Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **516,** 1829–1845 (2022) https://doi.org/10.1093/mnras/stac2351
Advance Access publication 2022 August 22

Three-point intrinsic alignments of dark matter haloes in the IllustrisTNG simulation

Susan Pyne,* Ananth Tenneti and Benjamin Joachimi
Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

Accepted 2022 August 17. Received 2022 July 16; in original form 2022 April 12

ABSTRACT

We use the IllustrisTNG suite of cosmological simulations to measure intrinsic alignment (IA) bispectra of dark matter subhaloes between redshifts 0 and 1. We decompose the intrinsic shear field into E- and B- modes and find that the bispectra $B_{\delta\delta\mathrm{E}}$ and $B_{\delta\mathrm{EE}}$, between the matter overdensity field, δ, and the E-mode field, are detected with high significance. We also model the IA bispectra analytically using a method consistent with the two-point non-linear alignment model. We use this model and the simulation measurements to infer the IA amplitude A_{IA} and find that values of A_{IA} obtained from IA power spectra and bispectra agree well at scales up to $k_{\mathrm{max}} = 2\,h\mathrm{Mpc}^{-1}$. For example at $z = 1$, $A_{\mathrm{IA}} = 2.13 \pm 0.02$ from the cross power spectrum between the matter overdensity and E-mode fields and $A_{\mathrm{IA}} = 2.11 \pm 0.03$ from $B_{\delta\delta\mathrm{E}}$. This demonstrates that a single physically motivated model can jointly model two-point and three-point statistics of IAs, thus enabling a cleaner separation between IAs and cosmological weak lensing signals.

Key words: gravitational lensing: weak – methods: numerical – large-scale structure of Universe – cosmology: theory.

1 INTRODUCTION

Gravitational tidal effects cause dark matter haloes and the galaxies within them to preferentially align with the cosmological large-scale structure. Thus, the shapes and orientations of galaxies and haloes may be correlated across cosmological distances. This phenomenon is known as intrinsic alignment (IA) (Kirk et al. 2012; Joachimi et al. 2015; Kiessling et al. 2015; Kirk et al. 2015; Troxel & Ishak 2015).

IAs have been widely studied, mainly because the IA effects mimic cosmic shear and thus are an undesirable contaminant of weak lensing shape measurements. Controlling this systematic uncertainty has been a major concern for recent weak lensing surveys, such as the Kilo-Degree Survey (KiDS-1000)[1] (Joachimi et al. 2021) and the Dark Energy Survey[2] (Secco et al. 2022a). For forthcoming surveys such as *Euclid*[3] (Laureijs et al. 2011) and the Vera C. Rubin Observatory Legacy Survey of Space and Time,[4] control of systematics will become even more pressing. These surveys will measure the shapes of billions of galaxies, reducing statistical uncertainties, so that uncontrolled systematics will be a limiting issue. Looking further ahead, as understanding of IAs develops it is also possible to see this signal as a valuable cosmological probe in its own right that future surveys will be able to exploit (Chisari & Dvorkin 2013; Taruya & Okumura 2020). All these considerations suggest that it is worthwhile to explore IAs as broadly as possible.

So far, most theoretical and observational studies of IAs have used two-point statistics; relatively few have taken the further step of considering three-point measurements. However, there is evidence that IAs affect two-point and three-point weak lensing statistics differently (Huterer et al. 2006; Semboloni et al. 2008; Troxel & Ishak 2011, 2012; Pyne & Joachimi 2021). For this reason, most work that has investigated three-point statistics aimed to use them to self-calibrate IA contamination in weak lensing data.

Three-point IA statistics have already been successfully measured from both survey data and simulations. On the observational side, Semboloni et al. (2010) measured three-point aperture mass statistics from the Cosmic Evolution Survey (Scoville et al. 2007) and Fu et al. (2014) built on this work using the larger Canada–France–Hawaii Telescope Lensing Survey (Heymans et al. 2012). Both these studies showed that using three-point statistics could help improve constraints on cosmological parameters. Early simulation results were described by Semboloni et al. (2008) who measured three-point IA aperture mass statistics from simulations described in Heymans et al. (2006). The main focus of this work was a comparison between the amplitudes of IA and weak lensing statistics.

Another strand of work has involved analytical modelling of three-point IA statistics, building on methods developed for two-point statistics, in particular the non-linear alignment (NLA) model (Hirata & Seljak 2004; Bridle & King 2007). This postulates that the IA of haloes is related to the tidal gravitational field at an earlier redshift. This methodology can been extended to three-point statistics in a natural but empirical way (Troxel & Ishak 2012; Merkel & Schäfer 2014; Pyne & Joachimi 2021). Other IA models, for example the halo-model of Schneider & Bridle (2010) and tidal alignment models such as Blazek, Vlah & Seljak (2015), have generally only been implemented for two-point statistics. An exception is the development by Vlah, Chisari & Schmidt (2020, 2021) of effective field theory models of galaxy alignments. These authors modelled two-point statistics at next-to-leading order and three-point statistics at leading order.

Despite the body of work described above, an analytical model for three-point IA statistics has never been tested against simulations. In this work we aim to fill this gap. As a starting point, we consider dark matter subhaloes rather than galaxies. Since the IA of galaxies

* E-mail: ucapsep@ucl.ac.uk
[1] http://kids.strw.leidenuniv.nl/index.php
[2] https://www.darkenergysurvey.org
[3] http://sci.esa.int/euclid/
[4] https://www.lsst.org

is related to that of their host haloes, our work can be seen as a first step towards a model for galaxies.

Our approach to measurement from simulations is based on that of Kurita et al. (2021) who developed a novel method to measure three-dimensional IA power spectra of dark matter haloes which they applied to the high-resolution DarkQuest suite of simulations (Nishimichi et al. 2019; Miyatake et al. 2021). This has a resolution of 2048^3 particles in a periodic cubic box size of 1 h^{-1} Gpc. Using the same methodology, we measure all theoretically non-zero IA bispectra for equilateral and specific isosceles triangles from the IllustrisTNG[5] simulation suite.

Kurita et al. (2021) compared their simulation results with the two-point NLA model and concluded that at linear scales the model matches the simulation results, with a scale-independent IA amplitude, but that at non-linear scales the model breaks down and the amplitude is no longer constant. We similarly compare our bispectrum measurements with an analytical model based on perturbation theory which is in keeping with the two-point NLA model.

In Section 2, we describe the simulation suite and the methods used to measure dark matter subhalo shapes. Section 3 explains how we measure IA power spectra and bispectra from simulations, and Section 4 discusses the NLA model and our analytical model for three-point IA statistics. In Section 5, we present our IA power spectrum and bispectrum measurements from simulations and compare these with analytical results, and in Section 6 we discuss how these compare with previous work. We summarize and discuss possible further work in Section 7. Appendices discuss other details, including signal-to-noise ratios and a possible phenomenological modification to the NLA model.

2 SIMULATIONS

2.1 Characteristics of the simulations

We use the publicly available cosmological simulation suite IllustrisTNG (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018, 2019; Pillepich et al. 2018; Springel et al. 2018). Each simulation in the suite uses the moving-mesh code AREPO (Springel 2010)[6] and self-consistently solves for the coupled evolution of dark matter, cosmic gas, luminous stars, and supermassive black holes from the starting redshift $z = 127$ to the present day, based on a cosmology consistent with *Planck* (Ade et al. 2016): $\Omega_\Lambda = 0.6911$, $\Omega_\mathrm{m} = 0.3089$, $\Omega_\mathrm{b} = 0.0486$, $\sigma_8 = 0.8159$, $n_\mathrm{s} = 0.9667$, $h = 0.6774$. Specifically, we use the IllustrisTNG300-1 hydrodynamic simulation which has the largest simulation box size within the suite of simulations available from the public data release (Nelson et al. 2019). The simulation box volume is 300 Mpc^3 (comoving) with 2500^3 dark matter particles and an equal initial number of gas cells.

The publicly available data include catalogues of haloes identified using the Friends-of-Friends algorithm (Davis et al. 1985). The SUBFIND algorithm (Springel et al. 2001) is used to identify substructures within the haloes. This algorithm defines the centre of the subhalo as the location of the most bound particle, and provides the positions of all dark matter particles in the halo.

For our analysis, we select dark matter subhaloes from the simulations at three redshifts, $z = 0.0$, 0.5, 1.0. The subhalo mass is the total sum of the masses of all individual particles belonging to a given subhalo as identified by the SUBFIND algorithm. The maximum subhalo masses in the simulation are approximately 3.4×10^{14}, 7.4×10^{14}, and $1.2 \times 10^{15}\,\mathrm{M}_\odot$ at $z = 1.0$, 0.5, and 0.0, respectively. We select subhaloes within the mass range 4×10^{10} to $10^{14}\,h^{-1}\mathrm{M}_\odot$ for our analysis.

2.2 Measurement of 3D halo shapes from simulations

We start by measuring 3D ellipsoidal subhalo shapes. For this, we adopt the widely used method based on the inertia tensor I_{ij}. This is defined as

$$I_{ij} = \frac{\sum_n m_n x_{ni} x_{nj}}{\sum_n m_n}\,, \tag{1}$$

where m_n is the mass of the nth particle and x_{ni}, x_{nj} are its position coordinates relative to the centre of the halo. The semi-axes a, b, c of the ellipsoid are obtained from the eigenvalues λ_a, λ_b, λ_c of the inertia tensor, with $a = \sqrt{\lambda_a}$ and so on. We set $c \leq b \leq a$ and define axis ratios $s = c/a$ and $q = b/a$. The eigenvectors of the inertia tensor determine the orientation of the axes.

In order to obtain well-resolved shapes, we choose only subhaloes with a minimum of 1000 dark matter particles, which is consistent with the particle number threshold adopted in earlier studies, for example Tenneti et al. (2015). With this threshold the minimum subhalo mass is about $4 \times 10^{10}\,h^{-1}\,\mathrm{M}_\odot$.

As an improvement on equation (1), we use the reduced inertia tensor (Tenneti et al. 2015) which gives more weight to particles that are closer to the centre of a subhalo, thus avoiding potential problems with defining its outer edge. The reduced inertia tensor is defined as

$$\tilde{I}_{ij} = \frac{\sum_n m_n \frac{x_{ni} x_{nj}}{r_n^2}}{\sum_n m_n}\,, \tag{2}$$

where

$$r_n^2 = \sum_{i=1}^{3} x_{ni}^2\,. \tag{3}$$

Rather than taking the 'raw' axis ratios defined above we use the iterative approach described in Tenneti et al. (2015) to recover the shape of an isodensity surface (Schneider, Frenk & Cole 2012). In this method, the principal axes of the ellipsoids are rescaled iteratively while the enclosed volume is kept constant. After each rescaling, particles outside the ellipsoidal volume are discarded. The process is repeated until the fractional change in the axis ratios is below a predefined limit, in our case 1 per cent.

3 INTRINSIC ALIGNMENT SPECTRA FROM SIMULATIONS

3.1 Ellipticity and tidal shear

The ellipticity, ϵ, of a subhalo shape is a spin-2 quantity (it is invariant under rotations of integer multiples of π). It can be parametrized in several ways in terms of the shape and orientation of the subhalo. All parametrizations are essentially equivalent so we can make a choice that suits the problem at hand. In many cases, it is convenient to express the ellipticity as $\epsilon = \epsilon_+ + \mathrm{i}\epsilon_\times$ where ϵ_+ represents stretching along a defined axis and $\epsilon_\times$ represent stretching along an axis at 45° to this. It is always possible to find such a decomposition (Stebbins 1996).

We can then express the two components of the ellipticity as

[5] https://www.tng-project.org
[6] https://arepo-code.org

$$\epsilon_+ = \left(\frac{a-b}{a+b}\right)\cos 2\theta, \tag{4}$$

$$\epsilon_\times = \left(\frac{a-b}{a+b}\right)\sin 2\theta\ , \tag{5}$$

where a and b are the semimajor and semiminor axes of the ellipse. Here we take the line of sight to be the z-axis and consider ellipses projected on to the plane perpendicular to this (Kurita et al. 2021; Shi et al. 2021). The angle θ is an arbitrary choice that we take to be the position angle of the major axis with respect to the x-axis in the x–y plane, as determined by the relevant eigenvector. Equations (4) and (5) clearly satisfy the requirements for a spin-2 quantity. Moreover, under a parity transformation ϵ_+ is unchanged but $\epsilon_\times \to -\epsilon_\times$ (Schneider, van Waerbeke & Mellier 2002).

This definition of ϵ, in terms of $(a-b)/(a+b)$, is commonly used for weak lensing shear because it provides an unbiased estimator of the shear and does not depend on the ellipticity distribution of source galaxies (Seitz & Schneider 1997; Viola, Kitching & Joachimi 2014). Even though we are considering intrinsic ellipticity rather than cosmic shear we adopt this definition for consistency. It has the added advantage that the tidal shear $\gamma = \gamma_+ + \mathrm{i}\gamma_\times$ can be assumed to be directly proportional to the ellipticity.

Other studies of IAs, for example Blazek et al. (2015) and Shi et al. (2021), used an alternative definition of ellipticity that replaces $(a-b)/(a+b)$ with $(a^2-b^2)/(a^2+b^2)$ in equations (4) and (5). This version is more readily comparable with observations but does not provide an unbiased shear estimator (Schneider & Seitz 1995). Moreover, rather than the shear being directly related to the ellipticity, an extra responsivity factor $\mathcal{R} = 1 - \epsilon_{\rm rms}^2$ is required (Bernstein & Jarvis 2002), where $\epsilon_{\rm rms} = \sqrt{\langle\epsilon_+^2\rangle} = \sqrt{\langle\epsilon_\times^2\rangle}$. The responsivity measures the average response of $\epsilon_{+,\times}$ to $\gamma_{+,\times}$ so that $\gamma_{+,\times} = \epsilon_{+,\times}/2\mathcal{R}$. With our definition of the ellipticity we do not need to consider the responsivity and can assume that ϵ directly traces the tidal shear field $\boldsymbol{\gamma}$. Thus γ_+ and $\gamma_\times$ are given by equations (4) and (5).

3.2 Decomposition into E- and B-modes

For studies of IAs, it is convenient to go a step further and decompose the shear field into a curl-free (E-mode) component $\gamma_{\rm E}$ and a gradient-free (B-mode) component $\gamma_{\rm B}$ (Kamionkowski et al. 1998; Crittenden et al. 2002). We define these by the equations

$$\nabla^2\gamma_{\rm E}(\boldsymbol{x}) = (\partial_x\partial_x - \partial_y\partial_y)\gamma_+(\boldsymbol{x}) + 2\partial_x\partial_y\gamma_\times(\boldsymbol{x}), \tag{6}$$

$$\nabla^2\gamma_{\rm B}(\boldsymbol{x}) = (\partial_x\partial_x - \partial_y\partial_y)\gamma_\times(\boldsymbol{x}) - 2\partial_x\partial_y\gamma_+(\boldsymbol{x})\ , \tag{7}$$

where $\boldsymbol{x}$ is the configuration space position and $\partial_x\partial_x \equiv \partial^2/\partial x^2$, and so on.

This decomposition takes a simpler form in Fourier space. We define Fourier space coordinates to be $\boldsymbol{k} = (k_x, k_y, k_z)$ and choose the k_z-axis to be along the line of sight. The derivatives in equations (6) and (7) change to multiplicative factors so we get

$$k_{xy}^2\tilde{\gamma}_{\rm E}(\boldsymbol{k}) = \left(k_x^2 - k_y^2\right)\tilde{\gamma}_+(\boldsymbol{k}) + 2k_xk_y\tilde{\gamma}_\times(\boldsymbol{k}), \tag{8}$$

$$k_{xy}^2\tilde{\gamma}_{\rm B}(\boldsymbol{k}) = \left(k_x^2 - k_y^2\right)\tilde{\gamma}_\times(\boldsymbol{k}) - 2k_xk_y\tilde{\gamma}_+(\boldsymbol{k})\ , \tag{9}$$

where $k_{xy}^2 = k_x^2 + k_y^2$ and $\tilde{\gamma}_+$ and $\tilde{\gamma}_\times$ are the Fourier transforms of equations (4) and (5). Alternatively, these expressions can be written in terms of the angle $\phi = \tan^{-1}(k_x/k_y)$ giving (Kurita et al. 2021; Shi et al. 2021)

$$\tilde{\gamma}_{\rm E}(\boldsymbol{k}) = \tilde{\gamma}_+(\boldsymbol{k})\cos 2\phi + \tilde{\gamma}_\times(\boldsymbol{k})\sin 2\phi, \tag{10}$$

$$\tilde{\gamma}_{\rm B}(\boldsymbol{k}) = \tilde{\gamma}_\times(\boldsymbol{k})\cos 2\phi - \tilde{\gamma}_+(\boldsymbol{k})\sin 2\phi\ . \tag{11}$$

Even though $\tilde{\gamma}_{\rm E}$ and $\tilde{\gamma}_{\rm B}$ are measured in the plane perpendicular to the line of sight, the wavevector $\boldsymbol{k}$ is three-dimensional: $\boldsymbol{k} = k(\sqrt{1-\mu^2}\cos\phi, \sqrt{1-\mu^2}\sin\phi, \mu)$, where $\cos^{-1}(\mu)$ is the angle between $\boldsymbol{k}$ and the k_z-axis (Blazek et al. 2015; Kurita et al. 2021).

3.3 Intrinsic alignment power spectra and bispectra

Equations (10) and (11) lead directly to the IA power spectrum between the Fourier transforms of the E-mode shear, $\tilde{\gamma}_{\rm E}$, and matter density contrast, $\tilde{\delta}$, and also the auto power spectra of $\tilde{\gamma}_{\rm E}$ and $\tilde{\gamma}_{\rm B}$:

$$\left\langle\tilde{\delta}(\boldsymbol{k})\tilde{\gamma}_{\rm E}(\boldsymbol{k}')\right\rangle = (2\pi)^3\delta_{\rm D}^3(\boldsymbol{k}+\boldsymbol{k}')P_{\delta{\rm E}}(\boldsymbol{k}), \tag{12}$$

$$\left\langle\tilde{\gamma}_{\rm E}(\boldsymbol{k})\tilde{\gamma}_{\rm E}(\boldsymbol{k}')\right\rangle = (2\pi)^3\delta_{\rm D}^3(\boldsymbol{k}+\boldsymbol{k}')P_{\rm EE}(\boldsymbol{k}), \tag{13}$$

$$\left\langle\tilde{\gamma}_{\rm B}(\boldsymbol{k})\tilde{\gamma}_{\rm B}(\boldsymbol{k}')\right\rangle = (2\pi)^3\delta_{\rm D}^3(\boldsymbol{k}+\boldsymbol{k}')P_{\rm BB}(\boldsymbol{k})\ , \tag{14}$$

where $\langle\,\rangle$ denotes the ensemble average and $\delta_{\rm D}^3$ is the three-dimensional Dirac delta function. From parity considerations, these are the only possible non-zero intrinsic shear power spectra (Stebbins 1996; Kamionkowski et al. 1998; Crittenden et al. 2002; Schneider et al. 2002). Power spectra involving a single B-mode shear will switch sign if $\boldsymbol{k} \to -\boldsymbol{k}$ which is physically impossible unless the spectra are zero. Theoretically, the B-mode auto power spectrum also vanishes to first order but in practice it may be non-zero because of Poisson shot noise due to the finite sampling of the halo positions, which is also present in the E-mode auto power spectrum. [See Blazek et al. (2019) and Kurita et al. (2021) for more detailed discussions of shape noise.]

The formalism of equations (12)–(14) can be extended to IA bispectra. Again, any bispectrum that includes an odd number of B modes can be expected to be zero by parity arguments so there are five possible non-zero bispectra:

$$\left\langle\tilde{\delta}(\boldsymbol{k}_1)\tilde{\delta}(\boldsymbol{k}_2)\tilde{\gamma}_{\rm E}(\boldsymbol{k}_3)\right\rangle = (2\pi)^3\delta_{\rm D}^3(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_{\delta\delta{\rm E}}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3), \tag{15}$$

$$\left\langle\tilde{\delta}(\boldsymbol{k}_1)\tilde{\gamma}_{\rm E}(\boldsymbol{k}_2)\tilde{\gamma}_{\rm E}(\boldsymbol{k}_3)\right\rangle = (2\pi)^3\delta_{\rm D}^3(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_{\delta{\rm EE}}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3), \tag{16}$$

$$\langle\tilde{\gamma}_{\rm E}(\boldsymbol{k}_1)\tilde{\gamma}_{\rm E}(\boldsymbol{k}_2)\tilde{\gamma}_{\rm E}(\boldsymbol{k}_3)\rangle = (2\pi)^3\delta_{\rm D}^3(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_{\rm EEE}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3), \tag{17}$$

$$\left\langle\tilde{\delta}(\boldsymbol{k}_1)\tilde{\gamma}_{\rm B}(\boldsymbol{k}_2)\tilde{\gamma}_{\rm B}(\boldsymbol{k}_3)\right\rangle = (2\pi)^3\delta_{\rm D}^3(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_{\delta{\rm BB}}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3), \tag{18}$$

$$\langle\tilde{\gamma}_{\rm E}(\boldsymbol{k}_1)\tilde{\gamma}_{\rm B}(\boldsymbol{k}_2)\tilde{\gamma}_{\rm B}(\boldsymbol{k}_3)\rangle = (2\pi)^3\delta_{\rm D}^3(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_{\rm EBB}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\ . \tag{19}$$

3.4 Methodology for measuring power spectra and bispectra from simulations

To measure the power spectra and bispectra, we follow the methodology in Kurita et al. (2021). We refer the reader to Kurita et al. (2021) and Shi et al. (2021) for more detailed descriptions of the measurement methodology. To measure power spectra, we use the publicly available package NBODYKIT[7] (Hand et al. 2018) which

[7] https://github.com/bccp/nbodykit

provides a wide range of tools to analyse cosmological simulations. For bispectra, we adapt BSKIT,[8] developed by Foreman et al. (2020). This package is based on NBODYKIT together with the Fast Fourier Transform-based bispectrum measurement algorithm presented in Tomlinson, Jeong & Kim (2019). In both cases, we incorporate equations (10) and (11) into the existing code in order to measure the IA spectra.

In measuring from simulations, we align all axes in Fourier space with the simulation box sides. We then define a regular grid of size 512^3 within the box and assign the subhalo shape measurements to the grid using the cloud-in-cell interpolation algorithm. We investigated the alternative triangle shaped cloud assignment method but found it made little difference to the results. We confine our bispectrum measurements to equilateral triangles and to a representative isosceles configuration whose sides have magnitudes in the ratio 2:2:1. To quantify uncertainty in the estimates, we divide the simulation box into 3^3 subboxes, and estimate standard errors using jackknife sampling, excluding each subbox in turn.

In all subsequent results we show IA auto power spectra with Poisson shot noise subtracted, but without any allowance for non-linear effects due to IAs of shapes. Kurita et al. (2021) explored the latter in detail but found it to be only 5–10 per cent of Poisson shot noise in their halo sample from the Dark Quest simulations. Thus we measure shot noise in the IA power spectra as $\epsilon^2_{\rm rms}/n_{\rm eff}$ where $n_{\rm eff} = n_{\rm h}/L^3_{\rm box}$ is the effective number density within the simulation box with side length $L_{\rm box}$ and $n_{\rm h}$ is the number of subhaloes in the box.

The measured shear 'fields' are weighted by the number density of haloes. Density weighting is important because haloes are biased tracers of the matter density field and ellipticity/shear measurements can only be made at the positions where haloes exist.

4 ANALYTICAL MODELLING

Two-point IA statistics are commonly modelled by the linear alignment model (Hirata & Seljak 2004). This model assumes that the intrinsic ellipticity of a halo is linearly related to the local quadrupole of the gravitational potential at the redshift at which the halo formed. Thus, in Fourier space we can write

$$\tilde{\gamma}_{(+/\times)}(\boldsymbol{k}, z) = -A_{\rm IA} f_{(+/\times)} \frac{C_1 \Omega_{\rm m} \rho_{\rm cr}}{D(z)} \tilde{\delta}(\boldsymbol{k}, z) \,, \tag{20}$$

where $\Omega_{\rm m}$ is the total matter density parameter, $\rho_{\rm cr}$ is the critical density at the present day, and $D(z)$ is the linear growth factor. The functions $f_{(+/\times)}$ are defined as

$$f_+ = \left(1-\mu^2\right)\cos 2\phi \,, \tag{21}$$

$$f_\times = \left(1-\mu^2\right)\sin 2\phi \,, \tag{22}$$

where, as before, $\cos^{-1}(\mu)$ is the angle between $\boldsymbol{k}$ and the k_z-axis and $\phi = \tan^{-1}(k_x/k_y)$. The parameter C_1 in equation (20) is a normalization factor which in principle can be determined from observations or simulations. The amplitude $A_{\rm IA}$ quantifies the magnitude of the IA effect. This quantity, commonly used in cosmological inference, is what we are particularly interested in.

Substituting from equation (20) into equations (10) and (11), we have

$$\tilde{\gamma}_{\rm E}(\boldsymbol{k}) = f_{\rm IA}\tilde{\delta}(\boldsymbol{k}) \,, \tag{23}$$

$$\tilde{\gamma}_{\rm B}(\boldsymbol{k}) = 0 \,, \tag{24}$$

[8] https://github.com/sjforeman/bskit

where we define $f_{\rm IA}$ as

$$f_{\rm IA} = -A_{\rm IA}\left(1-\mu^2\right)\frac{C_1\Omega_{\rm m}\rho_{\rm cr}}{D(z)} \,. \tag{25}$$

From equation (23), we can obtain the three-dimensional E-mode IA power spectra

$$P_{\delta{\rm E}}(k) = f_{\rm IA} P_{\rm NL}(k) \,, \tag{26}$$

$$P_{\rm EE}(k) = f^2_{\rm IA} P_{\rm NL}(k) \,, \tag{27}$$

where we have used the non-linear matter power spectrum, $P_{\rm NL}(k)$, as suggested by Bridle & King (2007). This modification, known as the NLA model, has been found empirically to improve the fit of the model at non-linear scales.

Theoretically the model predicts that B-mode power spectra are zero to first order, although as discussed previously, $P_{\rm BB}$ may be non-zero in practice due to shape noise or to higher order non-linear contributions, although these would be small (Blazek et al. 2019).

Note that equations (26) and (27) also depend on redshift but for brevity we have omitted the z-dependence, here and in all subsequent related equations.

4.1 Extension of linear alignment model to bispectra

To develop a similar analytical model for three-point statistics, we need to relate the IA bispectra to the non-linear matter bispectrum. It is most straightforward to use a fitting function for the matter bispectrum based on tree-level perturbation theory, for example those given in Scoccimarro & Couchman (2001) and Gil-Marín et al. (2012). These have the generic form

$$B_{\delta\delta\delta}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = 2\left[F_2^{\rm eff}(\boldsymbol{k}_1, \boldsymbol{k}_2) P_{\rm NL}(k_1) P_{\rm NL}(k_2) + 2 \text{ perms.}\right] , \tag{28}$$

where $F_2^{\rm eff}(\boldsymbol{k}_1, \boldsymbol{k}_2)$ is a modification of the standard perturbation theory kernel (Bernardeau et al. 2002).

This formulation is easily extended to include IA power spectra in place of the non-linear matter power spectrum, leading directly to expressions for the IA bispectra that are in the spirit of the two-point NLA model (Pyne & Joachimi 2021):

$$\begin{aligned} B_{\delta\delta{\rm E}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = 2\Big[& f^2_{\rm IA} F_2^{\rm eff}(\boldsymbol{k}_1, \boldsymbol{k}_2) P_{\rm NL}(k_1) P_{\rm NL}(k_2) \\ & + f_{\rm IA} F_2^{\rm eff}(\boldsymbol{k}_2, \boldsymbol{k}_3) P_{\rm NL}(k_2) P_{\rm NL}(k_3) \\ & + f_{\rm IA} F_2^{\rm eff}(\boldsymbol{k}_3, \boldsymbol{k}_1) P_{\rm NL}(k_3) P_{\rm NL}(k_1)\Big] \,, \end{aligned} \tag{29}$$

$$\begin{aligned} B_{\delta{\rm EE}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = 2\Big[& f^3_{\rm IA} F_2^{\rm eff}(\boldsymbol{k}_1, \boldsymbol{k}_2) P_{\rm NL}(k_1) P_{\rm NL}(k_2) \\ & + f^2_{\rm IA} F_2^{\rm eff}(\boldsymbol{k}_2, \boldsymbol{k}_3) P_{\rm NL}(k_2) P_{\rm NL}(k_3) \\ & + f^3_{\rm IA} F_2^{\rm eff}(\boldsymbol{k}_3, \boldsymbol{k}_1) P_{\rm NL}(k_3) P_{\rm NL}(k_1)\Big] \,, \end{aligned} \tag{30}$$

$$B_{\rm EEE}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = f^4_{\rm IA} B_{\delta\delta\delta}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \,. \tag{31}$$

The bispectrum $B_{\rm EEE}$ is positive but the signs of $B_{\delta\delta{\rm E}}$ and $B_{\delta{\rm EE}}$ depend on the triangle configuration.

For equilateral triangles, equations (29)–(31) reduce to

$$B_{\delta\delta{\rm E}} = \tfrac{1}{3}\left[f^2_{\rm IA} + 2 f_{\rm IA}\right] B_{\delta\delta\delta} \,, \tag{32}$$

$$B_{\delta{\rm EE}} = \tfrac{1}{3}\left[2 f^3_{\rm IA} + f^2_{\rm IA}\right] B_{\delta\delta\delta} \,, \tag{33}$$

$$B_{\rm EEE} = f^4_{\rm IA} B_{\delta\delta\delta} \,, \tag{34}$$

where we have omitted the $\boldsymbol{k}$ arguments for brevity.

In this work we use the fitting formula from Gil-Marín et al. (2012) in equations (29)–(31) with the non-linear matter power spectrum

estimated by the fitting formula in Takahashi et al. (2012). We note, however, that the formulas of both Scoccimarro & Couchman (2001) and Gil-Marín et al. (2012) are known to have deficiencies: they are fitted over limited k scales and have been found to be inaccurate for squeezed triangles (Namikawa et al. 2019). These issues were explored in detail by Takahashi et al. (2020) who developed a new formula, Bihalofit, based on the halo model. It is more accurate than the perturbation theory based fitting functions over a wider range of wavenumbers and redshifts. However, it does not lend itself to use in our model because it is the sum of 1-halo and 3-halo terms that involve halo model integrals and a large number of fitted parameters. Thus, it cannot easily be related to the NLA model. However, as we show in Appendix A, for scales up to at least $k \approx 3\,h\mathrm{Mpc}^{-1}$ there is good agreement between the matter bispectrum estimated from Gil-Marín et al. (2012) and from Takahashi et al. (2020) and between these results and our measurements from simulations.

5 RESULTS

We present our results as follows. First, in Section 5.1, we discuss the halo ellipticity distributions that underlie our IA measurements. In Section 5.2, we present our measurements of IA power spectra and bispectra from simulations, and discuss consistency with previous power spectrum results and between our two-point and three-point measurements. In the rest of Section 5, we use these simulation measurements to validate our analytical models. In Sections 5.3 and 5.4, we focus on IA power spectra, using the NLA model to estimate IA amplitudes and explore their mass dependence. These power spectrum results establish our methodology and provide context for our principal results from bispectra based on our three-point analytical IA model. Results for equilateral triangles are given in Section 5.5 and compared to power spectrum results in Section 5.6. Finally, in Section 5.7 we discuss non-equilateral IA bispectra.

5.1 Distribution of halo ellipticities by mass

Previous measurements of halo IAs from simulations, for example Jing (2002), Lee et al. (2008), Xia et al. (2017), Piras et al. (2018), and Kurita et al. (2021), have found that IA increases with increasing halo mass and to a lesser extent with redshift. Similar trends have been noted from measurements of galaxy shapes from simulations (Tenneti et al. 2014, 2015), and in survey data for galaxies (Joachimi et al. 2011, 2013; Singh, Mandelbaum & More 2015) and clusters (van Uitert & Joachimi 2017). In view of this, we split our measured subhaloes into four mass bins, with each bin spanning one order of magnitude from 10^{10} to $10^{14}\,h^{-1}\,\mathrm{M}_\odot$.

Fig. 1 shows the distribution of ellipticity, defined as $\epsilon = \sqrt{\epsilon_+^2 + \epsilon_\times^2}$, in each of the four bins at $z = 0.5$. (Similar distributions are found for other redshifts.) Dotted vertical lines show the median ellipticity in each bin. This confirms previous authors' findings: higher-mass haloes are more elliptical. It is noticeable that the two lowest-mass bins have similar ellipticities, which are lower than those of the two high-mass bins.

5.2 IA power spectra and bispectra measured from simulations

Fig. 2 shows our measured IA power spectra, $P_{\delta\mathrm{E}}$, P_{EE}, and P_{BB}, defined by equations (12)–(14), at three redshifts, with the non-linear matter power spectrum $P_{\delta\delta}$ shown for comparison. In this and all similar figures, we show the absolute value of the spectra. P_{BB} is essentially equal to the shot noise at all but the largest scales considered here. The non-vanishing cross-power spectrum

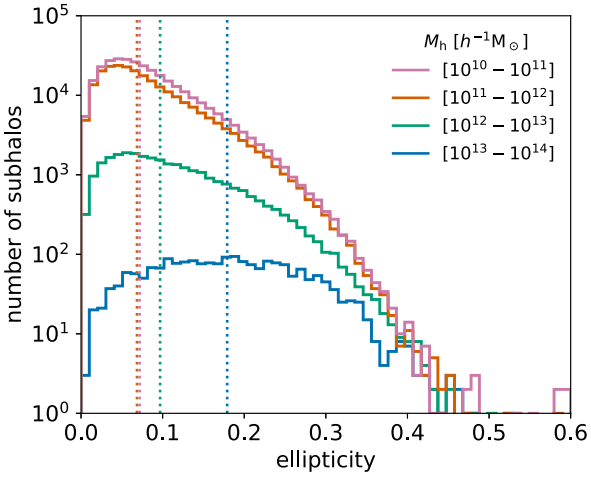

Figure 1. Distribution of dark matter subhaloes by projected ellipticity $\epsilon = \sqrt{\epsilon_+^2 + \epsilon_\times^2}$ in four mass bins at $z = 0.5$. Vertical dotted lines indicate the median ellipticity in each bin.

$P_{\delta\mathrm{E}}$ confirms that subhalo shapes are correlated with the matter overdensity field at all scales. The power spectra show little variation across redshifts. These results are consistent with IA power spectrum measurements in Kurita et al. (2021). (See their figs 2 and 4.)

Figs 3 and 4 show similar results for measured E-mode and B-mode bispectra respectively [equations (15)–(19)], for equilateral triangles and for isosceles triangles with sides in the ratio 2:2:1. The B-mode bispectra have very low signal-to-noise ratios (see Appendix B) and we do not consider them further in this work. However, the E-mode bispectra, in particular $B_{\delta\delta\mathrm{E}}$, have relatively strong signals, especially for isosceles configurations.

There is no convenient benchmark with which to compare our bispectrum results. The only previous measurements of three-point IAs from simulations were by Semboloni et al. (2008, 2010). Unfortunately, it is difficult to compare our results with theirs for a number of reasons: they worked in configuration space and measured aperture mass statistics, and they focused on the magnitudes of IAs relative to lensing signals, rather than on the strength of the IA signal itself. This is discussed further in Section 6.

In Fig. 5, we split the IA power spectra $P_{\delta\mathrm{E}}$ and P_{EE} and the bispectrum $B_{\delta\delta\mathrm{E}}$ between mass bins. In all three cases the magnitudes of the spectra increase with mass. Again we note that the spectra in the two lowest-mass bins are similar to each other and generally smaller than those in higher-mass bins.

5.3 Intrinsic alignment amplitude from power spectra

Having shown that E-mode spectra are detectable from the simulations, we now estimate the IA amplitude A_{IA} in equation (20) from simulation measurements. We first consider power spectra. From equations (26) and (27) we can obtain approximate estimates of f_{IA}, and hence A_{IA}, from the ratio $P_{\delta\mathrm{E}}(k)/P_{\delta\delta}(k)$ or from $\sqrt{P_{\mathrm{EE}}(k)/P_{\delta\delta}(k)}$. Alternatively, following Kurita et al. (2021), we can use least-squares minimization to fit A_{IA} from these equations. To do this we find the value of the parameter $\hat{A}_{\mathrm{IA}}$ which minimizes χ^2 given by

$$\chi^2 = \sum_k \frac{[R(k) - F(\hat{A}_{\mathrm{IA}})]^2}{\sigma_R^2(k)} \,. \tag{35}$$

Here, in the first case, based on $P_{\delta\mathrm{E}}$, $R(k) = P_{\delta\mathrm{E}}(k)/P_{\delta\delta}(k)$ and

$$F(\hat{A}_{\mathrm{IA}}) = \int_0^1 \left(1 - \mu^2\right) \mathrm{d}\mu\, c(z)\hat{A}_{\mathrm{IA}} = \tfrac{2}{3} c(z)\hat{A}_{\mathrm{IA}} \,, \tag{36}$$

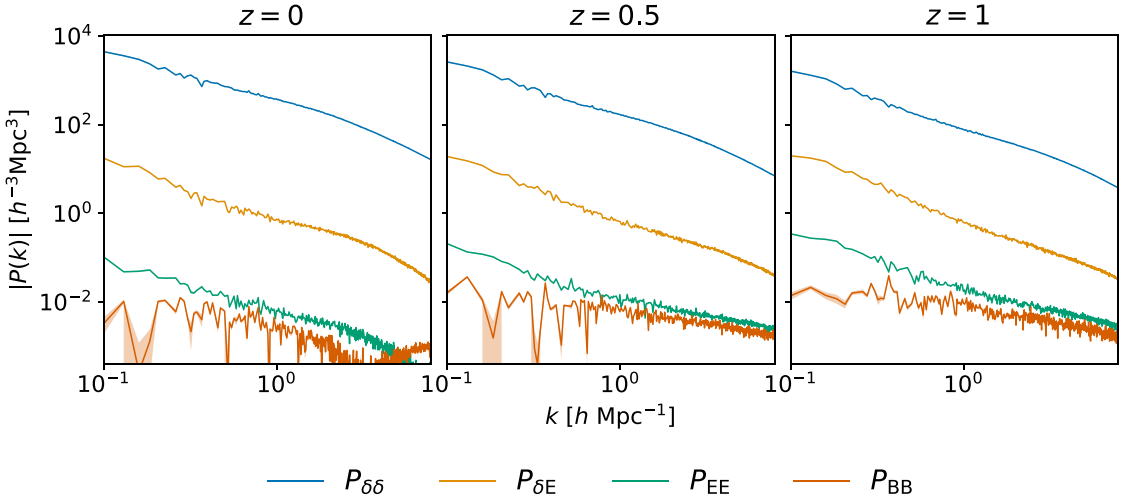

Figure 2. IA power spectra $P_{\delta E}$, P_{EE}, and P_{BB} measured from simulations. The matter power spectrum $P_{\delta\delta}$ is also shown for reference. Shot noise has been subtracted from the auto power spectra. Shaded areas are 68 per cent confidence intervals.

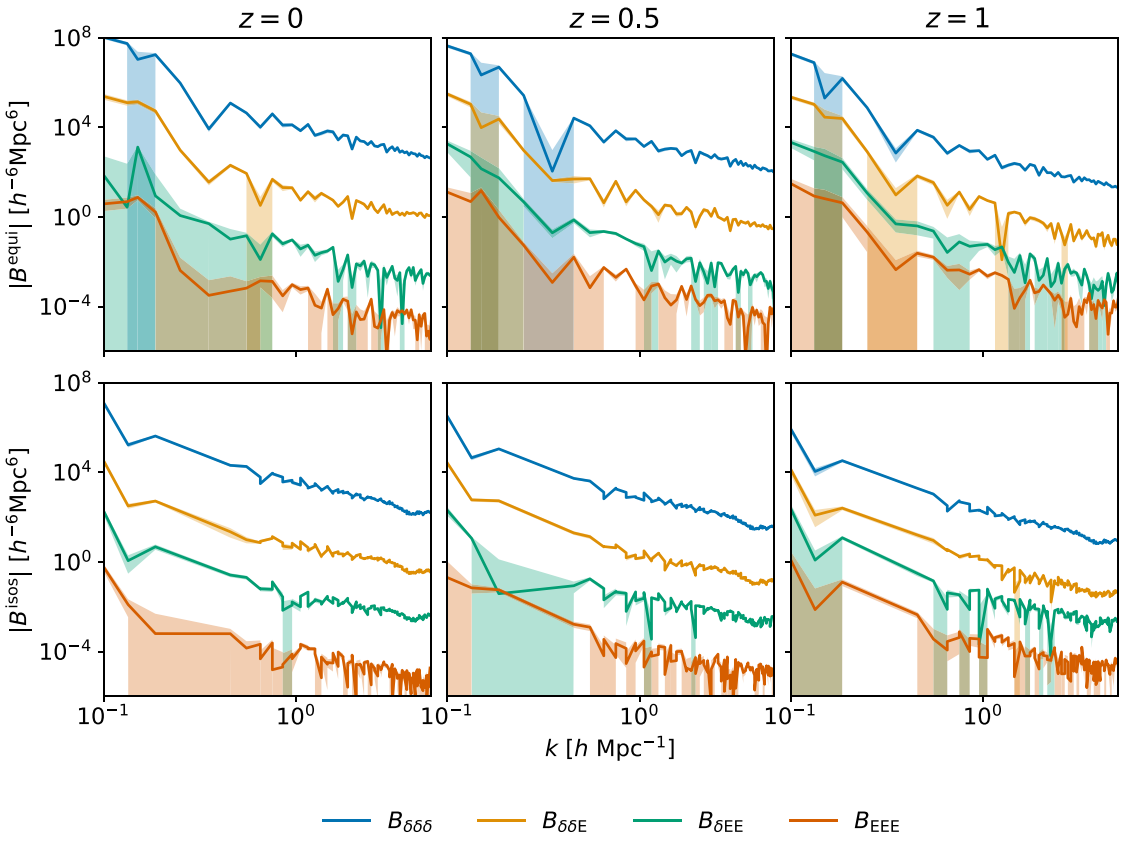

Figure 3. IA bispectra $B_{\delta\delta E}$, $B_{\delta EE}$, and B_{EEE} measured from simulations. *Top*: Equilateral triangles. *Bottom*: Isosceles triangles with sides in the ratio 2:2:1. The bispectrum $B_{\delta\delta\delta}$ is also shown for reference. Shaded areas are 68 per cent confidence intervals.

where $c(z) = C_1\Omega_m\rho_{cr}/D(z)$ [see equation (25)] and $\sigma_R^2(k)$ is the variance of $R(k)$ calculated using jackknife sampling from 27 simulation subboxes. To calculate $F(\hat{A}_{IA})$, we obtain the growth factor from the cosmological parameter estimation code CosmoSIS[9] (Zuntz et al. 2015). We use the value of C_1 derived by Bridle & King (2007) which is $5 \times 10^{-14}\, h^{-2}\mathrm{M}_\odot^{-1}\mathrm{Mpc}^3$, leading to $C_1\rho_{cr} = 0.0134$ (Joachimi et al. 2011). All cosmological parameters values are identical to those used in the simulations.

In the second case, based on P_{EE}, $R(k) = \sqrt{P_{EE}(k)/P_{\delta\delta}(k)}$ and

$$F(\hat{A}_{IA}) = \int_0^1 \left(1-\mu^2\right)^2 \mathrm{d}\mu\, c(z)\hat{A}_{IA} = \tfrac{8}{15} c(z)\hat{A}_{IA}\,, \tag{37}$$

with $c(z)$ and $\sigma_R^2(k)$ defined as before.

[9] https://github.com/joezuntz/cosmosis

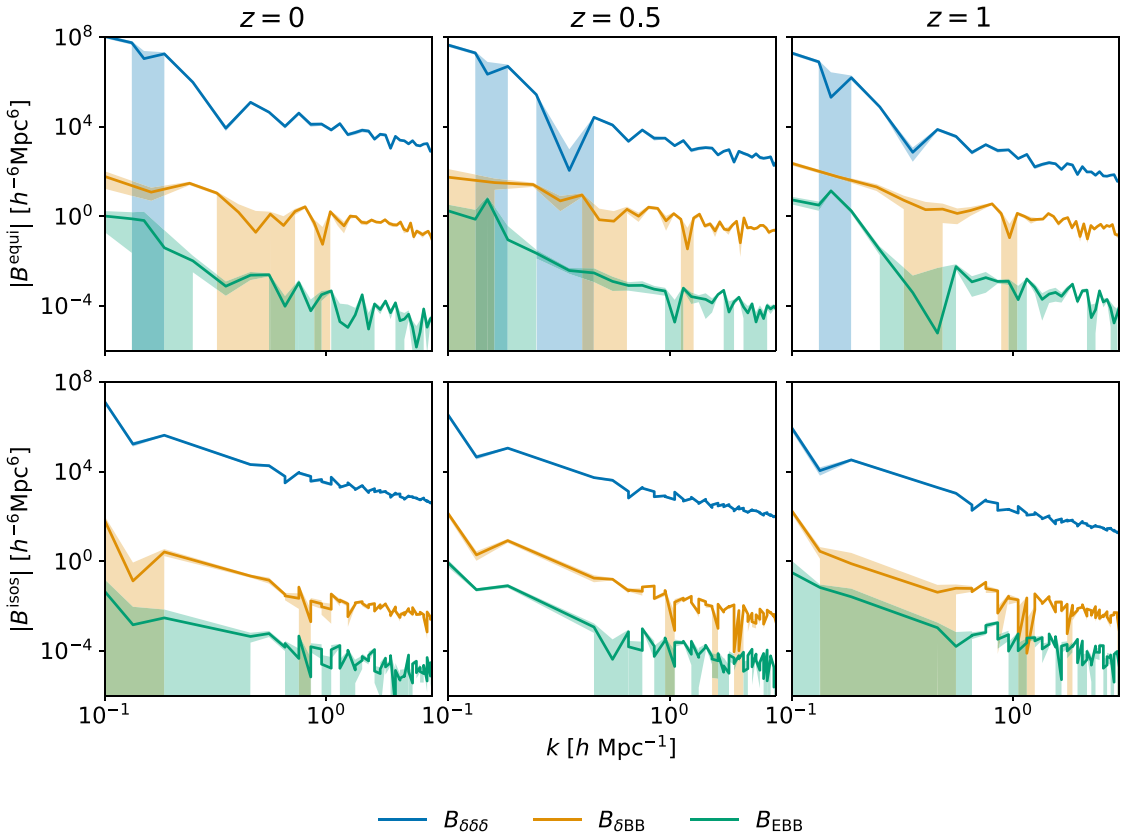

Figure 4. IA bispectra $B_{\delta\delta\mathrm{B}}$ and B_{EBB} measured from simulations. *Top*: Equilateral triangles. *Bottom*: Isosceles triangles with sides in the ratio 2:2:1. The bispectrum $B_{\delta\delta\delta}$ is also shown for reference. Shaded areas are 68 per cent confidence intervals.

Fig. 6 compares the values of A_{IA} obtained from ratios of power spectra to those obtained using equation (35). Results are for $k_{\mathrm{min}} = 0.1\,h\,\mathrm{Mpc}^{-1}$ and varying k_{max}. There is good agreement for all values of k_{max}. This validates the NLA model. The results for P_{EE} are consistently higher than those for $P_{\delta\mathrm{E}}$, suggesting a possible unexplained systematic trend.

5.4 Dependence of IA amplitude on mass

Since the IA power spectra depend strongly on subhalo mass, we also expect A_{IA} to depend on mass. We therefore obtain estimates of A_{IA} from $P_{\delta\mathrm{E}}$ and P_{EE} for each of our four mass bins. We assume that the same linear alignment model applies for all halo masses, with any mass-dependence being absorbed into the estimated amplitude. Alternative ways in which the mass-dependence has been incorporated include inserting a mass-dependent halo bias term into the linear alignment model (Xia et al. 2017) and using a virial argument to derive a scaling by mass (Piras et al. 2018).

We perform the χ^2 minimization using equation (35) for each of the four mass bins and for three redshifts, and also vary the maximum k value used. The results are shown in Fig. 7. The values of A_{IA} obtained from $P_{\delta\mathrm{E}}$ are consistent with fig. 6 in Kurita et al. (2021), given slightly different mass bins and k and z ranges.

We also fit a power law of the form

$$A_{\mathrm{IA}} \propto M_{\mathrm{h}}^{\beta}\,, \tag{38}$$

where M_{h} is the mean halo mass per bin. We estimate this relationship using both $P_{\delta\mathrm{E}}$ and P_{EE}. The results are shown in Fig. 8. At all redshifts the two estimates are reasonably consistent for $k_{\mathrm{max}} \leq 2\,h\,\mathrm{Mpc}^{-1}$, but show some variability for higher k_{max}.

We can compare our estimated power spectrum slopes in Fig. 8 with those found by Piras et al. (2018) from the Millennium simulation. These authors obtained $\beta \approx 0.35$ for $z = 0.46$, with a slightly lower-mass sample of haloes ($10^{11.36} < M_{\mathrm{h}} < 10^{13.36}\,h^{-1}\,\mathrm{M}_{\odot}$). Our value $\beta \approx 0.43$ for $k_{\mathrm{max}} \approx 2\,h\,\mathrm{Mpc}^{-1}$ at $z = 0.5$ is somewhat higher. We note, however, that Piras et al. (2018) repeated their analysis using observational data and found $\beta \approx 0.56$. They suggested several possible reasons for this difference, including the effect of baryons on halo shapes and the relative strength of the stellar and dark matter signals in different mass bins.

5.5 Intrinsic alignment amplitude from equilateral bispectra

We now turn to our main aim: to investigate whether our three-point analytical model is consistent with the two-point NLA model over the non-linear scales of interest. We start with the equilateral bispectrum $B_{\delta\delta\mathrm{E}}^{\mathrm{equi}}$ given by equation (32). We later expand this to non-equilateral versions of $B_{\delta\delta\mathrm{E}}$, but do not consider other IA bispectra because we find these have insufficient signal to provide robust estimates of the amplitude A_{IA}.

We again use equation (35) to fit a parameter $\hat{A}_{\mathrm{IA}}$ but in this case $R(k) = B_{\delta\delta\mathrm{E}}^{\mathrm{equi}}/B_{\delta\delta\delta}$. Also, using equation (32), we now have

$$F(\hat{A}_{\mathrm{IA}}) = \int_0^1 \left(1-\mu^2\right)\,\mathrm{d}\mu\,\tfrac{[f_{\mathrm{IA}}^2+2f_{\mathrm{IA}}]}{3} \tag{39}$$

$$= \tfrac{2}{9}[(c(z)\hat{A}_{\mathrm{IA}})^2 + 2c(z)\hat{A}_{\mathrm{IA}}]\,, \tag{40}$$

where $c(z)$ and $\hat{A}_{\mathrm{IA}}$ are defined as in Section 5.3. In contrast to the power spectra, there is no simple relationship between the ratio $B_{\delta\delta\mathrm{E}}^{\mathrm{equi}}/B_{\delta\delta\delta}$ and the IA amplitude so we cannot produce a plot similar to Fig. 6.

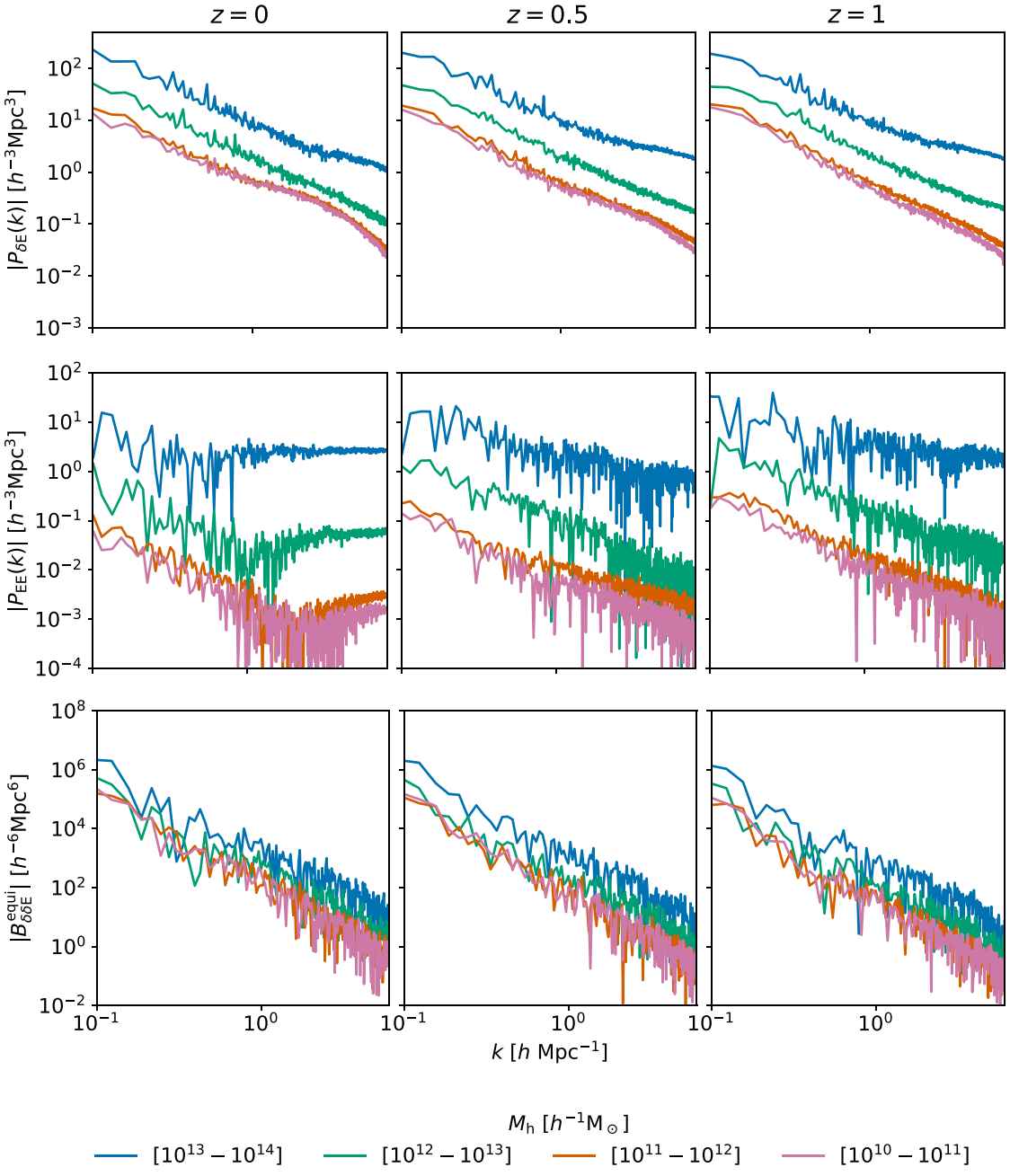

Figure 5. IA spectra in four mass bins. *Top*: $P_{\delta E}$. *Centre*: P_{EE}. *Bottom*: $B^{\mathrm{equi}}_{\delta\delta E}$.

As shown in Fig. 5, $B^{\mathrm{equi}}_{\delta\delta E}$ increases with mass, albeit not as consistently as $P_{\delta E}$, so we again obtain estimates of A_{IA} for each mass bin, and fit a power law to this. Figs 7 and 8 show the resulting estimates of A_{IA} and of the slope β. Note that we fit the bispectrum measurements only for $k \gtrsim 0.7\ h\,\mathrm{Mpc}^{-1}$ in contrast to $k \gtrsim 0.1\ h\,\mathrm{Mpc}^{-1}$ for the power spectrum.

5.6 Comparison of results from power spectra and equilateral bispectra

Table 1 summarizes our main results. It shows estimates of A_{IA} derived from $P_{\delta E}$, P_{EE}, and $B^{\mathrm{equi}}_{\delta\delta E}$ for each mass bin. Also in this table are the estimated power-law slopes β from equation (38) and the mean value of A_{IA}, weighted by the number of subhaloes per mass bin. For illustrative purposes, we choose $k_{\mathrm{max}} = 2\ h\,\mathrm{Mpc}^{-1}$ in this table, well within non-linear scales, since our models appear reliable up to this value. This range also covers the typical fit range for cosmic shear studies. The overall picture would be qualitatively similar with a different k_{max}.

The IA amplitudes obtained from $P_{\delta E}$ and $B_{\delta\delta E}$ are completely consistent, underlining the validity of the analytical modelling up to at least $k_{\mathrm{max}} = 2\ h\,\mathrm{Mpc}^{-1}$ and confirming that our modelling of three-point IA statistics is consistent with the two-point NLA model. The power-law relationship that we obtain using $B^{\mathrm{equi}}_{\delta\delta E}$ is broadly consistent with the power spectrum results but is rather lower in most cases and more dependent on the value of k_{max}.

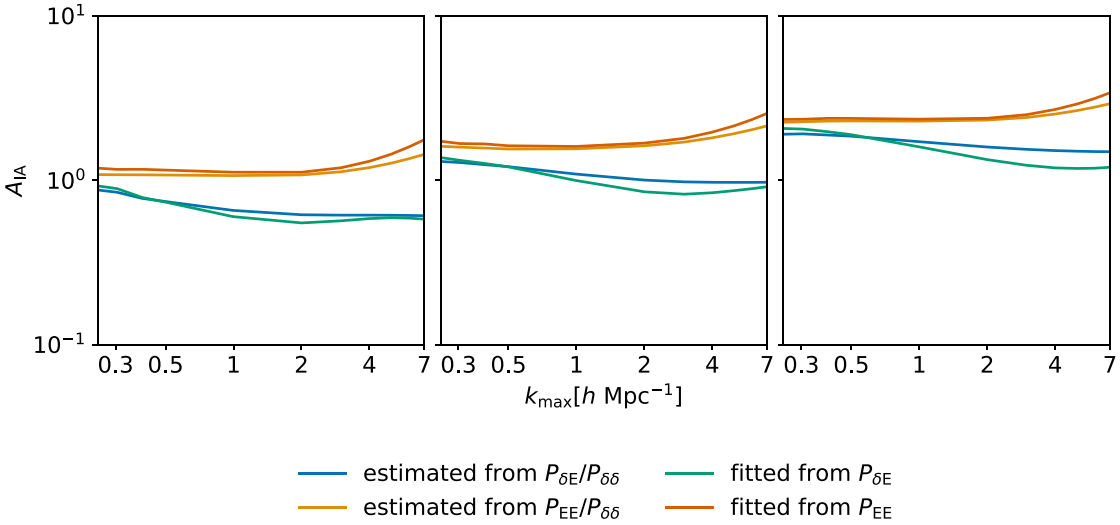

Figure 6. The mean IA amplitude $A_{\rm IA}$ as a function of $k_{\rm max}$ estimated from the relationship between the IA power spectra and the matter power spectrum given by equations (26) and (27) and also fitted using equation (35) over the range $[0.1, k_{\rm max}]\,h\,{\rm Mpc}^{-1}$.

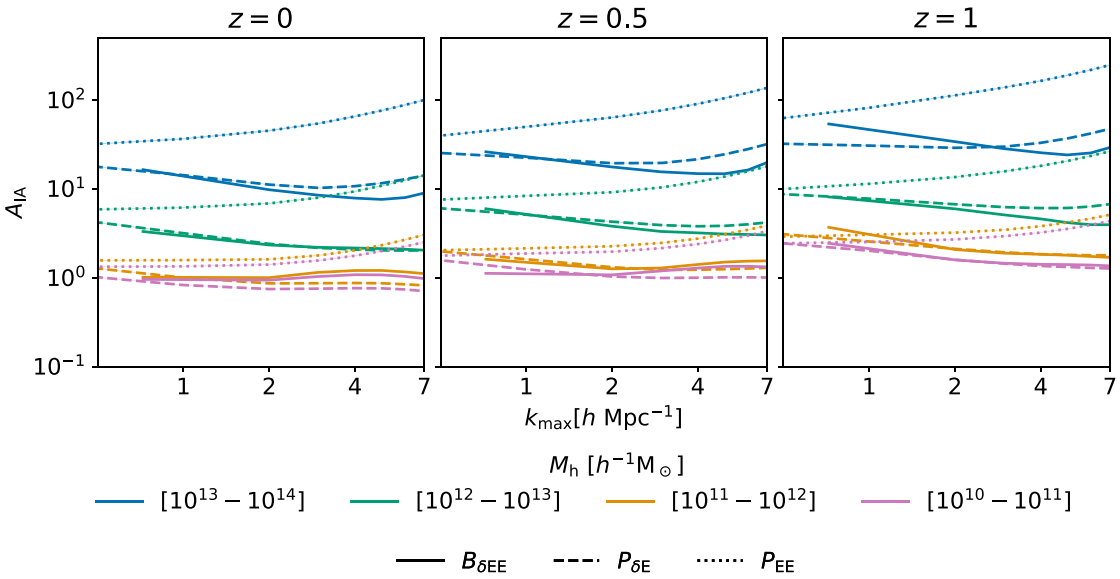

Figure 7. The IA amplitude $A_{\rm IA}$ estimated from $P_{\delta{\rm E}}$, $P_{\rm EE}$, and $B^{\rm equi}_{\delta\delta{\rm E}}$ for four subhalo mass ranges, as a function of $k_{\rm max}$. Typical uncertainties in the fits are given in Table 1.

5.7 Analytical predictions for isosceles bispectra

It is worth exploring whether our analytical model can also be applied to non-equilateral bispectra since these can contain more information than equilateral configurations (Appendix B). In this case, it is not possible to obtain simplified expressions similar to those in equations (32)–(34) because the perturbation theory kernels do not cancel out. Instead, we consider whether the IA amplitudes predicted from $B^{\rm equi}_{\delta\delta{\rm E}}$ are also valid for non-equilateral bispectra. For illustration, we again consider isosceles triangles with one side with magnitude k and two sides with magnitude $2k$.

We take the estimates of $A_{\rm IA}$ obtained from $B^{\rm equi}_{\delta\delta{\rm E}}$ in the final column of Table 1, noting that these are very similar to the estimates obtained from $P_{\delta{\rm E}}$. We calculate $f_{\rm IA}$ from equation (25) using the same assumptions as in Section 5.3, obtain the non-linear matter power spectrum from Halofit (Takahashi et al. 2012), and insert the resulting values into equations (29)–(31). In Fig. 9, we compare these predicted bispectra with those measured from simulations, finding a reasonable fit across scales and redshifts. This confirms that our empirical IA bispectrum model works for both equilateral and isosceles triangle configurations.

5.8 Decorrelation of matter overdensity and E-mode fields

In Appendix C, we discuss a possible phenomenological change to our two-point and three-point analytical models that adjusts them for the correlation between the matter overdensity field and the E-mode 'field'. We measure the correlation between the fields from simulations, and introduce the resulting correlation coefficient into the analytical model. In most cases, this improves the fit of the models considerably. This is not central to our main results but may be worth considering for future IA modelling.

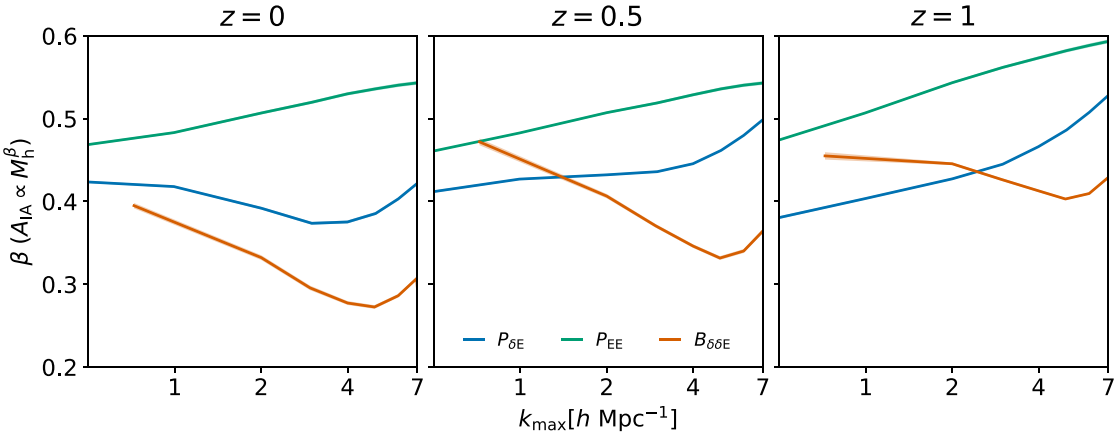

Figure 8. Slope β from a power law relationship $A_{\rm IA} \propto M_{\rm h}^{\beta}$, estimated from $P_{\delta \rm E}$, $P_{\rm EE}$, and $B_{\delta\delta \rm E}^{\rm equi}$. Typical uncertainties in the fits are given in Table 1.

Table 1. Estimates of the IA amplitude $A_{\rm IA}$ and the power-law exponent β in equation (38) for subhaloes in four mass ranges, with 68 per cent confidence intervals, from the power spectra $P_{\delta \rm E}$ and $P_{\rm EE}$ and the bispectrum $B_{\delta\delta \rm E}^{\rm equi}$ with $k_{\rm max} = 2\ h\,{\rm Mpc}^{-1}$. Also shown are the mean $A_{\rm IA}$ for all masses, calculated as the average over mass bins weighted by the number of haloes per bin. For both power spectra, $k_{\rm min} \approx 0.1\ h\,{\rm Mpc}^{-1}$ and for the bispectrum $k_{\rm min} \approx 0.7\ h\,{\rm Mpc}^{-1}$.

		Subhalo mass range ($h^{-1}\,{\rm M}_\odot$)					
	Redshift	$10^{10}-10^{11}$	$10^{11}-10^{12}$	$10^{12}-10^{13}$	$10^{13}-10^{14}$	Slope β	Mean $A_{\rm IA}$
From power spectrum $P_{\delta \rm E}$	0.0	0.74 ± 0.01	0.86 ± 0.01	2.40 ± 0.03	11.13 ± 0.12	0.392 ± 0.002	0.91 ± 0.01
	0.5	1.02 ± 0.01	1.31 ± 0.01	4.26 ± 0.04	19.38 ± 0.14	0.432 ± 0.001	1.36 ± 0.01
	1.0	1.59 ± 0.01	2.10 ± 0.02	6.68 ± 0.06	28.87 ± 0.21	0.427 ± 0.001	2.13 ± 0.02
From power spectrum $P_{\rm EE}$	0.0	1.40 ± 0.02	1.61 ± 0.02	6.84 ± 0.09	45.08 ± 0.81	0.507 ± 0.002	1.90 ± 0.02
	0.5	1.96 ± 0.02	2.25 ± 0.02	9.14 ± 0.09	63.43 ± 1.01	0.507 ± 0.001	2.64 ± 0.02
	1.0	2.69 ± 0.02	3.21 ± 0.02	13.54 ± 0.15	112.53 ± 1.47	0.543 ± 0.001	3.81 ± 0.03
From bispectrum $B_{\delta\delta \rm E}$	0.0	0.94 ± 0.01	0.99 ± 0.01	2.33 ± 0.03	9.74 ± 0.16	0.332 ± 0.003	1.06 ± 0.01
	0.5	1.07 ± 0.01	1.25 ± 0.01	3.79 ± 0.05	17.59 ± 0.23	0.407 ± 0.002	1.34 ± 0.01
	1.0	1.58 ± 0.02	2.07 ± 0.02	5.94 ± 0.09	33.88 ± 0.39	0.446 ± 0.002	2.11 ± 0.03

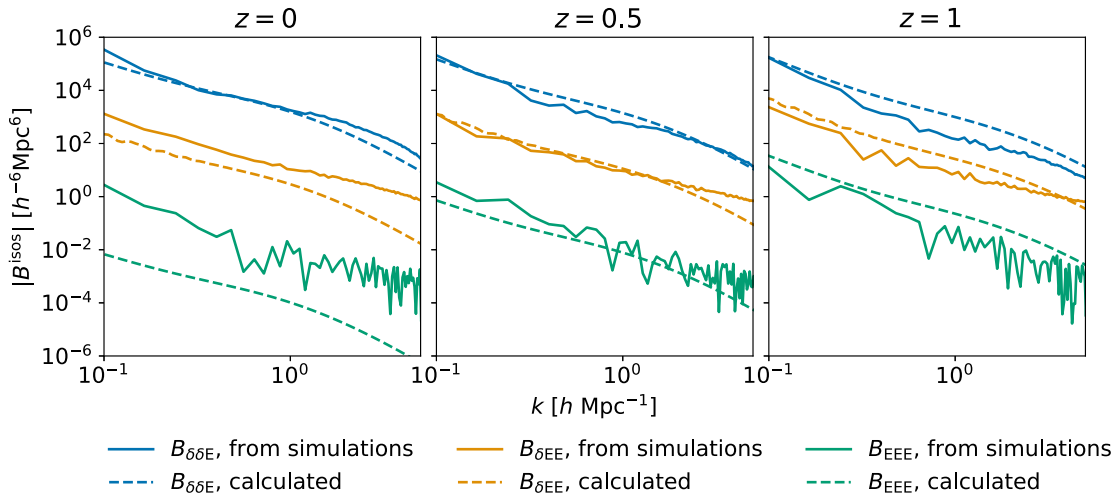

Figure 9. IA bispectra for isosceles triangles with one side of magnitude k and two sides of magnitude $2k$. *Solid lines*: Measured from simulations. *Dashed lines*: Calculated from equations (29)–(31) with best-fitting values of $A_{\rm IA}$ estimated from $B_{\delta\delta \rm E}^{\rm equi}$.

6 COMPARISON WITH PREVIOUS WORK

The most relevant previous work is by Kurita et al. (2021) who measured IA power spectra from simulations; Semboloni et al. (2008) who measured three-point IA statistics from simulations and compared them with cosmic shear; and Pyne & Joachimi (2021) who used analytical models to investigate the effect of IAs on the weak lensing power spectrum and bispectrum.

Our present work is most directly comparable with Kurita et al. (2021) who used the same methodology but a different simulation suite. Both studies explore the correlation between the intrinsic shear field and the matter overdensity field, and both consider dark matter haloes rather than galaxies. The power spectrum results presented in the current work are consistent with those in Kurita et al. (2021), giving confidence that the bispectrum measurements we present here are also sound.

Semboloni et al. (2008) built on the work of Heymans et al. (2006) who measured two-point (GI) correlations between the shapes of foreground galaxies and the weak lensing shear of source galaxies in a suite of *N*-body simulations, and also shape–shape (II) correlations. Semboloni et al. (2008) used the same simulations to measure three-point statistics (GGG, GGI, GII, and III correlations). Their main findings were that for surveys whose median redshift was around 0.7 the II and III terms were consistent with zero, but that for shallower surveys with $z_{\rm med} \approx 0.3$ the II/GG and III/GGG ratios were non-zero, and the III ratio could be a factor of 10 higher than the II ratio. Later Semboloni et al. (2010) also used these simulations to validate measurements of three-point shear statistics from observations. These studies are relevant to this work in that they measured three-point IA statistics, but not easily comparable because they focused on IA contamination of the cosmic shear signal.

A likely reason for any inconsistencies between the findings in this paper and those of Semboloni et al. (2008) is the different mass resolution of the simulations. The simulations used in Semboloni et al. (2008) and related works were state of the art at the time, with 512^3 particles in a periodic cubic box measuring 300 $h\,{\rm Mpc}^{-1}$ per side. However, the mass resolution was low with a particle mass of $1.7 \times 10^{10}\,h^{-1}\,{\rm M}_\odot$, compared with $4 \times 10^{7}\,h^{-1}\,{\rm M}_\odot$ in IllustrisTNG300. As a result, the smallest bound haloes that Heymans et al. (2006) could identify had masses several times $10^{11}\,h^{-1}\,{\rm M}_\odot$ whereas in the current work we use haloes with mass as low as around $4 \times 10^{10}\,h^{-1}\,{\rm M}_\odot$. This is despite the fact that we include only haloes with at least 1000 particles, in contrast to Heymans et al. (2006) who allowed a considerably smaller minimum particle number. Semboloni et al. (2008) explicitly stated that the lack of low-mass haloes in their sample was a limitation. The results in Heymans et al. (2006) and Semboloni et al. (2008) also depend on models that they used to populate each halo with a single spiral or elliptical galaxy. Although it would be possible to do an approximate lensing calculation to compare our results with Semboloni et al. (2008), in view of the many differences between the two studies we consider this would involve too many assumptions to be useful.

Pyne & Joachimi (2021) modelled the GGI, GII, and III correlations analytically using the same approach as in this paper (the NLA model extended to three-point statistics). This work also found that IAs affected two-point and three-point weak lensing statistics differently, although not consistently with the results in Semboloni et al. (2008). Again, this disagreement may be attributable to the low mass resolution of their simulations.

7 SUMMARY AND DISCUSSION

We have measured IA bispectra of dark matter subhaloes from the IllustrisTNG300-1 cosmological simulation suite, building on the power spectrum methodology developed by Kurita et al. (2021). We also measured the IA power spectra $P_{\delta{\rm E}}$ and $P_{\rm EE}$ and confirmed that they are consistent with results obtained by Kurita et al. (2021) using the DarkQuest simulation suite.

At all redshifts, the cumulative signal-to-noise ratios of the IA bispectra were well below those we obtained for power spectra – for example, 10–15 per cent at $k_{\rm max} \approx 4\,h\,{\rm Mpc}^{-1}$ compared with around 40 per cent for power spectra. The bispectrum $B_{\rm EEE}$ and measured bispectra involving B-modes have very low signal-to-noise ratios and provide no useful information. However, we found that the E-mode bispectra $B_{\delta\delta{\rm E}}$ and $B_{\delta{\rm EE}}$ do have useful information content. Signal-to-noise ratios for the non-equilateral triangles that we studied are notably higher than for equilateral triangles, suggesting that the common simplification of using only equilateral triangles in bispectrum analyses is sub-optimal (Yankelevich et al. 2022).

All the IA power spectra and bispectra we studied showed similar strong relationships with subhalo mass. This can be traced back to the relationship between halo ellipticity and mass. As discussed in Section 5.1, this mass dependence has been noted by others in both simulations and observations. Its origin is debated. Smith & Watts (2005) attributed it to higher mass haloes forming later so they have had less time to virialize and therefore retain more memory of the tidal fields at the time they formed. More recently, Xia et al. (2017) postulated that the IA strength depends independently on both halo formation time and mass, whereas Piras et al. (2018) suggested that higher-mass haloes experience stronger tidal fluctuations.

We used the standard NLA model to estimate the IA amplitude from both $P_{\delta{\rm E}}$ and $P_{\rm EE}$, and found that our estimates are consistent with corresponding results obtained by Kurita et al. (2021). This validated our two-point modelling.

For IA bispectra, we used the analytical model from Pyne & Joachimi (2021), which is in the spirit of the NLA model. From this we again estimated the IA amplitude, this time from equilateral bispectra. We found that the best-fitting amplitudes $A_{\rm IA}$ obtained using $B^{\rm equi}_{\delta\delta{\rm E}}$ were extremely close to those obtained from $P_{\delta{\rm E}}$. It is not possible to use the same methodology to estimate $A_{\rm IA}$ from non-equilateral triangles but we showed that the predicted $A_{\rm IA}$ from $B^{\rm equi}_{\delta\delta{\rm E}}$ produced an acceptable fit to simulation measurements of IA bispectra of specific isosceles triangles.

We fitted power-law relationships between the estimated $A_{\rm IA}$ and subhalo mass, obtaining almost identical relationships from $P_{\delta{\rm E}}$ and $B^{\rm equi}_{\delta\delta{\rm E}}$. These also agreed approximately with the relationship found by Piras et al. (2018) using power spectrum measurements from the Millennium simulation. It is interesting that all our estimated IA amplitudes are similar for the two lowest-mass bins and larger for the high-mass bins. This hints at a possible broken power-law relationship with mass similar to the luminosity relationship found by Fortuna et al. (2021) for a sample of luminous red galaxies from KiDS-1000 (Kuijken et al. 2019).

Our bispectrum measurements are an advance on the early three-point IA measurements from simulations reported in Semboloni et al. (2008) since we have been able to take advantage of the improved resolution and methodology of the IllustrisTNG simulation suite. Semboloni et al. (2008) focused on the magnitude of the IA effect relative to cosmic shear, whereas we considered the correlation between the IA shear and the matter overdensity field. There are also several other detailed differences between the two studies and, as discussed in Section 6, many assumptions would need to be made to compare our results directly with theirs. It is more useful to consider the results from Pyne & Joachimi (2021) who estimated ratios between IA and lensing signals using the same analytical models as in this paper. Like Semboloni et al. (2008), this work found that two-point and three-point weak lensing statistics are affected differently by IA. The present work has shown that our analytical models agree well with measurements from IllustrisTNG which validates the results reported in Pyne & Joachimi (2021).

The fits between the measured and modelled IA power spectra and bispectra are not perfect. In Appendix C, we suggest a possible phenomenological modification of the analytical models based on the correlation between the matter overdensity and E-mode fields measured in simulations. This may be worth considering for future IA modelling.

Overall our results demonstrate that a single physically motivated analytical approach can be applied to both two-point and three-point IA statistics, enabling a cleaner separation between IA and weak lensing signals. This opens up the prospect of using such a model in joint power spectrum–bispectrum analysis. Pyne & Joachimi (2021) showed that such analysis can allow self-calibration to mitigate IA contamination of weak lensing data in forthcoming surveys. This is particularly pertinent in the light of advances in the measurement of three-point statistics from survey data. Secco et al. (2022b) recently reported high signal-to-noise detections of three-point shear correlations and aperture mass statistics in the first three years of data from the Dark Energy Survey. These measurements lay the foundations for joint two- and three-point cosmological analyses which will of course need tight control of systematics such as IAs.

This work has considered only dark matter haloes and so has only limited application to observational data. In future work, we plan to build on the measurement techniques and modelling in this paper to confirm that the three-point approach can be extended to galaxies and to investigate the impact of galaxy characteristics and environment.

ACKNOWLEDGEMENTS

We thank the IllustrisTNG team for making their simulation data publicly available. We are also grateful to our anonymous reviewer for many constructive suggestions for improving the paper. SP also thanks Harry Johnston and Anik Halder for helpful discussions. This work was partially enabled by funding from the University College London Cosmoparticle Initiative.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Ade P. A. et al., 2016, A&A, 594, A13
Bernardeau F., Colombi S., Gaztanaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Bernstein G., Jarvis M., 2002, AJ, 123, 583
Blazek J., Vlah Z., Seljak U., 2015, J. Cosmol. Astropart. Phys., 2015, 015
Blazek J. A., MacCrann N., Troxel M., Fang X., 2019, Phys. Rev. D, 100, 103506
Bridle S., King L., 2007, New J. Phys., 9, 444
Chisari N. E., Dvorkin C., 2013, J. Cosmol. Astropart. Phys., 2013, 029
Crittenden R. G., Natarajan P., Pen U.-L., Theuns T., 2002, ApJ, 568, 20
Davis M., Efstathiou G., Frenk C. S., White S. D., 1985, ApJ, 292, 371
Foreman S., Coulton W., Villaescusa-Navarro F., Barreira A., 2020, MNRAS, 498, 2887
Fortuna M. C. et al., 2021, A&A, 654, A76
Fu L. et al., 2014, MNRAS, 441, 2725
Gil-Marín H., Wagner C., Fragkoudi F., Jimenez R., Verde L., 2012, J. Cosmol. Astropart. Phys., 2012, 047
Hand N., Feng Y., Beutler F., Li Y., Modi C., Seljak U., Slepian Z., 2018, AJ, 156, 160
Heymans C., White M., Heavens A., Vale C., Van Waerbeke L., 2006, MNRAS, 371, 750
Heymans C. et al., 2012, MNRAS, 427, 146
Hirata C. M., Seljak U., 2004, Phys. Rev. D, 70, 063526
Huterer D., Takada M., Bernstein G., Jain B., 2006, MNRAS, 366, 101
Jing Y., 2002, MNRAS, 335, L89
Joachimi B., Mandelbaum R., Abdalla F., Bridle S., 2011, A&A, 527, A26
Joachimi B., Semboloni E., Bett P., Hartlap J., Hilbert S., Hoekstra H., Schneider P., Schrabback T., 2013, MNRAS, 431, 477
Joachimi B. et al., 2015, Space Sci. Rev., 193, 1
Joachimi B. et al., 2021, A&A, 646, A129
Kamionkowski M., Babul A., Cress C. M., Refregier A., 1998, MNRAS, 301, 1064
Kiessling A. et al., 2015, Space Sci. Rev., 193, 67
Kirk D., Rassat A., Host O., Bridle S., 2012, MNRAS, 424, 1647
Kirk D. et al., 2015, Space Sci. Rev., 193, 139
Kuijken K. et al., 2019, A&A, 625, A2
Kurita T., Takada M., Nishimichi T., Takahashi R., Osato K., Kobayashi Y., 2021, MNRAS, 501, 833
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lee J., Springel V., Pen U.-L., Lemson G., 2008, MNRAS, 389, 1266
Marinacci F. et al., 2018, MNRAS, 480, 5113
Merkel P. M., Schäfer B. M., 2014, MNRAS, 445, 2918
Miyatake H. et al., 2021, preprint (arXiv:2101.00113)
Naiman J. P. et al., 2018, MNRAS, 477, 1206
Namikawa T., Bose B., Bouchet F. R., Takahashi R., Taruya A., 2019, Phys. Rev. D, 99, 063511
Nelson D. et al., 2018, MNRAS, 475, 624
Nelson D. et al., 2019, Comput. Astrophys. Cosmol., 6, 1
Nishimichi T. et al., 2019, ApJ, 884, 29
Pillepich A. et al., 2018, MNRAS, 475, 648
Piras D., Joachimi B., Schäfer B. M., Bonamigo M., Hilbert S., van Uitert E., 2018, MNRAS, 474, 1165
Pyne S., Joachimi B., 2021, MNRAS, 503, 2300
Schneider M. D., Bridle S., 2010, MNRAS, 402, 2127
Schneider P., Seitz C., 1995, A&A, 294, 411
Schneider P., van Waerbeke L., Mellier Y., 2002, A&A, 389, 729
Schneider M. D., Frenk C. S., Cole S., 2012, J. Cosmol. Astropart. Phys., 2012, 030
Scoccimarro R., Couchman H., 2001, MNRAS, 325, 1312
Scoville N. et al., 2007, ApJS, 172, 1
Secco L. et al., 2022a, Phys. Rev. D, 105, 023515
Secco L. et al., 2022b, Phys. Rev. D, 105, 103537
Seitz C., Schneider P., 1997, A&A, 318, 687
Semboloni E., Heymans C., Van Waerbeke L., Schneider P., 2008, MNRAS, 388, 991
Semboloni E., Schrabback T., van Waerbeke L., Vafaei S., Hartlap J., Hilbert S., 2010, MNRAS, 410, 143
Shi J., Kurita T., Takada M., Osato K., Kobayashi Y., Nishimichi T., 2021, J. Cosmol. Astropart. Phys., 2021, 030
Singh S., Mandelbaum R., More S., 2015, MNRAS, 450, 2195
Smith R. E., Watts P., 2005, MNRAS, 360, 203
Springel V., 2010, MNRAS, 401, 791
Springel V., White S. D., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Springel V. et al., 2018, MNRAS, 475, 676
Stebbins A., 1996, preprint (astro-ph/9609149)
Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152
Takahashi R., Nishimichi T., Namikawa T., Taruya A., Kayo I., Osato K., Kobayashi Y., Shirasaki M., 2020, ApJ, 895, 113
Taruya A., Okumura T., 2020, ApJ, 891, L42
Tenneti A., Mandelbaum R., Di Matteo T., Feng Y., Khandai N., 2014, MNRAS, 441, 470
Tenneti A., Singh S., Mandelbaum R., Matteo T. D., Feng Y., Khandai N., 2015, MNRAS, 448, 3522
Tomlinson J., Jeong D., Kim J., 2019, AJ, 158, 116
Troxel M. A., Ishak M., 2011, MNRAS, 419, 1804

Troxel M. A., Ishak M., 2012, MNRAS, 423, 1663
Troxel M., Ishak M., 2015, Phys. Rep., 558, 1
van Uitert E., Joachimi B., 2017, MNRAS, 468, 4502
Viola M., Kitching T., Joachimi B., 2014, MNRAS, 439, 1909
Vlah Z., Chisari N. E., Schmidt F., 2020, J. Cosmol. Astropart. Phys., 2020, 025
Vlah Z., Chisari N. E., Schmidt F., 2021, J. Cosmol. Astropart. Phys., 2021, 061
Xia Q., Kang X., Wang P., Luo Y., Yang X., Jing Y., Wang H., Mo H., 2017, ApJ, 848, 22
Yankelevich V., McCarthy I. G., Kwan J., Stafford S. G., Liu J., 2022, preprint (arXiv:2202.07680)
Zuntz J. et al., 2015, Astron. Comput., 12, 45

APPENDIX A: COMPARISON BETWEEN SIMULATIONS AND ANALYTICAL CALCULATIONS OF THE MATTER BISPECTRUM

To confirm the suitability of the perturbation theory-based formula from Gil-Marín et al. (2012) for our purposes, Fig. A1 compares it with our matter bispectrum measurements at $z = 1$ from the IllustrisTNG300-1 simulations. Also shown are more recent analytical estimates using Bihalofit (Takahashi et al. 2020) as well as tree-level perturbation theory bispectra (Bernardeau et al. 2002). We show results for equilateral triangles and for isosceles triangles with sides in the ratio 2:2:1. All the separate results for the non-linear bispectrum are consistent at the scales that we are interested in, with the simulation measurements and analytical estimates from Gil-Marín et al. (2012) only diverging significantly for $k \gtrsim 3\,h\,\mathrm{Mpc}^{-1}$.

APPENDIX B: SIGNAL-TO-NOISE RATIOS

We approximate the cumulative signal-to-noise ratio (SNR) of a power spectrum $P(k)$ as

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2 = \sum_{k_i=k_{\min}}^{k_{\max}} \frac{[P(k_i)]^2}{\sigma_i^2}\,, \tag{B1}$$

where σ_i^2 is the variance in the ith bin, measured using jackknife sampling, and we consider only diagonal terms of the covariance. Similar definitions apply to the IA power spectra and bispectra. Equation (B1) is a simplification that somewhat overestimates the signal-to-noise ratio. Nevertheless, it allows a useful comparison between the information content of different spectra.

Fig. B1 shows, as a function of $k_{\max}$, the cumulative SNR for the matter power spectrum and also for E-mode IA power spectra. Fig. B2 shows similar information for the E-mode IA bispectra, for equilateral triangles and for isosceles triangles with sides in the ratio 2:2:1.

The main messages to take from these figures are that the SNR for P_{EE} is much weaker than for $P_{\delta\mathrm{E}}$; the power spectra contain much more information than the bispectra; and the IA spectra $B_{\delta\mathrm{EE}}$ and B_{EEE} contain very little signal in the IllustrisTNG300-1 volume. For this reason, we use only $B_{\delta\delta\mathrm{E}}$ in most parts of this work. The $P_{\delta\mathrm{E}}$ SNR appears surprisingly strong in comparison to that for $P_{\delta\delta}$ but this could be due to the simplifications in our calculations. We also confirmed that in each case the cumulative SNRs for individual mass bins are typically within 80 per cent of the SNRs for the whole sample.

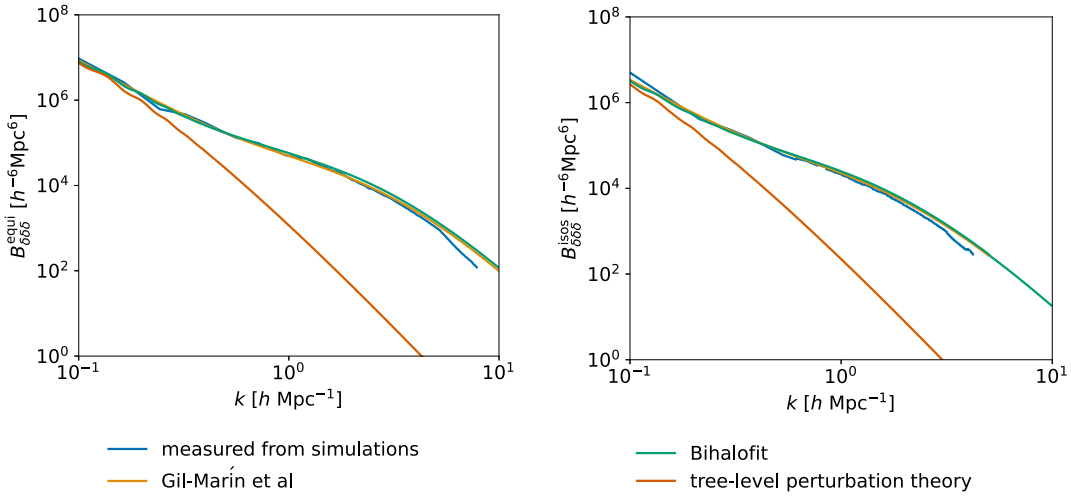

Figure A1. Comparison between the matter bispectrum at $z = 1$ measured from the IllustrisTNG300-1 simulation, calculated using the fitting formula from Gil-Marín et al. (2012), and calculated using the more accurate Bihalofit fitting formula (Takahashi et al. 2020). Also shown are the bispectra calculated from tree-level perturbation theory (Bernardeau et al. 2002). *Left*: Equilateral triangles, *right*: isosceles triangles with sides in the ratio 2:2:1.

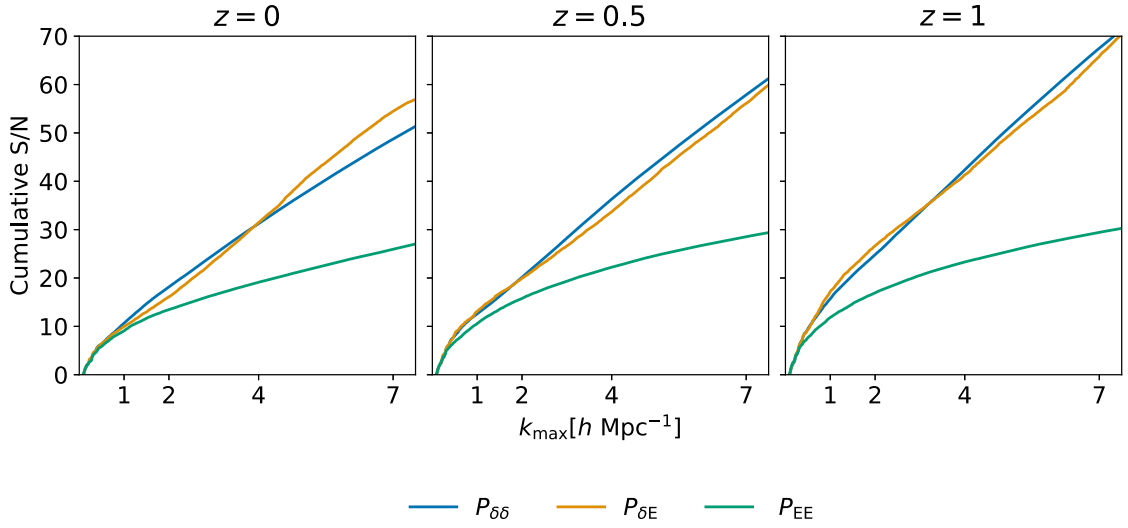

Figure B1. Cumulative signal-to-noise ratios for the matter power spectrum and E-mode IA power spectra measured from simulations.

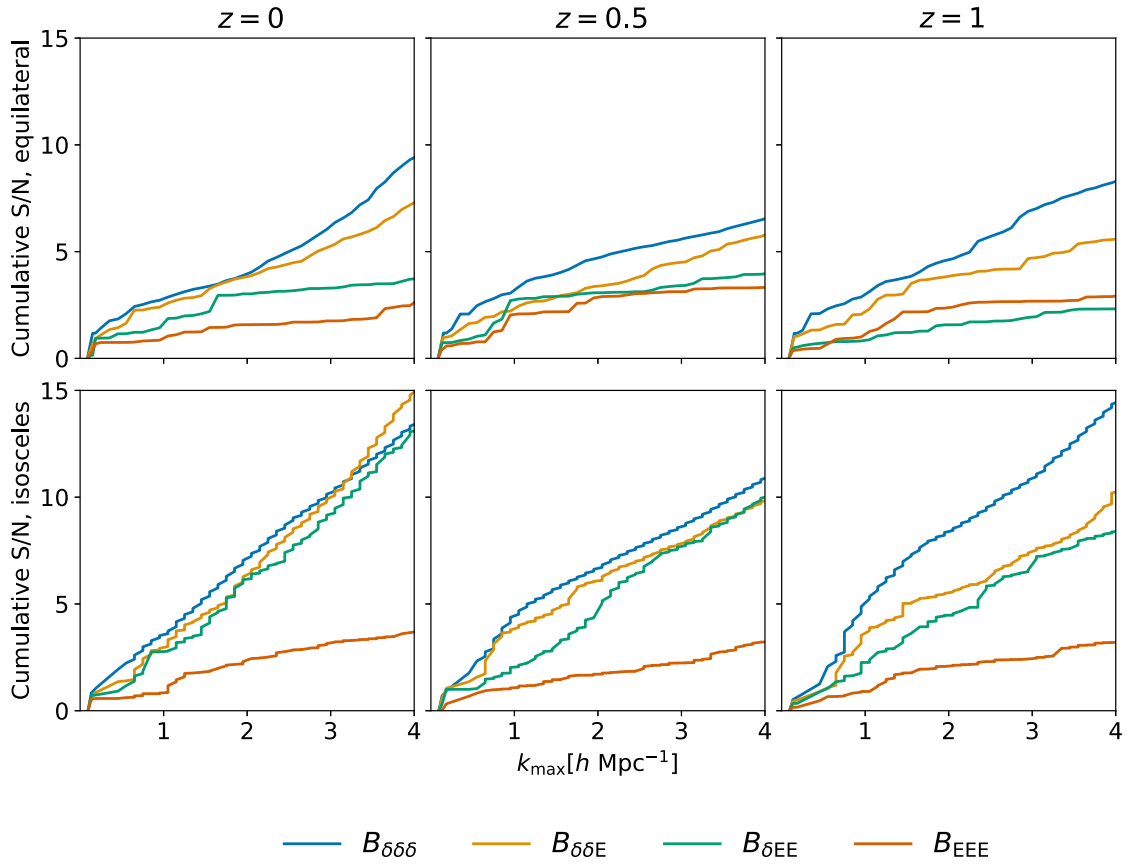

Figure B2. Cumulative signal-to-noise ratios for the matter bispectrum and E-mode IA bispectra measured from simulations. *Top*: Equilateral triangles. *Bottom*: Isosceles triangles with sides in ratio 2:2:1.

APPENDIX C: PHENOMENOLOGICAL ADJUSTMENT TO ANALYTICAL MODELS

In Fig. 9, we compare IA isosceles bispectra measurements from simulations with analytical predictions and show that there is a close but not perfect match between the two estimates. Figs C1 and C2 respectively show analogous results for IA power spectra and equilateral bispectra. Again the fits are close but not exact.

Looking more closely at Fig. C1, the simulations and analytical calculations match well for $P_{\rm EE}$; the discrepancy is only in $P_{\delta {\rm E}}$ which correlates the matter overdensity and E-mode fields (the latter is not a true field but we treat it as such). From Figs 9 and C2, we see that the fit is not perfect for any of the bispectra. The measured $B_{\rm EEE}$ (isosceles and equilateral) are noisy and in this case part of the discrepancy may

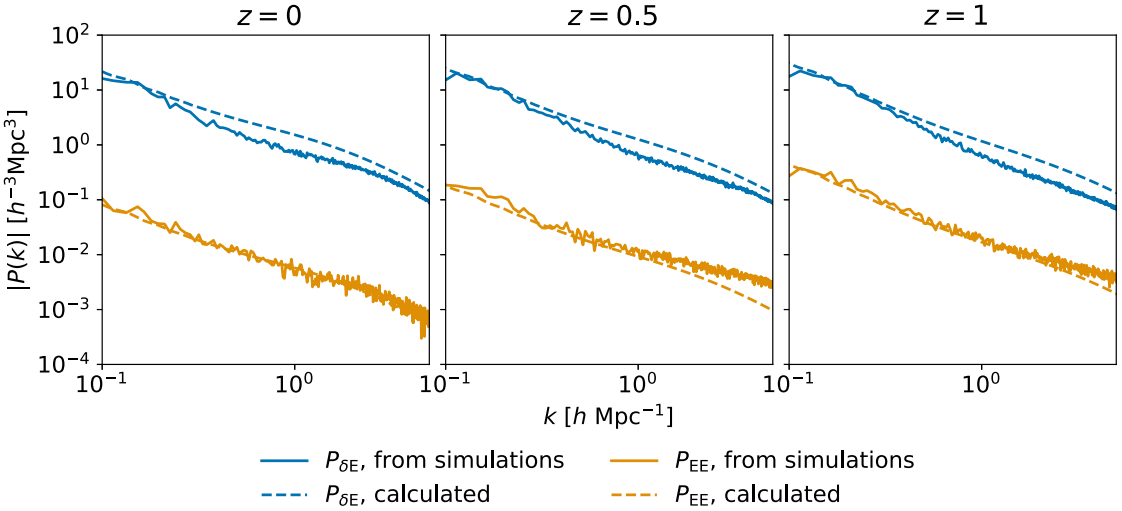

Figure C1. IA power spectra $P_{\delta\mathrm{E}}$ and P_{EE} estimated from simulations (solid lines) and calculated from the NLA model with best-fitting values of A_{IA} estimated from $P_{\delta\mathrm{E}}$ (dashed lines).

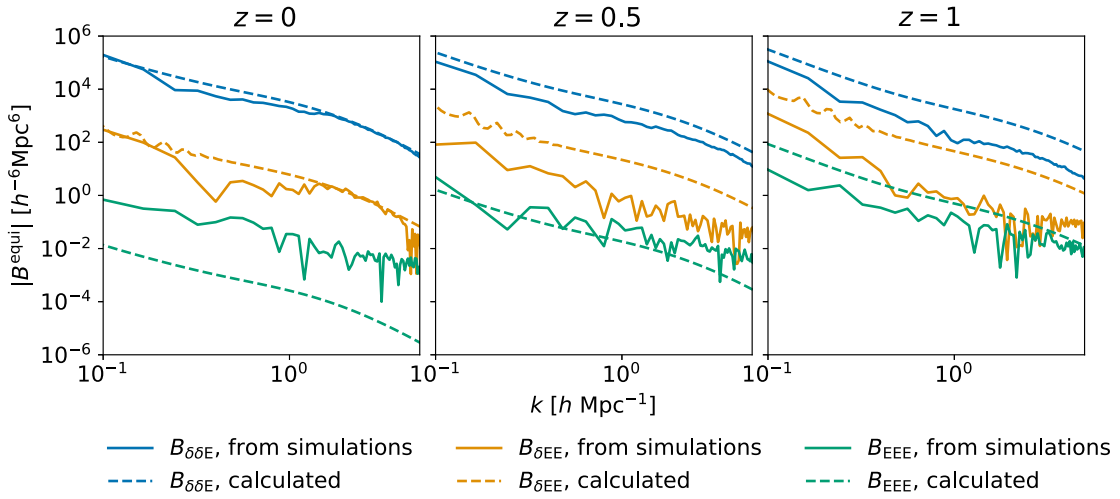

Figure C2. IA bispectra for equilateral triangles. *Solid lines*: Measured from simulations. *Dashed lines*: Calculated from equations (29)–(31) with best-fitting values of A_{IA} estimated from $B^{\mathrm{equi}}_{\delta\delta\mathrm{E}}$.

be due simply to the uncertainty in the measurements. We therefore consider here how to improve the fits for the spectra that depend on both the matter density and E-mode fields. We tackle this by considering the correlation between the measured power spectra of the two fields. The correlation coefficient $r(k)$ between the power spectra can be defined as (Kurita et al. 2021)

$$r^2(k) = \frac{6}{5}\frac{P_{\delta\mathrm{E}}(k)^2}{P_{\delta\delta}(k)P_{\mathrm{EE}}(k)} \,. \tag{C1}$$

The normalizing factor 6/5 arises from integration over μ of the terms in $(1-\mu^2)$ in equations (26) and (27):

$$\frac{\left[\int_0^1 \left(1-\mu^2\right)\,\mathrm{d}\mu\right]^2}{\int_0^1 \left(1-\mu^2\right)^2\,\mathrm{d}\mu} = \frac{5}{6} \,. \tag{C2}$$

With this normalization, $r(k)$ should be equal to unity if the NLA model is valid. Our calculated correlation coefficients at three redshifts are shown in blue in Fig. C3. The coefficients asymptote to approximately one as $k \to 0$ but reduce to a lower, redshift-dependent value for $k \gtrsim 1$, apart from $z = 0$ where the correlation coefficient is very noisy at large values of k.

To model the correlation coefficients at each redshift we fit a generalized logistic curve of the form

$$r(k) = r_0 + \frac{r_1 - r_0}{\left[1 + a\exp\left(-b(k-k_0)\right)\right]^{1/\nu}} \,. \tag{C3}$$

This has six free parameters: r_0 and r_1 are the lower and upper asymptotes of the curve, k_0 determines the k value at which the curve starts to decline, and the remaining parameters, a, b, and ν, control the shape of the curve. In particular, b controls the rate at which the curve decreases as k increases. We normalize the fitted curves to be equal to one at k_{min}. For $z = 0$, because of the noisy data at larger k, we set $r(k) = r(1.0)$ for $k \geq 1.0\,h\mathrm{Mpc}^{-1}$. The fitted curves are shown in green in Fig. C3.

It is difficult to fit a single model which takes redshift into account since we only have data for three redshifts. Instead, we fit the model separately for each redshift, leading to the best-fitting parameter values shown in Table C1. The values for $z = 0.5$ and $z = 1$ are similar so it is plausible that single model could in fact be constructed for both these redshifts.

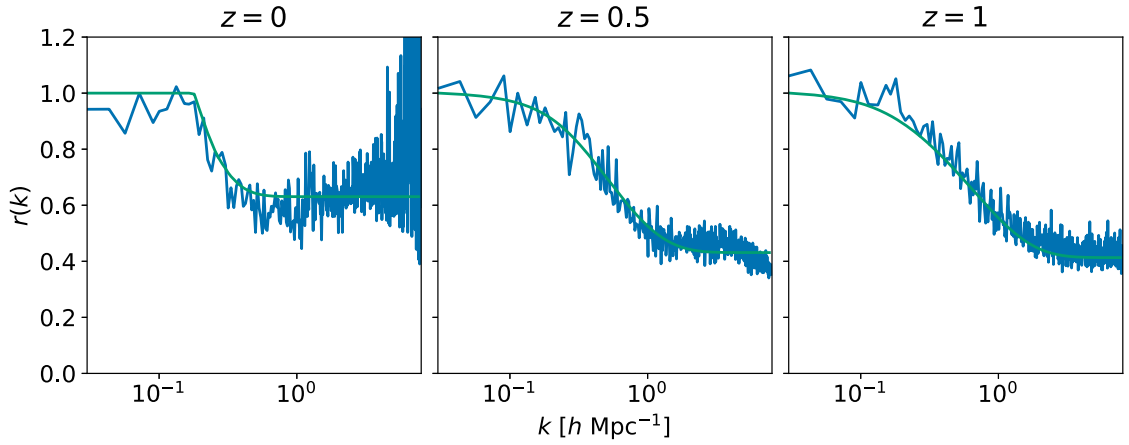

Figure C3. *Blue*: Correlation coefficients $r(k)$ between the E-mode field and the matter density field as defined by equation (C1), for three redshifts. *Green*: Logistic curve fitted to the calculated correlation coefficients using equation (C3) with parameter values from Table C1.

Table C1. Best-fitting parameter values and normalization factor for the modelled correlation coefficients given by equation (C3).

Redshift	r_0	r_1	a	b	ν	k_0	Norm
0.0	0.6	0.95	1.03	614	57.8	0.19	0.95
0.5	0.43	1.02	1.72	12.5	6.00	0.16	1.00
1.0	0.43	1.06	2.20	19.9	13.0	0.08	1.04

We now introduce the correlation coefficient into the two-point analytical model by rewriting equation (26) as

$$P_{\delta\mathrm{E}}(k) = r(k) f_{\mathrm{IA}} P_{\mathrm{NL}}(k) \,, \tag{C4}$$

for each redshift. Thus the model now includes the simulation-based correlation between the fields. Similarly, we incorporate correlation coefficients into the three-point model given by equations (29) and (30). So, since P_{EE} is unchanged, we have

$$\begin{aligned} B_{\delta\delta\mathrm{E}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = 2 \Big[& f_{\mathrm{IA}}^2 F_2^{\mathrm{eff}}(\boldsymbol{k}_1, \boldsymbol{k}_2) P_{\mathrm{NL}}(k_1) P_{\mathrm{NL}}(k_2) \\ & + r(k_2) f_{\mathrm{IA}} F_2^{\mathrm{eff}}(\boldsymbol{k}_2, \boldsymbol{k}_3) P_{\mathrm{NL}}(k_2) P_{\mathrm{NL}}(k_3) \\ & + r(k_3) f_{\mathrm{IA}} F_2^{\mathrm{eff}}(\boldsymbol{k}_3, \boldsymbol{k}_1) P_{\mathrm{NL}}(k_3) P_{\mathrm{NL}}(k_1) \Big] \,, \end{aligned} \tag{C5}$$

and similarly for $B_{\delta\mathrm{EE}}$.

Fig. C4 compares the unadjusted and adjusted analytical results to the simulation measurements for $P_{\delta\mathrm{E}}$. In this case, by construction, the phenomenological modification virtually eliminates the discrepancy between the simulations and analytical calculations.

Fig. C5 shows similar results for the bispectra $B_{\delta\delta\mathrm{E}}$ and $B_{\delta\mathrm{EE}}$ for equilateral and isosceles triangles. Here, the benefits of introducing the correlation coefficient are less clear, especially for $B_{\delta\mathrm{EE}}$ where neither version of the analytical model fits the simulations very well, especially at higher redshifts. Nevertheless, we conclude that it may be worth considering this decorrelation between the density field and the intrinsic shape in future two-point and three-point IA models.

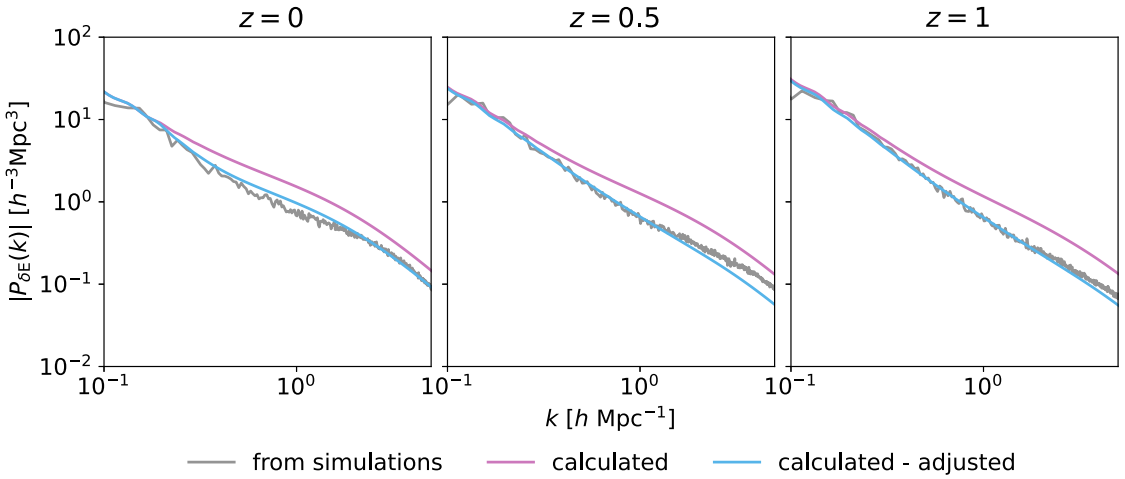

Figure C4. IA power spectrum $P_{\delta\mathrm{E}}$ measured from simulations and calculated from the NLA model with and without adjustment by the correlation coefficient $r(k)$ given by equation (C1).

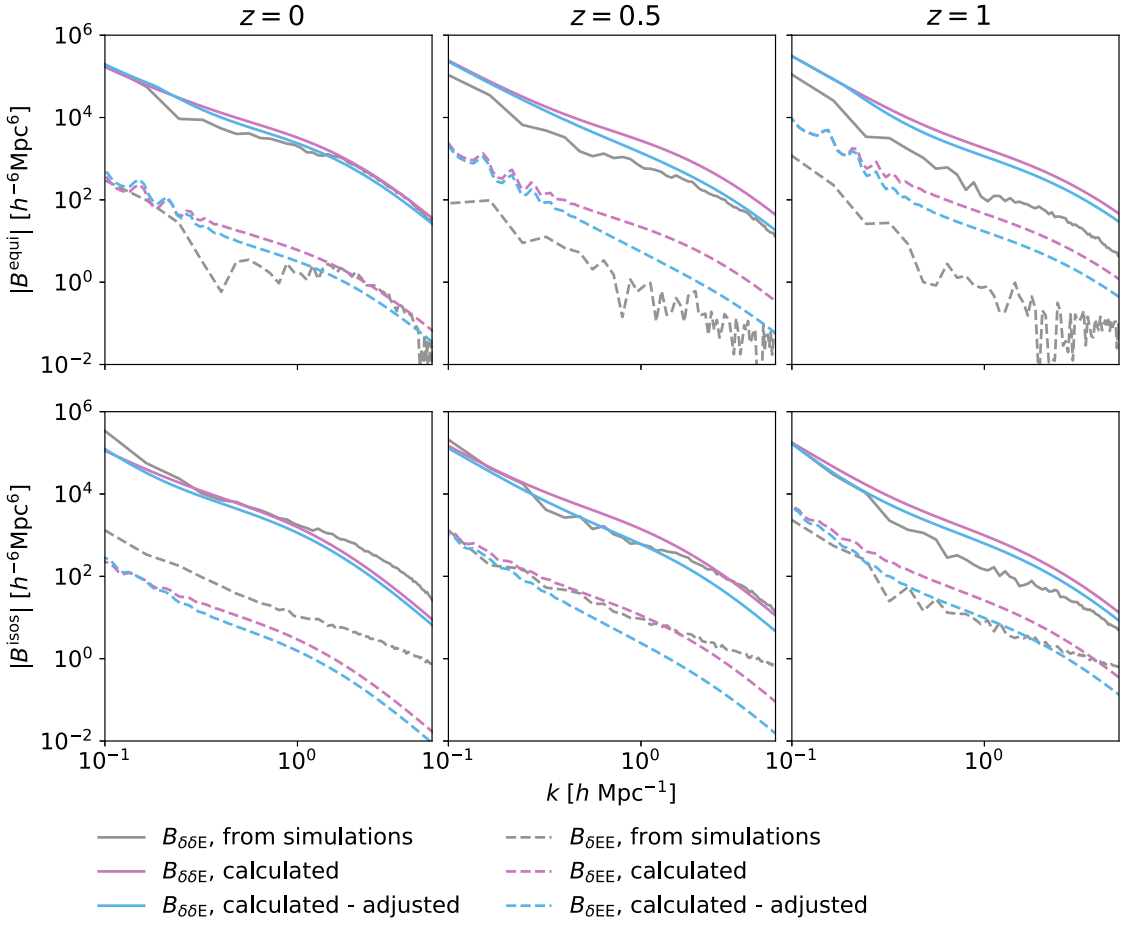

Figure C5. IA bispectra $B_{\delta\delta\rm E}$ and $B_{\delta\rm EE}$ estimated from simulations and calculated analytically with and without adjustment by the correlation coefficient $r(k)$ given by equation (C1). *Top*: Equilateral triangles. *Bottom*: Isosceles triangles with sides in ratio 2:2:1.

This paper has been typeset from a TEX/LATEX file prepared by the author.