Cosmology from HSC Y1 weak lensing data with combined higher-order statistics and simulation-based inference

Camila P. Novaes,[1,2,3,*] Leander Thiele,[2,3,†] Joaquin Armijo,[2,3] Sihao Cheng,[4,5] Jessica A. Cowell,[2,3,6] Gabriela A. Marques,[7,8] Elisa G. M. Ferreira,[2,3] Masato Shirasaki,[9,10] Ken Osato,[11,12,2] and Jia Liu[2,3]

[1]*Instituto Nacional de Pesquisas Espaciais, Avenida dos Astronautas 1758, Jardim da Granja, São José dos Campos, São Paulo, Brazil*
[2]*Kavli IPMU (WPI), UTIAS, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8583, Japan*
[3]*Center for Data-Driven Discovery, Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan*
[4]*Institute for Advanced Study, 1 Einstein Dr., Princeton, New Jersey 08540, USA*
[5]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada*
[6]*Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, United Kingdom*
[7]*Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, Illinois 60510, USA*
[8]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, Illinois 60637, USA*
[9]*National Astronomical Observatory of Japan (NAOJ), National Institutes of Natural Sciences, Osawa, Mitaka, Tokyo 181-8588, Japan*
[10]*The Institute of Statistical Mathematics, Tachikawa, Tokyo 190-8562, Japan*
[11]*Center for Frontier Science, Chiba University, 1-33 Yayoicho, Inage, Chiba 263-8522, Japan*
[12]*Department of Physics, Graduate School of Science, Chiba University, 1-33 Yayoicho, Inage, Chiba 263-8522, Japan*

(Received 12 September 2024; accepted 13 March 2025; published 7 April 2025)

We present cosmological constraints from weak lensing with the Subaru Hyper Suprime-Cam (HSC) first-year data, using a simulation-based inference (SBI) method. We explore the performance of a set of higher-order statistics (HOS) including the Minkowski functionals, counts of peaks and minima, and the probability distribution function and compare them to the traditional two-point statistics. The HOS, also known as non-Gaussian statistics, can extract additional non-Gaussian information that is inaccessible to the two-point statistics. We use a neural network to compress the summary statistics, followed by an SBI approach to infer the posterior distribution of the cosmological parameters. We apply cuts on angular scales and redshift bins to mitigate the impact of systematic effects. Combining two-point and non-Gaussian statistics, we obtain $S_8 \equiv \sigma_8\sqrt{\Omega_m/0.3} = 0.804^{+0.041}_{-0.040}$ and $\Omega_m = 0.344^{+0.083}_{-0.090}$, similar to that from non-Gaussian statistics alone. These results are consistent with previous HSC analyses and Planck 2018 cosmology. Our constraints from non-Gaussian statistics are ∼25% tighter in S_8 than two-point statistics, where the main improvement lies in Ω_m, with ∼40% tighter error bar compared to using the angular power spectrum alone ($S_8 = 0.766^{+0.054}_{-0.056}$ and $\Omega_m = 0.365^{+0.148}_{-0.141}$). We find that, among the non-Gaussian statistics we studied, the Minkowski functionals are the primary driver for this improvement. Our analyses confirm the SBI as a powerful approach for cosmological constraints, avoiding any assumptions about the functional form of the data's likelihood.

DOI: 10.1103/PhysRevD.111.083510

I. INTRODUCTION

There is no doubt about the importance of surveying the large-scale structure of the Universe and mapping the distribution of baryonic matter, as measurements of its clustering serve as a powerful cosmological probe. The matter clustering gives rise to the weak lensing signal, which can be measured as a distortion of galaxy shape and provide a direct measurement of the total matter density contrast. In fact, valuable cosmological constraints have been possible through analyses of lensing shear catalogs provided by stage-III[1] surveys, such as the Kilo-Degree

*Contact author: camila.novaes@inpe.br
†Contact author: leander.thiele@ipmu.jp

[1]Definition introduced by the Dark Energy Task Force report [1].

 © 2025 American Physical Society

Survey (KiDS) [2], the Dark Energy Survey (DES) [3] and the Subaru Hyper Suprime-Cam (HSC) [4]. Still, with the next generation of experiments, including Euclid [5,6], and Vera Rubin Observatory Legacy Survey of Space and Time (LSST) [7], we expect to be able to map the distribution of matter deeper in the Universe, with higher source density and better shape measurements. This would allow us to investigate tensions in the current cosmology scenario, in particular the apparent discrepancy identified between the growth-of-structure parameter, S_8, from CMB measurements, as well as CMB lensing, and weak lensing data [8] (for studies using different probes, see, e.g., Refs. [9–13]).

Unlike the primary CMB temperature fluctuations, a nearly Gaussian field with imprints of the primordial density perturbations, weak lensing fields carry non-Gaussian (NG) information introduced by the nonlinear evolution of cosmic structures. As a consequence, the traditional two-point statistics are insufficient to capture all the cosmological information contained in the data. To maximize the amount of information extracted from such datasets, it is crucial to make use of NG statistics, also known as higher-order statistics (HOS). With such motivation, several studies have been employing NG statistics for weak lensing probes [14–19]. Several NG statistics have been considered in the context of weak lensing, such as moments [20,21], density-split statistic [22–24], and three-point statistics [25,26]. In particular, HSC first-year (Y1) weak lensing data have been explored using peaks and minimum count distributions (PM) [27], probability distribution function (PDF) [28], scattering transform [29], and Minkowski functionals (MFs) [30], while the first three have also been studied in the context of baryonic feedback [31].

In this work, we use the MFs, PM, and PDF of the Subaru Hyper Suprime-Cam (HSC) first-year (HSC Y1) data, and their combination with the angular power spectrum (PS), to constrain the cosmological parameters Ω_m, the total matter density today, and $S_8 = \sigma_8\sqrt{\Omega_m/0.3}$, written as a function of the linear matter fluctuation within $8h^{-1}$ Mpc radius, σ_8. Since NG statistics, in general, have no analytical expression or a known closed-form theoretical model, it is necessary to rely on simulation-based models to describe a given summary statistic. By forward modeling the statistics allows, for example, to build emulators for cosmological constraints, as implemented by previous analyses using NG statistics of the HSC Y1 data [27–31]. However, this type of analysis requires making approximations, as the assumption of a Gaussian likelihood for the data. Here, instead, we forward model the summary statistics of HSC Y1 and perform a simulation-based inference (SBI) [32], also known as likelihood-free or implicit-likelihood inference. The SBI approach, then, might be considered as an optimal solution given that it does not require a functional form for the likelihood. A similar approach is presented in Refs. [15,33–37], which uses different Gaussian and non-Gaussian statistics to extract the cosmological information, while Refs. [37–39] rely on convolutional neural networks (CNN) for feature extraction (see Ref. [40] for an application to HSC Y1).

The pipeline implemented for our analyses consists of two main steps:

(1) a neural network (NN) compresses the summary statistics (neural compression), or, in other words, it maps between summary statistics and cosmological parameters (Ω_m, S_8), providing the predicted (compressed) values for these parameters (following Ref. [41]); and
(2) a neural density estimator (NDE), trained over the compressed data (from simulations), infers the posterior distribution of Ω_m and S_8 parameters when evaluated on the measured data (from simulations or observation) after the same compression.

We note that data vectors comprising a summary statistic, or a combination of statistics, can be large enough to make the SBI process computationally prohibitively expensive. Also, SBI does not handle large data vectors well with a finite number of simulations [42]. This motivates the first step of our pipeline, including a massive compression of the summary statistics. Our methodology brings a novel approach in comparison to previous weak lensing NG analyses, showing that the combination of several summary statistics is promising and feasible in terms of SBI analyses. In particular, we demonstrate the feasibility of the SBI approach with a limited number of simulations, a common limitation in weak lensing analyses, given their computational cost.

The paper is organized as follows. In Sec. II, we briefly describe the HSC Y1 data and simulations. Section III presents the summary statistics employed in our analyses, while Sec. IV presents our methodology for cosmological analyses, describing the NN algorithm employed for data compression and the likelihood-free approach for parameter inference. In Sec. V, we show the robustness of our results against systematic effects. Our results are presented in Sec. VI, and we conclude in Sec. VII.

II. DATA AND SIMULATIONS

In this section, we briefly describe the HSC Y1 lensing convergence data, as well as the simulations employed in our analysis. We refer the reader to Ref. [27] for additional information.

A. HSC Y1 data

As a precursor to stage IV surveys, HSC is a deep, large-scale lensing survey with a high density of source galaxies. Here, we utilize the shear catalog from the first public data release, HSC Y1 [43], [2] selecting galaxies with reliable

[2]Publicly available at https://hsc-release.mtk.nao.ac.jp/doc/index.php/s16a-shape-catalog-pdr2/.

shape measurements covering (after aplying the masks) 136.9 deg^2 of the sky in six disjoint patches and spanning a redshift range from 0.3 to 1.5. In our analysis, we define the redshift bins based on photometric redshifts determined using the MLZ code [44]. We split the source galaxies into four tomographic redshift bins, namely, $0.3 < z_1 < 0.6$, $0.6 < z_2 < 0.9$, $0.9 < z_3 < 1.2$, and $1.2 < z_4 < 1.5$, from which we use only the first three, given possible contamination by unknown systematics in the highest z-bin [28,45]. After creating pixelized maps from the shear catalog for each redshift bin and patch, they are smoothed with a Gaussian kernel, defined as $W_G = [1/(\pi\theta_s^2)]\exp(-\theta^2/\theta_s^2)$. Our main analysis employs the smoothing scale of $\theta_s = 4$ arcmin (as discussed in Sec. V), although we also evaluate our constraining power when adding $\theta_s = 8$ arcmin. We then use the Kaiser-Squires (KS) inversion method [46] to convert the shear fields into the convergence maps. To avoid artifacts introduced in the KS process by missing data due to the survey mask, the masked pixels of the shear maps are inpainted before the inversion (see Ref. [27] for details).

B. Simulations

To forward model the summary statistics, we use the suite of N-body simulations introduced in Ref. [47], produced for 100 different cosmologies sampled in the ranges $0.1 < \Omega_m < 0.7$ and $0.23 < S_8 < 1.1$, while fixing the Hubble parameter, baryon density, dark energy equation-of-state parameter, and spectral index, respectively, at $h = 0.6727$, $\Omega_b = 0.049$, $w = -1$, and $n_s = 0.9645$. These simulations are based on the parallel tree-particle mesh code Gadget-2, with 512^3 particles. A typical value of the thickness of the mass sheet is about 150, 200, and $300h^{-1}$ Mpc at $z < 0.5$, $0.5 < z < 1$, and $1 < z < 2$, respectively. To create convergence maps through the ray-tracing technique, we adopt a flat-sky approximation and the multiple lens-plane algorithm [48,49]. For each pair of (Ω_m, S_8) values, a total of 50 realizations are constructed by randomly selecting observer positions in the simulation box, providing a set of quasi-independent realizations from only one simulation. This set of simulations is hereafter called *cosmology-varied*.

We also use a second set of simulations comprised by 2268 realizations generated from 108 quasi-independent full-sky simulations [50] with cosmology fixed at the Wilkinson Microwave Anisotropy Probe nine-year data best-fit [51]: $\Omega_b = 0.046$, $\Omega_m = 0.279$, $\Omega_\Lambda = 0.721$, $h = 0.7$, $S_8 = 0.79$, and $n_s = 0.97$. We refer to this set of simulations at a fixed cosmology as our *fiducial* set.

Both sets of simulations are tailored to HSC Y1 data, accounting for realistic aspects such as the survey geometry, multiplicative biases, redshift uncertainty, spatial inhomogeneities of source galaxies, shape noise levels, among other observational characteristics, following Refs. [27,52]. We reinforce that all simulations account for the same survey mask as the data, to avoid any inconsistency between them (a comprehensive description of the masking procedure is provided in Ref. [27]). The simulations also account for the same smoothing scales applied to the data.

III. SUMMARY STATISTICS

Our analyses assess the constraining power of one Gaussian statistic, the PS, and three NG statistics, the MFs, PM, and PDF. Below, we summarize the main details of their estimates from the convergence maps.

Angular power spectrum. We estimate the PS using the pseudo-C_ℓ method implemented in the NaMaster public code [53],[3] accounting for the geometry of each sky patch, that is, deconvolving the survey mask. We estimate the pseudo-C_ℓ, $C_\ell^{\kappa\kappa}$, in 14 logarithmically spaced multipole bands in the total range of $81 < \ell < 6500$. Out of this range, we use $300 < \ell < 1000$, i.e., four multipole bands in our analyses. Following Refs. [27,28], we consider the lower limit as established by Ref. [54], according to assessments of unmodeled systematic errors on large scales. The maximum multipole is defined considering possible bias introduced by baryonic effects (see Ref. [31] for detailed analyses on the impact of these effects on Gaussian and non-Gaussian statistics). For a given redshift bin, a single $C_\ell^{\kappa\kappa}$ data vector is obtained by a weighted average over the six sky patches, $C_\ell^{\kappa\kappa} = \sum_i w_i C_\ell^{\kappa\kappa}|_i$, where w_i is proportional to the inverse of the number of galaxies in the ith sky patch.

Minkowski functionals. Widely used in CMB [55–59] and large scale structure analyses [60–64], the Minkowski functionals (MFs) [65,66] carry morphological information of structures, requiring $N+1$ functionals to completely characterize an N-dimensional field. Here, we use three MFs, the area (V_0), perimeter (V_1), and genus (V_2), to access the NG features in the 2D lensing convergence maps, $\nu = \kappa/\sigma_0$, normalized by the average of the standard deviation σ_0 of the κ fields composing the 2268 fiducial simulations. The three quantities are given by Refs. [66,67],

$$V_0(\nu_t) = \frac{1}{4\pi}\int_\Sigma d\Omega, \tag{1}$$

$$V_1(\nu_t) = \frac{1}{4\pi}\frac{1}{4}\int_{\partial\Sigma} dl, \tag{2}$$

$$V_2(\nu_t) = \frac{1}{4\pi}\frac{1}{2\pi}\int_{\partial\Sigma} \zeta dl, \tag{3}$$

and calculated for excursion sets $\Sigma \equiv \{\kappa > \nu_t\sigma_0\}$, i.e., sets of connected pixels whose values exceed a ν_t threshold, with boundaries $\partial\Sigma \equiv \{\kappa = \nu_t\sigma_0\}$. The quantities $d\Omega$ and dl are the solid angle and line elements, respectively, and ζ is the geodesic curvature. Here, we estimate the MFs for

[3] https://namaster.readthedocs.io/en/latest/.

19 ν_t thresholds linearly spaced in $[-4, 4]$. More information on MFs and their application to lensing data can be found in, e.g., Refs. [19,21,52,68–70]. The MFs have also been used for HSC Y1 analyses by Ref. [30]. Note that, for each redshift bin, we obtain the total MFs by summing the results from each of the six HSC sky patches. The same procedure is employed for the other NG statistics.

Peaks and minima count distributions. Peaks counts have been widely explored for weak lensing analyses [71–73] and, more recently, the minimum counts have also been shown to be efficient in cosmological constraints from weak lensing data [74]. A peak is defined as a pixel whose value $\nu \equiv \kappa/\sigma_0$ exceeds the values from the eight neighboring pixels, while a minimum is the opposite, exhibiting a lower value compared to the eight neighboring pixels. We count the peaks and minima (PM) within bins defined by the same 19 ν_t values as the MFs, following previous applications of these statistics to HSC Y1 data [27,31].

Probability distribution function. Sensitive to NG signatures of a field, the probability distribution function (PDF) is also a promising summary statistic for cosmological constraints [75], as demonstrated by Ref. [28] from analysis of the HSC Y1 data. For details on weak lensing convergence PDF, see [76–79]. Here, similarly to MFs and PM, we estimate this one-point statistic as the histogram of the pixels in the normalized convergence maps, $\nu \equiv \kappa/\sigma_0$, for the same 19 ν_t bins as previously defined in this section. We emphasize that, as for the other NG statistics, we normalize the convergence fields (data and simulations) by the same σ_0 factor, given by the average standard deviation from fiducial simulations, before estimating the PDF. This is done to allow a more direct comparison between the results from different NG statistics. Note that this is different from the PDF analysis presented in Ref. [28], where each κ map was normalized by its own standard deviation, removing most of the two-point information, focusing in the NG information.

IV. SIMULATION-BASED INFERENCE

Our purpose for this paper is to infer cosmology from HSC Y1 data following a procedure fully based on simulations. We rely on the forward modelling of both Gaussian and non-Gaussian statistics without assuming a functional form for the likelihood (likelihood-free inference). The combination of the summary statistics considered here, estimated from three redshift bins, results in data vectors of length 354 (see discussion in the next subsection). The performance of a neural density estimator to infer the posterior distribution directly from such a large data vector is limited by the number of simulations, besides being highly computationally expensive. Therefore, our SBI pipeline includes a first step of data compression. This is intended to reduce the dimensionality of the data vector while keeping as much information as possible. The posterior estimation is then performed over the compressed data.

A. Neural network and data compression

The data compression can be performed following different approaches [37,80,81]. The method adopted in this paper uses an NN[4] (neural compression; examples can be found in Refs. [15,37,83]). In this case, the NN works as a function $\mathcal{F}$ applied to the N-dimensional summary statistics X and returns a compressed data vector,

$$x = \mathcal{F}(X), \tag{4}$$

which keeps the maximum information about the cosmological parameters of interest, θ [83]. The dimension of the compressed data can be chosen to match that of θ. We note that an alternative way of looking at this problem is to take the NN as a way of mapping the summary statistics, a set of features, to the cosmological parameters, taken as targets for the loss function, so that x would be the prediction of the cosmological parameters [41,57,84,85].

We summarize below our implementation of NNs for neural compression. We closely follow the methodology for the training process and the choice of hyperparameters as done in Ref. [41], to which we refer the readers for further details.

Training and test. We use the summary statistics estimated from the cosmology-varied set of simulations for training and testing the NN algorithm: 50 realizations for each of the 100 cosmologies, or pairs $\{\Omega_m, S_8\}$, sampling the parameter space. The summary statistics, X, feed the NNs and the corresponding cosmological parameters are the outputs, x (values to be predicted). Since we have a restricted number of cosmologies, we maximize the information fed into the NN during training by following the idea of "leave-one-out" [86]. This means that we train an NN 100 times, keeping fixed the hyperparameters, each time leaving one cosmology out. We use the corresponding set of 50 realizations as our test set. In this way, we end up with 100 NN trained models, i.e., $\mathcal{F}_i$ for $i = 1, \ldots, 100$, and each of them is tested over the summary statistics from the 50 realizations of the excluded cosmology, i, ensuring that none of the models is tested over cosmologies used for training. The 99 cosmologies used for each training are split so that 20% of them is reserved for validation. Although, for simplicity, we refer to the summary statistics from the 99 cosmologies as the training set, a cross-validation procedure is always employed (see Ref. [41]).

The training set is defined as $\mathcal{T}(X, \theta) = \mathcal{T}(X_j, \theta_j)$, allowing the NN to learn the relation between a given summary statistic (or combination of them), X, and the true cosmological parameters, θ. Then, the X_j data vector for the

[4]For a comparison between different methods, see Ref. [82], where the authors investigate several aspects of their application to weak lensing data and evaluate the best techniques in the context of SBI.

jth realization is given by

$$X_j = \{(\mathtt{S})^z\}_j, \tag{5}$$

where S is composed by one or more summary statistics and z runs over the redshift bins, z_1, z_2, and z_3. For example, combining all four statistics, we have a data vector $(\mathtt{S}) = (C_\ell^{\kappa\kappa}, \mathrm{MFs}_\nu, \mathrm{PM}_\nu, \mathrm{PDF}_\nu)$ of length 118 for an individual redshift z, and length 354 when combining all three redshift bins. The training is performed for each cosmological parameter at a time, so that an NN learns the relation between X and $\theta = \{\Omega_m\}$, and X and $\theta = \{S_8\}$, separately.

Architecture. We use a fully connected NN optimizing the set of hyperparameters for each individual test performed (e.g., different combinations of summary statistics or redshift bins). For this, we use the OPTUNA package [87], which allows us to automatically sample the hyperparameter space, fixing the maximum number of trials (the number of tested architectures) to 600, although the optimized set is generally chosen between the 350 and 500 trials. We use the mean square error (MSE) as our loss function,

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \theta_i)^2, \tag{6}$$

where N denotes the number of simulations in the validation set and x_i, θ_i the prediction and true values of a cosmological parameter (Ω_m or S_8). Note that we use all the 100 cosmologies to search for the optimized set of hyperparameters. After that, we fix the hyperparameters[5] and train the NN 100 times, each of them excluding one of the cosmologies, obtaining 100 $\mathcal{F}_i$ models.

Compressed (predicted) data. Each NN model $\mathcal{F}_i$ obtained is then applied to the summary statistics estimated from the 50 realizations of ith cosmology excluded from the training, The average and standard deviation from the 50 predicted (compressed) values, x, for $i = 1, \ldots, 100$, are shown in Fig. 1 as a function of the true values, for both cosmological parameters. We summarize the results for three different combinations of summary statistics: all of them together, $(\mathtt{S}) = (C_\ell^{\kappa\kappa}, \mathrm{MFs}_\nu, \mathrm{PM}_\nu, \mathrm{PDF}_\nu)$, only the NG statistics, $(\mathtt{S}) = (\mathrm{MFs}_\nu, \mathrm{PM}_\nu, \mathrm{PDF}_\nu)$, and the PS alone $(\mathtt{S}) = (C_\ell^{\kappa\kappa})$. All of our tests combine the three redshift bins unless stated otherwise. Following the same color code, we show the root-mean-square error [RMSE $= \sqrt{\mathrm{MSE}}$, with MSE given by Eq. (6)], quantifying the error (accuracy) of the

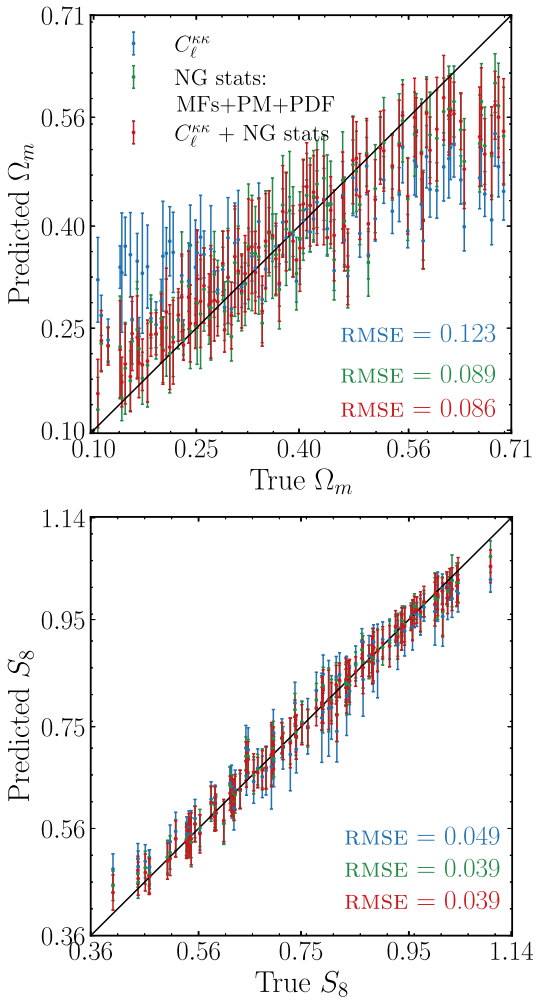


FIG. 1. Scatter plots of the average of the 50 predicted (compressed) values of Ω_m (top) and S_8 (bottom) as a function of the true values of the cosmological parameters. The error bars correspond to the standard deviation of the respective set of 50 values. The different colors represent the different sets of summary statistics used for neural compression: the PS alone (blue), the combination of all three NG statistics (green), and the combination of the PS and NG statistics (red). Following the same color scale, we show in the bottom right of each panel the RMSE value for each case, estimated from the 5000 predictions (100 cosmologies × 50 realizations) and the respective true values. The black diagonal line represents the identity $x = \theta$.

NN predictions. These values show that combining Gaussian and NG statistics allows Ω_m and S_8 parameter predictions to improve by 30% and 21%, respectively, when compared to using PS only. The tighter scatter of S_8 predictions, compared to Ω_m, indicates that the data are more informative about S_8 parameter relative to the prior range. The predictions for Ω_m are poorer than for S_8, as also found from a similar analysis of DES data, as discussed in Ref. [88]. In other words, the deviation from the identity, represented by the black solid lines in Fig. 1, happens when the features (summary statistics) have not enough information to allow

[5]By fixing the optimizer to `Adam` with $\beta_1, \beta_2 = 0.9, 0.999$, OPTUNA searches for the main hyperparameters, with ranges/options for each of them as follows: (1) number of neurons in the hidden layers: [1, 500]; (2) number of layers: [1, 3]; (3) activation function: `ReLu`, `tanh`; (4) learning rate: [10^{-4},10^{-2}]; (5) batch size: [50, 500]; and (6) number of epochs: [50, 500]. The 600 trials for one individual test usually take no more than 24 hours in a single Intel(R) Xeon(R) Gold 6230R CPU @ 2.10 GHz.

the NN to efficiently map them in terms of the targets (cosmological parameters). As a common behavior, the NN tends to predict values close to the mean of the parameters range, as doing so artificially reduces the loss function [41,85,89]. Additionally, by the central limit theorem, one would expect the likelihood of the data to be Gaussianized by the compression process. Indeed, we find the compressed data, in particular for the S_8 parameter, to be well approximated by a Gaussian distribution (see Appendix B).

Below, we discuss how the neural compressed data are employed to infer the posterior distribution for each cosmological parameter.

B. Parameter inference

The SBI approach, rather than making assumptions about the form of the likelihood, relies on simulations to implicitly reconstruct (learn) the likelihood $p(x|\theta)$ of the data [32]. Here, we use a neural posterior estimation [90,91] method implemented in the SBI package[6] [92] to train an NDE, using a given compressed data x, to learning the posterior distributions of the parameters of interest, θ. As our NDE, we chose the Gaussian mixture density network[7] (MDN; [94]) architecture, trained to learn a density $q_\phi(\theta|x)$ that approximates the underlying posterior distribution, $p(\theta|x)$, such that $q_\phi(\theta|x) \approx p(\theta|x)$.

Again, the training is performed 100 times, each of them excluding the 50 (x,θ) pairs from one of the cosmologies; the 99 cosmologies employed are split with an 80/20 ratio for training and validation. The training is performed by varying the neural network parameters, ϕ, and maximizing the log-likelihood, $\mathcal{L} = \sum_{i\in T} \log q_\phi(\theta_i|x_i)$, with i running over the training set. In other words, it minimizes the Kullback-Leibler divergence between q and p [95]. The main hyperparameters of the MDN architecture and the training process are chosen through experimentation. We verified that most of these values do not strongly depend on the summary statistics, so that, for all our inferences, we use the `Adam` optimizer, a number of hidden units of the NN of 20 and batch size of 50 (keep other parameters fixed at the default definitions of SBI). The learning rate, on the other hand, varies from 10^{-3} to 10^{-2} depending on the combination of summary statistics. Our choice of the learning rate also relies on tests for posterior validation, as we discuss below in this section. To help prevent overfitting, we use early stopping, interrupting the training after the validation log-likelihood does not increase after 20 epochs.

Since the randomness of the initialization of an NDE can lead to different posterior approximations, for the ith cosmology removed from the training+validation set, we perform the training 20 times and, among the 20 q^i_ϕ estimators, we choose the one leading to the largest validation log-likelihood. This way, at the end we have a set of 100 q^i_ϕ models, one for each cosmology excluded. We note that, when performing parameter inference, for a given combination of summary statistics, over the fiducial set of simulations and over the HSC Y1 data, we choose the 5 q^i_ϕ models with greater log-likelihood. We sample the posterior distribution inferred by each of them, generating 10,000 samples each, which are finally combined in one chain[8] of length 50,000 used to derive the parameter constraint.

Individual versus joint posterior estimation. Our analysis has two cosmological parameters of interest, Ω_m and S_8, for which we want to apply our inference pipeline to obtain the best approximated posterior and final parameters constraints, as previously discussed. There are two possible ways of training our NDE for this purpose: *1D case*: estimating $q^{i,p}_\phi$ for one parameter, $p = \Omega_m$ and $p = S_8$, at a time (marginalizing over the other paramenter), such that both x and θ are one-dimensional data vectors; and *2D case*: estimating q^i_ϕ to approximate the joint posterior of the parameters of interest, with x and θ being two-dimensional. We evaluated both scenarios. Given that 1D and 2D cases provide results in very good agreement, and the former leads to slightly tighter constraints from the fiducial set of simulations, we employ the 1D case for our baseline analysis (see Sec. VI B for more details). All tests and results presented here consider the 1D case unless stated otherwise.

Posterior validation. Before applying our method to the data, it is important to ensure that the posterior inference is being performed accurately and robustly. To validate our posteriors, we use the simulation-based calibration, as discussed in Ref. [96], and construct the rank statistics for each cosmological parameters of interest. Employing a given combination of summary statistics, we define the compressed data, x^n_i, from each of the 50 realizations at a time $(n = 1, \ldots, 50)$ as our "observation." We sample the posterior distribution for each realization of the ith cosmology using the respective q^i_ϕ estimator, for $i = 1, \ldots, 100$. We draw 10,000 samples for each x^n_i "observation" and use them to calculate the rank. The rank is computed as the fractional number of samples whose values fall below the true cosmological parameter, θ_i. We show in Fig. 2 the rank plot obtained for different combinations of summary statistics. As we mention before, the learning rate of the MDN architecture are chosen through experimentation; here we rely on the uniformity of the rank plots to define it in each case. For all tests presented, we find approximately uniform rank plots, with very few spikes exceeding the 2σ intervals, although those for the S_8 parameter are more close to flat.

[6] https://github.com/mackelab/sbi.

[7] We also tested the masked autoregressive flow (MAF; [93]) architecture as the NDE and the performance was similar.

[8] We verified that averaging over the median values of each posterior of length 10,000, as well as over their 1σ range, provide very similar result as taking the median and 1σ from the combined posterior with 50,000 samples.

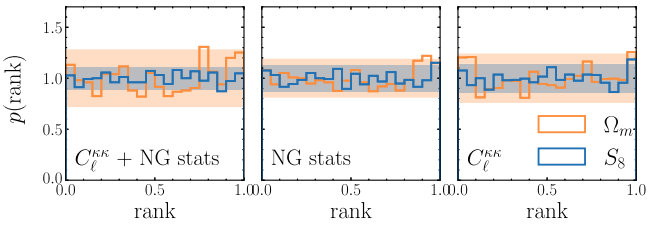


FIG. 2. Rank distributions constructed using each of the 5000 (50 realizations × 100 cosmologies) realizations of the cosmology-varied set as one "observation." The histogram is constructed over the 5000 rank values, each of them calculated from 10,000 samples drawn for each realization using the corresponding q_ϕ^i model. Each panel shows results for Ω_m (orange) and S_8 (blue) using different (combinations of) summary statistics. The colored bands show the respective 2σ intervals. The rank distribution is expected to be uniform for an accurate (valid) posterior estimate. See text for details.

Evaluating the shape of the rank distribution and how it deviates from uniform can help diagnose why a calculated posterior is incorrect [96], for example due to an insufficient number of simulations or even a bias. Rank distributions with ∪-like shape indicate that the posteriors are overconfident, meaning that the calculated posteriors are narrower than the true posterior. The ∩-like shape, on the other hand, indicates that the inferred posteriors are broader than the true one; i.e., they are underconfident and could be improved by more efficient training.

V. SYSTEMATICS

We assess the impact of different systematic effects not accounted for in our simulations, quantifying potential bias introduced by each of them, besides evaluating the accuracy of our SBI pipeline when applied outside the training set. For this, we perform cosmological inference of fiducial realizations accounting for observational and astrophysical systematics as described below (detailed discussions on the simulations can be found in Ref. [27]):

(i) Photometric redshifts: We employ two alternative methods, namely, `Mizuki` and `Frankenz` [44], to estimate the redshift, constructing, using each of them, a subset of 100 fiducial realizations.

(ii) Multiplicative bias: Considering the (percent level) uncertainty expected for the multiplicative bias employed for image calibration [97], we introduce a miscalibration of $\Delta m = \pm 0.01$ when generating a subset of 100 fiducial realizations.

(iii) Baryonic feedback: We use the hydrodynamic simulation IllustrisTNG, with and without including baryons, to model the impact of baryonic feedback on the summary statistics. We include its effect directly on the data vectors of the fiducial realizations by rescaling them by the ratio between the summary statistics calculated from the hydrodynamic and gravity-only simulations.

(iv) Intrinsic alignment: The nonlinear tidal alignment model (NLA) quantify the strength of this effect by the coupling parameter $A_{\rm IA}$ [98]. Given the alignment amplitude estimated for HSC data [99,100], we choose two different values for this parameter ($\pm 1\sigma$ shift from inferred value), $A_{\rm IA} = 1.18$ and -0.32, to contaminate our data vectors. Similarly to the baryonic feedback case, we rescale them by the ratio of the summary statistics estimated from simulations with the chosen $A_{\rm IA}$ and with $A_{\rm IA} = 0$ [101].

For each realization of the fiducial set, taking all four summary statistics as our data vector, we employ the $\mathcal{F}_i$ trained models, for $i = 1, \ldots, 100$, for data compression and use the mean value (over the 100 compressions) as our "observation," sampling the posterior distribution for the S_8 parameter, as detailed in previous section. We follow the exact same procedure for parameter inference from each simulation accounting for a systematic effect.

We show in Fig. 3 the impact on S_8 constraints from the fiducial set of simulations due to the contamination by different systematic effects. The impact is evaluated in two cases: combining all four summary statistics (employed in our baseline analysis) and using only the NG statistics

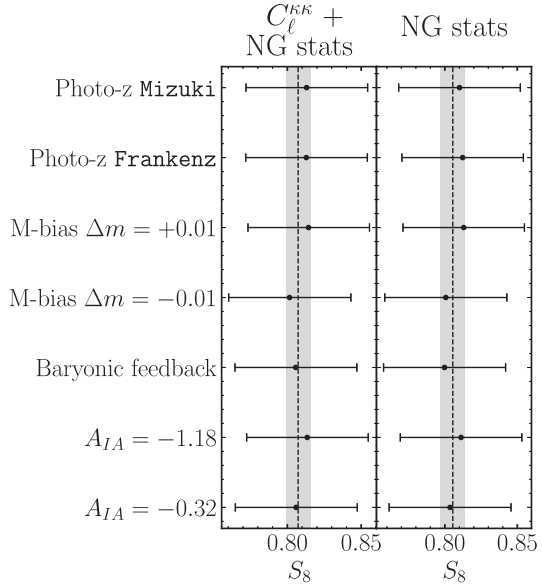


FIG. 3. Impact of the different systematic effects on S_8 constraints from fiducial simulations, using the combination of all four summary statistics ($C_\ell^{\kappa\kappa}$ + NG stats; left) and the NG statistics alone (excluding $C_\ell^{\kappa\kappa}$; right). The horizontal axis shows the average S_8 constraints from data vectors contaminated by systematics. The vertical axis indicates the individual systematic effect tested. The error bar corresponds to the average 68% confidence interval of the posteriors obtained for each contaminated data vector. The dashed vertical line in the left (right) panel represents the average from original fiducial realizations, $S_8 = 0.807 \pm 0.041$ ($S_8 = 0.805 \pm 0.042$) for $C_\ell^{\kappa\kappa}$ + NG stats (for NG stats), while the gray region illustrates the $\pm 0.2\sigma$ range from the same set.

(excluding the $C_\ell^{\kappa\kappa}$). Notice that, in both cases, the vertical dashed lines represent the average constraint from the fiducial set, showing a slight deviation from the true value ($\sim 0.016\sigma$). We find that, combining all the summary statistics, the bias introduced by systematic effects is no larger than 0.17σ, obtained for the case of a multiplicative bias miscalibration of $\Delta m = +0.01$. Very similar results are found using only the NG statistics. The rather small bias introduced by the evaluated systematic effects confirms the robustness and reliability of our results (see next section). In fact, our choices for the combination of all summary statistics, the smoothing scale of 4 arcmin, the exclusion of the highest redshift bin (z_4), and the scale cuts on $C_\ell^{\kappa\kappa}$, defining our baseline analysis, are made in such a way to assure the small amplitude of these biases, as we show here.[9] In other words, the small biases we show in Fig. 3 corroborate the applicability of our pipeline out of the training set, given that none of the systematic effects is considered in the cosmology-varied simulations employed during training processes.

Finally, it is also worth noticing that no cuts are applied to the NG statistics, meaning that the analyses employing these statistics (individually or combined) use the complete data vectors (all 19 ν_t bins from each of them). The reason for this choice is that using our methodology, unlike the $C_\ell^{\kappa\kappa}$, the NG statistics (combined or not with $C_\ell^{\kappa\kappa}$; Fig. 3) perform better without any cut, and no significant bias is observed. Therefore, our choice of smoothing scale balances minimizing the impact of systematic effects while preserving as much small-scale information as possible to maximize the non-Gaussian information accessible through NG statistics.

VI. RESULTS AND DISCUSSIONS

In this section, we present our main results and findings analyzing the HSC Y1 data, presenting our cosmological constraints using the combination of all summary statistics, our baseline analysis, and results from other combinations of statistics. We also present and discuss some consistency tests, reinforcing the accuracy of our main result.

A. HSC Y1 analysis

We present in Fig. 4 results of our method applied to the HSC Y1 data, showing Ω_m and S_8 posteriors and final constraints obtained using a combination of Gaussian and NG summary statistics, namely, $C_\ell^{\kappa\kappa}$, MFs, PM, and PDF using the analysis described above. The constraints obtained from $C_\ell^{\kappa\kappa}$ individually and the NG statistics are also shown for comparison, showing that all three SBI constraints are statistically consistent. We find that the combined analysis can improve over the PS constraints by 26% for S_8 and by 40% for Ω_m, while slightly shifting the posterior median to higher and smaller values, respectively. The individual usage of NG statistics seems to prefer higher (smaller) values of S_8 (Ω_m) parameter and allows 25% (30%) tighter constraints, with respect to PS only. It is worth noticing that a similar trend is observed from the RMSE values shown in Fig. 1. Also, the NN predictions for HSC Y1 data, as well as the error bars given by the RMSE amplitudes, are very consistent ($< 1\sigma$) with the constraints from Ω_m and S_8 posteriors (Fig. 4), in particular, when combining all summary statistics. Such consistency not only provides further confirmation of the reliability of our constraints, but also corroborates an expected equivalence between the Bayesian analysis, through posterior estimation, and the frequentist approach, using directly the NN predictions.

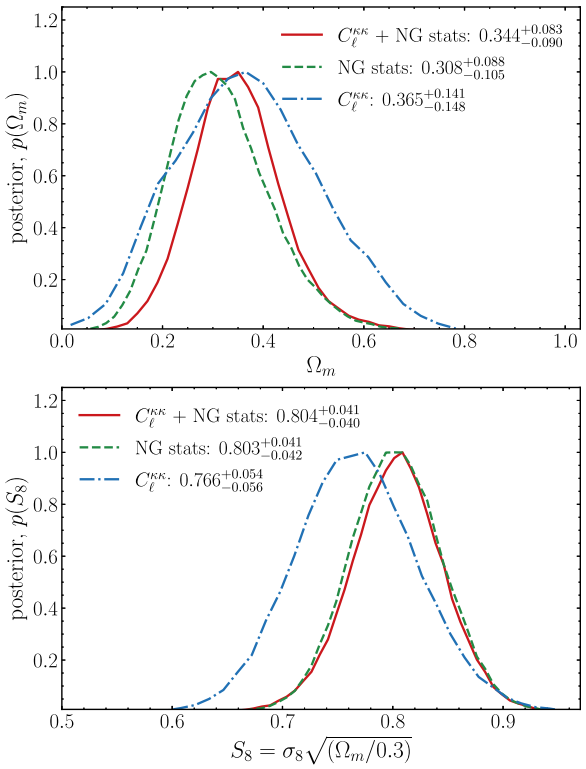


FIG. 4. Cosmological constraints on HSC Y1 data. The Ω_m and S_8 posterior are shown in the top and bottom panels, respectively. Their median and 16 and 84 percentiles are quoted in each panel, showing cosmological constraints using different combinations of summary statistics.

A summary of the S_8 cosmological constraints for the different NG statistics individually and each of them combined with the PS is presented in Fig. 5; results from Fig. 4 are also included for comparison. Figure 5 shows that all our tests lead to statistically consistent estimates of the S_8 parameter, with an apparent preference of $C_\ell^{\kappa\kappa}$ for a smaller value, but still inside the 1σ confidence interval. All the results employing NG statistics only lead to tighter

[9]See in Appendix A a brief discussion on how we adapt here the blind strategy followed by previous analyses of HSC Y1 data [27,28].

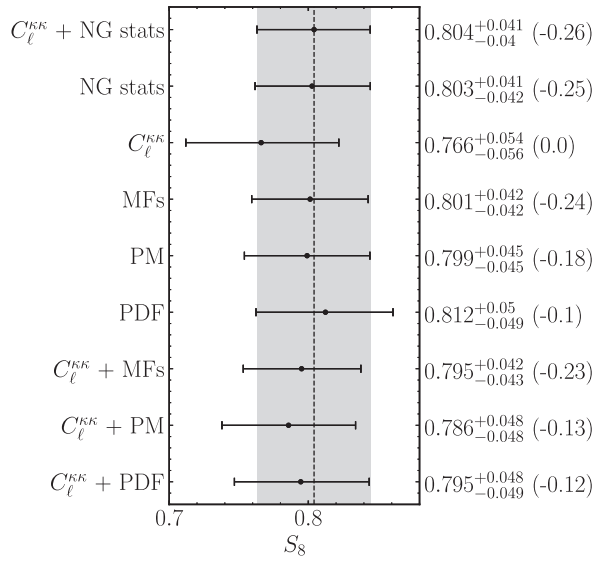


FIG. 5. S_8 constraints for HSC Y1 data using different summary statistics and combinations of them, where the error bars correspond to 16 and 84 percentiles. The right vertical axis also indicates, in parentheses, the fractional improvement of each summary statistic for $C_\ell^{\kappa\kappa}$. The dashed vertical line and gray region represent the inference for our baseline analysis ($C_\ell^{\kappa\kappa}$ + NG stats).

constraints than the PS, following the order of constraining power: MFs > PM > PDF. The individual combination of MFs and PDF with the PS yields a tighter error bar. For PM this is not observed, and its combination with PS yields a slightly wider error compared to PM only, possibly a statistical fluctuation. In addition, results from MFs indicate that this NG statistic seems to be primarily responsible for the improvement over the PS-only results, shrinking the S_8 error bar by ∼24% in comparison to the PS only constraints. It is worth noting that, following an SBI approach, the NG statistics combined with the PS seem to have very similar constraining power as the NG statistics only. However, one should point out that our conservative usage of the $C_\ell^{\kappa\kappa}$, with severe scale cuts, could result in the NG statistics capturing most of the information contained in the two-point statistic while retaining almost all the constraining power. Therefore, in addition to the improvement of the NG statistics in comparison to the PS only constraints, the robustness against systematics (see Sec. V), which allows keeping all ν_t bins, corroborates the significant contribution of these statistics. We emphasize that our conclusions on the relevance of the PS are driven from our analysis constraining only two cosmological parameters. Including more parameters might lead to a different constraining power of the PS.

Our SBI analysis leads to constraints that are statistically consistent with previous results derived from HSC Y1 data. In terms of two-point information, the official HSC Y1 analyses using the PS and the two-point correlation function find $S_8 = 0.780^{+0.030}_{-0.033}$ [99] and $S_8 = 0.804^{+0.032}_{-0.029}$ [100], respectively; notice the preference of the PS for a slightly smaller value, similarly to what we find here (see Fig. 5). Interestingly, constraints obtained by Ref. [18] indicate that a halo-based theoretical peaks model also prefers slightly smaller values of S_8. Using a CNN, Ref. [40] also provides a S_8 constraint quite consistent with our median estimates, besides also pointing to an improvement up to 24% compared to their PS analysis, similar to our finding including NG statistics. It is important to emphasize here that differences with respect to our findings are expected since we implement other choices of priors, scale cuts (for PS analyses), and treatment of systematics. Also, we employ a completely distinct methodology. While analyses from Refs. [99,100] are conducted using perturbation theory for modeling the two-point information, and Ref. [18] considers theoretical models and approximations for the likelihood, we employ an inference pipeline fully based on simulations. Also, we use an NDE to learn the posterior distribution of the parameters, while Ref. [40], although also forward modeling the observation, still make assumptions to calculate the likelihood of their data vector (the output from the CNN).

B. Consistency tests

To further verify the robustness and reliability of our results, we evaluate our S_8 inference against some modifications, one at a time, on our SBI pipeline. As we describe in Sec. IV, our main results are derived through posterior estimation individually for each cosmological parameter of interest (*1D case*). Figure 6 shows the S_8, Ω_m constraints

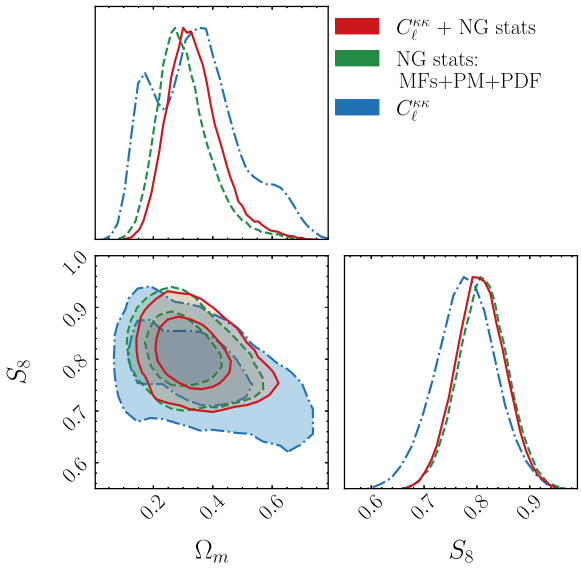


FIG. 6. Posterior distributions (68% and 95%) of Ω_m and S_8 with HSC Y1 data using the PS (blue contour), the combination of the three NG statistics (green contour), and the joint analysis combining all four summary statistics (red contour). The results are obtained following the 2D case approach, as discussed in the text (see also Sec. IV for definition of the 2D case).

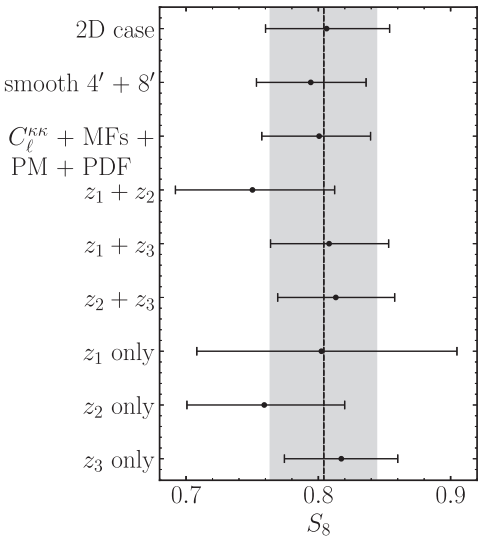


FIG. 7. Consistency tests showing the impact on the posterior approximations for S_8 due to modifications of our SBI pipeline as employed for the baseline analysis. We show the median of the posterior distributions, with error bars corresponding to the 16 and 84 percentiles. The dashed line and gray region represent the results from the baseline analysis. See text for details on each test performed.

from the joint estimation, the *2D case*. Combining all summary statistics, we find[10] $\Omega_m = 0.331^{+0.098}_{-0.078}$ and $S_8 = 0.806^{+0.048}_{-0.046}$, highly consistent with results from the *1D case* (Fig. 4). The comparison of this 2D estimate of S_8 parameter is directly compared to our 1D result in Fig. 7, visually confirming the agreement between the median values, obtaining a slightly tighter constraint from the 1D analysis.

We also evaluate whether the inclusion of compressed information originating from convergence fields with a different smoothing scale would impact our S_8 posterior approximation. In this case, x is a concatenation of the data vectors obtained for $\theta_s = 4$ arcmin and $\theta_s = 8$ arcmin individually, both resulting from the compression maximizing the information on S_8 parameter; θ is still one-dimensional vector filled with the S_8 true values. The constraint obtained is consistent with our baseline analysis, as also seen from Fig. 7 (labeled as "smooth $4' + 8'$"), although we conclude that such additional information does not contribute to a reduction of the uncertainties.

For all these tests, and our baseline analysis, we apply the data compression over the combination of all four summary statistics combined. Our next consistency test evaluates the performance of applying the data compression step to each summary statistic, $C_\ell^{\kappa\kappa}$, MFs, PM, and PDF, individually, and then concatenating the four data vectors into x, keeping the same one-dimensional θ. Again, we find a median and error bar very consistent with that from our baseline analysis, as shown in Fig. 7 (results labeled as "$C_\ell^{\kappa\kappa}$ + MFs + PM + PDF").

Finally, we evaluate how the removal of individual tomographic bins from analysis, as well as the usage of individual bins (nontomographic analysis), impact the S_8 constraints. Results are presented in Fig. 7 and show that the exclusion of z_1 or z_2 redshift bins do not impact significantly our main result (baseline analysis), leading to a slightly increased amplitude of the error bars (similar to using z_3 alone). The exclusion of the third redshift bin, z_3, on the other hand, has a more significant impact, resulting in a wider posterior with a median shifted to a smaller S_8 value, although still consistent within 1σ with our baseline analysis. The individual usage of z_2 shows a similar impact on the S_8 constraint, while z_1 alone, although not shifting the S_8 value, increases significantly the error bars. A similar impact was observed by previous analyses of HSC Y1 using NG statistics [27,28]; in particular, Ref. [29] found, regardless of the summary statistic, a lowering in S_8 due to the absence of z_3, pointing to a limitation of S_8 constraints by photometric redshift estimation when using HSC. For further discussion on this, we refer the reader to Ref. [29]. In summary, all tests indicate the robustness of our results, showing consistency within 1σ for all modifications of the pipeline. A discussion on the tomographic approach in the particular context of NG statistics is presented in Appendix D.

VII. CONCLUSIONS

This work presents the first cosmological inference with the HSC Y1 NG statistics using an approach fully based on simulations and no assumptions on the likelihood of the data. We forward model a set of Gaussian and NG statistics, namely, the PS, MFs, PM, and PDF, using a suite of N-body simulations tailored to the data. We also preprocess our convergence maps (data and simulations) and data vectors by applying smoothing kernels and cuts on scales (for the PS) and redshift bins so that we mitigate possible systematic effects not accounted in the simulations. Our SBI pipeline for cosmological inference is implemented in two steps: first, a neural network performs a massive compression of the summary statistics; then, the output (the compressed data) is used to train an NDE to learn the posterior distribution of the parameters of interest. The full pipeline is carefully validated on simulations prior to its application to data.

Combining all four summary statistics, we find $S_8 = 0.804^{+0.041}_{-0.040}$ and $\Omega_m = 0.344^{+0.083}_{-0.090}$, which represents an improvement of 26% and 40%, respectively, over the PS-only constraints, namely, $S_8 = 0.766^{+0.054}_{-0.056}$ and

[10]The *2D case* results in $\Omega_m = 0.294^{+0.100}_{-0.074}$ and $S_8 = 0.812^{+0.047}_{-0.047}$, using NG statistics (excluding $C_\ell^{\kappa\kappa}$), and $\Omega_m = 0.340^{+0.156}_{-0.155}$ and $S_8 = 0.779^{+0.056}_{-0.058}$, using $C_\ell^{\kappa\kappa}$ alone, also in very good agreement with results from the *1D case* (Fig. 4).

$\Omega_m = 0.365^{+0.148}_{-0.141}$ (68% CL). Similar constraints are obtained directly from the NN predictions, reinforcing the robustness of our estimates while indicating a consistency between Bayesian and frequentist analyses from the SBI perspective. These results are in good agreement (within 1σ) with previous analyses of the same data set, using both Gaussian and NG statistics, as well as with CMB measurements by Planck [102]. We also find that the combination of all four summary statistics has very similar constraining power as the NG statistics only (PS absent). A possible reason for that might be the severe scale cuts applied to $C_\ell^{\kappa\kappa}$, intended to minimize the impact of systematics. This suggests that most of the information in the PS is also captured by the NG statistics, besides the exclusive NG information they are capable of extracting from the convergence fields. In particular, using each statistic individually, we also conclude that the information is primarily coming from the MFs, while finding good agreement between the statistics. Also, for our SBI approach, no cuts seem to be necessary for the NG statistics, given our choice of ν range, indicating a certain robustness to systematic effects.

We verify the robustness of our pipeline against systematic effects. These effects are the photometric redshift uncertainties, multiplicative bias, baryonic feedback, and intrinsic alignments, not accounted for in the training simulations. Such evaluation not only guided choices for the baseline analysis, such as the smoothing scale, redshift bins, and combination of summary statics, but also allowed to confirm the robustness of our final pipeline outside the training set. Consistency tests, modifying some of our choices for the baseline analysis, show robust agreement with our constraints, reinforcing the reliability of the results.

Our study corroborates previous findings on the high constraining power of NG statistics, alone and jointly to standard two-point statistics, for weak lensing analyses. We also demonstrate how competitive (and computationally feasible) an SBI approach can be for cosmological inference using as many summary statistics as needed, without making assumptions about the likelihood of the data. We show this to be achievable even with a limited number of simulations, a common challenge in weak lensing analyses. However, to ensure the minimum impact from unknown systematics, we take actions that result in the loss of information, such as scale cuts and large smoothing scales. Future analyses, with improvements in the modeling of systematics, and the use of the latest HSC releases, would increase the constraining power of our methodology. Additionally, in the context of the next generation of extragalactic experiments, the straightforward combination of summary statistics from different probes can contribute in exploring additional parameters and alternative models, thereby helping to address the challenges of future surveys.

We acknowledge the use of many public Python packages not cited along the main text: NumPy [103], Astropy [104] a community-developed core Python package for Astronomy [105,106], Matplotlib [107], IPython [108], SciPy [109], SCIKIT-LEARN [110] and TensorFlow [111].

ACKNOWLEDGMENTS

The authors thank Tilman Tröster, Daniela Grandón, and Natalí de Santi for enlightening discussions. We also thank an anonymous referee for insightful comments. C. P. N. and E. F. thanks Instituto Serrapilheira for financial support. This work was initiated at the CD3 x Simons Foundation workshop "AI-driven discovery in physics and astrophysics" from January 22–26, 2024, at Kavli IPMU, Kashiwa, Japan. This work was supported by JSPS KAKENHI Grant No. JP23K19064 (J. A.). This work was supported by JSPS KAKENHI Grants No. 23K13095 and No. 23H00107 (to J. L.). This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. This research used computing resources at Kavli IPMU. This research used resources at the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231. The Kavli IPMU is supported by the WPI (World Premier International Research Center) Initiative of the MEXT (Japanese Ministry of Education, Culture, Sports, Science and Technology).

DATA AVAILABILITY

The data that support the findings of this article are openly available [43,50].

APPENDIX A: BLINDING PROCEDURE

Given that HSC Y1 official analyses using two-point statistics [99,100] were already public when studies using NG statistics by Refs. [27,28] were conducted, the authors followed a blinding approach as an honor system. Here, we employ the same simulations and NG statistics as these studies, in addition to the MFs [30], and, for this reason, we take some of their findings as our starting points, skipping or modifying some steps of the same blinding procedure, as we describe below:

(1) We employ a suite of simulations tailored to HSC Y1 data to forward model a the summary statistics. Our whole SBI pipeline, data compression and posterior inference, is constructed on these synthetic data vectors and, then, tested and validated taking individual realizations from the fiducial set as "observation," assuring the recovery of the input cosmology.
(2) We selected smoothing scales using as "observation" the data vectors contaminated by systematic effects,

so that we minimize their impact. Our choice of redshift bins and scale cuts on the PS follow findings from Refs. [19,28].

(3) Unblinding:
   (a) B-modes stage: The summary statistics estimated from B-modes of the fiducial set of realizations are compared to those from real data. References [19,28,30] confirmed that both have noiselike behavior, being statistically consistent. For this reason, we skip this step.
   (b) PS stage: We apply the SBI pipeline to HSC Y1 data, using the C_ℓ only, and compared the results to those previously published by the official analyses by Refs. [99,100], as well as to results from Refs. [27,28]. We find consistency between the results. We note that at this stage [27], based on results from HSC year 3 [45,112], decided to exclude the highest redshift bin due to possible problems with redshift calibration. We adopted the same choice since the beginning of our analysis.
   (c) NG statistics stage: Finally, we apply our SBI pipeline to the HSC Y1 data, using both the NG statistics only (MFs + PM + PDF) and their combination with the PS, finding good agreement (within 1σ) with the same previous results.

APPENDIX B: LIKELIHOOD OF THE COMPRESSED DATA

Here, we investigate the likelihood of the compressed data. In Fig. 8, the top panels show the histogram distribution of the compressed data, or predicted values of Ω_m and S_8, indicating them to be well fitted by Gaussian curves. The lower panels show the probability plots of the compressed data against the quantiles of a normal distribution, confirming them to be well approximated by a Gaussian.

APPENDIX C: ROBUSTNESS OF THE COMBINED ANALYSIS

As an additional robustness test of the combined analysis, using Gaussian and NG summary statistics, we investigate here if the difference of the posterior among $C_\ell^{\kappa\kappa}$ alone case and the NG statistics case could be explained by a statistical fluke. For this, we do the S_8 inference with our SBI pipeline over all mock data, comparing, in the top panel of Fig. 9, the median values of the posteriors obtained using $C_\ell^{\kappa\kappa}$ alone to those resulting from using the NG statistics.

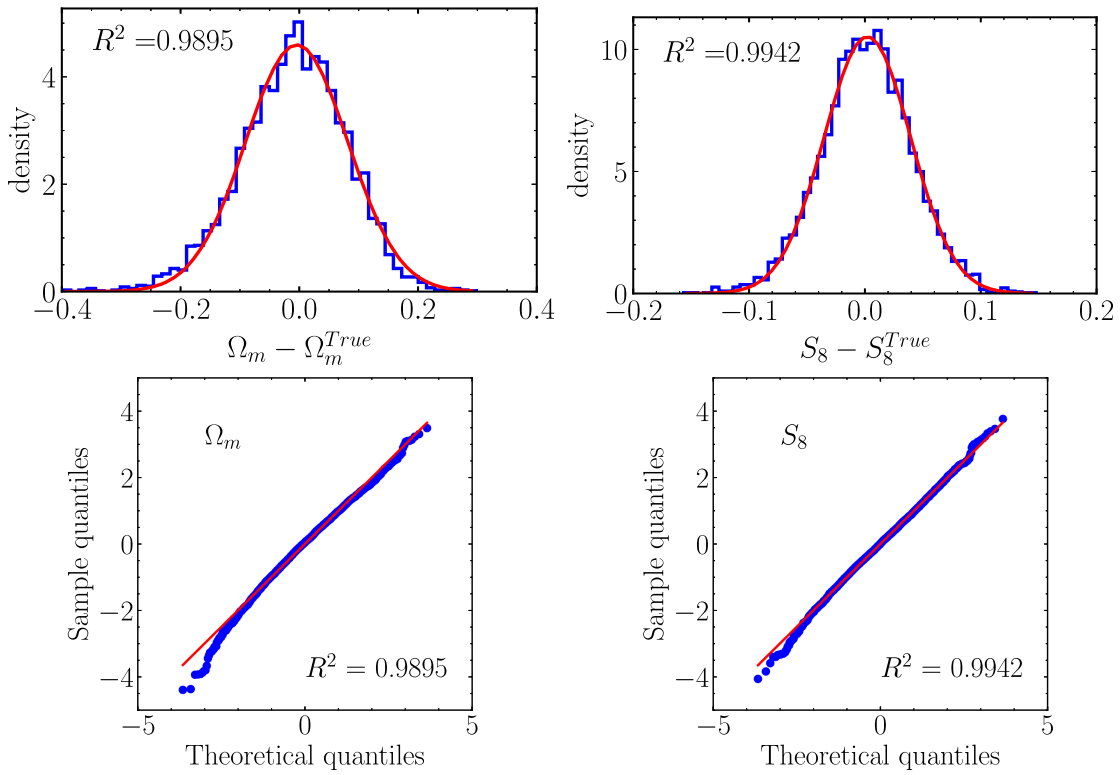


FIG. 8. Top panels: histogram distribution of the compressed data (predicted values) from the cosmology-varied simulations, a sample with 5000 elements (50 realizations × 100 cosmologies), after subtracting the respective true values. The best-fit of a Gaussian curve is shown as a red line; coefficients of determination, R^2 score, are also included. Bottom panels: normal probability plots, or quantile-quantile (Q-Q) plots, comparing the distribution of the compressed data against a (theoretical) Gaussian distribution. The red lines represent the linear fit of the data points and R^2 the respective coefficients of determination.

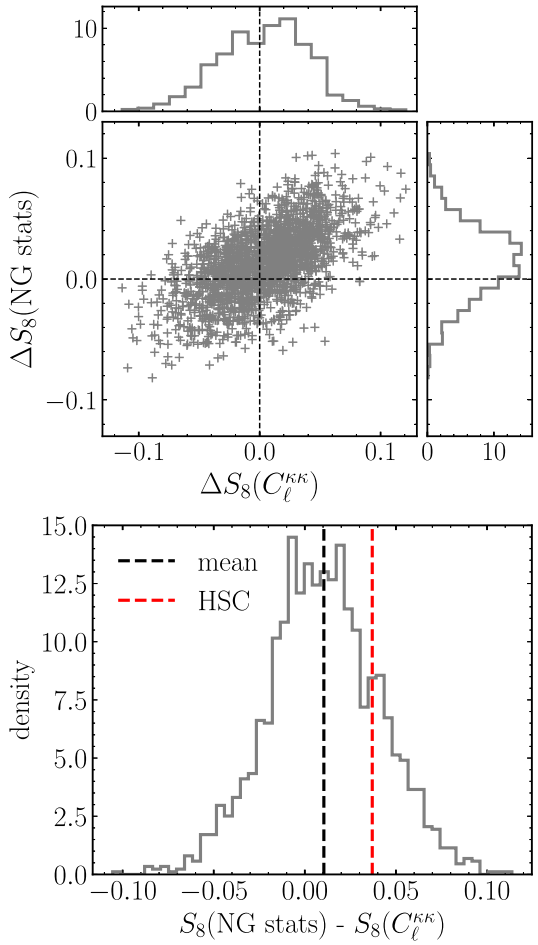


FIG. 9. Top: scatter plot of the median values of the posterior distributions of S_8 derived from applying the SBI pipeline to each of the 2268 fiducial simulations, after subtracting the true value, $\Delta S_8 = S_8 - S_8^{\rm true}$. The constraints obtained using the angular power spectrum (x-axis) are compared to those using the NG statistics (y-axis); the dashed lines correspond to $\Delta S_8 = 0$. Bottom: histogram distribution of the difference S_8(NG stats)-$S_8(C_\ell^{\kappa\kappa})$. The black and red dashed lines represent the average difference from the simulations (0.01) and the difference obtained from HSC analyses (0.037; lower panel of Fig. 4 of the paper).

The scatter plot shows a clear correlation between them, indicating that the NG statistics should provide results consistent with $C_\ell^{\kappa\kappa}$.

Additionally, the lower panel of Fig. 9 shows the histogram distribution of the difference between the median values of the posterior from each case, $C_\ell^{\kappa\kappa}$ and NG statistics. Given the reasonably high frequency with which we find in the simulations the same shift, S_8(NG stats)-$S_8(C_\ell^{\kappa\kappa})$, observed from the HSC Y1 data, it could indeed be explained by a statistical fluke. This indicates that our analyses does not seem to suffer from any serious systematic error, and we can safely perform and trust the constraints combining $C_\ell^{\kappa\kappa}$ and NG statistics.

APPENDIX D: TOMOGRAPHY FROM NG STATISTICS

Figure 7 shows the impact of removing each redshift bin, as well as using individual bins, employing the combination of all four summary statistics. From this case we find that the third redshift bin, z_3, provides the tightest constraint on S_8 parameter, while a tomographic approach including z_1 and z_2 tightens its error bar in 5%. Motivated by this, and by previous works showing benefits of using a tomographic approach [19,27,75,113], we followed a tomographic analysis in the main text. Here, we evaluate the tomographic approach in the context of the NG statistics alone. We show in Fig. 10 that the S_8 constraints obtained from NG statistics when using individual redshift bins and different combinations of pairs of them. The results are very similar to what is show in Fig. 7 for the combination of all summary statistics. In particular, we find that the result obtained using the last redshift bin, z_3, alone, namely, the lowest point in Fig. 10, has an error bar with same amplitude as the result combining all three redshift bins, represented by the gray region. This indicates that in our specific case, the combination of MFs, PM, and PDF to analyze the HSC Y1 data would not benefit significantly from a tomographic approach, unless they are combined with two-point statistics.

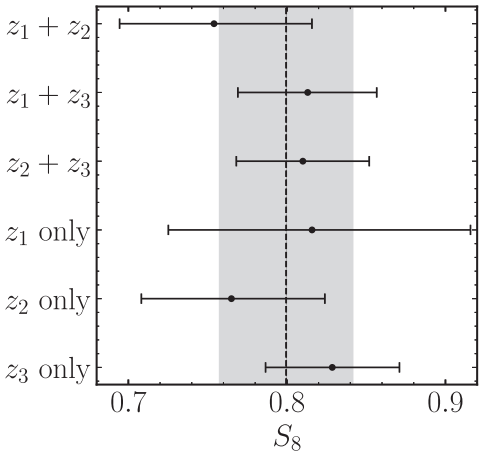


FIG. 10. Tomographic analyses, using only NG statistics, showing the impact on the posterior approximations for S_8 due to different combinations of redshif bins feeding our SBI pipeline. We show the median of the posterior distributions, with error bars corresponding to the 16 and 84 percentiles. The dashed line and gray region represent the results from using NG statistics alone and combining all three redshift bins, $z_1 + z_2 + z_3$.

[1] A. Albrecht, G. Bernstein, R. Cahn, W. L. Freedman, J. Hewitt, W. Hu, J. Huth, M. Kamionkowski, E. W. Kolb, L. Knox *et al.*, Report of the dark energy task force, arXiv:astro-ph/0609591.

[2] K. Kuijken, C. Heymans, H. Hildebrandt, R. Nakajima, T. Erben, J. T. de Jong, M. Viola, A. Choi, H. Hoekstra, L. Miller *et al.*, Gravitational lensing analysis of the kilo-degree survey, Mon. Not. R. Astron. Soc. **454**, 3500 (2015).

[3] T. M. C. Abbott *et al.* (DES Collaboration), The Dark Energy Survey data release 2, Astrophys. J. Suppl. Ser. **255**, 20 (2021).

[4] H. Aihara, R. Armstrong, S. Bickerton, J. Bosch, J. Coupon, H. Furusawa, Y. Hayashi, H. Ikeda, Y. Kamata, H. Karoji *et al.*, First data release of the Hyper Suprime-Cam Subaru Strategic Program, Publ. Astron. Soc. Jpn. **70**, S8 (2018).

[5] R. Scaramella, J. Amiaux, Y. Mellier, C. Burigana, C. Carvalho, J.-C. Cuillandre, A. Da Silva, A. Derosa, J. Dinis, E. Maiorano *et al.*, Euclid preparation-I. The Euclid wide survey, Astron. Astrophys. **662**, A112 (2022).

[6] Y. Mellier, J. Barroso, A. Achúcarro, J. Adamek, R. Adam, G. Addison, N. Aghanim, M. Aguena, V. Ajani, Y. Akrami *et al.*, Euclid. I. Overview of the Euclid mission, arXiv:2405.13491.

[7] A. Abate *et al.* (LSST Dark Energy Science Collaboration), Large synoptic survey telescope: Dark Energy Science Collaboration, arXiv:1211.0310.

[8] E. Di Valentino, L. A. Anchordoqui, Ö. Akarsu, Y. Ali-Haimoud, L. Amendola, N. Arendse, M. Asgari, M. Ballardini, S. Basilakos, E. Battistelli *et al.*, Cosmology intertwined III: $f\sigma_8 S_8$ and, Astropart. Phys. **131**, 102604 (2021).

[9] G. Marques, M. Madhavacheril, O. Darwish, S. Shaikh, M. Aguena, O. Alves, S. Avila, D. Bacon, E. Baxter, K. Bechtol *et al.*, Cosmological constraints from the tomography of DES-Y3 galaxies with CMB lensing from act DR4, J. Cosmol. Astropart. Phys. 01 (2024) 033.

[10] M. S. Madhavacheril, F. J. Qu, B. D. Sherwin, N. MacCrann, Y. Li, I. Abril-Cabezas, P. A. Ade, S. Aiola, T. Alford, M. Amiri *et al.*, The Atacama Cosmology Telescope: DR6 gravitational lensing map and cosmological parameters, Astrophys. J. **962**, 113 (2024).

[11] M. White, R. Zhou, J. DeRose, S. Ferraro, S.-F. Chen, N. Kokron, S. Bailey, D. Brooks, J. Garcia-Bellido, J. Guy *et al.*, Cosmological constraints from the tomographic cross-correlation of DESI luminous red galaxies and Planck CMB lensing, J. Cosmol. Astropart. Phys. 02 (2022) 007.

[12] F. Avila, A. Bernui, A. Bonilla, and R. C. Nunes, Inferring $S_8(z)$ and $\gamma(z)$ with cosmic growth rate measurements using machine learning, Eur. Phys. J. C **82**, 594 (2022).

[13] C. Franco, J. Oliveira, M. Lopes, F. Avila, and A. Bernui, Measuring the matter fluctuations in the local Universe with the ALFALFA catalogue, Mon. Not. R. Astron. Soc. **537**, 897 (2025).

[14] V. Ajani, M. Baldi, A. Barthelemy, A. Boyle, P. Burger, V. Cardone, S. Cheng, S. Codis, C. Giocoli, J. Harnois-Déraps *et al.*, Euclid preparation-XVIII. Forecasts for ten different higher-order weak lensing statistics, Astron. Astrophys. **675**, A120 (2023).

[15] M. Gatti, N. Jeffrey, L. Whiteway, J. Williamson, B. Jain, V. Ajani, D. Anbajagane, G. Giannini, C. Zhou, A. Porredon *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).

[16] D. Zürcher, J. Fluri, V. Ajani, S. Fischbacher, A. Refregier, and T. Kacprzak, Towards a full *w*CDM map-based analysis for weak lensing surveys, Mon. Not. R. Astron. Soc. **525**, 761 (2023).

[17] D. Anbajagane, C. Chang, A. Banerjee, T. Abel, M. Gatti, V. Ajani, A. Alarcon, A. Amon, E. Baxter, K. Bechtol *et al.*, Beyond the 3rd moment: A practical study of using lensing convergence CDFs for cosmology with DESY3, Mon. Not. R. Astron. Soc. **526**, 5530 (2023).

[18] X. Liu, S. Yuan, C. Pan, T. Zhang, Q. Wang, and Z. Fan, Cosmological studies from HSC-SSP tomographic weak-lensing peak abundances, Mon. Not. R. Astron. Soc. **519**, 594 (2023).

[19] G. A. Marques, J. Liu, J. M. Z. Matilla, Z. Haiman, A. Bernui, and C. P. Novaes, Constraining neutrino mass with weak lensing Minkowski functionals, J. Cosmol. Astropart. Phys. 06 (2019) 019.

[20] M. Gatti, B. Jain, C. Chang, M. Raveri, D. Zürcher, L. Secco, L. Whiteway, N. Jeffrey, C. Doux, T. Kacprzak *et al.*, Dark Energy Survey Year 3 results: Cosmology with moments of weak lensing mass maps, Phys. Rev. D **106**, 083509 (2022).

[21] A. Petri, J. Liu, Z. Haiman, M. May, L. Hui, and J. M. Kratochvil, Emulating the CFHTLenS weak lensing data: Cosmological constraints from moments and Minkowski functionals, Phys. Rev. D **91**, 103511 (2015).

[22] O. Friedrich, D. Gruen, J. DeRose, D. Kirk, E. Krause, T. McClintock, E. Rykoff, S. Seitz, R. Wechsler, G. Bernstein *et al.*, Density split statistics: Joint model of counts and lensing in cells, Phys. Rev. D **98**, 023508 (2018).

[23] D. Gruen, O. Friedrich, E. Krause, J. DeRose, R. Cawthon, C. Davis, J. Elvin-Poole, E. Rykoff, R. Wechsler, A. Alarcon *et al.*, Density split statistics: Cosmological constraints from counts and lensing in cells in DES Y1 and SDSS data, Phys. Rev. D **98**, 023507 (2018).

[24] P. A. Burger, O. Friedrich, J. Harnois-Déraps, P. Schneider, M. Asgari, M. Bilicki, H. Hildebrandt, A. H. Wright, T. Castro, K. Dolag *et al.*, KiDS-1000 cosmology: Constraints from density split statistics, Astron. Astrophys. **669**, A69 (2023).

[25] E. Semboloni, T. Schrabback, L. van Waerbeke, S. Vafaei, J. Hartlap, and S. Hilbert, Weak lensing from space: First cosmological constraints from three-point shear statistics, Mon. Not. R. Astron. Soc. **410**, 143 (2011).

[26] L. Fu, M. Kilbinger, T. Erben, C. Heymans, H. Hildebrandt, H. Hoekstra, T. D. Kitching, Y. Mellier, L. Miller, E. Semboloni *et al.*, CFHTLenS: Cosmological constraints from a combination of cosmic shear two-point and three-point correlations, Mon. Not. R. Astron. Soc. **441**, 2725 (2014).

[27] 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).

[28] L. Thiele, G. A. Marques, J. Liu, and M. Shirasaki, Cosmological constraints from the Subaru Hyper Suprime-Cam year 1 shear catalogue lensing convergence probability distribution function, Phys. Rev. D **108**, 123526 (2023).

[29] S. Cheng, G. A. Marques, D. Grandón, L. Thiele, M. Shirasaki, B. Ménard, and J. Liu, Cosmological constraints from weak lensing scattering transform using HSC Y1 data, J. Cosmol. Astropart. Phys. 01 (2025) 006.

[30] J. Armijo *et al.*, Cosmological constraints using Minkowski functionals from the first year data of the Hyper Suprime-Cam, Mon. Not. R. Astron. Soc. **537**, 3553 (2025).

[31] D. Grandón, G. A. Marques, L. Thiele, S. Cheng, M. Shirasaki, and J. Liu, Impact of baryonic feedback on HSC Y1 weak lensing non-Gaussian statistics, Phys. Rev. D **110**, 103539 (2024).

[32] K. Cranmer, J. Brehmer, and G. Louppe, The frontier of simulation-based inference, Proc. Natl. Acad. Sci. U.S.A. **117**, 30055 (2020).

[33] K. Lin, M. von Wietersheim-Kramsta, B. Joachimi, and S. Feeney, A simulation-based inference pipeline for cosmic shear with the kilo-degree survey, Mon. Not. R. Astron. Soc. **524**, 6167 (2023).

[34] E. Massara, C. Hahn, M. Eickenberg, S. Ho, J. Hou, P. Lemos, C. Modi, A. M. Dizgah, L. Parker, and B. R.-S. Blancard, SimBIG: Cosmological constraints using simulation-based inference of galaxy clustering with marked power spectra, arXiv:2404.04228.

[35] M. von Wietersheim-Kramsta, K. Lin, N. Tessore, B. Joachimi, A. Loureiro, R. Reischke, and A. H. Wright, KiDS-SBI: Simulation-based inference analysis of KiDS-1000 cosmic shear, Astron. Astrophys. **694**, A223 (2025).

[36] G. F. Abellán, G. C. Herrera, M. Martinelli, O. Savchenko, D. Sciotti, and C. Weniger, Fast likelihood-free inference in the LSS stage IV era, J. Cosmol. Astropart. Phys. 11 (2024) 057.

[37] 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 (2021).

[38] P. Lemos, L. Parker, C. Hahn, S. Ho, M. Eickenberg, J. Hou, E. Massara, C. Modi, A. M. Dizgah, B. R.-S. Blancard *et al.*, Field-level simulation-based inference of galaxy clustering with convolutional neural networks, Phys. Rev. D **109**, 083536 (2024).

[39] X. Zhao, Y. Mao, C. Cheng, and B. D. Wandelt, Simulation-based inference of reionization parameters from 3d tomographic 21 cm light-cone images, Astrophys. J. **926**, 151 (2022).

[40] T. Lu, Z. Haiman, and X. Li, Cosmological constraints from HSC survey first-year data using deep learning, Mon. Not. R. Astron. Soc. **521**, 2050 (2023).

[41] C. P. Novaes, E. J. de Mericia, F. B. Abdalla, C. A. Wuensche, L. Santos, J. Delabrouille, M. Remazeilles, V. Liccardo, E. Abdalla, L. Barosi *et al.*, Cosmological constraints from low redshift 21 cm intensity mapping with machine learning, Mon. Not. R. Astron. Soc. **528**, 2078 (2024).

[42] J. Hermans, A. Delaunoy, F. Rozet, A. Wehenkel, V. Begy, and G. Louppe, A trust crisis in simulation-based inference? Your posterior approximations can be unfaithful, arXiv:2110.06581.

[43] R. Mandelbaum, H. Miyatake, T. Hamana, M. Oguri, M. Simet, R. Armstrong, J. Bosch, R. Murata, F. Lanusse, A. Leauthaud *et al.*, The first-year shear catalog of the Subaru Hyper Suprime-Cam Subaru Strategic Program survey, Publ. Astron. Soc. Jpn. **70**, S25 (2018).

[44] M. Tanaka, J. Coupon, B.-C. Hsieh, S. Mineo, A. J. Nishizawa, J. Speagle, H. Furusawa, S. Miyazaki, and H. Murayama, Photometric redshifts for hyper Suprime-Cam Subaru Strategic Program data release 1, Publ. Astron. Soc. Jpn. **70**, S9 (2018).

[45] R. Dalal, X. Li, A. Nicola, J. Zuntz, M. A. Strauss, S. Sugiyama, T. Zhang, M. M. Rau, R. Mandelbaum, M. Takada *et al.*, Hyper Suprime-Cam year 3 results: Cosmology from cosmic shear power spectra, Phys. Rev. D **108**, 123519 (2023).

[46] N. Kaiser and G. Squires, Mapping the dark matter with weak gravitational lensing, Astrophys. J. **404**, 441 (1993).

[47] M. Shirasaki, K. Moriwaki, T. Oogi, N. Yoshida, S. Ikeda, and T. Nishimichi, Noise reduction for weak lensing mass mapping: An application of generative adversarial networks to Subaru Hyper Suprime-Cam first-year data, Mon. Not. R. Astron. Soc. **504**, 1825 (2021).

[48] B. Jain, U. Seljak, and S. White, Ray-tracing simulations of weak lensing by large-scale structure, Astrophys. J. **530**, 547 (2000).

[49] T. Hamana and Y. Mellier, Numerical study of the statistical properties of the lensing excursion angles, Mon. Not. R. Astron. Soc. **327**, 169 (2001).

[50] R. Takahashi, T. Hamana, M. Shirasaki, T. Namikawa, T. Nishimichi, K. Osato, and K. Shiroyama, Full-sky gravitational lensing simulation for large-area galaxy surveys and cosmic microwave background experiments, Astrophys. J. **850**, 24 (2017).

[51] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. Bennett, J. Dunkley, M. Nolta, M. Halpern, R. Hill, N. Odegard *et al.*, Nine-year Wilkinson microwave anisotropy probe (WMAP) observations: Cosmological parameter results, Astrophys. J. Suppl. Ser. **208**, 19 (2013).

[52] M. Shirasaki and N. Yoshida, Statistical and systematic errors in the measurement of weak-lensing Minkowski functionals: Application to the Canada-France-Hawaii lensing survey, Astrophys. J. **786**, 43 (2014).

[53] D. Alonso, J. Sanchez, and A. Slosar (LSST Dark Energy Science Collaboration), A unified pseudo-C_ℓ framework, Mon. Not. R. Astron. Soc. **484**, 4127 (2019).

[54] M. Oguri, S. Miyazaki, C. Hikage, R. Mandelbaum, Y. Utsumi, H. Miyatake, M. Takada, R. Armstrong, J. Bosch, Y. Komiyama *et al.*, Two-and three-dimensional wide-field weak lensing mass maps from the Hyper Suprime-Cam

Subaru strategic program S16A data, Publ. Astron. Soc. Jpn. **70**, S26 (2018).
[55] E. Komatsu, A. Kogut, M. Nolta, C. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, M. Limon, S. Meyer, L. Page *et al.*, First-year Wilkinson microwave anisotropy probe (WMAP)* observations: Tests of Gaussianity, Astrophys. J. Suppl. Ser. **148**, 119 (2003).
[56] H. Modest, C. Räth, A. Banday, G. Rossmanith, R. Sütterlin, S. Basak, J. Delabrouille, K. Górski, and G. Morfill, Scale-dependent non-Gaussianities in the CMB data identified with Minkowski functionals and scaling indices, Mon. Not. R. Astron. Soc. **428**, 551 (2013).
[57] C. Novaes, A. Bernui, I. Ferreira, and C. Wuensche, A neural-network based estimator to search for primordial non-Gaussianity in Planck CMB maps, J. Cosmol. Astropart. Phys. 09 (2015) 064.
[58] C. Novaes, A. Bernui, G. Marques, and I. Ferreira, Local analyses of Planck maps with Minkowski functionals, Mon. Not. R. Astron. Soc. **461**, 1363 (2016).
[59] Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. Barreiro, N. Bartolo, S. Basak, K. Benabed *et al.*, Planck 2018 results-VII. Isotropy and statistics of the CMB, Astron. Astrophys. **641**, A7 (2020).
[60] E. Saar, V. J. Martínez, J.-L. Starck, and D. L. Donoho, Multiscale morphology of the galaxy distribution, Mon. Not. R. Astron. Soc. **374**, 1030 (2007).
[61] M. Kerscher and A. Tikhonov, Morphology of the local volume, Astron. Astrophys. **509**, A57 (2010).
[62] Y.-Y. Choi, J. Kim, G. Rossi, S. S. Kim, and J.-E. Lee, Topology of luminous red galaxies from the Sloan Digital Sky Survey, Astrophys. J. Suppl. Ser. **209**, 19 (2013).
[63] C. P. Novaes, A. Bernui, H. S. Xavier, and G. A. Marques, Tomographic local 2D analyses of the WISE*x*SuperCOSMOS all-sky galaxy catalogue, Mon. Not. R. Astron. Soc. **478**, 3253 (2018).
[64] D. Munshi, T. Namikawa, J. McEwen, T. Kitching, and F. Bouchet, Morphology of weak lensing convergence maps, Mon. Not. R. Astron. Soc. **507**, 1421 (2021).
[65] H. Minkowski, Volumen und Oberfläche, Math. Ann. **57**, 447 (1903).
[66] D. Novikov, H. A. Feldman, and S. F. Shandarin, Minkowski functionals and cluster analysis for CMB maps, Int. J. Mod. Phys. D **08**, 291 (1999).
[67] A. Ducout, F. Bouchet, S. Colombi, D. Pogosyan, and S. Prunet, Non-Gaussianity and Minkowski functionals: Forecasts for Planck, Mon. Not. R. Astron. Soc. **429**, 2104 (2013).
[68] M. Shirasaki, T. Nishimichi, B. Li, and Y. Higuchi, The imprint of $f(r)$ gravity on weak gravitational lensing–II. Information content in cosmic shear statistics, Mon. Not. R. Astron. Soc. **466**, 2402 (2017).
[69] J. M. Z. Matilla, Z. Haiman, A. Petri, and T. Namikawa, Geometry and growth contributions to cosmic shear observables, Phys. Rev. D **96**, 023513 (2017).
[70] J. M. Kratochvil, E. A. Lim, S. Wang, Z. Haiman, M. May, and K. Huffenberger, Probing cosmology with weak lensing Minkowski functionals, Phys. Rev. D **85**, 103513 (2012).
[71] J. Harnois-Déraps, N. Martinet, T. Castro, K. Dolag, B. Giblin, C. Heymans, H. Hildebrandt, and Q. Xia, Cosmic shear cosmology beyond two-point statistics: A combined peak count and correlation function analysis of DES-Y1, Mon. Not. R. Astron. Soc. **506**, 1623 (2021).
[72] X. Liu, C. Pan, R. Li, H. Shan, Q. Wang, L. Fu, Z. Fan, J.-P. Kneib, A. Leauthaud, L. Van Waerbeke *et al.*, Cosmological constraints from weak lensing peak statistics with Canada–France–Hawaii telescope stripe 82 survey, Mon. Not. R. Astron. Soc. **450**, 2888 (2015).
[73] J. Liu, A. Petri, Z. Haiman, L. Hui, J. M. Kratochvil, and M. May, Cosmology constraints from the weak lensing peak counts and the power spectrum in CFHTLenS data, Phys. Rev. D **91**, 063507 (2015).
[74] W. R. Coulton, J. Liu, I. G. McCarthy, and K. Osato, Weak lensing minima and peaks: Cosmological constraints and the impact of baryons, Mon. Not. R. Astron. Soc. **495**, 2531 (2020).
[75] J. Liu and M. S. Madhavacheril, Constraining neutrino mass with the tomographic weak lensing one-point probability distribution function and power spectrum, Phys. Rev. D **99**, 083508 (2019).
[76] A. Barthelemy, A. Halder, Z. Gong, and C. Uhlemann, Making the leap. Part I. Modelling the reconstructed lensing convergence PDF from cosmic shear with survey masks and systematics, J. Cosmol. Astropart. Phys. 03 (2024) 060.
[77] B. Giblin, Y.-C. Cai, and J. Harnois-Déraps, Enhancing cosmic shear with the multiscale lensing probability density function, Mon. Not. R. Astron. Soc. **520**, 1721 (2023).
[78] A. Boyle, C. Uhlemann, O. Friedrich, A. Barthelemy, S. Codis, F. Bernardeau, C. Giocoli, and M. Baldi, Nuw CDM cosmology from the weak-lensing convergence PDF, Mon. Not. R. Astron. Soc. **505**, 2886 (2021).
[79] K. Patton, J. Blazek, K. Honscheid, E. Huff, P. Melchior, A. J. Ross, and E. Suchyta, Cosmological constraints from the convergence 1-point probability distribution, Mon. Not. R. Astron. Soc. **472**, 439 (2017).
[80] A. F. 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).
[81] T. Charnock, G. Lavaux, and B. D. Wandelt, Automatic physical inference with information maximizing neural networks, Phys. Rev. D **97**, 083004 (2018).
[82] M. Park, M. Gatti, and B. Jain, Dimensionality reduction techniques for statistical inference in cosmology, Phys. Rev. D **111**, 063523 (2025).
[83] N. Jeffrey, L. Whiteway, M. Gatti, J. Williamson, J. Alsing, A. Porredon, J. Prat, C. Doux, B. Jain, C. Chang *et al.*, Dark Energy Survey Year 3 results: Likelihood-free, simulation-based *w*CDM inference with neural compression of weak-lensing map statistics, Mon. Not. R. Astron. Soc. **536**, 1303 (2024).
[84] F. Villaescusa-Navarro, B. D. Wandelt, D. Anglés-Alcázar, S. Genel, J. M. Z. Matilla, S. Ho, and D. N. Spergel, Neural networks as optimal estimators to marginalize over baryonic effects, Astrophys. J. **928**, 44 (2022).
[85] L. A. Perez, S. Genel, F. Villaescusa-Navarro, R. S. Somerville, A. Gabrielpillai, D. Anglés-Alcázar, B. D.

Wandelt, and L. Yung, Constraining cosmology with machine learning and galaxy clustering: The CAMELS-SAM suite, Astrophys. J. **954**, 11 (2023).
[86] A. Geroldinger, L. Lusa, M. Nold, and G. Heinze, Leave-one-out cross-validation, penalization, and differential bias of some prediction model performance measures–A simulation study, Diagn. Prognostic Res. **7**, 9 (2023).
[87] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, OPTUNA: A next-generation hyperparameter optimization framework, in *Proceedings of the 25rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Association for Computing Machinery, New York, NY, 2019).
[88] M. Gatti, G. Campailla, N. Jeffrey, L. Whiteway, A. Porredon, J. Prat, J. Williamson, M. Raveri, B. Jain, V. Ajani, G. Giannini *et al.*, Dark Energy Survey Year 3 results: Simulation-based cosmological inference with wavelet harmonics, scattering transforms, and moments of weak lensing mass maps II. Cosmological results, Phys. Rev. D **111**, 063504 (2025).
[89] T. Lu, Z. Haiman, and J. M. Zorrilla Matilla, Simultaneously constraining cosmology and baryonic physics via deep learning from weak lensing, Mon. Not. R. Astron. Soc. **511**, 1518 (2022).
[90] G. Papamakarios and I. Murray, Fast ϵ-free inference of simulation models with Bayesian conditional density estimation, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., 2016), Vol. 29.
[91] D. Greenberg, M. Nonnenmacher, and J. Macke, Automatic posterior transformation for likelihood-free inference, in International Conference on Machine Learning (PMLR, 2019), pp. 2404–2414.
[92] A. Tejero-Cantero, J. Boelts, M. Deistler, J.-M. Lueckmann, C. Durkan, P. J. Gonçalves, D. S. Greenberg, and J. H. Macke, SBI—A toolkit for simulation-based inference, arXiv:2007.09114.
[93] G. Papamakarios, T. Pavlakou, and I. Murray, Masked autoregressive flow for density estimation, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., 2017), Vol. 30.
[94] C. M. Bishop, Mixture density networks, Technical Report, Aston University, 1994, https://publications.aston.ac.uk/id/eprint/373/.
[95] S. Kullback and R. A. Leibler, On information and sufficiency, Ann. Math. Stat. **22**, 79 (1951).
[96] S. Talts, M. Betancourt, D. Simpson, A. Vehtari, and A. Gelman, Validating Bayesian inference algorithms with simulation-based calibration, arXiv:1804.06788.
[97] R. Mandelbaum, F. Lanusse, A. Leauthaud, R. Armstrong, M. Simet, H. Miyatake, J. E. Meyers, J. Bosch, R. Murata, S. Miyazaki *et al.*, Weak lensing shear calibration with simulations of the HSC survey, Mon. Not. R. Astron. Soc. **481**, 3170 (2018).
[98] 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).
[99] C. Hikage, M. Oguri, T. Hamana, S. More, R. Mandelbaum, M. Takada, F. Köhlinger, H. Miyatake, A. J. Nishizawa, H. Aihara *et al.*, Cosmology from cosmic shear power spectra with Subaru Hyper Suprime-Cam first-year data, Publ. Astron. Soc. Jpn. **71**, 43 (2019).
[100] T. Hamana, M. Shirasaki, S. Miyazaki, C. Hikage, M. Oguri, S. More, R. Armstrong, A. Leauthaud, R. Mandelbaum, H. Miyatake *et al.*, Cosmological constraints from cosmic shear two-point correlation functions with HSC survey first-year data, Publ. Astron. Soc. Jpn. **72**, 16 (2020).
[101] J. Harnois-Déraps, N. Martinet, and R. Reischke, Cosmic shear beyond 2-point statistics: Accounting for galaxy intrinsic alignment with projected tidal fields, Mon. Not. R. Astron. Soc. **509**, 3868 (2022).
[102] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. Banday, R. Barreiro, N. Bartolo, S. Basak *et al.*, Planck 2018 results-VI. Cosmological parameters, Astron. Astrophys. **641**, A6 (2020).
[103] S. Van Der Walt, S. C. Colbert, and G. Varoquaux, The NumPy array: A structure for efficient numerical computation, Comput. Sci. Eng. **13**, 22 (2011).
[104] http://www.astropy.org
[105] A. M. Price-Whelan, B. Sipőcz, H. Günther, P. Lim, S. Crawford, S. Conseil, D. Shupe, M. Craig, N. Dencheva, A. Ginsburg *et al.*, The Astropy project: Building an open-science project and status of the v2. 0 core package, Astron. J. **156**, 123 (2018).
[106] T. P. Robitaille, E. J. Tollerud, P. Greenfield, M. Droettboom, E. Bray, T. Aldcroft, M. Davis, A. Ginsburg, A. M. Price-Whelan, W. E. Kerzendorf *et al.*, Astropy: A community Python package for astronomy, Astron. Astrophys. **558**, A33 (2013).
[107] J. D. Hunter, Matplotlib: A 2d graphics environment, Comput. Sci. Eng. **9**, 90 (2007).
[108] F. Pérez and B. E. Granger, IPython: A system for interactive scientific computing, Comput. Sci. Eng. **9**, 21 (2007).
[109] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in Python, Nat. Methods **17**, 261 (2020).
[110] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, SCIKIT-LEARN: Machine learning in Python, J. Mach. Learn. Res. **12**, 2825 (2011).
[111] M. Abadi *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems, software available from www.tensorflow.org (2015).
[112] X. Li, T. Zhang, S. Sugiyama, R. Dalal, R. Terasawa, M. M. Rau, R. Mandelbaum, M. Takada, S. More, M. A. Strauss *et al.*, Hyper Suprime-Cam year 3 results: Cosmology from cosmic shear two-point correlation functions, Phys. Rev. D **108**, 123518 (2023).
[113] L. Castiblanco, C. Uhlemann, J. Harnois-Déraps, and A. Barthelemy, Unleashing cosmic shear information with the tomographic weak lensing PDF, Open J. Astrophys. **7** (2024).