Density split statistics:
joint model of counts and lensing in cells

O. Friedrich,[1,2,*] D. Gruen,[3,4,†] J. DeRose,[5,3] D. Kirk,[6] E. Krause,[3] T. McClintock,[7] E. S. Rykoff,[3,4] S. Seitz,[2,1]
R. H. Wechsler,[5,3,4] G. M. Bernstein,[8] J. Blazek,[9,10] C. Chang,[11] S. Hilbert,[12,13] B. Jain,[8] A. Kovacs,[14] O. Lahav,[6]
F. B. Abdalla,[6,15] S. Allam,[16] J. Annis,[16] K. Bechtol,[17] A. Benoit-Lévy,[18,6,19] E. Bertin,[18,19] D. Brooks,[6]
A. Carnero Rosell,[20,21] M. Carrasco Kind,[22,23] J. Carretero,[14] C. E. Cunha,[3] C. B. D'Andrea,[8] L. N. da Costa,[20,21]
C. Davis,[3] S. Desai,[24] H. T. Diehl,[16] J. P. Dietrich,[12,13] A. Drlica-Wagner,[16] T. F. Eifler,[25,26] P. Fosalba,[27] J. Frieman,[16,11]
J. García-Bellido,[28] E. Gaztanaga,[27] D. W. Gerdes,[29,30] T. Giannantonio,[31,32,1] R. A. Gruendl,[22,23] J. Gschwend,[20,21]
G. Gutierrez,[16] K. Honscheid,[9,33] D. J. James,[34] M. Jarvis,[8] K. Kuehn,[35] N. Kuropatkin,[16] M. Lima,[36,20] M. March,[8]
J. L. Marshall,[37] P. Melchior,[38] F. Menanteau,[22,23] R. Miquel,[39,14] J. J. Mohr,[13,12,2] B. Nord,[16] A. A. Plazas,[26] E. Sanchez,[40]
V. Scarpine,[16] R. Schindler,[4] M. Schubnell,[30] I. Sevilla-Noarbe,[40] E. Sheldon,[41] M. Smith,[42] M. Soares-Santos,[16]
F. Sobreira,[43,20] E. Suchyta,[44] M. E. C. Swanson,[23] G. Tarle,[30] D. Thomas,[45] M. A. Troxel,[9,33] V. Vikram,[46] and J. Weller[13,2,1]

(DES Collaboration)

[1]*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany*
[2]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany*
[3]*Kavli Institute for Particle Astrophysics & Cosmology,*
P. O. Box 2450, Stanford University, Stanford, CA 94305, USA
[4]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[5]*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*
[6]*Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK*
[7]*Department of Physics, University of Arizona, Tucson, AZ 85721, USA*
[8]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*
[9]*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*
[10]*Institute of Physics, Laboratory of Astrophysics, École Polytechnique Fédérale*
de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
[11]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*
[12]*Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany*
[13]*Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany*
[14]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of*
Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain
[15]*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa*
[16]*Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA*
[17]*LSST, 933 North Cherry Avenue, Tucson, AZ 85721, USA*
[18]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[19]*Sorbonne Universités, UPMC Univ Paris 06, UMR 7095,*
Institut d'Astrophysique de Paris, F-75014, Paris, France
[20]*Laboratório Interinstitucional de e-Astronomia - LIneA,*
Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
[21]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[22]*Department of Astronomy, University of Illinois, 1002 W. Green Street, Urbana, IL 61801, USA*
[23]*National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA*
[24]*Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India*
[25]*Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA*
[26]*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA*
[27]*Institute of Space Sciences, IEEC-CSIC, Campus UAB,*
Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
[28]*Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, 28049 Madrid, Spain*
[29]*Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA*
[30]*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[31]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[32]*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[33]*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
[34]*Astronomy Department, University of Washington, Box 351580, Seattle, WA 98195, USA*
[35]*Australian Astronomical Observatory, North Ryde, NSW 2113, Australia*
[36]*Departamento de Física Matemática, Instituto de Física,*
Universidade de São Paulo, CP 66318, São Paulo, SP, 05314-970, Brazil
[37]*George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy,*
and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA
[38]*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*
[39]*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*

[40]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain*
[41]*Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA*
[42]*School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK*
[43]*Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-859, Campinas, SP, Brazil*
[44]*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831*
[45]*Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK*
[46]*Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA*
(Dated: October 25, 2017)

We present density split statistics, a framework that studies lensing and counts-in-cells as a function of foreground galaxy density, thereby providing a large-scale measurement of both 2-point and 3-point statistics. Our method extends our earlier work on trough lensing and is summarized as follows: given a foreground (low redshift) population of galaxies, we divide the sky into subareas of equal size but distinct galaxy density. We then measure lensing around uniformly spaced points separately in each of these subareas, as well as counts-in-cells statistics (CiC). The lensing signals trace the matter density contrast around regions of fixed galaxy density. Through the CiC measurements this can be related to the density profile around regions of fixed matter density. Together, these measurements constitute a powerful probe of cosmology, the skewness of the density field and the connection of galaxies and matter.

In this paper we show how to model both the density split lensing signal and CiC from basic ingredients: a non-linear power spectrum, clustering hierarchy coefficients from perturbation theory and a parametric model for galaxy bias and shot-noise. Using N-body simulations, we demonstrate that this model is sufficiently accurate for a cosmological analysis on year 1 data from the Dark Energy Survey.

I. INTRODUCTION

The large-scale structure (LSS) observed today is thought to originate from almost perfectly Gaussian density perturbations in the early Universe. This means that there was a complete symmetry in the abundance and amplitude of underdense and overdense regions in very early times. Gravitational attraction then caused initial overdensities to collapse to small but highly overdense structures such as galaxy clusters, while initial underdensities expanded but stayed moderately underdense and e.g. became voids. As a consequence the majority of the volume in the late-time Universe is underdense, compensated by the presence of few highly overdense spots. Or, in other words, a positive skewness in the distribution of density fluctuations emerges due to gravitational collapse.

A variety of probes have been used to study the statistical properties of the late-time density field and to thereby understand the physics of gravitational collapse as well as the processes responsible for the properties of the initial density fluctuations. So far, the most extensive studies have been carried out on the 2-point statistics of density fluctuations, i.e. on measuring the variance of density fluctuations as a function of scale. This has e.g. been done through measurements of cosmic shear 2-point correlation functions [e.g. 6, 8, 34, 39, 44, 47, 62, 74, 77, 80], galaxy clustering [e.g. 1, 26, 31] and galaxy-galaxy lensing [e.g. 16, 17, 19, 41, 60, 79, 83] as well as combined measurements thereof [e.g. 27, 53, 84].

While 2-point statistics are only sensitive to the overall amplitude of density fluctuations, higher-order statistics also know about the skewness arising from the different behaviour of underdense and overdense regions. This does not necessarily mean that higher-order statistics are better than 2-point statistics in discriminating between particular choices of cosmological parameters [3]. But they scale differently with parameters such as Ω_m, σ_8, galaxy bias and galaxy stochasticity than their 2-point counterparts. Hence, in a cosmological analysis that varies a large number of parameters, probes that are sensitive to both 2-point and higher order statistics have the power to break degeneracies between these parameters [11, 58, 71, 75].

Observations of higher-order statistical features of the density field include measurements of three point correlation functions [63], shear peak statistics [46, 49, 50] and the cluster mass function [54]. Also, a number of probes have been suggested (and in some cases measured in data) that study the correlation of 2-point statistics and background density. Chiang et al. [18] have measured this by means of the integrated bispectrum. Simpson et al. [65, 66, 67] have proposed a clipped power spectrum approach, where 2-point statistics are measured on the sky after excluding high density regions. They have shown that these measurements contain information complementary to the corresponding measurements on the full sky.

A similar direction was investigated by Gruen et al. [36] who separately measured the lensing power spectrum in underdense and overdense lines of sight. The framework presented in this paper is based on their concept of trough lensing. We will call it *density split lensing* when only lensing measurements are involved and *density split statistics* when it is combined with counts-in-cells measurements. This method can be summarized as follows: we consider a foreground (low redshift) population of galaxies and smooth their position field with a circular top-hat aperture. This smoothed density field is then used to divide the sky into sub-areas of equal size but distinct galaxy density. In this paper we consider in particular 5 sub-areas and call them quintiles of galaxy density. As a next step, we use a background (high redshift) population of galax-

* Corresponding author: oliverf@usm.uni-muenchen.de
† Einstein Fellow

ies to measure the tangential shear of these galaxies around a set of uniformly spaced points within the area of each density quintile. The resulting lensing signals trace the matter density contrast around regions of fixed foreground galaxy density. This data vector is then complemented by the histogram of counts-in-cells of the foreground galaxies to pin down their bias and stochasticity. As we show in this paper, a cosmological analysis based on this density split data vector has a number of desirable features:

- it allows for an accurate analytic modeling with the help of cosmological perturbation theory and a non-linear power spectrum,

- it yields high signal-to-noise measurement,

- it avoids systematics common to cosmic shear such as additive shear biases or intrinsic alignments (as long as tracer sample and source sample do not overlap in redshift),

- it has a very intuitive interpretation.

This paper is a companion paper of Gruen et al. [37], where we present our actual data analysis, including tests for systematic effects as well as a description of how we estimate the covariance of our signal. This paper is presenting the modeling framework used in that analysis. Our section II gives a general overview of density split statistics: we describe our data vector, explain how it can be modelled and also present forecasts on cosmological parameter constraints, both for a ΛCDM model and an extended model that allows gravitational collapse to behave differently than within general relativity. In section III we describe the simulated data used in this work. Section IV explains details of the model presented in section II. There we also compare individual components of this model directly to measurements in N-body simulations. In section V we show that our model for a data vector combining density split lensing and counts-in-cells statistics is accurate enough to recover the cosmology underlying our N-body simulations. Any possible deviation between our model and the simulations is shown to be well within statistical uncertainties of year 1 data of the Dark Energy Survey (DES Y1). In appendix A, we review a number of differential equations that govern gravitational collapse. In appendix B, we review the leading order perturbative calculation of the 3-point statistics of the cosmic density field for a general ΛCDM model. And in appendix C we qualitatively compare our model of the cosmic density PDF to a second set of N-body simulations as a complement to the comparison carried out in the main text.

II. DENSITY SPLIT STATISTICS: DATA VECTOR, MODELING AND FORECASTS

This section provides an introduction to the program of density split statistics. In section II A we describe how we obtain the density split lensing signal and how this signal can be further complemented with information on galaxy bias and stochasticity from counts-in-cells. In section II B we outline our modeling of this signal (but postponing technical details

of this model to section IV). In section II C we provide forecasts on the cosmological information that can be obtained with a measurement of density split statistics in year-1 data of the Dark Energy Survey (DES Y1).

A. Measuring density split statistics

Density split lensing is a generalization of trough lensing [36] and can be described in three steps:

1.) Splitting the sky into quantiles of different foreground galaxy density

Consider a sample of low-redshift galaxies that are tracing the line-of-sight density of matter with some redshift distribution $n_l(z)$. We will call these galaxies the foreground sample. For an angular radius θ_T, which we will call the *top-hat aperture radius*, we define $N_T(\hat{\mathbf{n}})$ to be the number of galaxies found within a radius θ_T around the point on the sky specified by the vector $\hat{\mathbf{n}}$ on the unit sphere. The field $N_T(\hat{\mathbf{n}})$ can be used to divide the sky into regions of different galaxy density. Gruen et al. [36] have done this by discretizing the sky with a HEALPIX[1] grid and sorting the pixels according to their value of $N_T(\hat{\mathbf{n}})$. Then they considered the 20% of the pixels with the lowest values of N_T, calling them troughs. In the limit of a fine pixelization these pixels can be considered the most underdense quintile of the sky area. This can be generalized to the second most underdense quintile, the third most underdense quintile etc. or even to finer splits using more then just 5 quantiles.

We stick to a division into 5 quantiles (quintiles) throughout this paper. The upper panel of figure 1 illustrates such a subdivision on a patch of a simulated sky (from the *Buzzard flock*, see section III and especially DeRose et al. [28] for details). There we use a top-hat aperture radius $\theta_T = 20'$ and the tracer galaxies have the redshift distribution that is displayed by the solid line in Figure 2. Figure 1 shows the most underdense quintile of the simulated patch in cyan, the most overdense quintile in red and the three intermediate quintiles in blue, green and orange.

2.) Tracing the mean dark matter density in each sky quintile with gravitational lensing

Now consider a second sample of galaxies at higher redshifts than the foreground sample (the *source sample*, see e.g. the dashed and dotted redshift distributions in Figure 2). As the light of these galaxies passes the large-scale structure of the foreground density distribution it undergoes gravitational lensing effects such as gravitational shear [see e.g. 4]. The density split lensing signal around each quintile of the sky is obtained by measuring the stacked radial profile of tangential shear around random points located within that quintile.

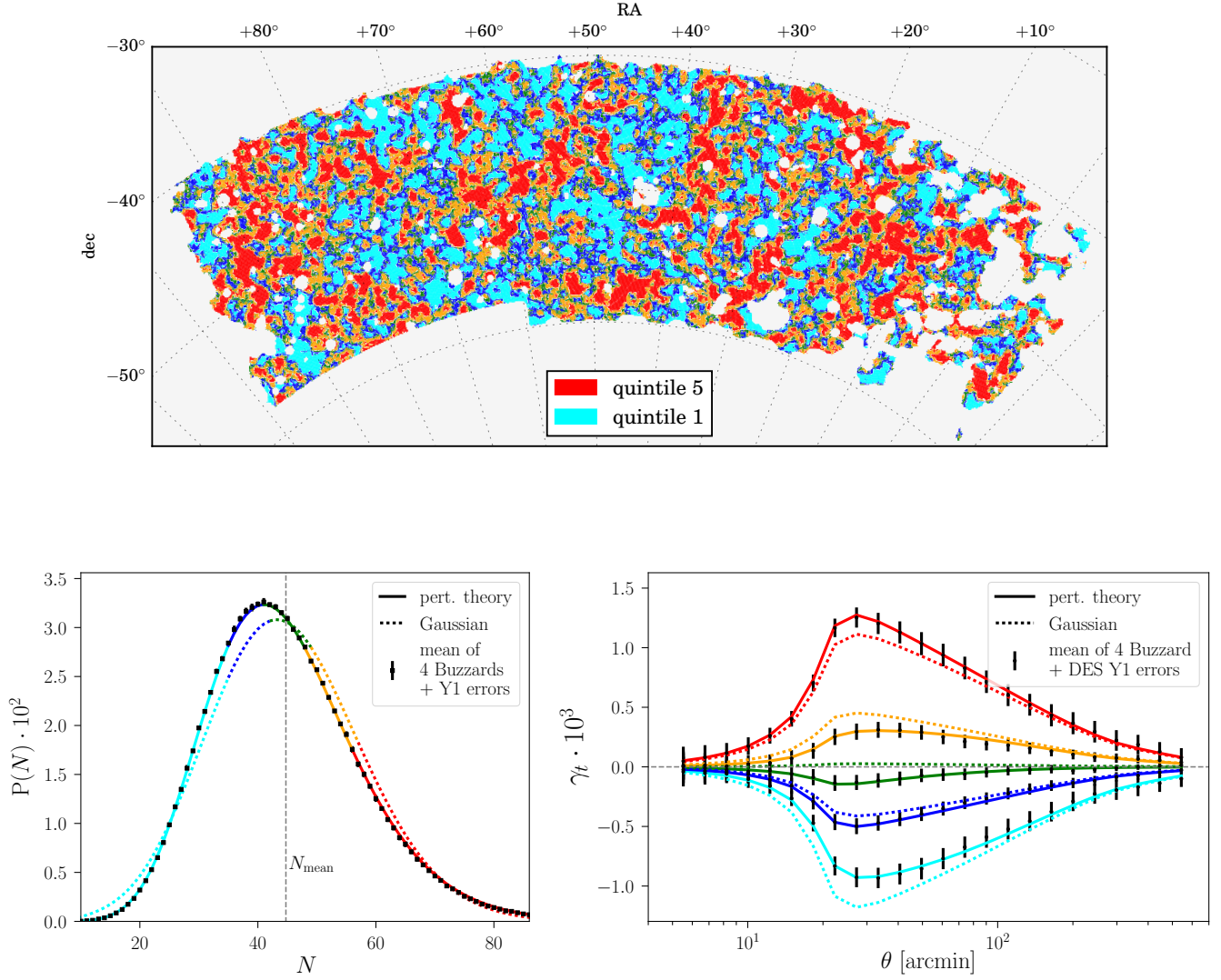[1] See Górski et al. [35] for details on HEALPIX.

FIG. 1. Top panel: splitting the lines of sight in one DES-Y1 like Buzzard simulation into 5 quantiles of galaxy density (color coding from cyan, most underdense, to red, most overdense). The map uses a 20 arcmin top-hat radius and REDMAGIC galaxies with a redshift range of $0.2 \lesssim z \lesssim 0.45$. Bottom left: histogram of REDMAGIC galaxy counts in 20 arcmin radii (counts-in-cells). We show the mean histogram from 4 Buzzard realisations of DES-Y1 (black points), our model based on perturbation theory and cylindrical collapse (solid line) and a model that assumes the projected density contrast to be a Gaussian random field (dotted line). The color coding corresponds exactly to the density quantiles in the top panel. Bottom right: Lensing signals around random points split by the density quantile in which these points are located. We show the mean measurement from 4 Buzzard realisations (black points), our perturbation theory model (solid line) and a model that assumes projected density contrast and lensing convergence to be joint Gaussian random variables (dotted line). Color coding is the same as in the other panels. The asymmetry between the lensing signals around the most underdense and most overdense lines-of-sight indicates the skewness of the cosmic density PDF.
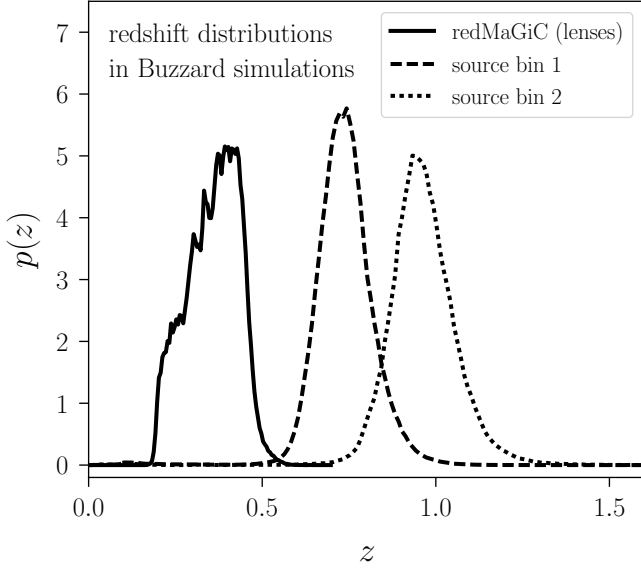
FIG. 2. Redshift distributions of the tracer galaxy sample and the source samples of our N-body realizations of DES-Y1.

These points are constrained to lie within the part of the sky covered by a certain quintile of galaxy density but are otherwise random in their location. Because these random points are split according to the density quintile they are located in, their stacked shear signals trace the average profile of density contrast around each quintile.

In the lower right panel of Figure 1 we show the signals measured for each density quintile in our mock data. The points show the average measurement from 4 Buzzard realisations of DES year-1 data (using the highest redshift source population shown in Figure 2) and the solid lines show predictions by the model presented in this paper. The error bars are derived from a set of log-normal realisations (using the FLASK tool by Xavier et al. [81]; in Gruen et al. [37] we describe in detail how we configured FLASK to generate our mock catalogs). Two main features of the density split lensing signals are apparent: first, the amplitude of the radial shear around the 20% most underdense pixels is lower than the amplitude of the tangential shear around the 20% most overdense pixels. This is reflecting the skewness of the cosmic density PDF. Secondly, the signal around points in the third quintile is still significantly negative, which reflects the fact that the median universe is underdense. A more subtle feature is the fact that the underdense signals fall off less rapidly with increasing scale than the overdense signals. This is because on large scales the density field becomes Gaussian and hence recovers its initial symmetry between overdensities and underdensities.

3.) Measuring the average counts-in-cells in each density quintile to obtain additional information on galaxy bias and stochasticity

If galaxy counts and the matter density field were perfectly correlated, then a split of the sky by galaxy density would be identical to a split by matter density. Hence, in this limit the density split lensing signals would be independent of the bias of the tracer galaxies. In a realistic scenario however, shot-noise of the galaxies smears out our attempts to divide the sky into areas of different matter density. Hence the density split lensing signals obtain a dependence on galaxy bias, but also on galaxy stochasticity. Increasing the linear bias of galaxy clustering will sharpen the tracers' ability to distinguish between overdensities and underdensities. Thus, increasing this bias will increase the amplitudes of the signals. This means that linear bias is to some degree degenerate with the amplitude of density fluctuations, σ_8. But σ_8 and bias influence the third moments of the density field in different ways and their degeneracy is not complete. As a consequence, it is possible to obtain constraints on cosmological parameters from the lensing signals alone (cf. section II C and the blue contour in the left panel of figure 3).

But additional information on bias and stochasticity nevertheless helps to tighten these constraints. In this paper we decided to add that information in the form of normalized quantiles of the counts-in-cells (CiC) histogram of the tracer galaxies: we measure the histogram of tracer counts within the same aperture that was used to identify our density quintiles. Then we identify the parts of this histogram that correspond to these quintiles (cf. lower left panel of Figure 1). For each quintile q we then determine the mean galaxy count in that quintile, N_q, and normalize it by the overall mean galaxy count in our aperture, $\bar{N}$. i.e. for each quintile we add $N_q/\bar{N}$ to our data vector. This indeed helps to tighten constraints on cosmological parameters (cf. section II C and the green contour in the left panel of figure 3).

B. Modelling density split statistics

We now outline a general framework for modeling the data vector described above, leaving details of this framework to section IV. Unless stated differently, we will assume a flat ΛCDM universe throughout this paper.

Let us start by introducing the quantities whose relations need to be modelled. First, we denote with $\delta_{m,2D}$ the line-of-sight projection of the 3D density contrast according to the redshift distribution $n_l(z)$ of our foreground galaxy sample, i.e.

$$\delta_{m,2D}(\hat{\mathbf{n}}) = \int dw \, q_l(w) \, \delta_{m,3D}(w\hat{\mathbf{n}}, w) \qquad (II.1)$$

where $\hat{\mathbf{n}}$ denotes a unit vector on the sky, w is co-moving distance and the projection kernel $q_l(w)$ is given in terms of $n_l(z)$ as

$$q_l(w) = n_l(z[w]) \frac{dz[w]}{dw} . \qquad (II.2)$$

We furthermore define $\delta_{m,T}$ to be the average of $\delta_{m,2D}$ over top-hat filters with aperture radius θ_T, i.e.

$$\delta_{m,T}(\hat{\mathbf{n}}) = \int_{|\hat{\mathbf{n}},\hat{\mathbf{n}}'|<\theta_T} d\Omega' \, \frac{\delta_{m,2D}(\hat{\mathbf{n}}')}{2\pi(1-\cos\theta_T)} . \qquad (II.3)$$

Here $|\cdot, \cdot|$ denotes the angular distance between two points on the sky.

We identify regions of different density by means of our foreground galaxy sample. When smoothed with a top-hat filter of radius θ_T, these galaxies are biased and possibly stochastic tracers of $\delta_{m,T}(\hat{\mathbf{n}})$. Hence our model also needs to include a description of how $N_T(\hat{\mathbf{n}})$, the number of tracer galaxies found within an angular radius θ_T around the line-of-sight $\hat{\mathbf{n}}$, relates to $\delta_{m,T}(\hat{\mathbf{n}})$.

Finally, in order to describe the density split lensing signal, we need to consider the lensing convergence field for our population of source galaxies. Given the source redshift distribution $n_s(z)$, the convergence κ is given by the line-of-sight projection

$$\kappa(\hat{\mathbf{n}}) = \int \mathrm{d}w \, W_s(w) \, \delta_{m,\mathrm{3D}}(w\hat{\mathbf{n}}, w) \,, \qquad \text{(II.4)}$$

where W_s is the lensing efficiency, which is defined by

$$W_s(w) = \frac{3\Omega_m H_0^2}{2c^2} \int_w^\infty \mathrm{d}w' \, \frac{w(w'-w)}{w' \, a(w)} \, q_s(w') \,, \qquad \text{(II.5)}$$

and $q_s(w) = n_s(z[w])\frac{\mathrm{d}z[w]}{\mathrm{d}w}$ is the line-of-sight density of the sources. Smoothing the convergence field with a circular aperture of radius θ results in a field which we will denote by $\kappa_{<\theta}(\hat{\mathbf{n}})$.

Because of the isotropy of the universe, we will now omit the the the dependence of the above quantities on $\hat{\mathbf{n}}$. To model the density split lensing signal one needs to answer the following questions:

• Given the number of galaxies N_T found around a line-of-sight $\hat{\mathbf{n}}$, what distribution can be inferred for the matter density contrast $\delta_{m,T}$ in that line-of-sight? i.e. what is the expectation value $\langle \delta_{m,T}|N_T \rangle$?

• Given the matter density contrast $\delta_{m,T}$ in the line-of-sight $\hat{\mathbf{n}}$, what lensing convergence $\kappa_{<\theta}$ is expected inside an angular distance θ from that line-of-sight? i.e. what is the expectation value $\langle \kappa_{<\theta}|\delta_{m,T} \rangle$? The tangential shear profile around that line-of-sight can then be inferred from the convergence profile as

$$\begin{aligned} \langle \gamma_t(\theta)|\delta_{m,T} \rangle &= \langle \kappa_{<\theta}|\delta_{m,T} \rangle - \langle \kappa_\theta|\delta_{m,T} \rangle \\ &= \frac{\cos\theta - 1}{\sin\theta} \frac{\mathrm{d}}{\mathrm{d}\theta} \langle \kappa_{<\theta}|\delta_{m,T} \rangle \,, \qquad \text{(II.6)} \end{aligned}$$

where κ_θ is the average convergence at the radius θ.

The first of the above questions can be answered in the form of a conditional PDF of $\delta_{m,T}$ given a certain value of N_T, i.e. $p(\delta_{m,T}|N_T)$. Using Bayes' theorem this can be written as

$$p(\delta_{m,T}|N_T) = \frac{P(N_T|\delta_{m,T}) \, p(\delta_{m,T})}{P(N_T)} \,, \qquad \text{(II.7)}$$

where $P(N_T|\delta_{m,T})$ is the probability of finding a number of galaxies N_T given that the density contrast is $\delta_{m,T}$ and where $p(\delta_{m,T})$ and $P(N_T)$ are the total PDF of $\delta_{m,T}$ and the total

probability of finding N_T tracer galaxies. The average convergence profile around a circle with N_T galaxies is then given by

$$\begin{aligned} \langle \kappa_{<\theta}|N_T \rangle &= \int \mathrm{d}\delta_{m,T} \, \langle \kappa_{<\theta}|\delta_{m,T}, N_T \rangle \, p(\delta_{m,T}|N_T) \\ &\approx \int \mathrm{d}\delta_{m,T} \, \langle \kappa_{<\theta}|\delta_{m,T} \rangle \, p(\delta_{m,T}|N_T) \,, \quad \text{(II.8)} \end{aligned}$$

where in the second step we have assumed that the expected convergence within θ only depends on the total matter density contrast within θ_T.

We now divide the sky into different quintiles of tracer galaxy density. Let us denote with $Q[0.0, 0.2]$ the 20% of the lines-of-sight on the sky that have the lowest value of N_T. There will be a maximal value $N_T \leq N_{\max}$ in this quintile and the stacked convergence profile around these lines-of-sight is given by

$$\langle \kappa_{<\theta}|Q[0.0, 0.2] \rangle =$$
$$\frac{1}{0.2} \left(\sum_{N < N_{\max}} P(N)\langle \kappa_{<\theta}|N_T = N \rangle + \alpha \, \langle \kappa_{<\theta}|N_T = N_{\max} \rangle \right) .$$
$$\text{(II.9)}$$

Here the factor α in the second term accounts for the fact that the lines-of-sight with exactly $N_T = N_{\max}$ might have to be split between the quintile $Q[0.0, 0.2]$ and the quintile $Q[0.2, 0.4]$. It is given by

$$\alpha = 0.2 - \sum_{N < N_{\max}} P(N) \,. \qquad \text{(II.10)}$$

This can be easily generalized to the other quintiles $Q[q_{\min}, q_{\max}]$ and also to the case of dividing the sky into more than 5 density regimes. Finally, we also add the average of the counts-in-cells in each quintile normalized by the mean galaxy count $\bar{N}$ to our data vector. For the quintile $Q[0.0, 0.2]$ this is given by

$$\frac{\langle N_T|Q[0.0, 0.2] \rangle}{\bar{N}} = \frac{1}{0.2\bar{N}} \left(\sum_{N < N_{\max}} P(N) \, N + \alpha \, N_{\max} \right) , \qquad \text{(II.11)}$$

which is also straightforward to generalize to other quantiles $Q[q_{\min}, q_{\max}]$.

The probabilities $P(N)$ can be computed from the normalization of equation II.7. Hence, our model for the density split lensing signal needs the following three ingredients:

(i) The PDF of matter density contrast, smoothed with a top-hat filter of radius θ_T,

$$p(\delta_{m,T}) \,. \qquad \text{(II.12)}$$

(ii) The expectation value of convergence inside a distance θ given the density contrast inside θ_T,

$$\langle \kappa_{<\theta}|\delta_{m,T} \rangle \,. \qquad \text{(II.13)}$$

(iii) The distribution of galaxy counts inside the top-hat radius θ_T given the density contrast within that radius,

$$P(N_T|\delta_{m,T}) . \qquad \text{(II.14)}$$

Gruen et al. [36] assumed $\delta_{m,T}$ and $\kappa_{<\theta}$ to have a joint Gaussian distribution. This allowed them to compute (i) and (ii) solely from the dark matter clustering power spectrum. To compute (iii) they assumed a linear galaxy bias and Poissonian shot-noise of the tracer galaxies. These assumptions allowed a sufficient model for their measurements made on DES Science Verification data. But as can be seen from the dotted lines in the lower panels of Figure 1, a Gaussian model for the density PDF is not sufficient within the much smaller uncertainty of DES-Y1. Also, in section IV we demonstrate that the shot-noise of the tracer galaxies can not necessarily be assumed to be Poissonian. In this work we hence want to revise their model.

For each of the model components (i) and (ii) we investigate two different modeling approaches - a baseline approach and an approach with increased complexity used to check the validity of the baseline model. In the following we briefly outline each approach. The reader interested in details of each modeling ansatz is referred to section IV. Readers who are not interested in this technical part of the paper should feel free to skip section IV.

(i) <u>Baseline model for $p(\delta_{m,T})$:</u>
In our fiducial model we assume $\delta_{m,T}$ to be a log-normal random field [38]. The PDF of such a variable can e.g. be fixed by specifying the variance $\langle\delta_{m,T}^2\rangle$ and skewness $\langle\delta_{m,T}^3\rangle$. We predict the variance of $\delta_{m,T}$ from the the non-linear matter power spectrum [cf. 36]. The latter is computed using *halofit* [68, 72] and an analytic transfer function [30]. We then use leading order perturbation theory to compute a scaling relation between the bispectrum and the power spectrum of the density field. Together with our power spectrum this fixes the skewness of $\delta_{m,T}$.

<u>Alternative model for $p(\delta_{m,T})$:</u>
As an alternative we compute the PDF $p(\delta_{m,T})$ from its cumulant generating function. To model this function we use a cylindrical collapse approach based on the work of Bernardeau [9], Bernardeau & Valageas [10] and Valageas [76].

(ii) <u>Baseline model for $\langle\kappa_{<\theta}|\delta_{m,T}\rangle$:</u>
In our fiducial model we assume that $\kappa_{<\theta}$ can be expressed as the sum of two random variables,

$$\kappa_{<\theta} = \kappa_{<\theta,\text{corr.}} + \kappa_{<\theta,\text{uncorr.}} , \qquad \text{(II.15)}$$

where $\kappa_{<\theta,\text{uncorr.}}$ is assumed to be completely uncorrelated to $\delta_{m,T}$ and hence doesn't contribute to the density split lensing signal. As a result we have

$$\langle\kappa_{<\theta}|\delta_{m,T}\rangle \equiv \langle\kappa_{<\theta,\text{corr.}}|\delta_{m,T}\rangle . \qquad \text{(II.16)}$$

We assume a joint log-normal PDF for the two random variables $\delta_{m,T}$ and $\kappa_{<\theta,\text{corr.}}$. The expectation value $\langle\kappa_{<\theta,\text{corr.}}|\delta_{m,T}\rangle$ is then fixed by specifying the moments

$$\langle\delta_{m,T}^2\rangle , \ \langle\delta_{m,T}^3\rangle \qquad \text{(II.17)}$$

as well as

$$\langle\kappa_{<\theta}\delta_{m,T}\rangle \equiv \langle\kappa_{<\theta,\text{corr.}}\delta_{m,T}\rangle \qquad \text{(II.18)}$$

and

$$\langle\kappa_{<\theta}\delta_{m,T}^2\rangle \equiv \langle\kappa_{<\theta,\text{corr.}}\delta_{m,T}^2\rangle . \qquad \text{(II.19)}$$

Second order moments are again computed from our non-linear power spectrum while third order moments are inferred from perturbation theory. (The introduction of $\kappa_{<\theta,\text{uncorr.}}$ is only necessary for consistency reasons: a joint log-normal PDF of $\delta_{m,T}$ and $\kappa_{<\theta}$ characterized by the above moments would be inconsistent with the variance $\langle\kappa_{<\theta}^2\rangle$ derived from our power spectrum.)

<u>Alternative model for $\langle\kappa_{<\theta}|\delta_{m,T}\rangle$:</u>
As an alternative we compute $\langle\kappa_{<\theta}|\delta_{m,T}\rangle$ from the joint cumulant generating function of the variables $\delta_{m,T}$ and $\kappa_{<\theta}$. This function can also be modelled in a cylindrical collapse approach.

For model component (iii) we also investigate two different modeling approaches - one ansatz introducing 2 free parameters and one ansatz introducing 3 free parameters. We find that our simulated tracer catalogs are well described by the 2-parametric model. But - anticipating real galaxies to behave more complicated then simulated ones - we do not consider either of these models as our baseline model and instead apply both approaches to DES data in Gruen et al. [37]. We summarize both ansatzes here, but the interested reader is again referred to section IV for details of each ansatz.

(iii) <u>Model 1 for $P(N_T|\delta_{m,T})$:</u>
In our fiducial model we introduce an auxilliary field $\delta_{g,T}$ such that our foreground galaxies are Poissonian tracers of that field, i.e.

$$P(N_T|\delta_{g,T}) = \frac{\left[\bar{N}(1+\delta_{g,T})\right]^{N_T}}{N_T!} e^{-\bar{N}(1+\delta_{g,T})} . \qquad \text{(II.20)}$$

$\delta_{g,T}$ can be thought of as a smooth (shot noise free) galaxy density contrast. We then assume that $\delta_{g,T}$ and $\delta_{m,T}$ are joint log-normal random variables with

$$\langle\delta_{g,T}^2\rangle = b^2\langle\delta_{m,T}^2\rangle , \ \langle\delta_{g,T}^3\rangle = b^3\langle\delta_{m,T}^3\rangle \qquad \text{(II.21)}$$

and

$$\langle\delta_{g,T}\delta_{m,T}\rangle = br\langle\delta_{m,T}^2\rangle . \qquad \text{(II.22)}$$

The parameters b and r will be called *galaxy bias* and *galaxy stochasticity* and are free parameters of the model.

Model 2 for $P(N_T|\delta_{m,T})$:

As an alternative we assume $P(N_T|\delta_{m,T})$ to be a generalization of the Poisson distribution, that allows for

$$\langle N_T^2|\delta_{m,T}\rangle \neq \langle N_T|\delta_{m,T}\rangle + \langle N_T|\delta_{m,T}\rangle^2 , \qquad \text{(II.23)}$$

i.e. for a shot-noise that is either enhanced or suppressed wrt. the Poisson case. The enhancement of shot-noise is also allowed to be a function of $\delta_{m,T}$ of (approximately) the form

$$\frac{\langle N_T^2|\delta_{m,T}\rangle - \langle N_T|\delta_{m,T}\rangle^2}{\langle N_T|\delta_{m,T}\rangle} \approx \alpha_0 + \alpha_1\delta_{m,T} . \qquad \text{(II.24)}$$

This model introduces an alternative bias parameter $\tilde{b}$ such that

$$\langle N_T|\delta_{m,T}\rangle = \bar{N}[1 + \tilde{b}\delta_{m,T}] . \qquad \text{(II.25)}$$

For the model components (i) and (ii) and on the scales considered in this paper, the baseline and alternative modeling approaches yield almost indistinguishable predictions (cf. section IV). For component (iii) the modeling ansatzes 1 and 2 are not necessarily identical, because they introduce a different number of degrees of freedom. Figure 1 as well as all parameter contours shown in this paper are using the baseline model for components (i) and (ii) and ansatz 1 for component (iii). The predictions derived from different modelling approaches are however almost indistinguishable (cf. figure 4).

C. Data vector and forecasts on parameter constraints

1. Binning and scales

Throughout this paper, we assume that one sample of tracer galaxies is used to identify density quintiles and that the lensing profiles around these quintiles are measured with two source redshift bins (cf. redshift distributions in Figure 2). To identify the different density quintiles, we use a top-hat filter with fiducial smoothing radius of $\theta_T = 20'$.

We measure the density split lensing signal in 24 log-spaced angular bins with

$$5' < \theta < 600' . \qquad \text{(II.26)}$$

But in the parameter forecasts and likelihood contours shown in the following, we exclude bins with scales $\leq \theta_T$. This reduces all lensing signals to 17 log-spaced angular bins with

$$20' \lesssim \theta < 600' . \qquad \text{(II.27)}$$

The scales with $\theta < \theta_T$ are excluded from our analysis for a number of reasons: first, the signal-to-noise ratio of our lensing profiles drops quickly below θ_T (cf. Figure 1). Second, we chose our fiducial value of $\theta_T = 20'$ because we still trust the

modeling of the density PDF described in section IV on this scale, and we do not want smaller angular scales to contribute to our fiducial data vector. Third, the approximation made in equation II.8 might fail at scales smaller than our aperture.

The sum of all 5 lensing signals will be very close to zero (though not exactly zero because of masking effects), so that they are not an independent set of observations. Hence, we only include the 2 most overdense and the 2 most underdense quintiles in our analysis, i.e. the cyan, blue, orange and red lines in Figure 1. For the same reason, we also use the normalized mean galaxy count only in four of the five quintiles of the counts-in-cells histogram to complement the lensing measurement. The total number of data points in our data vector is thus

$$\begin{aligned} N_{\text{data}} &= N_{\text{lens}} + N_{\text{CiC}} \\ &= [N_{\text{quant.}} - 1] \cdot N_{\text{source}} \cdot N_{\text{ang}} + [N_{\text{quant.}} - 1] \\ &= 4 \cdot 2 \cdot 17 + 4 \\ &= 140 . \qquad \text{(II.28)} \end{aligned}$$

2. Model parameters and forecast on constraining power

We now investigate what constraints on model parameters can be expected from the above data vector when measured in DES-Y1 data. For this, we assume the optimistic case that component (ii) of our model is sufficiently described by two parameters, i.e. by modeling approach 1 in the previous section.[2] This means that our model is determined by the following 7 parameters:

1.) Ω_m: present total matter density in units the critical density of the universe,

2.) σ_8: amplitude of present day density fluctuations in spheres of $8\text{Mpc}/h$ radius as predicted by the linear power spectrum,

3.) Ω_b: present density of baryonic matter in units of the critical density of the universe,

4.) n_s: the spectral index of the primordial power spectrum,

5.) h_{100}: the present day Hubble parameter in units of 100Mpc/(km/s),

6.) b: linear bias of the tracers w.r.t. matter density (cf. equation II.21),

7.) r: correlation coefficient between $\delta_{m,T}$ and $\delta_{g,T}$ (cf. equation II.22).

We summarize these parameters and our fiducial values for them in Table I. Throughout this paper, we assume the universe to be flat.

If π_α and π_β are any two of the above parameters then let $\hat{\pi}_{\alpha,\text{ML}}$ and $\hat{\pi}_{\beta,\text{ML}}$ be maximum likelihood estimates of these

———

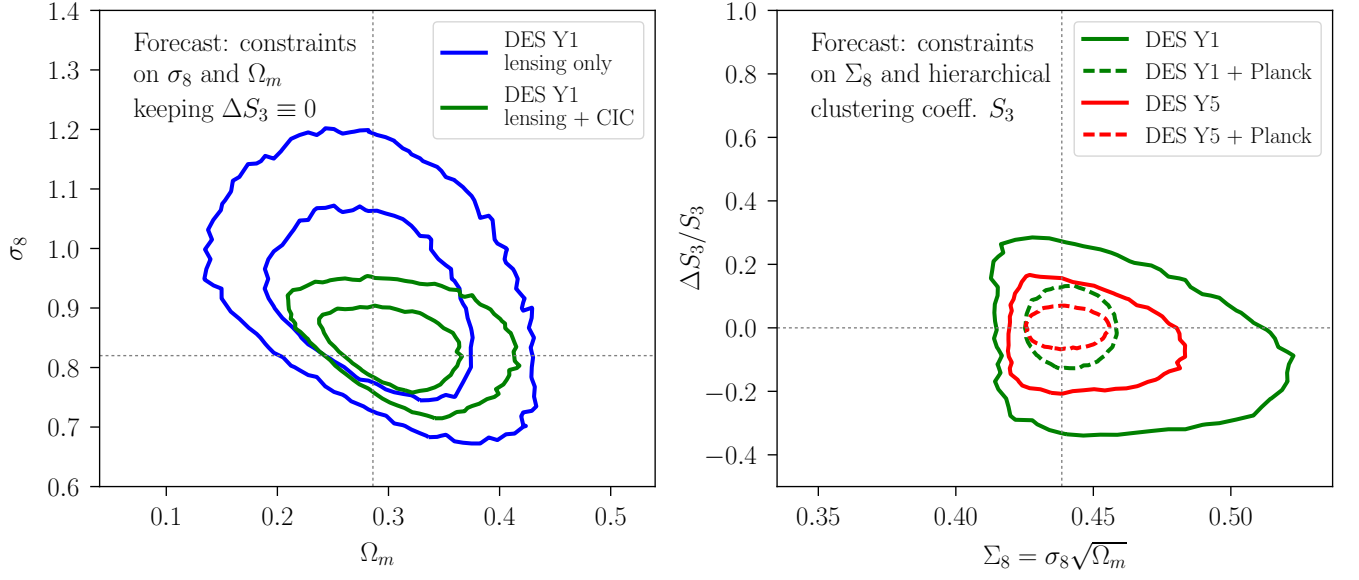[2] Gruen et al. [37] will also consider modeling approach 2.

FIG. 3. Left panel: Forecast of 1σ and 2σ constraints on Ω_m and σ_8 achievable with density split lensing and counts-in-cells in DES Y1 data. The constraints are marginalized over Ω_b, n_s, h_{100}, REDMAGIC galaxy bias b and galaxy-matter correlation coefficient r. For the parameters Ω_b, n_s, h_{100} we have assumed the same flat priors as used in the DES Y1 combined probes analysis presented in DES Collaboration [27]. Right panel: 1σ constraints on 3rd order statistics of the density field achievable with density split lensing and counts-in-cells in DES data alone (solid lines) and using additional information on cosmology from Planck (dashed lines, no lensing). The sharp cut-off of the contours for low values of Σ_8 is caused by the requirement that matter density and a shot-noise free galaxy density field must have a correlation coefficient $r \leq 1$. All likelihoods are centered around our fiducial cosmology, i.e. the parameters describing the Buzzard simulations.

parameter	fiducial value (in Buzzard)	prior in likelihood analysis	DES-Y1 constraints without $\Delta S_3/S_3$	DES-Y1 constraints with $\Delta S_3/S_3$
σ_8	0.82	$[0.2, 1.6]$	± 0.05	± 0.10
Ω_m	0.286	$[0.1, 0.9]$	± 0.04	± 0.06
Ω_b	0.047	$[0.01, 0.07]$	-	-
n_s	0.96	$[0.7, 1.3]$	-	-
h_{100}	0.7	$[0.55, 0.91]$	-	-
b	1.618	$[0.8, 2.5]$	± 0.11	± 0.27
r	0.956	$[0.0, 1.0]$	± 0.10	± 0.11
$\Delta S_3/S_3$	0.0	$[-1.0, 2.0]$	-	± 0.20

TABLE I. Model parameters of the forecast described in section II C 2. The second column shows our fiducial values (the values describing the Buzzard simulations). The third column shows the parameter priors used to cut our prediction for the posterior distribution of best-fit parameters. The priors on Ω_b, n_s and h_{100} are informative and chosen to be the same as used by the DES Collaboration [27]. The prior on r is needed for mathematical consistency. And the other priors only have to be introduced formally since we are approximating our analytic posterior by an MCMC. The fourth column shows the standard deviation of each parameter (as computed from the MCMC) after marginalization over all other parameters. The fifth column shows the same standard deviations but for the case where also $\Delta S_3/S_3$ is introduced as a free parameter of our model. These forecasts can be compared to the actual errors we find in Gruen et al. [37, their tables 2 and 3] which are close despite marginalizing over systematic uncertainties.

parameters based on a measurement of density split statistics. The covariance of $\hat{\pi}_{\alpha,\mathrm{ML}}$ and $\hat{\pi}_{\beta,\mathrm{ML}}$ can be estimated by

$$\mathrm{Cov}[\hat{\pi}_{\alpha,\mathrm{ML}}, \hat{\pi}_{\beta,\mathrm{ML}}]^{-1} \approx \frac{\partial \mathbf{d}_{\mathrm{th}}^T}{\partial \pi_\alpha} \cdot \mathbf{C}_d^{-1} \cdot \frac{\partial \mathbf{d}_{\mathrm{th}}}{\partial \pi_\beta} \ , \qquad \text{(II.29)}$$

where $\mathbf{d}_{\mathrm{th}}$ is our model of the density split data vector and $\mathbf{C}_d$ is the covariance matrix of a measurement of this signal in DES-Y1. We will in the following use an estimate of $\mathbf{C}_d$ from log-normal mocks and real DES Y1 shape noise [for details cf. 37]. The parameter covariance computed with equation II.29 can then be used to approximate the expected distribution of our best-fit parameter estimates with a multivariate Gaussian distribution.

Since the three parameters Ω_b, n_s and h_{100} are only poorly constrained by our data vector we are forced to assume prior knowledge on them. To do so, we cut the Gaussian posterior predicted from the parameter covariance with flat informative priors. These priors are chosen to be the same used by the DES Collaboration [27]. For reasons of mathematical consistency we furthermore have to demand that $r \in [0,1]$. These hard cuts of our originally Gaussian approximation to the posterior distribution of best-fit parameters make it difficult to marginalize over individual parameters. We hence numerically evaluate our analytic posterior with a Monte-Carlo Markov-Chain (MCMC). This chain is used in the following visualizations. Since we are using an MCMC to trace our analytic posterior, we have to formally define priors for all other model parameters. These are chosen to be flat and to extend well beyond the single-parameter standard deviations of the posterior. In the third column of table I, we summarize the priors chosen for each model parameter.

The left panel of Figure 3 shows the 1σ and 2σ constraints achievable in the Ω_m-σ_8 plane. These contours are marginalized over the other model parameters, using the priors mentioned above. The blue contours assume that only the density split lensing signal has been used while the green contours allow for complementary information from the tracer counts-in-cells histogram. In the fourth column of table I we show the standard deviation of each parameter as found in our approximation to the posterior (and assuming the full data vector, including lensing and counts-in-cell).

Density split statistics is complementary to an analysis based on 2-point statistics not just because it has a different dependence on the connection of galaxies and matter, but also since it is sensitive to higher order moments of the density field. We demonstrate this by introducing an additional degree of freedom in our model, described by an additional parameter:

8.) $\Delta S_3/S_3$: a factor multiplying all third order statistics in our predictions. The notation for this parameter is based on the ratio

$$S_3 \equiv \frac{\langle \delta^3 \rangle}{\langle \delta^2 \rangle^2} \qquad \text{(II.30)}$$

which connects third and second moments of the density contrast and hence characterizes the amplitude of the density bispectrum (see section IV A 2 for details). In our fiducial setup

we compute it from leading order perturbation theory and $\Delta S_3/S_3$ hence describes a relative deviation from that result.

Within the ΛCDM model and at leading order in perturbation theory, the scaling between 2-point and 3-point statistics of the density field is almost independent of the cosmological parameters Ω_m and σ_8 [12]. Hence, a value of $\Delta S_3/S_3 \neq 0$ would allow for deviations from the leading order result that cannot be compensated by changing Ω_m or σ_8. Such deviations could be caused non-standard physics of dark matter and dark energy that affect overdensities and underdensities differently (see e.g. Lue et al. [51], Multamäki et al. [57]; though $f(R)$ modified gravity theories have been shown to largely preserve the ΛCDM scaling, cf. Borisov & Jain [15], Jain & Zhang [43]). Alternatively, they could indicate a break down the the perturbative scaling relations due to highly nonlinear evolution of the density field or any small scale Baryonic physics that do not follow the scaling relations of perturbation theory [cf. 12, 45, 75].

In the right panel of Figure 3 we show how density split statistics including lensing and counts-in-cells can simultaneously constrain $\Delta S_3/S_3$ and the parameter

$$\Sigma_8 = \sigma_8 \sqrt{\Omega_m} \ , \qquad \text{(II.31)}$$

even after marginalization over the other model parameters. We also project how these constraints will improve when moving to year 5 (Y5) data of DES or when adding cosmological information from the cosmic microwave background. For the latter we estimated the parameter covariance in a Planck chain[3] and added this covariance as an additional Gaussian prior around our fiducial cosmology.

With DES Y1 alone, we will be able to constrain the aplitude of third order statistics of the density field to about 20% accuracy (cf. last column of table I). Combining DES Y5 and Planck, this improves to about 5%. And this is even underestimating the power of DES-Y5: to project our constraints onto year-5 we simply divided our covariance by a factor of 4 in order to account for the larger area of the final DES survey. But this does not take into account the fact that DES Y5 will also be deeper than DES-Y1, which reduces shape noise and opens up the possibility of analyzing a larger number of redshift bins.

III. SIMULATED DATA

In this work we use two sets of simulated data: the Buzzard galaxy catalogs which are constructed from high-resolution N-body simulations [28, 52, 78] and simulated random fields on the sky generated by the FLASK tool [81]. We briefly describe these data sets in the following sections.

---

[3] plikHM_TT_lowTEB in COM_CosmoParams_base-plikHM_R2.00.tar.gz from the Planck legacy archive https://pla.esac.esa.int/pla/, no lensing, cf. Planck Collaboration et al. [59]

A. Buzzard mock galaxy catalogs

Here we describe the key aspects of the Buzzard galaxy catalogs for the purposes of this work and refer the reader to more detailed descriptions elsewhere [28, 52, 78]. We use four independent realizations of a DES Y1-like survey in version Buzzard-v1.1. These catalogs were constructed from N-body simulations run using L-GADGET2 [69], a version of GADGET2 modified for memory efficiency. Second-order Lagrangian perturbation theory initial conditions [25] were employed using 2LPTIC [25], and lightcones were output on the fly. Each galaxy catalog is built from a set of three nested lightcones using progressively larger volume and lower resolution at higher redshifts. The force resolution in each box is 20, 35 and 53 kpc/h with the boundaries between the lightcones falling at redshifts $z = 0.34$ and $z = 0.9$.

The galaxy catalogs are constructed from the lightcones using the ADDGALS algorithm [78] which assigns galaxy luminosities and positions based on the relation between redshift, r-band absolute magnitude, and large-scale density, $p(\delta|M_r, z)$, found in a subhalo abundance matching (SHAM) model [23, 61, e.g.], in a high resolution N-body simulation. Spectral Energy Distributions (SEDs) are given to each simulated galaxy by finding a SDSS DR7 galaxy [24] that has a close match in M_r and distance to its fifth nearest neighbor galaxy and assigning the SDSS galaxy's SED to the simulated galaxy. Galaxy sizes and ellipticities were assigned by drawing from distributions fit to high resolution SuprimeCam i^--band data [55]. Observed magnitudes in $griz$ were generated by redshifting the SEDs to the observer frame and integrating over the DES passbands. Photometric errors were added using the DES Y1 Multi Object Fitting (MOF) depth maps.

The effects of weak gravitational lensing are calculated using the multiple-plane raytracing algorithm CALCLENS [5]. The raytracing is done on a $n_{side} = 4096$ HEALPIX [35] grid yielding an effective angular resolution of $0.85'$. At each lens plane, the inverse magnification matrix of the ray closest to every galaxy is interpolated to the galaxy's 3D position and used to shear and magnify the galaxy.

With galaxy catalogs with magnitudes, sizes and lensing effects in hand, REDMAGIC and METACALIBRATION [42, 64, 82] samples of galaxies are selected from the simulations in an effort to approximate the selections done in the data. In the case of REDMAGIC, the same algorithm which is used for selection in the data is applied to the simulations, resulting in a tracer galaxy catalog of equivalent volume density and at least comparable bias. For the METACALIBRATION sample, as we do not run full image simulations, we must make approximate cuts on signal to noise of the galaxies to create a facsimile of the source sample in the data. For in depth comparisons of these simulated samples to their data counterparts see DeRose et al. [28].

For the density split analysis in this work, we use RED-MAGIC high density tracer galaxies ($L > 0.5L_{\text{star}}, \rho = 10^{-3}\text{Mpc}^{-3}$ comoving density) selected at a REDMAGIC photometric redshift estimate of $0.2 < z < 0.45$. For the source redshift split lensing signals, we bin source galaxies by the expectation value of their $p(z)$ as estimated with BPZ [7, 40] from the Buzzard mock photometry. Bin limits are chosen such that the true mean redshifts of the bins match the mean redshifts of the two highest redshift source samples defined in Hoyle et al. [40].

B. Simulated density and convergence fields from FLASK

For testing the numerical implementation of the model described in the following section and for estimating a covariance matrix of the density split lensing and counts-in-cells signals in the Buzzard simulations and the DES data, we use log-normal realizations of matter density and convergence fields. We summarize their properties here, with details given in appendix A of Gruen et al. [37].

We generate these fields as all-sky HEALPIX maps of matter density and convergence using the FLASK software [81]. For the matter field, we choose the true redshift distribution of REDMAGIC galaxies in the Buzzard simulations, selected as described in section III A. The matter field is sampled by a tracer population with REDMAGIC density, bias of $b = 1.54$, and Poissonian noise, from which lines of sight of different density are identified by the same algorithm as in Buzzard or in the data. For the convergence fields, we choose the estimated redshift distributions of DES source galaxies in the two highest redshift bins of Hoyle et al. [40]. Log-normal parameters of the density and convergence fields are set by the perturbation theory formalism described in the following section.

960 of these realizations are used to estimate large-scale structure and shot noise contributions to the covariance matrix of the signals modeled herein. The contribution of shape noise is estimated by measuring the lensing signals in actual DES Y1 data with 960 realizations of the METACALIBRATION shape catalog [82] in which each galaxy ellipticity was rotated by a random angle.

IV. MODELLING DETAILS AND COMPARISON TO SIMULATIONS

In this section we present the approximations used to compute the ingredients (i), (ii) and (iii) of our model that are listed at the end of section II B. We also test the model ingredients (i) and (iii) directly against our N-body simulations.

Section IV A describes our model for the PDF of projected density contrast within the top-hat smoothing radius θ_T, $p(\delta_{m,T})$. Section IV B describes how we model the convergence profile $\langle \kappa_{<\theta}|\delta_{m,T}\rangle$ around apertures of fixed density contrast $\delta_{m,T}$. And in section IV C we describe our modeling of the probability $P(N_T|\delta_{m,T})$ of finding N_T tracer galaxies in an aperture with density contrast $\delta_{m,T}$.

Section IV D summarizes our fiducial model and the approximations used therein.

A. Projected density PDF

The computation of the PDF of the density field when smoothed by top-hat filters has e.g. been adressed by Bernardeau [9], Bernardeau & Valageas [10] and Valageas [76] (see also more recent developments in [13, 14, 21, 22, 75] which however do not yet enter our formalism). Bernardeau & Valageas [10] demonstrated how to extend methods for the computation of the 3D density PDF to the weak lensing aperture mass which is a projected quantity. In the following we show how to modify their formalism in order to compute the line-of-sight density PDF in angular circles of radius θ_T. To do so, we have to consider the *cumulant generating function* (CGF) of the field $\delta_{m,T}(\hat{\mathbf{n}})$. The *moment generating function* (MGF) of $\delta_{m,T}(\hat{\mathbf{n}})$ is defined as

$$\psi(y) = \sum_k \frac{\langle \delta_{m,T}(\hat{\mathbf{n}})^k \rangle}{k!} y^k . \tag{IV.1}$$

Due to the isotropy of the universe it does not depend on $\hat{\mathbf{n}}$. The CGF $\varphi(y)$ is given in terms of the MGF $\psi(y)$ as

$$\begin{aligned}\varphi(y) &= \ln \psi(y) \\ &\equiv \sum_k \frac{\langle \delta_{m,T}(\hat{\mathbf{n}})^k \rangle_c}{k!} y^k ,\end{aligned} \tag{IV.2}$$

where in the last line we have defined the *connected moments* or *cumulants* $\langle \delta_{m,T}(\hat{\mathbf{n}})^k \rangle_c$ of $\delta_{m,T}(\hat{\mathbf{n}})$. The CGF of $\delta_{m,T}(\hat{\mathbf{n}})$ is related to its PDF via

$$e^{\varphi(y)} = \int d\delta_{m,T} \, e^{y\delta_{m,T}} \, p(\delta_{m,T}) , \tag{IV.3}$$

which is the Laplace transform of the PDF. Hence, if $\varphi(y)$ is a known analytic function, then $p(\delta_{m,T})$ can be computed by the inverse Laplace transform

$$p(\delta_{m,T}) = \int_{-\infty}^{\infty} \frac{dy}{2\pi} \, e^{-iy\delta_{m,T} + \varphi(iy)} . \tag{IV.4}$$

This integral in the complex plain is most efficiently evaluated along the path of steepest descent (cf. [13], especially their appendix B). The cumulants of $\delta_{m,T}(\hat{\mathbf{n}})$ can be approximated as (cf. Bernardeau & Valageas [10])

$$\langle \delta_{m,T}(\hat{\mathbf{n}})^k \rangle_c \approx \int dw \, \frac{\left\langle \left\{ \delta_{\text{cy.},\theta w,L}(w\hat{\mathbf{n}}, w) q_g(w) L \right\}^k \right\rangle_c}{L} , \tag{IV.5}$$

where $\delta_{\text{cy.},R,L}$ is the average of $\delta_{m,3D}$ over a cylinder of length L and radius R, and where L has to be chosen such that correlations of $\delta_{m,3D}$ over a distance L vanish for all practical purposes. Equation IV.5 employs a small angle approximation and a Limber-like approximation [48]. This means it assumes that any n-point correlation function between density constrast at different redshifts z_i, $i = 1...n$, varies much more quickly with the redshift differences Δz_{ij} then the projection kernel

q_g. Comparing IV.2 and IV.5 we see that the CGF of $\delta_{m,T}$ can be computed in terms of the CGF of $\delta_{\text{cy.},R,L}$ as

$$\varphi(y) \approx \int dw \, \frac{\varphi_{\text{cy.},\theta w,L}(q_g(w)Ly, w)}{L} . \tag{IV.6}$$

Hence, we have reduced the task of computing $p(\delta_{m,T})$ to the task of computing the connected moments of matter contrast in a long 3D cylinder.

To proceed we consider two different ansatzes. The first is to approximate $p(\delta_{m,T})$ by a log-normal distribution which is fixed by the first three connected moments of $\delta_{m,T}(\hat{\mathbf{n}})$. The second ansatz is to compute $\varphi_{\text{cy.},\theta w,L}(y)$ as a whole in a way similar to the one of Bernardeau [9] and Valageas [76] for the matter contrast in a 3-dimensional sphere. Using IV.6 we can them attempt to solve the integral in IV.4 directly. We present details of both approaches in the following subsections.

1. Log-normal approximation for the PDF

Instead of computing the complete cumulant generating function of $\delta_{m,T}$ via equation IV.6 we start with an approach that only requires knowledge of the second and third cumulant, i.e. the moments $\langle \delta_{m,T}^2 \rangle_c$ and $\langle \delta_{m,T}^3 \rangle_c$ (implicitly we also assume $\langle \delta_{m,T} \rangle_c \equiv 0$ for the first cumulant). To do so, we approximate $\delta_{m,T}$ as a shifted log-normal random variable [38, 81]. In this case the PDF of $\delta_{m,T}$ is given by

$$p(\delta_{m,T}) = \frac{\exp\left(-\frac{[\ln(\delta_{m,T}/\delta_0 - 1) + \sigma^2/2]^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma(\delta_{m,T} - \delta_0)} \tag{IV.7}$$

if $\delta_{m,T} > \delta_0$ and $p(\delta_{m,T}) = 0$ otherwise. The expectation value of this PDF is zero, as is appropriate. The variance and skewness of $\delta_{m,T}$ are given in terms of the parameters parameters δ_0 and σ by[38]

$$\begin{aligned}\langle \delta_{m,T}^2 \rangle_c &= \delta_0^2 \left(e^{\sigma^2} - 1 \right) \\ \langle \delta_{m,T}^3 \rangle_c &= \frac{3}{\delta_0} \langle \delta_{m,T}^2 \rangle_c^2 + \frac{1}{\delta_0^3} \langle \delta_{m,T}^3 \rangle_c^2 .\end{aligned} \tag{IV.8}$$

The ansatz of a log-normal PDF has been shown to be consistent with early DES data by Clerkin et al. [20]. It can be reasonably motivated - at least for the 3-dimensional density contrast $\delta_{m,3D}$ - by observing the scaling of the skewness of a log-normal random field with its variance. Let $\hat{\delta}_{m,3D}$ be log-normal random variable with PDF given by IV.7. From equation IV.8 it can then be seen that

$$\langle \hat{\delta}_{m,3D}^3 \rangle_c = \frac{3}{\delta_0} \langle \hat{\delta}_{m,3D}^2 \rangle_c^2 + \mathcal{O}(\langle \hat{\delta}_{m,3D}^2 \rangle_c^3) . \tag{IV.9}$$

Choosing the parameter δ_0 appropriately, this is exactly the scaling relation that is observed at leading order in cosmological perturbation theory [e.g. 12]. For a power law power spectrum, the same kind of scaling is observed also for 2-dimensional, projected versions of the density field. Hence, one of the ansatzes used in this paper to compute the PDF of $\delta_{m,T}$ is to derive its variance and skewness as described in appendix B and fix δ_0 and σ by demanding that the PDF in IV.7 has the same second and third moments.

2. Tree level computation of $\varphi_{\mathrm{cy.},\theta w, L}(y, w)$ in the cylindrical collapse model

Let us first consider the cumulant generating function $\varphi_{3D}(y, \tau)$ of the 3-dimensional density contrast $\delta_{3D}(\mathbf{x}, \tau)$. To compute φ_{3D} at tree-level in perturbation theory it is sufficient to assume spherical symmetry around a particular point $\mathbf{x}$ [see e.g. 76]. Doing so, we can directly compute $\delta_{3D}(\mathbf{x}, \tau)$ as a function of the linear density contrast $\delta_{3D,\mathrm{lin.}}(\mathbf{x}, \tau)$ by means of the spherical collapse model [32, 76], i.e.

$$\delta_{3D}(\mathbf{x}, \tau) = F[\delta_{3D,\mathrm{lin.}}(\mathbf{x}, \tau), \tau] \qquad \text{(IV.10)}$$

where F is determined by one of the differential equations presented in appendix A. Hence, using the assumption that the linear density contrast has a Gaussian distribution with variance $\sigma^2_{3D,\mathrm{lin.}}(\tau)$ we can express the cumulant generating function as (cf. equation IV.3)

$$\exp\{\varphi_{3D}(y, \tau)\}$$
$$= \int \mathrm{d}\delta_{3D}\, e^{y\delta_{3D}}\, p(\delta_{3D}, \tau)$$
$$= \int \frac{\mathrm{d}\delta_{3D,\mathrm{lin.}}}{\sqrt{2\pi\sigma^2_{3D,\mathrm{lin.}}(\tau)}} \exp\left(yF[\delta_{3D,\mathrm{lin.}}, \tau] - \frac{\delta^2_{3D,\mathrm{lin.}}}{2\sigma^2_{3D,\mathrm{lin.}}(\tau)}\right),$$
$$\text{(IV.11)}$$

where in the last step we simply performed a change of variables from δ_{3D} to $\delta_{3D,\mathrm{lin.}}$. We now employ Laplace's method, which states that a function $f(x)$ which strongly peaks around x_0 fulfils

$$\int \mathrm{d}x\, e^{f(x)} \approx \sqrt{\frac{2\pi}{|f''(x_0)|}} e^{f(x_0)}. \qquad \text{(IV.12)}$$

This way we can approximate the last line of IV.11 as

$$e^{\varphi_{3D}(y, \tau)} \approx \sqrt{\frac{1}{|1 - yF''[\delta^*, \tau]\, \sigma^2_{3D,\mathrm{lin.}}(\tau)|}} \times$$
$$\times \exp\left(yF[\delta^*, \tau] - \frac{\delta^{*2}}{2\sigma^2_{3D,\mathrm{lin.}}(\tau)}\right) \text{(IV.13)}$$

where δ^* is the linear density contrast that maximizes the exponent in IV.11 and $'$ denotes derivation wrt. δ. In the quasi-linear limit of $\sigma^2_{3D,\mathrm{lin.}} \to 0$ this gives

$$\varphi_{3D}(y, \tau) \approx yF[\delta^*, \tau] - \frac{\delta^{*2}}{2\sigma^2_{3D,\mathrm{lin.}}(\tau)}, \qquad \text{(IV.14)}$$

where δ^* has to be determined by the implicit equation

$$\frac{\delta^*}{\sigma^2_{3D,\mathrm{lin.}}(\tau)} = yF'[\delta^*, \tau]. \qquad \text{(IV.15)}$$

It should be noted that equations IV.14 and IV.15 reproduce *exactly* the tree-level results for the cumulant generating function [cf. 9, 12, 14, 76] ! As described in Bernardeau et al. [12] the coefficients

$$S_n = \frac{\langle \delta^n \rangle_c}{\langle \delta^2 \rangle_c^{n-1}} \qquad \text{(IV.16)}$$

are given quite accurately by the lowest order of perturbation theory. Hence, using *halofit* [68, 72] and an analytic transfer function [30] to compute the non-linear matter power spectrum, we can compute the non-linear variance $\langle \delta^2 \rangle_{c,\mathrm{non.lin}}$ and then rescale the leading order CGF to its non-linear version via

$$\varphi_{\mathrm{non.lin}}(y, \tau) = \frac{\langle \delta^2 \rangle_{c,\mathrm{lin}}}{\langle \delta^2 \rangle_{c,\mathrm{non.lin}}} \varphi_{\mathrm{lead}}\left(y\frac{\langle \delta^2 \rangle_{c,\mathrm{non.lin}}}{\langle \delta^2 \rangle_{c,\mathrm{lin}}}, \tau\right). \qquad \text{(IV.17)}$$

To perform the projection integral in equation IV.6 we need to know the cumulant generating function of the density contrast in a long cylinder of radius R and length $L >> R$, $\delta_{\mathrm{cy.},R,L}$. Bernardeau [1994; see also 76, and the other references above] have generalized equation IV.14 to the case of matter density in a spherical aperture. Their results can easily be transferred to cylindrical apertures, yielding

$$\varphi_{\mathrm{cy.},R,L}(y, \tau) \approx yF_{\mathrm{cyl.}}[\delta^*, \tau] - \frac{\delta^{*2}}{2\sigma^2_{R\sqrt{1+F_{\mathrm{cyl.}}[\delta^*]}, L,\mathrm{lin.}}(\tau)}, \qquad \text{(IV.18)}$$

where $\sigma^2_{R,L,\mathrm{lin.}}$ is the variance of linear density contrast in a cylinder, $F_{\mathrm{cyl.}}$ is now determined by cylindrical collapse and δ^* has to be determined by the implicit equation

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\delta^*}\frac{\delta^{*2}}{\sigma^2_{R\sqrt{1+F_{\mathrm{cyl.}}[\delta^*]}, L,\mathrm{lin.}}(\tau)} = yF'[\delta^*, \tau]. \qquad \text{(IV.19)}$$

Using equation IV.17 we can again rescale this leading order result for the generating function to its non-linear counterpart.

The validity of equation IV.17 is ultimately limiting the accuracy of our model for the distribution of the density contrast inside the aperture radius θ_T, $p(\delta_{m,T})$. In figure 4 we compare the PDF measured in the Buzzard simulations to both the log-normal model and the full CGF computation of the PDF for aperture radii $\theta_T = 10', 20', 30'$. For both $\theta_T = 20'$ and $30'$ our model PDF's and the Buzzard simulations agree within DES-Y1 cosmic variance as can be seen in the residuals shown in the lower panels of each plot. Also, the difference between log-normal approximation and full CGF computation is completely negligible. To investigate the agreement of Buzzard and our models more quantitatively, we also compare the variance and skewness of each PDF. In figure 5 we show the ratio of these moments as found in Buzzard to our predictions. For $\theta_T = 10'$, the density field in Buzzard seems to have a significantly higher skewness then predicted by our model. We attribute this indeed to the failing of the Quasi-linear rescaling, eq. IV.17. For the other radii the skewness values agree to within $2-3\%$. A similar relative agreement is achieved for the variance of the distributions. At our fiducial radius $\theta_T = 20'$ the variances of Buzzard and our model differ by about 2.4%. This corresponds to a disagreement in the fluctuation amplitude σ_8 of about 1.2%.

For comparison, we also show a Gaussian model for the density PDF in figure 4, using the same variance as for the other PDF models. It can clearly be seen, that $p(\delta_{m,T})$ cannot be well described by a Gaussian distribution for any of the considered smoothing radii.
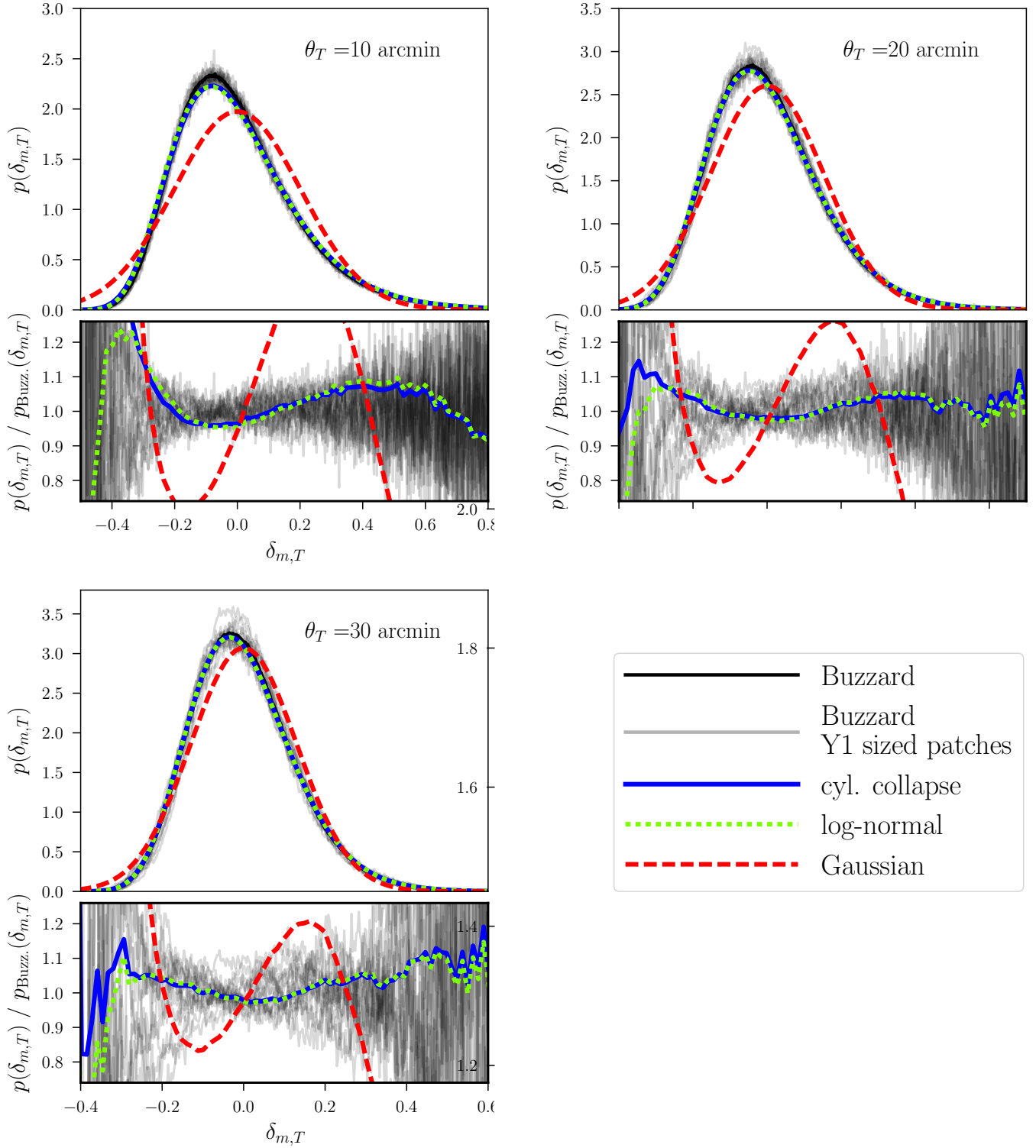
FIG. 4. The PDF of projected density contrast in Buzzard compared to several models. In the upper panel of each plot the black line shows a histogram of $\delta_{m,T}$ measured in an all-sky map from Buzzard. The grey lines show histograms measured in 14 DES year1 shaped patches of that all-sky map. The blue lines show the PDF predicted by our tree-level computation of the cumulant generating function, the green lines show the PT-motivated log-normal model and the red lines show a Gaussian PDF with the same variance as the other two models. The bottom panels of each plot are showing the ratio of each PDF to the one measured in the Buzzard all-sky. For all aperture radii our halofit power spectrum is predicting a standard deviation of $\delta_{m,T}$ that is $\gtrsim 2\%$ higher than what we find in Buzzard (cf. figure 5). For $\theta_T = 20'$ and $30'$ this is the dominant source of mismatch.
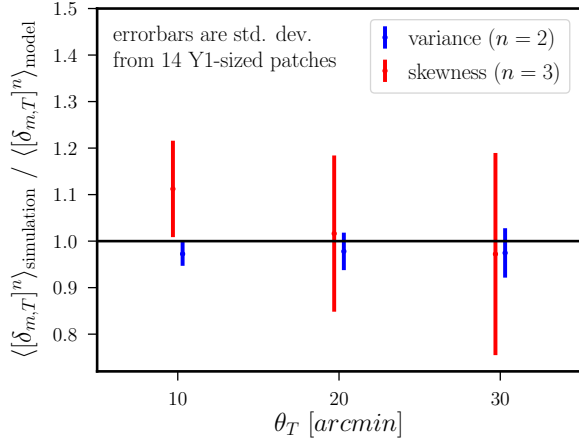
FIG. 5. This figure shows the ratio of moments of $\delta_{m,T}$ measured in a Buzzard all-sky density map to the moments predicted by our model as a function of θ_T. The errorbars represent the standard deviations of the moments found in a set of 14 DES Y1-sized patches in the Buzzard map. At our fiducial radius $\theta_T = 20'$ we find a $\sim 2.4\%$ deviation between the variance of $\delta_{m,T}$ measured in Buzzard and that predicted by our model. This would correspond to a $\sim 1.2\%$ deviation in σ_8.

B. Convergence profile around lines-of-sight of fix density contrast $\delta_{m,T}$

We now want to know the average lensing convergence $\kappa_{<\theta}$ inside a radius θ around a line-of-sight with a given value of $\delta_{m,T}$. By means of equation II.6 this can be turned into a prediction of the density split lensing signal. We start by looking at the joint moment generating function of $\kappa_{<\theta}$ and $\delta_{m,T}$,

$$\psi_\theta(y,z) = \sum_{k,l} \frac{\langle \delta_{m,T}^k \kappa_{<\theta}^l \rangle}{k!\, l!} y^k\, z^l\,, \qquad (\text{IV.20})$$

and their joint cumulant generating function defined by

$$\begin{aligned} \varphi_\theta(y,z) &= \ln \psi_\theta(y,z) \\ &\equiv \sum_{k,l} \frac{\langle \delta_{m,T}^k \kappa_{<\theta}^l \rangle_c}{k!\, l!} y^k\, z^l\,. \end{aligned} \qquad (\text{IV.21})$$

Using a Limber-like approximation similar to the one employed in eq. IV.6, one can write $\varphi_\theta(y,z)$ as a line-of-sight projection of the CGF of matter density contrasts that are smoothed over concentrical cylindrical apertures with length L and radii R_1, R_2, $\varphi_{\text{cyl.},R_1,R_2,L}(y,z,w)$:

$$\varphi_\theta(y,z) \approx \int \mathrm{d}w \, \frac{\varphi_{\text{cyl.},\theta_T w,\theta w,L}(q_g(w)Ly, W_s(w)Lz, w)}{L}\,. \qquad (\text{IV.22})$$

Here, $q_g(w)$ is again the line-of-sight density of lens galaxies and $W_s(w)$ is the lensing efficiency defined in eq. II.5.

The joint PDF of $\kappa_{<\theta}$ and $\delta_{m,T}$ is related to φ_θ via

$$p(\delta_{m,T}=s,\kappa_{<\theta}=r) = \int_{-\infty}^{\infty} \frac{\mathrm{d}y\,\mathrm{d}z}{(2\pi)^2} \, e^{-iys-izr+\varphi_\theta(iy,iz)}\,. \qquad (\text{IV.23})$$

The expectation value of $\kappa_{<\theta}$ given a certain value of $\delta_{m,T}$ is then given by

$$\begin{aligned} \langle \kappa_{<\theta} | \delta_{m,T} = s \rangle &= \int \mathrm{d}r \, p(\delta_{m,T}=s,\kappa_{<\theta}=r) \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{d}y\,\mathrm{d}z}{(2\pi)^2} \, e^{-iys+\varphi_\theta(iy,iz)} \int \mathrm{d}r\, r\, e^{-izr} \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{d}y\,\mathrm{d}z}{2\pi} \, e^{-iys+\varphi_\theta(iy,iz)} \, i\frac{\mathrm{d}}{\mathrm{d}z}\delta_D(z) \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{d}y}{2\pi} \, e^{-iys+\varphi_{m,T}(iy)} \, G_\theta(iy)\,, \end{aligned} \qquad (\text{IV.24})$$

with

$$\begin{aligned} G_\theta(y) &= \left.\frac{\mathrm{d}}{\mathrm{d}z}\varphi_\theta(y,z)\right|_{z=0} \\ &= \sum_k \frac{\langle \delta_{m,T}^k \kappa_{<\theta}\rangle_c}{k!} y^k\,. \end{aligned} \qquad (\text{IV.25})$$

Using equation IV.22 we can express $G_\theta(y)$ by the corresponding cylindrical, 3-dimensional quantity:

$$G_\theta(y) \approx \int \mathrm{d}w \, W_s(w) \, G_{\text{cyl.},\theta_T w,\theta w,L}(q_g(w)Ly, w)\,. \qquad (\text{IV.26})$$

We again pursue two ansatzes for computing $G_\theta(y)$: one involving a log-normal model for the joint cumulants of $\kappa_{<\theta}$ and $\delta_{m,T}$ and one involving the model of cylindrical collapse to compute a leading order perturbation theory prediction for $G_{\text{cyl.},\theta_T w,\theta w,L}(y,w)$. We are detailing these ansatzes in the following subsections.

1. Log-normal model for the joint cumulants of $\kappa_{<\theta}$ and $\delta_{m,T}$

From equation IV.25 one can see that only joint cumulants of the form

$$\langle \delta_{m,T}^k \kappa_{<\theta}\rangle_c \qquad (\text{IV.27})$$

enter the convergence profile around a given density contrast. Hence, in a spirit similar to section IV A 1 we only compute the leading order cumulants $\langle \delta_{m,T} \kappa_{<\theta}\rangle_c$ and $\langle \delta_{m,T}^2 \kappa_{<\theta}\rangle_c$ as described in appendix B and use these moments to fix a joint log-normal PDF for $\kappa_{<\theta}$ and $\delta_{m,T}$. Note that this is indeed *not* assuming, that $\kappa_{<\theta}$ is a log-normal random variable. It rather assumes that

$$\kappa_{<\theta} = \kappa_{\text{log-normal}} + \kappa_{\text{uncorr.}}\,, \qquad (\text{IV.28})$$

where $\kappa_{\text{log-normal}}$ is log-normal and $\kappa_{\text{uncorr.}}$ is an unspecified random variable that is uncorrelated with $\delta_{m,T}$. Only $\kappa_{\text{log-normal}}$ will actually contribute to the density split lensing signal and we can model the expectation value $\langle \kappa_{<\theta} | \delta_{m,T} = s\rangle$ by the following relation holding for two joint log-normal variables:

$$\frac{\langle \kappa_{<\theta} | \delta_{m,T} = s\rangle}{\kappa_0} = \exp\left(\frac{C(2\log(1+s/\delta_0)+V-C)}{2V}\right)-1 \qquad (\text{IV.29})$$

with

$$V = \log\left(1.0 + \frac{\langle \delta_{m,T}^2 \rangle_c}{\delta_0^2}\right)$$

$$C = \log\left(1.0 + \frac{\langle \delta_{m,T}\, \kappa_{<\theta} \rangle_c}{\delta_0\, \kappa_0}\right)$$

$$\kappa_0 = \frac{\langle \delta_{m,T}\, \kappa_{<\theta} \rangle_c^2\, e^V}{\langle \delta_{m,T}^2\, \kappa_{<\theta} \rangle_c - 2\langle \delta_{m,T}\, \kappa_{<\theta} \rangle_c \langle \delta_{m,T}^2 \rangle_c/\delta_0} \quad \text{(IV.30)}$$

and δ_0 determined as described in section IV A 1.

It should be noted that the log-normal parameter κ_0 which we use to approximate the contribution of $\kappa_{<\theta}$ to the lensing signal is dependent on the the smoothing scale θ. This indicates even further, that we do indeed not approximate κ as a log-normal field.

2. Tree level computation of $G_{\mathrm{cyl.},R_1,R_2,L}(y,w)$ in the cylindrical collapse model

For convenience we will shorten the notation of section IV A 2 by defining

$$G(y) \equiv G_{\mathrm{cyl.},R_1,R_2,L}(y,w)$$
$$\varphi(y,z) \equiv \varphi_{\mathrm{cyl.},R_1,R_2,L}(y,z,w)$$
$$F[\delta] \equiv F_{\mathrm{cyl.}}[\delta,\tau]$$
$$\sigma_R \equiv \sigma_{R,L,\mathrm{lin.}}$$
$$R_{i,\mathrm{L}}(\delta) \equiv R_i\sqrt{1 + F[\delta]}\,, \; i = 1,2\,. \quad \text{(IV.31)}$$

The joint cumulant generating function of density contrast in concentric cylinders is then [in complete analogy to equation IV.18; see also 10, who present very similar calculations] given by

$$\varphi(y,z) \approx yF[\delta_1^*] + zF[\delta_2^*] - \frac{1}{2}\sum_{i,j}\delta_i^*\delta_j^*\left(\mathcal{C}^{-1}\right)_{ij}\,, \quad \text{(IV.32)}$$

where the elements of the matrix $\mathcal{C}$ are given by $\mathcal{C}_{11} = \sigma_{R_1,\mathrm{L}(\delta_1^*)}^2$, $\mathcal{C} = \sigma_{R_2,\mathrm{L}(\delta_2^*)}^2$ and $\mathcal{C}_{12} = \mathcal{C}_{21}$ is the linear covariance of density contrasts in concentric cylinders of radii $R_{1,\mathrm{L}}(\delta_1^*)$ and $R_{2,\mathrm{L}}(\delta_2^*)$. This time the critical linear density contrasts δ_2^* and δ_2^* are given by the implicit equations

$$\frac{1}{2}\frac{\partial}{\partial\delta_1^*}\sum_{i,j}\delta_i^*\delta_j^*\left(\mathcal{C}^{-1}\right)_{ij} = yF'[\delta_1^*] \quad \text{(IV.33)}$$

$$\frac{1}{2}\frac{\partial}{\partial\delta_2^*}\sum_{i,j}\delta_i^*\delta_j^*\left(\mathcal{C}^{-1}\right)_{ij} = zF'[\delta_2^*]\,. \quad \text{(IV.34)}$$

Note that these conditions force each δ_i^* to be a function of both y and z, i.e. $\delta_i^* = \delta_i^*(y,z)$.

To predict the convergence profile around apertures of a given density contrast $\delta_{m,T}$ by means of equations IV.24 and IV.26 we are interested in computing the function

$$G(y) = \left.\frac{\partial}{\partial z}\varphi(y,z)\right|_{z=0}\,. \quad \text{(IV.35)}$$

Using the conditions IV.33 and IV.34 one can see right away that

$$G(y) = F[\delta_2^*(y,0)]\,. \quad \text{(IV.36)}$$

Furthermore, for $z = 0$ equations IV.33 and IV.34 can be simplified to

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\delta_1^*}\frac{\delta_1^{*2}}{\mathcal{C}_{11}} = yF'[\delta_1^*] \quad \text{(IV.37)}$$

$$\delta_2^* = \frac{\mathcal{C}_{12}}{\mathcal{C}_{11}}\delta_1^*\,. \quad \text{(IV.38)}$$

This way we obtain a solution for $G(y)$ at leading order in perturbation theory. In appendix B we argue that for $R_2 \geq R_1$ the cumulants $\langle \delta_{R_1}^k\, \delta_{R_2} \rangle_c$ approximately follow the scaling relation

$$\langle \delta_{R_1}^k\, \delta_{R_2} \rangle_c \sim \langle \delta_{R_1}\, \delta_{R_2} \rangle_c \langle \delta_{R_1}^2 \rangle_c^{k-1}\,. \quad \text{(IV.39)}$$

This can be used to correct the tree-level approximation of $G(y)$ for the non-linear evolution of the power spectrum. To do so, we first determine the proportionality factors of the relation IV.39 at leading order by fitting a polynomial in y to the function $G(y)$ and extracting the cumulants $\langle \delta_{R_1}^k\, \delta_{R_2} \rangle_c$ from the polynomial coefficients. In practice, we do this with a polynomial of degree 10, but already a polynomial of degree 5 gives almost identical results[4]. Then, we use the revised halofit of Takahashi et al. [72] to compute a late-time version of the right-hand-side of IV.39. This, together with the tree-level proportionality factors determined before, yields a non-linear approximation of the polynomial coefficients representing $G(y)$. We use those to re-compute $G(y)$ and then carry out the projection integral in equation IV.26.

Our rescaling of the coefficient corresponding to the cumulant $\langle \delta_{R_1}^2\, \delta_{R_2} \rangle_c$ is infact more complicated than described here, cf. appendix B 5. But we find our prediction of the density split lensing signal to be insensitive to the details of the rescaling procedure.

C. Shot-noise, stochasticity and Counts-in-Cells

We now want to model the conditional probability $P(N_T|\delta_{m,T})$ of finding N_T galaxies in an angular radius of θ_T, when the projected density contrast in that radius is $\delta_{m,T}$. This is the third ingredient of the framework described in section II and completes our modeling of the density split lensing signal as well as the counts-in-cells histogram.

To analyze the relation of N_T and $\delta_{m,T}$ in a systematic way, let us introduce the auxiliary field $\delta_{g,T}$. We assume that $\delta_{g,T}(\hat{\mathbf{n}})$ is a smooth field in the sky and that N_T is a Poissonian tracer of this field. This means we will assume that

$$P(N_T = N|\delta_{g,T}) = \frac{\left[\bar{N}(1 + \delta_{g,T})\right]^N}{N!}\,e^{-\bar{N}(1+\delta_{g,T})}\,, \quad \text{(IV.40)}$$

---

[4] The coefficients of linear and quadratic order in y are always obtained from the exact perturbation theory computation of appendix B.

where $\bar{N} \equiv \langle N_T \rangle$. A consequence of this assumption is that the expectation value of N_T for fixed $\delta_{g,T}$ is given by

$$\langle N_T | \delta_{g,T} \rangle = \bar{N}(1 + \delta_{g,T}) \qquad \text{(IV.41)}$$

and that the variance of N_T for fixed $\delta_{g,T}$ fulfils

$$\frac{\text{Var}\left[N_T | \delta_{g,T}\right]}{\langle N_T | \delta_{g,T} \rangle} = 1 \ . \qquad \text{(IV.42)}$$

To connect the galaxy field to the lensing convergence we however need to know the relation between N_T and $\delta_{m,T}$. Assuming a generic joint PDF of $\delta_{m,T}$ and $\delta_{g,T}$ we can write the expectation values of N_T for fixed $\delta_{m,T}$ as

$$\langle N_T | \delta_{m,T} \rangle = \int \text{d}\delta_{g,T} \, p(\delta_{g,T} | \delta_{m,T}) \langle N_T | \delta_{g,T} \rangle$$
$$= \bar{N}(1 + \langle \delta_{g,T} | \delta_{m,T} \rangle) \ . \qquad \text{(IV.43)}$$

Also, it can be shown that the variance of N_T for a fixed value of $\delta_{m,T}$ is given by

$$\text{Var}\left[N_T | \delta_{m,T}\right] = \langle N_T | \delta_{m,T} \rangle + \bar{N}^2 \text{Var}\left[\delta_{g,T} | \delta_{m,T}\right] \ . \qquad \text{(IV.44)}$$

From equation IV.44 we can see that the distribution of N_T given $\delta_{m,T}$ can only be a Poisson distribution if $\text{Var}\left[\delta_{g,T} | \delta_{m,T}\right] \equiv 0$. This is the simplest model for the connection of N_T and $\delta_{m,T}$ that we test in this work and in Gruen et al. [37]. If $\text{Var}\left[\delta_{g,T} | \delta_{m,T}\right] \neq 0$ we will say that there is a stochasticity between the galaxy field and the matter density field, and we cannot assume a Poisson distribution for $P(N_T | \delta_{m,T})$. We note that "stochasticity" in this context could arise from a nonlinear biasing relationship between $\delta_{g,T}$ and $\delta_{m,T}$, including e.g. a dependence on higher powers of $\delta_{m,T}$ or effects from halo exclusion [2], or from physical stochasticity in galaxy formation.

We explore two ways to account for a possible stochasticity (see also Dekel & Lahav [29], who have discussed similar concepts). In our first approach we introduce a free parameter to our model - a Pearson correlation coefficient $r \neq 1$ between the random fields $\delta_{g,T}$ and $\delta_{m,T}$. Within our lognormal framework we show that this automatically leads to a $\delta_{m,T}$-dependence of the ratio in eq. IV.46. We explain the details of this in section IV C 1.

In our second approach we employ a generalized Poisson distribution for $P(N_T | \delta_{m,T})$ that allows for

$$\frac{\text{Var}\left[N_T | \delta_{m,T}\right]}{\langle N_T | \delta_{m,T} \rangle} \neq 1 \ . \qquad \text{(IV.45)}$$

In this approach we introduce 2 parameters, α_0 and α_1, to our model such that

$$\frac{\text{Var}\left[N_T | \delta_{m,T}\right]}{\langle N_T | \delta_{m,T} \rangle} \approx \alpha_0 + \alpha_1 \, \delta_{m,T} \ . \qquad \text{(IV.46)}$$

The details of this are explained in section IV C 2.

Both of our approaches match our simulated data equally well (cf. Figure 7). This means that, for the galaxies in these realizations, the model based on the correlation coefficient r

is a sufficient description. It will thus be the fiducial model in this paper, used in all figures unless otherwise noted. In Gruen et al. [37] we will nevertheless apply both this and the two-parametric model to account for the possibility that the shot-noise of real galaxies behaves in a more complicated way than that of our simulated galaxies.

1. Shot-noise model 1: correlation $r \neq 1$ between galaxy density and matter density

In our fiducial model of $P(N_T | \delta_{m,T})$ we approximate the joint distribution of both $\delta_{m,T}$ and $\delta_{g,T}$ with a joint lognormal distribution (cf. eq. IV.7 and Hilbert et al. [38] for properties of joint log-normal distributions). The joint PDF of two log-normal random variables is characterized by 5 parameters, e.g. by the variance and skewness of each variable and the covariance between the two variables.

In our case, we compute the variance and skewness of $\delta_{m,T}$ as described in section IV A 1 and set the variance and skewness of $\delta_{g,T}$ to

$$\langle \delta_{g,T}^2 \rangle_c = b^2 \langle \delta_{m,T}^2 \rangle_c$$
$$\langle \delta_{g,T}^3 \rangle_c = b^3 \langle \delta_{m,T}^3 \rangle_c \qquad \text{(IV.47)}$$

where the *galaxy bias* b is a free parameter. The covariance of $\delta_{m,T}$ and $\delta_{g,T}$ is parametrized by their correlation coefficient

$$r = \frac{\langle \delta_{g,T} \, \delta_{m,T} \rangle_c}{\sqrt{\langle \delta_{g,T}^2 \rangle_c \, \langle \delta_{m,T}^2 \rangle_c}}$$
$$= \frac{\langle \delta_{g,T} \, \delta_{m,T} \rangle_c}{b \, \langle \delta_{m,T}^2 \rangle_c} \ , \qquad \text{(IV.48)}$$

i.e.

$$\langle \delta_{g,T} \, \delta_{m,T} \rangle_c = rb \, \langle \delta_{m,T}^2 \rangle_c \ . \qquad \text{(IV.49)}$$

The log-normal model for the joint PDF of $\delta_{m,T}$ and $\delta_{g,T}$ now allows us to compute the variance of galaxy counts as a function of $\delta_{m,T}$ and more generally to compute $P(N_T | \delta_{m,T})$. We present the necessary derivations in detail in appendix D.

In our data analysis we consider b and r as free parameters. The only restrictions we impose on them are

$$0 < b \, , \, 0 \leq r \leq 1 \ . \qquad \text{(IV.50)}$$

To test how accurately this model describes the behaviour of our mock REDMAGIC catalogs based on the Buzzard N-body simulations we nevertheless want to determine what values of b and r are underlying our simulations. To this end, we generate HEALPIX maps of $\delta_{m,T}$ with different top-hat aperture radii $\theta_T = 10, 20, 30$ arcmin, based on particle count maps at resolution $N_{\text{side}} = 8192$ in slices of comoving $50h^{-1}$Mpc thickness. We co-add these maps to reproduce a redshift range close to that of our fiducial analysis, $z = 0.2100 \ldots 0.4453$. We then select REDMAGIC galaxies with true redshift in this range and determine their counts around the same HEALPIX pixel centers and within the same aperture radii. The REDMAGIC mock catalogs have a complex mask similar to that

of real DES data, which adds complication because the fraction of masked area in each aperture must be equal in order to meaningfully sort lines of sight by galaxy count. To this end, we convert all counts to a masking fraction of 20 per-cent of area within the aperture radius using the stochastic method of Gruen et al. [37, their section 2.1]. This leaves us with simulated 2D maps of $\delta_{m,T}$ and N_T within a DES-Y1 shaped mask.

We can then measure the variances of these maps, $\mathrm{Var}(\delta_{m,T})$ and $\mathrm{Var}(N_T)$, as well as their covariance $\mathrm{Cov}(N_T, \delta_{m,T})$. These fulfill the relations

$$\mathrm{Var}(N_T) = \bar{N} + \bar{N}^2 \, b^2 \mathrm{Var}(\delta_{m,T}) \, . \tag{IV.51}$$

and

$$\mathrm{Cov}(N_T, \delta_{m,T}) = \bar{N} \, b \, r \, \mathrm{Var}(\delta_{m,T}) \tag{IV.52}$$

which then fixes b and r. The values determined in this way are shown in table II.

We now need to check whether these value for b and r together with our assumption of a log-normal PDF for $\delta_{m,T}$ and $\delta_{g,T}$ describe the properties of our tracer galaxies well. Using our simulated maps of $\delta_{m,T}$ and N_T we can measure the expectation value

$$\langle \delta_{g,T} | \delta_{m,T} \rangle = \frac{\langle N_T | \delta_{m,T} \rangle}{\bar{N}} - 1 \tag{IV.53}$$

as a function of $\delta_{m,T}$. Within the log-normal model (cf. appendix D for the relevant formulae) this is very well approximated by

$$\langle \delta_{g,T} | \delta_{m,T} \rangle \approx r b \, \delta_{m,T} \tag{IV.54}$$

which becomes exact for Gaussian random variables. In Figure 6 we compare measurements of $\langle \delta_{g,T} | \delta_{m,T} \rangle$ with the different smoothing radii $\theta_T = 10, 20, 30$ arcmin to the prediction of the log-normal model. We find that in our simulations $\langle \delta_{g,T} | \delta_{m,T} \rangle$ is consistent with a linear relation in $\delta_{m,T}$. Interestingly, the scale dependence of b and r we find in table II almost perfectly cancels to give a scale independent proportionality coefficient

$$r b \approx 1.54 \, . \tag{IV.55}$$

Next, we also measure the variance of galaxy counts N_T as a function of $\delta_{m,T}$ in our simulated maps and compare to the prediction of the log-normal model. In Figure 7 we indeed find that

$$\frac{\mathrm{Var}\,[N_T | \delta_{m,T}]}{\langle N_T | \delta_{m,T} \rangle} \neq 1 \tag{IV.56}$$

and that the $\delta_{m,T}$-dependence is very well described by the log-normal model and the values of b and r we determined before. Finally, we already showed in Figure 1 that also the overall distribution of counts-in-cells N_T is well reproduced by our model.

Smoothing Scale [arcmin]	b	r
10	1.644 ± 0.008	0.938 ± 0.001
20	1.618 ± 0.008	0.956 ± 0.001
30	1.603 ± 0.008	0.961 ± 0.001

TABLE II. Best-fit values galaxy bias and correlation coefficients of our simulated tracer galaxies within the model presented in section IV C 1. Error bars are estimated from a jackknife approach.

Smoothing Scale [arcmin]	$\tilde{b}$	α_0	α_1
10	1.54 ± 0.001	1.15 ± 0.001	0.22 ± 0.003
20	1.54 ± 0.002	1.26 ± 0.002	0.29 ± 0.010
30	1.54 ± 0.002	1.39 ± 0.003	0.45 ± 0.020

TABLE III. Best-fit values galaxy bias and shot-noise parameters of our simulated tracer galaxies within the model presented in section IV C 2. Error bars are again estimated from a jackknife approach.

2. Shot-noise model 2: Parametric model for super-Poissonianity

Our model of shot-noise based on galaxy bias b and galaxy matter correlation coefficient r describes our simulated tracer catalog well. But it contains the arbitrary assumption that both the variance and the skewness of $\delta_{m,T}$ and $\delta_{g,T}$ are related through the bias parameter b (cf. equation IV.47). To account for the possibility that real galaxies might behave in a more complicated way, we also consider a more flexible model of the conditional distribution $P(N_T | \delta_{m,T})$.

Gruen et al. [36] assumed that there is no stochasticity in the relation of $\delta_{m,T}$ and $\delta_{g,T}$ and that galaxies trace the matter density with a linear bias and Poissonian shot noise. This means they set

$$P(N_T = N | \delta_{m,T} = s) =$$
$$\exp\left(N \ln[\bar{N}(1 + \tilde{b}s)] - [\bar{N}(1 + \tilde{b}s)] - \ln \Gamma(N+1) \right) , \tag{IV.57}$$

where $\tilde{b}$ is the galaxy bias, and where we now use a generalizable definition of the Poisson distribution based on the Gamma function Γ, for reasons that will appear later. The galaxy bias $\tilde{b}$ defined this way is not identical to the definition in our fiducial model. We now rather have

$$\delta_{g,T} \equiv \tilde{b} \, \delta_{m,T} \, . \tag{IV.58}$$

We test this in our simulated maps of N_T and $\delta_{m,T}$ by fitting a linear relation to the mean smoothed galaxy contrast as a function of dark matter contrast that was shown in Figure 6. We indeed find that this linear biasing model describes the simulations very well and that $\tilde{b} \approx br$ as expected from our arguments in section IV C 1.

The model used by Gruen et al. [36] however predicts that

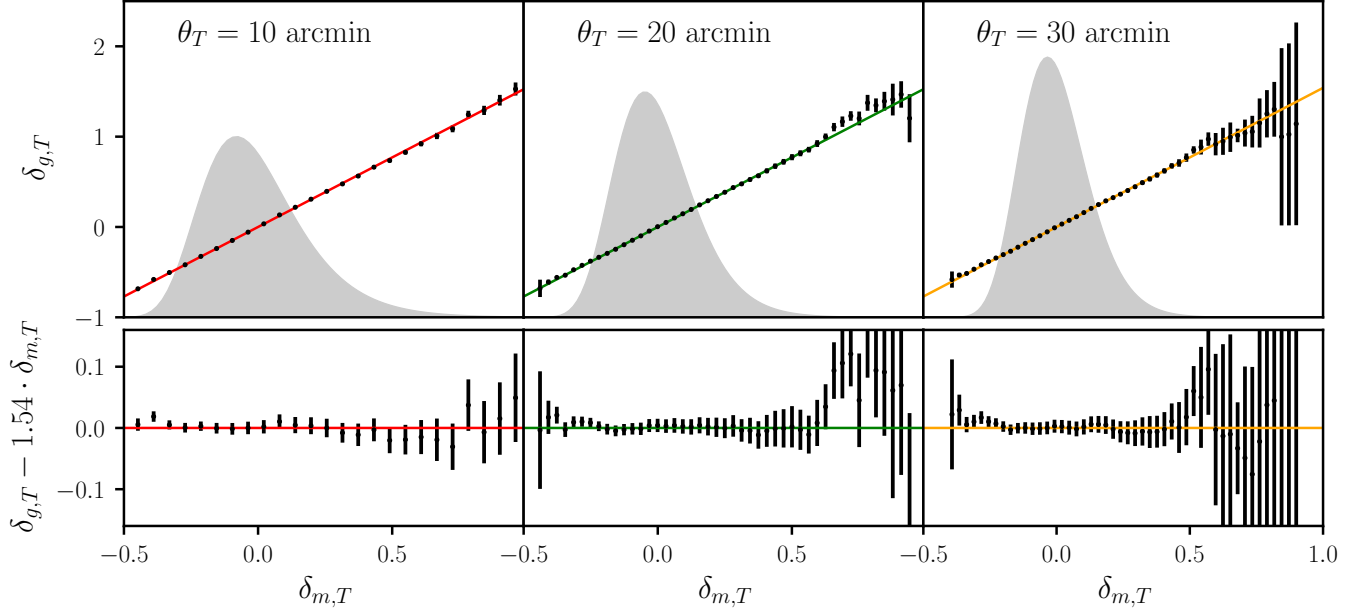$$\frac{\mathrm{Var}\,[N_T | \delta_{m,T}]}{\langle N_T | \delta_{m,T} \rangle} \equiv 1 \tag{IV.59}$$

FIG. 6. Average galaxy overdensity $\delta_{g,T}$ as a function of matter overdensity, $\delta_{m,T}$, for our simulated maps at different smoothing scales: 10 arcmin [left], 20 arcmin [middle] and 30 arcmin [right]. Solid lines show a linear bias model with the bias parameters obtained from maximizing the likelihood in equation IV.63 and the residual between the two are shown in the bottom panels. To indicate the range of $\delta_{m,T}$ that is relevant to our computation, we also show the density PDFs of figure 4 as shaded regions. The errorbars were derived from a jackknife approach.
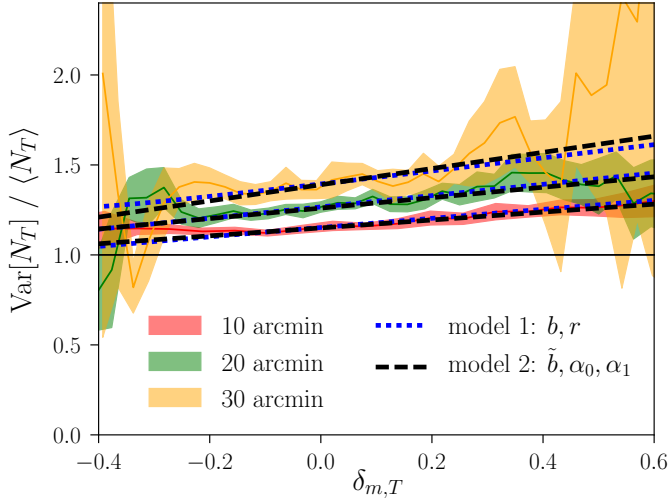
FIG. 7. Ratio of the variance in the galaxy count distribution to average galaxy counts in our simulated maps as a function of matter overdensity. Differently coloured solid lines show the result for each smoothing scale. The dashed black lines show the predictions of the 2-parametric shot-noise model described in section IV C 2. The blue dotted lines show the corresponding prediction of the alternative, 1-parametric model described in section IV C 1. The horizontal solid line shows the expectation if shot noise was purely Poissonian. The coloured regions show the 95% confidence limits derived from jackknife resampling.

which is not what we find in Figure 7. To account for the deviations we observe from pure Poissonian shot-noise, we hence model the distribution of N_T given $\delta_{m,T}$ as

$$P(N_T = N|\delta_{m,T} = s) \sim \mathcal{N} \times$$
$$\exp\left\{\frac{N}{\alpha}\ln\left[\frac{\bar{N}_T}{\alpha}(1+bs)\right] - \ln\Gamma\left(\frac{N+1}{\alpha}\right) - \frac{\bar{N}_T}{\alpha}(1+bs)\right\},$$
$$\text{(IV.60)}$$

where the parameter $\alpha > 0$ generalizes the distribution to one where groups of α galaxies appear with Poissonian noise and where the normalization coefficient $\mathcal{N}$ is needed to ensure that $\int P(N_T = N)\,\mathrm{d}N = 1$. We find $\mathcal{N}$ to be very close to α^{-1} and identical to α^{-1} in the case where α is an integer value.

To account for the observed increase of super-Poissonianity with density, we allow α to depend on $\delta_{m,T}$,

$$\alpha(\delta_{m,T}) = \alpha_0 + \alpha_1 \times \delta_{m,T}. \qquad \text{(IV.61)}$$

This indeed leads to a $\delta_{m,T}$-dependence of the variance of galaxy counts that is close to the relation mentioned in eq. IV.46. In our analysis we treat α_0 and α_1 as free parameters, only requiring that

$$0 < \alpha_0. \qquad \text{(IV.62)}$$

To compare this parametric shot-noise model to our simulations, we are nevertheless interested in the particular value of α_0 and α_1 that describe these simulations. From the tuples of $(N_T, \delta_{m,T})_i$ measured in our simulated maps, we can constrain bias and the $\alpha_{0/1}$ parameters with a likelihood L that is

simply the product of the probabilities of the individual tuples from Equation IV.60,

$$
\begin{aligned}
\ln L = \sum_i & \left[N_i/\alpha(\delta^i_{m,T}) \right] \ln \left[(\bar{N}_i/\alpha(\delta^i_{m,T}))(1 + b \times \delta^i_{m,T}) \right] \\
& - \left[\bar{N}_i/\alpha(\delta^i_{m,T}) \right] (1 + b \times \delta^i_{m,T}) \\
& - \ln G \left[N_i/\alpha(\delta^i_{m,T}) + 1 \right] - \ln \alpha(\delta^i_{m,T}) \, .
\end{aligned}
$$

$$(IV.63)$$

Because the tuples have correlated counts and densities, this is not an exact expression for the likelihood of our measurements, but it should be sufficient to obtain reasonable best fit values for b, α_0 and α_1. We determine the uncertainties of these best-fit values by finding the maximum of Equation IV.63 on jackknife resamplings of the simulations. The resulting parameter values are shown in Table III and displayed in Figure 6 and Figure 7. We find that our simulated REDMAGIC galaxies are indeed well described as linearly biased tracers of the density field with a small, but significant, scale and density dependent super-Poissonian shot-noise.

D. Summary of fiducial model and approximations therein

For each ingredient (i) to (iii) of the framework described in section II we have introduced at least two different modeling ansatzes. We want to once more describe our baseline model built from these ansatzes (cf. also section II B). This is the model we consider in section V and which we use in the data analysis presented in Gruen et al. [37].

(i) $p(\delta_{m,T})$: We find that the log-normal model (section IV A 1) and our model based in cylindrical collapse (section IV A 2) describe the PDF of projected density contrast equally well. The computations based on the log-normal model are however significantly faster. Hence in our fiducial analysis we employ the log-normal model.

(ii) $\langle \kappa_{<\theta} | \delta_{m,T} \rangle$: We also introduced a log-normal model (section IV B 1) and a model based on cylindrical collapse (section IV B 2) for the convergence profile around lines-of-sight with fixed density contrast $\delta_{m,T}$. Both models lead to almost identical predictions for the density split lensing signal. Hence we again choose the log-normal model for our fiducial analysis, because of the shorter computation time.

(iii) $P(N_T | \delta_{m,T})$: We introduced two models for the distribution of tracer counts N_T in lines-of-sight of matter density $\delta_{m,T}$. The first was based on linear galaxy bias b and galaxy-matter-correlation coefficient r (section IV C 1). The second was based on an alternative definition of galaxy bias and on two parameters α_0 and α_1 describing density dependent deviations from Poissonian shot-noise (section IV C 2). Both models describe the behaviour of our simulated tracer galaxies in Buzzard-v1.1 similarly well. But anticipating that real galaxies might behave in a more complicated way, we will consider both ansatzes in our fiducial analysis.

In the following list, we are summarizing the approximations that went into the derivation of our baseline model.

1.) We assumed, that for fixed value of $\delta_{m,T}$ the convergence within angular radius θ is not dependent on N_T (cf. equation II.8).

2.) All second order moments in our formalism are computed with a *halofit* power spectrum [72] using an analytic approximation for the transfer function [30].

3.) Equations IV.5, IV.6, IV.22 and IV.26 employ a small angle and a Limber-like approximation [following 10].

4.) We compute the cumulant generating function of density contrast in long cylinders by means of the cylindrical collapse approximation (cf. section IV A 2).

5.) We assume that the tree-level result of the cumulant generating function can be corrected for the full non-linear evolution of the density field by means of equations IV.16 and IV.17.

6.) We approximate the PDF of $\delta_{m,T}$ resulting from a cylindrical collapse approximation by a log-normal PDF (cf. section IV A 1).

7.) We employed approximations similar to 4.), 5.) and 6.) for the joint distribution of $\delta_{m,T}$ and $\kappa_{<\theta}$ (cf. section IV B 2, equation IV.39 and section IV B 1).

8.) We assume that galaxies are linearly biased tracers of the density field. We consider two different models for shot-noise (resp. stochasticity), assuming that the full distribution $P(N_T | \delta_{m,T})$ is well described bei either two parameters (b, r) or three parameters $(\tilde{b}, \alpha_0, \alpha_1)$.

Despite this long list of approximations, this baseline model describes our measurements in the Buzzard simulations well within DES Y1 errorbars (cf. figure 1). As shown in the next section, the model is also accurate enough to recover the true cosmology of our simulation within DES Y1 uncertainties in a simulated likelihood analysis. In Gruen et al. [37] (and using an extended set of simulations) we furthermore show that the values of χ^2 found between our fiducial model and individual simulation measurements are consistent with the χ^2-distribution expected from our number of data points, and that the coverage (i.e. the fraction of times the true simulation cosmology is within the confidence interval) matches expectations.

V. RECOVERING COSMOLOGY IN N-BODY SIMULATIONS

In this section we want to test whether the modeling that was described in sections II and IV is sufficient to recover the cosmology underlying a density split data vector measured in N-body simulations. The simulations we use are described in section III A. They are the same simulations against which we tested the ingredients of our model in the previous section. A likelihood analysis based on a density split data vector measured in these simulations in presented in section V A.

We only run a cosmological analysis on the mean data vector measured on 4 DES-Y1 realisations. The goal of this is to show that any possible systematic deviations between our modeling of density split statistics and the behaviour of our N-body simulations is smaller than the statistical uncertainties of DES-Y1. A more extensive validation of our likelihood pipeline is presented in Gruen et al. [37].

A. Simulated likelihood analysis

We now measure the data vector that was described in section II C in 4 N-body realisations of DES-Y1. We always use the mean of these 4 data vectors. In order to further reduce the noise of this measurement, we turn off the shape noise in our simulated source catalogs, i.e. we measure our signal directly from the gravitational shear acting each galaxy. We then run Monte-Carlo Markov Chains of our model around this data vector. For this we assume a Gaussian likelihood function with the full covariance (i.e. including shape noise) that was estimated by Gruen et al. [37] for a DES-Y1 data set. The goal of this analysis is to test whether the fiducial cosmology of the Buzzard simulations is well contained within the 1σ contraints derived from this likelihood analysis. A more extensive validation of our likelihood pipeline is presented in Gruen et al. [37].

In the left panel of Figure 8 we show the 1σ and 2σ contraints obtained from our simulated likelihood analysis in the Ω_m-σ_8 plane after marginalizing over different sets of parameters. First, we only vary Ω_m and σ_8, setting other cosmological parameters to the inputs of the Buzzard simulations and the parameters connecting galaxy count and matter density to the values we found from the Buzzard galaxy and density maps. Note that those values would be inaccessible in a real measurement. The corresponding constraints are very tight, but the fiducial values of our parameters are still well contained in the 1σ contour. Then, we also marginalize over the galaxy bias $\tilde{b}$ and the galaxy stochasticity r, demanding that $0 < r \leq 1$. The contours now widen, and the fiducial values of Ω_m and σ_8 are still located well within the corresponding 1σ contour. Finally, we also marginalize over Ω_b, n_s and h_{100}, using the same informative priors that have been used in the DES-Y1 combined probes analysis [27]. The contours widen further, but our model and our simulations still agree well within 1σ uncertainty.

In the right panel of Figure 8 we repeat this analysis, but now also vary the parameter $\Delta S_3/S_3$ that was introduced in section II C. This parameter allows for deviations of the 3-point statistics of the density field from our fiducial model. Within our statistical uncertainties we find that the scaling between 3-point and 2-point statistics in our simulations is well described by our fiducial assumptions ($\Delta S_3/S_3 = 0$).

VI. DISCUSSION & CONCLUSIONS

In this work we introduced density split statistics, a technique to separately measure contributions to weak lensing and
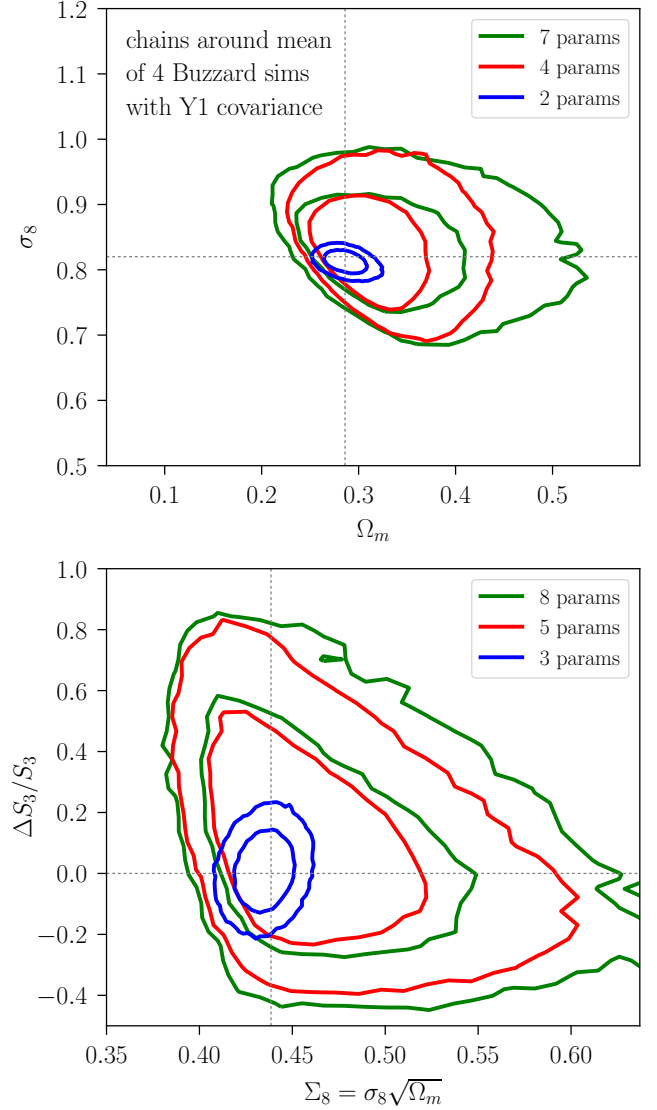

FIG. 8. To test our model for possible systematic deviations from N-body simulations, we try to recover the Buzzard cosmology in a simulated likelihood analysis. <u>Top panel:</u> 1σ and 2σ contours in the Ω_m-σ_8 plane from a likelihood computed around the mean of 4 shape noise free realisations of DES Y1 (but assuming the full covariance matrix for a single DES Y1). The green contours are marginalized over Ω_b, n_s, h_{100}, redMaGiC galaxy bias b and galaxy-matter correlation coefficient r. For the parameters Ω_b, n_s, h_{100} we have assumed the same flat priors as used in the DES Y1 combined probes analysis presented in Abbott et al. (in prep.). The red contours are marginalized only over bias and shot-noise parameters and the blue contours only vary Ω_m and σ_8. Even when going to this small parameter space, our model agrees with Buzzard within 1σ errors of DES Y1. <u>Bottom panel:</u> Same contours but in the Σ_8-$\Delta S_3/S_3$ plane and varying one additional parameter, $\Delta S_3/S_3$.

counts-in-cells from regions of different foreground galaxy density. Based on the pioneering work of Bernardeau [9], Bernardeau & Valageas [10] and Valageas [76] (see also references therein) on modeling the cosmic density PDF we were able to model the density split lensing signal as well as the counts-in-cells histogram from basic principles. With the help of this model, we then showed that density split statistics has two features that make it a potentially powerful cosmological probe:

- it is able to constrain the cosmological parameters Ω_m and σ_8 even if the relation of galaxy density and matter density is assumed to have 2 degrees of freedom: galaxy bias and galaxy-matter-correlation coefficient,

- it is able to constrain the amplitude of 3-point statistics of the density field with almost no degeneracy to constraints on the amplitude of 2-point statistics.

In our fiducial model we predict 3-point statistics from cosmological perturbation theory. Deviations from that fiducial prediction may hint to non-standard physics, that affect overdense and underdense parts of the matter field differently, or to any non-linear dynamics or small scale physics that break the scaling relations of ΛCDM perturbation theory. We showed that a DES-Y5 data set combined with data form the cosmic microwave background can measure the amplitude of 3-points statistics to a 1σ accuracy of $\lesssim 5\%$. This is a conservative estimate since our projections neglect the fact that DES-Y5 will be a deeper data set than DES-Y1. Also, we so far neglected the possibility of a combined analysis including density split statistics and measurements of 2-point correlation functions.

Using measurements in high-resolution N-body simulations we showed that our model of the density split lensing signal and the counts-in-cells histogram is accurate to well within the statistical uncertainties of the DES-Y1 data set. Especially, in a mock likelihood analysis we were able to recover the input cosmology of our simulations to well within DES-Y1 parameter errors. Cosmological constraints from DES-Y1 data based on density split statistics are presented in Gruen et al. [37].

[1] Alam S., et al., 2017, MNRAS, 470, 2617
[2] Baldauf T., Seljak U., Smith R. E., Hamaus N., Desjacques V., 2013, Phys. Rev. D, 88, 083507
[3] Barreira A., Bose S., Li B., Llinares C., 2017, JCAP, 2, 031
[4] Bartelmann M., Schneider P., 2001, Physics Reports, 340, 291
[5] Becker M. R., 2013, MNRAS, 435, 115

[6] Becker M. R., Troxel M. A., MacCrann N., Krause E., Eifler T. F., Friedrich O., Nicola A., The DES Collaboration 2015, preprint, (arXiv:1507.05598)

[7] Benítez N., 2000, ApJ, 536, 571

[8] Benjamin J., et al., 2007, MNRAS, 381, 702

[9] Bernardeau F., 1994, A&A, 291, 697

[10] Bernardeau F., Valageas P., 2000, A&A, 364, 1

[11] Bernardeau F., van Waerbeke L., Mellier Y., 1997, A&A, 322, 1

[12] Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Physics Reports, 367, 1

[13] Bernardeau F., Pichon C., Codis S., 2014, Phys. Rev. D, 90, 103519

[14] Bernardeau F., Codis S., Pichon C., 2015, MNRAS, 449, L105

[15] Borisov A., Jain B., 2009, Phys. Rev. D, 79, 103506

[16] Brainerd T. G., Blandford R. D., Smail I., 1996, ApJ, 466, 623

[17] Brimioulle F., Seitz S., Lerchster M., Bender R., Snigula J., 2013, MNRAS, 432, 1046

[18] Chiang C.-T., Wagner C., Sánchez A. G., Schmidt F., Komatsu E., 2015, Journal of Cosmology and Astroparticle Physics, 9, 028

[19] Clampitt J., et al., 2017, MNRAS, 465, 4204

[20] Clerkin L., et al., 2017, MNRAS, 466, 1444

[21] Codis S., Pichon C., Bernardeau F., Uhlemann C., Prunet S., 2016a, MNRAS, 460, 1549

[22] Codis S., Bernardeau F., Pichon C., 2016b, MNRAS, 460, 1598

[23] Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201

[24] Cooper M. C., et al., 2011, ApJS, 193, 14

[25] Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369

[26] Crocce M., et al., 2016, Monthly Notices of the Royal Astronomical Society, 455, 4301

[27] DES Collaboration 2017, preprint, (arXiv:1708.01530)

[28] DeRose J., Wechsler R., Rykoff E., et al., 2017, in prep.

[29] Dekel A., Lahav O., 1999, ApJ, 520, 24

[30] Eisenstein D. J., Hu W., 1998, ApJ, 496, 605

[31] Elvin-Poole J., et al., 2017, preprint, (arXiv:1708.01536)

[32] Fosalba P., Gaztanaga E., 1998, MNRAS, 301, 503

[33] Friedrich O., Seitz S., Eifler T. F., Gruen D., 2016, MNRAS, 456, 2662

[34] Fu L., et al., 2008, A&A, 479, 9

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

[36] Gruen D., Friedrich O., The DES Collaboration 2016, Monthly Notices of the Royal Astronomical Society, 455, 3367

[37] Gruen D., Friedrich O., et al., 2017, submitted to PRD

[38] Hilbert S., Hartlap J., Schneider P., 2011, Astronomy & Astrophysics, 536, A85

[39] Hildebrandt H., et al., 2017, MNRAS, 465, 1454

[40] Hoyle B., et al., 2017, preprint, (arXiv:1708.01532)

[41] Hudson M. J., Gwyn S. D. J., Dahle H., Kaiser N., 1998, ApJ, 503, 531

[42] Huff E., Mandelbaum R., 2017, preprint, (arXiv:1702.02600)

[43] Jain B., Zhang P., 2008, Phys. Rev. D, 78, 063503

[44] Jee M. J., Tyson J. A., Hilbert S., Schneider M. D., Schmidt S., Wittman D., 2016, ApJ, 824, 77

[45] Jeong D., Komatsu E., 2009, ApJ, 703, 1230

[46] Kacprzak T., et al., 2016, MNRAS, 463, 3653

[47] Kilbinger M., et al., 2013, Monthly Notices of the Royal Astronomical Society, 430, 2200

[48] Limber D. N., 1953, ApJ, 117, 134

[49] Lin C.-A., Kilbinger M., 2015, A&A, 576, A24

[50] Liu X., et al., 2015, MNRAS, 450, 2888

[51] Lue A., Scoccimarro R., Starkman G., 2004, Phys. Rev. D, 69, 044005

[52] MacCrann N., DeRose J., Wechsler R., et al., 2017, to be submitted to PRD

[53] Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2013, MNRAS, 432, 1544

[54] Mantz A. B., et al., 2016, MNRAS, 463, 3582

[55] Miyazaki S., et al., 2002, PASJ, 54, 833

[56] Mukhanov V., 2005, Physical Foundations of Cosmology, doi:10.2277/0521563984.

[57] Multamäki T., Gaztañaga E., Manera M., 2003, MNRAS, 344, 761

[58] Pires S., Leonard A., Starck J.-L., 2012, MNRAS, 423, 983

[59] Planck Collaboration et al., 2015, preprint, (arXiv:1502.01589)

[60] Prat J., et al., 2017, preprint, (arXiv:1708.01537)

[61] Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, ApJ, 771, 30

[62] Schrabback T., et al., 2010, A&A, 516, A63

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

[64] Sheldon E. S., Huff E. M., 2017, preprint, (arXiv:1702.02601)

[65] Simpson F., James J. B., Heavens A. F., Heymans C., 2011, Physical Review Letters, 107, 271301

[66] Simpson F., Heavens A. F., Heymans C., 2013, Phys. Rev. D, 88, 083510

[67] Simpson F., Harnois-Déraps J., Heymans C., Jimenez R., Joachimi B., Verde L., 2016, Monthly Notices of the Royal Astronomical Society, 456, 278

[68] Smith R. E., et al., 2003, Monthly Notices of the Royal Astronomical Society, 341, 1311

[69] Springel V., 2005, MNRAS, 364, 1105

[70] Springel V., et al., 2005, Nature, 435, 629

[71] Takada M., Jain B., 2002, MNRAS, 337, 875

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

[73] Takahashi R., Soma S., Takada M., Kayo I., 2014, Monthly Notices of the Royal Astronomical Society, 444, 3473

[74] Troxel M. A., et al., 2017, preprint, (arXiv:1708.01538)

[75] Uhlemann C., et al., 2017, preprint, (arXiv:1705.08901)

[76] Valageas P., 2002, A&A, 382, 412

[77] Van Waerbeke L., et al., 2000, A&A, 358, 30

[78] Wechsler R., DeRose J., Busha M., et al., 2017, in prep.

[79] Wilson G., Kaiser N., Luppino G. A., Cowie L. L., 2001, ApJ, 555, 572

[80] Wittman D. M., Tyson J. A., Kirkman D., Dell'Antonio I., Bernstein G., 2000, Nature, 405, 143

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

[82] Zuntz J., et al., 2017, submitted to Mon. Not. R. Astron. Soc.

[83] van Uitert E., Hoekstra H., Velander M., Gilbank D. G., Gladders M. D., Yee H. K. C., 2011, A&A, 534, A14

[84] van Uitert E., et al., 2017, preprint, (arXiv:1706.05004)

Appendix A: Friedmann equations, linear growth, spherical collapse and cylindrical collapse

Throughout this section we set $G = 1 = c$ and we assume a flat ΛCDM universe. In proper co-moving time t the Fried-

mann equations take the form

$$H^2 = \frac{8\pi}{3} \left(\bar{\rho}_m + \bar{\rho}_\Lambda \right) \tag{A.1}$$

$$\frac{\mathrm{d}H}{\mathrm{d}t} + H^2 = -\frac{4\pi}{3} \left(\bar{\rho}_m - 2\bar{\rho}_\Lambda \right) , \tag{A.2}$$

where $H = \frac{\mathrm{d}}{\mathrm{d}t} \ln a$. In conformal time, defined by $\mathrm{d}t = a\mathrm{d}\tau$, this changes to

$$\mathcal{H}^2 = \frac{8\pi}{3} a^2 \left(\bar{\rho}_m + \bar{\rho}_\Lambda \right) \tag{A.3}$$

$$\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}\tau} = -\frac{4\pi}{3} a^2 \left(\bar{\rho}_m - 2\bar{\rho}_\Lambda \right) , \tag{A.4}$$

where $\mathcal{H} = \frac{\mathrm{d}}{\mathrm{d}\tau} \ln a$. We will from now put $\frac{\mathrm{d}}{\mathrm{d}\tau} \equiv \dot{} $.

In the Newtonian approximation, i.e. on scales much smaller that the curvature horizon of the universe, the evolution of a spherical, cylindrical or planar perturbation δ is given

by the equation

$$\frac{\mathrm{d}^2\delta}{\mathrm{d}t^2} + 2H\frac{\mathrm{d}\delta}{\mathrm{d}t} - \frac{N+1}{N}\frac{1}{1+\delta}\left(\frac{\mathrm{d}\delta}{\mathrm{d}t}\right)^2 = 4\pi\bar{\rho}_m\delta(1+\delta) , \tag{A.5}$$

where $N = 3$ for a spherical perturbation, $N = 2$ for a cylindrical perturlation and $N = 1$ for a planar perturbation (see Mukhanov [56] where this is demonstrated for $N = 3$ and $N = 1$).

In conformal time this equation reads

$$\ddot{\delta} + \mathcal{H}\dot{\delta} - \frac{N+1}{N}\frac{\dot{\delta}^2}{1+\delta} = 4\pi\bar{\rho}_m a^2\delta(1+\delta) . \tag{A.6}$$

To linear order in δ this becomes

$$\ddot{\delta} + \mathcal{H}\dot{\delta} = 4\pi\bar{\rho}_m a^2\delta , \tag{A.7}$$

which is indeed independent of the particular shape of the perturbation.

Appendix B: ΛCDM perturbation theory

Consider the matter density contrast δ and the divergence of the velocity field $\theta = \boldsymbol{\nabla}\mathbf{v}$. In the Newtonian approximation the Fourier space equations of motion of δ and θ are [cf. 12]

$$\frac{\partial\tilde{\delta}(\mathbf{k},\tau)}{\partial\tau} + \tilde{\theta}(\mathbf{k},\tau) = -\int \mathrm{d}^3k_1\mathrm{d}^3k_2 \, \delta_D(\mathbf{k}-\mathbf{k}_{12}) \, \alpha(\mathbf{k}_1,\mathbf{k}_2) \, \tilde{\delta}(\mathbf{k}_1,\tau) \, \tilde{\theta}(\mathbf{k}_2,\tau)$$

$$\frac{\partial\tilde{\theta}(\mathbf{k},\tau)}{\partial\tau} + \mathcal{H}\,\tilde{\theta}(\mathbf{k},\tau) + \frac{3\Omega_m^0 H_0^2}{2a}\tilde{\delta}(\mathbf{k},\tau) = -\int \mathrm{d}^3k_1\mathrm{d}^3k_2 \, \delta_D(\mathbf{k}-\mathbf{k}_{12}) \, \beta(\mathbf{k}_1,\mathbf{k}_2) \, \tilde{\theta}(\mathbf{k}_1,\tau) \, \tilde{\theta}(\mathbf{k}_2,\tau) , \tag{B.1}$$

where $\mathbf{k}_{12} = \mathbf{k}_1 + \mathbf{k}_2$ and α and β are given by

$$\alpha(\mathbf{k}_1,\mathbf{k}_2) = 1 + \frac{1}{2}\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right)$$

$$\beta(\mathbf{k}_1,\mathbf{k}_2) = \frac{1}{2}\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} . \tag{B.2}$$

In the following we will abbreviate the integrals involving α and β as $\alpha[\tilde{\delta}, \tilde{\theta}, \mathbf{k}]$ and $\beta[\tilde{\theta}, \tilde{\theta}, \mathbf{k}]$. In perturbation theory we write $\tilde{\delta}$ and $\tilde{\theta}$ as

$$\tilde{\delta}(\mathbf{k},\tau) = \sum_{n=1}^{\infty} \delta_n(\mathbf{k},\tau) \;\; \text{and} \;\; \tilde{\theta}(\mathbf{k},\tau) = -\frac{\partial \ln D_+(\tau)}{\partial\tau}\sum_{n=1}^{\infty} \theta_n(\mathbf{k},\tau) , \tag{B.3}$$

where δ_n and θ_n are of order n in the linearly approximated fields δ_1 and θ_1 and $D_+(\tau)$ is the linear growth factor. (We will ignore the decaying mode of linear growth here.) At linear order we have

$$\delta_1(\mathbf{k},\tau) = \theta_1(\mathbf{k},\tau) = \frac{D_+(\tau)}{D_+(\tau_0)}\delta_1(\mathbf{k},\tau_0) \equiv D_+(\tau)\delta_{1,1}(\mathbf{k}) , \tag{B.4}$$

where we have assumed that $D_+(\tau_0) = 1$ at present time τ_0 and introduced the notation $\delta_{1,1}(\mathbf{k}) = \delta_1(\mathbf{k},\tau_0)$ whose purpose will become clear at the end of this section. To get $\tilde{\delta}$ at second order we first note that

$$\frac{\partial\tilde{\theta}(\mathbf{k},\tau)}{\partial\tau} + \mathcal{H}\,\tilde{\theta}(\mathbf{k},\tau) = \frac{1}{a}\frac{\partial}{\partial\tau}\left(a\tilde{\theta}(\mathbf{k},\tau)\right) . \tag{B.5}$$

Hence, multiplying the first of equations B.1 with a and differentiating wrt. τ and then multiplying with $1/a$ we get

$$\frac{\partial^2 \tilde{\delta}(\mathbf{k},\tau)}{\partial^2 \tau} + \mathcal{H}\frac{\partial \tilde{\delta}(\mathbf{k},\tau)}{\partial \tau} + \frac{\partial \tilde{\theta}(\mathbf{k},\tau)}{\partial \tau} + \mathcal{H}\,\tilde{\theta}(\mathbf{k},\tau) = -\alpha\left[\frac{\partial \tilde{\delta}}{\partial \tau}, \tilde{\theta}, \mathbf{k}\right] - \alpha\left[\tilde{\delta}, \frac{\partial \tilde{\theta}}{\partial \tau}, \mathbf{k}\right] - \mathcal{H}\,\alpha\left[\tilde{\delta}, \tilde{\theta}, \mathbf{k}\right] . \tag{B.6}$$

Now the second of equations B.1 can be used to eliminate $\tilde{\theta}$ from the right-hand-side, giving

$$\frac{\partial^2 \tilde{\delta}(\mathbf{k},\tau)}{\partial^2 \tau} + \mathcal{H}\frac{\partial \tilde{\delta}(\mathbf{k},\tau)}{\partial \tau} - \frac{3\Omega_m^0 H_0^2}{2a}\tilde{\delta}(\mathbf{k},\tau) = \beta[\tilde{\theta}, \tilde{\theta}, \mathbf{k}] - \alpha[\frac{\partial \tilde{\delta}}{\partial \tau}, \tilde{\theta}, \mathbf{k}] - \alpha[\tilde{\delta}, \frac{\partial \tilde{\theta}}{\partial \tau}, \mathbf{k}] - \mathcal{H}\,\alpha[\tilde{\delta}, \tilde{\theta}, \mathbf{k}] . \tag{B.7}$$

At second order in perturbation theory this equation becomes

$$\frac{\partial^2 \delta_2(\mathbf{k},\tau)}{\partial^2 \tau} + \mathcal{H}\frac{\partial \delta_2(\mathbf{k},\tau)}{\partial \tau} - \frac{3\Omega_m^0 H_0^2}{2a}\delta_2(\mathbf{k},\tau) = \left(\frac{\partial D_+}{\partial \tau}\right)^2 \beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] + \left(D\frac{\partial^2 D_+}{\partial \tau^2} + D\mathcal{H}\frac{\partial D_+}{\partial \tau} + \left(\frac{\partial D_+}{\partial \tau}\right)^2\right)\alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}]$$

$$= \left(\frac{\partial D_+}{\partial \tau}\right)^2 \beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] + \left(\frac{3\Omega_m^0 H_0^2}{2a}D^2 + \left(\frac{\partial D_+}{\partial \tau}\right)^2\right)\alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}]$$

$$= \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}]\left(\frac{3\Omega_m^0 H_0^2}{2a}D^2 + 2\left(\frac{\partial D_+}{\partial \tau}\right)^2\right)$$

$$+ (\beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] - \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}])\left(\frac{\partial D_+}{\partial \tau}\right)^2 . \tag{B.8}$$

This is solved by

$$\delta_2(\mathbf{k},\tau) = D_{2,1}(\tau)\delta_{2,1}(\mathbf{k}) + D_{2,2}(\tau)\delta_{2,2}(\mathbf{k}) \tag{B.9}$$

where

$$D_{2,1}(\tau) \equiv D_+^2(\tau) \,,\; \delta_{2,1}(\mathbf{k}) = \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] \,,\; \delta_{2,2}(\mathbf{k}) = \beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] - \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] \tag{B.10}$$

and $D_{2,2}$ is given by the differential equation

$$\frac{\partial^2 D_{2,2}(\tau)}{\partial^2 \tau} + \mathcal{H}\frac{\partial D_{2,2}(\tau)}{\partial \tau} - \frac{3\Omega_m^0 H_0^2}{2a}D_{2,2}(\tau) = \left(\frac{\partial D_+}{\partial \tau}\right)^2 . \tag{B.11}$$

1. Second order of δ in Einstein-de Sitter universe

Let us define $1 - \mu \equiv D_{2,2}/D_+^2$. Then the general solution to B.8 is given by

$$\delta_2(\mathbf{k},\tau) = D_+^2 \left([1 - \mu]\beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] + \mu\alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}]\right) . \tag{B.12}$$

In an Einstein-de Sitter Universe where $\Omega_m^0 = 1$ and $D \equiv a$ we have

$$D_{2,2} = \frac{2}{7}D_+^2 \,,\; \mu = \frac{5}{7} \tag{B.13}$$

and δ_2 is hence given by

$$\delta_2(\mathbf{k},\tau) = D_+^2 \left(\frac{2}{7}\beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] + \frac{5}{7}\alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}]\right)$$

$$= \int d^3k_1 d^3k_2\, \delta_D(\mathbf{k} - \mathbf{k}_{12})\, F_2(\mathbf{k}_1, \mathbf{k}_2)\, \delta_{1,1}(\mathbf{k}_1)\, \delta_{1,1}(\mathbf{k}_2) \tag{B.14}$$

with

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7}\alpha(\mathbf{k}_1, \mathbf{k}_2) + \frac{2}{7}\beta(\mathbf{k}_1, \mathbf{k}_2)$$

$$= \frac{5}{7} + \frac{1}{2}\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{2}{7}\frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} . \tag{B.15}$$

2. Second order of δ in ΛCDM universe

In a general ΛCDM universe the function F_2 becomes time dependent. It is given by

$$F_2(\mathbf{k}_1, \mathbf{k}_2, \tau) = \mu(\tau) + \frac{1}{2} \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + [1 - \mu(\tau)] \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} \ . \tag{B.16}$$

A useful property or this kernel is that

$$F_2(\mathbf{k}, -\mathbf{k}, \tau) = \mu(\tau) + \frac{1}{2} \frac{-k^2}{k^2} (1 + 1) + [1 - \mu(\tau)] \frac{k^4}{k^4} = \mu(\tau) - 1 + 1 - \mu(\tau) = 0 \ . \tag{B.17}$$

Denoting the angle between $\mathbf{k}_1$ and $\mathbf{k}_2$ with ϕ one can also arive at the following form of $F_2(\mathbf{k}_1, \mathbf{k}_2, \tau)$ which will be useful when computing the skewness of matter inside a long cylinder:

$$F_2(\mathbf{k}_1, \mathbf{k}_2, \tau) = \frac{1}{2} \left\{ \left(1 + \frac{k_1}{k_2} \cos \phi \right) + \left(1 + \frac{k_2}{k_1} \cos \phi \right) \right\} + [1 - \mu(\tau)](\cos^2 \phi - 1) \ . \tag{B.18}$$

3. Bispectrum and 3-point function at leading order

The bispectrum $B(k_1, k_2, k_3, \tau)$ of δ is defined by

$$\langle \tilde{\delta}(\mathbf{k}_1, \tau) \tilde{\delta}(\mathbf{k}_2, \tau) \tilde{\delta}(\mathbf{k}_3, \tau) \rangle = \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \, B(k_1, k_2, k_3, \tau) \ . \tag{B.19}$$

At leading order in perturbation theory this can be calculated as

$$\langle \tilde{\delta}(\mathbf{k}_1, \tau) \tilde{\delta}(\mathbf{k}_2, \tau) \tilde{\delta}(\mathbf{k}_3, \tau) \rangle_{\text{2nd.}} = D_+^2 \int \mathrm{d}^3 q_1 \mathrm{d}^3 q_2 \, \delta_D(\mathbf{k}_3 - \mathbf{q}_{12}) \, F_2(\mathbf{q}_1, \mathbf{q}_2, \tau) \, \langle \delta_{1,1}(\mathbf{k}_1) \, \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \, \delta_{1,1}(\mathbf{q}_2) \rangle$$
$$+ \text{cycl.} \ , \tag{B.20}$$

where $'$cycl.$'$ indicates that the integral on the right-hand-side should appear for all possible permutations of $\mathbf{k}_1$, $\mathbf{k}_3$ and $\mathbf{k}_3$. Since we assume the linear density field to be a Gaussian random field, the expectation value on the left-hand-side factorizes as

$$\langle \delta_{1,1}(\mathbf{k}_1)\delta_{1,1}(\mathbf{k}_2)\delta_{1,1}(\mathbf{q}_1)\delta_{1,1}(\mathbf{q}_2) = \langle \delta_{1,1}(\mathbf{k}_1)\delta_{1,1}(\mathbf{k}_2)\rangle \langle \delta_{1,1}(\mathbf{q}_1)\delta_{1,1}(\mathbf{q}_2)\rangle + \langle \delta_{1,1}(\mathbf{k}_1)\delta_{1,1}(\mathbf{q}_1)\rangle \langle \delta_{1,1}(\mathbf{k}_2)\delta_{1,1}(\mathbf{q}_2)\rangle$$
$$+ \langle \delta_{1,1}(\mathbf{k}_1)\delta_{1,1}(\mathbf{q}_2)\rangle \langle \delta_{1,1}(\mathbf{k}_2)\delta_{1,1}(\mathbf{q}_1)\rangle$$
$$= \delta_D(\mathbf{k}_1 + \mathbf{k}_2)\delta_D(\mathbf{q}_1 + \mathbf{q}_2)P_{\text{lin},0}(k_1)P_{\text{lin},0}(q_1) + \delta_D(\mathbf{k}_1 + \mathbf{q}_1)\delta_D(\mathbf{k}_2 + \mathbf{q}_2)P_{\text{lin},0}(k_1)P_{\text{lin},0}(q_2)$$
$$+ \delta_D(\mathbf{k}_1 + \mathbf{q}_2)\delta_D(\mathbf{k}_2 + \mathbf{q}_1)P_{\text{lin},0}(k_1)P_{\text{lin},0}(q_1) \tag{B.21}$$

Because of equation B.17 the contribution of the first term to the bispectrum is zero. Using the symmetry $1 \leftrightarrow 2$ between the second and third term we hence get

$$\langle \tilde{\delta}(\mathbf{k}_1, \tau) \tilde{\delta}(\mathbf{k}_2, \tau) \tilde{\delta}(\mathbf{k}_3, \tau) \rangle_{\text{2nd.}} = 2D_+^2 \, \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \, F_2(\mathbf{k}_1, \mathbf{k}_2, \tau) P_{\text{lin},0}(k_1) P_{\text{lin},0}(k_2)$$
$$+ \text{cycl.} \ . \tag{B.22}$$

4. Variance and skewness of long cylinder at leading order in perturbation theory

Consider a cylinder with radius R and length L. In Fourier space the tophat filter for this cylinder is given by

$$W_{R,L}(\mathbf{k}) = \frac{1}{(2\pi)^3} W_L(k_\parallel) W_R(\mathbf{k}_\perp) \tag{B.23}$$

where we denote the component of $\mathbf{k}$ parallel to the cylinder axis with $k_\parallel$ and the components orthogonal to it are represented by the two-dimensional vector $\mathbf{k}_\perp$ and W_L and W_R given by

$$W_L(k_\parallel) = \frac{\sin(Lk_\parallel/2)}{Lk_\parallel/2} \ , \ W_R(\mathbf{k}_\perp) = \frac{2J_1(Rk_\perp)}{Rk_\perp} \ . \tag{B.24}$$

Here $k_\perp = |\mathbf{k}_\perp|$ and J_ν are the cylindrical bessel functions. At leading order or tree level in perturbation theory the variance of matter contrast within the cylinder is then given by

$$\langle \delta_{R,L}^2 \rangle_{\text{tree}}(\tau) = D_+^2(\tau) \int \mathrm{d}k_{\parallel,1} \mathrm{d}k_{\parallel,2} \mathrm{d}^2 k_{\perp,1} \mathrm{d}^2 k_{\perp,2} \, W_L(k_{\parallel,1}) \, W_L(k_{\parallel,2}) \, W_R(\mathbf{k}_{\perp,1}) \, W_R(\mathbf{k}_{\perp,2}) \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \rangle$$

$$= D_+^2(\tau) \int \mathrm{d}k_\parallel \mathrm{d}^2 k_\perp \, W_L(k_\parallel)^2 \, W_R(\mathbf{k}_\perp)^2 \, P_{\text{lin},0}(\mathbf{k}) \,. \tag{B.25}$$

Here $P_{\text{lin},0}(\mathbf{k})$ is today's linear power spectrum. For $L \gg R$ we can actually approximate W_L by

$$W_L(k_\parallel)^2 \approx \frac{2\pi}{L} \delta_D(k_\parallel) \tag{B.26}$$

such that in this limit we get

$$\langle \delta_{R,L}^2 \rangle_{\text{tree}}(\tau) \approx \frac{2\pi D_+^2}{L} \int \mathrm{d}k \, k \, W_R(k)^2 \, P_{\text{lin},0}(k) \,. \tag{B.27}$$

The third moment at tree level is given by

$$\langle \delta_{R,L}^3 \rangle_{\text{tree}}(\tau) = 3D_+^2 \int \mathrm{d}k_{\parallel,1} \mathrm{d}k_{\parallel,2} \mathrm{d}q_\parallel \mathrm{d}^2 k_{\perp,1} \mathrm{d}^2 k_{\perp,2} \mathrm{d}^2 q_\perp \, W_L(k_{\parallel,1}) \, W_L(k_{\parallel,2}) \, W_L(q_\parallel) \, W_R(\mathbf{k}_{\perp,1}) \, W_R(\mathbf{k}_{\perp,2}) \, W_R(\mathbf{q}_\perp)$$

$$\langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \delta_2(\mathbf{q}, \tau) \rangle$$

$$= 3D_+^4 \int \mathrm{d}k_{\parallel,1} \mathrm{d}k_{\parallel,2} \mathrm{d}q_\parallel \mathrm{d}^2 k_{\perp,1} \mathrm{d}^2 k_{\perp,2} \mathrm{d}^2 q_\perp \mathrm{d}^3 q_1 \mathrm{d}^3 q_2 \, W_L(k_{\parallel,1}) \, W_L(k_{\parallel,2}) \, W_L(q_\parallel) \, W_R(\mathbf{k}_{\perp,1}) \, W_R(\mathbf{k}_{\perp,2}) \, W_R(\mathbf{q}_\perp)$$

$$\delta_D(\mathbf{q} - \mathbf{q}_1 - \mathbf{q}_2) F_2(\mathbf{q}_1, \mathbf{q}_2, \tau) \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \delta_{1,1}(\mathbf{q}_2) \rangle$$

$$= 3D_+^4 \int \mathrm{d}k_{\parallel,1} \mathrm{d}k_{\parallel,2} \mathrm{d}^2 k_{\perp,1} \mathrm{d}^2 k_{\perp,2} \mathrm{d}^3 q_1 \mathrm{d}^3 q_2 \, W_L(k_{\parallel,1}) \, W_L(k_{\parallel,2}) \, W_R(\mathbf{k}_{\perp,1}) \, W_R(\mathbf{k}_{\perp,2})$$

$$W_R(\mathbf{q}_{\perp,1} + \mathbf{q}_{\perp,2}) \, W_L(q_{\parallel,1} + q_{\parallel,2}) F_2(\mathbf{q}_1, \mathbf{q}_2, \tau) \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \delta_{1,1}(\mathbf{q}_2) \rangle \,. \tag{B.28}$$

Since we assume the linear density field to be a Gaussian random field, the expectation value on the left-hand-side factorizes as

$$\langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \delta_{1,1}(\mathbf{q}_2) = \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \rangle \langle \delta_{1,1}(\mathbf{q}_1) \delta_{1,1}(\mathbf{q}_2) \rangle + \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{q}_1) \rangle \langle \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_2) \rangle$$

$$+ \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{q}_2) \rangle \langle \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \rangle$$

$$= \delta_D(\mathbf{k}_1 + \mathbf{k}_2) \delta_D(\mathbf{q}_1 + \mathbf{q}_2) P_{\text{lin},0}(k_1) P_{\text{lin},0}(q_1) + \delta_D(\mathbf{k}_1 + \mathbf{q}_1) \delta_D(\mathbf{k}_2 + \mathbf{q}_2) P_{\text{lin},0}(k_1) P_{\text{lin},0}(q_2)$$

$$+ \delta_D(\mathbf{k}_1 + \mathbf{q}_2) \delta_D(\mathbf{k}_2 + \mathbf{q}_1) P_{\text{lin},0}(k_1) P_{\text{lin},0}(q_1) \tag{B.29}$$

Because of equation B.17 the contribution of the first term to the skewness is zero. Using the symmetry $1 \leftrightarrow 2$ between the second and third term we hence get

$$\langle \delta_{R,L}^3 \rangle_{\text{tree}}(\tau) \rangle = 6D_+^4 \int \mathrm{d}q_{\parallel,1} \mathrm{d}q_{\parallel,2} \mathrm{d}^2 q_{\perp,1} \mathrm{d}^2 q_{\perp,2} \, W_L(q_{\parallel,1}) \, W_L(q_{\parallel,2}) \, W_L(q_{\parallel,1} + q_{\parallel,2}) \, W_R(\mathbf{q}_1) \, W_R(\mathbf{q}_2) \, W_R(\mathbf{q}_1 + \mathbf{q}_2)$$

$$P_{\text{lin},0}(q_1) P_{\text{lin},0}(q_2) F_2(\mathbf{q}_1, \mathbf{q}_2, \tau) \,. \tag{B.30}$$

For $L \gg R$ we can use the approximation

$$W_L(q_{\parallel,1}) \, W_L(q_{\parallel,2}) \, W_L(q_{\parallel,1} + q_{\parallel,2}) \approx \frac{1}{L^2} \delta_D^2(q_{\parallel,1}, q_{\parallel,2}) \,. \tag{B.31}$$

This gives

$$\langle \delta_{R,L}^3 \rangle_{\text{tree}}(\tau) \rangle = \frac{6D_+^4}{L^2} \int \mathrm{d}^2 q_1 \mathrm{d}^2 q_2 \, W_R(\mathbf{q}_1) \, W_R(\mathbf{q}_2) \, W_R(\mathbf{q}_1 + \mathbf{q}_2) \, P_{\text{lin},0}(q_1) P_{\text{lin},0}(q_2) F_2(\mathbf{q}_1, \mathbf{q}_2, \tau) \,, \tag{B.32}$$

where we will consider all vectors to be 2-dimensional from now on. Using equation B.18 to express F_2 interms of q_1, q_2 and ϕ we can simplify this to

$$\langle \delta_{R,L}^3 \rangle_{\text{tree}}(\tau) \rangle = \frac{12\pi D_+^4}{L^2} \int \mathrm{d}q_1 \mathrm{d}q_2 \, q_1 W_R(q_1) \, q_2 W_R(q_2) \, P_{\text{lin},0}(q_1) P_{\text{lin},0}(q_2) \int \mathrm{d}\phi \, W_R \left[\sqrt{q_1^2 + q_2^2 + 2q_1 q_2 \cos \phi} \right] F_2(q_1, q_2, \phi, \tau) \,. \tag{B.33}$$

Using relations given in Bernardeau (1995) or Buchalter et al. (2000) one can simplify the integral over ϕ as

$$
\int \mathrm{d}\phi \, W_R\left[\sqrt{q_1^2 + q_2^2 + 2q_1 q_2 \cos\phi}\right] F_2(q_1, q_2, \phi, \tau)
$$

$$
= \frac{1}{2} \int \mathrm{d}\phi \, W_R\left[\sqrt{q_1^2 + q_2^2 + 2q_1 q_2 \cos\phi}\right] \left\{ \left(1 + \frac{k_1}{k_2}\cos\phi\right) + \left(1 + \frac{k_2}{k_1}\cos\phi\right) \right\}
$$

$$
+ [1 - \mu(\tau)] \int \mathrm{d}\phi \, W_R\left[\sqrt{q_1^2 + q_2^2 + 2q_1 q_2 \cos\phi}\right] (\cos^2\phi - 1)
$$

$$
= 2\pi W_R(q_1) W_R(q_2) + \frac{\pi}{2}\left\{ W_R(q_1)\, Rq_2\, \left.\frac{\partial W_R(x)}{\partial x}\right|_{x=Rq_2} + W_R(q_2)\, Rq_1\, \left.\frac{\partial W_R(x)}{\partial x}\right|_{x=Rq_1} \right\} - \pi[1 - \mu(\tau)] W_R(q_1) W_R(q_2)
$$

$$
= \pi[1 + \mu(\tau)] W_R(q_1) W_R(q_2) + \frac{\pi}{2}\frac{\partial}{\partial \ln R}\left\{ W_R(q_1) W_R(q_2) \right\} \ . \tag{B.34}
$$

For the third moment of $\delta_{R,L}$ this gives

$$
\langle \delta_{R,L}^3 \rangle_{\mathrm{tree}}(\tau) = [1 + \mu(\tau)]\frac{12\pi^2 D_+^4(\tau)}{L^2} \int \mathrm{d}q_1 \mathrm{d}q_2 \, q_1 q_2 \, W_R(q_1)^2 \, W_R(q_2)^2 \, P_{\mathrm{lin},0}(q_1) P_{\mathrm{lin},0}(q_2)
$$

$$
+ \frac{6\pi^2 D_+^4(\tau)}{L^2} \int \mathrm{d}q_1 \mathrm{d}q_2 \, q_1 q_2 \, W_R(q_1)\, W_R(q_2) \frac{\partial}{\partial \ln R}\left\{ W_R(q_1) W_R(q_2) \right\} P_{\mathrm{lin},0}(q_1) P_{\mathrm{lin},0}(q_2)
$$

$$
= 3[1 + \mu(\tau)]\left(\frac{2\pi D_+^2(\tau)}{L} \int \mathrm{d}q_1 \, q \, W_R(q)^2 \, P_{\mathrm{lin},0}(q)\right)^2
$$

$$
+ \frac{3\pi^2 D_+^4(\tau)}{L^2}\frac{\partial}{\partial \ln R} \int \mathrm{d}q_1 \mathrm{d}q_2 \, q_1 q_2 W_R(q_1)^2 \, W_R(q_2)^2 \, P_{\mathrm{lin},0}(q_1) P_{\mathrm{lin},0}(q_2)
$$

$$
= 3[1 + \mu(\tau)]\left(\frac{2\pi D_+^2(\tau)}{L} \int \mathrm{d}q_1 \, q \, W_R(q)^2 \, P_{\mathrm{lin},0}(q)\right)^2
$$

$$
+ \frac{3}{4}\frac{\partial}{\partial \ln R}\left(\frac{2\pi D_+^2(\tau)}{L^2} \int \mathrm{d}q_1 \, q \, W_R(q)^2 \, P_{\mathrm{lin},0}(q)\right)^2
$$

$$
= 3[1 + \mu(\tau)]\left(\langle\delta_{R,L}^2\rangle_{\mathrm{tree}}(\tau)\right)^2 + \frac{3}{4}\frac{\partial}{\partial \ln R}\left(\langle\delta_{R,L}^2\rangle_{\mathrm{tree}}(\tau)\right)^2 \ . \tag{B.35}
$$

Especially we have

$$
S_3 \equiv \frac{\langle\delta_{R,L}^3\rangle_{\mathrm{tree}}(\tau)}{\langle\delta_{R,L}^2\rangle_{\mathrm{tree}}(\tau)^2} = 3[1 + \mu(\tau)] + \frac{3}{2}\frac{\partial \ln\langle\delta_{R,L}^2\rangle_{\mathrm{tree}}(\tau)}{\partial \ln R} \ . \tag{B.36}
$$

For an Einstein-de Sitter universe and a power law power spectrum $P(k) \sim k^n$ this gives $S_3 = 36/7 - 3/2 \, (n + 2)$. The leading order prediction for S_3 is surprisingly good, even in the mildly non-linear regime [see 12, and references therein]. Hence in order to predict the non-linear skewness, we simply employ the approximation

$$
\langle\delta_{R,L}^3\rangle_{\mathrm{non-lin.}}(\tau) \approx S_3 \, \langle\delta_{R,L}^2\rangle_{\mathrm{non-lin.}}(\tau)^2 \ , \tag{B.37}
$$

where we compute the non-linear variance with the use of halofit as detailed in Takahashi et al. [73] which is a revised version of Smith et al. [68].

5. The moment $\langle \delta_{R_A,L}^2 \, \delta_{R_B,L} \rangle_{\mathrm{tree}}$

For predicting the density split lensing signal we are also interested in the moment $\langle \delta_{R_A,L}^2 \, \delta_{R_B,L} \rangle_{\mathrm{tree}}$, where R_A and R_B are two different Radii. The above derivations can be generalized to give

$$
\langle \delta_{R_A,L}^2 \, \delta_{R_B,L} \rangle_{\mathrm{tree}}(\tau) = \mathrm{Var}(R_A)\, \mathrm{Cov}(R_A, R_B) \left\{ 2[1 + \mu(\tau)] + \frac{1}{2}\frac{\partial \ln \mathrm{Var}(R_A)}{\partial \ln R_A} + \frac{\partial \ln \mathrm{Cov}(R_A, R_B)}{\partial \ln R_A} \right\}
$$

$$
+ \mathrm{Cov}(R_A, R_B)^2 \left\{ [1 + \mu(\tau)] + \frac{\partial \ln \mathrm{Cov}(R_A, R_B)}{\partial \ln R_B} \right\} \tag{B.38}
$$

where we defined

$$\text{Var}(R_A) = \langle \delta^2_{R_A,L} \rangle_{\text{tree}} \, , \quad \text{Cov}(R_A, R_B) = \langle \delta_{R_A,L} \, \delta_{R_B,L} \rangle_{\text{tree}} \, . \tag{B.39}$$

To correct this expression for the non-linear evolution of the power spectrum, we compute $\text{Var}(R_A)$ and $\text{Cov}(R_A, R_B)$ with our halofit power spectrum whenever they appear outside of the logarithmic derivatives. This is a generalization of the rescaling of $\text{Var}(R_A)$ by means of S_3.

For $R_B \gg R_A$ this rescaling is dominated by first term on the right hand side of equation B.38. For $R_B \approx R_A$ it reduces to equation B.37. As a consequence, using the procedure described around IV.39 to correct for the non-linear power spectrum evolution yields a predciton for the density split lensing signal that is almost identical to the procedure described here. Also, it can be considered accurate to the extend that equationB.37 is accurate. We nevertheless rescale the 3rd order moments in the more elaborate way described here.

6. The moment $\langle \delta^n_{R_A,L} \, \delta_{R_B,L} \rangle_{\text{tree}}$

Using a diagrammatic representation of perturbation theory (see e.g. [12]) one can see that the tree-level result for the moment $\langle \delta^n_{R_A,L} \, \delta_{R_B,L} \rangle_c$ will consist of terms that scale as

$$\sim \text{Cov}(R_A, R_B)^k \, \text{Var}(R_A)^{n-k} \, , \quad 1 \le k \le n \, . \tag{B.40}$$

For $R_B \approx R_A$ each of these scalings reduces to $\sim \text{Var}(R_A)^n$ (cf. B.37 and the definition of S_{n+1} in IV.16). On the other hand, for $R_B \gg R_A$ the terms scaling as $\sim \text{Cov}(R_A, R_B) \, \text{Var}(R_A)^{n-1}$ are the dominant contributions (cf. the last section for the case $n = 2$). This is why we use IV.39 when rescaling moments with $n > 2$ in $G_{\text{cyl}.,\theta_T w, \theta w, L}(q_g(w) L y, w)$ (see also IV.26).

Appendix C: Comparison with Millennium simulation

In Figure 9, we compare our model for the PDF of projected density contrast to another set of N-body simulations, the Millennium Run [MR 70]. The MR has a smaller simulation volume of only $(500h^{-1}\,\text{Mpc})^3$ co-moving, but a force resolution of $5h^{-1}$ kpc that is 4-10 times higher than that of the Buzzard simulations. The fiducial model and the log-normal model describe the distribution of $\delta_{m,T}$ measured from the MR well considering the large statistical uncertainty on $p(\delta_{m,T})$ due to the limited simulated sky area.

Appendix D: Galaxy stochasticity

Consider the field of galaxy density contrast $\delta_{g,T}$ and the field of matter density contrast $\delta_{m,T}$, where both fields are assumed to be smoothed over a fix circular aperture. The number of galaxies found inside such an aperture is assumed to be a Poissonian random variable with first and second moments for a given value of $\delta_{g,T}$ are given by

$$\langle \hat{N} | \delta_{g,T} \rangle = \bar{N}(1 + \delta_{g,T}) \tag{D.1}$$

and

$$\langle \hat{N}^2 | \delta_{g,T} \rangle = \bar{N}(1 + \delta_{g,T}) + \bar{N}^2 (1 + \delta_{g,T})^2 \, . \tag{D.2}$$

Let Var_m be the variance of $\delta_{m,T}$ and $\text{Var}_g = b^2 \text{Var}_m$ the variance of $\delta_{g,T}$, where b is the galaxy bias. Then the galaxy stochasticity r is defined by $\text{Cov}_{mg} = rb\text{Var}_m$, i.e. it is the correlation coefficient of $\delta_{g,T}$ and $\delta_{m,T}$.

We will now assume both $\delta_{g,T}$ and $\delta_{m,T}$ to be joint log-normal random variables, i.e.

$$\delta_{m,T} = [e^{n_m} - 1]\,\delta_{m,0}$$
$$\delta_{g,T} = [e^{n_g} - 1]\,\delta_{g,0} \, , \tag{D.3}$$

where n_m and n_g have a joint Gaussian distribution and $\delta_{g,0} = b\delta_{m,0}$. The variances of n_m and n_g are given by

$$\sigma^2_m = \ln\left\{ 1 + \frac{\text{Var}_m}{\delta^2_{m,0}} \right\}$$

$$\sigma^2_g = \ln\left\{ 1 + \frac{\text{Var}_g}{\delta^2_{g,0}} \right\}$$

$$= \sigma^2_m \tag{D.4}$$

and their covariance is given by

$$\xi_{mg} = \ln\left\{ 1 + \frac{\text{Cov}_{mg}}{\delta_{m,0}\delta_{g,0}} \right\}$$

$$= \ln\left\{ 1 + r\frac{\text{Var}_m}{\delta^2_{m,0}} \right\} \, . \tag{D.5}$$

Let us denote the correlation coefficient of the Gaussian field by

$$\rho = \frac{\xi_{mg}}{\sigma^2_m}$$

$$= \frac{\ln\left\{ 1 + r\frac{\text{Var}_m}{\delta^2_{m,0}} \right\}}{\ln\left\{ 1 + \frac{\text{Var}_m}{\delta^2_{m,0}} \right\}} \, . \tag{D.6}$$

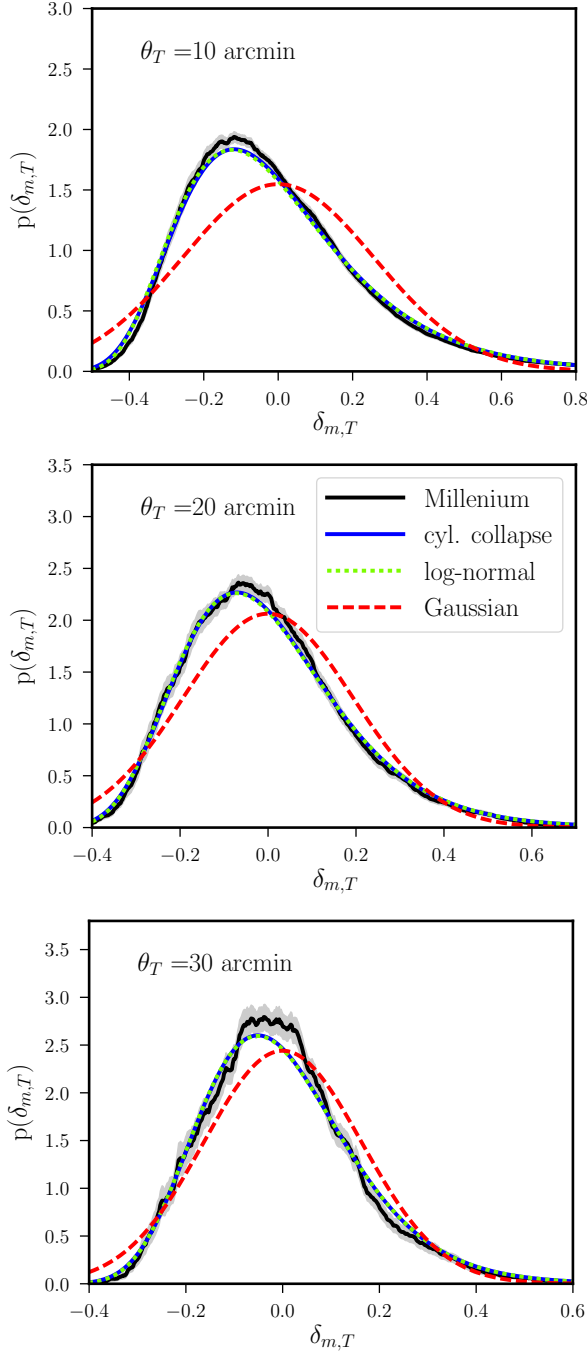Note that ρ will depend on scale even of b and r do not.

FIG. 9. The PDF of projected density contrast $\delta_{m,T}$ in the Millennium Run (MR) compared to our model. In each plot, the black line shows a histogram of $\delta_{m,T}$ measured from 64 patches of $4 \times 4 \deg^2$ made from the MR by projecting the 3D density contrast with a constant selection function q_l between $0.19 \lesssim z \lesssim 0.43$, i.e. with a constant co-moving density between those redshifts. The blue lines display the PDF predicted by our PT-motivated log-normal model, and the red lines show a Gaussian PDF with the same variance. The grey band is using the subsample covariance to estimate the error on the mean of all patches [33].

Now we want to compute the conditional moments $\langle \delta_{g,T}|\delta_{m,T}\rangle$ and $\langle \delta_{g,T}^2|\delta_{m,T}\rangle$. First,

$$\langle e^{n_g}|n_m\rangle = e^{\langle n_g|n_m\rangle + \sigma_g^2(1-\rho^2)/2}$$
$$= e^{\rho(n_m+\sigma_m^2/2)-\sigma_g^2\rho^2/2}$$
$$= e^{\sigma_g^2(\rho-\rho^2)/2}e^{\rho n_m} . \tag{D.7}$$

Second,

$$\mathrm{Var}\left(e^{n_g}|n_m\right) = \left(e^{\sigma_g^2(1-\rho^2)} - 1\right)e^{2\langle n_g|n_m\rangle + \sigma_g^2(1-\rho^2)}$$
$$= \left(e^{\sigma_g^2(1-\rho^2)} - 1\right)e^{\sigma_g^2(\rho-\rho^2)}e^{2\rho n_m} . \tag{D.8}$$

Now what is $\mathrm{Var}(\hat{N}|\delta_{m,T})$?

$$\langle \hat{N}^2|\delta_{g,T}\rangle = \langle \hat{N}^2|n_m\rangle$$
$$= \int d\delta_{g,T}\, p(\delta_{g,T}|n_m) \times$$
$$\times \left(\bar{N}[1+\delta_{g,T}] + \bar{N}^2[1+\delta_{g,T}]^2\right)$$
$$= \bar{N} + \bar{N}^2 + \int d\delta_{g,T}\, p(\delta_{g,T}|n_m) \times$$
$$\times \left(\delta_{g,T}[\bar{N} + 2\bar{N}^2] + \bar{N}^2\delta_{g,T}^2\right)$$
$$= \bar{N} + \bar{N}^2 + [\bar{N} + 2\bar{N}^2]\langle\delta_{g,T}|n_m\rangle +$$
$$+ \bar{N}^2\left(\mathrm{Var}\left(\delta_{g,T}|n_m\right) + \langle\delta_{g,T}|n_m\rangle^2\right)$$

$$\Rightarrow \mathrm{Var}(\hat{N}|\delta_{m,T}) = \bar{N}(1 + \langle\delta_{g,T}|n_m\rangle) + \bar{N}^2\mathrm{Var}\left(\delta_{g,T}|n_m\right)$$
$$= \bar{N}(1 + \langle\delta_{g,T}|n_m\rangle) + \bar{N}^2\delta_{g,0}^2\mathrm{Var}\left(e^{n_g}|n_m\right)$$
$$\tag{D.9}$$

The probability $P(N_A|\delta_{m,T})$ can be computed in a similar way, by numerically evaluating

$$P(N_A|\delta_{m,T}) = \int d\delta_{g,T}\, p(\delta_{g,T}|\delta_{m,T})\, P(N_A|\delta_{g,T}) , \tag{D.10}$$

where $p(\delta_{g,T}|\delta_{m,T})$ can be computed from basic relations for joint log-normal random variables.