

Euclid preparation

Calibrated intrinsic galaxy alignments in the Euclid Flagship simulation

Euclid Collaboration: K. Hoffmann^{*1}, R. Paviot², B. Joachimi³, N. Tessaore³, P. Tallada-Crespi^{4,5}, N. E. Chisari^{6,7}, E. J. Gonzalez^{8,5,9}, A. Loureiro^{10,11}, P. Fosalba^{12,1}, J. Blazek¹³, C. Laigle¹⁴, Y. Dubois¹⁴, C. Pichon^{14,15}, B. Altieri¹⁶, S. Andreon¹⁷, N. Auricchio¹⁸, C. Baccigalupi^{19,20,21,22}, M. Baldi^{23,18,24}, S. Bardelli¹⁸, F. Bernardeau^{25,14}, A. Biviano^{20,19}, E. Branchini^{26,27,17}, M. Brescia^{28,29}, S. Camera^{30,31,32}, G. Cañas-Herrera^{33,34,7}, V. Capobianco³², C. Carbone³⁵, V. F. Cardone^{36,37}, J. Carretero^{4,5}, S. Casas³⁸, F. J. Castander^{1,12}, M. Castellano³⁶, G. Castignani¹⁸, S. Cavuoti^{29,39}, K. C. Chambers⁴⁰, A. Cimatti⁴¹, C. Colodro-Conde⁴², G. Congedo⁴³, L. Conversi^{44,16}, Y. Copin⁴⁵, F. Courbin^{46,47}, H. M. Courtois⁴⁸, A. Da Silva^{49,50}, H. Degaudenzi⁵¹, G. De Lucia²⁰, H. Dole⁵², F. Dubath⁵¹, C. A. J. Duncan^{43,53}, X. Dupac¹⁶, S. Dusini⁵⁴, S. Escoffier⁵⁵, M. Farina⁵⁶, R. Farinelli¹⁸, S. Farrens², S. Ferriol⁴⁵, F. Finelli^{18,57}, N. Fourmanoit⁵⁵, M. Frailis²⁰, E. Franceschi¹⁸, M. Fumana³⁵, S. Galeotta²⁰, K. George⁵⁸, B. Gillis⁴³, C. Giocoli^{18,24}, J. Gracia-Carpio⁵⁹, A. Grazian⁶⁰, F. Grupp^{59,61}, S. V. H. Haugan⁶², H. Hoekstra⁷, W. Holmes⁶³, F. Hormuth⁶⁴, A. Hornstrup^{65,66}, K. Jahnke⁶⁷, M. Jhabvala⁶⁸, E. Keihänen⁶⁹, S. Kermiche⁵⁵, A. Kiessling⁶³, M. Kilbinger², B. Kubik⁴⁵, M. Kümmel⁶¹, M. Kunz⁷⁰, H. Kurki-Suonio^{71,72}, A. M. C. Le Brun⁷³, S. Ligori³², P. B. Lilje⁶², V. Lindholm^{71,72}, I. Lloro⁷⁴, G. Mainetti⁷⁵, D. Maino^{76,35,77}, E. Maiorano¹⁸, O. Marcin⁷⁸, O. Marggraf⁷⁹, M. Martinelli^{36,37}, N. Martinet⁸⁰, F. Marulli^{81,18,24}, R. J. Massey⁸², E. Medinaceli¹⁸, S. Mei^{83,84}, Y. Mellier^{85,14}, M. Meneghetti^{18,24}, E. Merlin³⁶, G. Meylan⁸⁶, A. Mora⁸⁷, M. Moresco^{81,18}, L. Moscardini^{81,18,24}, C. Neissner^{8,5}, S.-M. Niemi³³, C. Padilla⁸, S. Paltani⁵¹, F. Pasian²⁰, K. Pedersen⁸⁸, V. Pettorino³³, S. Pires², G. Polenta⁸⁹, M. Poncet⁹⁰, L. A. Popa⁹¹, L. Pozzetti¹⁸, F. Raison⁵⁹, R. Rebolo^{42,92,93}, A. Renzi^{94,54}, J. Rhodes⁶³, G. Riccio²⁹, E. Romelli²⁰, M. Roncarelli¹⁸, R. Saglia^{61,59}, Z. Sakr^{95,96,97}, A. G. Sánchez⁵⁹, D. Sapone⁹⁸, B. Sartoris^{61,20}, P. Schneider⁷⁹, T. Schrabback⁹⁹, A. Secroun⁵⁵, E. Sefusatti^{20,19,21}, G. Seidel⁶⁷, S. Serrano^{12,100,1}, P. Simon⁷⁹, C. Sirignano^{94,54}, G. Sirri²⁴, A. Spurio Mancini¹⁰¹, L. Stanco⁵⁴, J. Steinwagner⁵⁹, A. N. Taylor⁴³, I. Tereno^{49,102}, S. Toft^{103,104}, R. Toledo-Moreo¹⁰⁵, F. Torradeflat^{5,4}, I. Tutusaus^{1,12,96}, L. Valenziano^{18,57}, J. Valiviita^{71,72}, T. Vassallo²⁰, A. Veropalumbo^{17,27,26}, Y. Wang¹⁰⁶, J. Weller^{61,59}, G. Zamorani¹⁸, F. M. Zerbi¹⁷, E. Zucca¹⁸, M. Ballardini^{107,108,18}, E. Bozzo⁵¹, C. Burigana^{109,57}, R. Cabanac⁹⁶, M. Calabrese^{110,35}, A. Cappi^{18,111}, D. Di Ferdinando²⁴, J. A. Escartin Vigo⁵⁹, L. Gabarra¹¹², W. G. Hartley⁵¹, S. Matthew⁴³, M. Maturi^{95,113}, N. Mauri^{41,24}, R. B. Metcalf^{81,18}, A. Pezzotta¹⁷, M. Pöntinen⁷¹, C. Porciani⁷⁹, I. Rizzo^{17,27}, V. Scottez^{85,114}, M. Sereno^{18,24}, M. Tenti²⁴, M. Viel^{19,20,22,21,115}, M. Wiesmann⁶², Y. Akrami^{116,117}, S. Alvi¹⁰⁷, I. T. Andika^{118,119}, S. Anselmi^{54,94,120}, M. Archidiacono^{76,77}, F. Atrio-Barandela¹²¹, D. Bertacca^{94,60,54}, M. Bethermin¹²², A. Blanchard⁹⁶, L. Blot^{123,73}, M. Bonici^{124,35}, S. Borgani^{125,19,20,21,115}, M. L. Brown⁵³, S. Bruton¹²⁶, A. Calabro³⁶, B. Camacho Quevedo^{19,22,20}, F. Caro³⁶, C. S. Carvalho¹⁰², T. Castro^{20,21,19,115}, F. Cogato^{81,18}, S. Conseil⁴⁵, A. R. Cooray¹²⁷, O. Cucciati¹⁸, S. Davini²⁷, G. Desprez¹²⁸, A. Díaz-Sánchez¹²⁹, J. J. Diaz⁴², S. Di Domizio^{26,27}, J. M. Diego¹³⁰, M. Y. Elkhashab^{20,21,125,19}, A. Enia^{23,18}, Y. Fang⁶¹, A. G. Ferrari²⁴, A. Finoguenov⁷¹, A. Fontana³⁶, A. Franco^{131,132,133}, K. Ganga⁸³, J. García-Bellido¹¹⁶, T. Gasparetto³⁶, V. Gautard¹³⁴, R. Gavazzi^{80,14}, E. Gaztanaga^{1,12,135}, F. Giacomini²⁴, F. Gianotti¹⁸, G. Gozaliasl^{136,71}, M. Guidi^{23,18}, C. M. Gutierrez¹³⁷, A. Hall⁴³, S. Hemmati¹³⁸, H. Hildebrandt¹³⁹, J. Hjorth⁸⁸, S. Joudaki⁴, J. J. E. Kajava^{140,141}, Y. Kang⁵¹, V. Kansal^{142,143}, D. Karagiannis^{107,144}, K. Kiiveri⁶⁹, J. Kim¹¹², C. C. Kirkpatrick⁶⁹, S. Kruk¹⁶, J. Le Graet⁵⁵, L. Legrand^{145,146}, M. Lembo¹⁴, F. Lepori¹⁴⁷, G. Leroy^{148,82}, G. F. Lesci^{81,18}, J. Lesgourges³⁸, L. Leuzzi¹⁸, T. I. Liaudat¹⁴⁹, J. Macias-Perez¹⁵⁰, G. Maggio²⁰, M. Magliocchetti⁵⁶, F. Mannucci¹⁵¹, R. Maoli^{152,36}, C. J. A. P. Martins^{153,154}, L. Maurin⁵², M. Miluzio^{16,155}, P. Monaco^{125,20,21,19,115}, A. Montoro^{1,12}, C. Moretti^{20,19,21,22}, G. Morgante¹⁸, S. Nadathur¹³⁵, K. Naidoo^{135,3}, A. Navarro-Alsina⁷⁹, S. Nesseris¹¹⁶, D. Paoletti^{18,57}, F. Passalacqua^{94,54}, K. Paterson⁶⁷, L. Patrizii²⁴, A. Pisani⁵⁵, D. Potter¹⁴⁷, S. Quai^{81,18}, M. Radovich⁶⁰, P.-F. Rocci⁵², G. Rodighiero^{94,60}, S. Sacquegna^{156,132,131}, M. Sahlén¹⁵⁷, D. B. Sanders⁴⁰, E. Sarpa^{22,115,21}, J. Schaye⁷, A. Schneider¹⁴⁷, M. Schultheis¹¹¹, D. Sciotti^{36,37}, E. Sellentin^{158,7}, L. C. Smith¹⁵⁹, J. G. Sorce^{160,52}, K. Tanidis¹¹², C. Tao⁵⁵, G. Testera²⁷, R. Teyssier¹⁶¹, S. Tosi^{26,27,17}, A. Troja^{94,54}, M. Tucci⁵¹, C. Valieri²⁴, A. Venhola¹⁶², D. Vergani¹⁸, F. Vernizzi²⁵, G. Verza¹⁶³, P. Vielzeuf⁵⁵, and N. A. Walton¹⁵⁹

(Affiliations can be found after the references)

Received ???; accepted ???

ABSTRACT

Intrinsic alignments of galaxies are potentially a major contaminant of cosmological analyses of weak gravitational lensing. We construct a semi-analytic model of galaxy ellipticities and alignments in the *Euclid* Flagship simulation to predict this contamination in *Euclid*'s weak lensing observations. Galaxy shapes and orientations are determined by the corresponding properties of the host haloes in the underlying N -body simulation, as well as the relative positions of galaxies within their halo. Alignment strengths are moderated via stochastic misalignments, separately for central and satellite galaxies and conditional on the galaxy's redshift, luminosity, and rest-frame colour. The resulting model is calibrated against galaxy ellipticity statistics from the COSMOS Survey, selected alignment measurements based on Sloan Digital Sky Survey samples, and galaxy orientations extracted from the Horizon-AGN hydrodynamic simulation at redshift $z = 1$. The best-fit model has a total of 12 alignment parameters and generally reproduces the calibration data sets well within the 1σ statistical uncertainties of the observations and the Flagship simulation, with notable exceptions for the most luminous sub-samples on small physical scales. The statistical power of the calibration data and the volume of the single Flagship realisation are still too small to provide informative prior ranges for intrinsic alignment amplitudes in relevant galaxy samples. As a first application, we predict that *Euclid* end-of-mission tomographic weak gravitational lensing two-point statistics are modified by up to order 10% due to intrinsic alignments.

Key words. Cosmology – large-scale structure of Universe – Gravitational lensing: weak – Galaxies: statistics – Methods: numerical

1. Introduction

The intrinsic alignments of galaxy ellipticities (IA) constitute a potentially limiting factor in the cosmological interpretation of weak gravitational lensing by the large-scale structure (Kirk et al. 2015; Krause et al. 2016; Troxel & Ishak 2015). Alongside baryonic feedback on the dark matter distribution (van Daalen et al. 2011; Sembolini et al. 2011), IA are determined by astrophysical processes in and around galaxies that need to be understood to enable inference on cosmological parameters from weak gravitational lensing surveys. As opposed to baryonic feedback, however, IA cannot be readily isolated from the cosmological signal by selection on the dataset, e.g. through small-scale cuts.

Model-independent removal of intrinsic alignment (IA) is possible by marginalizing over IA model parameters due to its distinctive scaling with the redshifts of the source galaxy samples (e.g., Wright et al. 2025). However, this comes at a significant cost to the cosmological constraining power, for example, reducing the figure of merit of dynamic dark energy constraints by an order of magnitude (Joachimi & Schneider 2009). Self-calibration algorithms, which rely on symmetries of the signal in redshift space, are being explored as an alternative path to IA removal, but are challenging to apply on photometric redshift samples as those used in *Euclid* (e.g., Yao et al. 2020; Chisari 2025).

Therefore, current surveys rely on physically motivated models to account for IA contamination (Asgari et al. 2021; Secco et al. 2022; Dalal et al. 2023). The vast majority of models rely on the assumption that galaxy ellipticities and their orientations are determined by the local tidal field of the surrounding dark matter distribution (Catelan et al. 2001; Hirata & Seljak 2004; Blazek et al. 2019; Vlah et al. 2020; Bakx et al. 2023; see Joachimi et al. 2015 for a review and discussion of alternative mechanisms and Lamman et al. 2024 for a practical introduction). Simple tidal field-based models have shown some success at explaining the scale and mass dependencies of IA signals for certain types of galaxies (Joachimi et al. 2011; Piras et al. 2018; Johnston et al. 2019; Fortuna et al. 2021; Samuroff et al. 2019), but observational constraints remain scarce for the typical galaxy types, masses, and redshift ranges targeted by modern galaxy imaging surveys. Depending on the properties and statistical power of each survey, a careful balance has to be struck between conservative modelling of IA which minimises residual

bias on cosmology and including robust prior knowledge on IA to preserve a maximum of cosmological information.

This work is the first of a series in which we explore this balance in the context of the ESA *Euclid* mission, which features weak gravitational lensing (WL) as one of its primary cosmological probes (Euclid Collaboration: Mellier et al. 2025)¹. Here, we present a semi-analytic model of galaxy ellipticities and alignments for the *Euclid* Flagship Simulation (Euclid Collaboration: Castander et al. 2025, Flagship hereafter). We introduce a semi-analytic model for galaxy ellipticities and alignments based on the *Euclid* Flagship Simulation (Euclid Collaboration: Castander et al. 2025, Flagship hereafter). Unlike simple tidal field-based models, this semi-analytic approach utilizes the high-resolution matter field as well as detailed galaxy photometry provided in Flagship to attribute IA characteristics to each galaxy in the simulation. It thereby matches multiple IA constraints simultaneously, enabling more physically grounded predictions for *Euclid*-like samples. Using this model, we build mock datasets that will be used in companion papers to analyse the range of validity and the performance of analytic IA models to inform analysis choices particularly for *Euclid*'s first data release, which will be based on one year of observations.

Our semi-analytic model follows a similar approach to previous work on different simulations (Joachimi et al. 2013a,b; Hoffmann et al. 2022b, H22 hereafter), but advances beyond these in three aspects: 1. we implement a flexible model for controlling the colour, luminosity, and redshift dependencies of the IA signal; 2. the model is calibrated against a selection of multiple constraints from observations, as well as from a hydro-dynamic simulation where observational data are not available; and 3. the resulting mock catalogue is unprecedented in its volume and depth, covering one octant of the sky up to redshift $z = 3$, and containing 3.4 billion galaxies with magnitudes down to $H_E = 26$, thus allowing for the construction of realistic *Euclid*-like galaxy samples.

This paper is organised as follows: in Sect. 2 we introduce the Flagship simulation and relevant data products for intrinsic alignment calibration, followed by details on the galaxy samples used for calibrating the IA model in Sect. 3. Section 4 provides an overview of the summary statistics employed in the calibration process. In Sects. 5 and 6, the modelling of the mock galaxy shapes and orientations are described, respectively. The calibra-

¹[www.esa.int/Science_Exploration/Space_Science/
Euclid](http://www.esa.int/Science_Exploration/Space_Science/Euclid)

* e-mail: kai.d.hoffmann@gmail.com

tion of the orientations are described in Sect. 7. In Sect. 8, we investigate the statistical uncertainty of the semi-analytic model before computing the expected IA and WL signals for a *Euclid*-like survey in Sect. 9. We summarise our findings and conclude in Sect. 10.

2. Flagship simulation

2.1. Dark matter simulation

The Flagship simulation is one of the largest cosmological N -body simulations of the Universe produced so far, with (16 000)³ particles over a simulation box of $3600 h^{-1}$ Mpc, leading to a mass resolution of $10^9 h^{-1} M_\odot$. The simulation was run using PKDGRAV3 (Potter et al. 2017) with cosmological parameters similar to those of the *Planck* 2015 cosmology (Planck Collaboration et al. 2016). These are the density parameters of cold dark matter, $\Omega_{\text{CDM}} = 0.270$, of baryons, $\Omega_b = 0.049$, and of dark energy, $\Omega_\Lambda = 0.681 - \Omega_r - \Omega_\nu$, where Ω_r and Ω_ν are the density parameters of radiation and massive neutrinos, respectively (where the neutrino normal mass hierarchy is assumed). The other cosmological parameters are the equation of state parameter of dark energy, $w = -1$, the reduced Hubble constant, $h = 0.67$, the (linear) matter density fluctuation amplitude on scales of $8 h^{-1}$ Mpc, $\sigma_8 = 0.83$, and the slope of the primordial power spectrum of scalar density fluctuations, $n_s = 0.96$. A light-cone up to $z = 3$ was produced on the fly during the simulation, covering one octant of the sky, around 5157 deg^2 , and centred at approximately the North Galactic Pole.

Weak lensing maps All-sky lensing convergence maps were constructed by decomposing and rotating the dark matter light-cone into a set of all-sky concentric spherical shells around the observer. Lensing observables were then derived from the two-dimensional dark matter density maps of the different shells, using the Born approximation (Fosalba et al. 2008, 2015). The shells have a width of $\Delta_r \approx 95.22$ Myr in lookback time, and an angular resolution of $\Delta_\theta \approx 0.43$. Further details on the construction of the lensing maps are given in (Fosalba et al. 2015).

By combining the dark-matter “onion shells” that make up the light-cone, we can easily derive lensing observables, as explained in Fosalba et al. (2008). This approach, based on approximating the observables by a discrete sum of two-dimensional dark-matter density maps multiplied by the appropriate lensing weights, agrees with the much more complex and CPU time-consuming ray-tracing technique in the Born approximation limit, i.e. in the limit where lensing deflections are calculated using unperturbed light paths (see e.g., Hilbert et al. 2020). Following this technique, one can produce all-sky maps of the convergence field (which is simply related to the lensing potential in harmonic space), as well as maps for other lensing fields obtained from covariant derivatives of the lensing potential, such as the deflection angle, shear, flexion, etc. In particular, using the spherical transform of the lensing potential all-sky map one can obtain the corresponding maps for other lensing observables through simple relations (see Hu 2000). These weak lensing maps allow us to predict the relative contribution of IA to the observed shear statistics.

2.2. Halo detection and halo shape measurement

Dark matter halos and sub-halos were identified in the particle distribution in phase space using ROCKSTAR (Behroozi et al.

2013), which provides positions, velocities, as well as three-dimensional shapes and orientations for each object. The shapes and orientations are defined via the eigenvectors and eigenvalues of the moment of inertia,

$$I_{i,j} = \sum_n^{N_p} \frac{r_{n,i} r_{n,j}}{R_n^2}, \quad (1)$$

where N_p is the number of halo particles and $r_{n,i}$ are the components of the three-dimensional position vector of the n th particle with respect to the halo centre (e.g. Dubinski & Carlberg 1991; Allgood et al. 2006). R_n is a normalisation factor, which is set to $R_n = 1$ in the standard definition of the moment of inertia and to $R_n = (r_{n,1}^2 + r_{n,2}^2 + r_{n,3}^2)^{1/2}$ for the so-called reduced moment of inertia. The lengths of the halo’s principal axes are given by the square root of the eigenvalues λ of $I_{i,j}$, $(A_{3D}, B_{3D}, C_{3D}) = (\sqrt{\lambda_A}, \sqrt{\lambda_B}, \sqrt{\lambda_C})$, and satisfy $A_{3D} \geq B_{3D} \geq C_{3D}$. The halo shapes are characterized by two of the three axis ratios

$$q_{3D} \equiv \frac{B_{3D}}{A_{3D}}, \quad r_{3D} \equiv \frac{C_{3D}}{B_{3D}}, \quad s_{3D} \equiv \frac{C_{3D}}{A_{3D}}, \quad (2)$$

while the halo orientations are given by the normalized eigenvectors $\hat{\mathbf{A}}_{3D}$, $\hat{\mathbf{B}}_{3D}$, and $\hat{\mathbf{C}}_{3D}$. In this study, we use halo orientations measured via the reduced moment of inertia, as it assigns more weight to the mass distribution in the halo’s centre, where the central galaxies - whose orientations we aim to model - are located. Measurements of galaxy orientations from the stellar particle distribution in the hydrodynamical simulation (described in Sect. 3.4), on the other hand, are based on the standard moment of inertia, as it provides a higher signal-to-noise ratio (Chisari et al. 2015).

Orientations measured from discrete particle distributions are subject to sampling noise, which increases for decreasing numbers of particles. In the galaxy and halo samples used in this work, the average number of particles per object ranges from about 1000 to 18000, for which we expect only moderate randomization of below 5 degree (Herle et al. 2025). However, the minimum number of particles per object can be as low as 10, in which case a strong randomization of the orientation is expected. H22 demonstrated that randomly down-sampling halos to 10 particles reduces the alignment signal by roughly 30 percent, while increasing the particle limit to 100 lowers the impact of noise on the alignment signal to the percent level. For this reason, we do not expect noise in the measured orientations to have a dominant effect on the predicted IA statistics.

2.3. Galaxy modelling

Based on the dark matter halos, a galaxy catalogue was built using the SciPIC Scientific pipeline at Port d’Informació Científica (Carretero et al. 2017a). SciPIC combines the halo occupation distribution technique with halo abundance matching (Carretero et al. 2015, Euclid Collaboration: Castander et al. 2025) to distribute central and satellite galaxy positions in the halos and then assign velocities as well as initial magnitudes and colours. The parameters of the model were calibrated such that the simulation reproduces Sloan Digital Sky Survey (York et al. 2000, SDSS) observations of the luminosity function, the colour-magnitude distribution and the clustering in redshift space as a function of colour and luminosity (Blanton et al. 2003; Zehavi et al. 2011). In a subsequent step, spectral energy distributions (SEDs) are derived for each galaxy based on its initial magnitude and colour. From these SEDs, additional magnitudes can

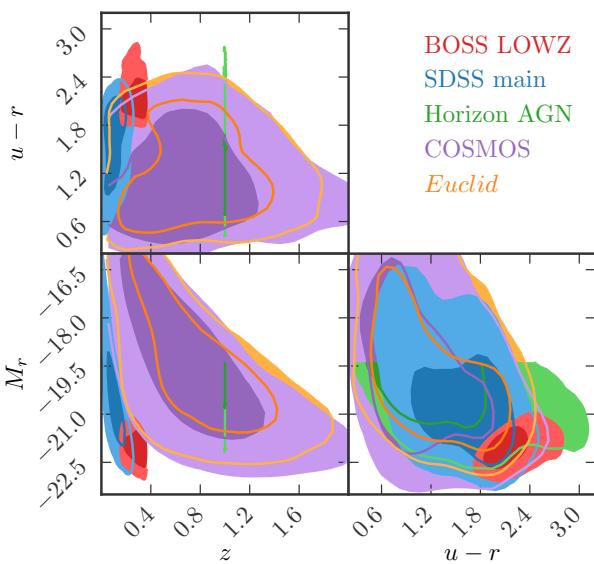


Fig. 1. Redshift, magnitude, colour distributions of mock samples, used for calibrating the IA model parameters for galaxy misalignment. Contours enclose 68% and 95% of the distributions.

then be derived for any filter, which enables the construction of mock galaxy catalogues of any desired galaxy survey.

3. Mock Galaxy Catalogs

We construct mock galaxy catalogues to calibrate the IA simulation against measurements from reference samples. Specifically, we use a mock catalogue of the COSMOS (Scoville et al. 2007) survey to adjust the definition of red and blue galaxies, used in the shape component of our IA model and to validate the shape distribution predicted by that model. In addition, we use mock catalogues of the SDSS main sample and the BOSS LOWZ sample for calibrating the galaxy orientation model component at low redshifts. These low-redshift samples are complemented by a mock catalogue of the Horizon-AGN simulation (Dubois et al. 2014; Kaviraj et al. 2017), which we use for the calibration of the orientations at $z = 1$. In Fig. 1 we compare the colour-magnitude-redshift distribution of these different mock catalogues, while the main characteristics of each mock sample used in this work are summarised in Table 1. In the following, we give an overview of how the reference samples are selected and detail the construction of the corresponding mocks.

3.1. COSMOS

The COSMOS survey provides space-based imaging from the Hubble Space Telescope (HST) together with ground-based photometry in various bands which allows for highly accurate measurements of galaxy shapes and redshifts. Primarily due to its relatively small area of roughly one square degree, two-point IA statistics have not been measured in COSMOS to date. However, it is large enough to provide constraints on galaxy shape distributions over a wide range of galaxy magnitudes, colours, and redshifts (e.g., Joachimi et al. 2013a). The reference sample of the COSMOS survey used in this work has been composed from the COSMOS2015 catalogue² (Laigle et al. 2016

Table 1. Mean redshift z , mean absolute r -band magnitude M_r , mean rest-frame colour index $u - r$ and satellite fraction f_{sat} of the mock samples used to calibrate the Flagship simulation.

sample		N_{gal}	$\langle z \rangle$	$\langle M_r \rangle$	$\langle u - r \rangle$	f_{sat}
SDSS	main	495502	0.10	-20.50	1.59	0.35
	blue	198674	0.09	-19.96	1.16	0.25
	red	296828	0.11	-20.87	1.87	0.43
BOSS	main	155290	0.28	-22.06	2.17	0.25
	L1	31058	0.29	-22.60	2.08	0.22
	L2	31058	0.28	-22.25	2.11	0.23
	L3	31058	0.28	-22.05	2.15	0.24
	L4	62116	0.26	-21.69	2.26	0.28
HAGN	main	757638	1.00	-20.41	1.76	0.35
	H1	33311	1.00	-22.49	2.41	0.20
	H2	183915	1.00	-21.35	2.08	0.31
	H3	540412	1.00	-19.96	1.62	0.38

and the Advanced Camera for Surveys General Catalog (ACS-GC³, Griffith et al. 2012). The COSMOS15 catalogue provides photometry in 30 bands together with photometric redshift estimates. The ACS-GC catalogue provides measurements of galaxy shapes from imaging in the HST ACS F814W i -band which are corrected for effects of the HST point spread function. For our analysis, we select objects with apparent magnitudes brighter than $i = 24$ to ensure that the redshifts and shape measurements are reliable (Griffith et al. 2012; Laigle et al. 2016). The corresponding mock sample is constructed from Flagship by applying the same cut. Based on this mock, we expect this matched COSMOS catalogue to overlap almost entirely with the *Euclid* sample in the colour-magnitude-redshift space, as we show in Fig. 1. A more detailed summary of the COSMOS15 and the ACS-GC catalogues, as well as the procedure for matching these catalogues, can be found in Hoffmann et al. (2022a).

3.1.1. Colour-selected sub-samples

To introduce a colour dependence of galaxy shapes in our model, we assign a colour type to each object, classifying it as either red or blue. Following H22, we define this colour type by a cut in the $u - r := M_u - M_r$ colour, where M_u and M_r are absolute rest-frame magnitudes in the CFHT u -band and the Subaru r -band, respectively. We compare the colour distribution in Flagship and COSMOS catalogues in the top panel of Fig. 2, for *Euclid*-like samples of galaxies with $i < 24$ and $z < 2.0$ and find a good agreement between both datasets.

For a more detailed validation of Flagship, we display the joint distribution of the $u - r$ colour and the r -band magnitude in three redshift bins in the central and bottom panels of Fig. 2. We find an overall good agreement between both datasets, with both showing similar trends of increasing brightness and blueness with redshift as well as comparable shapes of the distributions. However, our comparison also reveals a shift of the distribution in Flagship towards fainter magnitudes at all redshifts compared to COSMOS, while the colours in Flagship are slightly bluer in the lowest redshift bin ($0.1 < z < 0.3$) and significantly redder in the highest redshift bin ($1.9 < z < 2.1$).

For the colour type definition in COSMOS, we adopt the cut from H22 at $u - r = 1.2$, which leads to a global fraction of blue galaxies of $f_{\text{blue}} = 0.72$ for galaxies with $i < 24$. We adjust this cut in Flagship to $u - r = 1.32$, such that f_{blue} has the same value

²<https://www.eso.org/qi>

³vizier.u-strasbg.fr/viz-bin/VizieR-3?-source=J/ApJS/200/9/acs-gc

as in COSMOS. In Fig. 2 we show that these simple cuts provide a reasonable separation between the red and blue sequences in colour-magnitude space in COSMOS as well as in Flagship within the redshift range probed by *Euclid*. A more quantitative validation of the colour cut is shown in Fig. 3, where we compare f_{blue} in Flagship and COSMOS in different redshift bins, showing that Flagship replicates f_{blue} from COSMOS typically to within 10 percent or less.

3.2. SDSS Main sample

The SDSS Main sample⁴ consists of galaxies with an apparent SDSS r -band magnitude brighter than $r_{\text{lim}} = 17.77$. Spectroscopic redshifts from this sample as well as estimates of the absolute rest-frame magnitudes are publicly available in the New York University Value-Added Galaxy Catalog of the SDSS Data Release 7 (Blanton et al. 2005, NYU-VAGC)⁵, which we use in this work as an observational reference when constructing the mock sample. This catalogue consists of 741 874 galaxies and covers 7966 square degree, which results in a galaxy number density of 93 deg^{-2} .

When constructing the mock SDSS main sample from Flagship, we adjust the limiting magnitude by adding the constant $\Delta r = -0.15$ (i.e. $r_{\text{lim}} + \Delta r$), which is chosen such that the mock catalogue has the same number density as the reference sample for objects with redshifts above $z = 0.15$. Lower redshifts have been excluded for this adjustment to reduce the impact of sampling variance when comparing the number densities.

In the top right panel of Fig. 4 we compare the differential redshift distributions of the galaxy number density in NYU-VAGC to the Flagship mock sample. We find a reasonable agreement between both datasets for $z > 0.15$ with absolute deviations around 15%. At lower redshifts both distributions appear to be similar but are strongly affected by fluctuations, which we attribute to sampling variance. The top left panel of Fig. 4 shows a comparison between the differential galaxy number densities in the NYU-VAGC and Flagship as a function of the apparent r -band magnitude for objects above $z = 0.15$. We find the differential number density to be higher in the Flagship than in NYU-VAGC, which explains the necessity of reducing the limiting magnitude in order to obtain the observed cumulative number density.

3.2.1. Colour-selected sub-samples

In order to calibrate our IA model against the IA statistics measured in the red and blue sub-samples of the SDSS Main sample used by Johnston et al. (2019, hereafter referred to as J19), we adopt the authors' methodology by splitting our mock SDSS main sample with a cut on the $g - r := M_g - M_r$ rest-frame colour. While J19 applied a cut at $g - r = 0.66$, we again adjust our cut in Flagship to $g - r = 0.61$ in order to obtain the same fraction of blue objects of $f_{\text{blue}} = 0.4$ as J19. In Fig. 5 we compare the distribution of the colours in the NYU-VAGC with the Flagship mock, showing that both datasets agree well with each other.

As an additional validation of the Flagship mock selection, we compare the two-point galaxy clustering statistics w_{gg} (described in Sect. 4) with the observational measurements of the red and blue sub-samples from J19, as shown in Fig. 7. We find good agreement between the two datasets, indicating that

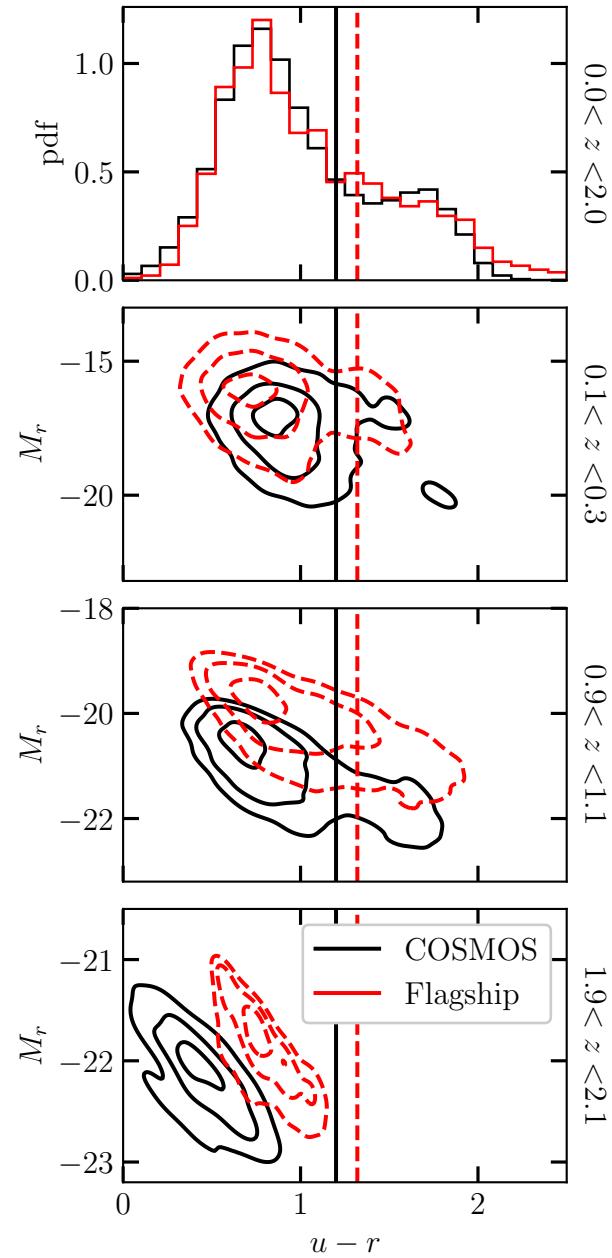


Fig. 2. Top: Distribution of the rest-frame colour $u - r \equiv M_u - M_r$ for galaxies in a *Euclid*-like sample, selected by $i < 24$ from COSMOS (black) and Flagship (red) across the full redshift range covered by *Euclid*. Center to Bottom: Joint distribution of the rest-frame colour and the absolute rest-frame magnitude in the Subaru r -band in three redshift bins. Contours enclose the central 10, 40, and 70 percent of the distributions, while black solid and red dashed lines represent results for COSMOS and Flagship respectively. Vertical lines at $u - r = 1.2$ and $u - r = 1.32$ indicate the colour cuts used to define red and blue subsamples in COSMOS and Flagship, respectively. The cut for Flagship is shifted such that the global fraction of blue galaxies matches that of COSMOS.

our mock selection is suitable for calibrating our IA model; see Sect. 4.2 for a detailed discussion of the clustering comparison.

⁴<https://classic.sdss.org>

⁵<http://sdss.physics.nyu.edu/vagc>

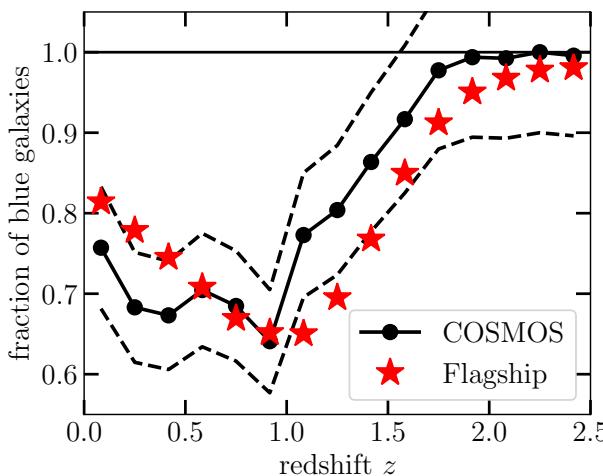


Fig. 3. Fraction of blue galaxies in Flagship and COSMOS, selected by the colour cuts shown in Fig. 2. The dashed lines mark $\pm 10\%$ deviations from the COSMOS data.

3.3. BOSS LOWZ sample

While the SDSS Main sample provides IA constraints over a wide range of absolute magnitudes and colours, it is limited to relatively low redshifts. We therefore use direct IA measurements based on the LOWZ sample of the BOSS survey from [Singh & Mandelbaum \(2016\)](#), hereafter referred to as SM16), which provide constraints at higher redshifts when calibrating the IA model (see Fig. 1 and 4). The LOWZ sample consists of luminous red galaxies that were selected by multiple cuts in the apparent colour-magnitude space,

$$r < 13.5 + \frac{c_{\parallel}}{0.3} + \Delta r ; \quad (3)$$

$$16.0 < r < 19.6 + \Delta r ;$$

$$|c_{\perp}| < 0.2 ,$$

with

$$c_{\parallel} = 0.7(g - r) + 1.2[(r - i) - 0.18] ; \quad (4)$$

$$c_{\perp} = (r - i) - (g - r)/4.0 - 0.18 .$$

In addition to these cuts, we select galaxies within the redshift range $0.16 < z < 0.36$ in which SM16 performed their analysis. As for the SDSS selection, we introduce a parameter Δr , which is zero in the selection of observed data and is adjusted in the simulation to match the observed galaxy number density. For that adjustment, we use as a reference the LOWZ sample from the BOSS Data Release 12, which covers an area of 8337 deg^2 (here we use the effective area from [Reid et al. 2016](#), table 2). It contains 249 938 galaxies within the SM16 redshift range, which leads to a galaxy number density of 29.98 deg^{-2} . In order to achieve the same density in Flagship we set $\Delta r = -0.17$, which leads to 155 179 galaxies in the mock catalogue in the 5157 deg^2 octant of the Flagship simulation. In the bottom right panel of Fig. 4 we compare the differential redshift distributions of the galaxy number density in the LOWZ sample from BOSS and the corresponding Flagship mock sample. We find both data sets to agree reasonably well with absolute deviations around 15% within the SM16 redshift range, given the complexity of the selection cuts and the fact that the galaxy colours and magnitudes in Flagship have not been calibrated against the LOWZ sample.

The bottom left panel of Fig. 4 shows a comparison between the differential galaxy number densities in BOSS and Flagship as a function of the apparent r -band magnitude for objects within the redshift range analysed by SM16. For apparent magnitudes brighter than $m_r < 18$ we find the differential galaxy number density in Flagship to be slightly below the observational data, while it is significantly higher at dimmer magnitudes. This figure illustrates how reducing the limiting magnitude by Δr allows us to match the observed cumulative number density.

3.3.1. Magnitude-selected sub-samples

Following SM16, we split the LOWZ samples into four luminosity sub-samples that are labelled from bright to faint L1 to L4. The samples are selected by cuts in the absolute SDSS r -band magnitude, which are chosen so that each of the L1 to L3 samples contains 20% of all galaxies, while L4 contains the faintest 40%. The distribution of absolute magnitudes is shown for the four sub-samples from the Flagship mock in Fig. 6 while the main characteristics of these sub-samples are given in Table 1.

To further validate the Flagship mock selection, we compare the two-point galaxy clustering statistics w_{gg} (see Sect. 4) with observational measurements of the luminosity sub-samples from SM16, as shown in Fig. 7. We find overall good agreement, although significant deviations appear at small scales for the brightest samples (L1 and L2), which are taken into account when calibrating our IA model. Further details on this clustering comparison are provided in Sect. 4.2.

3.4. Horizon AGN

The observational samples described in the previous subsections provide IA constraints up to redshifts of $z = 0.36$, which is well below the bulk of redshifts probed by *Euclid* (see Fig. 1). Since observations currently do not provide tight constraints at higher redshifts, we use data from the Horizon-AGN (HAGN, [Dubois et al. 2014](#); [Kaviraj et al. 2017](#)) simulation at $z = 1.0$ instead.

HAGN is a hydrodynamic simulation of cosmic structure formation within a $100 h^{-1}\text{Mpc}$ cube. It takes into account key processes in the formation of galaxies: cooling, heating, and the chemical enrichment of gas, the formation and evolution of stars and black holes, as well as feedback from stellar winds, supernovae, and active galactic nuclei as described in [Dubois et al. \(2014\)](#). Galaxies were identified in the simulation as groups of gravitationally bound stellar particles using the AdaptaHOP finder ([Aubert et al. 2004](#)).

The HAGN simulation has been used for extensive studies on IA ([Chisari et al. 2015](#); [Codis et al. 2015](#); [Chisari et al. 2016, 2017](#); [Codis et al. 2018](#); [Soussana et al. 2020](#); [Bate et al. 2020](#)). Galaxy shapes and orientations were derived from the standard moment of inertia of the galaxies' stellar particle distribution, analogously to the measurements of halo shapes and orientations in Flagship from the reduced moment of inertia, as detailed in Sect. 2.

For the calibration of the Flagship IA model against the predictions from HAGN, we select a main reference sample of galaxies from HAGN that consists of objects at $z = 1.0$ with absolute SDSS r -band magnitudes brighter than $M_r = -20$. We further require stellar particle groups to be reliably detected at the AdaptaHOP tree level equal to 1. This selection leads to a minimum of 621 stellar particles per galaxy. Note that the HAGN simulation output is available at multiple redshifts, which could

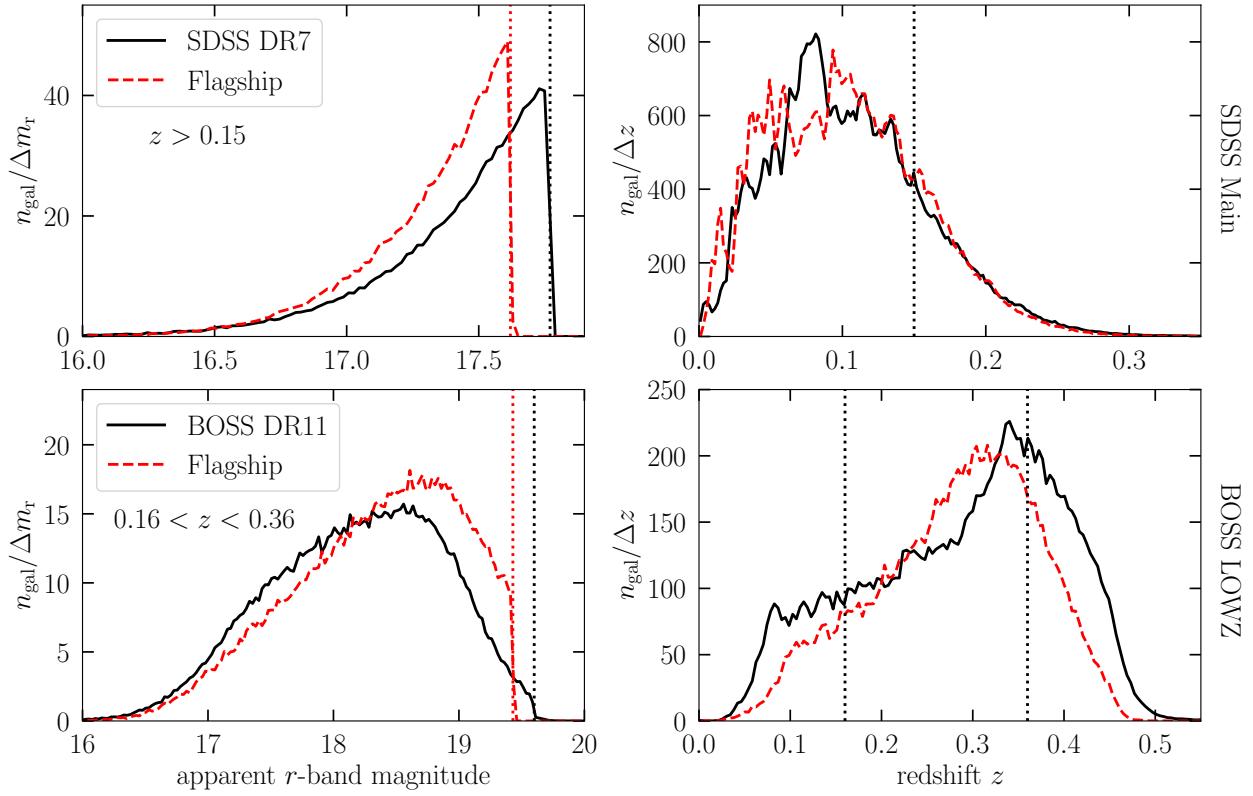


Fig. 4. Distribution of galaxy counts per square degree (n_{gal}) as a function of apparent r -band magnitude (m_r , left) and redshift (z , right). Results are normalized by the bin widths Δm_r and Δz to allow comparisons independent of bin size. Top and bottom panels show results for the SDSS Main and BOSS LOWZ samples, respectively. Black solid and red dashed lines represent observational data and the Flagship mock samples, respectively. Vertical dotted lines indicate the selection cuts in magnitude (left) and redshift (right) applied in our sample selection. The redshift cuts in Flagship coincide with those used for the observed samples.

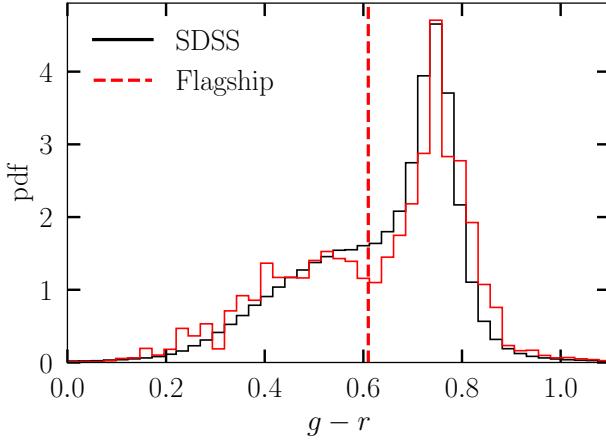


Fig. 5. Distribution of the rest-frame colour $g - r := M_g - M_r$ rest-frame colour in the SDSS Main sample and the corresponding Flagship mock catalogue. The vertical red line shows the cut at $g - r = 0.61$ where we split the mock SDSS catalogue into a red and blue sub-sample. This cut is chosen such that we reproduce the fraction of blue galaxies from J19 of $f_{\text{blue}} = 0.4$.

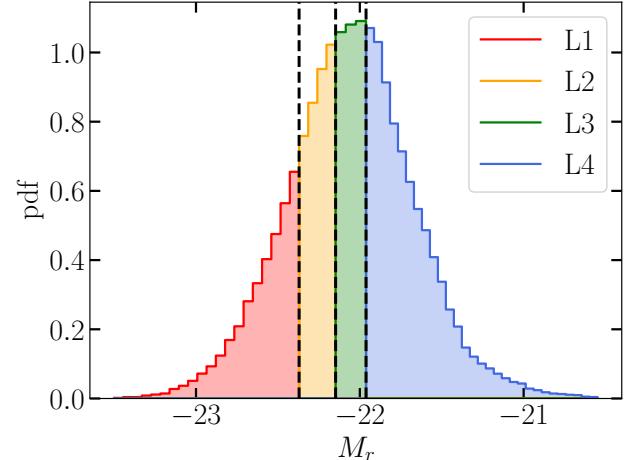


Fig. 6. Distribution of absolute SDSS r -band magnitude of the mock BOSS LOWZ sample constructed from Flagship. Vertical lines indicate the limits of the four luminosity sub-samples L1 to L4.

have been incorporated into the calibration for a more comprehensive coverage of the *Euclid* light-cone. However, due to the limited resources available for this work, we focus on the output at $z = 1.0$, as it is the redshift closest to the mean redshift

of our mock *Euclid* sample when selecting objects at redshifts higher than those probed by the observational reference samples ($z > 0.36$).

We construct a mock HAGN sample from Flagship by first selecting galaxies in a thin shell of the light-cone around $z = 1.0$ with the width of $\Delta z = 0.02$. The comoving volume of this shell

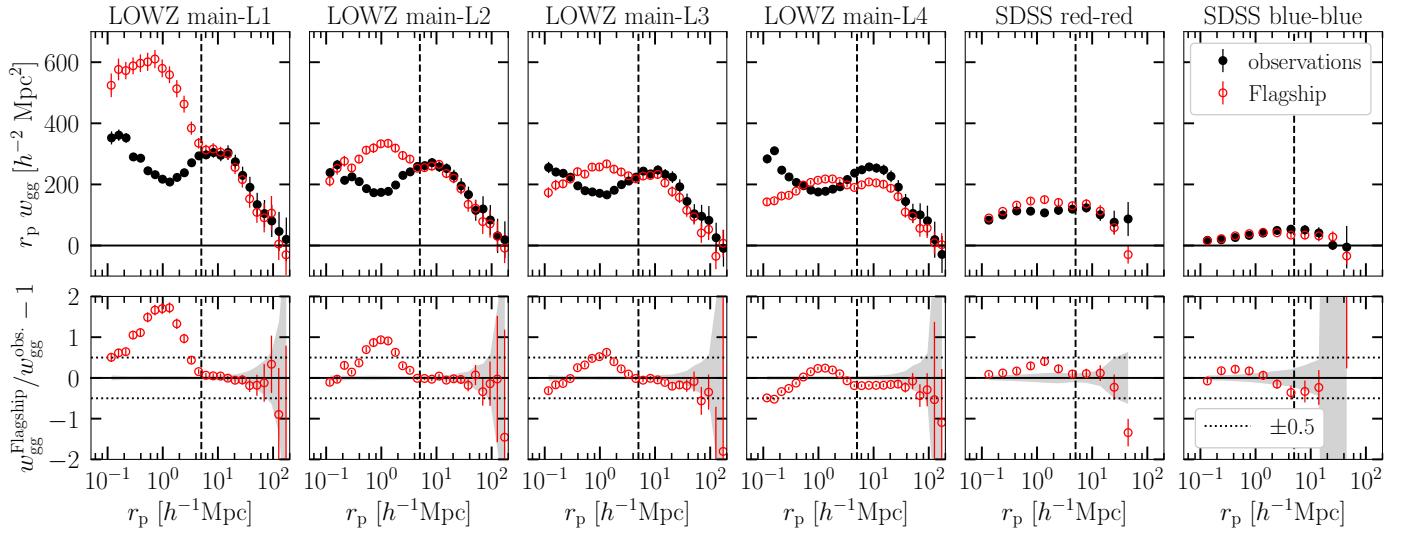


Fig. 7. Top: Projected galaxy-galaxy cross-correlations between different density and shape samples against the projected scale for different observations and corresponding mock samples from the calibrated Flagship simulation (black and red symbols, respectively). The four left panels show results for the BOSS LOWZ sample, where the density sample is the LOWZ main sample, while the shape samples are the magnitude-selected sub-samples L1 to L4 (from bright to faint). The two right panels show results for SDSS, where both, density and shape samples are red and blue sub-samples. Error bars indicate jackknife estimates of the standard deviation. Bottom: Relative deviations of the Flagship results from the observations. Grey areas show the standard deviation of the observational data. Dotted lines mark 50 % deviations. The vertical dashed lines in the top and bottom panels at $5 h^{-1}\text{Mpc}$ indicate the approximate scale below (above) which we expect the alignment signal to be dominated by the alignment of satellites (centrals).

is 33 times larger than the volume covered by HAGN, leading to significantly smaller errors on the IA statistics measured in Flagship compared to HAGN. Galaxies in the Flagship redshift shell are required to be brighter than the absolute SDSS r -band magnitude $M_r = -20 + \Delta M_r$. Here, we have introduced the constant $\Delta M_r = 0.7$ which we adjust to match the comoving galaxy number density of the HAGN main reference sample of $0.023 h^3 \text{Mpc}^{-3}$, analogously to the adjustment of the sample selections of the mock observational samples.

In Fig. 8 we compare the luminosity functions of HAGN and Flagship at $z = 1$ to each other as functions of M_r . At magnitudes brighter than $M_r = -22$, we find a similar decrease of the luminosity functions with magnitude for both simulations. At fainter magnitudes the luminosity function in HAGN is significantly higher than in Flagship, which explains why the limiting magnitude had to be set fainter in Flagship in order to reproduce the cumulative number density of HAGN. Please note here that Flagship as well as HAGN have been calibrated against luminosity functions at low redshifts, while the high-redshift luminosity functions are predictions which may differ due to the fundamental differences in the methods used in these simulations.

In order to probe the luminosity dependence of the IA signal in HAGN we split the main sample into three luminosity sub-samples at $M_r = -21$ and -22 , as illustrated in Fig. 8. The resulting samples, labelled H1 (bright) through H3 (faint) contain 4.4%, 24.3%, and 71.3% of the galaxies, respectively. The corresponding selection cuts in Flagship are adjusted to $M_r = -20.8$ and -22.1 in order to obtain the same fractions of galaxies per sub-sample. The main characteristics of the mock HAGN samples are summarised in Table 1.

4. Clustering and IA statistics

For the calibration of the Flagship IA model, we quantify the IA signal using two types of two-point statistics. First, we use the projected two-point correlation function w_{g+} to compare

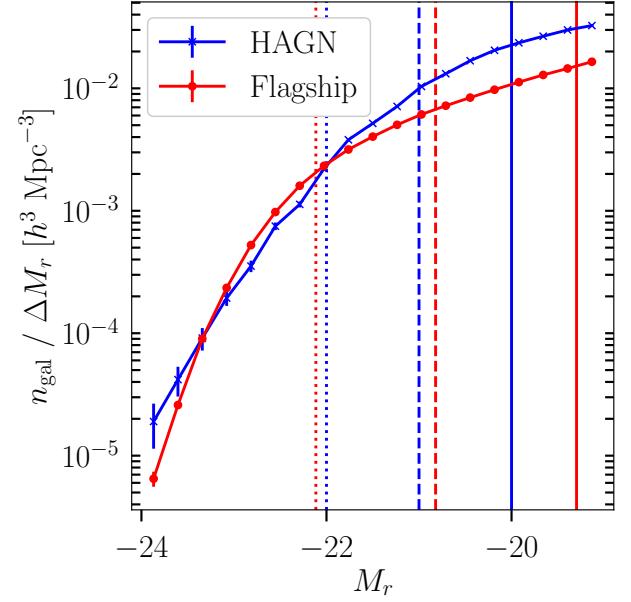


Fig. 8. Comoving galaxy number density, n_{gal} , as a function of absolute r -band magnitude (M_r). Values are normalised by the bin width ΔM_r to remove bin-size dependence. Results are shown at redshift $z = 1.0$ for Flagship and HAGN (red and blue lines, respectively). Error bars enclose the central 68% of a Poisson distribution. Blue vertical lines at $M_r = [-20, -21, -22]$ denote the lower limits of HAGN sub-samples H1 to H3, used as high-redshift references for IA model calibration. The corresponding limits in Flagship (red vertical lines) are adjusted to $M_r \in [-19.3, -20.8, -22.1]$ to match the number densities in HAGN.

the IA signal predicted by the Flagship mock with direct measurements from observational data. For the calibration based on the HAGN simulation, we compute the two-point correlation of

three-dimensional galaxy orientations, η , and compare these directly to the corresponding measurements in HAGN.

4.1. Projected correlation functions w_{gg} and $w_{\text{g+}}$

We measure the projected galaxy-ellipticity correlation to calibrate the IA model against observational constraints derived from the BOSS LOWZ sample by SM16 and the SDSS Main sample by J19. In addition, we measure projected galaxy-galaxy correlations to validate the galaxy clustering in the mock samples constructed from Flagship that are used for the calibration. When measuring these correlations we follow SM16 and J19 by studying the cross-correlation between a shape sample S, consisting of the galaxies whose IA signal we want to measure and a density sample D, which is used as a tracer for the underlying matter distribution.

The galaxy-galaxy cross-correlation function between two samples is defined as $\xi_{\text{gg}}(r_p, \Pi) := \langle \delta_g^S \delta_g^D \rangle(r_p, \Pi)$, where δ_g^S and δ_g^D are the galaxy density contrasts of the shape and density samples, respectively, and $\langle \dots \rangle$ indicates an ensemble average. Measurements are done in bins centred on a transverse comoving separation r_p and a line-of-sight comoving separation Π . We measure this correlation from the data using the estimator from Landy & Szalay (1993),

$$\hat{\xi}_{\text{gg}}(r_p, \Pi) = \frac{\text{SD} - \text{DR}_S - \text{R}_D \text{RS} + \text{R}_D \text{R}_S}{\text{RS} \text{R}_D}. \quad (5)$$

Each term in the numerator and denominator on the right-hand side of this equation stands for the counts of galaxy pairs in bins separated by r_p and Π . The symbols R_S and R_D denote samples of randomly distributed points that are constructed to follow the redshift distributions and sky footprints of the S and D samples, respectively.

Analogously to the galaxy-galaxy correlation one can define the galaxy-ellipticity correlation function as $\xi_{\text{g+},x}(r_p, \Pi) = \langle \delta_g^D \epsilon_{+,x} \rangle(r_p, \Pi)$. The ellipticity is defined for each position-ellipticity galaxy pair in the average $\langle \dots \rangle$ such that the orientation is the angle between the galaxies' major axis and the separation vector on the sky. An estimator for this correlation function can be defined as (Mandelbaum et al. 2006)

$$\hat{\xi}_{\text{g+},x}(r_p, \Pi) = \frac{\text{S}_{+,x}\text{D} - \text{S}_{+,x}\text{R}_D}{\text{R}_S \text{R}_D}, \quad (6)$$

with

$$\text{S}_{+,x}\text{X} = \sum_{i,j} \epsilon_{+,x}(i|j), \quad (7)$$

where $\epsilon_{+}(i|j) = \Re[\epsilon \exp(-2i\phi_{ij})]$ is the (+)-component of the ellipticity of a galaxy i in sample S, defined with respect to the separation vector pointing to position j in sample X. ϕ_{ij} is the position angle of this separation vector at the position i and X refers to either D or R_D. The (\times)-component is defined analogously using the imaginary part of the expression above. Since ξ_{gx} changes sign under parity transformation, it does not contain a physical signal, and we therefore do not consider it further in this study.

For the measurements in the mock BOSS LOWZ samples we follow SM16 by using 25 logarithmic bins in the interval $0.1 < |r_p|/(h^{-1}\text{Mpc}) < 200$ and 20 linear bins in the range $|\Pi| < \Pi_{\text{max}}$. Similarly, for the measurements in the mock SDSS Main samples we follow J19 by using 11 logarithmic bins in the interval $0.1 < |r_p|/(h^{-1}\text{Mpc}) < 60$ and 30 linear bins within

$|\Pi| < \Pi_{\text{max}}$. In both cases $\Pi_{\text{max}} = 60 h^{-1}\text{Mpc}$ was chosen. The projected correlation is then given by

$$w_{\text{gx}}(r_p) = \int_{-\Pi_{\text{max}}}^{\Pi_{\text{max}}} \xi_{\text{gx}}(r_p, \Pi) d\Pi, \quad (8)$$

where w_{gx} stands for w_{gg} or $w_{\text{g+}}$. In the measurement the integral above is approximated by a Riemann sum over the line-of-sight bins.

Errors on the measurements of w_{gg} and $w_{\text{g+}}$ are estimated using jackknife resampling. The Flagship octant is therefore split into $N_{\text{JK}} = 88$ HEALPix regions. The covariance is then estimated as

$$C_{ij}^{\text{JK}} = (N_{\text{JK}} - 1)(\Delta_i \Delta_j), \quad (9)$$

with $\Delta_i = w_i^{\text{JK}} - w_i$, where w_i is the projected correlation measured in the r_p bin i on the full sky area, w_i^{JK} is the same measurement, but dropping one jackknife sub-region at a time and $\langle \dots \rangle$ is the average over the N_{JK} measurements of Δ_i . Further details on the jackknife sampling, the construction of random catalogues and the code for performing the measurements are given in H22.

4.2. Galaxy clustering: Flagship versus observations

In Fig. 7 we compare the measurements of w_{gg} in the SDSS Main and BOSS LOWZ samples, derived from observed data by J19 and SM16, respectively, to the corresponding measurements from Flagship as an additional validation of the mock samples described in Sect. 3. In the case of LOWZ w_{gg} refers to the cross-correlations between the full LOWZ sample and the four luminosity sub-samples, while for SDSS w_{gg} is the auto-correlation of the red and blue sub-samples, respectively.

We find the deviations between Flagship and observations at scales above $5 h^{-1}\text{Mpc}$ to be consistent with the 1σ jackknife errors. However, at smaller scales we find significant deviations, which are highest at roughly $1 h^{-1}\text{Mpc}$ and increase with the brightness of the sample, reaching almost 200% for the LOWZ-L1 sample. It is interesting to note here that H22 did a similar comparison between w_{gg} in the LOWZ sub-sample and measurements from the MICE simulation (Fosalba et al. 2015), which is based on a similar combined HOD and abundance matching model as Flagship. This comparison revealed smaller deviations below 60%, indicating that the stronger deviations which we find in Flagship result from the choice of model parameters for the galaxy population, rather than from the simulation method itself. It is important to note here that the HOD–abundance matching model in Flagship was calibrated against the SDSS Main Sample, but not the LOWZ sample.

4.3. Three-dimensional correlation function of orientation vectors η

When comparing the IA signal in Flagship with the HAGN simulation, we take advantage of the fact that we have access to the three-dimensional galaxy positions and orientations in both simulations. This allows us to measure the three-dimensional alignment without losing information due to projections, as in the analysis of observational samples. In addition, we can define the statistics to depend solely on galaxy orientations but not on the shapes. We expect the shapes of galaxies in HAGN to be, on average, rounder than those found in observations, due to an absence of thin disks, which results from resolution effects in the simulation (see Hoffmann et al. 2022a, and references therein).

This deficiency of disks may bias correlations of intrinsic ellipticity components, such as w_{g+} , when considering samples that include galaxies of all morphological types, as is the case in the present work. We measure the IA signal in HAGN independently of galaxy shapes as the two-point alignment statistics of galaxy orientation with their surrounding large-scale structure in three dimensions, as

$$\eta_X(r) = \langle |\hat{\mathbf{X}}_1 \cdot \hat{\mathbf{r}}_{12}| \rangle(r) - 1/2, \quad (10)$$

where $\hat{\mathbf{X}}_1$ refers to the normalised three-dimensional major or minor axis vectors ($\hat{\mathbf{A}}_{3D}$ or $\hat{\mathbf{C}}_{3D}$, respectively, as defined in Sect. 2) of a galaxy at position 1 and $\hat{\mathbf{r}}_{12}$ is the normalised distance vector that points to a galaxy at position 2, and $\langle \dots \rangle$ is the average over all galaxy pairs that are separated by the distance $r = |\mathbf{r}_{12}|$. The subtraction of 1/2 leads to $\eta_X(r) = 0$ if the alignment between $\hat{\mathbf{X}}_1$ and \mathbf{r}_{12} is random.⁶ Errors are estimated by jackknife resampling, using 88 HEALPix regions for measurements in the Flagship light-cone and 64 cubical regions in HAGN. In the process the jackknife regions are used to organise the galaxy catalogue in a simple tree structure, which allowed for a significant acceleration of the search for galaxy pairs. As in the computation of w_{g+} we increase the signal-to-noise by computing the cross-correlation between the positions of galaxies in a density sample and the orientations of galaxies in magnitude-limited sub-samples. The construction of these samples is described in Sect. 3.4. The code used for the measurements in the Flagship light-cone and the HAGN simulation box (including error estimation) is publicly available⁷. A potential concern for IA constraints from AMR simulations such as HAGN is a systematic effect arising from the alignment between galaxy shapes and the axes of the Cartesian simulation grid, commonly referred to as grid locking. We do not expect the η_X statistic to be affected by this effect, based on the results of Chisari et al. (2015), who showed that grid locking has no significant impact on position–orientation correlations in HAGN.

5. Modelling galaxy shapes

Modelling the intrinsic 3D galaxy shapes is the first out of the two main steps in our IA model. Together with the intrinsic 3D galaxy orientations, which are discussed in Sect. 6, the 3D shapes allow us to derive the 2D intrinsic galaxy shear components by projection along the observer’s line of sight.

We model the 3D shapes following the methodology from H22 which takes into account observed dependencies of galaxy shapes on redshift, magnitude, and colour. In that model, galaxies are approximated as 3D ellipsoids, whose shape is fully characterised by the axis ratios q_{3D} and r_{3D} , which we have defined in Eq. (2). At a fixed point in redshift-magnitude-colour space, the 3D axis ratio distribution is described by a Gaussian model,

$$\tilde{P}(q_{3D}, r_{3D}) = \exp\left\{-\frac{1}{2}\left[\left(\frac{q_{3D} - q_0}{\sigma_{qr}}\right)^2 + \left(\frac{r_{3D} - r_0}{\sigma_{qr}}\right)^2\right]\right\}, \quad (11)$$

where q_0 , r_0 , and σ_{qr} are free model parameters. The distribution is truncated as

$$P = \begin{cases} \tilde{P}_{3D}/N & \text{if } q_{3D}, r_{3D} \in (0, 1]; \\ 0 & \text{else,} \end{cases} \quad (12)$$

⁶In the literature, the 3D alignment statistic is commonly quantified as $\langle \hat{\mathbf{X}}_1 \cdot \hat{\mathbf{r}}_{12} \rangle(r) - 1/3$. Here, we instead adopt Eq. 10, as we find it more intuitive: for a random distribution, $|\hat{\mathbf{X}}_1 \cdot \hat{\mathbf{r}}_{12}|$ is uniformly distributed between 0 and 1 with an average value of 1/2. Both definitions are valid measures of 3D alignment and convey essentially the same information.

⁷<https://github.com/kaidhoffmann/covo>

with

$$N = \int_0^1 \int_0^1 \tilde{P}_{3D}(q_{3D}, r_{3D}) \, dr_{3D} \, dq_{3D} \quad (13)$$

to ensure $A_{3D} \geq B_{3D} \geq C_{3D}$. This model is motivated by results from Hoffmann et al. (2022a), who showed that the distribution of q_{3D} and r_{3D} of disk galaxies in HAGN and Illustris TNG (Pillepich et al. 2018) can be well approximated by a truncated diagonal Gaussian distribution⁸. H22 simplified this Gaussian model by assuming identical dispersion parameters σ_{qr} for q_{3D} and r_{3D} . This simplification was motivated by their finding that such an approach allows for more robust parameter estimation from small, noisy samples, thereby facilitating the model calibration described below. Furthermore, H22 demonstrated that this model is sufficiently accurate for reproducing the observed two-dimensional axis ratio distributions from COSMOS which included galaxies of all types.

The 3D shape of a given galaxy in the simulation can now be obtained from Eq. (12) by randomly drawing the axis ratios q_{3D} and r_{3D} from P . In order to introduce a dependence of galaxy shapes on galaxy properties, an additional modelling step is necessary. For that purpose, we modify the axis ratio distribution parameters based on each galaxy’s redshift, magnitude, and colour. This parameter dependence has been calibrated at distinct points in redshift-colour-magnitude space, such that the distribution of projected 2D axis ratios for an ensemble of randomly oriented ellipsoids matches the observed distribution from the COSMOS catalogue, described in Sect. 3.1. The parameters for a specific galaxy in the simulation are then obtained by interpolation between the points where the model was calibrated. Further details on the method, the motivation and the calibration of the galaxy shape model are given in H22.

The simulated galaxy shapes are publicly available as part of the *Euclid* Flagship Catalogue on CosmoHub⁹ (Carretero et al. 2017b; Tallada et al. 2020). The code used to generate these simulations is available on GitHub.¹⁰

In Fig. 9 we compare shape distributions of the projected 2D axis ratios in Flagship with the COSMOS reference data for a set of ten volume-limited samples of the red and blue galaxies, covering absolute magnitudes in the range of $-19.5 < M_r < -23.0$ and redshifts in the range of $0.2 < z < 1.5$ as detailed in Sect. 3.1. We find an overall good agreement between Flagship and COSMOS in all samples, demonstrating that the simulated 2D galaxy shape distribution follows the observed colour-magnitude-redshift dependence.

6. Modelling galaxy orientations

We model the galaxy 3D orientations in two steps. In the first step we set initial orientations, which are then randomised in a second step, such that the IA statistics in the simulations match the reference measurements.

6.1. Initial orientations

The initial orientations are set separately for central and satellite galaxies, using two distinct methods for each type.

⁸Note that a diagonal Gaussian distribution may not be a suitable approximation for other combinations of shape parameters. For instance, the joint distribution of q_{3D} and s_{3D} must satisfy the constraint $s_{3D} \lesssim q_{3D}$, which is not captured by a diagonal Gaussian distribution.

⁹<https://cosmohub.pic.es>

¹⁰<https://github.com/kaidhoffmann/genIAL>

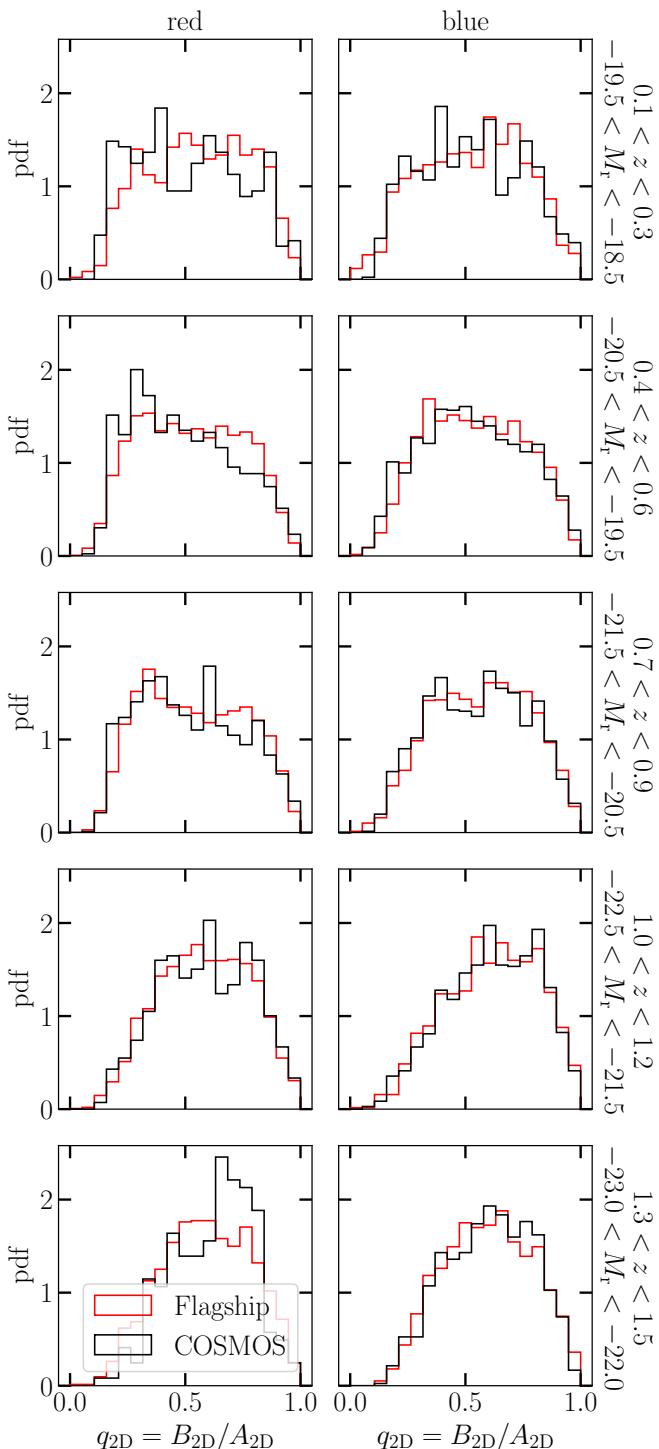


Fig. 9. 2D galaxy axis ratios measured for red and blue galaxies in different volume-limited samples. The redshift and absolute Subaru r -band limits of each sample are indicated on the right. Red and black histograms show results from the Flagship simulation and COSMOS observations, respectively.

Centrals: All central galaxies are initially oriented by aligning their principle axes with those of their host dark matter halos, i.e. $(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}})_{\text{gal}} = (\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}})_{\text{halo}}$. This part of the modelling deviates from previous semi-analytic alignment models (e.g. J13, H22). In these models only red central galaxies are aligned with the principle axes of their host halos while blue centrals are assumed to be disks, whose minor axis is initially aligned with their host

halo's angular momentum vector. We simplify this mixed modelling for red and blue galaxies for three reasons:

1) a colour-dependent modelling of the orientations relies on a colour-based classification of galaxies into disks and ellipticals. Such a classification is known to be inaccurate, in particular when using a simple colour cut, e.g. disks can be red when being oriented edge-on to the observer due to dust extinction or when having a large central bulge (see e.g. the discussion in Xiao et al. 2012; Lackner & Gunn 2012; Hoffmann et al. 2022a; Uzeirbegovic et al. 2022). The mixed modelling further neglects the significant fraction of galaxies with irregular morphologies (e.g. Conselice et al. 2005; Huertas-Company et al. 2016). For these reasons it is not obvious that the mixed model is more accurate than the more simple modelling used in this work and described below.

2) The minor axis direction of a dark matter halo correlates with the direction of its angular momentum (Bailin & Steinmetz 2005; Bett et al. 2007). In our model this alignment will be weaker compared to the mixed alignment models, where it is set explicitly. However, since we randomise galaxy orientations in the second step of our modelling, taking galaxy colour into account, we expect the final alignment signal for blue galaxies to be similar to that obtained with the mixed modelling approach.

3) A consistent description of all galaxies reduces the total number of parameters needed for introducing a redshift, magnitude, and colour dependence of the galaxy-halo misalignment.

Satellites: All satellite galaxies are initially oriented by pointing their major axis towards their host halo's centre, which is motivated by findings in hydrodynamic simulations (Pereira et al. 2008; Pereira & Bryan 2010; Welker et al. 2017) and observations (Johnston et al. 2019; Georgiou et al. 2019), as well as the success of this modelling approach in reproducing the observed alignment statistics at small scales (Fortuna et al. 2021, H22).

Note that this model component could be advanced by aligning satellites with their sub-halos or by introducing a radial dependence to the alignment model (Georgiou et al. 2019). However, in a comparison of these advanced models with the simpler radial alignment model used in this work, Van Alfen et al. (2024) found that the latter is sufficient for reproducing second-order IA statistics as measured in IllustrisTNG.

6.2. Random misalignment

It has been shown in the literature that semi-analytic IA models, which assume perfect alignment between galaxies and their host halos, predict IA statistics that are significantly higher than those observed in the data (Heymans et al. 2004). This discrepancy can be removed by introducing a random misalignment between galaxies and their host dark matter halos (e.g., Heymans et al. 2006; Okumura et al. 2009), which is also predicted by hydrodynamic simulations (e.g., van den Bosch et al. 2002).

To randomise the 3D major and minor axes ($\hat{\mathbf{A}}_{3D}$ and $\hat{\mathbf{C}}_{3D}$, respectively) we follow the approach of H22 by first drawing two misalignment angles θ_A and θ_C from the von Mises–Fisher distribution (Fisher et al. 1993),

$$P(\cos \theta) = \frac{1}{2\sigma_{MF}^2 \sinh(\sigma_{MF}^{-2})} \exp\left(\frac{\cos \theta}{\sigma_{MF}^2}\right), \quad (14)$$

for each galaxy. The randomised principle axis vectors $\hat{\mathbf{A}}_{3D}^r$ and $\hat{\mathbf{C}}_{3D}^{r'}$ are then constructed such that $\hat{\mathbf{A}}_{3D} \cdot \hat{\mathbf{A}}_{3D}^r = \cos(\theta_A)$ and $\hat{\mathbf{C}}_{3D} \cdot \hat{\mathbf{C}}_{3D}^{r'} = \cos(\theta_C)$. Here $\hat{\mathbf{C}}_{3D}^{r'}$ is a temporary vector which is

in general not perpendicular to $\hat{\mathbf{A}}_{3D}^r$ since θ_A and θ_C are drawn independently from each other. The final randomised minor axis orientation is therefore obtained as $\mathbf{C}_{3D}^r = (\hat{\mathbf{A}}_{3D}^r \times \hat{\mathbf{C}}_{3D}^{r'}) \times \hat{\mathbf{A}}_{3D}^r$ and is then normalised to $\hat{\mathbf{C}}_{3D}^r = \mathbf{C}_{3D}^r / |\mathbf{C}_{3D}^r|$.

We adopt the von Mises–Fisher distribution, since it successfully employed by H22 to reproduce the observed IA signal from BOSS LOWZ data using a semi-analytical IA model similar to the one adopted in this work. Bett (2012) showed that this distribution provides a good fit to the misalignment angles between halo and galaxy angular momentum vectors in simulated haloes including baryons and galaxy formation physics. However, a direct validation of the von Mises–Fisher distribution against misalignment angles between the principal axes of central galaxies and haloes, as well as between the major axes of satellites and the halo-centre direction, remains to be done. Velliscig et al. (2015) analysed such distributions in the EAGLE and cosmo-OWLS simulations and proposed an alternative fitting model. While its functional form is similar to the von Mises–Fisher distribution, it requires four free parameters, which would make the parameter calibration in our work substantially more challenging. A further single-parameter alternative to the von Mises–Fisher distribution is the Dimroth–Watson distribution. Van Alfen et al. (2024) demonstrated that a semi-analytical IA model based on this distribution successfully reproduced IA statistics of galaxies in the IllustrisTNG hydrodynamical simulation.

The width of the distribution σ_{MF} is a free parameter that provides control over the alignment amplitude. Increasing (decreasing) σ_{MF} leads to a higher (lower) randomisation of the initial orientations and hence to a lower (higher) IA signal in the simulation. We can hence control the dependence of the predicted IA signal on galaxy properties by introducing a dependence of σ_{MF} on galaxy redshift z , absolute rest-frame magnitude M_r and rest-frame colour index $u - r$ (defined in Sect. 3),

$$\sigma_{MF} = p_0 \underbrace{\left(\frac{z}{z_0} + 1 \right)^{p_1}}_{\sigma_z(p_1)} \underbrace{\left(\frac{M_r}{M_0} + p_2 \right)^{p_3}}_{\sigma_{mag}(p_2, p_3)} \underbrace{\left(\frac{u - r}{(u - r)_0} + p_4 \right)^{p_5}}_{\sigma_{col}(p_4, p_5)}, \quad (15)$$

The constants $M_0 = -22$, $(u - r)_0 = 1$ and $z_0 = 1$ are chosen to avoid large values of the bases of the σ_z , σ_{mag} and, σ_{col} terms¹¹. We chose the functional form of these terms to allow for a large range of dependencies of the alignment signal on galaxy properties while avoiding negative values for σ_{MF} , as can arise for instance when extrapolating linear relations. In addition to the dependence on galaxy properties, we add a dependence on galaxy type, which can be either central or satellite by using a separate set of the parameters p_i for each type.

As a result, we have two von Mises–Fisher distributions in our model, $\sigma_{MF}^{\text{cent}}$ and σ_{MF}^{sat} , with a total number of 12 free parameters. By introducing a type dependence to the model we obtain control over the scale dependence of the final IA signal in the simulation (see Sect. 6.3). Another motivation for using two separate sets of misalignment parameters for centrals and satellites is that the initial orientations were set with different methods. Using a type-independent misalignment would therefore be an inconsistency in the model. The final form of Eq. (15) is a result of a development process, in which we gradually added complexity to the model until we were able to match the different constraints simultaneously with the simulation. This matching is achieved by calibrating the parameters p_i as described in Sect. 7.

¹¹The M_0 value in particular was motivated by the literature (Joachimi et al. 2011; Singh et al. 2015).

6.3. Response of IA statistics to galaxy misalignment

In preparation for the calibration of the σ_{MF} parameters from Eq. (15) we study our IA model in a simplified setting, fixing $\sigma_{MF}^{\text{cent}}$ and σ_{MF}^{sat} to constant values that are independent of galaxy redshift, magnitude, and colour. In Fig. 10 we show w_{g+} versus the scale r_p for the LOWZ and SDSS sub-samples, comparing the measurements in the observed reference samples with those from the Flagship mock samples for three cases of fixed $\sigma_{MF}^{\text{cent}}$ and σ_{MF}^{sat} values.

In the first case we apply a minimal misalignment to both, centrals and satellites, by setting $\sigma_{MF}^{\text{cent}} = \sigma_{MF}^{\text{sat}} = 0.05$ (red lines in Fig. 10). The resulting amplitudes are roughly between 2 and 20 times higher than those of the reference samples, depending on the sample and the scale. Besides these strong differences, we also find a similarity between the simulation and the reference samples in that the amplitudes increase the brighter and the redder the samples are. This finding might result from the similar magnitude and colour dependence of the galaxy clustering in the observations and the simulation (see Fig. 7).

In the second case of our simplified model we keep the low misalignment for centrals in Flagship while increasing misalignment for satellites by setting $\sigma_{MF}^{\text{cent}} = 0.05$, $\sigma_{MF}^{\text{sat}} = 0.75$ (green lines in Fig. 10). We find that the amplitudes strongly decrease at small scales ($r_p \lesssim 5 h^{-1} \text{Mpc}$), while we see no significant change at large scales in the LOWZ sub-samples and a relatively weak decrease for the SDSS-sub-samples. In the third case we keep the low misalignment for satellites, while increasing the misalignment for centrals, setting $\sigma_{MF}^{\text{cent}} = 0.75$, $\sigma_{MF}^{\text{sat}} = 0.05$ (blue lines in Fig. 10). We now find a strong decrease of the amplitudes at large scales ($r_p \gtrsim 5 h^{-1} \text{ Mpc}$) compared to the first case and only mild or no significant changes at small scales compared to the first case, depending on the sample. These trends persist across a range of satellite fractions in the galaxy samples; see Table 1.

We extend this investigation in Fig. 11, where we compare the 3D alignment statistics η_A and η_C , measured in the three magnitude-limited samples H1 to H3 from HAGN (from bright to faint, as defined in Sect. 3) with the corresponding measurements derived from the mock HAGN samples from Flagship. The Flagship results are shown for the same three $\sigma_{MF}^{\text{cent}}$, σ_{MF}^{sat} combinations discussed above.

For the measurements in HAGN we find the behaviour expected from results in the literature (Chisari et al. 2015), i.e. η_A (η_C) have positive (negative) amplitudes, indicating a preference for the galaxies’ 3D major (minor) axis to be oriented parallel (perpendicular) to the vector pointing to its neighbours.

The dependence of the total amplitudes in HAGN on scale and magnitude further follows our expectations, as it decreases with scale and increases with the samples brightness. The same dependencies on scale and magnitude can be seen in Flagship at large scales ($r_{12} > 5 h^{-1} \text{Mpc}$) for the three cases of our simplified IA model. However, at small scales the magnitude dependence in Flagship is inverted, as the amplitude decreases with the brightness of the sample. This finding is interesting, as we have not seen such an inverted small-scale magnitude dependence in the Flagship mocks of the observed samples, where we found w_{g+} amplitudes to be lower for the SDSS samples than for the brighter LOWZ samples.

This behaviour poses a challenge for the calibration of the magnitude dependence of the IA model, as we need to randomise the orientations of faint satellites more than those of bright satellites to obtain an increase of the signal with brightness in the mock HAGN samples. Besides this magnitude dependence, it is interesting to note in Fig. 11, that increasing the randomisa-

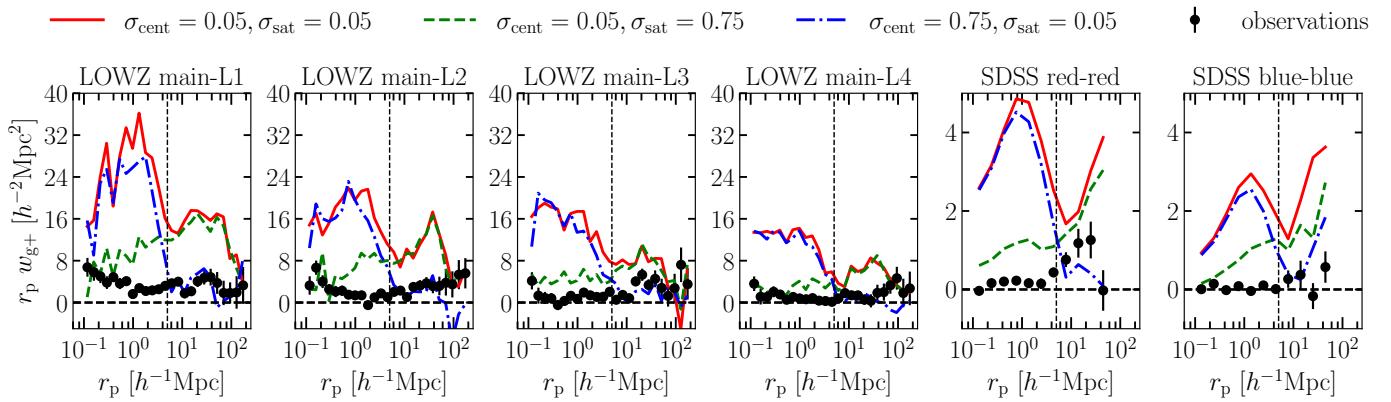


Fig. 10. Projected galaxy-ellipticity cross-correlation between a density sample and various shape samples. The four left panels show results based on shapes from the luminosity sub-samples, with the full LOWZ sample as the density sample. The two right panels display results for SDSS, using the Main sample as the density sample and colour sub-samples as shape samples. Black dots represent observational data, while coloured lines represent Flagship simulations for three simplified intrinsic alignment (IA) models, each using different fixed values for $\sigma_{\text{MF}}^{\text{cent}}$ and $\sigma_{\text{MF}}^{\text{sat}}$.

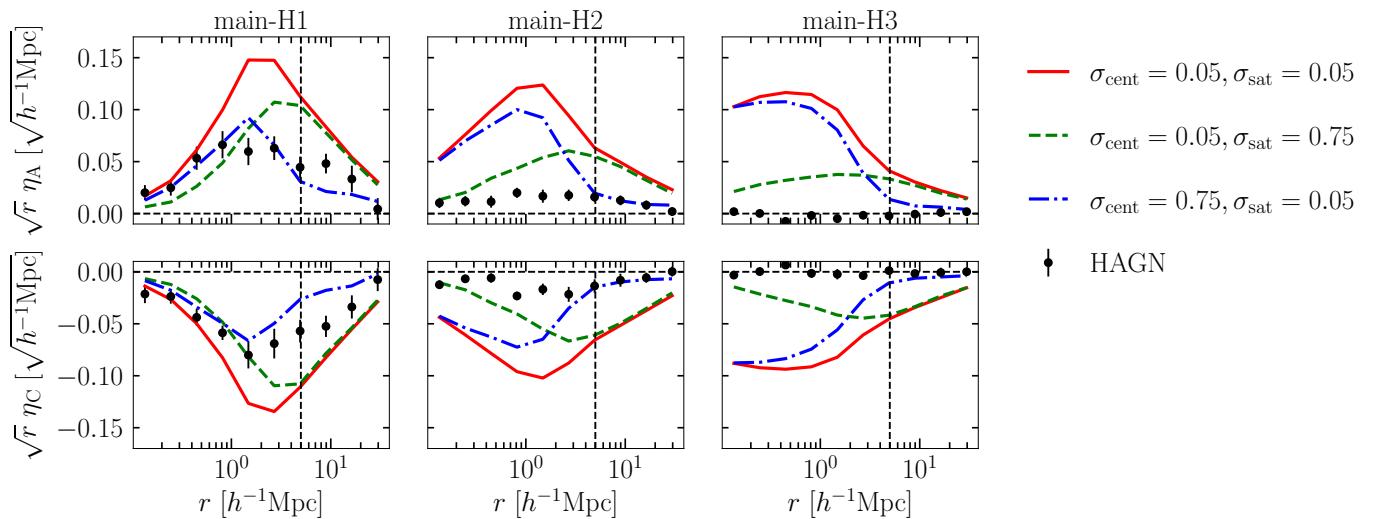


Fig. 11. 3D IA cross-correlations (η , as defined in Eq. 10) between a main density sample and the orientations of galaxies in three luminosity-limited sub-samples. Orientations are given by the 3D major and minor axes A and C . Black dots show measurements from the HAGN simulation at $z = 1.0$ with errors indicating the standard deviation. Coloured lines show the results from mock HAGN samples from Flagship for three simplified IA models with different combinations of fixed values for $\sigma_{\text{MF}}^{\text{cent}}$ and $\sigma_{\text{MF}}^{\text{sat}}$, analogous to Fig. 10. Errors for Flagship measurements are similar to those displayed in Fig. 16 and are not shown for clarity.

tion of orientations only for satellites decreases the signal only at scales below $r_{12} = 5h^{-1}\text{Mpc}$ (compare the red and green lines). Increasing the randomisation for centrals on the other hand decreases the signal at large and at small scales simultaneously (compare the red and blue lines). Overall we conclude from this preliminary investigation that our IA model provides a high flexibility, as it allows for a significant variation of the IA amplitude in the simulation. The goal is now to calibrate the dependence of $\sigma_{\text{MF}}^{\text{cent}}$ and $\sigma_{\text{MF}}^{\text{sat}}$ on galaxy properties, such that the signal matches all reference measurements at all scales simultaneously.

6.4. Variations across random realisations

The IA model parameters in Eqs. (11) and (15) determine the probability distributions from which we draw random shapes and orientations for each galaxy, respectively. For a fixed set of IA model parameters the shapes and orientations of individual galaxies will therefore vary across different random realisations.

This variation translates into a variation of the resulting IA signal that is measured in the simulation.

We illustrate this effect in Fig. 12, where we show w_{g+} , measured in five random realisations of the IA model for the mock BOSS LOWZ L1 sample. The IA parameters used for generating these realisations are those used for the production of the final Flagship IA catalogue. We find a substantial variation across random realisations, which can reach the same order of magnitude as the signal itself. This variation poses a problem when calibrating the model, as it introduces noise in the cost function that we aim to minimise. In order to mitigate the impact of that noise on our calibration, we run the IA model up to five times at each point in the parameter space during the calibration. We obtain the final IA statistics that enter the cost function by averaging over measurements from the N_i random realisations, i.e. $X_{\text{model}} = \sum_i X_{\text{model}}^i / N_i$, where i indicates a single random realisation. These averaged statistics are shown as red symbols in Fig. 12.

7. Misalignment parameter space exploration

We explore the parameter space of the galaxy misalignment model component following two objectives. Firstly, we aim to calibrate the model parameters by minimising the deviations between the IA statistics from the simulation and those from the reference samples. Our second objective is to trace the probability distribution of the IA model parameters, which enables us to estimate the uncertainties on the IA contamination in *Euclid*-like sample that is predicted by the simulation.

7.1. Cost functions

We calibrate the parameters p_i in the σ_{MF} model from Eq. (15) by minimizing the total deviation between the IA statistics measured in the different mock IA catalogues from Flagship and those obtained from the corresponding reference measurements. For a given sample s we define this deviation as

$$\chi_s^2 = \sum_i^{N_{\text{bin}}} \frac{(X_{\text{model}} - X_{\text{ref}})_i^2}{(\sigma_{\text{model}}^2 + \sigma_{\text{ref}}^2)_i}, \quad (16)$$

where X is either the projected statistics w_{g+} in the case of the observational samples from SDSS and LOWZ or the 3D statistic η_A in the case of the HAGN samples (defined in Sect. 4). The predictions of the model measured in Flagship and the measurements of the reference samples are denoted by X_{model} and X_{ref} , respectively. The corresponding standard deviations are denoted as σ_{model} and σ_{ref} , while N_{bin} is the number of scale bins in which the statistic X is measured.

When calibrating the misalignment parameters, our aim is to minimise the overall deviation between Flagship and the reference samples, which we define as the cost function

$$C = \sum_s^{N_s} w_s \chi_{s,r}^2, \quad (17)$$

where $N_s = 9$ is the number of galaxy sub-samples summarised in Table 1, $\chi_{s,r}^2 = \chi_s^2/N_{\text{bin}}$ is the reduced χ^2 of sample s , and w_s is a weight that we assign to each sub-sample. When calibrating the model parameters for central galaxies, we computed C from measurements of X at small and large scales, respectively, as explained in Sect. 7.2, using slightly different weights in both cases.

For large scales ($> 5 h^{-1}\text{Mpc}$), we assign weights of 0.1 to each of the four BOSS LOWZ sub-samples, resulting in a combined weight of 0.4. Similarly, we assign weights of 0.2 to each of the two SDSS samples, which leads to a combined weight of 0.4 for SDSS and a total observational weight of 0.8. For the three HAGN sub-samples, we set weights of 0.2/3 each, which leads to a total weight of 1.0 for all samples combined. This weighting strategy reduces the overall contribution of the HAGN samples, accounting for the higher confidence we place in the observational constraints and aiming to mitigate potential biases introduced by inaccuracies in the IA signal predicted by HAGN. When computing C at small scales ($< 5 h^{-1}\text{Mpc}$), we neglect the LOWZ samples L1 and L2 by setting their weights to zero. This is motivated by the strong deviations of the clustering in Flagship from the reference observations (see Fig. 7). Since these deviations in clustering also affect the w_{g+} predictions from the Flagship IA model, including L1 and L2 may bias the calibration results for satellite parameters. To maintain the same overall LOWZ contribution to C of 0.4 as on large scales, we increase the weights for the fainter sub-samples, L3 and L4, to 0.2 each.

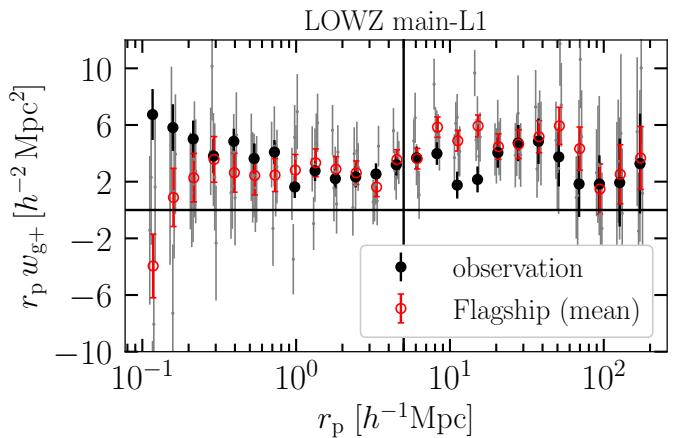


Fig. 12. Projected galaxy-ellipticity cross-correlation for the BOSS LOWZ sample L1. Grey dots represent results from five random model realisations using parameters from the final Flagship model. Red dots indicate the average across these random realisations with re-scaled errors, while black dots show observational results.

While the cost C is a valid measure for the deviations between the Flagship and the reference samples, it is not suitable for the error propagation technique, which we present in Sect. 8, since it is not directly related to the likelihood \mathcal{L} that σ_{model} and σ_{ref} are randomly sampled from the same underlying probability distribution. In order to establish this relation, we compute, in addition to C , the total deviation as

$$\chi_{\text{tot}}^2 = \sum_s^{N_s} w_s \chi_s^2. \quad (18)$$

Assuming a normal distribution of the likelihood for measuring X at a given point \mathbf{p} in the parameter space of the model, we can obtain the likelihood as $\mathcal{L} \propto \exp\left\{-\frac{1}{2}\chi_{\text{tot}}^2\right\}$. The weights w_s are the same as those we apply when computing C in Eq. (17).

In addition to weighting, our χ_{tot}^2 definition differs from the common statistics χ^2 in two aspects. 1) X_{model} is not an analytic prediction but a measurement with non-negligible errors that need to be taken into account when quantifying the significance of the difference between model and reference measurements. 2) We do not take into account the covariance between different scale bins. This simplification is valid if the signals are small and the errors are dominated by shot noise, but future calibration exercises should take into account the full covariance.

7.2. Parameter sampling method

A challenging aspect in the calibration is to efficiently sample the high-dimensional space spanned by the 2×6 parameters in Eq. (15) that describe the redshift, luminosity, and colour dependence of the misalignment parameters $\sigma_{\text{MF}}^{\text{cent}}$ and $\sigma_{\text{MF}}^{\text{sat}}$ in our IA model. This challenge is due to the numerical cost, which arises in part from generating the simulation, but mainly from the measurements of the various IA statistics at each sampling point.

In order to reduce the sampling volume in the parameter space and thus lower the computational costs of the calibration, we leverage a finding from our preliminary study in Sect. 6.3. There, we found that on large scales ($> 5 h^{-1}\text{Mpc}$), the alignment statistics w_{g+} and η_A are most sensitive to the parameters controlling the alignment of centrals (p_i^{cent}), while the impact of the parameters controlling satellite alignment (p_i^{sat}) is relatively weak. This result allows us to first explore only the six-

dimensional space spanned by p_i^{cent} , by comparing the IA statistics from Flagship to those of the reference samples only at large scales between 5 and $40 h^{-1} \text{Mpc}$ ¹². Once we have found a set of best-fit parameters for centrals, we fix these parameters and calibrate the parameters p_i^{sat} by comparing to the IA reference statistics at small scales, between 0.1 and $5 h^{-1} \text{Mpc}$.

In both cases, we generate an initial set of sampling points by uniformly sampling the parameter space within the limits $0 < p_0 < 10$, $-10 < p_1 < 10$, $-0.5 < p_2 < 10$, $-10 < p_3 < 10$, $0.5 < p_4 < 10$, $-10 < p_5 < 0$. The lower limits for p_2 and p_4 were chosen to ensure a positive basis of the σ_{mag} and σ_{col} terms in Eq. (15) for $M_r < -11$ and $u - r > -0.5$, respectively, hence avoiding imaginary values for σ_{MF} within the magnitude and colour range covered by the mock *Euclid* sample (see Fig. 1). The constraint $p_5 < 0$ is imposed to ensure that σ_{MF} decreases with the $u - r$ colour index, leading to higher misalignment for bluer galaxies compared to red galaxies, and consequently a lower IA signal.

To further reduce the computational cost of the calibration, we discard sampling points from the initial distribution for which $\sigma_{\text{MF}} > 2$ within the volume of the redshift-magnitude-colour space covered by Flagship, i.e. $0.0 < z < 2.0$, $-23.0 < M_r < -16.0$; $0.5 < u - r < 3.0$ (see Fig. 1). Our reasoning behind this choice is that values of $\sigma_{\text{MF}} > 2$ lead to a distribution of random galaxy orientations that is close to a uniform distribution. It is therefore not necessary for our σ_{MF} model to predict higher values in order to include a zero alignment signal for any object in our *Euclid*-like sample as a possible calibration outcome. This pre-selection allows us to reduce the uniform initial random distribution by over 98% percent. Note here that we apply this selection only on the initial sampling points, but not on those that we probe during the subsequent calibration.

Starting from an initial set of 1000 points, we continue sampling the parameter space with an iterative procedure, that consists of the following steps:

1. Run the IA model on the mock galaxy samples at each sampling point in the parameter space, measure the IA statistics and compute the corresponding deviation from the reference samples, quantified by the cost function C from Eq. (17).
2. Select the fraction f_{low} of points with the lowest cost and compute their covariance matrix.
3. Draw N_r random points from the Gaussian distribution that is described by the covariance matrix from step 2.

For centrals we repeat these steps 12 times, while gradually decreasing f_{low} from 5% to 0.5% from the first to the final iteration. For satellites we performed only 4 iterations, decreasing f_{low} from 5% to 1%, as we find a faster convergence of the results. In each step of the iteration we generate 250 new sampling points, which leads to a total of 4000 and 2000 points for centrals and satellites, respectively, including the initial guess distribution.

As detailed in Sect. 6.4, we average the IA statistics over N random realisations to mitigate the impact of noise on the model predictions. We thereby set $N = 3$ for an initial ‘burn-in’ phase, which lasts over the first seven iterations for centrals and is restricted to the initial sampling step for satellites. Once we find

¹²Note that the HAGN results may be affected by the limited box size of $100 h^{-1} \text{Mpc}$. However, we do not expect this effect to significantly impact our calibration results, as the contributions of measurements to the cost function χ_s^2 are down-weighted by their errors, which increase substantially with scale. In addition, we have reduced the weight of the HAGN constraints to the total cost C , as detailed in Sect. 7.1

Table 2. Parameters of the galaxy-halo misalignment model used in Flagship; see Eq. (15). Errors are 1 sigma uncertainties, derived from the distributions of sampling points.

parameter	centrals	satellites	model component
p_0	$1.02^{+2.52}_{-1.02}$	4.74 ± 1.72	overall amplitude
p_1	-0.41 ± 0.24	0.43 ± 0.34	σ_z
p_2	3.46 ± 1.85	5.72 ± 2.33	σ_{mag}
p_3	2.71 ± 0.92	2.26 ± 1.12	
p_4	4.28 ± 1.27	4.27 ± 2.34	σ_{col}
p_5	-2.43 ± 0.59	-3.07 ± 0.88	
$\chi_{\text{tot}}^2/\text{d.o.f.}$	0.95	1.32	

the region with the lowest cost C to be visually well covered by new sampling points, we end the ‘burn-in’ phase and increase the number of iterations to 5, gaining precision in the model predictions at the price of higher computational costs.

7.3. Calibration results

The set of galaxy-halo misalignment parameters for which we find the cost C to be minimal (hereafter referred to as best-fit parameters) is given in Table 2. These are the parameters which we use to compute the alignment on the full light-cone output of Flagship.

In order to obtain a physical understanding of this result, we show in Fig. 13 the redshift, magnitude, and colour components of our σ_{MF} model for the parameter combinations probed during the calibration. Results are shown for central and satellite galaxies on the top and bottom panels, respectively. They are normalised to the model values at $z_0 = 1.0$, $M_0 = -22$, $(u-r)_0 = 1.0$ for a better visual comparison of the scaling with galaxy properties. The thick black line in this figure shows the model components for our best-fit parameters. The thin lines show the results for all parameter combinations that were sampled during the calibration process. Their colours indicate the corresponding χ_{tot}^2 deviations between the IA statistics in Flagship and the reference samples, defined by Eq. (18).

Focusing on sampling points with the lowest χ_{tot}^2 deviations from the reference samples ($\hat{\chi}_{\text{tot}}^2 < 1.5$), we find for both centrals and satellites that the dependence of the misalignment on redshift and magnitude, quantified by σ_z and σ_{mag} , is weak compared to the colour dependence, which is quantified by σ_{col} . Results are distributed around unity, which is consistent with a galaxy misalignment that is independent on redshift and magnitude. This weak dependence may allow for a reduction of free parameters in a future version of the IA model, by assuming a constant relationship of σ_{MF} with redshift and magnitude. Note that such a constant relationship does not imply that the alignment signal is independent of these properties. This can be seen in Fig. 10 and 11, where we found a strong variation of the IA signal across different galaxy samples for fixed misalignment parameters. For centrals, a dependence on magnitude and redshift can be introduced solely through the mass and redshift dependence of the host halo alignment, where the former may be affected mass resolution effects, introducing mass dependent noise in the measured halo orientations (see discussion in Sect. 2.2 and H22). For satellites, such dependencies may instead be introduced via the mass- and redshift-dependence of host-halo concentrations (Dutton & Macciò 2014). Therefore, a more detailed physical interpretation of the results may be misleading.

In contrast to σ_z and σ_{mag} , σ_{col} deviates significantly from a constant relation, as the values decrease with larger $u - r$ colour. From our preliminary studies (Fig. 10 and 11), we expect this decrease of misalignment to translate into an increase of the alignment signal with galaxy redness. The trends that we find for the parameter sampling points with the lowest χ^2_{tot} values are also present for our best-fit model that is defined by the parameters for which the cost C is minimal. However, the latter shows a stronger dependence on redshift and colour. We attribute this offset to differences in the definitions of C and χ^2_{tot} , as well as the noise affecting both quantities (see Sect. 6.4), which may cause a displacement between the sampling point with the lowest C value and the minimum of the underlying distribution.

The impact of this noise on our constraints becomes more apparent in Fig. 14, which shows the distribution for the sampling points in the parameter space for centrals and satellites in the left and right panels, respectively. The colours indicate the χ^2_{tot} deviation between the IA statistics of the simulation and the reference samples. Red dots are those with the highest χ^2_{tot} values and indicate the volume covered by the initial distribution, whereas the bluest dots mark the region in the parameter space for which we found the best agreement between simulation and reference samples.

There are several interesting aspects to note in Fig. 14. Firstly, the χ^2_{tot} values are significantly higher for satellites than for centrals, which could result from the strong deviations between the 3D IA statistics of Flagship and HAGN at small scales that we see in Fig. 11. Second, the shape of the parameter regions with good fits (with $\chi^2_{\text{tot}} < 2$) for centrals and satellites is similar, despite the fact that the misalignment described by these parameters is defined for very different initial orientations in both cases, as we have discussed in Sect. 6.1. One possible explanation for this finding could be that the constraints are too weak to reveal significant deviations. Another potential reason could be a bias introduced by using the same initial guess distribution for centrals and satellites. However, the fact that the best-fit volumes are significantly smaller than the volume spanned by the initial distributions, and are well enclosed by them, argues against a bias imposed by the initial guess.

A third aspect to note is the large uncertainty in the parameter p_0 , which controls the overall amplitude of the misalignment and therefore directly the amplitude of the alignment signals. This is because χ^2_{tot} is affected by the errors of the IA statistics in the reference samples as well as those from the measurements in the Flagship simulation (Eq. 18). Correlations with the parameters p_1 and p_5 for centrals and p_4 for satellites further contribute to the uncertainty on p_0 . Finally, we observe a strong correlation between the parameters p_3 and p_5 , which are the exponents of the magnitude and colour terms of our misalignment model, and hence describe how strongly the galaxy misalignment scales with these galaxy properties. A possible explanation for the correlation of these parameters is the correlation between galaxy colours and magnitudes in Flagship, which can be seen in projection over redshift in Fig. 1.

However, we refrain from a detailed physical interpretation of the calibration results, as these are likely influenced by two systematic effects. First, the mass dependence of noise in the measured halo orientations (Sect. 2.2) can impact the inferred dependence of galaxy–halo misalignment on galaxy properties. In addition, differences in clustering between Flagship and the observations (Sect. 4.2) are expected to further affect the calibration parameters. Since the clustering also affects the w_{g+} measurements, an overly high clustering will be compensated by an overly low alignment amplitude, and vice versa. We mit-

igate this effect by excluding data for which we find the most significant deviations in w_{gg} from the calibration, i.e. the measurements from the BOSS LOWZ samples L1 and L2 at scales below $5 h^{-1} \text{Mpc}$ (Sect. 7.2). However, smaller deviations are still present in the remaining data and their strength depends not only on scale but also on the luminosity, colour and redshift of the samples. The deviations in the clustering may therefore affect the dependence of the IA signal on redshift and galaxy properties, as described by our calibration parameters. Since the statistical uncertainty of w_{g+} is generally much larger than that of w_{gg} , significant deviations in the clustering amplitude do not generally translate into significant bias in IA statistics. Nevertheless, the IA parameters controlling the galaxy–halo misalignment should be regarded as effective parameters that are specific to Flagship.

7.4. Validation

For validation we compare the predictions from our best-fit Flagship IA model, averaged over five realisations, against the reference samples used for calibration. In Fig. 15, we show the projected cross-correlation w_{g+} for the four luminosity sub-samples of the BOSS LOWZ sample and the colour sub-samples of the SDSS main sample (analogous to Fig. 10). For the LOWZ sub-samples, we find an overall good agreement between results from Flagship and the observational reference, as the discrepancies are within the standard deviations of the measurements. However, for the brightest samples L1 and L2 and at scales of less than $1 h^{-1} \text{Mpc}$, we find the Flagship results to be consistently below the observational results. Note that for these samples, we also found the strongest deviations from the reference samples in clustering and have therefore set their weight in the cost function to zero when calibrating the misalignment parameters for satellites (Sect. 7). The simulation is hence not calibrated to match the LOWZ observations in this regime.

For the red SDSS sub-sample we find the deviations between the Flagship results and the observations to be consistent overall within the errors, although we note that at scales above $5 h^{-1} \text{Mpc}$ the Flagship results are consistently below the observations. The Flagship results for the blue SDSS sample are consistent with zero at all scales and therefore below the observational reference. However, also here the deviations between Flagship and observations are consistent within the errors.

In Fig. 16 we compare the 3D IA statistics η , defined in Eq. (10), from Flagship with the reference measurements from HAGN for the three magnitude-limited samples H1 to H3. The top panels display η_A , which measures the average alignment between the major axis of a galaxy in a magnitude sub-sample with the direction of a neighbouring galaxy from the main sample. Similarly, the central panels show η_C , which quantifies the same alignment but for the galaxy’s minor axis (analogously to Fig. 11). The bottom panels show the relative deviations between Flagship and HAGN.

We find that Flagship matches the overall trends predicted by HAGN, showing a decrease in amplitude for fainter samples and at both small and large scales. However, our comparison also reveals several discrepancies: At scales above $5 h^{-1} \text{Mpc}$ we find the η_A measurements from Flagship to deviate from HAGN by up to 50% for the bright samples H1 and H2. While deviations in the largest-scale bin are greater, they are consistent with the increased measurement errors at that scale. For the faint sample, H3, we observe similar absolute deviations as for the brighter samples. However, due to the low amplitude of the signal, these absolute deviations result in large relative deviations of up to a factor of five. At scales below $5 h^{-1} \text{Mpc}$, we find the Flagship

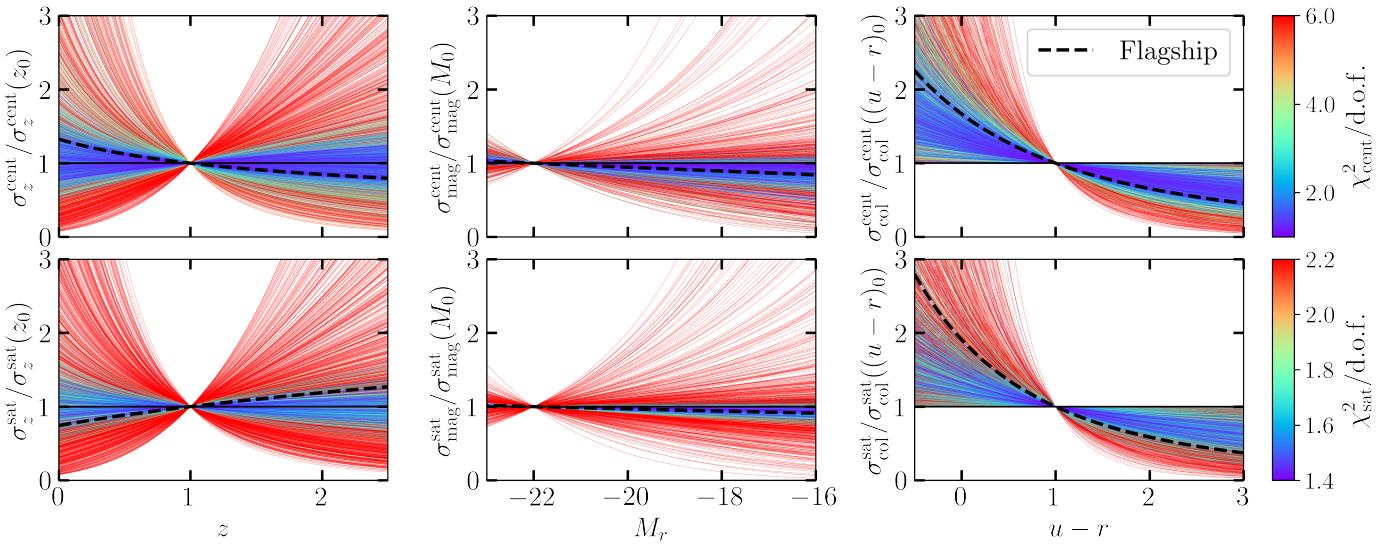


Fig. 13. Components of the model for the galaxy misalignment parameter σ_{MF} from Eq. (15) for the parameters sampled during the model calibration. Line colours indicate the χ^2_{tot} deviation between simulations and reference data (as defined in Eq. 18) per degree of freedom (d.o.f.). The black dashed line corresponds to the model used in the production of the Flagship simulation. Top and bottom panels show results for the calibration of satellite and central misalignment, respectively. All results are shown with respect to the values at $z_0 = 1.0$, $M_0 = -22$, $(u - r)_0 = 1.0$ to facilitate the comparison of the dependence of the misalignment amplitude on the different galaxy properties.

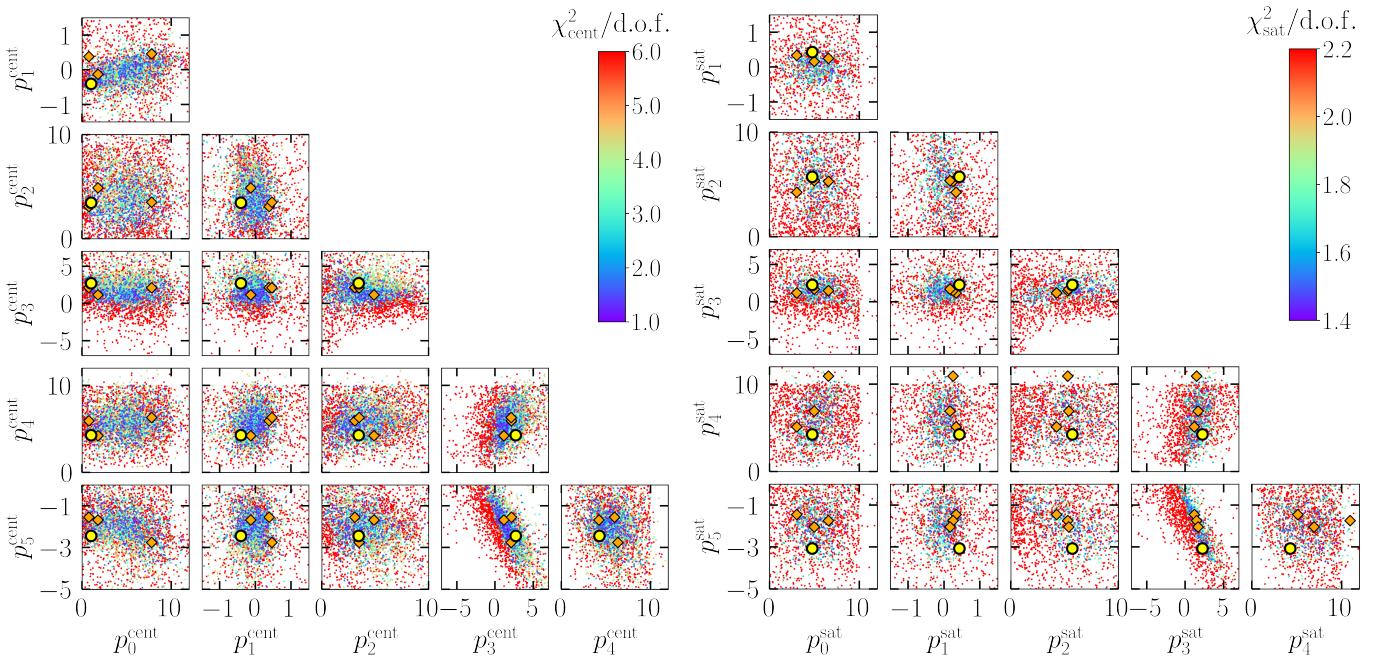


Fig. 14. Parameter combinations of the galaxy misalignment model (Eq. 15) at which the simulation was evaluated in the calibration process, described in Sect. 6. Colours indicate the χ^2_{tot} (averaged over five realisations, see Sect. 7.2) deviations per degree of freedom (d.o.f.) of different IA statistics measured in the simulation from the references measurements (see Eq. 18). Left panels show the parameter combinations for centrals for which the χ^2_{tot} values are derived from IA statistics at scales larger than $5h^{-1}\text{Mpc}$. Right panels show the corresponding results for parameters of satellites, where χ^2_{tot} is derived at scales below $5h^{-1}\text{Mpc}$. The parameters used for the IA model in Flagship (see Table 2) are marked as yellow dots. Orange diamonds show the three sampling points with the lowest χ^2_{tot} values.

results to be up to a factor 2 below those from HAGN, as the signal in Flagship decreases towards zero. This discrepancy is most notable for the brighter samples, H1 and H2, where the HAGN measurements exhibit a high signal-to-noise ratio.

Interestingly, we achieved better agreement between Flagship and HAGN for the brightest sample, H1, at small scales when fixing the misalignment parameters to $\sigma_{\text{MF}}^{\text{cent}} = 0.75$ and

$\sigma_{\text{MF}}^{\text{sat}} = 0.05$ in our preliminary analysis (see Fig. 11). This demonstrates that the IA model can replicate the scale dependence of the HAGN signal within a specific luminosity and redshift range. However, obtaining highly accurate matches between the constraining datasets and Flagship across all luminosity, colour, redshift, and scale ranges simultaneously, while using

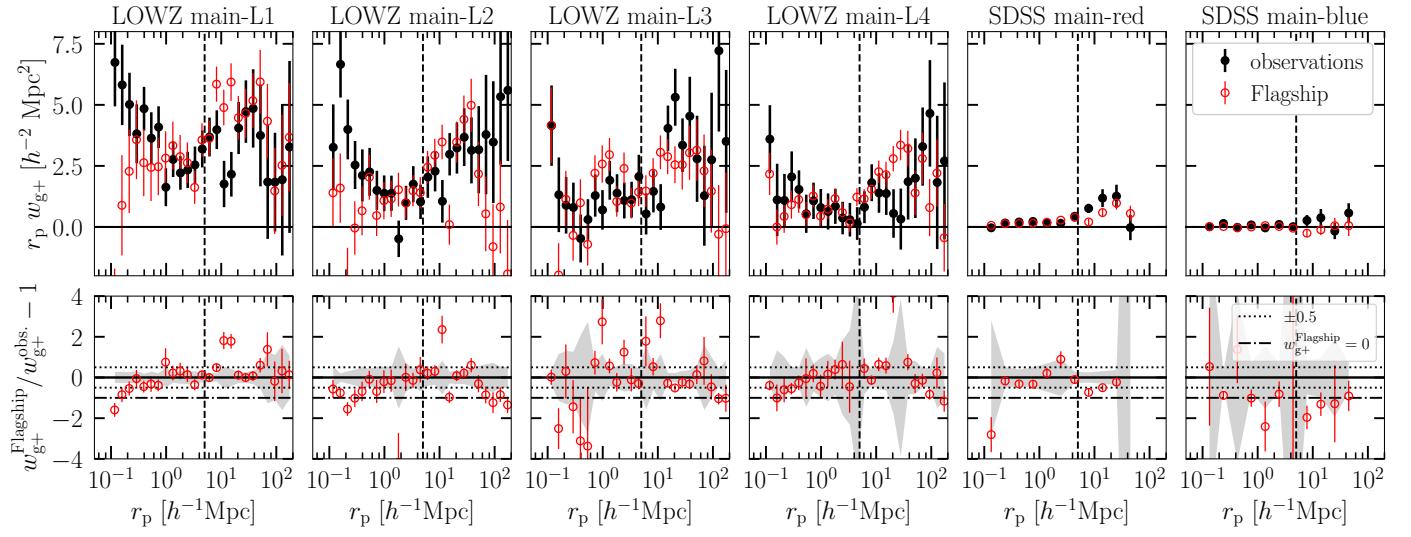


Fig. 15. Projected galaxy-ellipticity cross-correlation function w_{g+} as a function of transverse separation for the observational reference samples (black points) and the corresponding calibrated Flagship mock samples (red points), analogous to Fig. 7. The bottom panels show the relative deviation of the Flagship signals from the observations. Dashed-dotted horizontal lines at -1 in the bottom panels indicate the relative deviation expected for a null detection of the IA signal in Flagship; dotted horizontal lines indicate a relative deviation of ± 0.5 . The dashed vertical lines separate the regimes where central and satellite misalignment parameters were fitted in the simulation. The grey areas indicate the standard deviations of the observations.

misalignment parameters that vary based on galaxy properties, remains a challenge for our model.

This difficulty may stem from three primary factors. First, it is possible that the IA signal in HAGN is inconsistent with the observational constraints, making it difficult to match both HAGN and observational data, even with a realistic IA model. Second, there may be limitations in our modelling approach, including: a) insufficient flexibility in the model to capture the full redshift, magnitude, and colour dependence of galaxy misalignment, b) shortcomings in the model calibration process, or c) the relatively low weighting of the HAGN sample in the cost function used for calibration. Third, inaccuracies of the Flagship properties, such as the clustering, colour, or luminosity distributions may reduce the accuracy of the IA model predictions, even if the latter is sufficiently realistic. These issues, either individually or in combination, may contribute to the discrepancies we observe for the HAGN H1 sample at small scales. However, given the overall good agreement between Flagship and HAGN when considering all samples and all scales, we conclude that our IA model is sufficiently realistic to serve as a test-bed for analytical IA models.

The simulated galaxy orientations and the simulation code are publicly available.^{9, 10}

8. Uncertainty propagation

To eventually derive informative priors on IA model parameters, we need to propagate the uncertainty in the Flagship IA model inherited from the statistical uncertainty in the calibration observations (and simulations), as well as from the finite volume of Flagship. Both of these sources of uncertainty are included in the χ^2 of the calibration fits; see Eq. (16). Propagated into the IA signals measured from the Flagship mock, these uncertainties can then be translated into the permitted range of the parameters of analytic IA models fitted to the mock datasets, which can serve as informative priors for cosmological inference. Full error propagation, e.g. via Monte Carlo sampling, is computationally expensive because at each step the semi-analytic model needs

Table 3. Best-fit parameters of Eq. (20), describing the empirical relation between IA amplitude and the scatter of the von Mises–Fisher distribution, σ_{MF} , for the observational galaxy samples (both central and satellites) used for calibration in this work.

sample	type	<i>a</i>	<i>b</i>	<i>c</i>
LOWZ L1	cen	0.01	1.10	-2.40
	sat	8.75	68.26	-2.33
LOWZ L2	cen	0.01	0.77	-2.42
	sat	-0.33	52.86	-2.38
LOWZ L3	cen	0.01	0.56	-2.16
	sat	2.40	54.57	-2.49
LOWZ L4	cen	0.01	0.47	-2.19
	sat	-0.62	44.53	-2.26
SDSS red	cen	0.00	0.15	-2.18
	sat	0.29	12.96	-2.29
SDSS blue	cen	0.00	0.15	-2.04
	sat	0.33	5.65	-2.32

to be optimised, and IA summary statistics re-measured. In this work we limit ourselves to order-of-magnitude estimates via an approximate scheme.

As shown in Fig. 10, to good approximation we can control the small-(large-)scale IA amplitude via the satellite (central) model parameters of the semi-analytic model, with the transition at around $5 h^{-1}\text{Mpc}$. We derive the covariance $\text{Cov}(\mathbf{p})$ of the fit parameters $p_{0.5}$ for both centrals and satellites from the distributions of χ^2_{tot} shown in Fig. 14. We linearly propagate the uncertainty on $p_{0.5}$ into the parameter controlling the misalignment distribution,

$$\text{Var}(\sigma_{\text{MF}}) = \sum_{i,j=0}^5 \frac{\partial \sigma_{\text{MF}}}{\partial p_i} \text{Cov}(\mathbf{p})_{ij} \frac{\partial \sigma_{\text{MF}}}{\partial p_j}, \quad (19)$$

where the derivatives are determined analytically from Eq. (15).

We generate mock realisations of the SDSS samples used for calibration, systematically varying σ_{MF} separately for central

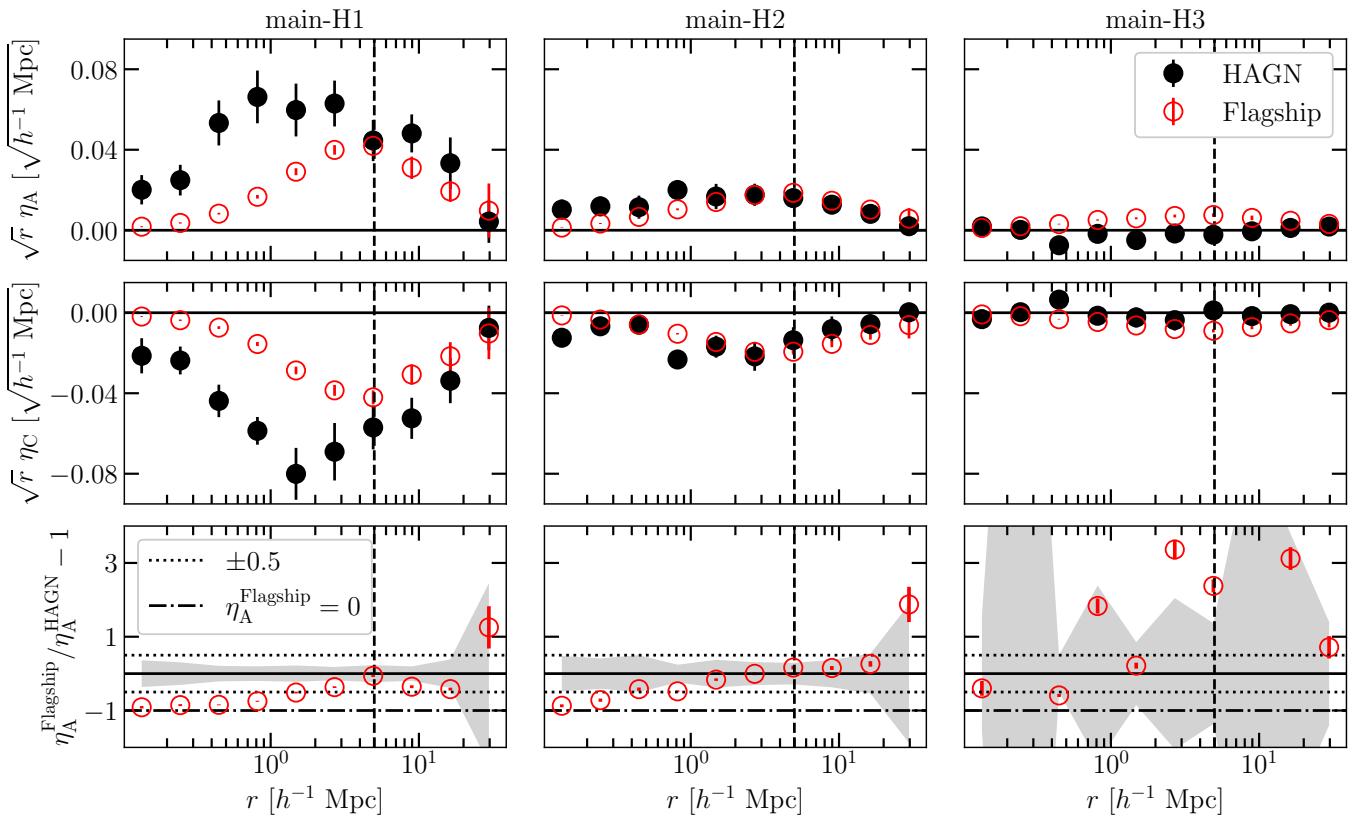


Fig. 16. Top: 3D correlations between galaxy major axis \mathbf{A} and the vector \mathbf{r} pointing to the position of a neighbouring galaxy (as defined in Eq. 10). Black and red symbols show measurements from the comoving output of HAGN at $z = 1.0$ and Flagship light-cone at the same redshift respectively. Results are shown for the magnitude limited samples H1 to H3 (from bright to dim), as defined in Table 1. Error bars indicate the Jack-knife estimates of the standard deviation. Center: Analogous to the top panel for the correlation of the galaxy minor axis \mathbf{C} . Bottom: Relative deviations of the Flagship measurements from HAGN. Grey areas indicate the standard deviations of the HAGN measurements. Dashed lines mark 100% deviations.

and satellites (while holding the other fixed) and measuring the amplitude of the resulting IA signal on scales above and below $5 h^{-1} \text{Mpc}$, which roughly separates the regimes where centrals and satellites drive the IA amplitude. The resulting relations are well fit by the empirical model

$$A_{\text{IA}} = a + b \exp(c\sigma_{\text{MF}}), \quad (20)$$

with the best-fit coefficients for the free parameters a, b, c per galaxy sample listed in Table 3. The offset a is generally small for all samples, while c shows only little variation with an average value around $c = -2.3$ because this scaling is a feature of the von Mises–Fisher distribution rather than a sample-dependent property. The parameter b scales with the luminosity of the sample, weakly for centrals, and approximately exponentially for satellites. Some residual luminosity trends are to be expected because the IA amplitude is not solely determined by either satellites or centrals, e.g. due to cross-correlations between centrals and satellites (cf. Fig. 10). Equations (19) and (20) can now be combined to propagate errors from the Flagship IA model to the amplitude of an intrinsic alignment correlation function, separately on small and large transverse scales. Through Eq. (15) this propagation depends on the galaxy sample’s typical redshift, luminosity, and colour.

Applying this scheme using typical values for the galaxy samples used for calibration, we find standard deviations on A_{IA} that exceed the parameter value, even for bright samples with strong signals and on both large and small transverse scales. This

is caused by the additional uncertainty due to the finite volume of the Flagship in Eq. (16). The two sources of statistical error in the denominator have roughly the same order of magnitude: while the real SDSS observations typically cover a larger volume, the Flagship measurements have been averaged over five shape noise realisations.

This means that, in order to obtain informative priors on IA parameters in the future, we need not only compile as many informative direct IA observations (and hydrodynamic simulation measurements) as possible over a wide range of galaxy sample properties, but also implement the IA models into simulations that exceed the constraining power of the observations by at least an order of magnitude to suppress additional statistical uncertainty beyond those of the observations to less than 10 %, e.g. by using multiple light-cone realisations. While the HOD parameters in Flagship are held fixed here, their scatter against observational calibrators is likely to constitute another significant source of uncertainty in the overall IA model, especially on small scales. In conclusion, we are currently unable to derive meaningful, informative prior constraints on the amplitude of IA signals in *Euclid*’s weak lensing samples directly from Flagship.

9. IA contamination in *Euclid*-like samples

As a first application of our semi-analytic IA model, we predict the expected contribution of IA signals to the observed cosmic shear signals in a *Euclid*-like survey. We select all galaxies

brighter than $I_E = 24.5$ and with valid photometric redshift estimates in the full octant of Flagship. The latter were obtained with a template-based method (see Paltani et al., in prep.) applied to the mock photometry of the near-infrared *Euclid* bands and LSST *ugriz* fluxes to a depth expected by the time of *Euclid*'s final data release. The resulting galaxy sample has a number density of 26 arcmin^{-2} on the sky, close to, but slightly higher than, *Euclid*'s nominal wide survey expectation.

We divide the sample into 13 equipopulated tomographic bins via the first mode of the photometric redshift probability density function. This choice ensures good redshift resolution and a fairly even distribution of statistical errors across the signals we extract. For every combination of tomographic bins we measure angular pseudo-power spectra of galaxy ellipticity correlations up to a maximum multipole of $\ell = 3000$. Details of the methodology are described in [Euclid Collaboration: Tessore et al. \(2025\)](#); briefly, the power spectra are calculated from Healpix ([Górski et al. 2005](#)) maps at a resolution of $N_{\text{side}} = 2048$, shape noise-subtracted, and not corrected for the effect of the limited footprint of the mock. Since data is available over a full octant, the mixing between multipoles, and between E - and B -modes, is limited to large scales and has minor impact, particularly in the power spectrum ratios that we consider.

We separate the contributions to the ellipticity correlations by the cosmological signal (i.e. the gravitational shear correlations, denoted by GG) and the combined signal from intrinsic ellipticity correlations (II) and shear-intrinsic cross-correlations (GI). Figure 17 shows the relative contribution of the combined II and GI signal to the total ellipticity correlation in each tomographic bin combination.

As expected, the B -mode contribution is consistent with zero throughout as our model is based on tidal, as well as radial satellite, alignments, which only generate B -modes at higher order ([Hirata & Seljak 2004](#); [Schneider & Bridle 2010](#)). In the E -mode, the IA contamination is fairly strong overall, reaching the 10 % level and beyond over a broad range of redshifts and angular scales. This result should be interpreted with caution, not only because of the expected uncertainties in the IA model parameters discussed in Sect. 8, but also owing to potential inaccuracies in the HOD model, for which we found indications when comparing the colour-magnitude distribution in Flagship with COSMOS observations (Sect. 3.1). Nevertheless, it is worth noting that our model predictions are of the same magnitude as the analytic forecasts made by [Fortuna et al. \(2021\)](#) who also considered a calibrated IA model for both central and satellite galaxies. Depending on modelling choices, they found a systematic shift in the parameter S_8 by up to 5 %, which roughly corresponds to an average bias of 10 % in power spectrum amplitude.

The figure shows that IA makes a positive contribution in tomographic auto-correlations due to the II term, and a negative contribution to tomographic cross-correlations where the GI term dominates. The GI term tends to be most significant when the tomographic bins are far apart in redshift so that the lensing efficiency of the foreground structures that align nearby galaxies is large. It is noteworthy that the opposing trends of the II and GI contributions are capable of breaking any degeneracy with shifts in the mean of the redshift distributions in the tomographic bins, which would move all signals coherently to larger or smaller amplitudes. This degeneracy breaking is not observed when binning more coarsely into fewer tomographic bins, as is done in current surveys mostly due to the lower overall galaxy number densities (see e.g. the discussion in [Dark Energy Survey and Kilo-Degree Survey Collaboration et al. 2023](#)). Nevertheless, cross-talk between residual uncertainty in the modelling of

IA and of tomographic redshift distributions will remain a major issue for *Euclid* and other Stage-IV surveys ([Fischbacher et al. 2023](#); [Leonard et al. 2024](#)).

10. Summary and conclusions

This work describes the addition of realistic intrinsic alignments (IA) of galaxies to *Euclid*'s Flagship simulation and mock galaxy catalogue ([Euclid Collaboration: Castander et al. 2025](#)), providing a test bed for the analysis choices and methods applied to forthcoming weak gravitational lensing analyses of *Euclid*.

Our semi-analytic IA model consists of two main components: firstly, we randomly sample galaxy ellipticities conditional on the redshift, luminosity, and colour in the simulation such that they match the corresponding distributions observed in the COSMOS Survey ([Scoville et al. 2007](#)). Second, to determine alignment, central galaxies follow the orientation of their host halo, while satellites align radially towards the halo centre. Unlike previous work (e.g. [Heymans et al. 2006](#); [Joachimi et al. 2013b](#); [Hoffmann et al. 2022b](#)), we apply the same model to both early- and late-type galaxies instead of aligning the minor axis of blue central galaxies with the halos' angular momentum vector, assuming they are rotationally supported discs. The strength of alignment is controlled separately for centrals and satellites via random misalignments, which depend via power laws on galaxy redshift, luminosity, and colour. To calibrate these dependencies, we employ constraints from multiple IA observations from the Sloan Digital Sky Survey (SDSS; [Singh & Mandelbaum 2016](#); [Johnston et al. 2019](#)), providing constraints up to redshift $z = 0.36$. Since *Euclid* weak lensing samples will reach $z \approx 2$ and beyond, we compensate for the current dearth of observational IA constraints at high redshifts by using measurements in three luminosity sub-samples from the Horizon AGN hydrodynamic simulation ([Dubois et al. 2014](#)) at $z = 1$. However, the latter samples are downweighted in the model optimisation as there is still considerable uncertainty in IA predictions from simulations.

Optimising a total of 12 alignment parameters, we found an overall good agreement between Flagship and the different reference data sets, with close matches in the distributions of axis ratios to all COSMOS samples and agreement in alignment statistics mostly within the 1σ errors. However, significant discrepancies occurred when comparing Flagship against the brightest Horizon AGN sub-sample on small scales. These deviations may result from the low weight that we assigned to Horizon AGN in the optimisation, or might indicate that the IA observations and Horizon AGN simulations are difficult to reconcile with a simple model, despite the long redshift baseline. A future, more detailed comparison of semi-analytic IA models against a range of hydrodynamic simulations will be required to shed light on these model limitations. Significant discrepancies also occur in the small-scale clustering of the brightest SDSS-LOWZ sub-sample. Since we exclude these scales from the IA parameter optimisation, low-level differences are also apparent in the corresponding alignment signals. Improvements here require a revision of the hybrid halo occupation and abundance matching approach for the Flagship galaxy mock.

For both centrals and satellites the model matches the constraints best for parameters that set a strong dependence of the galaxy misalignment on colour and a relatively weak dependence on magnitude and redshift, in agreement with the results from [Hoffmann et al. \(2022b\)](#) who employed a similar model. This result suggests that our IA model could potentially be simplified by eliminating the direct dependence on magnitude and

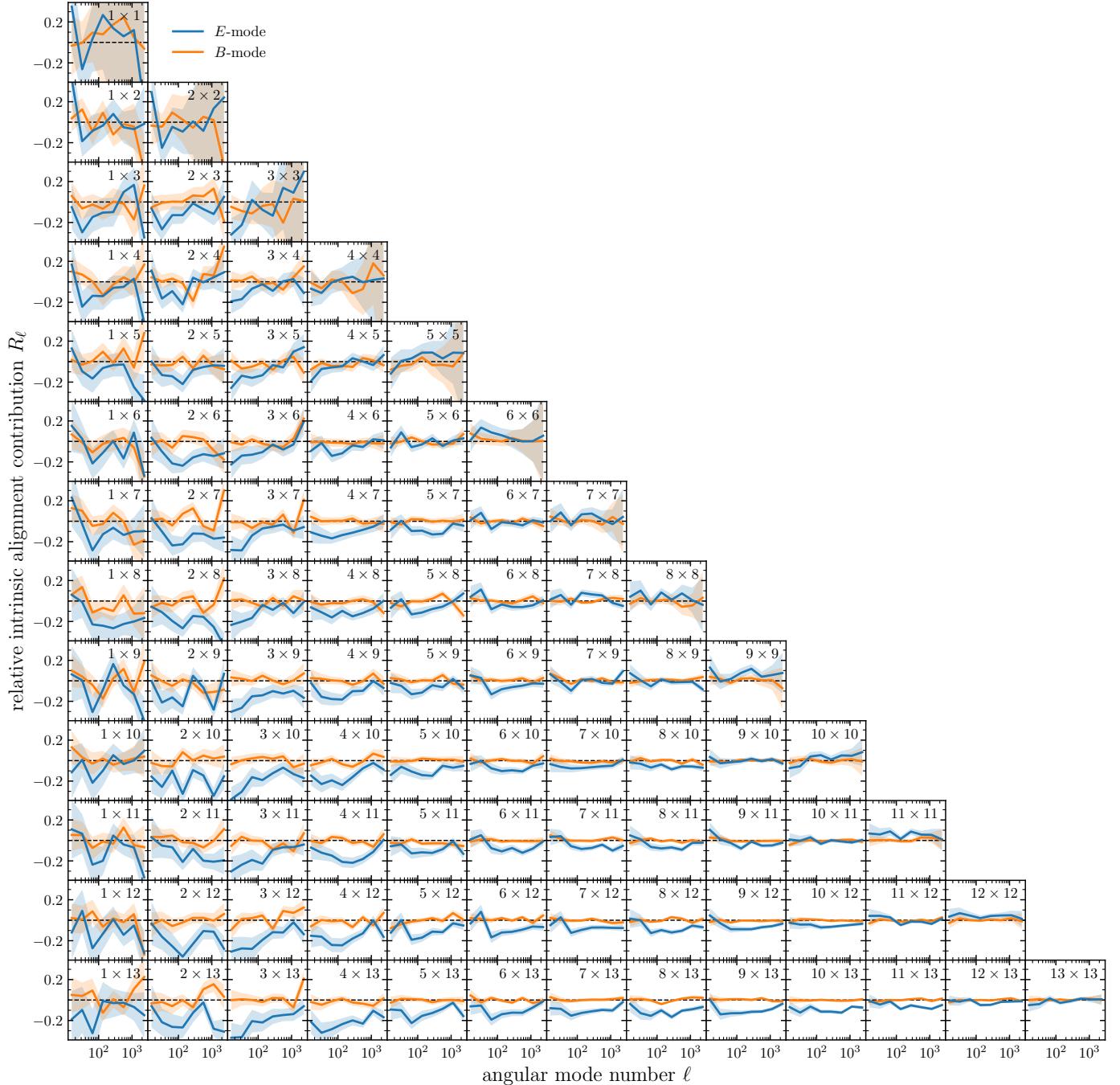


Fig. 17. Relative intrinsic alignment (IA) contribution $R_\ell^{ij} = C_{IA}^{ij}/(C_{IA}^{ij} + C_{GG}^{ij})$ to a tomographic cosmic shear analysis in a *Euclid*-like galaxy sample divided into 13 photometric redshift bins. Panels correspond to the different tomographic bin combinations, with auto-correlations along the diagonal and higher-redshift bins towards the bottom and right. Blue (orange) lines show the relative change to due intrinsic alignments to the E -(B)-mode angular shear power spectra. Bands around the signals indicate the standard deviations as determined from a Jackknife covariance.

redshift in the galaxy misalignment. Such a simplification would reduce the dimension of the model parameter space and remove parameter degeneracies. A simpler model is desirable to increase the predictiveness of the IA model given the limited number and constraining power of the calibration data sets.

Propagating the statistical uncertainty due to the calibration data and the limited precision of the Flagship mock measurements, we found the combined uncertainties on the IA amplitude currently still too large to provide informative bounds on IA parameter ranges. Nevertheless, we conclude that the best-fit calibrated IA model presented in this work is a useful tool for creat-

ing realistic mock IA data sets for a wide range of observed one- and two-point statistics of galaxy ellipticities. As a first application, we predicted the IA contamination to a tomographic weak lensing measurement akin to what can be expected for the final *Euclid* data release and found it to be significant, with $\sim 10\%$ IA contribution in some tomographic bin combinations, in line with earlier predictions.

A companion paper (Paviot et al., in prep.) will investigate how well common analytic IA models can fit the combined clustering and IA measurements from Flagship, informing some of the modelling choices for weak lensing inference from *Euclid*'s

first data release as well as comparing simulation predictions with observational samples not used in the calibration procedure. In line with forthcoming updates of the Flagship galaxy models, we will also improve the IA generation. This will include an expanded observational calibration data set, hydrodynamic simulation measurements from several projects (e.g. Nelson et al. 2019; Pakmor et al. 2023; Schaye et al. 2023) and over a wider redshift range, multiple lightcone realisations, and tests of current modelling assumptions such as the independence of galaxy shapes and orientations, and the spatial isotropy of misalignments.

Acknowledgements. KH acknowledges support by the Swiss National Science Foundation (Grant No. 173716, 198674), and from the Forschungskredit Grant of the University of Zurich (Projekt K-76106-01-01). The Euclid Consortium acknowledges the European Space Agency and a number of agencies and institutes that have supported the development of *Euclid*, in particular the Agenzia Spaziale Italiana, the Austrian Forschungsförderungsgesellschaft funded through BMIMI, the Belgian Science Policy, the Canadian Euclid Consortium, the Deutsches Zentrum für Luft- und Raumfahrt, the DTU Space and the Niels Bohr Institute in Denmark, the French Centre National d'Etudes Spatiales, the Fundação para a Ciéncia e a Tecnologia, the Hungarian Academy of Sciences, the Ministerio de Ciencia, Innovación y Universidades, the National Aeronautics and Space Administration, the National Astronomical Observatory of Japan, the Nederlandse Onderzoekschool Voor Astronomie, the Norwegian Space Agency, the Research Council of Finland, the Romanian Space Agency, the State Secretariat for Education, Research, and Innovation (SERI) at the Swiss Space Office (SSO), and the United Kingdom Space Agency. A complete and detailed list is available on the *Euclid* web site (www.euclid-ec.org/consortium/community/). This work has made use of CosmoHub, developed by PIC (maintained by IFAE and CIEMAT) in collaboration with ICE-CSIC. CosmoHub received funding from the Spanish government (MCIN/AEI/10.13039/501100011033), the EU NextGeneration/PRTR (PRTR-C17.II), and the Generalitat de Catalunya. This publication is part of the project “A rising tide: Galaxy intrinsic alignments as a new probe of cosmology and galaxy evolution” (with project number VI.Vidi.203.011) of the Talent programme Vidi which is (partly) financed by the Dutch Research Council (NWO).

References

- Allgood, B., Flores, R. A., Primack, J. R., et al. 2006, MNRAS, 367, 1781
 Asgari, M., Lin, C.-A., Joachimi, B., et al. 2021, A&A, 645, A104
 Aubert, D., Pichon, C., & Colombi, S. 2004, MNRAS, 352, 376
 Bailin, J. & Steinmetz, M. 2005, ApJ, 627, 647
 Bakx, T., Kurita, T., Elisa Chisari, N., Vlah, Z., & Schmidt, F. 2023, JCAP, 10, 005
 Bate, J., Chisari, N. E., Codis, S., et al. 2020, MNRAS, 491, 4057
 Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013, ApJ, 762, 109
 Bett, P. 2012, MNRAS, 420, 3303
 Bett, P., Eke, V., Frenk, C. S., et al. 2007, MNRAS, 376, 215
 Blanton, M. R., Hogg, D. W., Bahcall, N. A., et al. 2003, ApJ, 592, 819
 Blanton, M. R., Schlegel, D. J., Strauss, M. A., et al. 2005, AJ, 129, 2562
 Blazek, J. A., MacCrann, N., Troxel, M. A., & Fang, X. 2019, Phys. Rev. D, 100, 103506
 Carretero, J., Castander, F. J., Gaztañaga, E., Crocce, M., & Fosalba, P. 2015, MNRAS, 447, 646
 Carretero, J., Tallada, P., Casals, J., et al. 2017a, in Proceedings of the European Physical Society Conference on High Energy Physics. 5-12 July, 488
 Carretero, J., Tallada, P., Casals, J., et al. 2017b, in Proceedings of the European Physical Society Conference on High Energy Physics. 5-12 July, 488
 Catelan, P., Kamionkowski, M., & Blandford, R. D. 2001, MNRAS, 320, L7
 Chisari, N., Codis, S., Laigle, C., et al. 2015, MNRAS, 454, 2736
 Chisari, N., Laigle, C., Codis, S., et al. 2016, MNRAS, 461, 2702
 Chisari, N. E. 2025, arXiv e-prints, arXiv:2510.15738
 Chisari, N. E., Koukoufilippas, N., Jindal, A., et al. 2017, MNRAS, 472, 1163
 Codis, S., Gavazzi, R., Dubois, Y., et al. 2015, MNRAS, 448, 3391
 Codis, S., Jindal, A., Chisari, N. E., et al. 2018, MNRAS, 481, 4753
 Conselice, C. J., Blackburne, J. A., & Papovich, C. 2005, ApJ, 620, 564
 Dalal, R., Li, X., Nicola, A., et al. 2023, Phys. Rev. D, 108, 123519
 Dark Energy Survey and Kilo-Degree Survey Collaboration, Abbott, T. M. C., Aguena, M., et al. 2023, The Open Journal of Astrophysics, 6, 36
 Dubinski, J. & Carlberg, R. G. 1991, ApJ, 378, 496
 Dubois, Y., Pichon, C., Welker, C., et al. 2014, MNRAS, 444, 1453
 Dutton, A. A. & Macciò, A. V. 2014, MNRAS, 441, 3359
 Euclid Collaboration: Castander, F., Fosalba, P., Stadel, J., et al. 2025, A&A, 697, A5
 Euclid Collaboration: Mellier, Y., Abdurro'uf, Acevedo Barroso, J., et al. 2025, A&A, 697, A1
 Euclid Collaboration: Tessore, N., Joachimi, B., Loureiro, A., et al. 2025, A&A, 694, A141
 Fischbacher, S., Kacprzak, T., Blazek, J., & Refregier, A. 2023, JCAP, 1, 033
 Fisher, N. I., Lewis, T., & Embleton, B. J. J. 1993, Statistical Analysis of Spherical Data
 Fortuna, M. C., Hoekstra, H., Joachimi, B., et al. 2021, MNRAS, 501, 2983
 Fosalba, P., Gaztañaga, E., Castander, F. J., & Crocce, M. 2015, MNRAS, 447, 1319
 Fosalba, P., Gaztañaga, E., Castander, F. J., & Manera, M. 2008, MNRAS, 391, 435
 Georgiou, C., Chisari, N. E., Fortuna, M. C., et al. 2019, A&A, 628, A31
 Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
 Griffith, R. L., Cooper, M. C., Newman, J. A., et al. 2012, ApJS, 200, 9
 Herle, A., Chisari, N. E., Hoekstra, H., et al. 2025, arXiv e-prints, arXiv:2502.17135
 Heymans, C., Brown, M., Heavens, A., et al. 2004, MNRAS, 347, 895
 Heymans, C., White, M., Heavens, A., Vale, C., & van Waerbeke, L. 2006, MNRAS, 371, 750
 Hilbert, S., Barreira, A., Fabbian, G., et al. 2020, MNRAS, 493, 305
 Hirata, C. M. & Seljak, U. 2004, Phys. Rev. D, 70, 063526
 Hoffmann, K., Laigle, C., Chisari, N. E., et al. 2022a, MNRAS, 515, 3603
 Hoffmann, K., Secco, L. F., Blazek, J., et al. 2022b, Phys. Rev. D, 106, 123510
 Hu, W. 2000, Phys. Rev. D, 62, 043007
 Huertas-Company, M., Bernardi, M., Pérez-González, P. G., et al. 2016, MNRAS, 462, 4495
 Joachimi, B., Cacciato, M., Kitching, T. D., et al. 2015, Space Sci. Rev., 193, 1
 Joachimi, B., Mandelbaum, R., Abdalla, F. B., & Bridle, S. L. 2011, A&A, 527, A26
 Joachimi, B. & Schneider, P. 2009, A&A, 507, 105
 Joachimi, B., Semboloni, E., Bett, P. E., et al. 2013a, MNRAS, 431, 477
 Joachimi, B., Semboloni, E., Hilbert, S., et al. 2013b, MNRAS, 436, 819
 Johnston, H., Georgiou, C., Joachimi, B., et al. 2019, A&A, 624, A30
 Kaviraj, S., Laigle, C., Kimm, T., et al. 2017, MNRAS, 467, 4739
 Kirk, D., Brown, M. L., Hoekstra, H., et al. 2015, Space Sci. Rev., 193, 139
 Krause, E., Eifler, T., & Blazek, J. 2016, MNRAS, 456, 207
 Lackner, C. N. & Gunn, J. E. 2012, MNRAS, 421, 2277
 Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24
 Lamman, C., Tsaprazi, E., Shi, J., et al. 2024, The Open Journal of Astrophysics, 7, 14
 Landy, S. D. & Szalay, A. S. 1993, ApJ, 412, 64
 Leonard, C. D., Rau, M. M., & Mandelbaum, R. 2024, Phys. Rev. D, 109, 083528
 Mandelbaum, R., Hirata, C. M., Ishak, M., Seljak, U., & Brinkmann, J. 2006, MNRAS, 367, 611
 Nelson, D., Springel, V., Pillepich, A., et al. 2019, Computational Astrophysics and Cosmology, 6, 2
 Okumura, T., Jing, Y. P., & Li, C. 2009, ApJ, 694, 214
 Pakmor, R., Springel, V., Coles, J. P., et al. 2023, MNRAS, 524, 2539
 Pereira, M. J. & Bryan, G. L. 2010, ApJ, 721, 939
 Pereira, M. J., Bryan, G. L., & Gill, S. P. D. 2008, ApJ, 672, 825
 Pillepich, A., Springel, V., Nelson, D., et al. 2018, MNRAS, 473, 4077
 Piras, D., Joachimi, B., Schäfer, B. M., et al. 2018, MNRAS, 474, 1165
 Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
 Potter, D., Stadel, J., & Teyssier, R. 2017, Computational Astrophysics and Cosmology, 4, 2
 Reid, B., Ho, S., Padmanabhan, N., et al. 2016, MNRAS, 455, 1553
 Samuroff, S., Blazek, J., Troxel, M. A., et al. 2019, MNRAS, 489, 5453
 Schaye, J., Kugel, R., Schaller, M., et al. 2023, MNRAS, 526, 4978
 Schneider, M. D. & Bridle, S. 2010, MNRAS, 402, 2127
 Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
 Secco, L. F., Samuroff, S., Krause, E., et al. 2022, Phys. Rev. D, 105, 023515
 Semboloni, E., Hoekstra, H., Schaye, J., van Daalen, M. P., & McCarthy, I. G. 2011, MNRAS, 417, 2020
 Singh, S. & Mandelbaum, R. 2016, MNRAS, 457, 2301
 Singh, S., Mandelbaum, R., & More, S. 2015, MNRAS, 450, 2195
 Souissana, A., Chisari, N. E., Codis, S., et al. 2020, MNRAS, 492, 4268
 Tallada, P., Carretero, J., Casals, J., et al. 2020, Astronomy and Computing, 32, 100391
 Troxel, M. A. & Ishak, M. 2015, Phys. Rep., 558, 1
 Uzeirbegovic, E., Martin, G., & Kaviraj, S. 2022, MNRAS, 510, 3849
 Van Alfen, N., Campbell, D., Blazek, J., et al. 2024, The Open Journal of Astrophysics, 7, 45
 van Daalen, M. P., Schaye, J., Booth, C. M., & Dalla Vecchia, C. 2011, MNRAS, 415, 3649
 van den Bosch, F. C., Abel, T., Croft, R. A. C., Hernquist, L., & White, S. D. M. 2002, ApJ, 576, 21
 Velliscig, M., Cacciato, M., Schaye, J., et al. 2015, MNRAS, 454, 3328
 Vlah, Z., Chisari, N. E., & Schmidt, F. 2020, JCAP, 1, 025
 Welker, C., Power, C., Pichon, C., et al. 2017, arXiv:1712.07818

- Wright, A. H., Stölzner, B., Asgari, M., et al. 2025, arXiv e-prints, arXiv:2503.19441
 Xiao, T., Wang, T., Wang, H., et al. 2012, MNRAS, 421, 486
 Yao, J., Pedersen, E. M., Ishak, M., et al. 2020, MNRAS, 495, 3900
 York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579
 Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, ApJ, 736, 59

-
- ¹ Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
² Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191, Gif-sur-Yvette, France
³ Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
⁴ Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Avenida Complutense 40, 28040 Madrid, Spain
⁵ Port d'Informació Científica, Campus UAB, C. Albareda s/n, 08193 Bellaterra (Barcelona), Spain
⁶ Institute for Theoretical Physics, Utrecht University, Princetonplein 5, 3584 CE Utrecht, The Netherlands
⁷ Leiden Observatory, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands
⁸ Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona), Spain
⁹ Instituto de Astronomía Teórica y Experimental (IATE-CONICET), Laprida 854, X5000BGR, Córdoba, Argentina
¹⁰ Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, Stockholm, SE-106 91, Sweden
¹¹ Astrophysics Group, Blackett Laboratory, Imperial College London, London SW7 2AZ, UK
¹² Institut d'Estudis Espacials de Catalunya (IEEC), Edifici RDIT, Campus UPC, 08860 Castelldefels, Barcelona, Spain
¹³ Department of Physics, Northeastern University, Boston, MA, 02115, USA
¹⁴ Institut d'Astrophysique de Paris, UMR 7095, CNRS, and Sorbonne Université, 98 bis boulevard Arago, 75014 Paris, France
¹⁵ Kyung Hee University, Dept. of Astronomy & Space Science, Yongin-shi, Gyeonggi-do 17104, Republic of Korea
¹⁶ ESAC/ESA, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, 28692 Villanueva de la Cañada, Madrid, Spain
¹⁷ INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy
¹⁸ INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy
¹⁹ IFPU, Institute for Fundamental Physics of the Universe, via Beirut 2, 34151 Trieste, Italy
²⁰ INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34143 Trieste, Italy
²¹ INFN, Sezione di Trieste, Via Valerio 2, 34127 Trieste TS, Italy
²² SISSA, International School for Advanced Studies, Via Bonomea 265, 34136 Trieste TS, Italy
²³ Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
²⁴ INFN-Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
²⁵ Institut de Physique Théorique, CEA, CNRS, Université Paris-Saclay 91191 Gif-sur-Yvette Cedex, France
²⁶ Dipartimento di Fisica, Università di Genova, Via Dodecaneso 33, 16146, Genova, Italy
²⁷ INFN-Sezione di Genova, Via Dodecaneso 33, 16146, Genova, Italy
²⁸ Department of Physics "E. Pancini", University Federico II, Via Cinthia 6, 80126, Napoli, Italy
²⁹ INAF-Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131 Napoli, Italy
³⁰ Dipartimento di Fisica, Università degli Studi di Torino, Via P. Giuria 1, 10125 Torino, Italy
³¹ INFN-Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy
³² INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese (TO), Italy
³³ European Space Agency/ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
³⁴ Institute Lorentz, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands
³⁵ INAF-IASF Milano, Via Alfonso Corti 12, 20133 Milano, Italy
³⁶ INAF-Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monteporzio Catone, Italy
³⁷ INFN-Sezione di Roma, Piazzale Aldo Moro, 2 - c/o Dipartimento di Fisica, Edificio G. Marconi, 00185 Roma, Italy
³⁸ Institute for Theoretical Particle Physics and Cosmology (TTK), RWTH Aachen University, 52056 Aachen, Germany
³⁹ INFN section of Naples, Via Cinthia 6, 80126, Napoli, Italy
⁴⁰ Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
⁴¹ Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
⁴² Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain
⁴³ Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
⁴⁴ European Space Agency/ESRIN, Largo Galileo Galilei 1, 00044 Frascati, Roma, Italy
⁴⁵ Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, Villeurbanne, F-69100, France
⁴⁶ Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, 08028 Barcelona, Spain
⁴⁷ Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig de Lluís Companys 23, 08010 Barcelona, Spain
⁴⁸ UCB Lyon 1, CNRS/IN2P3, IUF, IP2I Lyon, 4 rue Enrico Fermi, 69622 Villeurbanne, France
⁴⁹ Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, PT1749-016 Lisboa, Portugal
⁵⁰ Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
⁵¹ Department of Astronomy, University of Geneva, ch. d'Ecogia 16, 1290 Versoix, Switzerland
⁵² Université Paris-Saclay, CNRS, Institut d'astrophysique spatiale, 91405, Orsay, France
⁵³ Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK
⁵⁴ INFN-Padova, Via Marzolo 8, 35131 Padova, Italy
⁵⁵ Aix-Marseille Université, CNRS/IN2P3, CPPM, Marseille, France
⁵⁶ INAF-Istituto di Astrofisica e Planetologia Spaziali, via del Fosso del Cavaliere, 100, 00100 Roma, Italy
⁵⁷ INFN-Bologna, Via Irnerio 46, 40126 Bologna, Italy
⁵⁸ University Observatory, LMU Faculty of Physics, Scheinerstr. 1, 81679 Munich, Germany
⁵⁹ Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, 85748 Garching, Germany
⁶⁰ INAF-Osservatorio Astronomico di Padova, Via dell'Osservatorio 5, 35122 Padova, Italy
⁶¹ Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstr. 1, 81679 München, Germany
⁶² Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, 0315 Oslo, Norway
⁶³ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, 91109, USA
⁶⁴ Felix Hormuth Engineering, Goethestr. 17, 69181 Leimen, Germany
⁶⁵ Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark
⁶⁶ Cosmic Dawn Center (DAWN), Denmark
⁶⁷ Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

- ⁶⁸ NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
⁶⁹ Department of Physics and Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, 00014 Helsinki, Finland
⁷⁰ Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland
⁷¹ Department of Physics, P.O. Box 64, University of Helsinki, 00014 Helsinki, Finland
⁷² Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, 00014 Helsinki, Finland
⁷³ Laboratoire d'étude de l'Univers et des phénomènes eXtremes, Observatoire de Paris, Université PSL, Sorbonne Université, CNRS, 92190 Meudon, France
⁷⁴ SKAO, Jodrell Bank, Lower Withington, Macclesfield SK11 9FT, UK
⁷⁵ Centre de Calcul de l'IN2P3/CNRS, 21 avenue Pierre de Coubertin 69627 Villeurbanne Cedex, France
⁷⁶ Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy
⁷⁷ INFN-Sezione di Milano, Via Celoria 16, 20133 Milano, Italy
⁷⁸ University of Applied Sciences and Arts of Northwestern Switzerland, School of Computer Science, 5210 Windisch, Switzerland
⁷⁹ Universität Bonn, Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany
⁸⁰ Aix-Marseille Université, CNRS, CNES, LAM, Marseille, France
⁸¹ Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, via Piero Gobetti 93/2, 40129 Bologna, Italy
⁸² Department of Physics, Institute for Computational Cosmology, Durham University, South Road, Durham, DH1 3LE, UK
⁸³ Université Paris Cité, CNRS, Astroparticule et Cosmologie, 75013 Paris, France
⁸⁴ CNRS-UCB International Research Laboratory, Centre Pierre Binétruy, IRL2007, CPB-IN2P3, Berkeley, USA
⁸⁵ Institut d'Astrophysique de Paris, 98bis Boulevard Arago, 75014, Paris, France
⁸⁶ Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
⁸⁷ Telespazio UK S.L. for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
⁸⁸ DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 155, 2200 Copenhagen, Denmark
⁸⁹ Space Science Data Center, Italian Space Agency, via del Politecnico snc, 00133 Roma, Italy
⁹⁰ Centre National d'Etudes Spatiales – Centre spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
⁹¹ Institute of Space Science, Str. Atomistilor, nr. 409 Măgurele, Ilfov, 077125, Romania
⁹² Consejo Superior de Investigaciones Científicas, Calle Serrano 117, 28006 Madrid, Spain
⁹³ Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain
⁹⁴ Dipartimento di Fisica e Astronomia "G. Galilei", Università di Padova, Via Marzolo 8, 35131 Padova, Italy
⁹⁵ Institut für Theoretische Physik, University of Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany
⁹⁶ Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, 31400 Toulouse, France
⁹⁷ Université St Joseph; Faculty of Sciences, Beirut, Lebanon
⁹⁸ Departamento de Física, FCFM, Universidad de Chile, Blanco Encalada 2008, Santiago, Chile
⁹⁹ Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, 6020 Innsbruck, Austria
¹⁰⁰ Satlantis, University Science Park, Sede Bld 48940, Leioa-Bilbao, Spain
¹⁰¹ Department of Physics, Royal Holloway, University of London, Surrey TW20 0EX, UK
¹⁰² Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal
¹⁰³ Cosmic Dawn Center (DAWN)
¹⁰⁴ Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark
¹⁰⁵ Universidad Politécnica de Cartagena, Departamento de Electrónica y Tecnología de Computadoras, Plaza del Hospital 1, 30202 Cartagena, Spain
¹⁰⁶ Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA
¹⁰⁷ Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy
¹⁰⁸ Istituto Nazionale di Fisica Nucleare, Sezione di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy
¹⁰⁹ INAF, Istituto di Radioastronomia, Via Piero Gobetti 101, 40129 Bologna, Italy
¹¹⁰ Astronomical Observatory of the Autonomous Region of the Aosta Valley (OAVdA), Loc. Lignan 39, I-11020, Nus (Aosta Valley), Italy
¹¹¹ Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice cedex 4, France
¹¹² Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH, UK
¹¹³ Zentrum für Astronomie, Universität Heidelberg, Philosophenweg 12, 69120 Heidelberg, Germany
¹¹⁴ ICL, Junia, Université Catholique de Lille, LITL, 59000 Lille, France
¹¹⁵ ICSC - Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing, Via Magnanelli 2, Bologna, Italy
¹¹⁶ Instituto de Física Teórica UAM-CSIC, Campus de Cantoblanco, 28049 Madrid, Spain
¹¹⁷ CERCA/ISO, Department of Physics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA
¹¹⁸ Technical University of Munich, TUM School of Natural Sciences, Physics Department, James-Franck-Str. 1, 85748 Garching, Germany
¹¹⁹ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany
¹²⁰ Laboratoire Univers et Théorie, Observatoire de Paris, Université PSL, Université Paris Cité, CNRS, 92190 Meudon, France
¹²¹ Departamento de Física Fundamental, Universidad de Salamanca, Plaza de la Merced s/n. 37008 Salamanca, Spain
¹²² Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 67000 Strasbourg, France
¹²³ Center for Data-Driven Discovery, Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan
¹²⁴ Waterloo Centre for Astrophysics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
¹²⁵ Dipartimento di Fisica - Sezione di Astronomia, Università di Trieste, Via Tiepolo 11, 34131 Trieste, Italy
¹²⁶ California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA
¹²⁷ Department of Physics & Astronomy, University of California Irvine, Irvine CA 92697, USA
¹²⁸ Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
¹²⁹ Departamento Física Aplicada, Universidad Politécnica de Cartagena, Campus Muralla del Mar, 30202 Cartagena, Murcia, Spain
¹³⁰ Instituto de Física de Cantabria, Edificio Juan Jordá, Avenida de los Castros, 39005 Santander, Spain
¹³¹ INFN, Sezione di Lecce, Via per Arnesano, CP-193, 73100, Lecce, Italy
¹³² Department of Mathematics and Physics E. De Giorgi, University of Salento, Via per Arnesano, CP-193, 73100, Lecce, Italy
¹³³ INAF-Sezione di Lecce, c/o Dipartimento Matematica e Fisica, Via per Arnesano, 73100, Lecce, Italy

- ¹³⁴ CEA Saclay, DFR/IRFU, Service d'Astrophysique, Bat. 709,
91191 Gif-sur-Yvette, France
- ¹³⁵ Institute of Cosmology and Gravitation, University of Portsmouth,
Portsmouth PO1 3FX, UK
- ¹³⁶ Department of Computer Science, Aalto University, PO Box
15400, Espoo, FI-00 076, Finland
- ¹³⁷ Instituto de Astrofísica de Canarias, E-38205 La Laguna; Universi-
dad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife,
Spain
- ¹³⁸ Caltech/IPAC, 1200 E. California Blvd., Pasadena, CA 91125,
USA
- ¹³⁹ Ruhr University Bochum, Faculty of Physics and Astronomy, As-
tronomical Institute (AIRUB), German Centre for Cosmological
Lensing (GCCL), 44780 Bochum, Germany
- ¹⁴⁰ Department of Physics and Astronomy, Vesilinnantie 5, University
of Turku, 20014 Turku, Finland
- ¹⁴¹ Serco for European Space Agency (ESA), Camino bajo del
Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de
la Cañada, 28692 Madrid, Spain
- ¹⁴² ARC Centre of Excellence for Dark Matter Particle Physics, Mel-
bourne, Australia
- ¹⁴³ Centre for Astrophysics & Supercomputing, Swinburne University
of Technology, Hawthorn, Victoria 3122, Australia
- ¹⁴⁴ Department of Physics and Astronomy, University of the Western
Cape, Bellville, Cape Town, 7535, South Africa
- ¹⁴⁵ DAMTP, Centre for Mathematical Sciences, Wilberforce Road,
Cambridge CB3 0WA, UK
- ¹⁴⁶ Kavli Institute for Cosmology Cambridge, Madingley Road, Cam-
bridge, CB3 0HA, UK
- ¹⁴⁷ Department of Astrophysics, University of Zurich, Winterthurer-
strasse 190, 8057 Zurich, Switzerland
- ¹⁴⁸ Department of Physics, Centre for Extragalactic Astronomy,
Durham University, South Road, Durham, DH1 3LE, UK
- ¹⁴⁹ IRFU, CEA, Université Paris-Saclay 91191 Gif-sur-Yvette Cedex,
France
- ¹⁵⁰ Univ. Grenoble Alpes, CNRS, Grenoble INP, LPSC-IN2P3, 53, Av-
enue des Martyrs, 38000, Grenoble, France
- ¹⁵¹ INAF-Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125,
Firenze, Italy
- ¹⁵² Dipartimento di Fisica, Sapienza Università di Roma, Piazzale
Aldo Moro 2, 00185 Roma, Italy
- ¹⁵³ Centro de Astrofísica da Universidade do Porto, Rua das Estrelas,
4150-762 Porto, Portugal
- ¹⁵⁴ Instituto de Astrofísica e Ciências do Espaço, Universidade do
Porto, CAUP, Rua das Estrelas, PT4150-762 Porto, Portugal
- ¹⁵⁵ HE Space for European Space Agency (ESA), Camino bajo del
Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de
la Cañada, 28692 Madrid, Spain
- ¹⁵⁶ INAF - Osservatorio Astronomico d'Abruzzo, Via Maggini, 64100,
Teramo, Italy
- ¹⁵⁷ Theoretical astrophysics, Department of Physics and Astronomy,
Uppsala University, Box 516, 751 37 Uppsala, Sweden
- ¹⁵⁸ Mathematical Institute, University of Leiden, Einsteinweg 55, 2333
CA Leiden, The Netherlands
- ¹⁵⁹ Institute of Astronomy, University of Cambridge, Madingley Road,
Cambridge CB3 0HA, UK
- ¹⁶⁰ Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRISAL, 59000
Lille, France
- ¹⁶¹ Department of Astrophysical Sciences, Peyton Hall, Princeton Uni-
versity, Princeton, NJ 08544, USA
- ¹⁶² Space physics and astronomy research unit, University of Oulu,
Pentti Kaiteran katu 1, FI-90014 Oulu, Finland
- ¹⁶³ Center for Computational Astrophysics, Flatiron Institute, 162 5th
Avenue, 10010, New York, NY, USA