Cosmology with second- and third-order shear statistics for the Dark Energy Survey: Methods and simulated analysis

R. C. H. Gomes,[1] S. Sugiyama,[1] B. Jain,[1] M. Jarvis,[1] D. Anbajagane,[2] M. Gatti,[3] D. Gebauer,[4,5] Z. Gong,[4,5] A. Halder,[6,7,5,8] G. A. Marques,[9,3] S. Pandey,[10] J. L. Marshall,[11] S. Allam,[12] O. Alves,[13] F. Andrade-Oliveira,[14] D. Bacon,[15] J. Blazek,[16] S. Bocquet,[17] D. Brooks,[18] A. Carnero Rosell,[19,20,21] J. Carretero,[22] L. N. da Costa,[20] P. Doel,[18] C. Doux,[1,23] S. Everett,[24] B. Flaugher,[12] J. Frieman,[2,12,3] J. García-Bellido,[25] E. Gaztanaga,[26,15,27] D. Gruen,[17] R. A. Gruendl,[28,29] G. Gutierrez,[12] K. Herner,[12] S. R. Hinton,[30] D. L. Hollowood,[31] K. Honscheid,[32,33] D. Huterer,[13] D. J. James,[34] N. Jeffrey,[18] J. Mena-Fernández,[35] R. Miquel,[36,22] J. Muir,[37,38] R. L. C. Ogando,[39] M. E. S. Pereira,[40] A. Pieres,[20,39] A. A. Plazas Malagón,[41,42] S. Samuroff,[16,22] E. Sanchez,[43] D. Sanchez Cid,[43,44] B. Santiago,[45,20] I. Sevilla-Noarbe,[43] M. Smith,[46] E. Suchyta,[47] M. E. C. Swanson,[28] G. Tarle,[13] C. To,[2] V. Vikram,[48] N. Weaverdyck,[49,50] and J. Weller[51,52]

(DES Collaboration)

[1]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*
[2]*Department of Astronomy and Astrophysics, University of Chicago, Chicago, Illinois 60637, USA*
[3]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, Illinois 60637, USA*
[4]*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstraße 1, 81679 München, Germany*
[5]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstraße 1, 85748 Garching, Germany*
[6]*Institute of Astronomy and Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, United Kingdom*
[7]*Jesus College, Jesus Lane, Cambridge, CB5 8BL, United Kingdom*
[8]*University Observatory, Ludwig-Maximilians-University of Munich, Scheinerstraße 1, 81679 Munich, Germany*
[9]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[10]*Columbia Astrophysics Laboratory, Columbia University, 550 West 120th Street, New York, New York 10027, USA*
[11]*George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, Texas 77843, USA*
[12]*Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, Illinois 60510, USA*
[13]*Department of Physics, University of Michigan, Ann Arbor, Michigan 48109, USA*
[14]*Physik-Institut, University of Zürich, Winterthurerstraße 190, CH-8057 Zürich, Switzerland*
[15]*Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, United Kingdom*
[16]*Department of Physics, Northeastern University, Boston, Massachusetts 02115, USA*
[17]*University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstraße 1, 81679 Munich, Germany*
[18]*Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, United Kingdom*
[19]*Instituto de Astrofisica de Canarias, E-38205 La Laguna, Tenerife, Spain*
[20]*Laboratório Interinstitucional de e-Astronomia-LIneA, Avenida Pastor Martin Luther King Jr., 126 Del Castilho, Nova América Offices, Torre 3000/sala 817 CEP: 20765-000, Brazil*
[21]*Universidad de La Laguna, Departamento Astrofísica, E-38206 La Laguna, Tenerife, Spain*
[22]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain*
[23]*Université Grenoble Alpes, CNRS, LPSC-IN2P3, 38000 Grenoble, France*
[24]*California Institute of Technology, 1200 East California Boulevard, MC 249-17, Pasadena, California 91125, USA*
[25]*Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, 28049 Madrid, Spain*
[26]*Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain*
[27]*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain*
[28]*Center for Astrophysical Surveys, National Center for Supercomputing Applications, 1205 West Clark Street, Urbana, Illinois 61801, USA*

[29]*Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 West Green Street, Urbana, Illinois 61801, USA*
[30]*School of Mathematics and Physics, University of Queensland, Brisbane, Queensland 4072, Australia*
[31]*Santa Cruz Institute for Particle Physics, Santa Cruz, California 95064, USA*
[32]*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, Ohio 43210, USA*
[33]*Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA*
[34]*Center for Astrophysics | Harvard and Smithsonian, 60 Garden Street, Cambridge, Massachusetts 02138, USA*
[35]*LPSC Grenoble-53, Avenue des Martyrs 38026 Grenoble, France*
[36]*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
[37]*Department of Physics, University of Cincinnati, Cincinnati, Ohio 45221, USA*
[38]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada*
[39]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[40]*Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany*
[41]*Kavli Institute for Particle Astrophysics and Cosmology, P. O. Box 2450, Stanford University, Stanford, California 94305, USA*
[42]*SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA*
[43]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain*
[44]*Physik-Institut, University of Zürich, Winterthurerstraße 190, CH-8057 Zürich, Switzerland*
[45]*Instituto de Física, UFRGS, Caixa Postal 15051, Porto Alegre, RS-91501-970, Brazil*
[46]*Physics Department, Lancaster University, Lancaster, LA1 4YB, United Kingdom*
[47]*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831*
[48]*Department of Physics, Central University of Kerala, 93VR+RWF, CUK Road, Kerala 671316, Índia*
[49]*Department of Astronomy, University of California, Berkeley, 501 Campbell Hall, Berkeley, California 94720, USA*
[50]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA*
[51]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstraße, 85748 Garching, Germany*
[52]*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstraße 1, 81679 München, Germany*

(Received 29 August 2025; accepted 11 November 2025; published 4 December 2025)

We present a new pipeline designed for the robust inference of cosmological parameters using both second- and third-order shear statistics. We build a theoretical model for rapid evaluation of three-point correlations using our FASTNC code and integrate it into the CosmoSIS framework. We measure the two-point functions $\xi_\pm$ and the full configuration-dependent three-point shear correlation functions across all auto- and cross-redshift bins. We compress the three-point functions into the mass aperture statistic $\langle \mathcal{M}_{\rm ap}^3 \rangle$ for a set of 796 simulated shear maps designed to model the Dark Energy Survey Year 3 data. We estimate from it the full covariance matrix and model the effects of intrinsic alignments, shear calibration biases and photometric redshift uncertainties. We apply scale cuts to minimize the contamination from the baryonic signal as modeled through hydrodynamical simulations. We find a significant improvement of 83% on the figure of merit in the $\Omega_{\rm m}$-S_8 plane when we add the $\langle \mathcal{M}_{\rm ap}^3 \rangle$ data to $\xi_\pm$. We present our findings for all relevant cosmological and systematic uncertainty parameters and discuss the complementarity of third-order and second-order statistics.

DOI: 10.1103/3f1p-spfp

I. INTRODUCTION

The current state of observational cosmology is one of tensions and expectations [1,2]. While discrepancies continue to be identified between standard cosmological predictions and observational data, the upcoming Stage IV surveys offer a promising hope for identifying where we need to refine or go beyond the Λ cold dark matter (ΛCDM) model and where we need to improve our understanding of observational systematics [3]. This context has been driving cosmologists towards the development of techniques to extract the most information out of the available data.

For weak lensing data, going beyond the two-point correlation function of the cosmic shear field allows us to probe the field's non-Gaussian features, which can both lead us to a tightening of the parameter space constraints

and offer improved characterization of systematics such as the intrinsic alignment of galaxies [4]. Cosmic shear two-point analyses have been performed with Dark Energy Survey Year 3 (DES-Y3) data [5,6], as well as with data from Kilo-Degree Survey (KiDS)-1000 [7] and Hyper Suprime-Cam Year 3 (HSC Y3) [8,9]. Many studies have been conducted using higher-order statistics of the cosmic shear field as well. An approach that uses moments of the weak lensing mass maps was conducted by Gatti *et al.* [10] with DES data. The use of third moments of the field (encoding non-Gaussian information) in addition to the second moments is shown to give an improvement of 15% in S_8 constraints and of 25% in Ω_m constraints relative to the use of second moments alone. A similar level of improvement was found by Gong *et al.* [11] on simulated data when adding the integrated shear three-point function to the shear two-point correlation functions $\xi_\pm$.

Additional state-of-the-art higher-order approaches are discussed and reviewed by Ajani [12], and include Minkowski functionals [13], weak lensing peak counts and minimum counts [14], scattering transforms [15], wavelet phase harmonics [16], homology statistics [17], and probability distribution function/cumulative distribution functions [18,19]. Finally, deep learning approaches at the field level have also been developed and applied to lensing mass maps [20–22].

Given that the full shear three-point correlation function has been measured in DES-Y3 data with high signal-to-noise [23], the question arises: How much additional information would this statistic provide to the two-point shear constraints from DES-Y3? Full shear three-point function measurements typically yield us data vectors with large dimensions, which are not optimal for covariance determination. However, a principal component analysis by Heydenreich *et al.* [24] shows that the information content of the full three-point function is similar to that of the mass aperture statistic ($\langle\mathcal{M}_{\rm ap}^3\rangle$). Thus, the latter can be interpreted as a physically motivated efficient compression of the former.

The usage of $\langle\mathcal{M}_{\rm ap}^3\rangle$ has been shown by Burger *et al.* [25] to lead to complementary constraints to those from two-point statistics. Their analysis with KiDS-1000 data provided a factor of two improvement in the joint Ω_m-S_8 constraints. Nonetheless, this was partially driven by the fact that their setup for the two-point analysis was less constraining than that of the fiducial KiDS-1000 cosmic shear analysis [7]. In light of this result, we can expect joint $\xi_\pm$ and $\langle\mathcal{M}_{\rm ap}^3\rangle$ constraints to yield significant improvement relative to two-point analyses in the context of other collaborations.

This paper is motivated by these recent developments and establishes a pipeline for the analysis of the third-order cosmic shear information on DES-Y3 data. We construct a theoretical model for the mass aperture statistic, estimate its covariance through simulations, and build a robust likelihood for cosmological analyses. These developments are complemented by a companion paper [26] that will describe the analysis with DES-Y3 data.

The paper is organized as follows: in Sec. II, we describe our theoretical model for the three-point correlation function and the mass aperture statistic, including the modeling of observational systematics and a description of our neural network emulator. In Sec. III, we present our data vector and covariance. In Sec. IV, we describe the parameter inference scheme and our analysis choices. Finally, in Sec. V, we show the parameter estimation results of our simulated analysis, validating therefore our pipeline for use with DES data. Our concluding remarks are made in Sec. VI.

II. THEORETICAL MODELING

A. Basics of weak lensing

The study of weak lensing allows us to probe the matter density field of the Universe in an unbiased way, because it is sensitive to the total matter density including both visible and dark matter. The convergence field on the sky coordinate $\boldsymbol{X}$, which characterizes the isotropic magnification of the background galaxy's shape, is obtained by the line-of-sight integration of the matter density contrast with the lensing efficiency [27]

$$\kappa(\boldsymbol{X}) = \frac{3\Omega_m H_0^2}{2c^2}\int_0^\infty \mathrm{d}\chi q_i(\chi)\frac{\delta_m(\chi\boldsymbol{X},\chi;z(\chi))}{a(\chi)} \tag{1}$$

$$q_i(\chi) = \int_\chi^\infty \mathrm{d}\chi' p_i(\chi')\frac{\chi'-\chi}{\chi'}, \tag{2}$$

where χ is the comoving distance, $z(\chi)$ is the redshift at a comoving distance χ from the observer, $a(\chi)$ is the scale factor, and $\delta_m(\boldsymbol{r};z)$ is the matter density contrast at the three-dimensional coordinate $\boldsymbol{r}$ and redshift z. The lensing efficiency $q_i(\chi)$ is defined for the ith source redshift bin, and characterizes the efficiency of contribution of the matter density field at comoving distance χ to the convergence field, and depends on the distribution of the source galaxies $p(\chi)$, which is normalized to unity by imposing $\int \mathrm{d}\chi p(\chi) = 1$.

The shear field of weak lensing in Cartesian frame is defined as $\gamma_c(\boldsymbol{X}) = \gamma_1(\boldsymbol{X}) + i\gamma_2(\boldsymbol{X})$, where γ_1 and γ_2 are the shears along the x-axis and an axis rotated 45 degrees from the x-axis, respectively. When the shear is projected onto another reference frame that is rotated by an angle ζ from the Cartesian frame, the shear field follows the spin-2 transformation

$$\begin{aligned}\gamma(\boldsymbol{X};\zeta) &\equiv \gamma_t(\boldsymbol{X};\zeta) + i\gamma_\times(\boldsymbol{X};\zeta)\\ &= -\gamma_c(\boldsymbol{X})e^{-2i\zeta}.\end{aligned} \tag{3}$$

where γ_t is the tangential component of each galaxy's shear signal, and $\gamma_\times$ its radial component.

We define the Fourier transformation of the shear field as

$$\gamma_c(\boldsymbol{X}) = \int \frac{d^2\boldsymbol{\ell}}{(2\pi)^2}\gamma_c(\boldsymbol{\ell})e^{-i\boldsymbol{\ell}\cdot\boldsymbol{X}}, \tag{4}$$

and the same for the convergence field. The shear and convergence fields can be related to each other in Fourier space as

$$\gamma_c(\boldsymbol{\ell}) = \kappa(\boldsymbol{\ell})e^{2i\beta}. \tag{5}$$

Here β is the polar angle of the Fourier mode $\boldsymbol{\ell}$.

B. Second- and third-order shear statistics

To capture the Gaussian information of the shear field, we traditionally rely on two-point statistics. In Fourier space, we write the power spectrum of the convergence field as

$$\langle\kappa(\boldsymbol{\ell}_1)\kappa(\boldsymbol{\ell}_2)\rangle = (2\pi)^2\delta^D(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2)P_\kappa(\ell_1). \tag{6}$$

The relation between $P_\kappa(\ell)$ and the matter power spectrum is given by integrating the latter over redshift with the lensing kernel. For redshift bins i and j, we make use of the Fourier space Limber approximation [28] and write

$$P_\kappa(\ell) = \frac{9\Omega_m^2H_0^4}{4c^4}\int_0^\infty d\chi\frac{q_i(\chi)q_j(\chi)}{a^2(\chi)}P_\delta\left(\frac{\ell}{\chi},z(\chi)\right), \tag{7}$$

Analogously, we can extract more information if we move beyond the Gaussian features of the field and consider its three-point statistics. In this way, we use the convergence bispectrum

$$\langle\kappa(\boldsymbol{\ell}_1)\kappa(\boldsymbol{\ell}_2)\kappa(\boldsymbol{\ell}_3)\rangle = (2\pi)^2\delta^D\left(\sum_{i=1}^3\boldsymbol{\ell}_i\right)B_\kappa(\ell_1,\ell_2,\ell_3). \tag{8}$$

The relation between the matter bispectrum and the convergence bispectrum can also be given under the Limber approximation [29]. We have

$$B_\kappa(\ell_1,\ell_2,\ell_3) = \frac{27\Omega_m^3H_0^6}{8c^6}\int_0^\infty d\chi\frac{q_i(\chi)q_j(\chi)q_k(\chi)}{a(\chi)^3\chi} \times B_\delta\left(\frac{\ell_1}{\chi},\frac{\ell_2}{\chi},\frac{\ell_3}{\chi},z(\chi)\right). \tag{9}$$

We theoretically model the matter power spectrum and the matter bispectrum in order to have a starting point to compute our n-point shear statistics. The nonlinear power spectrum is computed through the revised Halofit fitting formula, devised by the authors of Ref. [30]. This expression yields an accuracy of 5% for scales with $k \leq 1h$ Mpc^{-1} and 10% for scales between 1 and $10h$ Mpc^{-1} for all our redshifts of interest. Similarly, we model the matter bispectrum with the BiHalofit formula [31], which is shown to match the perturbation level calculation on *Planck* 2015 cosmology to an accuracy of 10% on scales with $k < 3h$ Mpc^{-1}. Our implementation of the BiHalofit model is described in Appendix A.

We can also write, in real space, the shear two-point and three-point correlation functions in terms of the power spectrum and bispectrum. The fact that the shear field is spin-2 gives us three distinct two-point functions by correlating $\langle\gamma_t\gamma_t\rangle$, $\langle\gamma_t\gamma_\times\rangle$, and $\langle\gamma_\times\gamma_\times\rangle$. The correlation between the tangential and radial components is zero for a universe with parity symmetry, leaving us with two independent functions [27]. We can combine the tangential and radial components and write

$$\xi_+(\theta) \equiv \langle\gamma_t\gamma_t\rangle + \langle\gamma_\times\gamma_\times\rangle, \tag{10}$$

$$\xi_-(\theta) \equiv \langle\gamma_t\gamma_t\rangle - \langle\gamma_\times\gamma_\times\rangle. \tag{11}$$

To express the relation between these functions and the convergence power spectrum, we decompose the convergence field into E and B modes. We have now two separate fields κ_E and κ_B given by

$$\nabla^2\kappa_E = \boldsymbol{\nabla}(\boldsymbol{\nabla}\kappa), \tag{12}$$

$$\nabla^2\kappa_B = \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\kappa). \tag{13}$$

The convergence power spectrum can be decomposed as well between $P_\kappa^E(\ell)$ and $P_\kappa^B(\ell)$ by replacing $\kappa(\boldsymbol{\ell})$ for its E and B components in Eq. (6). Following Kilbinger [27], this brings us to

$$\xi_+(\theta) = \int_0^\infty\frac{\ell d\ell}{2\pi}J_0(\ell\theta)[P_\kappa^E(\ell)+P_\kappa^B(\ell)], \tag{14}$$

$$\xi_-(\theta) = \int_0^\infty\frac{\ell d\ell}{2\pi}J_4(\ell\theta)[P_\kappa^E(\ell)-P_\kappa^B(\ell)], \tag{15}$$

where $J_{0/4}(x)$ is the 0th-/4th-order Bessel function of the first kind.

For three-point statistics, the traditional convention is to project the shear of a given point over the line that connects it to a chosen center of the triangle, which can be either the orthocenter or the centroid. From this approach, described in Schneider and Lombardi [32], we write eight components of the three-point correlation function (3PCF):

$$\gamma_{ijk} = \langle\gamma_i(\boldsymbol{X}_1)\gamma_j(\boldsymbol{X}_2)\gamma_k(\boldsymbol{X}_3)\rangle. \tag{16}$$

The so-called natural components of cosmic shear, which are invariant combinations of the above quantities, are defined by Schneider and Lombardi [32]. Its real and imaginary parts are

$$\begin{aligned}\mathbf{Re}(\Gamma^0) &= \gamma_{ttt} - \gamma_{t\times\times} - \gamma_{\times t\times} - \gamma_{\times\times t}\\ \mathbf{Im}(\Gamma^0) &= \gamma_{tt\times} + \gamma_{t\times t} + \gamma_{\times tt} - \gamma_{\times\times\times}\end{aligned} \tag{17}$$

$$\begin{aligned}\mathbf{Re}(\Gamma^1) &= \gamma_{ttt} - \gamma_{t\times\times} + \gamma_{\times t\times} + \gamma_{\times\times t}\\ \mathbf{Im}(\Gamma^1) &= \gamma_{tt\times} + \gamma_{t\times t} - \gamma_{\times tt} + \gamma_{\times\times\times}\end{aligned} \tag{18}$$

$$\begin{aligned}\mathbf{Re}(\Gamma^2) &= \gamma_{ttt} + \gamma_{t\times\times} - \gamma_{\times t\times} + \gamma_{\times\times t}\\ \mathbf{Im}(\Gamma^2) &= \gamma_{tt\times} - \gamma_{t\times t} + \gamma_{\times tt} + \gamma_{\times\times\times}\end{aligned} \tag{19}$$

$$\begin{aligned}\mathbf{Re}(\Gamma^3) &= \gamma_{ttt} + \gamma_{t\times\times} + \gamma_{\times t\times} - \gamma_{\times\times t}\\ \mathbf{Im}(\Gamma^3) &= -\gamma_{tt\times} + \gamma_{t\times t} + \gamma_{\times tt} + \gamma_{\times\times\times}.\end{aligned} \tag{20}$$

The functions Γ^0, Γ^1, Γ^2 and Γ^3 can be worked out as functions of the convergence bispectrum. This has been done by Schneider *et al.* [33], and is reproduced by Heydenreich *et al.* [24] for their three-point function integration code. The expressions involve integrations with J_2 and J_6 functions, which have a highly oscillatory behavior and thus require caution in direct numerical integration. We have

$$\begin{aligned}&\Gamma^0(x_1,x_2,x_3)\\ &= 2\pi\int_0^\infty \frac{\ell_1 d\ell_1}{(2\pi)^2}\int_0^\infty \frac{\ell_2 d\ell_2}{(2\pi)^2}\\ &\quad\times\int_0^{2\pi} d\varphi B_\kappa(\ell_1,\ell_2,\varphi)e^{2i\bar{\beta}}(e^{i(\phi_1-\phi_2-6\alpha_3)}J_6(A_3)\\ &\quad + e^{i(\phi_2-\phi_3-6\alpha_1)}J_6(A_1) + e^{i(\phi_3-\phi_1-6\alpha_2)}J_6(A_2))\end{aligned} \tag{21}$$

and

$$\begin{aligned}&\Gamma^1(x_1,x_2,x_3)\\ &= 2\pi\int_0^\infty \frac{\ell_1 d\ell_1}{(2\pi)^2}\int_0^\infty \frac{\ell_2 d\ell_2}{(2\pi)^2}\\ &\quad\times\int_0^{2\pi} d\varphi B_\kappa(\ell_1,\ell_2,\varphi)e^{2i\bar{\beta}}(e^{i(\phi_1-\phi_2+2\phi_3+2\bar{\beta}-2\varphi-2\alpha_3)}J_2(A_3)\\ &\quad + e^{i(\phi_3-\phi_2-2\bar{\beta}-2\alpha_1)}J_2(A_1) + e^{i(\phi_3-\phi_1-2\phi_2+2\bar{\beta}+2\varphi-2\alpha_2)}J_2(A_2)).\end{aligned} \tag{22}$$

Here, we express the convergence bispectrum as a function of ℓ_1, ℓ_2, and the opening angle between them φ, for which $\ell_3 = (\ell_1^2 + \ell^2 - 2\ell_1\ell_2\cos\varphi)^{1/2}$. We also define the quantity $\bar{\beta}$, which corresponds to the angle between ℓ_3 and the mean direction between ℓ_1 and ℓ_2. The values of Γ^2 and Γ^3 are obtained through cyclic permutation of the indices on the expression for Γ^1. The quantities A_3 and α_3 are defined by the pair

$$A_3\sin\alpha_3 = (\ell_1 x_2 - \ell_2 x_1)\sin\frac{\varphi+\phi_3}{2}, \tag{23}$$

$$A_3\cos\alpha_3 = (\ell_1 x_2 + \ell_2 x_1)\cos\frac{\varphi+\phi_3}{2}, \tag{24}$$

and their counterparts A_1 (A_2) and α_1 (α_2) are also given by cyclic permutation of indices.

The requirements for a parameter inference pipeline include being able to generate fast theoretical models to be used by a sampler. Though we can enhance speed with emulation, a faster code is still important to be able to generate a large enough sample for emulator training. The complexity of Eqs. (21) and (22) calls for a level of precision during integration that leads to an increased runtime. We propose in Sugiyama *et al.* [34] a fast algorithm to go from the matter bispectrum to the natural components of the shear three-point correlation function by performing a multipole expansion of the bispectrum. In order to do so, we move away from the conventional projection of the shear, and write our correlation functions relative to the so-called ×-projection, defined by Porth *et al.* [35] as

$$\zeta_1^\times = \frac{\varphi_1+\varphi_2}{2}, \tag{25}$$

$$\zeta_2^\times = \varphi_1, \tag{26}$$

$$\zeta_3^\times = \varphi_2, \tag{27}$$

where φ_1 and φ_2 correspond to the angles between each of two triangle sides and the x-axis.

Using this convention, we write the expression for the first component of the shear 3PCF as follows:

$$\begin{aligned}\Gamma_0^\times(\theta_1,\theta_2,\phi) = -\int \frac{\mathrm{d}^2\boldsymbol{\ell_1}}{(2\pi)^2}\frac{\mathrm{d}^2\boldsymbol{\ell_2}}{(2\pi)^2}e^{-i\boldsymbol{\ell_1}\cdot\boldsymbol{\theta}_1 - i\boldsymbol{\ell_2}\boldsymbol{\theta}_2}\\ \times B_\kappa(\ell_1,\ell_2,\varphi)e^{2i\sum_i\beta_i}e^{-3i(\varphi_1+\varphi_2)}.\end{aligned} \tag{28}$$

By replacing the bispectrum in Eq. (28) by its expansion in Legendre polynomials, we arrive at a way to compute $\Gamma_0^\times$ as a sum of a series of multipole components. The same procedure is done for $\Gamma_1^\times$, $\Gamma_2^\times$, and $\Gamma_3^\times$ and detailed in Sugiyama *et al.* [34]. We show that this method yields a significant speedup relative to brute-force integration. Our FASTNC code is used in our pipeline to model the shear three-point functions.

C. Mass aperture statistics

The mass aperture is a single measurement of the convergence signal within a circular patch. While the average convergence is, by construction, zero, we can get second-order and third-order shear information by analyzing the moments of the mass aperture. Particularly, the skewness of the mass aperture, given by $\langle\mathcal{M}_{\mathrm{ap}}^3\rangle$, is an efficient way to compress the third-order information that would rather be

scattered along multiple triangle configurations on the three-point correlation function.

The mass aperture statistic can be computed either as an integral of the bispectrum or directly from the natural components of the 3PCF. We opt for the second approach, which allows us to have a theory model consistent with TreeCorr [36] measurements of the mass aperture. In order to guarantee that our model will reproduce the measured statistic, we consistently apply scale cuts and bin averaging on our modeled 3PCF, using the same choices as the ones done for the measurements (detailed in Sec. III).

With the binned values of the Γ^i functions, the mass aperture is found through matrix multiplication. In integral form, the expression was determined by Jarvis *et al.* [36], with the T_0 and T_1 functions defined in their Eqs. (51) and (52), and the quantities s and $\mathbf{t}'$ given in their Eqs. (46)–(48) as combinations of the vectors $\mathbf{q}_i$ that connect each vertex of a triangle configuration to the centroid. Permuting the $\mathbf{q}_i$ indices in the T_1 expression gives us the functions T_2 and T_3. The final expression for $\langle \mathcal{M}_{\rm ap}^3 \rangle$ at an aperture radius θ can then be simplified as

$$\langle \mathcal{M}_{\rm ap}^3 \rangle(\theta) = \frac{3}{2}\mathrm{Re}\int \frac{sds}{s^2}\int_{s<t'<|t'-s|}\frac{d^2\boldsymbol{t}'}{2\pi\theta^2} \times \sum_{i=0,1,2,3}\Gamma^i(s,\boldsymbol{t}')T_i\left(\frac{s}{\theta},\frac{\boldsymbol{t}'}{\theta}\right). \tag{29}$$

In addition to computing the $\langle \mathcal{M}_{\rm ap}^3 \rangle(\theta)$ with the DES-Y3 redshift source redshift bin integration kernels, our pipeline also computes the redshift-dependent $\langle \mathcal{M}_{\rm ap}^3 \rangle(\theta, z)$, which can be transformed into the full mass aperture with a line-of-sight integral. This gives us the advantage of making all the three-point and mass aperture integration independent of the source galaxy redshift distribution. We take advantage of this feature to build our emulator, described in Sec. II F.

The $\langle \mathcal{M}_{\rm ap}^3 \rangle$, as a local measurement, is generally insensitive to partial-sky coverage. On survey data, which commonly have complex masking, the presence of holes and edges can bias a direct measurement of the mass aperture [23]. However, the approach of estimating $\langle \mathcal{M}_{\rm ap}^3 \rangle$ from the three-point correlation function integration removes the mask dependency [36].

It has been shown by Heydenreich *et al.* [24] that using $\langle \mathcal{M}_{\rm ap}^3 \rangle$ for cosmological parameter estimation yields results comparable to those obtained by using the most relevant principal components of the full three-point function, and that these first principal components already saturate the available information, making it unnecessary to continue adding the remaining ones.

D. Modeling observational systematics

An important step towards a robust estimate of the cosmological parameters is to model the effect of observational systematics. If left unmodeled, systematic effects on the data can bring significant biases to the final parameters. Although most of the observational systematics are calibrated at the catalog level in DES-Y3 data, we introduce nuisance parameters in our theoretical expressions in order to marginalize over the potential residual error from these systematics. Our formalism follows the one used by the two-point DES-Y3 analysis [5,37,38].

We first include in our model the photometric redshift error, which should account for eventual shifts in the computed redshift distribution of the survey data. To model the residual error in the estimated photometric redshift, we allow the mean redshift of the source distribution to vary by a shift parameter Δz_i for ith redshift bin, as follows:

$$p_i(z) \to p_i(z - \Delta z_i). \tag{30}$$

In terms of shear biases, the DES-Y3 shape catalog uses a self-calibration method to calibrate most of the bias based on the data itself. This is expected to eliminate additive biases on the shear values. However, the blending effect cannot be fully removed, which leaves percent-level systematic uncertainties in the multiplicative bias of the shear. To account for this residual bias in theory, we include one shear multiplicative bias parameter for each source redshift bin, and modulate the 2PCF and 3PCF by

$$\xi_{\pm}^{ij} \to (1+m_i)(1+m_j)\xi_{\pm}^{ij}, \tag{31}$$

$$\Gamma_{\mu}^{ijk} \to (1+m_i)(1+m_j)(1+m_k)\Gamma_{\mu}^{ijk}. \tag{32}$$

Here m_i is the multiplicative bias for ith source redshift bin, and is marginalized over in the parameter inference process with a proper prior following the Bayesian framework discussed in Sec. IV A.

Modeling the point spread function (PSF) in the observed image is another important aspect of shear estimation. Inaccuracies in the PSF model over the observed sky are known to introduce two additive bias terms in shear estimation: residual and leakage. These biases give rise to additive PSF correlation terms in the *N*PCF. Gatti *et al.* [39] and Secco *et al.* [23] measured the size of the PSF correction terms for the 2PCF and 3PCF, respectively, and concluded that they are subdominant. Therefore, we do not include the PSF terms in our model.

E. Intrinsic alignment

The measured signal from galaxy ellipticity catalogs comprises both the weak lensing signal and the intrinsic alignment (IA) signal. A careful consideration of the latter is important in order to get unbiased constraints from a cosmic shear analysis. Additionally, a consistent modeling of IA in two-point and three-point functions can be responsible, via self-calibration, for a significant gain in information on the cosmological parameters [4]. The nonlinear alignment

(NLA) model is based on the assumption that the intrinsic galaxy shapes are linearly aligned with the tidal field [40]. The NLA model can be easily included in the two-point and three-point function computations by replacing the lensing efficiency as follows [10,41]:

$$q_i(\chi) \to q_i(\chi) + f_{\rm IA}(z(\chi)) p_i(\chi) \frac{{\rm d}z}{{\rm d}\chi}. \tag{33}$$

Here $f_{\rm IA}(z)$ is the redshift dependence of the relation between the intrinsic galaxy alignment and the tidal field,

$$f_{\rm IA}(z) = -A_{\rm IA}\left(\frac{1+z}{1+z_0}\right)^{\alpha_{\rm IA}} \frac{c_1\rho_{\rm crit}\Omega_{\rm m,0}}{D(z)}, \tag{34}$$

where $z_0 = 0.62$ is the pivot redshift, $c_1\rho_{\rm crit} = 0.0134$ is a conventional constant value, $D(z)$ is the growth function normalized to unity at $z = 0$, and $A_{\rm IA}$ and $\alpha_{\rm IA}$ are the NLA model parameters.

In the context of two-point analyses, intrinsic alignment models with higher complexity have also been explored. Blazek *et al.* [42] introduce the tidal alignment and tidal torquing (TATT) model, motivated by the need to account for the behavior of spiral galaxies which do not necessarily follow the NLA assumption. Instead of performing a population split, they take a perturbative approach, introducing two new free parameters to probe the amplitude and redshift dependence of the tidal torquing alignment component.

In terms of statistical model selection, Secco *et al.* [5] find a preference for NLA over TATT in modeling DES-Y3 data. Analyses from other surveys have also found no definite need to move from NLA towards models with more parameters. Dalal *et al.* [8] use HSC Y3 data and determine the shift in S_8 when changing between IA models to be insignificant. In a joint DES + KiDS analysis, Abbott *et al.* [43] find a larger shift in S_8 but it is partially attributed to the fact that TATT allows their pipeline to explore lower S_8 values in a regime with high tidal torquing alignment, skewing the posterior mean. This is not, therefore, a preference for any particular model.

For higher-order statistics, it is important to maintain consistent IA modeling between second- and third-order correlations. Pyne *et al.* [44] determine that such consistency is possible within the NLA framework, which allows us to use the same A_1 and α_1 parameters for our two-point and three-point functions. Given these considerations, we opt for model simplicity and implement NLA in our pipeline, through Eqs. (33) and (34).

F. Emulator for redshift-dependent mass aperture

We build a neural network emulator for the mass aperture statistic, in order to speed up our cosmological inference pipeline. We use the functions provided by the CosmoPower framework [45] to create and train the network. Our model inputs are the five cosmological parameters $\Omega_{\rm m}$, S_8, h_0, Ω_b, and n_s. The sum of the neutrino masses m_ν is not included as an emulator input because its dependence is not captured by the BiHalofit modeling, which is calibrated on simulations with fixed $m_\nu = 0.06$ eV. Nonetheless, we do not expect this to yield any biases due to current lensing data lacking sensitivity to constrain neutrino masses.

TABLE I. Range of values used to sample our parameter space for emulator training and testing. Values were chosen in order to avoid at most the outlier cosmologies that can break down the $P(k)$ or the $B(k_1, k_2, k_3)$ computation. Our main emulator does not sample the dark energy equation of state w_0, leaving it fixed. For our wCDM analysis, we build a second emulator and sample it over w_0.

Parameter	Min	Max
$\Omega_{\rm m}$	0.12	0.5
S_8	0.65	0.9
h_0	0.55	0.91
Ω_b	0.03	0.07
n_s	0.87	1.07
w_0	−2	−0.5

For each set of cosmological parameters, the model is trained to compute the redshift-dependent (matter field) $\langle\mathcal{M}_{\rm ap}^3\rangle(\theta_i, z)$ for $\theta_i = 7', 14', 25', 40'$ and for a predefined set of $36z$ values, giving us an output of length 144 (see Appendix B). We verify that using this set of z values for line-of-sight integration gives us a precision of better than 2% on $\langle\mathcal{M}_{\rm ap}^3\rangle(\theta)$ relative to a fiducial theoretical model computed with 30 log bins from $z = 10^{-4}$ to $z = 0.1$ and 100 linear bins from $z = 0.1$ to $z = 3$. Generating a z-dependent output for $\langle\mathcal{M}_{\rm ap}^3\rangle$ allows us to have an emulator that is independent of the source redshift distributions and of the intrinsic alignment kernel.

To train and test our model, we generate a set of 1500 samples, drawn as a Sobol sequence on the five-dimensional parameter space. The range of values used for each parameter is found in Table I. We discard the samples on which the $\langle\mathcal{M}_{\rm ap}^3\rangle$ computation fails (due to limitations of the Boltzmann solver and of BiHalofit modeling on outlier cosmologies). We divide our remaining samples into one subset for training and validation (with a split of 75%/25% and total size of 1300), and a second set of size 171 for testing purposes. We investigate the set of 144×171 outputs for the testing set and find that the network error is below 0.29% for 99% of the samples, and below 1.02% for 100% the samples. The percent error on the testing samples is shown in Fig. 1.

In order to use the network predictions on our cosmological inference pipeline, we write the mass aperture $\langle\mathcal{M}_{\rm ap}^3\rangle(\theta)_{ijk}$ for each redshift bin combination (i, j, k), as follows:

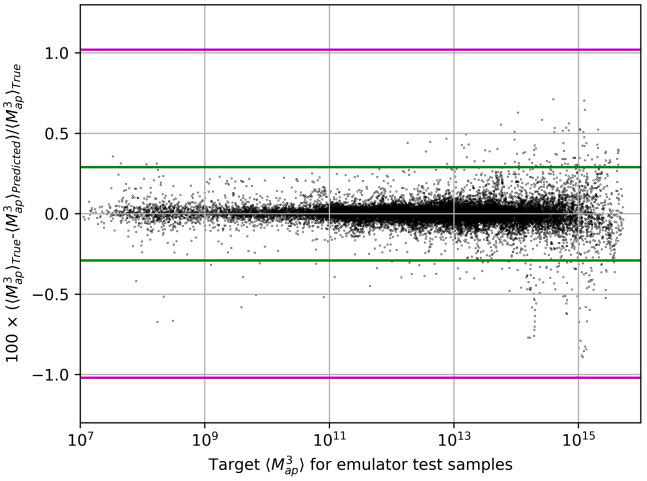

FIG. 1. Dispersion of mass aperture emulator predictions. On the y-axis we show the percentage error of the network predictions relative to their corresponding target values. The green lines mark the interval between −0.29% and 0.29%, and the magenta between −1.02% and 1.02%. A total of 99% of the samples lie between the green lines, and 100% of them between the magenta lines, showing that the emulator scatter is a negligible source of error for our $\langle \mathcal{M}_{\rm ap}^3 \rangle$ model.

$$\langle \mathcal{M}_{\rm ap}^3 \rangle(\theta)_{ijk} = \int \frac{d\chi}{\chi} \frac{q_i(\chi) q_j(\chi) q_k(\chi)}{a(\chi)^3} \langle \mathcal{M}_{\rm ap}^3 \rangle(\theta, z(\chi)). \tag{35}$$

Each integration kernel $q(\chi)$ includes both the lensing kernel and the intrinsic alignment NLA kernel for the selected redshift bin, as described in Sec. II E.

During integration, the errors from the network output are suppressed. The discrepancy between target and predicted $\langle \mathcal{M}_{\rm ap}^3 \rangle$ for all redshift bin combinations is of the order of 0.01%, being an insignificant contribution to our total model error budget.

We also developed one additional emulator model to allow for wCDM cosmologies. We generated 2000 Sobol samples with the same ranges for $\Omega_{\rm m}$, S_8, h_0, Ω_b, and n_s as in our fiducial model. We varied the dark energy equation of state w_0 parameter between −2 and −0.5. On this new set of samples, we selected 1700 for training and validation (with a 75%/25% split), and left the remaining 262 successful ones for testing. By analyzing the testing samples, we certify that the error is also not significant, remaining below 0.7% for 99% of the samples.

By introducing the emulator, the computation speed has significantly improved, reducing the computational time from approximately 40 seconds using FASTNC to just 0.03 seconds with the emulator—an acceleration by a factor of $\mathcal{O}(10^3)$. This remarkable speedup makes it feasible for posterior Monte Carlo sampling, enabling efficient exploration of high-dimensional parameter spaces while substantially reducing computational costs.

III. COVARIANCE ESTIMATION

A joint analysis of second- and third-order shear statistics requires modeling the covariance matrix of the joint data vector. In this paper, we use simulations to estimate the data covariance for $\xi_\pm$ and $\langle \mathcal{M}_{\rm ap}^3 \rangle$. The CosmoGridV1 suite [46] is a set of light cone simulations generated with the PkdGrav3 code [47]. It includes runs at a grid of varying cosmologies, as well as runs at a fiducial cosmology. For our covariance estimation, we use their fiducial runs to generate 199 full-sky convergence maps at the DES-Y3 source redshift distributions, which are then each transformed into cosmic shear maps and cut out into four DES-Y3 footprints, yielding us a total of 796 shear maps.

To add shape noise to our mock data, we randomly rotate the ellipticities of the Y3 shape catalog [39] at the positions of its galaxies. Next, we add to the randomly rotated e_1 and e_2 values the mock CosmoGridV1 shear values for each galaxy position. Finally, we create an `NSIDE` = 1024 HEALPix map [48] and take, for each pixel, the average over all the final shear values.

For each realization, we perform a measurement of the 2PCF, $\xi_\pm(\theta)_{ij}$, using TreeCorr on 20 log-spaced θ bins ranging from 2.5 to 250 arcminutes, which is the same angular binning as in the DES-Y3 key papers [49]. Here the subscript indicates the pair of redshift bins that are being correlated, and runs over the 10 possible combinations that arise from the four tomographic redshift bin setup.

Similarly, we perform a measurement of third order shear statistics for each realization. We first perform our 3PCF measurements with TreeCorr, using the Multipole binning scheme. In this scheme, each triangle configuration is described by two side lengths and a set of multipoles of the opening angle. We take as the maximum multipole $\max_n = 100$. For the side binning, we take 20 logarithmically spaced bins between $\theta_{\rm min} = 0.5'$ and $\theta_{\rm max} = 80'$. Next, we convert these measurements into the side-angle-side binning scheme, using 63 linearly spaced bins for the opening angle, which go from 0.016 to 3.126 radians. Finally, these measurements are transformed, using TreeCorr, into measurements of the skewness of the mass aperture statistic $\langle \mathcal{M}_{\rm ap}^3 \rangle(\theta)_{ijk}$. We use four equal-aperture filters, at $\theta = 7$, 14, 25, and 40 arcminutes. As in the case of 2PCF, the subscript indicates the triplet of the redshift bins to correlate, and runs over the 20 possible combinations of a four tomographic redshift bin setup. By restricting ourselves to equal-aperture filters, we avoid an unnecessary increase in our data vector length and retain around 90% of the full mass aperture information, as investigated by Burger *et al.* [25].

We concatenate the measured 2PCF and third-order mass aperture statistic in a joint data vector, $\boldsymbol{d} = [\xi_\pm(\theta)_{ij}, \langle \mathcal{M}_{\rm ap}^3 \rangle(\theta)_{ijk}]$ for each realization, and estimate the covariance by the sample covariance of all the simulation realizations as follows:

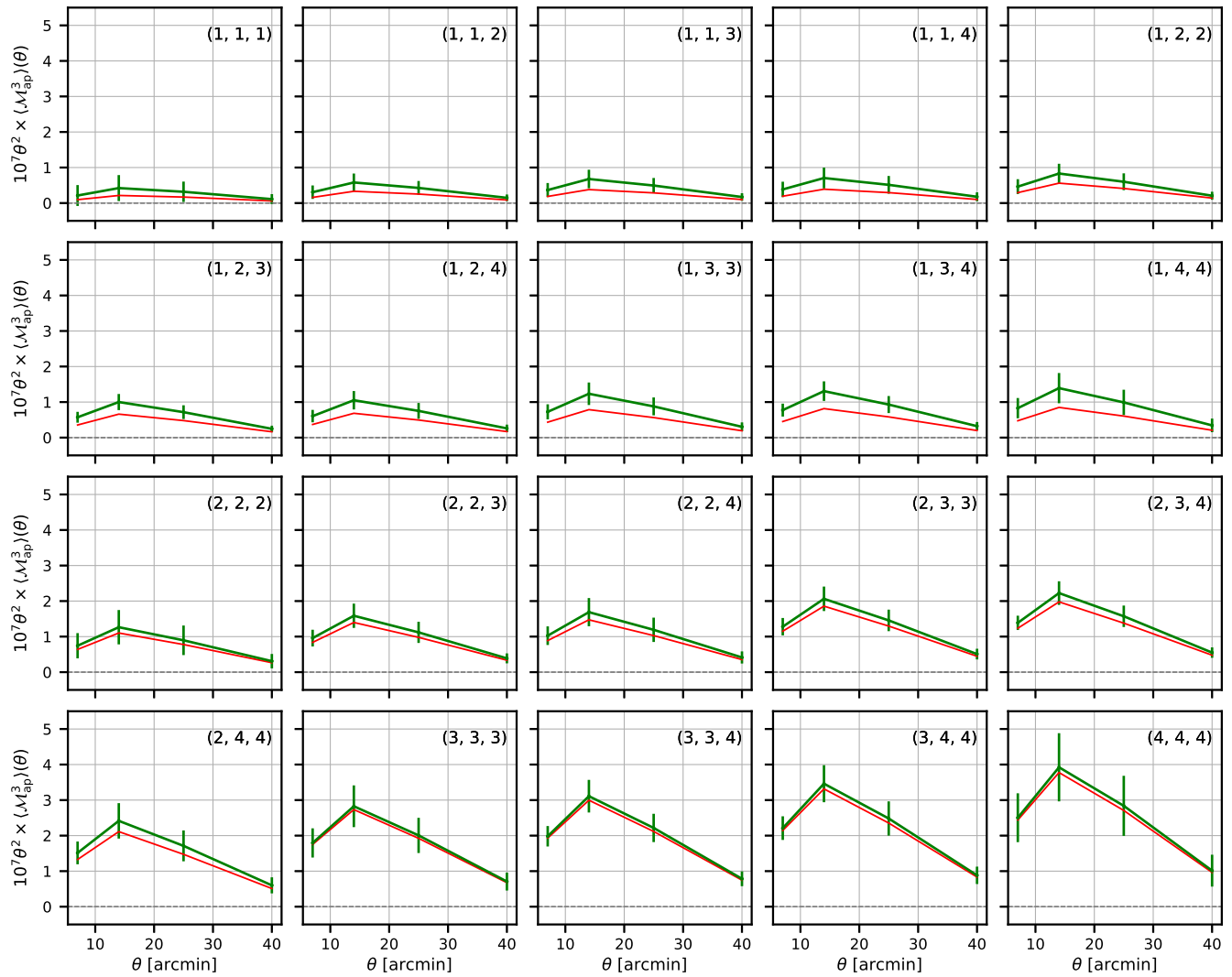

FIG. 2. Mass aperture statistic data vector for simulated analysis. Panels show the aperture mass statistic as a function of filter radii θ for different redshift bin combinations (i, j, k) indicated on the upper right corner of each panel. The green line indicates our synthetic data vector computed from our pipeline at CosmoGridV1 cosmology. The error bars are estimated from the measured covariance. We also include a red line indicating a theoretical calculation with an artificially increased intrinsic alignment signal to show the effect of this systematic on third-order statistics. We use $A_1 = 1.0$ and $\alpha_1 = 0.5$, and find that this is more relevant for the lower redshift bins.

$$\boldsymbol{C} = \frac{1}{N_{\rm real} - 1} \sum_{r=1}^{N_{\rm real}} [\boldsymbol{d}^r - \bar{\boldsymbol{d}}][\boldsymbol{d}^r - \bar{\boldsymbol{d}}]^{\rm T}, \tag{36}$$

where $\boldsymbol{d}^{\rm r}$ is the measurement of the joint data vector from the rth realization, $\bar{\boldsymbol{d}}$ is the mean over all realizations, and $N_{\rm real} = 796$.

Through a preliminary comparison of our simulated covariance with jackknife and bootstrap estimates from Y3 data, we verify that our current NSIDE introduces an inaccuracy in the mass aperture covariance if we do not cut our 3PCF measurements at $\theta \approx 8'$, which is slightly larger than two times the simulation pixel size. Once this low cut is implemented on the two triangle length parameters, there is no need to introduce additional cuts on the maximum multipole parameter. We implement this cut and recompute our $\langle \mathcal{M}_{\rm ap}^3 \rangle$ measurements and joint $\xi_\pm$ and $\langle \mathcal{M}_{\rm ap}^3 \rangle$ covariance with the reduced 3PCF data.

Our mass aperture data vector is presented in Fig. 2, and was computed with our theory pipeline at CosmoGridV1 cosmology. We use our simulated covariance to compute the total SNR over all redshift bin combinations, finding it to be SNR = 8.9. The highest contributions towards the total SNR can be traced to the redshift bin combinations $(z_1, z_2, z_3) = (3, 4, 4)$, (3, 3, 4), and (2, 3, 4), each yielding individual values of SNR between 6 and 6.3. For this reason, when we explore parameter constraints with subsets of the data, we include a scenario in which we focus on the high-redshift information and add to the autocorrelations the combinations (3, 3, 4) and (3, 4, 4).

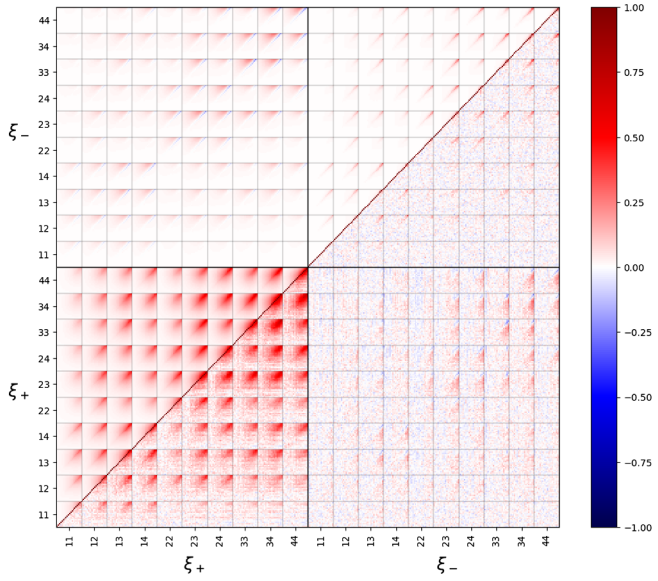

FIG. 3. Covariance of second-order shear correlation functions. The upper triangle is the analytic result, and the lower triangle is measured from 796 CosmoGridV1 simulations. The covariances are consistent, having sufficiently close amplitude and structure.

In order for our likelihood model to bring robust contours on parameter space, we must accurately model three components of our covariance: the two-point correlations, the mass aperture, and the cross-covariance between both (Figs. 3–5). We confirm the validity of our two-point covariance by comparing it with the analytic model by [5,6] (Fig. 3). We verify the similarity in structure between the measured and analytic scenarios. The amplitude of both covariances is also consistent, with the average ratio of their diagonal elements being sufficiently close to unity. The effect on the two-point function parameter constraints of switching from an analytic to a simulated covariance is presented in Fig. 6, where we also apply Massively Optimized Parameter Estimation and Data (MOPED) compression on the simulated covariance chains, as described in Sec. IV B.

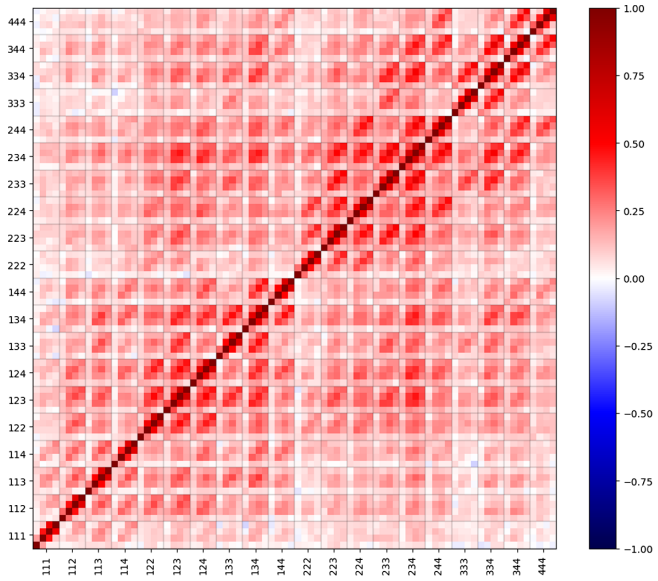

FIG. 4. Covariance of the aperture mass statistics estimated from 796 CosmoGridV1 simulations. The number indicated by the ticks on x- and y-axes are the triplets of the redshift bins. The off-diagonal structure amounts mostly to cross-correlations between different aperture filters for the same or nearby redshift bin combinations. This is expected, given that there is overlap between the 3PCF contributions to the $\langle \mathcal{M}_{\rm ap}^3 \rangle$ integral at different filters. We check the convergence of this covariance in Appendix C.

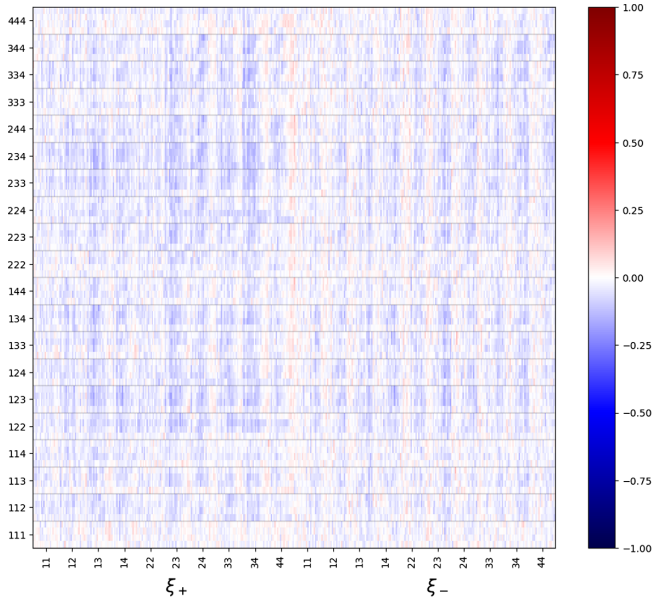

FIG. 5. Cross-covariance between the two-point functions and the aperture mass statistics estimated from 796 CosmoGridV1 simulations. The numbers indicated by the ticks on x/y axes are the pairs/triplets of the redshift bins. The lack of noticeable cross-correlation even for the same redshifts is due to the SNR discrepancy between the statistics.

The covariance of our $\langle \mathcal{M}_{\rm ap}^3 \rangle$ data vector is shown in Fig. 4. We verify that, within the same redshift bin combinations, there is high cross-correlation between the information from different aperture angles θ. Nonetheless, the cross-correlation over different z-bin combinations is low.

The cross-covariance between two-point and three-point information is also not significant, and there is no structure on the matrix, as can be seen in Fig. 5. Normalizing the full covariance matrix so that it has a diagonal equal to unity, we find that the largest value for the cross-covariance equals 0.13. We note that this absence of correlation extends even to data vector entries that probe the same redshift bins. Even though it would be reasonable to expect that the (i, i) sections of the $\xi_\pm$ data vector would correlate with the (i, i, i) section of the $\langle \mathcal{M}_{\rm ap}^3 \rangle$ vector, at least more than with the other parts of the $\langle \mathcal{M}_{\rm ap}^3 \rangle$ vector, this effect can

be suppressed by the SNR discrepancy between the statistics.

IV. METHODOLOGY

A. Parameter inference

For the parameter inference process, we attempt to fit our theoretical modeling to the data vector through Bayesian analysis, where the posterior distribution of the model parameters $\boldsymbol{p}$ for a given data vector $\boldsymbol{d}$ is proportional to the product of the data likelihood and the prior on the model parameters, as follows:

$$\mathcal{P}(\boldsymbol{p}|\boldsymbol{d}) \propto \mathcal{L}(\boldsymbol{d}|\boldsymbol{p})\Pi(\boldsymbol{p}). \tag{37}$$

We adopt a non-Gaussian likelihood in order to accurately propagate the variance of our covariance matrix into our final parameter constraints. It has been shown by Sellentin and Heavens [50] that, when using a covariance matrix estimated from simulations, a marginalization over the inverse-Wishart distribution of the true underlying covariance will yield a likelihood that follows a modified multivariate t-distribution. An alternative approach is to use a Gaussian likelihood and rescale the inverse covariance by the multiplicative factors introduced by Hartlap *et al.* [51] and Dodelson and Schneider [52]. We adopt the likelihood proposed by Percival *et al.* [53], which follows a t-distribution and is designed to yield Bayesian credible intervals for the model parameters that would match the frequentist-based confidence intervals of the multiplicative factor approach.

Our likelihood is written in terms of the χ-squared difference between the data and the theoretical prediction $\boldsymbol{t}(\boldsymbol{p})$ for a set of parameters $\boldsymbol{p}$, as follows:

$$\ln \mathcal{L}(\boldsymbol{d}|\boldsymbol{p}) = -\frac{m}{2}\ln\left(1+\frac{\chi^2}{N_{\rm real}-1}\right) + \text{const} \tag{38}$$

$$\chi^2 = [\boldsymbol{d}-\boldsymbol{t}(\boldsymbol{p})]^{\rm T}\boldsymbol{C}^{-1}[\boldsymbol{d}-\boldsymbol{t}(\boldsymbol{p})]. \tag{39}$$

Here $N_{\rm real} = 796$ is the number of simulations used to estimate the covariance, and m is a factor given by

$$m = N_{\boldsymbol{p}} + 2 + \frac{N_{\rm real}-1+f_D}{1+f_D}, \tag{40}$$

with $N_{\boldsymbol{p}}$ being the number of model free parameters and f_D the Dodelson-Schneider factor [52], which is a function of $N_{\rm real}$, $N_{\boldsymbol{p}}$, and the dimension of the data vector $N_{\boldsymbol{d}}$, as follows:

$$f_D = \frac{(N_{\boldsymbol{d}}-N_{\boldsymbol{p}})(N_{\rm real}-N_{\boldsymbol{d}}-2)}{(N_{\rm real}-N_{\boldsymbol{d}}-4)(N_{\rm real}-N_{\boldsymbol{d}}-1)}. \tag{41}$$

We compute the inverse covariance using the measured covariance matrix from our 796 CosmoGridV1 realizations described in Sec. III. In this paper, we consider several choices for the data vector $\boldsymbol{d}$, which are summarized in Table II along with their dimensions $N_{\boldsymbol{d}}$. The lower part of the table includes MOPED compression of the data vectors, which is discussed in Sec. IV B.

Our fiducial data vector includes the ξ_+ and ξ_- functions for two-point statistics, as well as the $\langle\mathcal{M}_{\rm ap}^3\rangle$ function for three-point correlations. For the $\xi_\pm$ part of our data vector, we used 20 log-spaced values of θ for each source redshift bin combination, with $\theta_{\rm min} = 2.5'$ and $\theta_{\rm max} = 250'$. This gives us a total length of 200 for ξ_+ and 200 for ξ_-. To avoid baryonic contamination at the level of the two-point functions, we performed the scale cuts validated by Secco *et al.* [5] for use on DES-Y3 shear data. After applying cuts, our total size for the $\xi_\pm$ data vector became equal to 227.

For the mass aperture part of the data vector, we used our full set of 80 data points, which correspond to four aperture values for each of the 20 redshift bin combinations. Our limits of $\theta_{\rm min} = 7'$ and $\theta_{\rm max} = 40'$ are informed by the tests described in Sec. IV C. Therefore, the length of the fiducial

TABLE II. Choices of data vector considered in this paper.

Label	Description	Data dim. $N_{\boldsymbol{d}}$
`2 pcf+map3`	$\xi_\pm$ and all $\langle\mathcal{M}_{\rm ap}^3\rangle$	307
`2 pcf+map3 (auto)`	$\xi_\pm$ and auto-z ($z_1 = z_2 = z_3$) of $\langle\mathcal{M}_{\rm ap}^3\rangle$	243
`2 pcf+map3 (auto+334+344)`	adding $(z_1, z_2, z_3) = (3,4,4)$ and $(3,4,4)$ of $\langle\mathcal{M}_{\rm ap}^3\rangle$ to `2 pcf+map3 (auto)`	251
`2 pcf`	only 2 pcf	227
`map3`	$\langle\mathcal{M}_{\rm ap}^3\rangle$	80
`2 pcf [MOPED] +map3`	MOPED compression of `2 pcf` and all `map3` (This is the fiducial)	96
`2 pcf [MOPED] +map3 (auto)`	MOPED compression of `2 pcf` and `map3 (auto)`	32
`2 pcf [MOPED] +map3 (auto+334 +344)`	MOPED compression of `2 pcf` and `map3 (auto+334+344)`	40
`2 pcf [MOPED]`	MOPED compression of `2 pcf`	16

TABLE III. List of the model parameters and their priors. The first section summarizes the cosmological parameters, and the following sections are nuisance parameters for the observational and astrophysical systematics: the residual photometric redshift uncertainty, the shear multiplicative bias, and the intrinsic alignment of the source galaxies. We employ Gaussian priors $\mathcal{N}$ for parameters of observational systematics (instead of uniform priors $\mathcal{U}$ which are used for cosmological and astrophysical parameters) as these are determined by the calibration pipelines [5].

Parameter	Prior
Cosmological parameters	
Ω_m	$\mathcal{U}(0.1, 0.5)$
S_8	$\mathcal{U}(0.7, 0.84)$
h	$\mathcal{U}(0.55, 0.91)$
Ω_b	$\mathcal{U}(0.03, 0.07)$
n_s	$\mathcal{U}(0.87, 1.07)$
m_ν [eV]	$\mathcal{U}(0.06, 0.6)$
w_0	$\mathcal{U}(-2.0, -0.3333)$
Photo-z errors	
Δz_1	$\mathcal{N}(-0.0, 0.018)$
Δz_2	$\mathcal{N}(-0.0, 0.015)$
Δz_3	$\mathcal{N}(-0.0, 0.011)$
Δz_4	$\mathcal{N}(-0.0, 0.017)$
Multiplicative shear calibration	
m_1	$\mathcal{N}(-0.0063, 0.0091)$
m_2	$\mathcal{N}(-0.0198, 0.0078)$
m_3	$\mathcal{N}(-0.0241, 0.0076)$
m_4	$\mathcal{N}(-0.0369, 0.0076)$
Intrinsic alignment parameters	
A_{IA}	$\mathcal{U}(-5, 5)$
α_{IA}	$\mathcal{U}(-5, 5)$

uncompressed two-point plus three-point data vector is 307.

Besides the full fiducial data vector, we also perform parameter inference with the $\xi_\pm$ vector alone, and also with the $\langle \mathcal{M}_{ap}^3 \rangle$ alone. Then, we introduce subsets of $\langle \mathcal{M}_{ap}^3 \rangle$ to the $\xi_\pm$ vector, in order to understand which redshift bin combinations add more information. We start with only the redshift autocorrelations, which give us a total data vector size of $4 \times 4 = 16$. Next, we add the correlations between bins (3, 3, 4) and (3, 4, 4), to include more information from the high-redshift sources, which contribute significantly to the total signal-to-noise. This gives us a data vector of length $6 \times 4 = 24$.

For our theory modeling, we considered the set of parameters listed on Table III. This gives us six cosmological parameters (plus one for our wCDM model), and a set of 10 nuisance parameters, over which our results are marginalized. We use priors consistent with the shear two-point analysis of DES-Y3 data described by Secco *et al.* [5], reproducing their prior ranges for most parameters and being informed by their posteriors for Ω_m and S_8, in order to avoid extrapolating the limits of our emulator sampling.

We build our analysis pipeline in the CosmoSIS package [54], which offers an easy interface to the various samplers and calculation modules. For the modeling of the 2PCF theoretical vector, we utilize the existing modules in the `cosmosis-standard-library`, a supplemental library tailored for cosmology analysis. We newly build the CosmoSIS modules for third-order shear statistics, using the FASTNC code [34].

We sample the posterior using MultiNest [55] through CosmoSIS. For the MultiNest hyperparameters, we use `nlive` $= 500$, `efficiency` $= 0.3$, `tolerance` $= 0.1$. Once the pipeline finishes sampling, we check if the chain covers most of the posterior volume by plotting the weight of samples as a function of sample identification numbers [56].

B. Data compression

As we have seen in the last section, the dimension of the joint analysis data vector grows up to 307, which is comparable to the number of CosmoGridV1 simulations used for covariance estimate. In this case, the Hartlap factor can be significant, and even with the Hartlap factor, the simulation-based inverse covariance fails to describe the true covariant structure of the data vector elements. In order to solve this difficulty, we reduce the dimension of the data vector using the MOPED [57,58] algorithm. In this algorithm, we transform the original data vector to a compressed data vector using a linear transformation,

$$\boldsymbol{d} \to \boldsymbol{B}^{\mathrm{T}} \cdot \boldsymbol{d}, \tag{42}$$

where the transformation matrix $\boldsymbol{B}$ is the two-dimensional matrix with shape $(N_{\boldsymbol{d}}, N_{\boldsymbol{p}})$, being $N_{\boldsymbol{p}}$ the dimension of the model parameter vector. Therefore MOPED compresses the original data vector down to the size of the model parameter vector, $N_{\boldsymbol{p}}$. The transformation matrix is designed to maximize the Fisher information after compression, and it is shown that the compression conserves the Fisher information. Specifically, the ith element of the data vector is designed to maximize the ith model parameter's Fisher information.

To obtain the MOPED transformation matrix, we calculate the Fisher matrix, which requires the inverse covariance. One possibility that emerges is to use the rough estimate of the inverse covariance matrix from a small number of simulations. However, we found this not to be an adequate approach. If we use a covariance derived with a simulation number comparable to the dimension of the data vector, the inverse variance of the data elements can be misestimated and the compression matrix can upweight the noisy data vector elements by mistake. This process leads to overconfidence in the parameter inference. Figure 6 shows how much our constraints from the two-point function can be overconfident when we use the simulation-based covariance from a small number of simulations to estimate the compression matrix.

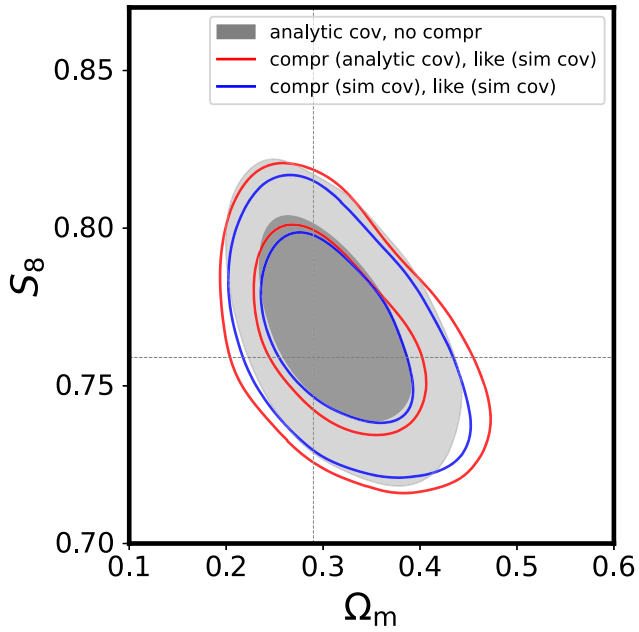

FIG. 6. Cosmological constraints with distinct MOPED data compression setups. The gray contour uses the analytic covariance for inference with no compression. The red contour is obtained with the MOPED compressed data vector, where the compression matrix is estimated with the analytic covariance while the simulation-based covariance is used in the likelihood during parameter inference. The blue contour is similar to the red contour, but the simulation-based covariance is used for the estimate of the compression matrix. Here, only the 2PCF data vector is considered. The red contour is slightly larger than the gray due to the difference between the CosmoGridV1 input cosmology and that of the analytic covariance. For our simulated analysis we use the analytic covariance to estimate the compression matrix (red contour). By doing this, we avoid the overconfidence present in the blue contour.

For this reason, we perform a partial data compression only on the 2PCF part of the data vector, for which we have an analytic model of the covariance matrix. We keep our mass aperture data vector uncompressed, and perform MOPED on ξ_+ and ξ_-, using for compression the analytic model described by Refs. [5,6]. The performance of the 2PCF compression is shown in Fig. 6, where we can see that the compression works well to preserve the information content of the 2PCF. The confidence region of the compressed data analysis is slightly larger than that of the original contour due to the larger cosmological amplitude parameter used in CosmoGridV1 simulations relative to that of the analytic covariance. With the estimated compression matrix for 2PCF, we define the compression matrix for the full data vector by forming the block diagonal matrix.

With the estimated MOPED transformation matrix, we compress the data vector as in Eq. (42), and compute the covariance of the compressed data vector as

$$\boldsymbol{C} \to \boldsymbol{B}^{\mathrm{T}} \cdot \boldsymbol{C} \cdot \boldsymbol{B}. \tag{43}$$

We compute the transformation matrix for all the data vector choices discussed in the last section. The compressed versions of the data vector are summarized in Table II.

C. Pipeline validation

To perform a robust inference of the cosmological parameters, we need to confirm that our analysis pipeline can recover the underlying true cosmological parameters from data. Thus, we validate our analysis pipeline by performing parameter inference on a simulated data vector at CosmoGridV1 cosmology. We present our results in Sec. V.

We also test the validity of our scale cuts against the systematic effect of baryonic feedback on matter clustering. The baryonic effect on matter clustering can be interpreted as the outward displacement of the matter particles in the host halo due to the baryonic physics, e.g., active galactic nuclei feedback, star formation or gas heating. As a result, the baryonic effect prevents matter clustering at small scales, leading to the suppression in the matter power- and bi-spectra [59]. The physics of these effects is studied in hydrodynamical simulations utilizing subgrid physical modeling [60,61], but its whole extension is still little understood when compared to gravity-only dynamics, which can lead to uncertain estimates of the resultant suppression.

Because of this difficulty in modeling the baryonic effect on the matter power- and bi- spectra, we do not include baryonic physics in our model, but rather try to remove from the data vector the scales that could be contaminated by the baryonic effect. To this end, we generate a synthetic data vector of the 2PCF and the mass aperture that includes a baryonic effect similar in amplitude to that of the OverWhelmingly large simulations (OWLS) simulations [62]. For the 2PCF, we estimate the suppression factor for $P(k)$ from the OWLS simulations, multiply it with the gravity-only nonlinear matter power spectrum, and then transform it to the $\xi_\pm$ observable. For third-order shear statistics, we start by computing the suppression factor for the matter bispectrum through the fitting formula in Appendix C of Takahashi *et al.* [31]. This expression was calibrated from the TNG-300 (IllustrisTNG) simulation, which has a baryonic effect smaller in amplitude than that of the OWLS simulations. In order to have a consistent baryonic effect between 2PCF and third-order shear statistics, we rescale the suppression factor accordingly. The parameter R_b is defined as the ratio between the matter bispectrum with (B_b) and without (B_{DM}) baryons, as follows:

$$R_b(k_1,k_2,k_3) \equiv \frac{B_b(k_1,k_2,k_3)}{B_{\mathrm{DM}}(k_1,k_2,k_3)}. \tag{44}$$

For scales in which the baryonic effect is important, we can approximate the expressions for the matter power- and bi- spectra to the 1-halo regime, in which $P(k) \propto u(k)^2$ and

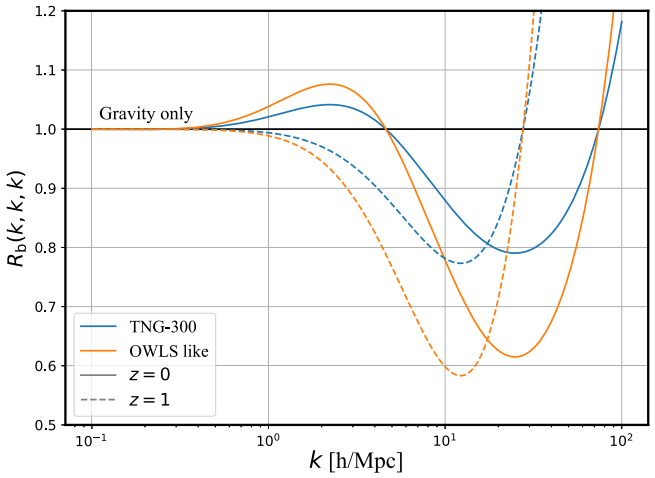

FIG. 7. Suppression factors on the matter bispectrum, R_b, due to baryonic feedback. Here we plot $R_b(k_1, k_2, k_3)$ for the equilateral configuration $k_1 = k_2 = k_3$. The blue lines represent the TNG-300 simulations, obtained with the fitting formula of Takahashi *et al.* [31]. The orange lines represent the OWLS simulations, obtained by rescaling the blue lines by a factor of $1.5^{3/2}$. The solid and dashed lines are at redshifts $z = 0$ and $z = 1$, respectively. We use the OWLS-like suppression for pipeline validation.

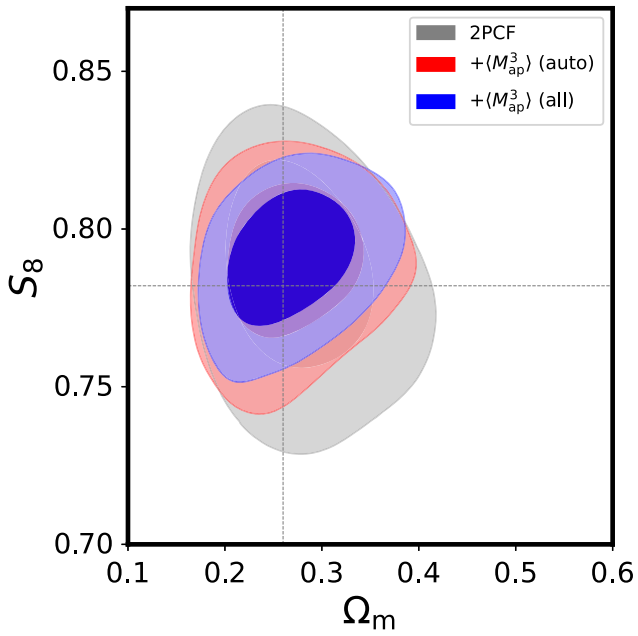

FIG. 8. Results of our simulated 2- and 3-point cosmological analysis. The gray contour is the constraint from $\xi_\pm$ alone, while the colored contours are constraints the addition of the third-order mass aperture statistic. The dotted lines indicate the input values of Ω_m and S_8. We find an improvement of 83% on the Ω_m-S_8 figure of merit when we add the full mass aperture data vector.

$B(k, k, k) \propto u(k)^3$, with $u(k)$ being the halo profile. Thus, we have

$$\frac{B_b(k, k, k)}{B_{\rm DM}(k, k, k)} \propto \left(\frac{P_b(k)}{P_{\rm DM}(k)}\right)^{3/2}. \tag{45}$$

We compare the peaks in the matter power spectra suppression factors from the OWLS and TNG-300 simulations, estimating them to differ by a factor of 1.5. Therefore, we rescale our original $R_b(k_1, k_2, k_3)$ parameter by a factor of $1.5^{3/2}$, so that it has an amplitude consistent with that of OWLS. The resultant suppression factor is shown in Fig. 7. Using the suppressed matter bispectrum, we compute the shear 3PCF and mass aperture statistics, and perform a run of our pipeline on the baryon-contaminated data vector. We present our results in the following section, along with those of the full simulated analysis.

V. RESULTS

In this section, we show the results of our analysis pipeline validation. We measure our level of joint improvement on the Ω_m and S_8 constraints through the figure of merit (FoM) defined by

$$\text{FoM} = \frac{1}{\sqrt{\text{Cov}(\Omega_m, S_8)}}. \tag{46}$$

Our full run of the pipeline uses synthetic $\xi_\pm$ and $\langle \mathcal{M}_{ap}^3 \rangle$ data at the CosmoGridV1 input cosmology. We use our CosmoGridV1 simulated covariance and perform MOPED compression on the $\xi_\pm$ data vector before concatenating it with the mass aperture information.

With ΛCDM modeling, we find that adding the autocorrelations of the mass aperture increases the FoM on the Ω_m-S_8 plane by 40%, while adding the full mass aperture data improves the 2PCF constraints on Ω_m-S_8 by 83%. The results are shown in Figs. 8 (focused on Ω_m-S_8) and 9 (which includes the constraint from mass aperture alone). The figure of merit for each considered scenario on Ω_m-S_8 is presented in Table IV. Adding the high signal-to-noise redshift bin combinations (3, 3, 4) and (3, 4, 4) of the mass aperture along with its autocorrelations only slightly increased the gain relative to the scenario restricted to the autocorrelations. This shows us that, despite those combinations bringing important contributions to the signal, it is still necessary to combine low and high redshift information to achieve the full constraining power of the mass aperture. Our incremental approach demonstrates the presence of valuable cosmological information at all redshifts.

We attribute the high FoM gain of 83% to the fact that the individual contours from $\langle \mathcal{M}_{ap}^3 \rangle$ and $\xi_\pm$ are nearly orthogonal to each other in the Ω_m-S_8 plane, inducing a significant reduction of the joint contour, as can be seen in Fig. 9. The contours for all cosmological parameters are presented in Appendix D for both ΛCDM and wCDM models.

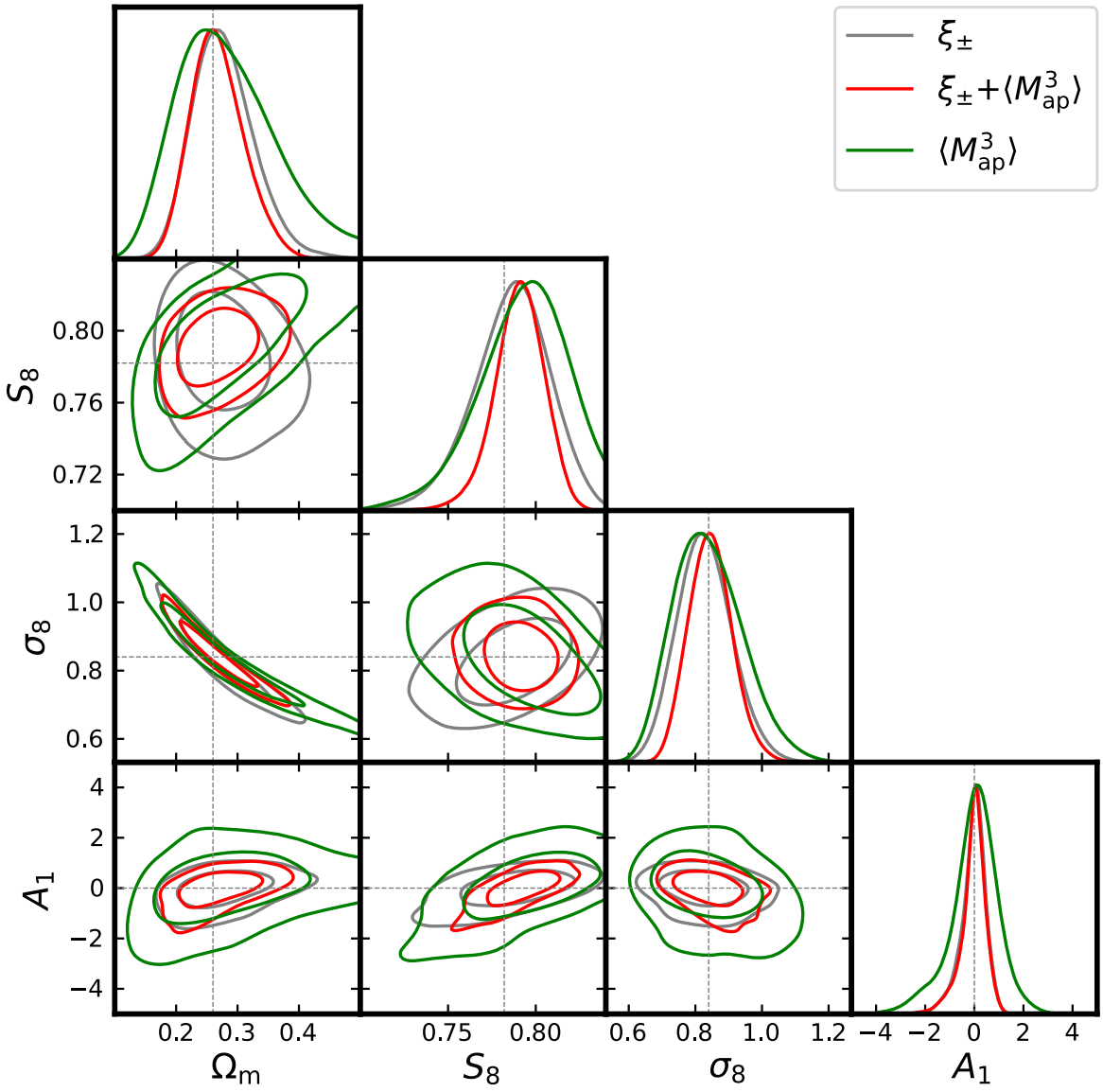

FIG. 9. Results of our simulated 2- and 3-point cosmological analysis. The gray constraint uses 2PCF data alone, the red constraint uses both 2PCF and mass aperture, and the green constraint uses mass aperture data alone. The dotted lines indicate the input values of the parameters. Note the difference in orientation of the $\Omega_{\rm m}$-S_8 contours from $\xi_\pm$ and $\langle \mathcal{M}_{\rm ap}^3 \rangle$, which contributes to the improvement in the joint constraint.

TABLE IV. Figure of merit on the $\Omega_{\rm m}$-S_8 plane for our different data vector choices. The rightmost column summarizes the gain in the FoM compared to the 2PCF-only analysis.

Data vector	FoM	FoM gain
`2 pcf`	930.8	⋯
`2 pcf+map3(auto)`	1299.2	40%
`2 pcf+map3(auto+334+344)`	1408.7	52%
`2 pcf+map3(all)`	1700.8	83%
`wCDM-2 pcf`	653.2	⋯
`wCDM-2 pcf+map3(all)`	928.1	42%

We verify the consistency between the two-point analysis and our full analysis by comparing the mean of the marginalized parameter posteriors. We also compare those values with the CosmoGridV1 input cosmological parameters. The results are shown in Table V.

To verify the adequacy of our 2PCF and mass aperture scale cuts in removing effects from baryonic physics, we run our pipeline on our baryon-contaminated data vector which mimics the baryonic effect present in the OWLS simulations (see Sec. IV C for the detail). The results are shown in Fig. 10. We measure the shift in the mean of the S_8 and $\Omega_{\rm m}$ posteriors caused by introducing baryons on the data vector. For the joint analysis, S_8 is shifted by 0.05σ and

TABLE V. Mean of parameter posteriors from 2PCF and 2PCF + $\langle \mathcal{M}_{\rm ap}^3 \rangle$ analysis.

Parameter	CosmoGridV1	$\xi_\pm$	$\xi_\pm + \langle \mathcal{M}_{\rm ap}^3 \rangle$	$\xi_\pm$ (wCDM)	$\xi_\pm + \langle \mathcal{M}_{\rm ap}^3 \rangle$ (wCDM)
$\Omega_{\rm m}$	0.26	$0.277^{+0.049}_{-0.048}$	0.268 ± 0.044	$0.267^{+0.052}_{-0.051}$	$0.254^{+0.041}_{-0.042}$
S_8	0.782	$0.786^{+0.022}_{-0.021}$	$0.790^{+0.014}_{-0.013}$	0.764 ± 0.032	0.768 ± 0.031
w_0	-1	⋯	⋯	-1.28 ± 0.35	-1.29 ± 0.35

$\Omega_{\rm m}$ by 0.0σ, which is completely insignificant. This indicates that our choice of scale cuts for 2PCF and mass aperture is large enough to cut out the small-scale data vector elements that are more sensitive to the baryonic effect.

We also note that while the presence of baryons slightly moves down the S_8 value on the 2PCF analysis, it slightly increases the S_8 value on the $\langle \mathcal{M}_{\rm ap}^3 \rangle$-only analysis. The net effect on a joint analysis is that of a smaller shift than in each of the individual analyses. As can be seen in Fig. 7, there is a regime of baryonic enhancement on the bispectrum for mid-to-low scales. Due to our cuts and filter choices, the enhancement dominates over the smaller-scale suppression regime on the $\langle \mathcal{M}_{\rm ap}^3 \rangle$ integral, leading to a net enhancement of the $\langle \mathcal{M}_{\rm ap}^3 \rangle$ signal.

Next, we show the improvement of the cosmological parameter constraints in the wCDM model. Here we use the same data vector as in our previous ΛCDM analyses, but we allow the equation-of-state parameter w_0 of dark energy to freely vary within the prior range specified in Table III. The constraints are shown in Fig. 11. We find no improvement on the marginalized w_0 uncertainty when adding third-order shear information to the 2PCF signals. The individual $\Omega_{\rm m}$ gain was found to be of 20%, while the S_8 gain was 2%. In contrast, the marginalized parameter gains in the ΛCDM scenario were of 10% for $\Omega_{\rm m}$ and 36% for S_8, as can be seen in Table V. Despite the negligible S_8 and w_0 individual gains, the improvement in the w_0-S_8 figure of merit was found to be of 36%, indicating that mass aperture data helps to break the existing w_0-S_8 degeneracy.

Finally, in addition to constraints on the cosmological parameters, we also find a slight improvement on the intrinsic alignment amplitude parameter as detailed in Fig. 12. With the $\xi_\pm$ data alone, we find $A_1 = -0.05^{+0.47}_{-0.48}$. When we add the mass aperture information, this constraint moves to $A_1 = -0.05^{+0.48}_{-0.45}$, being tightened by a factor of 3%. As the amplitude of gravitational lensing is larger for higher source redshift bins due to the larger line-of-sight coverage, the relative contribution of intrinsic alignment is larger at lower redshift bins. This can be seen in Fig. 2, where the predicted signals with and without intrinsic alignment are shown.

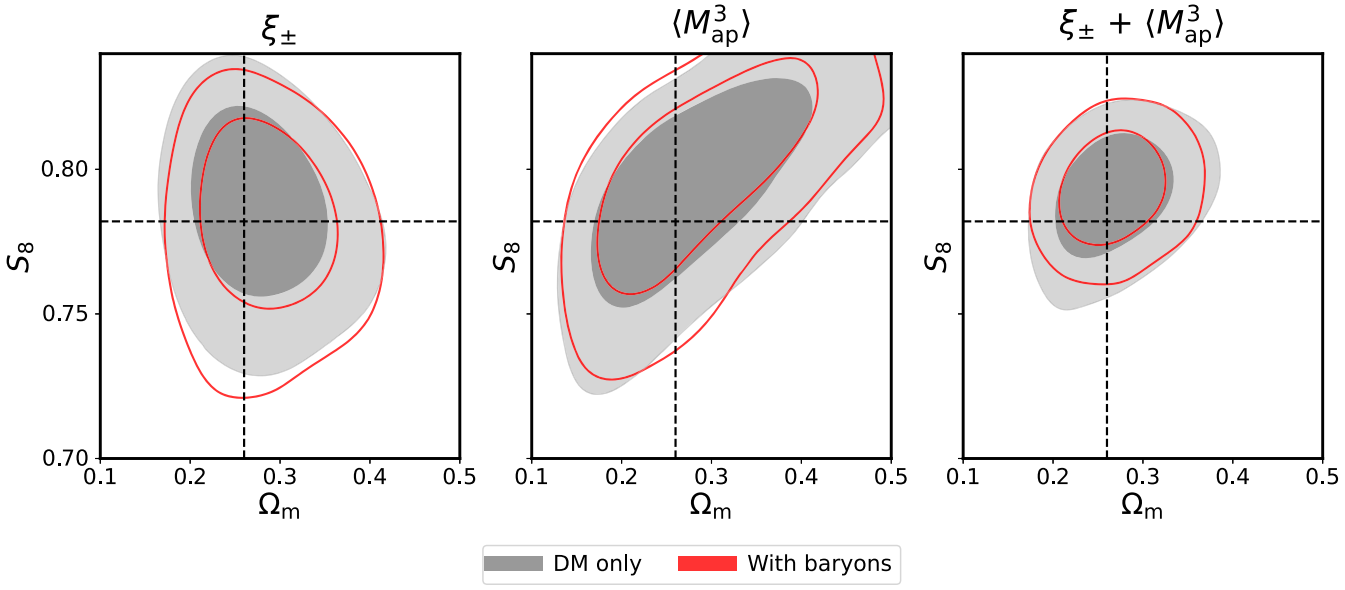

FIG. 10. Validation of our analysis pipeline against the effect of baryonic feedback. The gray contours are the results with the dark matter only data vector, while the red contours use the baryon contaminated data vector. The first panel includes only 2PCF data, the second panel only $\langle \mathcal{M}_{\rm ap}^3 \rangle$ data, and the third panel uses the full data vector. Our baryon-contaminated data vector uses an amplitude similar to that seen on the OWLS simulations (see also Fig. 7 for the baryonic effect on the bispectrum). We find no significant shift in $\Omega_{\rm m}$ and S_8 for any of the datasets.

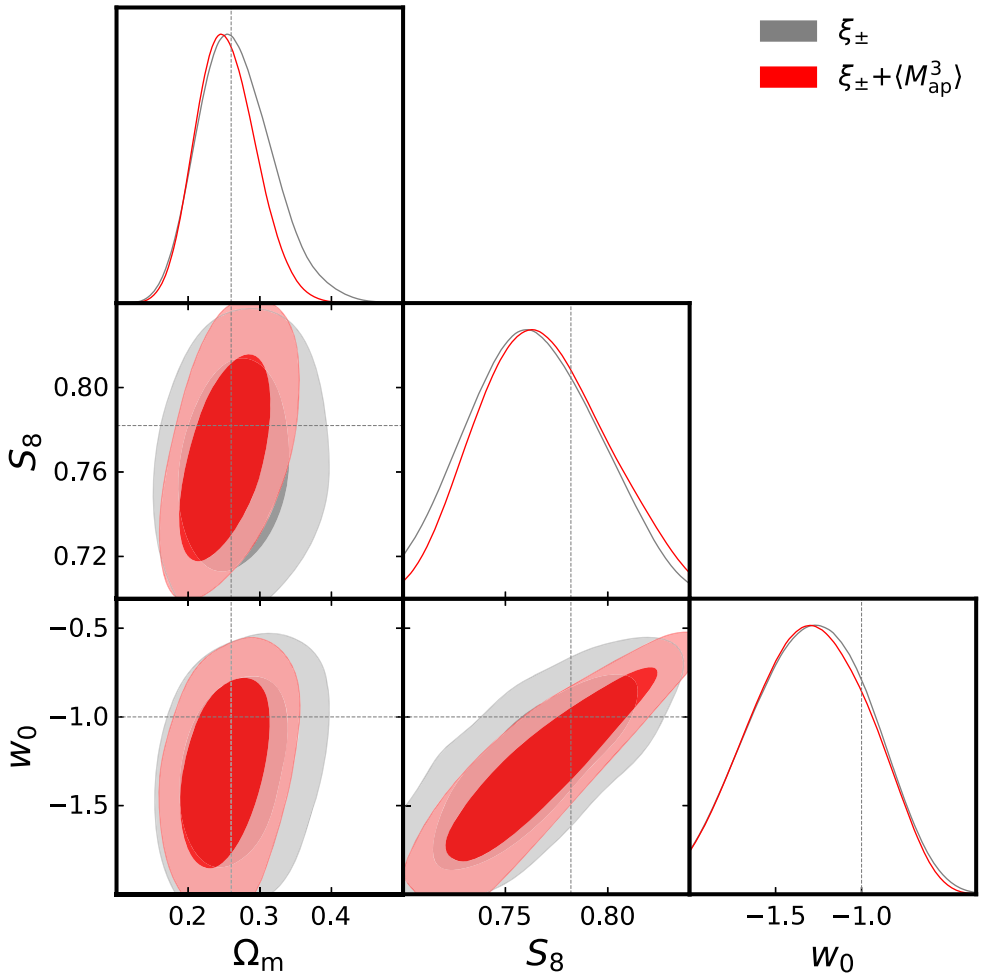

FIG. 11. Results of our simulated analysis for the wCDM cosmological model. The gray and red contours are from 2PCF-only and 2PCF + $\langle \mathcal{M}_{\rm ap}^3 \rangle$, respectively. The dotted lines are the input values of the parameters. We find an improvement of 36% on the joint w_0-S_8 constraint.

VI. CONCLUSIONS

We have presented results from a pipeline to obtain robust constraints on cosmological parameters using second- and third-order shear statistics. The full configuration and redshift-dependent shear three-point correlations are measured from simulated shear catalogs and compressed for computational efficiency into the third moment of the mass aperture statistic $\langle \mathcal{M}_{\rm ap}^3 \rangle$. We build a theoretical model for $\langle \mathcal{M}_{\rm ap}^3 \rangle$ and integrate it into the CosmoSIS framework. The methodology developed in our earlier work [34] enables rapid computation of $\langle \mathcal{M}_{\rm ap}^3 \rangle$ at every point in parameter space. Building upon this, we introduce an emulator that further accelerates the computation, reducing the runtime from approximately 40 seconds to just 0.03 seconds—an improvement by a factor of $\mathcal{O}(10^3)$. This substantial speedup makes posterior Monte Carlo sampling feasible and significantly enhances computational efficiency.

We compute the combined likelihood of the two-point functions $\xi_\pm$ and of $\langle \mathcal{M}_{\rm ap}^3 \rangle$, including the covariance matrix from a set of 796 CosmoGridV1 shear maps. We verify the robustness of our pipeline against observational systematics and show that our scale cuts sufficiently remove contamination from baryonic feedback. We thus complete the preparation of our pipeline for an analysis with DES-Y3 data.

We also determine the improvement in the $\Omega_{\rm m}$-S_8 figure of merit of our second- and third-order shear analysis relative to the $\xi_\pm$ data vector alone. We obtain a significant gain of 83%, which is much higher than the gains reported by Gatti *et al.* [63] in studies of the third moment with DES-Y3. While there are several previous analyses of third-order correlations, Burger *et al.* [25] is the closest to our work as they use the full three-point correlations and find a similar gain (93%) for a KiDS simulated analysis with complete orthogonal sets of E-/B-mode integrals and $\langle \mathcal{M}_{\rm ap}^3 \rangle$. We study the complementarity of three-point information in helping to break degeneracies between cosmological and systematic parameters. As noted in early studies of three-point correlations, their dependence on σ_8 and $\Omega_{\rm m}$ is different from that of two-point correlations (see

Ref. [64] for a perturbation theory description). Our Fig. 9 shows that in practice the degeneracy directions differ only subtly, but they do enable stronger constraints on both parameters from a joint analysis. More quantitatively, the two-point correlation function is sensitive to the combination of $S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5}$ [65], while we found the three-point is to the combination of $S_8' \equiv \sigma_8(\Omega_m/0.3)^{0.43}$. The full parameter constraints from our joint analysis can be seen in Figs. 14 and 15.

Interestingly, for our scale cuts on the three-point correlation function and aperture filters between $7'$ and $40'$, the lensing signal is slightly enhanced for $\langle \mathcal{M}_{ap}^3 \rangle$ due to the presence of baryons. Thus, the mass aperture analysis has the opposite bias on the S_8 constraint compared to the two-point analysis. The combined effect on a joint analysis is a negligible shift in S_8 (below 0.1σ), and thus one of robustness to baryonic feedback relative to that of the individual summary statistics. How generally this holds is an interesting question for future work.

The gain in the Ω_m-S_8 FoM that we find with the full three-point correlation raises interesting questions regarding the use of higher-order statistics. In general, various studies have reported complementary information in an array of higher-order statistics. In particular, Gatti *et al.* [63] compare the Ω_m-S_8 FoM gain from third moments, scattering transforms, and wavelet phase harmonics, with each yielding improvements of 15%, 38% and 53% relative to second moments alone. The combination of the three statistics yields an improvement of 92% relative to a second moments analysis. The gain in FoM they find with all higher-order statistics is very close to the gain we report from three-point correlations alone. While those numbers should not be compared in detail with ours (in particular because we use $\xi_\pm$ as our second-order statistic), they motivate the use of the full three-point function rather than smoothed moments, and of a careful examination of what information remains to be captured by other higher-order statistics. We do believe that the three-point function ought to be the first higher-order statistics to be included beyond two-point analyses, due to its ease of modeling and interpretation.

Interestingly, for three-dimensional clustering analyses, recent findings suggest that the power spectrum and bispectrum come close to saturating the information content [66]. Again, there are several differences in the analyses (in particular, the presence of galaxy bias which necessitates more conservative scale cuts), and the question may not be settled yet (see, e.g., Ref. [67] for the field-level inference argument).

Finally, we also consider the wCDM model, for which we present results for the three key cosmological parameters (S_8, Ω_m and w_0). We find an improvement of 36% in the joint w_0-S_8 constraint. The marginalized uncertainty on w_0 itself shows no improvement [68].

In conclusion, we have shown that using the full three-point correlation functions is a promising way to extract significant additional cosmological information from lensing surveys. The three-point function extracts non-Gaussian information while staying well above the small scales that are impacted by baryonic feedback. The application to DES-Y3 data will follow in a separate publication.

ACKNOWLEDGMENTS

We thank Ryuichi Takahashi, Gary Bernstein and Masahiro Takada for useful discussions. We also thank Kazuyuki Akitsu for providing references to studies on galaxy clustering. B. J. and R. C. H. G. are partially supported by the U.S. Department of Energy Grant No. DE-SC0007901 and S. S. is supported by the JSPS Overseas Research Fellowships. Part of this work was supported by the NASA ROSES Grant No. 22-ROMAN11-0011 via a JPL subaward. Funding for the DES projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF NOIRLab, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium. Based in part on observations at NSF Cerro Tololo Inter-American Observatory at NSF NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which

is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. The DES data management system is supported by the National Science Foundation under Grants No. AST-1138766 and No. AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under Grants No. PID2021-123012, No. PID2021-128989, No. PID2022-141079, No. SEV-2016-0588, No. CEX2020-001058-M and No. CEX2020-001007-S, some of which include ERDF funds from the European Union. I. F. A. E. is partially funded by the CERCA program of the Generalitat de Catalunya. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) do e-Universo (CNPq Grant No. 465376/2014-2). This document was prepared by the DES Collaboration using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, Office of High Energy Physics HEP User Facility. Fermilab is managed by Fermi Forward Discovery Group, LLC, acting under Contract No. 89243024CSC000002.

R. C. H. G. and S. S. developed the pipeline, performed the analysis, and wrote the manuscript. B. J. served as project advisor and contributed to the manuscript. M. J. developed novel TREECORR functionalities. D. A. generated the CosmoGridV1 mocks used in this work. M. G. contributed with the two-point analysis setup. D. G., Z. G., and A. H. contributed by testing the simulations. A. H., G. A. M. and S. P. served as collaboration internal reviewers, and J. L. M. as collaboration final reader.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: THE BIHALOFIT FORMULA

Here we follow the prescription detailed by Takahashi *et al.* [31] to implement the fitted nonlinear bispectrum. To start the BiHalofit computation of the bispectrum, we calculate the following from code for anisotropies in the microwave background (CAMB) [69]:

(i) The linear matter power spectrum $P(k,z)$ on a grid of k and z values.
(ii) The nonlinear scale k_{nl} as for each redshift value.
(iii) The comoving distance $\chi(z)$ and the derivative $dz/d\chi(z)$.

Next, we compute the fitting parameters, as functions of $\log\sigma_8$ and of $n_{\rm eff}$. We divide the parameters into global and dependent parameters, the former being k-independent, and the latter k-dependent. The global parameters are [31]

$$\begin{aligned}
\log b_n &= -3.428 - 2.681\log\sigma_8 \\
&\quad + 1.624\log{\sigma_8}^2 - 0.095\log{\sigma_8}^3 \\
\log c_n &= 0.159 - 1.107 n_{\rm eff} \\
\log \gamma_n &= 0.182 + 0.57 n_{\rm eff} \\
\log f_n &= -10.533 - 16.838 n_{\rm eff} \\
&\quad - 9.3048 n_{\rm eff}^2 - 1.8263 n_{\rm eff}^3 \\
\log g_n &= 2.787 + 2.405 n_{\rm eff} + 0.4577 n_{\rm eff}^2 \\
\log h_n &= -1.118 - 0.394 n_{\rm eff} \\
\log m_n &= -2.605 - 2.434\log\sigma_8 + 5.710\log{\sigma_8}^2 \\
\log n_n &= -4.468 - 3.080\log\sigma_8 + 1.035\log{\sigma_8}^2 \\
\log \mu_n &= 15.312 + 22.977 n_{\rm eff} \\
&\quad + 10.9579 n_{\rm eff}^2 + 1.6586 n_{\rm eff}^3 \\
\log \nu_n &= 1.347 + 1.246 n_{\rm eff} + 0.4525 n_{\rm eff}^2 \\
\log p_n &= 0.071 - 0.433 n_{\rm eff} \\
\log d_n &= -0.483 + 0.892\log\sigma_8 - 0.086\Omega_{\rm m} \\
\log e_n &= -0.632 + 0.646 n_{\rm eff}. \qquad \text{(A1)}
\end{aligned}$$

In order to compute the dependent parameters, we sort each set of (k_1,k_2,k_3) to find the set $(k_{\rm min},k_{\rm mid},k_{\rm max})$. Then, we define $r_1 \equiv k_{\rm min}/k_{\rm max}$ and $r_2 = (k_{\rm mid}+k_{\rm min}-k\,{\rm max})/k_{\rm max}$. Now we compute the dependent parameters [31], as follows:

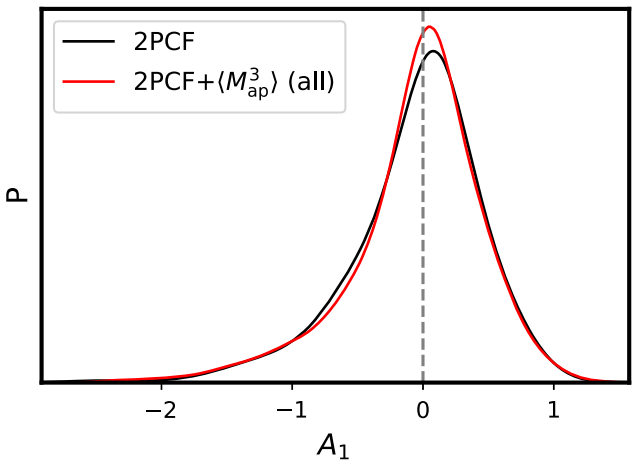

FIG. 12. Result of our simulated analysis focusing on the IA parameter A_1. The black and red contours are from 2PCF-only and 2PCF+$\langle\mathcal{M}_{\rm ap}^3\rangle$ analyses respectively. The vertical dashed line is the input value. A small tightening of the posterior distribution is visible, corresponding to an improvement of 3% on the 68% confidence interval of the A_1 parameter.

$$\log a_n = -2.167 - 2.944\log\sigma_8 - 1.106\log{\sigma_8}^2 - 2.865\log{\sigma_8}^3 - 0.310 r_1\gamma_n$$
$$\log\alpha_n = \min(-4.348 - 3.006 n_{\rm eff} - 0.5745 n_{\rm eff}^2 + 10^{-0.9+0.2 n_{\rm eff}} r_2^2, n_s)$$
$$\log\beta_n = -1.731 - 2.845 n_{\rm eff} - 1.4995 n_{\rm eff}^2 - 0.2811 n_{\rm eff}^3 + 0.007 r_2. \quad \text{(A2)}$$

From these parameters, we compute the one-halo term:

$$B_{1h}(k_1,k_2,k_3,z) = \prod_{i=1,2,3} \frac{1}{a_n q_i^{\alpha_n} + b_n q_i^{\beta_n}} \frac{1}{1+(c_n q_i)^{-1}}, \quad \text{(A3)}$$

where $q_i = k_i/k_{nl}$.

The 2-halo and 3-halo effects are put together in the BiHalofit model B_{3h} term [31]. To compute this part, we take the F_2 kernel defined as

$$F_2(k_1,k_2,k_3) = \frac{5}{7} + \frac{2}{7k_1^2k_2^2}\left(\frac{-k_3^2+k_1^2+k_2^2}{2}\right)^2 - \frac{-k_3^2+k_1^2+k_2^2}{2}\left(\frac{1}{2k_1^2}+\frac{1}{2k_2^2}\right). \quad \text{(A4)}$$

Next, we compute the enhanced power spectrum:

$$P_{\rm enhanced}(k_i,z) = \frac{1+f_n q_i^2}{1+g_n q_i + h_n q_i^2} P_L(k_i,z) + \frac{1}{m_n q_i^{\mu_n} + n_n q_i^{\nu_n}} \frac{1}{1+(p_n q_i)^{-3}}. \quad \text{(A5)}$$

Finally, we get

$$B_{3h}(k_1,k_2,k_3,z) = 2(F_2(k_1,k_2,k_3) + d_n q_3) \times \prod_{i=1,2,3} \frac{1}{1+e_n q_i} P_{\rm enhanced}(k_1,z) \times P_{\rm enhanced}(k_2,z) + \text{two permutations} \quad \text{(A6)}$$

and

$$B(k_1,k_2,k_3,z) = B_{1h}(k_1,k_2,k_3,z) + B_{3h}(k_1,k_2,k_3,z). \quad \text{(A7)}$$

APPENDIX B: ΛCDM AND *w*CDM MASS APERTURE EMULATORS

Here we specify the configurations of the emulators described in Sec. II F. For both ΛCDM and *w*CDM, we emulate the redshift-dependent $\langle\mathcal{M}_{\rm ap}^3\rangle(\theta,z)$ at a specified set of z values. In order to determine the ideal set, we examined possible spacings and attempted to minimize the error relative to our fiducial 130-bin computation. We found that spacing the values linearly between 0.1 and 1.5 with a step size of $\Delta z = 0.05$ captured most of the information from this redshift range. We added to the set seven initial values at 0.005, 0.01, 0.02, 0.03, 0.04, 0.05, and 0.07, in order to capture the low-redshift information. This yields us a total of 36 z values.

We then used this set of z values to generate the emulator training set. For ΛCDM, we used 1300 samples for training/validation and 171 for testing. For *w*CDM, we used 1700 samples for training/validation and 262 for testing. We perform a logarithmic transformation followed by a min-max scaling as preprocessing steps for both networks.

Finally, we obtain optimized network hyperparameters and train our models. For ΛCDM, we use six hidden layers with 64, 256, 1024, 1024, 384 and 192 nodes, respectively. We perform training with a sequence of decreasing learning rates (10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}), associating them with the progressively reduced batch sizes of 60, 30, 20, 10, and 5. Our final validation loss was equal to 6.36×10^{-5}.

For *w*CDM, we used six hidden layers with 64, 256, 1024, 1024, 256 and 192, respectively. Our learning rate and batch size progressions were equal to that of the ΛCDM emulator. The final validation loss was equal to 9.83×10^{-5}.

APPENDIX C: CONVERGENCE OF THE SIMULATED MASS APERTURE COVARIANCE

Here we demonstrate the convergence of our simulated covariance matrix. We select 700 out of our 796 CosmoGridV1 realizations to compute an alternative covariance matrix for the full $\xi_\pm$ and $\langle\mathcal{M}_{\rm ap}^3\rangle$ data vector. We compare it with our fiducial matrix and find both to be similar in structure and in magnitude. Next, we perform parameter estimation with the alternate covariance for both the $\xi_\pm$-only scenario and the $\xi_\pm + \langle\mathcal{M}_{\rm ap}^3\rangle$, following our MOPED compression scheme for the $\xi_\pm$ section of the data vector. We find only a minimal alteration of the contours and posteriors in both scenarios. Therefore, we consider that our set of simulations is large enough to accurately describe the covariance of our full data vector. Our results are shown in Fig. 13.

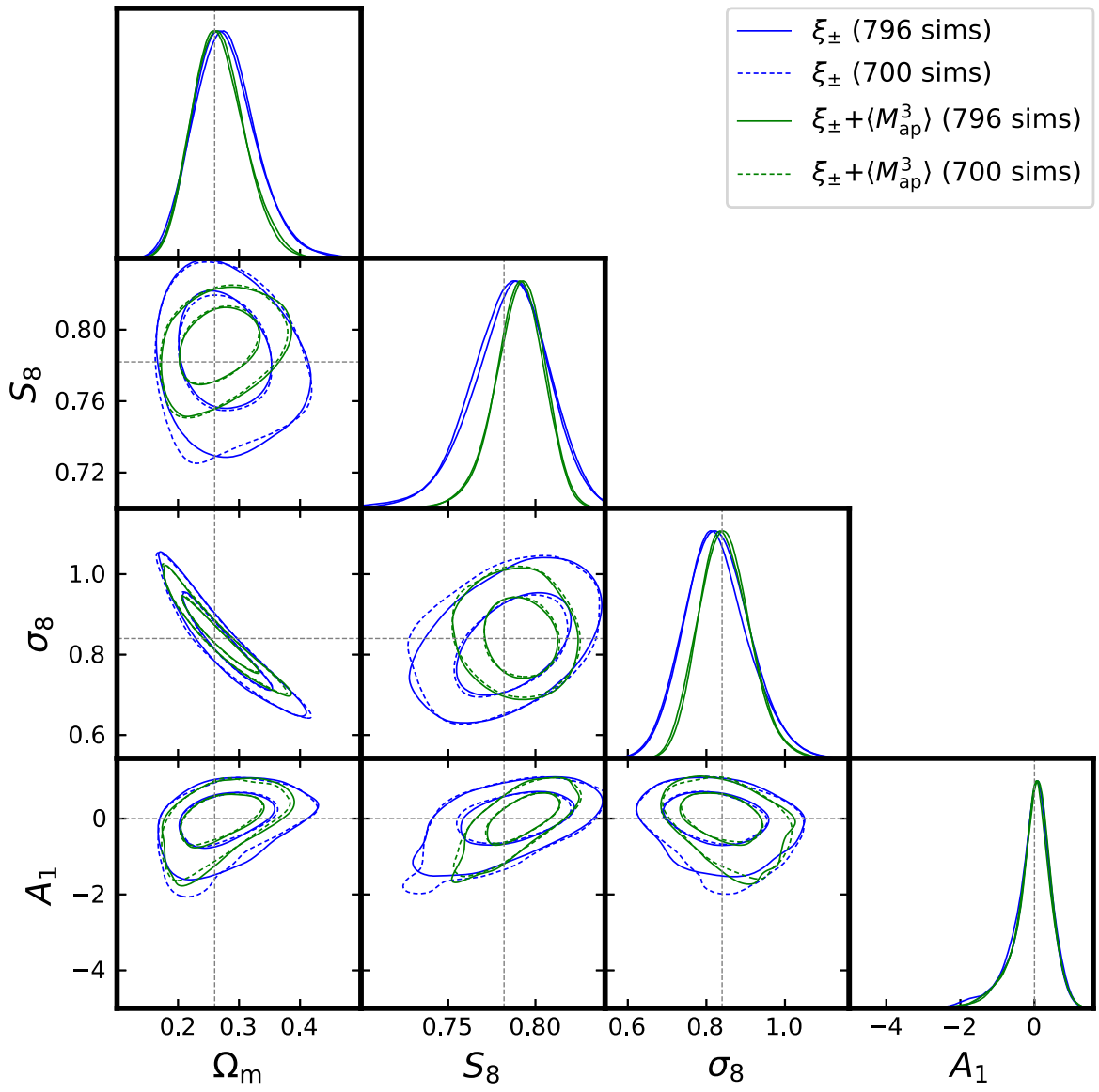

FIG. 13. Parameter constraints from $\xi_\pm$ and $\langle\mathcal{M}_{\rm ap}^3\rangle$ using covariance matrices computed with 700 and 796 CosmoGridV1 realizations. Our results are consistent and attest to the convergence of our simulated covariance matrix.

APPENDIX D: CONTOURS FOR ALL COSMOLOGICAL PARAMETERS

Here we present our ΛCDM and *w*CDM contours for the full set of cosmological parameters. We exclude the constraints on the sum of the neutrino masses because our $\langle \mathcal{M}_{\rm ap}^3 \rangle$ emulator assumes a fixed value of $\Sigma m_\nu = 0.06$. For ΛCDM, our results are shown in Fig. 14. For *w*CDM, we present our results in Fig. 15.

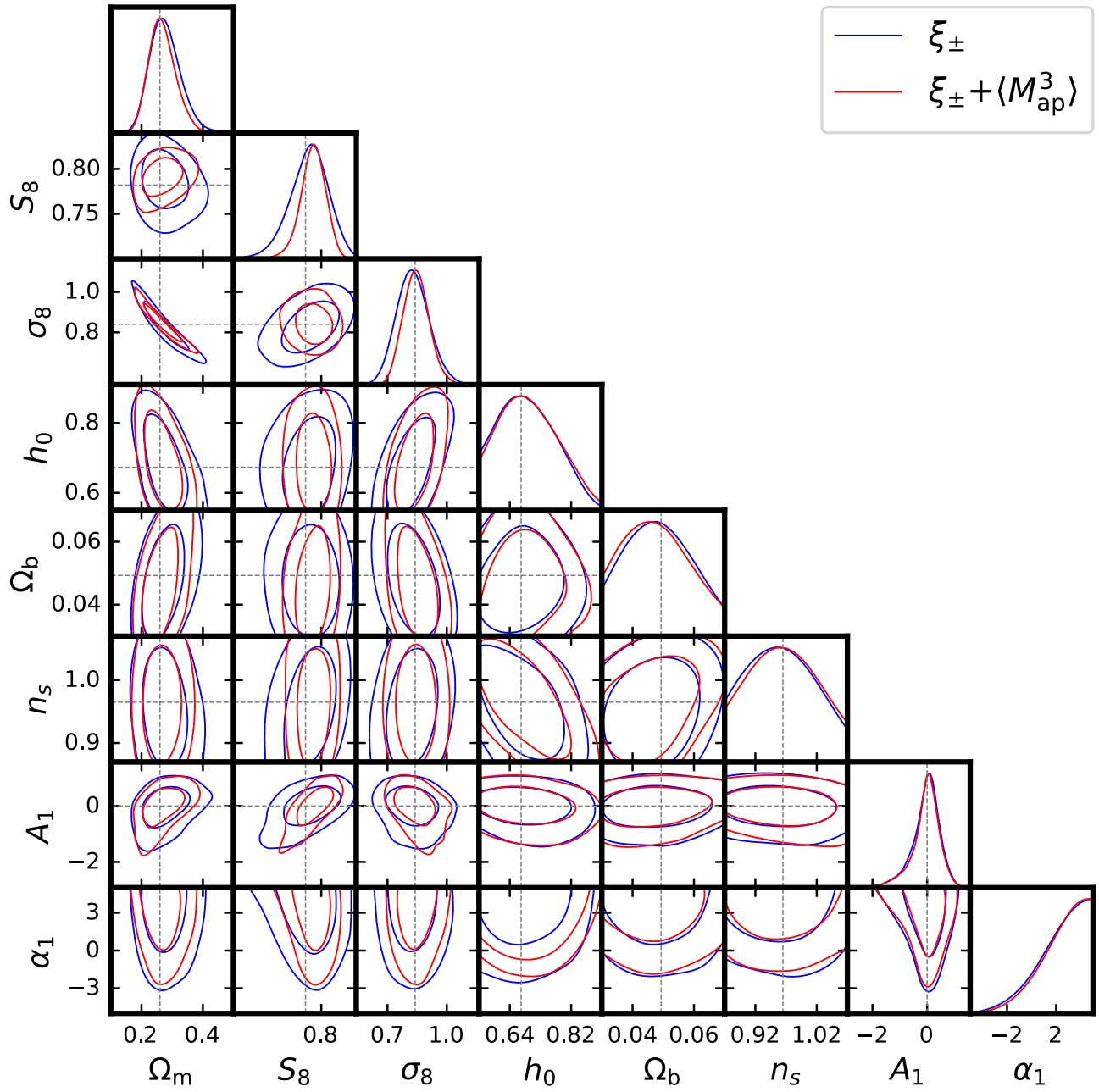

FIG. 14. Parameter constraints on the whole set of cosmological parameters from $\xi_\pm$ and $\langle \mathcal{M}_{\rm ap}^3 \rangle$ using ΛCDM modeling. The blue contours indicate results from the 2PCF alone, while the red contours indicate results from the 2PCF combined with the mass aperture statistic. The gray dotted lines represent the input values for the parameters. We do not indicate any value for α_1 since our input data vector had a zero intrinsic alignment amplitude.

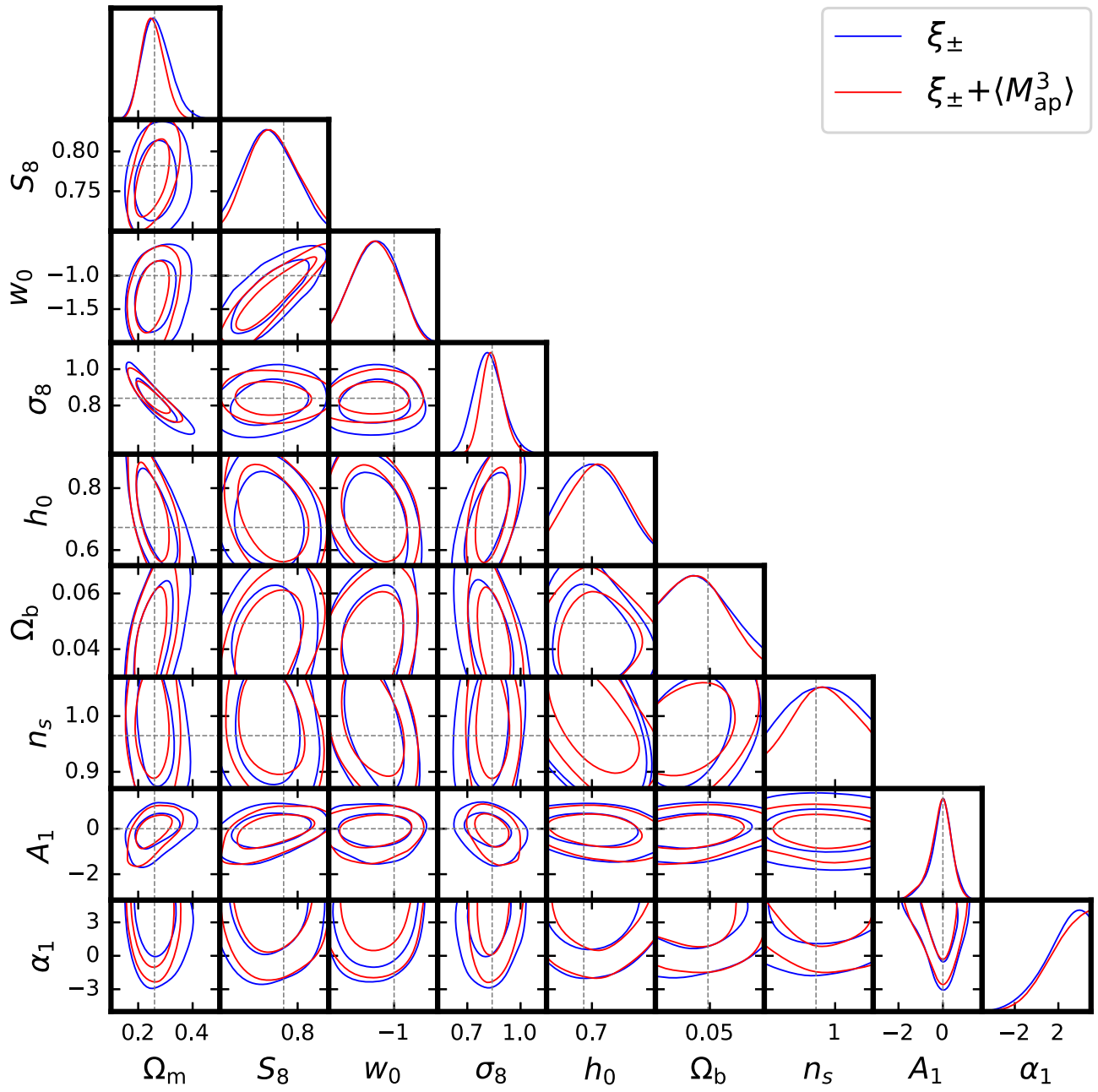

FIG. 15. Parameter constraints on the whole set of cosmological parameters from $\xi_\pm$ and $\langle \mathcal{M}_{\rm ap}^3 \rangle$ using wCDM modeling. The blue contours indicate results from the 2PCF alone, while the red contours indicate results from the 2PCF combined with the mass aperture statistic. The gray dotted lines represent the input values for the parameters. As in the previous figure, we do not indicate any value for α_1 since our input data vector had a zero intrinsic alignment amplitude.

[1] L. Perivolaropoulos and F. Skara, Challenges for ΛCDM: An update, New Astron. Rev. **95**, 101659 (2022).

[2] G. Efstathiou, Challenges to the λCDM cosmology, Phil. Trans. R. Soc. A **383**, 20240022 (2025).

[3] T. Eifler, M. Simet, E. Krause *et al.*, Cosmology with the roman space telescope: Synergies with the rubin observatory legacy survey of space and time, Mon. Not. R. Astron. Soc. **507**, 1514 (2021).

[4] S. Pyne and B. Joachimi, Self-calibration of weak lensing systematic effects using combined two- and three-point statistics, Mon. Not. R. Astron. Soc. **503**, 2300 (2021).

[5] L. F. Secco, S. Samuroff, E. Krause *et al.*, Dark energy survey year 3 results: Cosmology from cosmic shear and robustness to modeling uncertainty, Phys. Rev. D **105**, 023515 (2022).

[6] A. Amon, D. Gruen, M. A. Troxel, N. MacCrann, S. Dodelson *et al.*, Dark energy survey year 3 results: Cosmology from cosmic shear and robustness to data calibration, Phys. Rev. D **105**, 023514 (2022).

[7] M. Asgari, C.-A. Lin, B. Joachimi, B. Giblin, C. Heymans *et al.*, KiDS-1000 cosmology: Cosmic shear constraints and comparison between two point statistics, Astron. Astrophys. **645**, A104 (2021).

[8] R. Dalal, X. Li, A. Nicola, J. Zuntz, M. A. Strauss *et al.*, Hyper suprime-cam year 3 results: Cosmology from cosmic shear power spectra, Phys. Rev. D **108**, 123519 (2023).

[9] X. Li, T. Zhang, S. Sugiyama, R. Dalal, R. Terasawa *et al.*, Hyper suprime-cam year 3 results: Cosmology from cosmic shear two-point correlation functions, Phys. Rev. D **108**, 123518 (2023).

[10] M. Gatti, C. Chang, O. Friedrich, B. Jain, D. Bacon *et al.*, Dark energy survey year 3 results: Cosmology with moments of weak lensing mass maps—validation on simulations, Mon. Not. R. Astron. Soc. **498**, 4060 (2020).

[11] Z. Gong, A. Halder, A. Barreira, S. Seitz, and O. Friedrich, Cosmology from the integrated shear 3-point correlation function: Simulated likelihood analyses with machine-learning emulators, J. Cosmol. Astropart. Phys. 07 (2023) 040.

[12] V. Ajani, Higher order statistics for cosmology: Likelihood development for future surveys like Euclid, Ph.D. thesis, Universite Pierre et Marie Curie (Paris VI), France, 2021.

[13] A. Petri, Z. Haiman, L. Hui, M. May, and J. M. Kratochvil, Cosmology with Minkowski functionals and moments of the weak lensing convergence field, Phys. Rev. D **88**, 123002 (2013).

[14] G. A. Marques, J. Liu, M. Shirasaki, L. Thiele, D. Grandón, K. M. Huffenberger, S. Cheng, J. Harnois-Déraps, K. Osato, and W. R. Coulton, Cosmology from weak lensing peaks and minima with subaru hyper suprime-cam survey first-year data, Mon. Not. R. Astron. Soc. **528**, 4513 (2024).

[15] S. Cheng, Y.-S. Ting, B. Ménard, and J. Bruna, A new approach to observational cosmology using the scattering transform, Mon. Not. R. Astron. Soc. **499**, 5902 (2020).

[16] E. Allys, T. Marchand, J.-F. Cardoso, F. Villaescusa-Navarro, S. Ho, and S. Mallat, New interpretable statistics for large-scale structure analysis and generation, Phys. Rev. D **102**, 103506 (2020).

[17] S. Heydenreich, B. Brück, and J. Harnois-Déraps, Persistent homology in cosmic shear: Constraining parameters with topological data analysis, Astron. Astrophys. **648**, A74 (2021).

[18] D. Anbajagane, C. Chang, A. Banerjee *et al.*, Beyond the 3rd moment: A practical study of using lensing convergence CDFs for cosmology with DES Y3, Mon. Not. R. Astron. Soc. **526**, 5530 (2023).

[19] A. Barthelemy, A. Halder, Z. Gong, and C. Uhlemann, Making the leap I: Modelling the reconstructed lensing convergence PDF from cosmic shear with survey masks and systematics, J. Cosmol. Astropart. Phys. 03 (2024) 060.

[20] J. Fluri, T. Kacprzak, A. Refregier, A. Amara, A. Lucchi, and T. Hofmann, Cosmological constraints from noisy convergence maps through deep learning, Phys. Rev. D **98**, 123518 (2018).

[21] N. Jeffrey, J. Alsing, and F. Lanusse, Likelihood-free inference with neural compression of DES SV weak lensing map statistics, Mon. Not. R. Astron. Soc. **501**, 954 (2020).

[22] D. Ribli, B. A. Pataki, J. M. Zorrilla Matilla, D. Hsu, Z. Haiman, and I. Csabai, Weak lensing cosmology with convolutional neural networks on noisy data, Mon. Not. R. Astron. Soc. **490**, 1843 (2019).

[23] L. F. Secco, M. Jarvis, B. Jain, C. Chang, M. Gatti *et al.*, Dark energy survey year 3 results: Three-point shear correlations and mass aperture moments, Phys. Rev. D **105**, 103537 (2022).

[24] S. Heydenreich, L. Linke, P. Burger, and P. Schneider, A roadmap to cosmological parameter analysis with third-order shear statistics I: Modelling and validation, Astron. Astrophys. **672**, A44 (2023).

[25] P. A. Burger, L. Porth, S. Heydenreich, L. Linke, N. Wielders *et al.*, KiDS-1000 cosmology: Combined second- and third-order shear statistics, Astron. Astrophys. **683**, A103 (2024).

[26] R. C. H. Gomes *et al.*, companion paper, Dark Energy Survey Year 3 results: Cosmological constraints from second- and third-order shear statistics, Phys. Rev. D **112**, 123515 (2025).

[27] M. Kilbinger, Cosmology with cosmic shear observations: A review, Rep. Prog. Phys. **78**, 086901 (2015).

[28] N. Kaiser, Weak gravitational lensing of distant galaxies, Astrophys. J. **388**, 272 (1992).

[29] A. Buchalter, M. Kamionkowski, and A. H. Jaffe, The angular three-point correlation function in the quasi-linear regime, Astrophys. J. **530**, 36 (2000).

[30] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya, and M. Oguri, Revising the Halofit model for the nonlinear matter power spectrum, Astrophys. J. **761**, 152 (2012).

[31] R. Takahashi, T. Nishimichi, T. Namikawa, A. Taruya, I. Kayo, K. Osato, Y. Kobayashi, and M. Shirasaki, Fitting the nonlinear matter bispectrum by the Halofit approach, Astrophys. J. **895**, 113 (2020).

[32] P. Schneider and M. Lombardi, The three-point correlation function of cosmic shear: I. The natural components, Astron. Astrophys. **397**, 809 (2003).

[33] P. Schneider, M. Kilbinger, and M. Lombardi, The three-point correlation function of cosmic shear—II. Relation to the bispectrum of the projected mass density and generalized third-order aperture measures, Astron. Astrophys. **431**, 9 (2005).

[34] S. Sugiyama, R. C. H. Gomes, and M. Jarvis, Fast modeling of the shear three-point correlation function, arXiv:2407.01798.
[35] L. Porth, S. Heydenreich, P. Burger, L. Linke, and P. O. Schneider, A roadmap to cosmological parameter analysis with third-order shear statistics III: Efficient estimation of third-order shear correlation functions and an application to the KiDS-1000 data, Astron. Astrophys. **689**, A227 (2024).
[36] M. Jarvis, G. Bernstein, and B. Jain, The skewness of the aperture mass statistic, Mon. Not. R. Astron. Soc. **352**, 338 (2004).
[37] J. Myles, A. Alarcon, A. Amon *et al.*, Dark energy survey year 3 results: Redshift calibration of the weak lensing source galaxies, Mon. Not. R. Astron. Soc. **505**, 4249 (2021).
[38] N. MacCrann, M. R. Becker, J. McCullough *et al.*, Dark energy survey Y3 results: Blending shear and redshift biases in image simulations, Mon. Not. R. Astron. Soc. **509**, 3371 (2021).
[39] M. Gatti, E. Sheldon, A. Amon, M. Becker, M. Troxel *et al.*, Dark energy survey year 3 results: Weak lensing shape catalogue, Mon. Not. R. Astron. Soc. **504**, 4312 (2021).
[40] S. Bridle and L. King, Dark energy constraints from cosmic shear power spectra: Impact of intrinsic alignments on photometric redshift requirements, New J. Phys. **9**, 444 (2007).
[41] E. Krause, T. F. Eifler, J. Zuntz, O. Friedrich, M. A. Troxel *et al.*, Dark energy survey year 1 results: Multi-probe methodology and simulated likelihood analyses, arXiv:1706.09359.
[42] J. A. Blazek, N. MacCrann, M. A. Troxel, and X. Fang, Beyond linear galaxy alignments, Phys. Rev. D **100**, 103506 (2019).
[43] T. Abbott, M. Aguena, A. Alarcon *et al.*, DES Y3 + KiDS-1000: Consistent cosmology combining cosmic shear surveys, Open J. Astrophys. **6** (2023).
[44] S. Pyne, A. Tenneti, and B. Joachimi, Three-point intrinsic alignments of dark matter halos in the IllustrisTNG simulation, Mon. Not. R. Astron. Soc. **516**, 1829 (2022).
[45] A. Spurio Mancini, D. Piras, J. Alsing, B. Joachimi, and M. P. Hobson, CosmoPower: Emulating cosmological power spectra for accelerated Bayesian inference from next-generation surveys, Mon. Not. R. Astron. Soc. **511**, 1771 (2022).
[46] T. Kacprzak, J. Fluri, A. Schneider, A. Refregier, and J. o. Stadel, CosmoGridV1: A simulated *w*CDM theory prediction for map-level cosmological inference, J. Cosmol. Astropart. Phys. 02 (2023) 050.
[47] D. Potter, J. Stadel, and R. Teyssier, PKDGRAV3: Beyond trillion particle cosmological simulations for the next era of galaxy surveys, Comput. Astrophys. Cosmol. **4**, 2 (2017).
[48] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere, Astrophys. J. **622**, 759 (2005).
[49] T. Abbott, M. Aguena, A. Alarcon *et al.*, Dark energy survey year 3 results: Constraints on extensions to ΛCDM with weak lensing and galaxy clustering, Phys. Rev. D **107**, 083504 (2023).
[50] E. Sellentin and A. F. Heavens, Parameter inference with estimated covariance matrices, Mon. Not. R. Astron. Soc. **456**, L132 (2016).
[51] J. Hartlap, P. Simon, and P. Schneider, Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix, Astron. Astrophys. **464**, 399 (2006).
[52] S. Dodelson and M. D. Schneider, The effect of covariance estimator error on cosmological parameter constraints, Phys. Rev. D **88**, 063537 (2013).
[53] W. J. Percival, O. Friedrich, E. Sellentin, and A. Heavens, Matching Bayesian and frequentist coverage probabilities when using an approximate data covariance matrix, Mon. Not. R. Astron. Soc. **510**, 3207 (2021).
[54] J. Zuntz, M. Paterno, E. Jennings, D. Rudd, A. Manzotti, S. Dodelson, S. Bridle, S. Sehrish, and J. Kowalkowski, CosmoSIS: Modular cosmological parameter estimation, Astron. Comput. **12**, 45 (2015).
[55] F. Feroz, M. P. Hobson, and M. Bridges, MultiNest: An efficient and robust Bayesian inference tool for cosmology and particle physics, Mon. Not. R. Astron. Soc. **398**, 1601 (2009).
[56] P. Lemos, N. Weaverdyck, R. P. Rollins, J. Muir, A. Ferté *et al.*, Robust sampling for weak lensing and clustering analyses with the Dark Energy Survey, Mon. Not. R. Astron. Soc. **521**, 1184 (2023).
[57] A. Heavens, E. Sellentin, D. d. Mijolla, and A. Vianello, Massive data compression for parameter-dependent covariance matrices, Mon. Not. R. Astron. Soc. **472**, 4244 (2017).
[58] A. Heavens, R. Jimenez, and O. Lahav, Massive lossless data compression and multiple parameter estimation from galaxy spectra, Mon. Not. R. Astron. Soc. **317**, 965 (2000).
[59] Baryonic physics can also cause a bump on the matter bispectrum at an intermediate scale as well as suppression according to Takahashi *et al.* [31].
[60] J. Schaye, R. A. Crain, R. G. Bower *et al.*, The EAGLE project: Simulating the evolution and assembly of galaxies and their environments, Mon. Not. R. Astron. Soc. **446**, 521 (2014).
[61] R. Davé, D. Anglés-Alcázar, D. Narayanan, Q. Li, M. H Rafieferantsoa, and S. Appleby, SIMBA: Cosmological simulations with black hole growth and feedback, Mon. Not. R. Astron. Soc. **486**, 2827 (2019).
[62] J. Schaye, C. D. Vecchia, C. M. Booth, R. P. C. Wiersma, T. Theuns, M. R. Haas, S. Bertone, A. R. Duffy, I. G. McCarthy, and F. van de Voort, The physics driving the cosmic star formation history, Mon. Not. R. Astron. Soc. **402**, 1536 (2010).
[63] M. Gatti, N. Jeffrey, L. Whiteway, J. Williamson, B. Jain *et al.*, Dark energy survey year 3 results: Simulation-based cosmological inference with wavelet harmonics, scattering transforms, and moments of weak lensing mass maps I: Validation on simulations, Phys. Rev. D **109**, 063534 (2024).
[64] F. Bernardeau, L. Van Waerbeke, and Y. Mellier, Weak lensing statistics as a probe of Omega and power spectrum, Astron. Astrophys. **322**, 1 (1997), https://adsabs.harvard.edu/pdf/1997A%26A...322....1B.

[65] This is the conventional S_8 definition, while we found that $\sigma_8(\Omega_m/0.3)^{0.53}$ is the parameter combination that the two-point alone analysis in this paper constraints most tightly.
[66] G. Cabass, M. Simonović, and M. Zaldarriaga, Cosmological information in perturbative forward modeling, Phys. Rev. D **109**, 043526 (2024).
[67] N.-M. Nguyen, F. Schmidt, B. Tucci, M. Reinecke, and A. Kostić, How much information can be extracted from galaxy clustering at the field level?, Phys. Rev. Lett. **133**, 221006 (2024).
[68] This contrasts with the 30%–45% improvement found by Gong *et al.* [11] in their analysis of the integrated shear three-point function. One major difference between the studies is in the lower bounds of the w_0 prior. While we impose that $w_0 > -2$, Gong *et al.* [11] use $w_0 > -3.33$. This removes a region of parameter space which is well constrained by the combination of second- and third-order statistics.
[69] A. Lewis, CAMB notes, https://cosmologist.info/notes/CAMB.pdf.