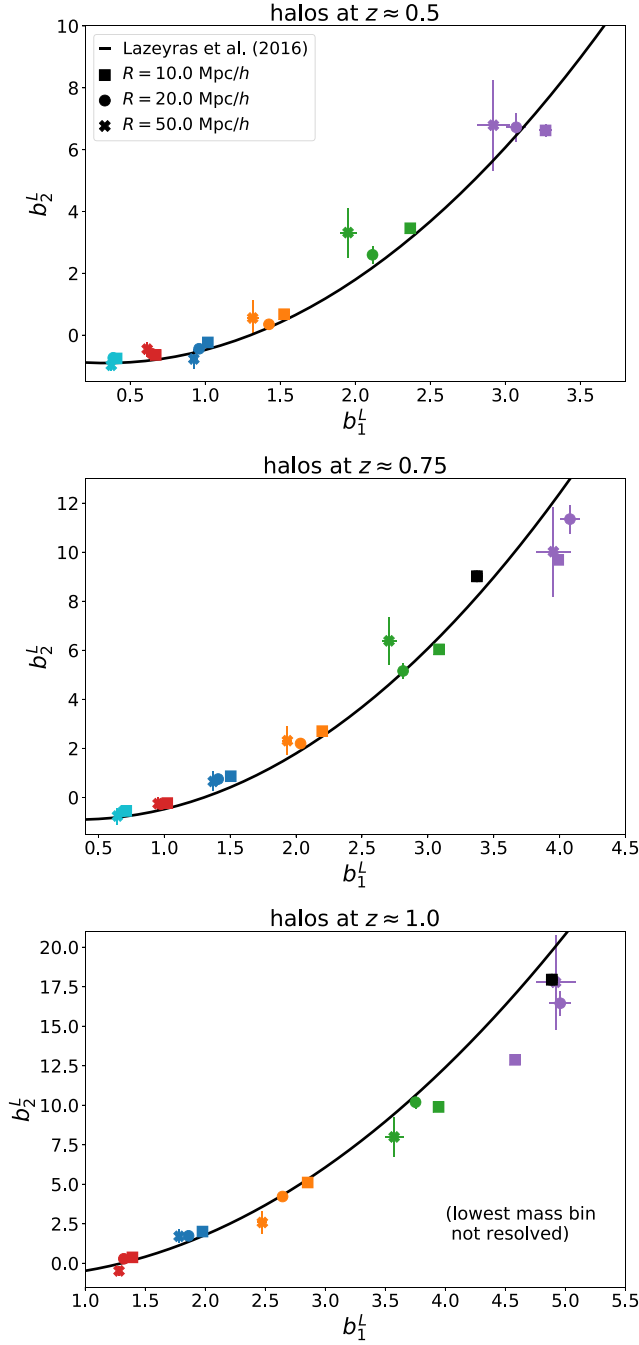
Figure 6. Displaying our measurements of b_2^L from $\langle\delta_{\text{halo}}|\delta_m\rangle$ as a function of the corresponding measurements of b_1^L for different halo mass bins, different radii of our smoothing aperture, and in different redshift shells of the T17 sims. The colour coding of the mass bins is identical to that of Fig. 5 (higher bias values correspond to higher masses). The solid lines represent an empirical relation between linear and quadratic bias found by Lazeyras et al. (2016) using a response approach in separate universe simulations (see main text for details).

Here $b_1^{2\text{pt}}$ is the linear Eulerian halo bias (in the two-point function context) and the term proportional to $b_{\text{NL}}^{2\text{pt}} \ell^2$ aims to capture corrections from non-linear (resp. scale dependent) bias. We fit the above relation to measurements of C_ℓ^{hm} and C_ℓ^{mm} in T17 data. This has the advantage that we do not need to employ any analytic modelling of the involved power spectra. Also, it removes the dependence of the
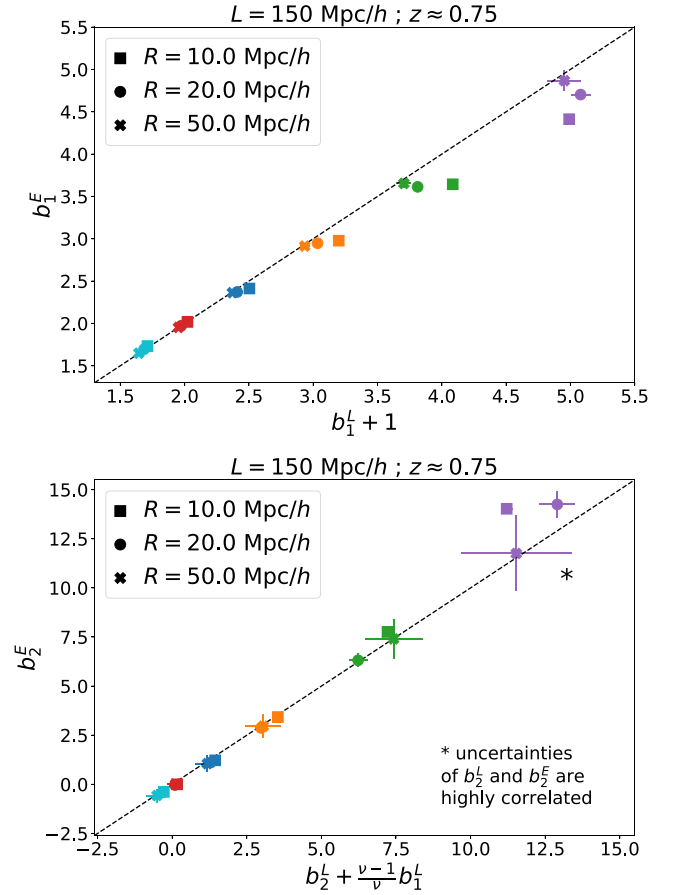

Figure 7. Testing whether standard relations between Lagrangian and Eulerian bias coefficients hold for our measurements of these coefficients from $\langle\delta_{\text{halo}}|\delta_m\rangle$ in the large-scale limit (see main text for details). Different colours again represent different halo mass bins and the colour coding is the same as that in Fig. 5. In the lower panel, the agreement between Eulerian and Lagrangian parameters looks suspiciously good, given the statistical uncertainties of our fits. This is caused by the fact that the measurement uncertainties for both sets of parameters are highly correlated.

fit on shot-noise. We follow Lazeyras et al. (2016) in restricting the fit to co-moving wave numbers below $k = 0.06$ h/Mpc. Coincidentally, this roughly corresponds to real space scales of $\pi/k \gtrsim 50$ Mpc/h, i.e. to about the largest aperture radius in which we have measured $\langle\delta_{\text{halo}}|\delta_m\rangle$. If w is the average co-moving distance of a shell, then the co-moving wavenumber k probed by an angular mode ℓ is approximately ℓ/w. Hence, we restrict ourselves to modes $\ell \leq 0.06 \cdot w$ h/Mpc.

To estimate the statistical uncertainties of this fit, let us assume that bias is perfectly linear, and that both the matter density and halo density field are Gaussian random fields. These assumptions are likely sufficient for our two-point analysis, since the power spectrum covariance at small scales (where the assumptions may break down) will be dominated by shot-noise (cf. Friedrich et al. 2021). This will especially be the case for the narrow bins in halo mass that we consider here.

A measurement of the matter auto power spectrum will be given by

$$\hat{C}_\ell^{mm} = \frac{1}{2\ell+1} \sum_{M=-\ell}^{\ell} |a_{\ell M}|^2 \tag{31}$$
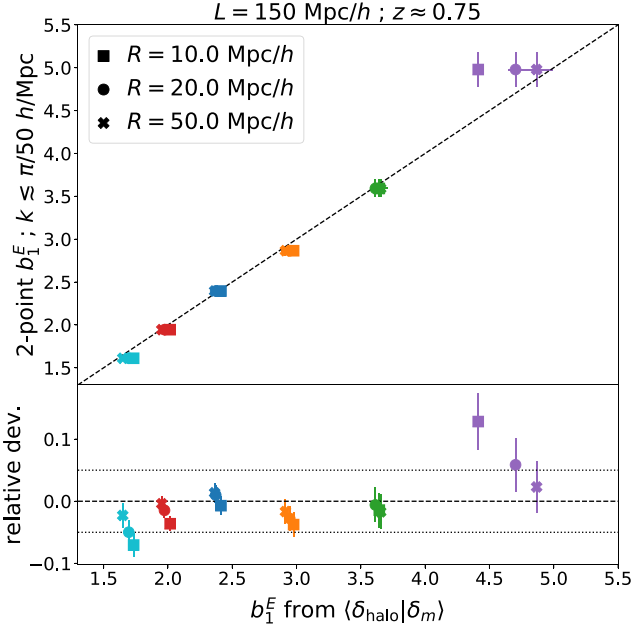
Figure 8. Testing whether the Eulerian linear bias b_1^E measured from $\langle \delta_{\mathrm{halo}} | \delta_{\mathrm{m}} \rangle$ agrees with the bias measured from comparing the auto power spectrum of matter density fluctuations to the cross power spectrum of matter and halo density fluctuations (see main text for details, the lower panel shows relative deviations between the two sets of measurements). Different colours again represent different halo mass bins (cf. Fig. 5 for the colour coding). To perform the power spectrum fits we only considered scales with $\pi/k \gtrsim 50$ Mpc/h.

where $a_{\ell M}$ are the spherical harmonics coefficients of the matter density field projected on to the sky. Similarly, a measurement of C_ℓ^{hm} will be given by

$$\hat{C}_\ell^{hm} = \frac{1}{2\ell + 1} \sum_{M=-\ell}^{\ell} a_{\ell M}^* \left(b_{\ell M} + \epsilon_{\ell M} \right) \qquad (32)$$

where $\epsilon_{\ell M}$ represents shot-noise and $b_{\ell M}$ are the spherical harmonics coefficients of the (hypothetical) shot-noise free halo density field. We need to know the covariance matrix of

$$\hat{C}_\ell^{hm} - b_1^{\mathrm{2pt}} \hat{C}_\ell^{mm} \approx \frac{1}{2\ell + 1} \sum_{M=-\ell}^{\ell} a_{\ell M}^* \, \epsilon_{\ell M}. \qquad (33)$$

Within our Gaussianity and linearity assumption it is easy to see that this covariance is diagonal and that the variances for each value of ℓ are given by

$$\mathrm{Var}\left(\hat{C}_\ell^{hm} - b_1^{\mathrm{2pt}} \hat{C}_\ell^{mm} \right) = \frac{C_\ell^{mm}}{(2\ell + 1) n_{\mathrm{halo}}} \approx \frac{\hat{C}_\ell^{mm}}{(2\ell + 1) n_{\mathrm{halo}}}. \qquad (34)$$

Here, n_{halo} is the number density of haloes (projected on to the sky) and we have assumed that the shot-noise is uncorrelated to the underlying matter density field. So the figure of merit that we are optimizing in order to fit for the bias parameters in equation (30) is

$$\chi^2[b_1^{\mathrm{2pt}}, b_{\mathrm{NL}}^{\mathrm{2pt}}] \approx \sum_{\ell < \ell_{\max}} \frac{\left(\hat{C}_\ell^{hm} - (b_1^{\mathrm{2pt}} + b_{\mathrm{NL}}^{\mathrm{2pt}} \, \ell^2) \, \hat{C}_\ell^{mm} \right)^2}{\hat{C}_\ell^{mm}} (2\ell + 1) n_{\mathrm{halo}}. \qquad (35)$$

The best-fitting reduced χ^2 values we obtain this way indeed scatter closely around 1. In Fig. 8 we display the corresponding best-fitting

values of b_1^{2pt} as a function of b_1^E obtained from the conditional expectation value $\langle \delta_{\mathrm{halo}} | \delta_{\mathrm{m}} \rangle$ in different halo mass bins (the same bins and colour coding as before). Different symbols again correspond to different aperture radii. One can indeed see that the two types of bias measurements agree in the large-scale limit. For our largest aperture radius, the relative agreement is better than 3 per cent in all mass bins and within the statistical uncertainties of the two-point fit. The systematic shift of b_1^E when going to smaller radii does not necessarily signify a systematic difference between two-point function and PDF biases but rather implies a general scale-dependence of bias when moving to smaller scales. Note especially, that we only measured our power spectra on scales of $\pi/k \gtrsim 50$ Mpc/h. Allowing the two-point fit to use even smaller scales leads to a shift in b_1^{2pt} similar to that observed in $\langle \delta_{\mathrm{halo}} | \delta_{\mathrm{m}} \rangle$.

Lazeyras et al. (2016) perform a similar test also for the quadratic bias coefficients. This would require us to either model the non-linear part of the halo power spectra or to measure complicated combinations of bispectra in the T17 shell. We do not attempt that because we take our comparison for the linear coefficients in combination with the results obtained for the quadratic coefficients in Figs 6 and 7 as sufficient indication that our language for quadratic bias in the PDF agrees with the parametrizations that appear in more standard contexts.

4.3 Shot-noise of haloes and galaxies

We conclude this section by investigating the shot-noise of our different tracer samples in more detail. The upper panel of Fig. 9 plots the ratio $\mathrm{Var}(N_{\mathrm{tracer}} | \delta_{\mathrm{m}}) / \langle N_{\mathrm{tracer}} | \delta_{\mathrm{m}} \rangle$ measured in the T17 shell with $z \approx 0.75$ and with cylindrical apertures of radius $R = 20$ Mpc/h. The bins in δ_{m} are the same as those we have considered for $\langle \delta_{\mathrm{tracer}} | \delta_{\mathrm{m}} \rangle$ in the previous subsections and the statistical uncertainties have been estimated using the same jackknife procedure as before. The dark blue band in the figure represents the ratio measured for the synthetic galaxy sample described in Section 3.2 while the semi-transparent bands represent the same halo mass bins as considered previously. For Poissonian shot-noise, the ratios $\mathrm{Var}(N_{\mathrm{tracer}} | \delta_{\mathrm{m}}) / \langle N_{\mathrm{tracer}} | \delta_{\mathrm{m}} \rangle$ should be equal to 1. For the different halo mass bins it is slightly below that, with the variances of halo counts being on average about 3 per cent below the Poisson value and with a slight increase of this effect towards higher matter densities. For our synthetic galaxies the situation is quite different: they show variances that are up to 8 per cent above the Poissonian value for negative δ_{m}, which then steeply fall to give sub-Poissonian variances for positive δ_{m}. Our HOD prescription of Section 3.2 should in principle return a weighted average over haloes of different masses and at a first glance it is surprising that this would give such a qualitatively and quantitatively different behaviour of shot-noise compared to the individual mass bins.

To understand this in more detail, note that for halo bins with a very narrow mass range the ratio $\mathrm{Var}(N | \delta_{\mathrm{m}}) / \langle N | \delta_{\mathrm{m}} \rangle$ will always tend to 1. This is because for a small enough mass range, there will always be either 0 or 1 halo in any of our apertures. And in that case, the shot-noise becomes binomial with number of trials $N_{\mathrm{trial}} = 1$ and probability of failure $1 - p \approx 1$, which results in $\mathrm{Var}(N | \delta_{\mathrm{m}}) / \langle N | \delta_{\mathrm{m}} \rangle \approx 1$. Now what happens to the ratio when summing over many of these narrow bins? Let $N_i, i = 1, \ldots, n$ be the counts of n of such narrow halo bins in our aperture and consider their sum

$$N = \sum_{i=1}^{n} N_i. \qquad (36)$$
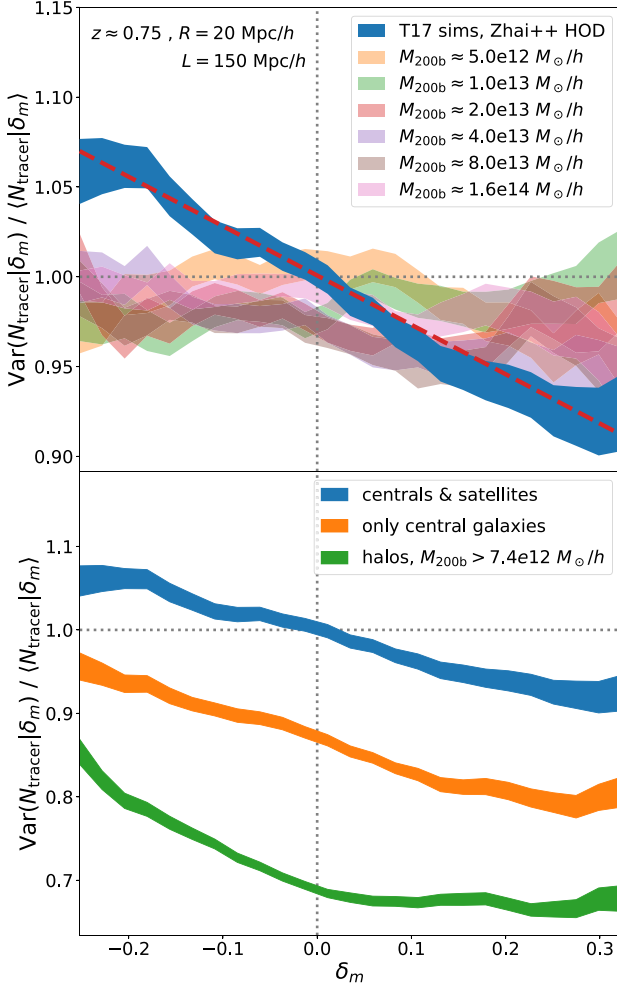
Figure 9. Upper panel: Ratio between the variance and expectation value of tracer counts in cylindrical apertures as a function of the matter density contrast in those apertures. For Poisson shot-noise this ration should be equal to 1. The red dashed line represents a linear fit to the ratio observed for our synthetic galaxy sample in the T17 shell. Lower panel: Same ratio but considering different tracer samples.

Obviously, the expectation value of N is just the sum of the expectation values of the N_i,

$$\langle N | \delta_{\mathrm{m}} \rangle = \sum_{i=1}^{n} \langle N_i | \delta_{\mathrm{m}} \rangle, \tag{37}$$

where we have inserted a dependence on δ_{m} to be closer to our situation of interest. For the variance of N the situation is more complicated since

$$\mathrm{Var}(N|\delta_{\mathrm{m}}) = \sum_{i,j} \mathrm{Cov}(N_i, N_j | \delta_{\mathrm{m}})$$

$$\approx \sum_i \langle N_i | \delta_{\mathrm{m}} \rangle + \sum_{i \neq j} \mathrm{Cov}(N_i, N_j | \delta_m)$$

$$\Rightarrow \frac{\mathrm{Var}(N|\delta_{\mathrm{m}})}{\langle N | \delta_{\mathrm{m}} \rangle} \approx 1 + \frac{\sum_{i \neq j} \mathrm{Cov}(N_i, N_j | \delta_{\mathrm{m}})}{\langle N | \delta_{\mathrm{m}} \rangle}. \tag{38}$$

Now each of the finite mass bins in the upper panel of Fig. 9 can be seen as a sum over many, even narrower mass bins. In order for the shot-noise of the finite mass bins to be sub-Poisson we would

hence need the covariance between the narrow bins to be negative. In the following intuitive sense this would indeed be expected: if there are already a lot of haloes of one bin in our aperture, one would expect less mass to be left for forming other haloes which would cause negative correlations among the shot-noise of the two mass bins. This is also in line with arguments of halo-exclusion (e.g. Baldauf et al. 2013, 2021) and with the finding that certain weighting schemes among halo masses can reduce tracer stochasticity (e.g. Hamaus et al. 2010; Jee et al. 2012; Uhlemann et al. 2018a). In particular, Ginzburg, Desjacques & Chan (2017) noted that within a halo model description halo shot-noise would also be present in the matter density field itself (cf. their equation 29). That shot-noise would be positively correlated with the noise of any tracer sample, such that the relative noise between tracers and matter would indeed be lower than naive expectations.

We could in principle estimate the covariance $\mathrm{Cov}(N_i, N_j | \delta_{\mathrm{m}})$ from our simulated data. Unfortunately, for very narrow mass bins such an estimate will be extremely noisy, because the standard deviation of off-diagonal elements of the estimate will be proportional to diagonal elements of the covariance (Taylor et al. 2013) which are significantly higher than the off-diagonal elements in the limit of narrow bins. Nevertheless, to qualitatively test our above considerations, we measure the covariance of the shot-noise of wide mass bins instead. We choose those to be centred around the same masses as our previous bins, but widen the mass ranges to touch each other (but we keep the binning logarithmic). We then re-fit the Eulerian bias model to these new bins and apply the best-fitting parameters to the dark matter density field δ_{m} in the T17 data. This way we effectively obtain a shot-noise free estimate of the halo density field which we can then subtract from the actual halo density field to obtain shot-noise-only maps.

In the upper panel of Fig. 10 we show the covariance matrix of these shot-noise-only maps, divided by the mean number count of all haloes (i.e. the matrix $\mathrm{Cov}(N_i, N_j)/\langle N \rangle$ appearing on the right-hand side of equation (38)), using again the T17 shell at $z \approx 0.75$ and filtering with $R_{\mathrm{cyl}} = 20$ Mpc/h. All off-diagonal elements of this matrix are indeed negative. We can furthermore split the T17 shell into regions of positive and negative δ_{m}. The middle panel of Fig. 10 shows $\mathrm{Cov}(N_i, N_j)/\langle N \rangle$ obtained only from overdense regions while the lower panel uses only underdense regions. Most of the elements of $\mathrm{Cov}(N_i, N_j)/\langle N \rangle$ are more negative for $\delta_{\mathrm{m}} > 0$ than they are for $\delta_{\mathrm{m}} < 0$. From this behaviour of the shot-noise correlation matrix we can draw the following qualitative conclusions: we expect the shot-noise of haloes with a wide mass range to be even more sub-Poissonian than what we observed for our narrow mass bins in the upper panel of Fig. 9. And we expect the shot-noise of wide halo bins to be more sub-Poissonian in overdense regions than in underdense regions.

These qualitative statements are indeed confirmed by the green band in the lower panel of Fig. 9, which shows the ratio $\mathrm{Var}(N_{\mathrm{tracer}})/\langle N_{\mathrm{tracer}} \rangle$ for a halo mass bin that includes all haloes that enter our HOD as described in Section 3.2 (i.e. all haloes with masses $M_{\mathrm{200b}} > 7.4 \times 10^{12}$ M$_{\odot}$/h). The shot-noise of that pure halo sample is strongly sub-Poissonian and the ratio $\mathrm{Var}(N_{\mathrm{tracer}})/\langle N_{\mathrm{tracer}} \rangle$ is lower in overdense regions than it is in underdense regions. In that panel, we also plot the shot-noise behaviour of the central galaxies within our mock galaxy sample (orange band; i.e. those galaxies that are central to their host halo, cf. Section 3.2) as well as the behaviour of the full synthetic galaxy sample (blue band; same as in upper panel). These bands show a subsequent increase in the ratio $\mathrm{Var}(N_{\mathrm{tracer}})/\langle N_{\mathrm{tracer}} \rangle$ with centrals being already less sub-Poisson than the haloes and satellites showing almost Poissonian noise again. One could think that this suggest that the randomness in the HOD is increasing the
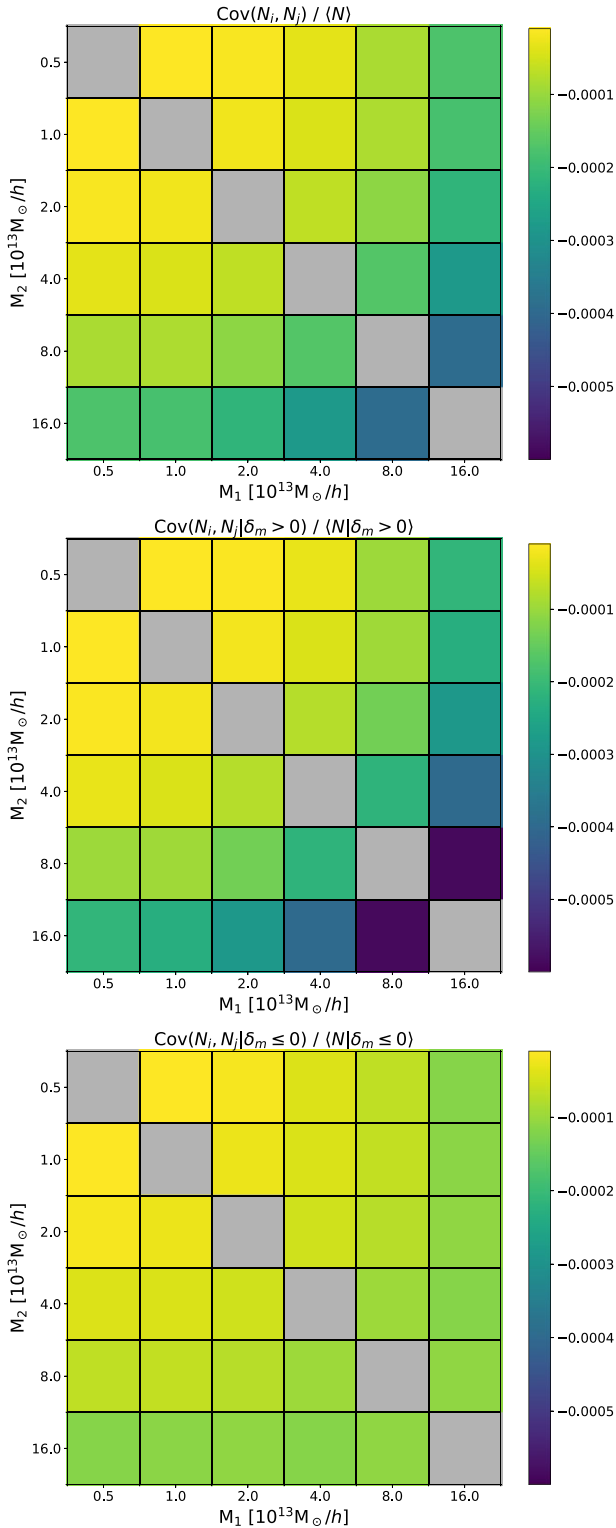
Figure 10. The covariance matrix of shot-noise-only maps for haloes in different mass bins, divided by the mean number count of all haloes (i.e. the matrix $\mathrm{Cov}(N_i, N_j)/\langle N \rangle$ appearing on the right-hand side of equation (38)). The figure uses our fiducial aperture of $R_{\mathrm{cyl}} = 20$ Mpc/h and the $z \approx 0.75$ shell of our T17 data. The upper panel measures the correlations from the full T17 shell, the middle panel only from parts of the shell where $\delta_{\mathrm{m}} > 0$ and the lower panel from parts where $\delta_{\mathrm{m}} \leq 0$.

shot-noise wrt. a pure halo sample and hence pushes the noise closer to Poisson again (or even beyond). However, the situation is more complicated as we explain in the following.

The behaviour of shot-noise in our mock LRG sample strongly differs from what has been observed in a different mock sample by Friedrich et al. (2018) or even for real Dark Energy Survey (DES) galaxies by Gruen et al. (2018). They have observed super-Poissonian noise for redMaGiC(-like) galaxies (Rozo et al. 2016) that increases with increasing matter density. Since this is so different from our findings, we want to cross check the latter wrt. data from a different N-body simulation – the Quijote suite (cf. Section 3.3). For that suite we only have a snapshot available at $z = 0.5$. So in order to compare our Quijote results to the T17 results we also repeat some of our measurements in the T17 shell at $z \approx 0.476$. We populate haloes in both of these data sets with the same HOD as before (see also Hahn & Villaescusa-Navarro 2021 for the general methodology) and we again consider the ratio $\mathrm{Var}(N_{\mathrm{tracer}})/\langle N_{\mathrm{tracer}} \rangle$ for haloes (with $M_{200\mathrm{b}} > 7.4 \times 10^{12}$ M$_\odot$/h), for central galaxies and for the full mock galaxy samples. This is not entirely realistic, since the HOD description of Zhai et al. (2017) has been specifically fit to LRGs at $z \approx 0.6$–0.9, but it should nevertheless suffice for a qualitative comparison. Note also that we have only had access to M_{vir} for the Quijote haloes, instead of $M_{200\mathrm{b}}$. But we find that a lower mass cut at $M_{\mathrm{vir}} = 6.986 \times 0^{12}$ M$_\odot$/h within the T17 sims gives a similar halo density as the cut in $M_{200\mathrm{b}}$, so we apply this M_{vir} cut in Quijote.

The upper panel of Fig. 11 shows the behaviour of shot-noise for the three different tracer samples in the Quijote data while the lower panel shows the measurements from the T17 data. One feature that persists in both data sets compared to what we found in Fig. 9 is that satellite galaxies show an increase of $\mathrm{Var}(N_{\mathrm{tracer}})/\langle N_{\mathrm{tracer}} \rangle$ wrt. central galaxies that is almost independent of the total matter density δ_m in the smoothing aperture. But the shot-noise behaviour of the halo samples is quite different both between Quijote and T17 and compared to the $z \approx 0.75$ shell of T17. For all halo samples we considered there is a significant curvature of $\mathrm{Var}(N_{\mathrm{tracer}})/\langle N_{\mathrm{tracer}} \rangle$ as a function of δ_{m}. But that curvature is strongest for the Quijote haloes and even causes them to be super-Poissonian at very high densities. We could not find an obvious explanation for this difference but assume that it is caused by the different cosmologies at which the simulations are run (cf. Appendix C for a comparison of the halo mass functions of the two simulations, and Fig. C1 where it is shown that Quijote has significantly more high-mass haloes). Given a precise measurement of the covariance $\mathrm{Cov}(N_i, N_j)$ of halo shot-noise in narrow mass bins as well as a model for the halo mass function and a given HOD we could in principle model $\mathrm{Var}(N|\delta_m)/\langle N|\delta_m \rangle$ exactly. There is, however, a number of practical reasons that prevent us from doing so:

(i) As mentioned earlier in this section, measuring $\mathrm{Cov}(N_i, N_j)$ in sufficiently narrow bins will require a prohibitively large amount of simulations. Alternatively one could attempt to model the shot-noise covariance, but as of now no such model is available.

(ii) HOD descriptions themselves make the assumption that satellite counts in a given halo are drawn from a Poisson distribution. This assumption is similarly *ad hoc* as the assumption that galaxies are Poissonian tracers of the matter density field (see e.g. Boylan-Kolchin et al. 2010; Mao, Williamson & Wechsler 2015, who indeed find non-Poissonianity in the occupation distribution of sub-haloes).

(iii) HOD descriptions also introduce a large number of free parameters which – unless they can be constrained *a priori* – may
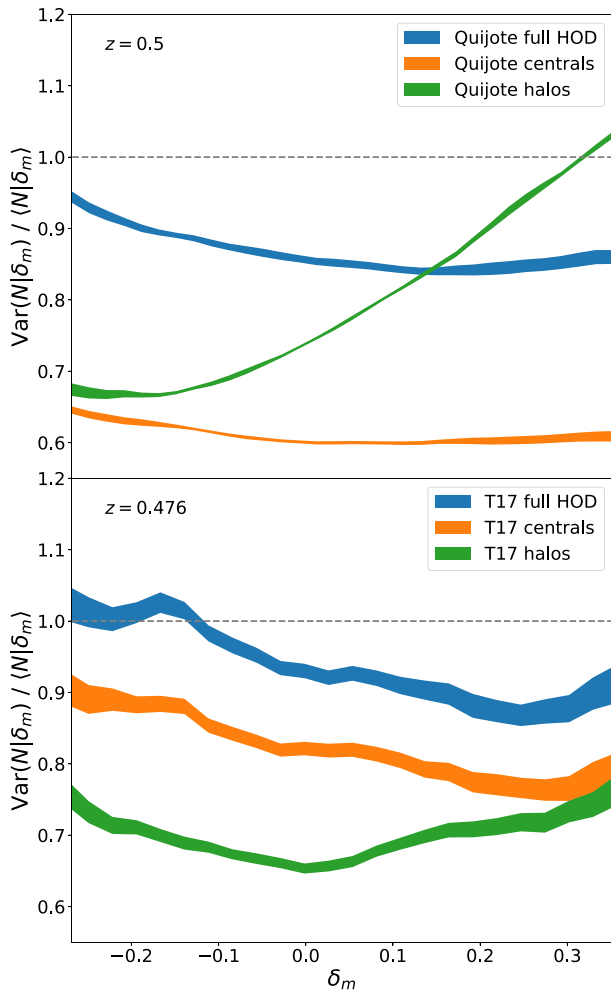
Figure 11. Same as the lower panel of Fig. 9 but at $z = 0.5$ in the Quijote simulations (upper panel) and $z \approx 0.476$ in our T17 mock galaxy catalogue (lower panel).

significantly dilute the cosmological constraining power of PDF analyses (or at least make them significantly more complicated).

Given our current (poor) understanding of shot-noise we hence conclude that effective parametrizations of non-Poisson shot-noise such as the one described in Section 2 are the most promising way forward for PDF analyses. In Fig. 12 we show that this parametrization can indeed capture the impact of the non-Poissonianity observed in Figs 9 and 11 on the joint PDF of δ_m and δ_g. The blue, filled contours in Fig. 12 represent measurements of the PDF in our different mock data sets. For the black dashed lines we have fit a linear slope to our measurements of $\mathrm{Var}(N|\delta_m)/\langle N|\delta_m\rangle$ in order to determine the parameters α_0 and α_1 of our fiducial model presented in Section 2 (cf. the red dashed line in the upper panel of Fig. 9 and analogous fits for the other tracer samples). The red dash–dotted contours represent an alternative model that assumes that shot-noise is exactly Poissonian. For the centrals in our fiducial T17 shell (i.e. $z \approx 0.75$) such a model clearly overestimates the vertical width of the distribution (cf. upper panel). But for the full sample the values of $\mathrm{Var}(N|\delta_m)/\langle N|\delta_m\rangle$ become close to Poissonian again. Hence, even the Poissonian model accurately captures the shape of the joint PDF for that sample (middle panel). This is, however, only coincidental, and for our alternative sample in the Quijote simulations ($z = 0.5$, full HOD) even the PDF
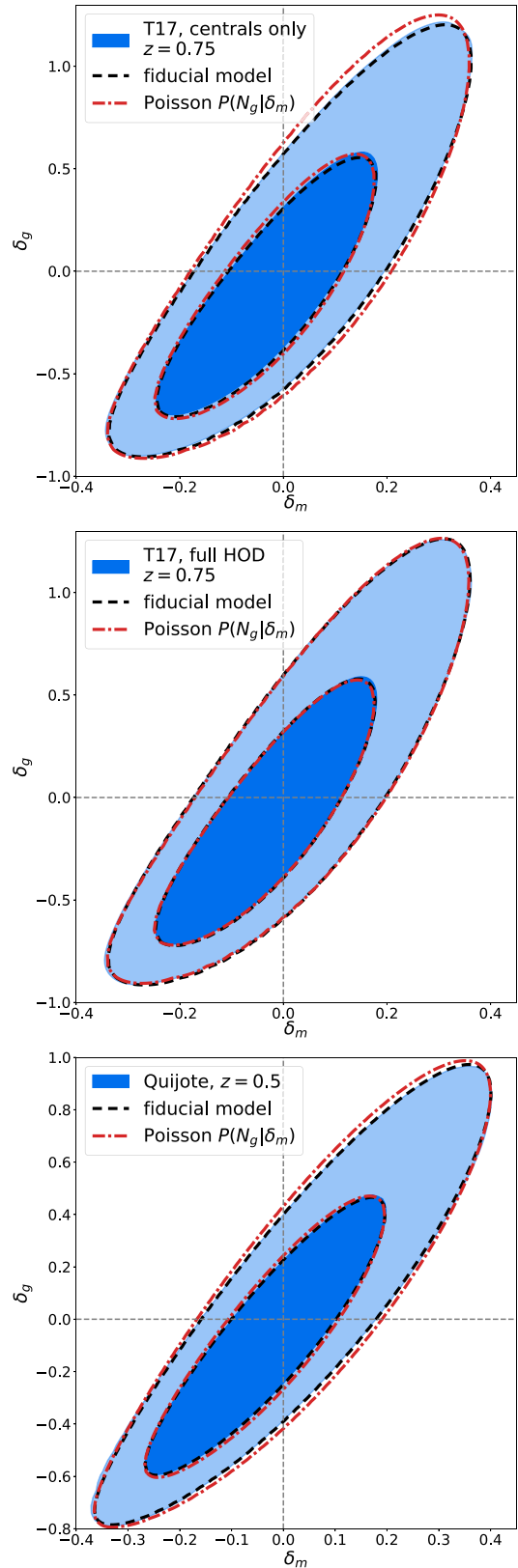
Figure 12. Impact of non-Poisson shot-noise on the joint PDF of δ_g and δ_m. The filled contours represent measurements in our mock data sets, the black dashed contours represent our fiducial model, and the red dash–dotted contours show a model that assumes Poissonian shot-noise (see main text for details).

Table 2. Best-fitting parameters of the shot-noise model from Section 2.4 for the different galaxy samples considered in Figs 11 and 12.

Tracer sample	Redshift	α_0	α_1
T17 centrals	0.751	0.872 ± 0.005	-0.329 ± 0.033
T17 all	0.751	1.015 ± 0.006	-0.310 ± 0.043
T17 centrals	0.476	0.829 ± 0.003	-0.234 ± 0.016
T17 all	0.476	0.945 ± 0.003	-0.274 ± 0.018
Quijote centrals	0.5	0.608 ± 0.001	-0.061 ± 0.004
Quijote all	0.5	0.868 ± 0.001	-0.142 ± 0.007

of the full sample is noticeably different from the Poisson model (lower panel). The best-fitting parameters of our shot-noise model for the different galaxy samples we considered are summarized in Table 2.

Understanding shot-noise remains one of the most crucial tasks in the program of fully harvesting the information content of PDF-type analyses. Our results can serve as a foundation and starting point for that but they remain qualitative. We have also only qualitatively shown that our shot-noise model is effective in capturing the behaviour of the joint PDF $p(\delta_m, \delta_g)$ (though see Friedrich et al. 2018, for a quantitative analysis of the performance of this parametrization for density split statistics). To determine the accuracy of our model quantitatively, we need to specify a target survey (and hence target statistical uncertainties) as well as an observable that can replace matter density in our pair of (δ_m, δ_g), since δ_m cannot directly be observed on real data. We leave this to the next step in our program and give a preliminary outlook in Section 5.

5 DISCUSSION

In this paper we investigated the relationship between the matter density field and its tracers from the PDF perspective, i.e. the impact of the matter-tracer connection on the joint PDF $p(\delta_{tracer}, \delta_m)$. To evaluate this PDF we considered the matter density and tracer density fields in (long) cylindrical apertures as opposed to the spherical filters that are more commonly used in the theoretical literature. This choice was motivated by the fact that the CGF of line-of-sight projected density fields can be expressed as Limber-type integral over CGFs of density fields in cylindrical apertures. Hence, it is only a small step to transfer our results to realistic observational situations of e.g. photometric galaxy surveys.

The matter-tracer connection in the PDF context can be viewed as consisting of two ingredients: the conditional expectation value of δ_{tracer} given δ_m, $\langle\delta_{tracer}|\delta_m\rangle$, and the scatter of δ_{tracer} around this expectation value which is usually referred to as shot-noise. The fiducial model for $p(\delta_{tracer}, \delta_m)$ which we present here then consists of

(i) a standard, LDT model for the PDF of matter density fluctuations $p(\delta_m)$, following the work of e.g. Bernardeau (1994), Bernardeau & Valageas (2000), Valageas (2002a), Friedrich et al. (2018), Uhlemann et al. (2018c), and Barthelemy et al. (2020);
(ii) a Lagrangian bias expansion for $\langle\delta_{tracer}|\delta_m\rangle$, incorporated into the LDT formalism;
(iii) a generalization of the Poisson distribution as proposed by Friedrich et al. (2018) and Gruen et al. (2018).

Our Figs 2 and 12 show that all of these aspects of our model are important for describing the full shape of $p(\delta_{tracer}, \delta_m)$. In the following two subsections we first summarize the results of our study and then briefly discuss open tasks for PDF cosmology.

5.1 Summary of results

We have added a number of tools and observations to the already rich subject of cosmic density PDFs:

(i) We consistently incorporated a Lagrangian bias expansion for the conditional expectation value $\langle\delta_{tracer}|\delta_m\rangle$ into the standard LDT formalism for modelling cosmic density PDFs. We also demonstrated that at the saddle point configuration of the initial density field which determines the LDT predictions (cf. the path integral in equation (8)) the operations of filtering and squaring the density field approximately commute (cf. Fig. 1). This makes it possible to evaluate the Lagrangian expansion up to second order with essentially no additional computational coast. An advantage of our Lagrangian model that we did not discuss here is that it allows one to consistently incorporate scale-dependent bias from primordial non-Gaussianity (e.g. Dalal et al. 2008; Desjacques, Seljak & Iliev 2009; Jeong & Komatsu 2009) into the LDT formalism. This is because our ansatz in Section 2.3 can be used to translate scale-dependence of b_1^L into a density dependence.

(ii) We fitted both the Langrangian and an Eulerian expansion to measurements of $\langle\delta_{halo}|\delta_m\rangle$ for different halo mass bins and at different redshifts and filtering scales in simulated data by Takahashi et al. (2017). In this way we could validate that the bias expansion we developed in Section 2 conforms to standard consistency relations between the Eulerian and Lagrangian perspective of halo bias. We also checked for the consistency of our best-fitting bias parameters with expectations from other methods: our values of b_1^L as a function of halo mass agree well with a number of different theoretical and empirical predictions; the relation we observe between b_1^L and b_2^L agrees with an empirical formula found by Lazeyras et al. (2016) in separate universe simulations; and for large smoothing scales our best-fitting linear bias converges to the corresponding parameter measured from the large-scale cross power spectrum of matter and galaxy density. This array of tests confirms that the theory we developed in Section 2 represents a sensible Lagrangian bias model and hence moves PDF analyses one step closer to being on equal footing with the more advanced field of N-point correlation functions. We also showed that for a synthetic galaxy sample mimicking eBOSS-like luminous red galaxies, the Lagrangian expansion yields a significantly better fit to $\langle\delta_{halo}|\delta_m\rangle$ than the Eulerian expansion at second order. This is, however, not a general statement and we saw indications that for very massive haloes the Eulerian expansion performs better.

(iii) We established that the deviation of shot-noise from Poisson noise in a sample of haloes with a wide mass range is determined by the covariance matrix $\text{Cov}(N_i, N_j)$ of the shot-noise of haloes in a very narrow binning of that mass range. Considering the ratio $\text{Var}(N_{tracer})/\langle N_{tracer}\rangle$ for different tracer samples in both the Quijote and T17 simulations we have found a wide variety of deviations from Poissonian shot-noise. We have however shown that our shot-noise model from Section 2.4 is effective in capturing the impact of these deviations on the joint PDF $p(\delta_g, \delta_m)$.

As mentioned in the previous section, our results on shot-noise remain qualitative and more insights may be needed to efficiently model that part of the PDF. We discuss this further in the following outline.

5.2 Open tasks for PDF cosmology

Cosmological analyses of the full shape of $p(\delta_{tracer}, \delta_m)$ can be seen as an extension of the density split statistics framework developed by

Friedrich et al. (2018) and Gruen et al. (2018). In year-1 data of the DES they have analysed a data vector consisting of (a compressed version of) the galaxy density PDF $p(\delta_{\text{tracer}})$ and a number of lensing signals that effectively probe the expectation values $\langle \delta_{\text{m}} | \delta_{\text{tracer}} \rangle$ as well as the slope of the lensing power spectrum. Moving away from these compressed statistics and directly analysing $p(\delta_{\text{tracer}}, \delta_{\text{m}})$ instead will, at any given smoothing scale, open up an entire 2D plane of data for cosmological analysis. There is a number of steps that still need to be completed to implement this program.

Cosmological constraining power: Numerous studies have shown that the cosmological information contained in the PDF of density fluctuations strongly complements the information obtained from more standard probes such as the two-point correlations of fluctuations – see e.g. Codis et al. (2016), Patton et al. (2017), Uhlemann et al. (2020), Friedrich et al. (2020), and Boyle et al. (2021) for recent examples. Some of their results, however, only apply to idealized situations where one has direct access to the matter density field and the question remains to what extent the cosmological power of the PDF carries over to realistic data sets. Boyle et al. (2021) have considered the PDF of lensing convergence, which can in principle be obtained from observations of cosmic shear. And Friedrich et al. (2018) and Gruen et al. (2018) analysed compressed statistics of the joint PDF $p(\delta_{\text{tracer}}, \delta_{\text{m}})$, showing that it has a competitive power to constrain cosmological models. But as mentioned above, their density split statistics are also sensitive to the slope of the lensing power spectrum, and that information would be lost if one would only consider $p(\delta_{\text{tracer}}, \delta_{\text{m}})$ at one smoothing scale. Two solutions to this problem would be to analyse the PDF at a number of different smoothing scales (as was e.g. done by Boyle et al. 2021) or to analyse the joint PDF of galaxy densities in apertures that are located at a finite distance (a two-point PDF, cf. Uhlemann et al. 2018a). Alternatively, one could consider combined analyses of the PDF and the two-point function. We have shown that at large scales the linear bias of a PDF analysis agrees with the large-scale bias of the tracer-matter cross power spectrum. This would suggest that a combination of a PDF-type analysis with measurements of the galaxy–galaxy lensing correlation function (gg-lensing; see e.g. Prat et al. 2018, and references therein) is a promising route to take. To efficiently analyse such a combined data vector, one will need to make contact between the shot-noise and higher order bias parameters of our PDF model and stochasticity effects and non-linear biasing in the gg-lensing correlation function. This leads us to the next point.

Improved modelling: In the model presented here, the galaxy–matter connection is described by four free parameters. While Friedrich et al. (2018) and Gruen et al. (2018) have shown that the rich information content of the PDF can constrain complex bias models, a more efficient modelling would be highly desirable. This can e.g. be achieved by choosing informative, physically motivated priors on our parameters (cf. Britt et al., Ried et al. in preparation), or by identifying consistency relations between them. For example, non-linear bias at a small scale will lead to an effective change in shot-noise at a larger scale (Philcox et al. 2020), which should lead to a relation between bias and the scale dependence of shot-noise. Understanding these kinds of relations will also enable a more fruitful combination of PDF and two-point function analyses, and the information present in the PDF may be able to constrain nuisance parameters in two-point function models.

Proof of concept: A more immediate goal that we envision as a follow-up to this study is a proof-of-concept study that demonstrates the feasibility of analysing the full shape of $p(\delta_{\text{tracer}}, \delta_{\text{m}})$ in real data. Since matter density is not directly directly observable, we aim at the joint PDF of lensing convergence and 2D-projected galaxy

density. The results of this paper can be readily generalized to such line-of-sight projected fields (see e.g. Bernardeau & Valageas 2000; Friedrich et al. 2018; Uhlemann et al. 2018c; Barthelemy et al. 2020; Boyle et al. 2021), so such an analysis is indeed within reach.

DATA AVAILABILITY

C++ and PYTHON tools to compute our model predictions are publicly available at https://github.com/OliverFHD/CosMomentum. The data for the T17 N-body simulations used in this article are publicly available at http://cosmo.phys.hirosaki-u.ac.jp/takahasi/all sky_raytracing/. Summary statistics measured in the Quijote N-body simulations are publicly available at https://github.com/franciscovi llaescusa/Quijote-simulations. The Molino mock galaxy catalogues are publicly available at https://changhoonhahn.github.io/molino/c urrent/.

REFERENCES

Baldauf T., Seljak U., Desjacques V., McDonald P., 2012, Phys. Rev., 86, 083540
Baldauf T., Seljak U., Smith R. E., Hamaus N., Desjacques V., 2013, Phys. Rev., 88, 083507
Baldauf T., Schaan E., Zaldarriaga M., 2016, J. Cosmol. Astropart. Phys., 2016, 017
Baldauf T., Codis S., Desjacques V., Pichon C., 2021, Phys. Rev., 103, 083530
Barthelemy A., Codis S., Uhlemann C., Bernardeau F., Gavazzi R., 2020, MNRAS, 492, 3420
Bautista J. E. et al., 2020, MNRAS, 500, 736
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ, 762, 109
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Bernardeau F., 1994, A&A, 291, 697
Bernardeau F., 1995, A&A, 301, 309
Bernardeau F., Reimberg P., 2016, Phys. Rev., 94, 063520
Bernardeau F., Valageas P., 2000, A&A, 364, 1
Bernardeau F., Codis S., Pichon C., 2015, MNRAS, 449, L105
Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., 2010, MNRAS, 406, 896
Boyle A., Uhlemann C., Friedrich O., Barthelemy A., Codis S., Bernardeau F., Giocoli C., Baldi M., 2021, MNRAS, 505, 2886
Clerkin L. et al., 2017, MNRAS, 466, 1444
Codis S., Pichon C., Bernardeau F., Uhlemann C., Prunet S., 2016, MNRAS, 460, 1549
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Coupon J. et al., 2012, A&A, 542, A5
Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369
Dalal N., Doré O., Huterer D., Shirokov A., 2008, Phys. Rev., 77, 123514
de Mattia A. et al., 2021, MNRAS, 501, 5616

DeRose J. et al., 2019, preprint (arXiv:1901.02401)
Desjacques V., Seljak U., Iliev I. T., 2009, MNRAS, 396, 85
Desjacques V., Jeong D., Schmidt F., 2018, Phys. Rep., 733, 1
Dvornik A. et al., 2018, MNRAS, 479, 1240
Efstathiou G., 1995, MNRAS, 276, 1425
Friedrich O., Seitz S., Eifler T. F., Gruen D., 2016, MNRAS, 456, 2662
Friedrich O. et al., 2018, Phys. Rev., 98, 023508
Friedrich O., Uhlemann C., Villaescusa-Navarro F., Baldauf T., Manera M., Nishimichi T., 2020, MNRAS, 498, 464
Friedrich O. et al., 2021, MNRAS, 508, 3125
Gil-Marín H. et al., 2020, MNRAS, 498, 2492
Ginzburg D., Desjacques V., Chan K. C., 2017, Phys. Rev., 96, 083528
Gruen D. et al., 2018, Phys. Rev., 98, 023507
Hahn C., Villaescusa-Navarro F., 2021, J. Cosmol. Astropart. Phys., 2021, 029
Hamana T., Sakurai J., Koike M., Miller L., 2015, PASJ, 67, 34
Hamaus N., Seljak U., Desjacques V., Smith R. E., Baldauf T., 2010, Phys. Rev., 82, 043515
Hamaus N., Seljak U., Desjacques V., 2011, Phys. Rev., 84, 083509
Hearin A. P. et al., 2017, ApJ, 154, 190
Ivanov M. M., Kaurov A. A., Sibiryakov S., 2019, J. Cosmol. Astropart. Phys., 2019, 009
Ivanov M. M., Simonović M., Zaldarriaga M., 2020, J. Cosmol. Astropart. Phys., 2020, 042
Jee I., Park C., Kim J., Choi Y.-Y., Kim S. S., 2012, ApJ, 753, 11
Jeong D., Komatsu E., 2009, ApJ, 703, 1230
Kaiser N., 1984, ApJ, 284, L9
Lazeyras T., Wagner C., Baldauf T., Schmidt F., 2016, J. Cosmol. Astropart. Phys., 2016, 018
Limber D. N., 1953, ApJ, 117, 134
Manera M., Gaztañaga E., 2011, MNRAS, 415, 383
Mao Y.-Y., Williamson M., Wechsler R. H., 2015, ApJ, 810, 21
Mukhanov V., 2005, Phys. Found. Cosmol., Cambridge University Press
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, MNRAS, 396, 19
Pandey S. et al., 2020, Phys. Rev. D, 102, 12
Patton K., Blazek J., Honscheid K., Huff E., Melchior P., Ross A. J., Suchyta E., 2017, MNRAS, 472, 439
Philcox O. H. E., Eisenstein D. J., O'Connell R., Wiegand A., 2020, MNRAS, 491, 3290
Prat J. et al., 2018, Phys. Rev. D, 98, 4
Press W. H., Schechter P., 1974, ApJ, 187, 425
Repp A., Szapudi I., 2020, MNRAS, 498, L125
Ross A. J. et al., 2020, MNRAS, 498, 2354
Rozo E. et al., 2016, MNRAS, 461, 1431
Salvador A. I. et al., 2019, MNRAS, 482, 1435
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Shirasaki M., Hamana T., Yoshida N., 2015, MNRAS, 453, 3043
Smith R. E. et al., 2003, MNRAS, 341, 1311

Springel V., 2005, MNRAS, 364, 1105
Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152
Takahashi R., Hamana T., Shirasaki M., Namikawa T., Nishimichi T., Osato K., Shiroyama K., 2017, ApJ, 850, 24
Tamone A. et al., 2020, MNRAS, 499, 5527
Taylor A., Joachimi B., Kitching T., 2013, MNRAS, 432, 1928
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
Uhlemann C., Codis S., Pichon C., Bernardeau F., Reimberg P., 2016, MNRAS, 460, 1529
Uhlemann C. et al., 2018a, MNRAS, 473, 5098
Uhlemann C., Pajer E., Pichon C., Nishimichi T., Codis S., Bernardeau F., 2018b, MNRAS, 474, 2853
Uhlemann C., Pichon C., Codis S., L'Huillier B., Kim J., Bernardeau F., Park C., Prunet S., 2018c, MNRAS, 477, 2772
Uhlemann C., Friedrich O., Villaescusa-Navarro F., Banerjee A., Codis S. r., 2020, MNRAS, 495, 4006
Valageas P., 2002a, A&A, 382, 412
Valageas P., 2002b, A&A, 382, 477
Villaescusa-Navarro F. et al., 2020, The Astrophysical Journal Supplement Series, 250, 20
Wagner C., Schmidt F., Chiang C.-T., Komatsu E., 2015, J. Cosmol. Astropart. Phys., 2015, 042
Zhai Z. et al., 2017, ApJ, 848, 76

APPENDIX A: EQUATIONS OF MOTION FOR CYLINDRICAL COLLAPSE

We repeat here an appendix of Friedrich et al. (2020) about the evolution of symmetric density perturbations. In the Newtonian approximation and setting $G = 1 = c$ the evolution of spherical, cylindrical or planar perturbations δ is described by

$$\ddot{\delta} + \mathcal{H}\dot{\delta} - \frac{N+1}{N}\frac{\dot{\delta}^2}{1+\delta} = 4\pi\bar{\rho}_m a^2 \delta(1+\delta), \qquad (A1)$$

where τ is conformal time, $\mathcal{H} = \mathrm{d}\ln a/\mathrm{d}\tau$ is the conformal expansion rate, and $N = 3$ for a spherical perturbation, $N = 2$ for a cylindrical perturlation, and $N = 1$ for a planar perturbation (see Mukhanov 2005 who demonstrates this for $N = 1$ and $N = 3$). To compute the evolution of the saddle point fluctuation in Section 2.2 we choose $N = 2$ and solve equation (A1) with the initial conditions

$$\delta_i = \delta^*_{\mathrm{lin}, R_{\mathrm{lin}}} D(z_i) , \quad \dot{\delta}_i = \delta_i \, \mathcal{H}(z_i), \qquad (A2)$$

where z_i is a redshift chosen during matter domination. (In fact, in our calculation of $D(z)$ we set the radiation density Ω_r to zero and then choose $z_i = 4000$.)

APPENDIX B: CYLINDRICAL AVERAGE OF THE SQUARED LINEAR SADDLE POINT

Adjusting the results of Valageas (2002a) and Friedrich et al. (2020) to cylindrical filters, the saddle point configuration of the linear density contrast, $\delta_{\rm lin}^*$, filtered with a cylindrical aperture of radius R is given by

$$\delta_{{\rm lin},R}^* = \delta_{{\rm lin},R_{\rm lin}}^* \frac{\langle \delta_{{\rm lin},R}\, \delta_{{\rm lin},R_{\rm lin}} \rangle}{\langle \delta_{{\rm lin},R_{\rm lin}}^2 \rangle}. \tag{B1}$$

Here we have assumed Gaussian initial conditions (see Friedrich et al. 2020, for general non-Gaussian initial conditions) and we have set $\lambda_h = 0$, which is the case that is of interest for the calculation of $\langle \delta_{\rm halo}|\delta_{\rm m} \rangle$ (cf. equation (5)).

At any point $\boldsymbol{r}$ the saddle point configuration is then given by

$$\delta_{\rm lin}^*(\boldsymbol{r}) = \delta_{{\rm lin},r}^* + \frac{r}{2} \left. \frac{{\rm d}\delta_{{\rm lin},R'}^*}{{\rm d}R'} \right|_{R'=r}. \tag{B2}$$

In equation (19) we need to know the average of $\delta_{\rm lin}^*(\boldsymbol{r})^2$ in cylindrical apertures. This average can be calculated as

$$[\delta_{\rm lin}^{*\,2}]_{R_{\rm lin}} = \frac{2}{R_{\rm lin}^2} \int_0^{R_{\rm lin}} {\rm d}r\, r \left\{ \delta_{{\rm lin},r}^* + \frac{r}{2} \left. \frac{{\rm d}\delta_{{\rm lin},R'}^*}{{\rm d}R'} \right|_{R'=r} \right\}^2. \tag{B3}$$

In Fig. 1 we show that on the scales we are interested in, this full computation is well approximated by simply squaring the cylindrically averaged saddle point configuration. This approximation will bias our values of quadratic Lagrangian bias by a couple of per cent wrt. other measures of bias, which does not significantly affect the conclusions of our study.

APPENDIX C: COMPARING THE MASS FUNCTIONS OF QUIJOTE AND T17

In Fig. C1 we compare the mass function $n(M_{\rm vir})$ of the two different N-body data sets considered in Section 4.3 at $z = 0.5$ (Quijote) and

$z = 0.476$ (T17). Our reason for using $M_{\rm vir}$ is that we do not have $M_{\rm 200b}$ available for Quijote. The differences in the mass functions are likely caused by the different cosmology of the simulations – $(\Omega_{\rm m}, \Omega_{\rm b}, \sigma_8, n_{\rm s}, h) = (0.3175, 0.049, 0.834, 0.9624, 0.6711)$ for Quijote and $(0.279, 0.046, 0.82, 0.97, 0.7)$ for T17. We think that this difference in cosmology and the mass function is at least in part responsible for the differences in shot-noise behaviour of the two data sets that we observed in Section 4.3.
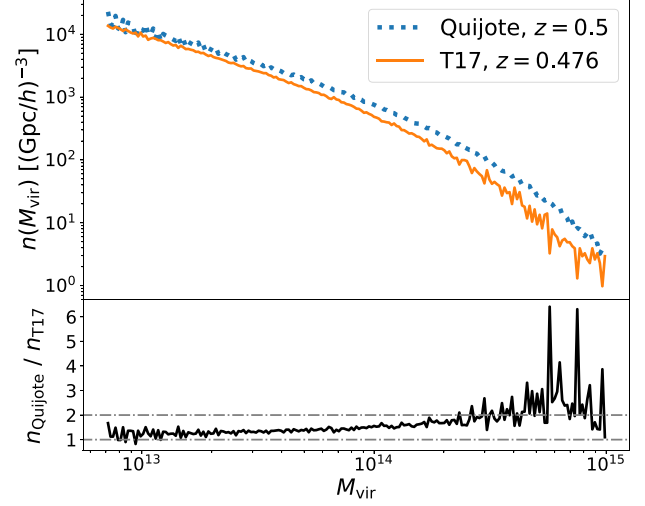
Figure C1. Upper panel: The mass function $n(M_{\rm vir})$ of the two different N-body data sets considered in Section 4.3 at $z = 0.5$ (Quijote) and $z = 0.476$ (T17). We are only plotting $n(M_{\rm vir})$ above the mass cut of $M_{\rm vir} = 6.986 \cdot 10^{12}\,{\rm M}_\odot/h$ that we considered in that section. Lower panel: Ratio of the mass functions in the two simulations.

This paper has been typeset from a TEX/LATEX file prepared by the author.

Chapter 9

Summary and future directions

The study of large scale structures (LSS) in our Universe has revealed intriguing mysteries that have necessitated the introduction of two enigmatic components in our cosmological models: *dark matter*, an invisible substance constituting majority of the Universe's mass, and *dark energy*, driving the Universe's recent accelerated expansion. Together, they constitute approximately 95% of our Universe's energy budget and significantly influence the distribution, growth and evolution of cosmic density perturbations. However, despite their exceptional importance, little is understood about their nature. Precisely pinning down their properties is therefore essential for deepening our understanding of the cosmos. One of the standard ways to do so is by measuring in wide-area galaxy surveys the 2-point correlation functions (2PCF) of the galaxy density and the cosmic shear field (minute distortions of background galaxy images by the weak gravitational lensing effect of intervening foreground LSS). Analysing 2PCFs with our cosmological models helps constrain parameters such as the amplitude of cosmic matter density fluctuations A_s, total amount of matter Ω_m, and the dark energy equation of state w_0. However, these 2PCFs are only sufficient for characterising the LSS if the density perturbations follow a *Gaussian* distribution. In reality, due to nonlinear gravitational and astrophysical processes shaping the cosmic web, the perturbations are *non-Gaussian* distributed. Consequently, the 2PCF (or the variance) alone cannot access information stored in the higher-order (beyond second) moments of LSS, rendering it inadequate. Therefore, to gain insights not captured by 2PCFs of the weak lensing and projected galaxy clustering fields, it is crucial to harness this information through higher-order statistical (HOS) methods and maximise the potential of data collected in ongoing and upcoming LSS surveys. Hence, together with my collaborators, we have pioneered novel and practical approaches to achieve this.

9.1   Key contributions of this thesis

Here, I briefly summarise the developments we have made with two higher-order statistics called the integrated 3-point correlation functions and the probability density function (PDF) of cosmic density fields in Halder et al. (2021); Halder & Barreira (2022); Gong

et al. (2023); Halder et al. (2023); Friedrich et al. (2022).

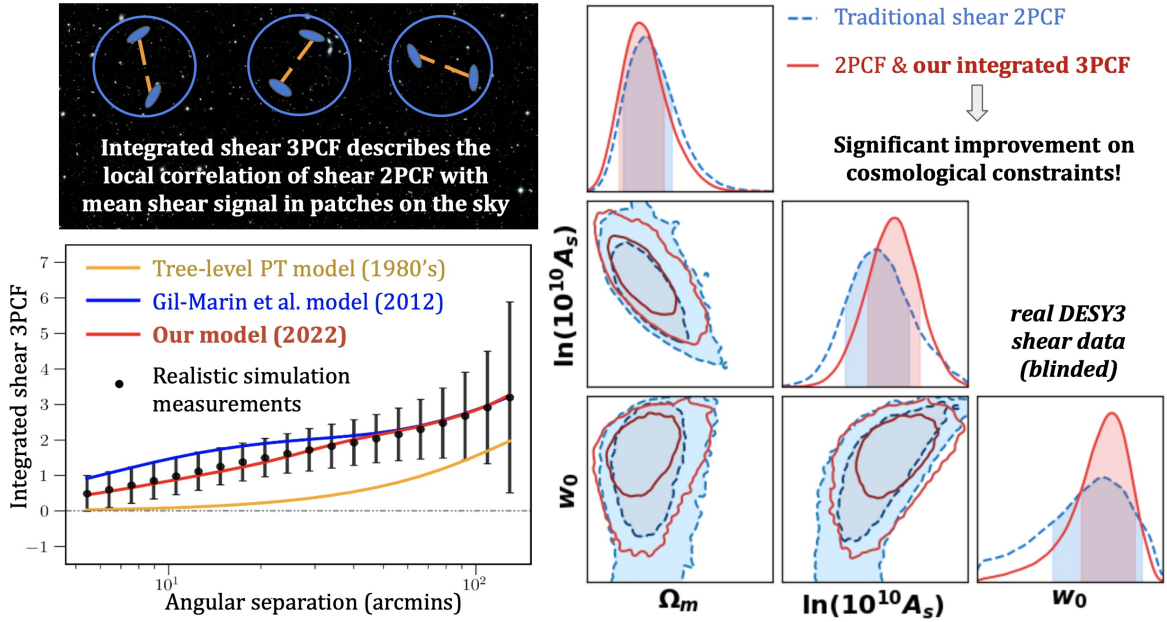## Integrated 3-point correlation functions of projected fields

Figure 9.1: Upper left: Depiction of the concept of integrated 3PCF $\zeta_\pm$. Lower left: Comparison of different ways of modelling $\zeta_\pm$ against measurements from realistic N-body weak lensing simulations (black dots with error bars reflecting DESY3 survey uncertainties) — red curve shows our accurate $\zeta_\pm$ model from Halder et al. (2023); the yellow and blue curves show other models which cannot describe $\zeta_\pm$ as accurately. Right: We have applied $\zeta_\pm$ to real DESY3 weak lensing cosmic shear data (*blinded*; Halder and DES Collaboration (in prep.)) and found that *adding $\zeta_\pm$ (red) to the shear 2PCFs $\xi_\pm$ (blue) significantly improves the precision on cosmological parameters such as the dark energy equation of state w_0.*

In chapters 3, 4, 5, we have introduced the integrated 3PCFs $\zeta_\pm$ of cosmic shear fields which are natural extensions of shear 2PCFs $\xi_\pm$ — we measure 2PCFs of the shear field locally inside patches and correlate them with the mean lensing aperture mass within the same patches. This admits a clear physical interpretation — *modulation of the small-scale 2PCFs by large-scale mean fluctuations of the cosmic shear field* which in turn probes an *integrated* form of the full 3PCF (correlations of triplets of points; or the third-order moment of the field). Using cosmological perturbation theory (PT) techniques we have developed and validated accurate theoretical models for predicting $\zeta_\pm$ (see Fig. 9.1 for an illustrative summary of the work) while accounting for various astrophysical and measurement systematic effects. In chapter 6, *we have also taken the major step to measure and analyse $\zeta_\pm$ alongside $\xi_\pm$ in the (blinded) weak gravitational lensing cosmic shear Year 3 dataset of the Dark Energy Survey DES (currently in preparation inside DES collaboration; Halder*

and DES Collaboration (in prep.)). The right panel of Fig. 9.1 illustrates one of the key results from the upcoming paper (also shown in chapter 6) demonstrating that the addition of $\zeta_\pm$ (red) to $\xi_\pm$ (blue) yields significant tightening of cosmological parameter Ω_m, A_s, w_0 constraints ($\sim$ **40% improvement on** w_0). The integrated shear 3PCF is one of the few theoretical and practical HOS that has been successfully applied to real data, thus elevating its applicability to a similar footing as the traditionally analysed 2PCF. Based on these
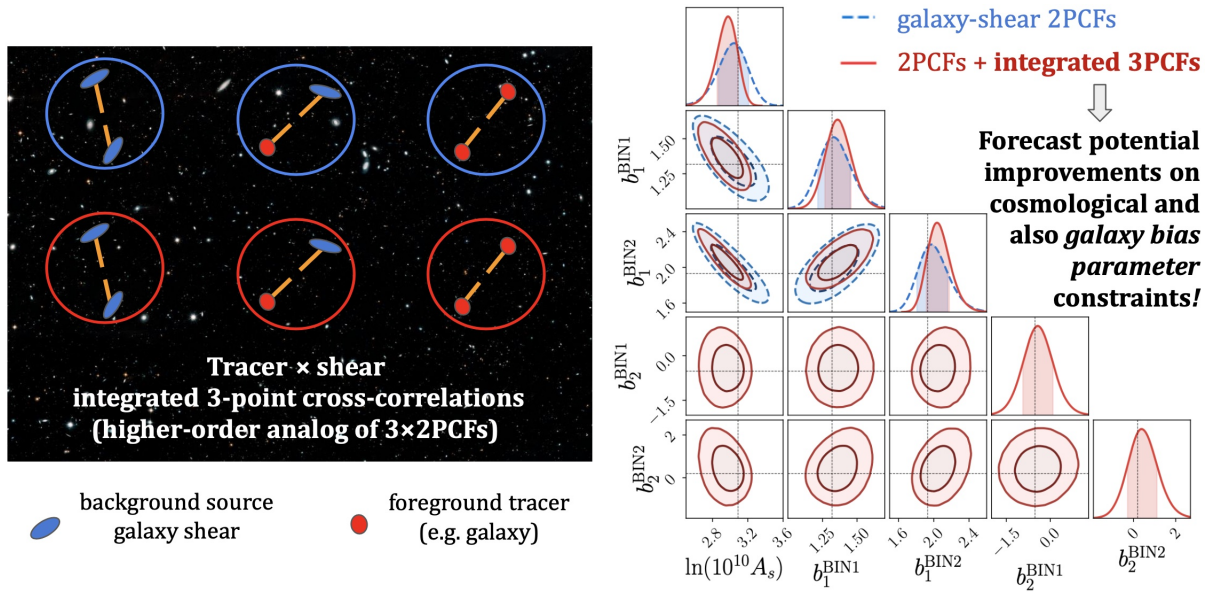

Figure 9.2: <u>Left:</u> Depiction of the concept of integrated 3PCFs involving both the background cosmic shear as well as a foreground tracer (e.g. galaxy) density field as proposed in Halder et al. (2023). <u>Right:</u> Forecast using simulated likelihood analysis on the constraints on cosmological and galaxy bias parameters for a Stage-III like lensing survey when analysing the galaxy-shear 2PCFs alone (blue dashed) and when adding the integrated 3PCFs involving the galaxy density field (red). *Addition of galaxy-shear integrated 3PCFs to the 2PCFs has the potential to significantly improve the precision on A_s as well as linear and quadratic galaxy bias parameters b_1, b_2.*

encouraging results obtained using $\zeta_\pm$, we have also proposed the integrated 3-point cross-correlations ζ between cosmic shear and the foreground galaxy density fields in chapter 7 (see Fig. 9.2 for an illustrative summary), extending the 3×2PCFs methodology to the practical higher-order integrated 3PCF framework — correlating the local measurements of 3×2PCFs with the mean signal (aperture mass or mean galaxy density) within patches. Using perturbation theory (PT) techniques validated on N-body simulations, we find that these galaxy-shear integrated 3PCFs have the potential to bring further *20-40% improvements on cosmological parameters* (e.g. A_s) even when utilising only conservatively large scales where the PT based galaxy bias modelling is valid. *Importantly, we find that these integrated 3PCFs can also extract information about the linear b_1 as well as higher-order quadratic galaxy bias b_2 parameters.* Interpreting the bias terms, for example within the

Halo-Occupation-Distribution framework, can further improve our understanding of the connection between dark matter halos and the empirical properties of the observed galaxy samples. Our results therefore motivate future applications of the galaxy-shear integrated 3PCF in the data of current and upcoming large area surveys.
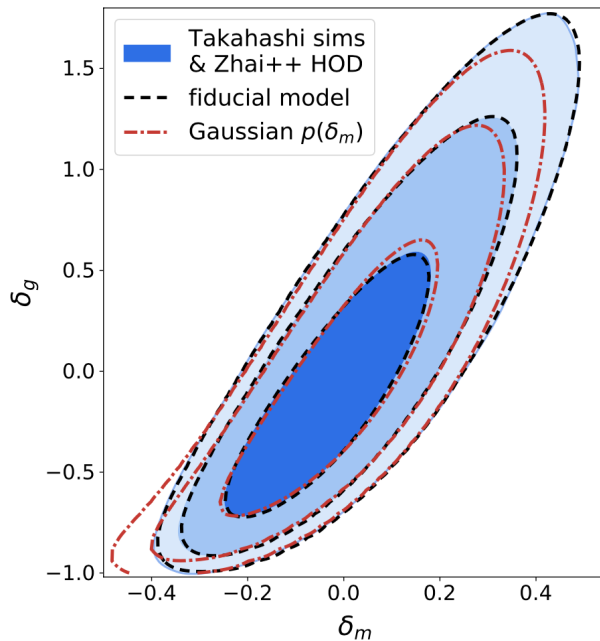
The joint PDF of matter and galaxy density fields

Figure 9.3: Illustration from Fig. 2 of chapter 8 comparing different models of the joint PDF of δ_m, δ_g fluctuations to the PDF measured in realistic N-body simulations (blue). The black-dashed contours represent our accurate theoretical modelling of the PDF, whereas the red contours show what happens to the joint PDF when assuming an incorrect Gaussian model for $p(\delta_m)$.

In chapter 8 we have put forward and accurately modelled another novel probe that is capable of obtaining a more complete view of the LSS than 2PCFs: the joint PDF $p(\delta_m, \delta_g)$ of local matter δ_m and galaxy density δ_g fluctuations (see Fig. 9.3). Unlike conventional 2-point techniques that reduce this PDF to just three quantities — the variance of δ_m, the variance of δ_g and their cross-covariance — we study the *full shape of the joint PDF* which is sensitive to information in *all* local higher-order 1-point joint-moments of these two fields. This can not only constrain cosmological parameters but also extract detailed information about the connection between the invisible dark matter and the observed galaxy density field that traces it. With applications to real data in mind, in Barthelemy et al. (2023) (not discussed in this thesis) we have also rigorously investigated the modelling of systematic effects (masking, astrophysical systematics, etc.) in the PDF $p(\kappa)$ of the lensing

convergence κ field — a projection of δ_m accessible through gravitational lensing measurements in galaxy surveys. Currently, we are in the process of measuring and analysing $p(\kappa)$ in DESY3 data.

9.2   Future directions

There are several directions in which the developments made in this thesis can be investigated further that may be the starting point of interesting follow-up analyses:

- The first straightforward direction would entail the analysis of the galaxy-shear integrated 3PCFs in real data. However a few steps remain to achieve this: (i) nonlinear redshift-space-distortions and gravitational lensing magnification effects are important systematic uncertainties which are known to impact galaxy clustering 2PCFs and hence also need to be modelled for the integrated 3PCF; (ii) as done for $\zeta_\pm$ in chapter 5, leveraging machine learning emulation techniques to obtain rapid model predictions, one can then perform a galaxy-shear integrated 3PCF analysis in galaxy and shear datasets e.g. DESY3 data. This would be one of the first efforts to extend the popular galaxy-shear 3×2PCFs to higher-order analyses in real data.

- As explained in chapters 4, 5, two major challenges currently confront cosmic shear 2PCF $\xi_\pm$ analyses — the modelling of two poorly understood astrophysical effects: *intrinsic alignment* (IA) of galaxies and *baryonic feedback* which redistributes matter on small scales. These *systematic* effects significantly contaminate the true cosmological lensing signal in $\xi_\pm$ and hence need to be carefully accounted for to obtain unbiased cosmological results. However, the constraining power of current $\xi_\pm$ data cannot distinguish between the various IA and feedback models proposed in literature, highlighting the need for more discerning methods (DES and KiDS collaborations, 2023). In this context, the power of higher-order statistics can prove to be extremely beneficial, for example, by incorporating different IA and baryonic recipes into the framework of $\zeta_\pm$ or the PDF $p(\kappa)$ and analysing them in data. The additional information from $\zeta_\pm$ or $p(\kappa)$ relative to $\xi_\pm$ will be useful to not only improve cosmological constraints but also to extract more detailed information about IA and small-scale baryonic feedback physics. This will be crucial in guiding the modelling choices of these two systematic effects in upcoming lensing surveys.

- One can extend the integrated 3PCF framework to perform higher-order cross-probe analysis beyond galaxies and cosmic shear to obtain new insights into our Universe. For example, to specifically constrain baryonic feedback, one can develop the integrated cross-probe 3PCF between LSS weak lensing and thermal Sunyaev–Zeldovich (tSZ) effects from Cosmic Microwave Background (CMB) experiments (e.g. this has been analysed with 2PCFs $\xi_\pm$ in Pandey & DES Collaboration (2022)). Such multi-probe methods are currently being actively pursued at the 2PCF level to combine datasets. The integrated 3PCF and also the PDF frameworks provide straightfor-

ward theoretical and practical avenues to study the combination of different probes and extract higher-order information beyond the reach of 2PCFs.

These are just a few examples of the directions one can pursue, but there are many further possibilities one can imagine. The overarching goal of this thesis has been to advance the development of methods to *practically* analyse and interpret the higher-order information content of weak lensing LSS surveys. We hope that these methods will pave the way for a systematic exploration of the non-Gaussian information from the rich galaxy imaging sky-survey data in the years to come.

Bibliography

Abbott T. M. C., et al., 2018, , 98, 043526

Barreira A., Schmidt F., 2017, Journal of Cosmology and Astroparticle Physics, 2017, 053–053

Barreira A., Lazeyras T., Schmidt F., 2021, Journal of Cosmology and Astroparticle Physics, 2021, 029

Bartelmann M., Schneider P., 2001, , 340, 291

Barthelemy A., Codis S., Bernardeau F., 2021, Monthly Notices of the Royal Astronomical Society, 503, 5204

Barthelemy A., Halder A., Gong Z., Uhlemann C., 2023, arXiv e-prints, p. arXiv:2307.09468

Baugh C. M., 2013, Publications of the Astronomical Society of Australia, 30, e030

Berlind A. A., Weinberg D. H., 2002, The Astrophysical Journal, 575, 587–616

Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Physics Reports, 367, 1–248

Blas D., Lesgourgues J., Tram T., 2011, , 2011, 034

Boyle A., Uhlemann C., Friedrich O., Barthelemy A., Codis S., Bernardeau F., Giocoli C., Baldi M., 2021, Monthly Notices of the Royal Astronomical Society, 505, 2886

Buchalter A., Kamionkowski M., Jaffe A. H., 2000, The Astrophysical Journal, 530, 36–52

Burger, Pierre A. et al., 2023, A&A, 669, A69

Burger P., Schneider P., Demchenko V., Harnois-Deraps J., Heymans C., Hildebrandt H., Unruh S., 2020, Astronomy & Astrophysics, 642, A161

Cabré A., Gaztañaga E., 2009, Monthly Notices of the Royal Astronomical Society, 396, 1119

Chiang C.-T., Wagner C., Schmidt F., Komatsu E., 2014, Journal of Cosmology and Astroparticle Physics, 2014, 048–048

Chiang C.-T., Wagner C., Sánchez A. G., Schmidt F., Komatsu E., 2015, JCAP, 09, 028

Coles P., Lucchin F., 2002, Cosmology: The Origin and Evolution of Cosmic Structure. Wiley, `https://books.google.de/books?id=uUFVb-DHtCwC`

DES Collaboration 2022, Physical Review D, 105

DES and KiDS collaborations 2023, The Open Journal of Astrophysics, 6

Dark Energy Survey Collaboration et al., 2016, , 460, 1270

Davies C. T., Cautun M., Giblin B., Li B., Harnois-Déraps J., Cai Y.-C., 2022, Monthly Notices of the Royal Astronomical Society, 513, 4729

Desjacques V., Jeong D., Schmidt F., 2018, Physics Reports, 733, 1–193

Dodelson S., Schmidt F., 2020, Modern Cosmology. Elsevier Science, `https://books.google.de/books?id=GGjfywEACAAJ`

Eisenstein D. J., Hu W., 1998, , 496, 605

Friedrich O., 2018, Statistical properties of the cosmic density field beyond 2-point statistics, `http://nbn-resolving.de/urn:nbn:de:bvb:19-234011`

Friedrich O., et al., 2018, Physical Review D, 98

Friedrich O., Uhlemann C., Villaescusa-Navarro F., Baldauf T., Manera M., Nishimichi T., 2020, Monthly Notices of the Royal Astronomical Society, 498, 464–483

Friedrich O., Halder A., Boyle A., Uhlemann C., Britt D., Codis S., Gruen D., Hahn C., 2022, Monthly Notices of the Royal Astronomical Society, 510, 5069

Fu L., et al., 2014, Monthly Notices of the Royal Astronomical Society, 441, 2725

Gatti M., DES Collaboration 2022, PRD, 106, 083509

Giblin B., Cai Y.-C., Harnois-Déraps J., 2023, , 520, 1721

Gong Z., Halder A., Barreira A., Seitz S., Friedrich O., 2023, , JCAP 2023, 040

Gruen D., DES Collaboration 2018, , 98, 023507

Halder A., Barreira A., 2022, Mon. Not. Roy. Astron. Soc., 515, 4639

Halder A., Friedrich O., Seitz S., Varga T. N., 2021, Monthly Notices of the Royal Astronomical Society, 506, 2780

Halder A., Gong Z., Barreira A., Friedrich O., Seitz S., Gruen D., 2023, , 2023, 028

Harnois-Déraps J., Martinet N., Castro T., Dolag K., Giblin B., Heymans C., Hildebrandt H., Xia Q., 2021, Monthly Notices of the Royal Astronomical Society, 506, 1623

Heydenreich S., Linke L., Burger P., Schneider P., 2022, A roadmap to cosmological parameter analysis with third-order shear statistics I: Modelling and validation, doi:10.48550/ARXIV.2208.11686, https://arxiv.org/abs/2208.11686

Hobson M. P., Efstathiou G. P., Lasenby A. N., 2006, General Relativity: An Introduction for Physicists. Cambridge University Press, doi:10.1017/CBO9780511790904

Hogg D. W., 1999, arXiv e-prints, pp astro–ph/9905116

Hubble E., 1929, Proceedings of the National Academy of Sciences, 15, 168

Isserlis L., 1918, Biometrika, 12, 134

Jain B., Seljak U., 1997, The Astrophysical Journal, 484, 560

Jarvis M., Bernstein G., Jain B., 2004, Monthly Notices of the Royal Astronomical Society, 352, 338

Kaiser N., 1992, , 388, 272

Kilbinger M., 2015, Reports on Progress in Physics, 78, 086901

Krause E., Fang X., Pandey S., DES collaboration 2021, Dark Energy Survey Year 3 Results: Multi-Probe Modeling Strategy and Validation (arXiv:2105.13548)

Lanzieri D., Lanusse F., Modi C., Horowitz B., Harnois-Déraps J., Starck J.-L., Collaboration T. L. D. E. S., 2023, Forecasting the power of Higher Order Weak Lensing Statistics with automatically differentiable simulations (arXiv:2305.07531)

Limber D. N., 1954, , 119, 655

Linke L., Simon P., Schneider P., Farrow D. J., Rödiger J., Wright A. H., 2022, Astronomy & Astrophysics, 665, A38

MacKenzie R., 2000, Path Integral Methods and Applications (arXiv:quant-ph/0004090)

Martinez V., Saar E., 2001, Statistics of the Galaxy Distribution. Taylor & Francis, https://books.google.de/books?id=GzkDngEACAAJ

Massey R., Kitching T., Richard J., 2010, Reports on Progress in Physics, 73, 086901

Mead A. J., Peacock J. A., Heymans C., Joudaki S., Heavens A. F., 2015, Monthly Notices of the Royal Astronomical Society, 454, 1958–1975

Mo H., van den Bosch F., White S., 2010, Galaxy Formation and Evolution. Cambridge University Press, doi:10.1017/CBO9780511807244

Muir J., et al., 2020, Monthly Notices of the Royal Astronomical Society, 494, 4454

Mukhanov V., 2005, Physical Foundations of Cosmology. Cambridge University Press, doi:10.1017/CBO9780511790553

Ntelis P., et al., 2017, , 2017, 019

Pandey S., DES Collaboration 2022, Physical Review D, 106

Peacock J. A., 1998, Cosmological Physics. Cambridge University Press, doi:10.1017/CBO9780511804533

Planck Collaboration et al., 2018, arXiv e-prints, p. arXiv:1807.06209

Riess A. G., et al., 1998, , 116, 1009

Schneider P., 2006, Gravitational Lensing: Strong, Weak and Micro, p. 269–451

Schneider, P. Watts, P. 2005, A&A, 432, 783

Schneider P., van Waerbeke L., Mellier Y., 2002, Astronomy & Astrophysics, 389, 729–741

Schneider P., Kilbinger M., Lombardi M., 2005, , 431, 9

Secco L. F., DES Collaboration 2022, PRD, 105, 103537

Seitz S., Schneider P., 1996, , 305, 383

Semboloni E., Hoekstra H., Schaye J., 2013, Monthly Notices of the Royal Astronomical Society, 434, 148–162

Spurio Mancini A., Piras D., Alsing J., Joachimi B., Hobson M. P., 2022, Monthly Notices of the Royal Astronomical Society, 511, 1771

Takada M., Jain B., 2004, Monthly Notices of the Royal Astronomical Society, 348, 897–915

Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, , 761, 152

Takahashi R., Hamana T., Shirasaki M., Namikawa T., Nishimichi T., Osato K., Shiroyama K., 2017, , 850, 24

Takahashi R., Nishimichi T., Namikawa T., Taruya A., Kayo I., Osato K., Kobayashi Y., Shirasaki M., 2020, The Astrophysical Journal, 895, 113

Troxel M., Ishak M., 2015, Physics Reports, 558, 1

Troxel M., et al., 2018, Physical Review D, 98

Uhlemann C., Friedrich O., Villaescusa-Navarro F., Banerjee A., Codis S., 2020, Monthly Notices of the Royal Astronomical Society, 495, 4006–4027

Voivodic R., Barreira A., 2021, Journal of Cosmology and Astroparticle Physics, 2021, 069

Wong K. C., et al., 2019, arXiv e-prints, p. arXiv:1907.04869

Xavier H. S., Abdalla F. B., Joachimi B., 2016, Mon. Not. Roy. Astron. Soc., 459, 3693

Zürcher D., DES Collaboration 2022, Monthly Notices of the Royal Astronomical Society, 511, 2075

Acknowledgements

Firstly, I express my gratitude to my advisor Stella Seitz for providing me with the opportunity to work in the Gravitational Lensing group at USM. I would also like to thank Ralf Bender for letting me be a part of the OPINAS group at MPE. Besides my PhD, my research trips abroad for different workshops were also supported by them which allowed me to interact with many excellent scientists and embrace new ideas. I am deeply grateful for the advice and unwavering support I have received from them both and also from a multitude of individuals during my doctoral study:

To Oliver Friedrich: I am deeply indebted to you. Over the past 5 years, you have always supported me and have patiently addressed and discussed all my questions. You have had an enormous influence on my approach to research and have instilled in me the courage to ponder questions beyond my immediate expertise and the determination to acquire new skills in pursuit of their answers. I am grateful to you for this invaluable lesson. I'll always remember the time when, amidst caring for Laura, you reviewed my first paper and provided comments just a day before the arrival of your little Henry. Beyond your brilliance as a scientist, you stand out as an even better human being. Thank you, Oliver, Laura, Henry, and Carlotta, for bringing immense joy to my life.

To Alex Barreira: You are an inspiration to me. When I think of your influence, the word that resonates is 'rigorous'. Thank you for teaching me how to do science methodically. I am grateful for the many hours we spent discussing science (and football) over Zoom, as well as the detailed correspondences over emails. The knowledge I have gained from you will always remain with me. Thank you so much, Alex.

To Daniel Gruen: I extend my gratitude for your invaluable suggestions on my projects and for the several pieces of advice you shared with me when I simply dropped by your office. You allowed me to be closely associated with the ACAI group and participate in the group activities for which I am extremely grateful.

To Alexandre Barthelemy: Thank you for spending several hours in the afternoons at USM sharing your theoretical expertise with me, patiently answering my questions and giving me valuable advice on numerous topics. Your wit and humour is endless and I must confess that it has really rubbed off on me. I probably finish my PhD a bit wittier than I was when I started it. And I have only you to thank for that!

To Tamas Varga: My sincere thanks for the several walks we took in the Englischer Garten and the lunches over which we discussed various topics. You gave me a different outlook to life in general and I am very grateful to you for your company.

To Laurence Gong: You are more than a friend; you are a brother to me. The significant portion of my PhD journey (research and non-research) that we spent working closely together holds a special place in my heart. I look back fondly on the hours I dedicated to training you, and (it is no secret) I am immensely proud of you. Now that you have become my valued collaborator, I am hopeful that in the future, we shall continue our scientific adventures together.

I would like to thank Eiichiro Komatsu, Ariel Sanchez and Jochen Weller for teaching the courses that I took during my Masters and PhD which aroused and deepened my interest in Cosmology. Eiichiro, Ariel, Jochen and many others including Marco Gatti, Nico Hamaus, Benjamin Joachimi, Elisabeth Krause, Marilena LoVerde, Susan Pyne, Fabian Schmidt, Alessio Spurio-Mancini, Ryuichi Takahashi and Cora Uhlemann took the time out to answer many of my questions which helped me towards my research. Thank you all. The many insightful discussions with members (especially of the higher-order statistics teams) of DES and LSST DESC collaborations also helped me significantly in my projects.

I am grateful to the many colleagues I have interacted with over the last few years. I would like to thank (in no particular order) Arno, Juan, Pierre, Matthias, Giacomo, Raphael, Luis, Dylan (my RMT partner), Pascal, Bernardita, Elisa, Patrick, Kevin, Bianca, Jed, Parth, Luca, Michael, Agne, Sophie, Yun-Hsin and the members of the Extragalactic Astronomy group at USM and OPINAS group at MPE for their company and friendship. To my friends from MPA: Beatriz, Ivana(s) (Sr. and Jr.) and Safak, thank you all for the wonderful times shared during workshops and dinners. In case I have forgotten any names, my apologies. To all the students I have tutored over the course of labs, tutorials, and projects and had the pleasure to discuss science with during our tea gatherings: thank you for being enthusiastic and making the discussions lively. To David, Kai, Rintaro, Bhashin, Leon and Yue: you all will become very successful one day! From my experience, I can say that science is fun when you do it together with the people you connect with. Therefore, make time for your collaborators and always treat them with respect and kindness.

No word of gratitude will ever be enough towards my family (Ma, Baba, Dada) — thank you for being there for me and supporting me always. I hope to make you all proud someday. I am very grateful to my extended family at home, my host family in Bremen, and all my dear friends for their continuous support. To Maria, what can I say besides thank you for all that you have done for me. I am a better, and wiser person because of you. I will always be grateful for the part of your life you shared with me. Thank you. All that I am today is due to the love and support of you all.

I will reserve my final words in this dissertation for my advisor, Stella. Throughout my LMU journey, you have supported me in all my endeavours (be it research or teaching), fostering my growth as an independent scientist. Today, I can proudly say that the researcher I have become is a direct result of your guidance. Despite facing many health challenges, each of which you have overcome with resilience, you have consistently demonstrated to me and all your students your passion and commitment to the science we pursue. If, one day, I can become a fraction of the scientist you are, I will consider myself truly successful. Words cannot adequately express my deepest gratitude to you. This work, I dedicate it to you, Stella.