Cosmological constraints from the nonlinear galaxy bispectrum

ChangHoon Hahn,[1,*] Michael Eickenberg,[2] Shirley Ho,[3] Jiamin Hou,[4,5] Pablo Lemos,[6,7,2] Elena Massara,[8,9] Chirag Modi,[2,3] Azadeh Moradinezhad Dizgah,[10] Liam Parker,[1] and Bruno Régaldo-Saint Blancard[2]

(SimBIG Collaboration)

[1]*Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey 08544, USA*
[2]*Center for Computational Mathematics, Flatiron Institute, 162 5th Avenue, New York, New York 10010, USA*
[3]*Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, New York 10010, USA*
[4]*Department of Astronomy, University of Florida, 211 Bryant Space Science Center, Gainesville, Florida 32611, USA*
[5]*Max-Planck-Institut für Extraterrestrische Physik, Postfach 1312, Giessenbachstrasse 1, 85748 Garching bei München, Germany*
[6]*Department of Physics, Université de Montréal, Montréal, 1375 Avenue Thérèse-Lavoie-Roux, Quebec H2V 0B3, Canada*
[7]*Mila-Quebec Artificial Intelligence Institute, Montréal, 6666 Rue Saint-Urbain, Quebec H2S 3H1, Canada*
[8]*Waterloo Centre for Astrophysics, University of Waterloo, 200 University University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada*
[9]*Department of Physics and Astronomy, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada*
[10]*Département de Physique Théorique, Université de Genève, 24 quai Ernest Ansermet, 1211 Genève 4, Switzerland*

(Received 10 July 2023; accepted 18 December 2023; published 30 April 2024)

We present the cosmological constraints from analyzing higher-order galaxy clustering on small nonlinear scales. We use SimBIG, a forward modeling framework for galaxy clustering analyses that employs simulation-based inference to perform highly efficient cosmological inference using normalizing flows. It leverages the predictive power of high-fidelity simulations and robustly extracts cosmological information from regimes inaccessible with current standard analyses. In this work, we apply SimBIG to a subset of the BOSS galaxy sample and analyze the redshift-space bispectrum monopole, $B_0(k_1, k_2, k_3)$, to $k_{\rm max} = 0.5\ h/\mathrm{Mpc}$. We achieve 1σ constraints of $\Omega_m = 0.293^{+0.027}_{-0.027}$ and $\sigma_8 = 0.783^{+0.040}_{-0.038}$, which are more than 1.2 and 2.4× tighter than constraints from standard power spectrum analyses of the same dataset. We also derive 1.4, 1.4, 1.7× tighter constraints on Ω_b, h, n_s. This improvement comes from additional cosmological information in higher-order clustering on nonlinear scales and, for σ_8, is equivalent to the gain expected from a standard analysis on a ∼4× larger galaxy sample. Even with our BOSS subsample, which only spans 10% of the full BOSS volume, we derive competitive constraints on the growth of structure: $S_8 = 0.774^{+0.056}_{-0.053}$. Our constraint is consistent with results from both cosmic microwave background and weak lensing. Combined with a ω_b prior from big bang nucleosynthesis, we also derive a constraint on $H_0 = 67.6^{+2.2}_{-1.8}\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ that is consistent with early Universe constraints.

DOI: 10.1103/PhysRevD.109.083534

I. INTRODUCTION

The three-dimensional spatial distribution of galaxies enables us to constrain the nature of dark matter and dark energy and measure the contents of the Universe. Along with other cosmological probes, it provides one of the most stringent tests of the standard Lambda cold dark matter (ΛCDM) cosmological model that can lead to discoveries of new physics. With this aim, spectroscopic galaxy surveys of the next decade, the Dark Energy Spectroscopic Instrument (DESI) [1–3], Subaru Prime Focus Spectrograph (PFS) [4,5], the ESA *Euclid* satellite mission [6], and

*changhoon.hahn@princeton.edu

the Nancy Grace Roman Space Telescope (Roman) [7,8], will probe galaxies over unprecedented cosmic volumes out to $z \sim 3$.

Current analyses of galaxy clustering focus on the power spectrum, the Fourier counterpart to the two-point correlation function, as the primary measurement of galaxy clustering (e.g. [9–14]). These standard analyses model the power spectrum using the perturbation theory (PT) of large-scale structure (see [15,16] for a review). As a result, they focus on large, mostly linear, scales ($k_{\max} \sim 0.2\ h/\mathrm{Mpc}$) where deviations from linear theory are small and PT remains valid. Accurate modeling of higher-order clustering statistics (e.g. bispectrum) with PT is progressively more complex and challenging. Furthermore, there are currently no PT-based models that describe new promising summary statistics (e.g. [17–20]).

Meanwhile, studies have now established that there is additional cosmological information in higher-order statistics (e.g. [21–25]). Forecasts have also long suggested that there may be even more information on small scales (e.g. [26]). Recently, [27,28] showed that constraints on ΛCDM cosmological parameters, $\Omega_m, \Omega_b, h, n_s, \sigma_8$, improve by a factor of ~2 by analyzing the bispectrum down to nonlinear scales ($k_{\max} = 0.5\ h/\mathrm{Mpc}$). References [18,19,29–34] found consistent improvements from forecasts of other summary statistics that extract non-Gaussian cosmological information from nonlinear scales. These improvements are further corroborated by recent small-scale clustering analyses using emulators [35,36].

Another major limitation of current analyses is robustly accounting for observational systematics in e.g. targeting, imaging, completeness that significantly impact clustering measurements [37,38]. Fiber collisions, for example, prevent galaxy surveys that use fiber-fed spectrographs (e.g. DESI, PFS) from successfully measuring redshifts from galaxies within some angular scale of one another [39]. They significantly bias the power spectrum measurement on scales smaller than $k > 0.1\ h/\mathrm{Mpc}$ [40–42]. While improved correction schemes for fiber collisions may be sufficient for power spectrum analyses [41–44], no correction scheme has yet been designed or demonstrated for other summary statistics.

Recently, [45,46][1] presented the SIMulation-Based Inference of Galaxies (SimBIG), a forward modeling framework for analyzing galaxy clustering. SimBIG uses simulation-based inference (SBI)[2] (see [47] for a review) to perform highly efficient cosmological parameter inference using neural density estimation (NDE) from machine learning (e.g. [48,49]). This enables SimBIG to use high-fidelity simulations that model the details and realism of the observations. In particular, the SimBIG forward model is based on cosmological N-body simulations that can more accurately model nonlinear structure formation to smaller scales than PT. It also includes observational systematics (e.g. survey geometry, masking, fiber collisions). With this approach, H22a analyzed the galaxy power spectrum from the Sloan Digital Sky Survey (SDSS)-III Baryon Oscillation Spectroscopic Survey (BOSS) [50,51]. This work demonstrated that they can rigorously analyze the power spectrum down to smaller scales than ever before, $k_{\max} = 0.5\ h/\mathrm{Mpc}$.

In this work, we extend the SimBIG analysis to the first higher-order statistic: the bispectrum. For a near-Gaussian galaxy distribution, the bispectrum extracts nearly all of its cosmological information (e.g. [52–54]). We present the first robust cosmological constraints from an analysis that exploits clustering information in both higher-order statistics and on nonlinear scales beyond $k_{\max} > 0.2\ h/\mathrm{Mpc}$. We begin in Sec. II by describing the observational galaxy sample that we analyze. We then briefly summarize the details of the SimBIG approach in Sec. III. We present and discuss our cosmological results in Sec. IV and compare them to constraints in the literature.

II. OBSERVATIONS: BOSS CMASS GALAXIES

We apply our SimBIG bispectrum analysis to the same observed galaxy sample as H22a, which is derived from the SDSS-III BOSS data release 12 [50,51]. More specifically, the sample consists of galaxies in the Southern Galactic Cap of BOSS CMASS galaxy sample that are within the redshift range $0.45 < z < 0.6$ and have Dec > -6 deg. and $-25 <$ RA < 28 deg. Overall, the galaxy sample covers ~3,600 deg^2 and includes 109,636 galaxies. This corresponds to 70% of the SGC footprint and ~10% of the full BOSS volume. We refer readers to H22a and H23 for further details on the observed galaxy sample.

III. SIMBIG WITH THE GALAXY BISPECTRUM

The SimBIG approach uses SBI to infer posteriors of ΛCDM cosmological parameters with only a forward model that can generate mock observations, i.e. the 3D galaxy distribution. In this section, we briefly describe the forward model, the SBI methodology, the bispectrum, and our posterior validation.

A. Forward model

The SimBIG forward model constructs simulated galaxy catalogs from Quijote N-body simulations run at different cosmologies in a Latin-hypercube configuration [55]. Each simulation has a volume of 1 $(h^{-1}\ \mathrm{Gpc})^3$ and is constructed using 1024^3 CDM particles gravitationally evolved from $z = 127$ to $z = 0.5$. From the N-body simulations, halos are identified using the phase-space information of dark matter particles with the Rockstar halo finder [56]. Afterwards, the halos are populated using the halo

[1]Hereafter H22a and H23.

[2]Also known as "likelihood-free inference" or "implicit likelihood inference".

occupation distribution (HOD) [57,58] framework, which provides a flexible statistical prescription for determining the number of galaxies as well as their positions and velocities within halos. SimBIG uses a state-of-the-art HOD model with 9 parameters that supplements the standard [58] model with assembly, concentration, and velocity biases.

From the HOD galaxy catalog, SimBIG adds a full BOSS survey realism by applying the exact survey geometry and observational systematics. The forward modeled catalogs have the same redshift range and angular footprint of the CMASS sample, including masking for bright stars, center post, bad field, and collision priority. Furthermore, SimBIG also includes fiber collisions, which systematically removes galaxies in galaxy pairs within an angular scale of 62″. We forward model fiber collisions because the standard correction schemes do not accurately correct for them [41]. In summary, the SimBIG forward model aims to generate mock galaxy catalogs that are statistically indistinguishable from the observations. For more details on the forward model, we refer readers to H22a and H23.

B. Simulation-based inference

From the forward modeled galaxy catalogs, we use the SimBIG SBI framework to infer posterior distributions of cosmological parameters, $\boldsymbol{\theta}$, for a given summary statistic, $\boldsymbol{x}$, of the observations: $p(\boldsymbol{\theta}|\boldsymbol{x})$. The SimBIG SBI framework enables cosmological inference with a limited number of forward modeled simulations. This in turn enables us to exploit cosmological information on small, nonlinear scales and in higher-order statistics that is inaccessible with standard cosmological analyses.

The SBI in SimBIG is based on NDE and uses "normalizing flow" models [59–61]. Normalizing flows use neural networks to learn an extremely flexible and bijective transformation, $f: x \mapsto z$, that maps a complex target distribution to a simple base distribution, $\pi(z)$, that is fast to evaluate. f is defined to be invertible and have a tractable Jacobian so that the target distribution can be evaluated from $\pi(z)$ by change of variables. Since $\pi(z)$ is easy to evaluate, this enables us to also easily evaluate the target distribution. In our case, the target distribution is the posterior and the base distribution is a multivariate Gaussian. Among various normalizing flow architectures, we use masked autoregressive flow (MAF) [49] models.[3]

Our goal is to train a normalizing flow with hyperparameters, $\boldsymbol{\phi}$, that best approximates the posterior, $q_{\boldsymbol{\phi}}(\boldsymbol{\theta}|\boldsymbol{x}) \approx p(\boldsymbol{\theta}|\boldsymbol{x})$. We do this by minimizing the forward Kullback-Leibler (KL) divergence between $p(\boldsymbol{\theta}, \boldsymbol{x}) = p(\boldsymbol{\theta}|\boldsymbol{x})p(\boldsymbol{x})$ and $q_{\boldsymbol{\phi}}(\boldsymbol{\theta}|\boldsymbol{x})p(\boldsymbol{x})$. In practice, we first split the forward modeled catalogs into a training and validation set with a 90/10 split. Then we maximize the total log-likelihood $\sum_i \log q_{\boldsymbol{\phi}}(\boldsymbol{\theta}_i|\boldsymbol{x}_i)$ over the training set, $\{(\boldsymbol{\theta}_i, \boldsymbol{x}_i)\}$ This is equivalent to minimizing the forward KL divergence. We use the Adam optimizer [66] with a batch size of 50. To prevent overfitting, we evaluate the total log-likelihood on the validation data at every training epoch and stop the training when the validation log-likelihood fails to increase after 20 epochs.

We determine the architecture of our normalizing flow, i.e. number of blocks, transforms, hidden features, and dropout probability, through experimentation. We train a large number of flows with architectures and learning rates determined using the Optuna hyperparameter optimization framework [67]. Afterwards, we select five normalizing flows with the lowest validation losses. Our final flow is an equally weighted ensemble of the flows: $q_{\boldsymbol{\phi}}(\boldsymbol{\theta}|\boldsymbol{x}) = \sum_{j=1}^{5} q^j_{\boldsymbol{\phi}}(\boldsymbol{\theta}|\boldsymbol{x})/5$. We find that ensembling flows with different initializations and architectures generally improves the robustness of our normalizing flow [68,69]. For the bispectrum, the posteriors predicted by each individual flow in the ensemble are in good agreement.

In $q_{\boldsymbol{\phi}}(\boldsymbol{\theta}|\boldsymbol{x})$, $\boldsymbol{\theta}$ represents the 5 cosmological and 9 HOD parameters. The prior of our posterior estimate is set by the parameter distribution of our training set. Since the N-body simulations used for our forward modeled catalogs are evaluated over a Latin-hypercube, we use uniform priors over the cosmological parameters, $\{\Omega_m, \Omega_b, h, n_s, \sigma_8\}$. The prior ranges fully encompass the Planck priors. For the HOD parameters, we use the same conservative priors from H22a and H23. Next, we describe our summary statistic $\boldsymbol{x}$.

C. Summary statistic: The galaxy bispectrum

With SimBIG we can derive robust cosmological constraints using any summary statistic of the galaxy distribution that we can accurately forward model. In this work, we apply SimBIG to the first higher-order statistic: the galaxy bispectrum. The bispectrum, $B(k_1, k_2, k_3)$, is the three-point correlation function in Fourier space and measures the excess probability of different triangle configurations (k_1, k_2, k_3) over a random distribution. In this work, we focus solely on the monopole of the redshift-space bispectrum, $B_0(k_1, k_2, k_3)$.

To measure B_0, for both observed and forward modeled galaxy samples, we use the [70] redshift-space bispectrum estimator, implemented in the pySpectrum PYTHON package.[4] The estimator uses fast Fourier transforms with grid size $N_{\rm grid} = 360$ and box size $(1800\ h^{-1}\,{\rm Mpc})^3$. The estimator accounts for the survey geometry using a random catalog that has the same radial and angular selection functions as the observed catalog but with a much larger number of objects ($> 4{,}000{,}000$). When measuring B_0, we include the same [71] weights as in H22a. For the observed galaxy sample, we also include angular systematic weights

[3]We use the MAF implementation in sbi PYTHON package [62,63], which is based on the nflows PYTHON package [64,65].

[4]https://github.com/changhoonhahn/pySpectrum.

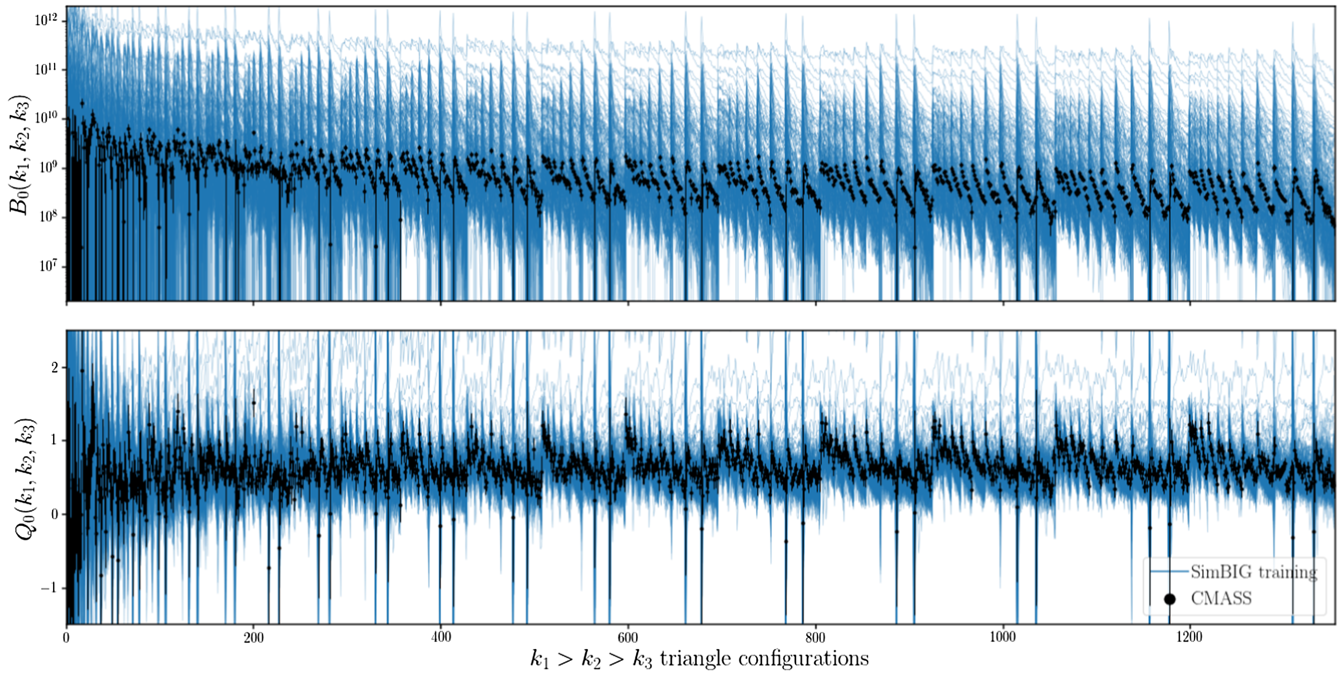

FIG. 1. The bispectrum monopole (B_0; top panel) and reduced bispectrum monopole (Q_0; bottom panel) of a subset of simulated galaxy catalogs in our training set. The catalogs are constructed using the SimBIG forward model from the Quijote N-body simulations and include BOSS survey realism. We randomly select 200 out of the 20,000 catalogs. We present a subset of 1,354 triangle configurations with $k_1, k_2, k_3 < k_{\rm max} = 0.25\ h/{\rm Mpc}$,for clarity. The configurations are ordered by looping through k_3 in the inner most loop and k_1 in the outer most loop with $k_1 \leq k_2 \leq k_3$. For reference, we include B_0 measured from the observed BOSS CMASS sample (black) with error bars estimated from the TEST0 simulations. The observed B_0 is well within our training dataset.

to account for stellar density and seeing conditions as well as redshift failure weights. We do not include weights for fiber collisions, since this effect is included in the SimBIG forward model.

We measure B_0 in triangle configurations defined by (k_1, k_2, k_3) bins of width $\Delta k = 0.0105\ h/{\rm Mpc}$, three times the fundamental mode $k_f = 2\pi/(1800\ h^{-1}\,{\rm Mpc})$. For $k_{\rm max} = 0.5\ h/{\rm Mpc}$, B_0 has 10,052 total triangle configurations. In practice, we use the reduced bispectrum instead of the bispectrum to reduce the dynamic range of the summary statistic[5]:

$$Q_0(k_1, k_2, k_3) = \frac{B_0(k_1, k_2, k_3)}{P_0(k_1)P_0(k_2) + P_0(k_1)P_0(k_3) + P_0(k_2)P_0(k_3)}, \quad (1)$$

where $P_0(k)$ represents the monopole of the power spectrum.

We present $B_0(k_1, k_2, k_3)$ and $Q_0(k_1, k_2, k_3)$ for 200 out of 20,000 randomly selected subsets of the training set in Fig. 1. We only show a subset of 1,354 triangle configurations with $k_1, k_2, k_3 \leq k_{\rm max} = 0.25\ h/{\rm Mpc}$ for clarity. We order the triangles by looping through k_3 in the inner most loop and k_1 in the outer most loop satisfying $k_1 \geq k_2 \geq k_3$. For reference, we include B_0 of the observed CMASS sample (black) with uncertainties estimated using the TEST0 simulations, which we describe in the next section. The B_0 of the training dataset has a broad range that fully encompasses the observed B_0.

D. Posterior validation

Before applying our SimBIG B_0 posterior estimator, $q_{\boldsymbol{\phi}}(\boldsymbol{\theta}|\boldsymbol{x})$, to observations, we validate that it can robustly infer unbiased posteriors of the ΛCDM cosmological parameters. First, we assess whether $q_{\boldsymbol{\phi}}$ accurately estimate the posterior across the parameter space of the prior. We call this the "NDE accuracy test". In principle, with a sufficiently large training set and successful minimization, q_ϕ is guaranteed to accurately estimate the true posterior, since we train it by minimizing the KL divergence with the true posterior. In our case, however, we have a limited number of simulations.

We use the 2,000 validation simulations that were excluded from the training of our posterior estimate (Sec. III B). In Fig. 2, we present the SBC [72] for the ΛCDM cosmological parameters. For each validation simulation, we apply q_ϕ to its $Q_0(k_{123} < 0.5\ h/{\rm Mpc})$ measurement to infer the posterior. Then for each

[5]For simplicity, we will use B_0 to refer to both the bispectrum and reduced bispectrum.

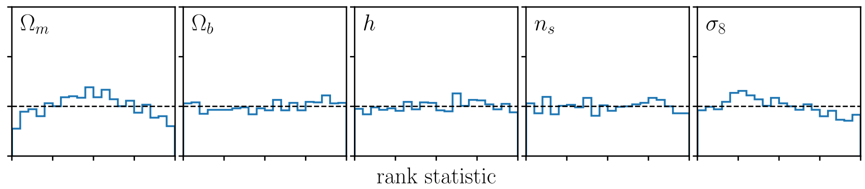

FIG. 2. The NDE accuracy test that shows the SBC validation of the SimBIG $B_0(k_{123} < 0.5\ h/\mathrm{Mpc})$ posterior estimate. We present the distribution of the rank statistics, which are derived by comparing the true parameter values to the inferred marginalized 1D posteriors. The rank statistics are calculated using 2,000 validation simulations that were excluded from training the posterior estimate. For an accurate estimate of the true posterior, the rank statistic would be uniformly distributed (black dashed). Overall, we estimate unbiased posteriors of all of the ΛCDM cosmological parameters.

cosmological parameter, we calculate the rank of the true parameter value within the marginalized 1D posterior estimate. A uniform rank distribution indicates that we accurately estimate the true posterior (black dashed). Overall, the rank distributions are close to uniform for all of the ΛCDM cosmological parameters. For Ω_m and σ_8, the distributions have a slight ∩ shape, which indicate that our Ω_m and σ_8 posterior estimates are slightly broader than the true posterior (i.e. underconfident). Since this means that our cosmological constraints will be conservative, we conclude that q_ϕ is sufficiently accurate.

Next, we verify the robustness of our B_0 posterior with the SimBIG "mock challenge." The SimBIG forward model, or *any* forward model, makes modeling choices and assumptions that, in detail, do not reflect the actual Universe. To account for this, SimBIG is designed to be highly flexible so that we can robustly marginalize over the complex physical processes that govern galaxy formation and the galaxy-halo connection. Nevertheless, a summary statistic may be sensitive to the specific choices made in the forward model. More importantly, this can bias the inferred cosmological parameters. We, therefore, assess whether this is the case for B_0 and validate that we can derive unbiased cosmological parameter constraints.

We use 2,000 test simulations in the three test sets described in H23: TEST0, TEST1, and TEST2. TEST0 consists of 500 "in distribution" simulations built using the same forward model as the training set: Quijote N-body, Rockstar halo finder, and the full SimBIG HOD. TEST1 and TEST2 are "out of distribution" simulations. TEST1 are constructed using Quijote N-body, the Friend-of-Friend halo finder [73], and a simpler HOD model. Lastly, TEST2 consists of 1,000 "out of distribution" simulations built using AbacusSummit N-body simulations [74], CompaSO halo finder [75], and the full SimBIG HOD. Each test set is constructed using a different forward model. Hence, they serve as a stringent test sets for the robustness of the SimBIG B_0 analysis.

We run q_ϕ on the B_0 of all of the test sets and derive a posterior for each simulation. In Fig. 3, we present the

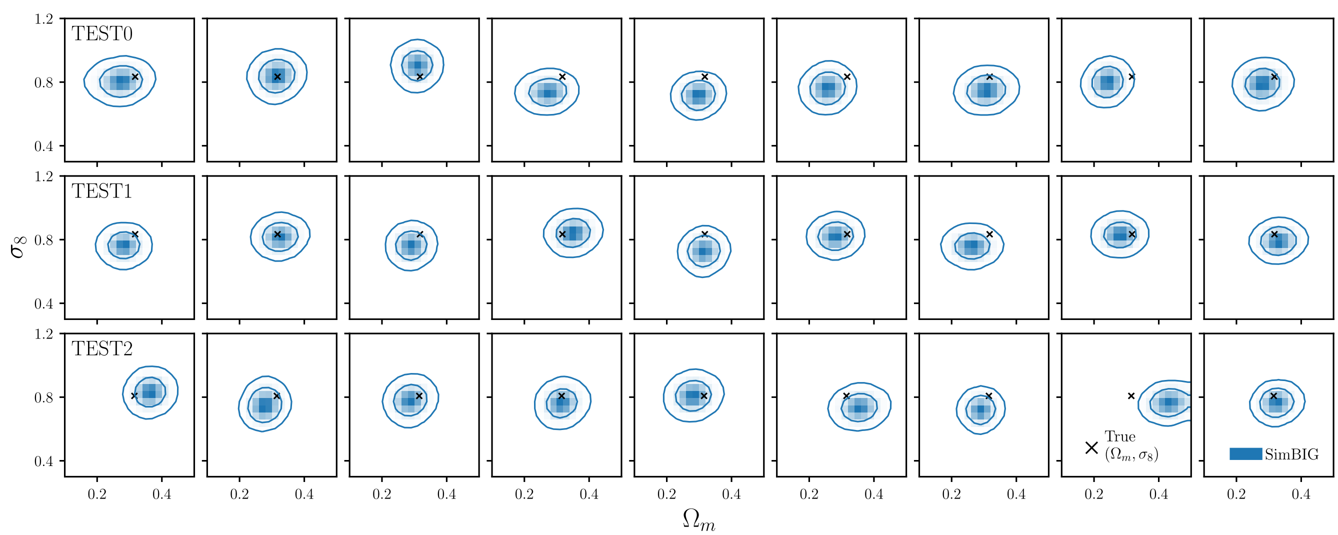

FIG. 3. Posteriors of (Ω_m, σ_8) inferred using the SimBIG bispectrum analysis for a random subset of the TESTO0 (top), TEST1 (center), and TEST2 (bottom) simulations. We mark the 68 and 84 percentiles of the posteriors with the contours. We also include the true (Ω_m, σ_8) of the test simulations in each panel (black ×). The comparison between the posteriors and the true parameter values qualitatively show good agreement for each test simulations.

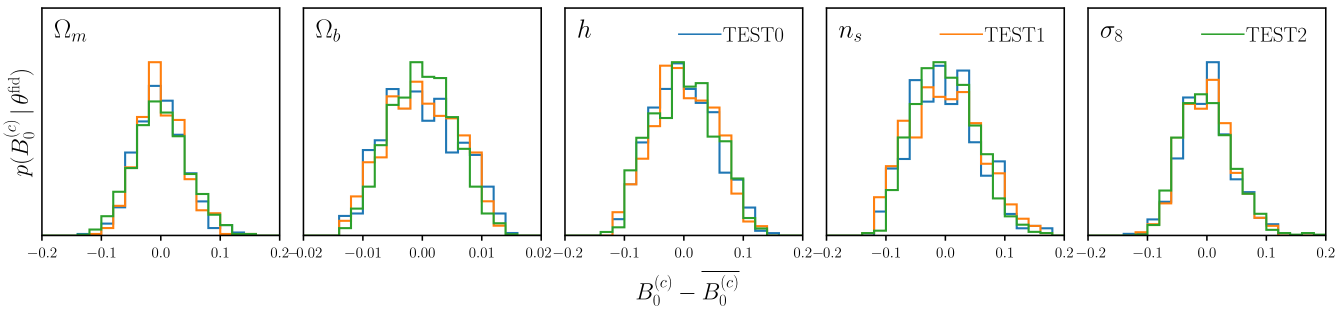

FIG. 4. Comparison of the compressed bispectrum likelihood, $p(B_0^{(c)}|\theta_{\rm fid})$, computed on the three sets of test simulations: TEST0 (blue), TEST1 (orange), and TEST2 (green). $B_0^{(c)}$ is derived by taking the mean of the marginalized 1D SimBIG $B_0(k_{123} < 0.5\ h/{\rm Mpc})$ posterior for the ΛCDM parameters, an optimal compression of the cosmological information in B_0. In each panel, we mark the corresponding ΛCDM parameters. The likelihoods are at the fixed fiducial cosmologies and parameter values of the test sets. We present the distribution of $B_0^{(c)} - \overline{B_0^{(c)}}$ because TEST2 simulations are constructed using different fiducial parameter values than the TEST0 and TEST1 simulations. Overall, we find excellent agreement among the likelihoods of the different test simulations and conclude that our B_0 analysis is robust to modeling choices in our forward model.

(Ω_m, σ_8) posteriors for a randomly selected subset of the test simulations. We present posteriors for TEST0, TEST1, and TEST2 simulations in the top, center, and bottom panels, respectively. The contours represent the 68 and 95 percentiles of the posteriors. In each panel, we mark the true (Ω_m, σ_8) value of the test simulation (black x). Each test simulation is a unique realization of a CMASS-like galaxy catalog subject to cosmic variance. We, therefore, do not expect the true (Ω_m, σ_8) value to lie at the center of each of the posteriors. Instead, we note that for the majority of the randomly selected test simulations, the true parameter values lie within the 68 and 95 percentiles SimBIG posteriors.

Next, we assess the robustness more quantitatively. In H23, we used SBC, or coverage, to assess the robustness of the posterior estimates. This assessment, however, requires that the parameters of the test simulations sample the full prior distribution. Otherwise, the distribution of the rank statistic is not guaranteed to be uniform, even for the true posterior. The test simulations are evaluated at fiducial values of the cosmological parameters. Consequently, we use a different approach and assess the robustness by comparing the B_0 likelihoods of the different test sets. If B_0 is sensitive to variations in the forward model, there will be significant discrepancies among the likelihoods of the test sets.

In practice, comparing the B_0 likelihoods is challenging since $B_0(k_{123} < 0.5\ h/{\rm Mpc})$ is 10,052-dimensional. We instead compare the likelihoods of the compressed B_0, $B_0^{(c)}$, as shown in Fig. 4 for TEST0 (blue), TEST1 (orange), and TEST2. For the compression, we use the mean of the marginalized 1D SimBIG B_0 posterior for the ΛCDM cosmological parameters: $B_0^{(c)} = \sum_{j=1}^{N} \boldsymbol{\theta}_j / N$ where $\boldsymbol{\theta}_j \sim q_{\boldsymbol{\phi}}(\boldsymbol{\theta}|B_0)$. We use $N = 10{,}000$ samples to estimate the mean. Each panel represents a dimension of $B_0^{(c)}$ that corresponds to one of the ΛCDM parameters. This is a near-optimal compression of the cosmological information in B_0, since $q_{\boldsymbol{\phi}}$ accurately estimates the true posterior.

We present the distribution of $B_0^{(c)} - \overline{B_0^{(c)}}$, where $\overline{B_0^{(c)}}$ is the average $B_0^{(c)}$ instead of $B_0^{(c)}$. This is because the TEST2 simulations are constructed using a different set of fiducial parameter values than the TEST0 and TEST1 simulations. Overall, we find excellent agreement among the $B_0^{(c)}$ likelihoods with no significant discrepancies. We also find similar levels of agreement when we use other summaries of the marginalized SimBIG B_0 posterior (e.g. standard deviation, 16th percentile) for the compression. Given the good agreement of $B_0^{(c)}$ likelihoods among the test sets, we conclude that our B_0 analysis is sufficiently robust to the modeling choices in our forward model.

IV. RESULTS

In Fig. 5, we present the posterior distribution of all parameters inferred from the CMASS bispectrum monopole with $k_{\rm max} < 0.5\ h/{\rm Mpc}$ using SimBIG. The top and bottom sets of panels present the posterior of the cosmological and halo occupation parameters, respectively. The diagonal panels present the 1D marginalized posteriors; the rest of the panels present marginalized 2D posteriors of different parameter pairs. The contours represent the 68 and 95 percentiles and the ranges of the panels match the prior. We also list the 50, 16, and 84th percentile constraints on the parameters above the diagonal panels.

Focusing on the ΛCDM cosmological parameters (Fig. 6), we find that the SIMBIG B_0 analysis tightly constrains *all* of them. This is without relying on any priors from big bang nucleosynthesis (BBN) or cosmic microwave background (CMB) experiments that are typically used in galaxy

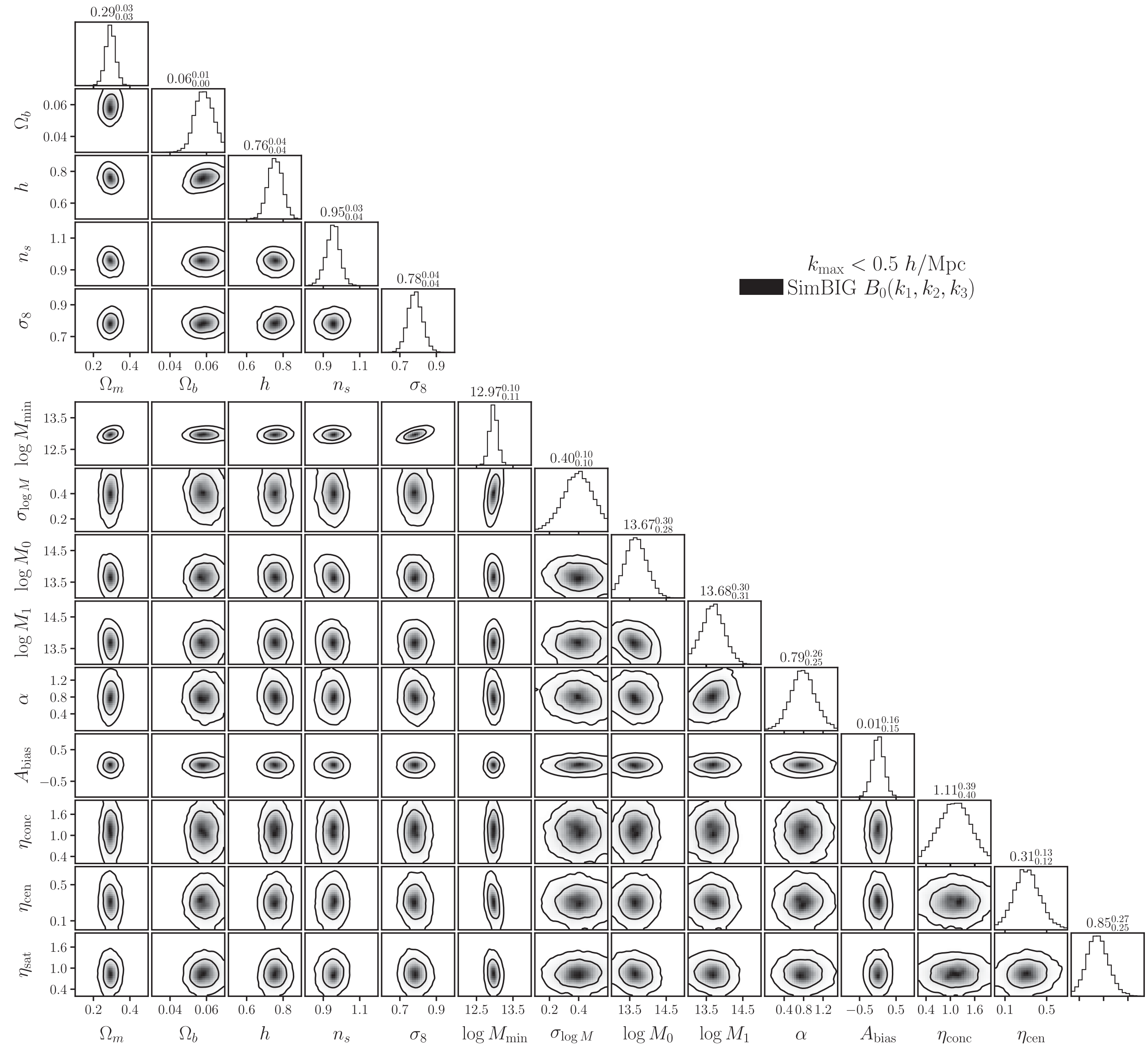

FIG. 5. Posterior distribution of all parameters inferred using the SimBIG B_0 analysis to $k_{\max} < 0.5\ h/\mathrm{Mpc}$ from BOSS CMASS SGC. In the top set of panels, we present the cosmological parameters. In the bottom, we present the halo occupation parameters. The axis ranges of the panels represent the prior range. We place significant constraints on all ΛCDM parameters and a number of the halo occupation parameters (e.g. $\log M_{\min}$, $\log M_0$, and η_{sat}).

clustering analyses (e.g. [11,13,25]). We derive $\Omega_b = 0.059^{+0.005}_{-0.005}$, $h = 0.756^{+0.040}_{-0.039}$, and $n_s = 0.954^{+0.033}_{-0.040}$. For the growth of structure parameters (right panels) we derive: $\Omega_m = 0.293^{+0.027}_{-0.027}$ and $\sigma_8 = 0.783^{+0.040}_{-0.038}$.

Our B_0 analysis places significantly tighter constraints than $P_\ell(k)$ for the same BOSS SGC sample from previous works. Compared to the H22a SimBIG $P_\ell(k < k_{\max} = 0.5)$ analysis, our Ω_m and σ_8 constraints are both 1.7× tighter. This P_ℓ analysis, however, goes beyond standard analyses and includes cosmological information on nonlinear scales. If we compare to a standard PT $P_\ell(k < k_{\max} = 0.25\ h/\mathrm{Mpc})$ analysis ([11] $\Omega_m = 0.317^{+0.031}_{-0.032}$ and $\sigma_8 = 0.719^{+0.100}_{-0.085}$; orange), our Ω_m and σ_8 constraints are 1.2 and 2.5× tighter. Our constraints are also 1.1 and 2.0× tighter than the $P_\ell(k < 0.25\ h/\mathrm{Mpc})$ constraints from [13] ($\Omega_m = 0.314^{+0.031}_{-0.030}$ and $\sigma_8 = 0.790^{+0.083}_{-0.072}$; green). They use a theoretical model based on a halo power spectrum emulator and a halo occupation framework. These comparisons clearly illustrate that the cosmological information in both higher-order statistics and nonlinear scales is *substantial*.

Next, we analyze B_0 to $k_{\max} = 0.3\ h/\mathrm{Mpc}$ to examine how much of the improvement in our B_0 constraints comes from the nonlinear scales alone. In Fig. 7, we present the SimBIG $B_0(k_{123} < 0.3\ h/\mathrm{Mpc})$ posterior (red dashed) on

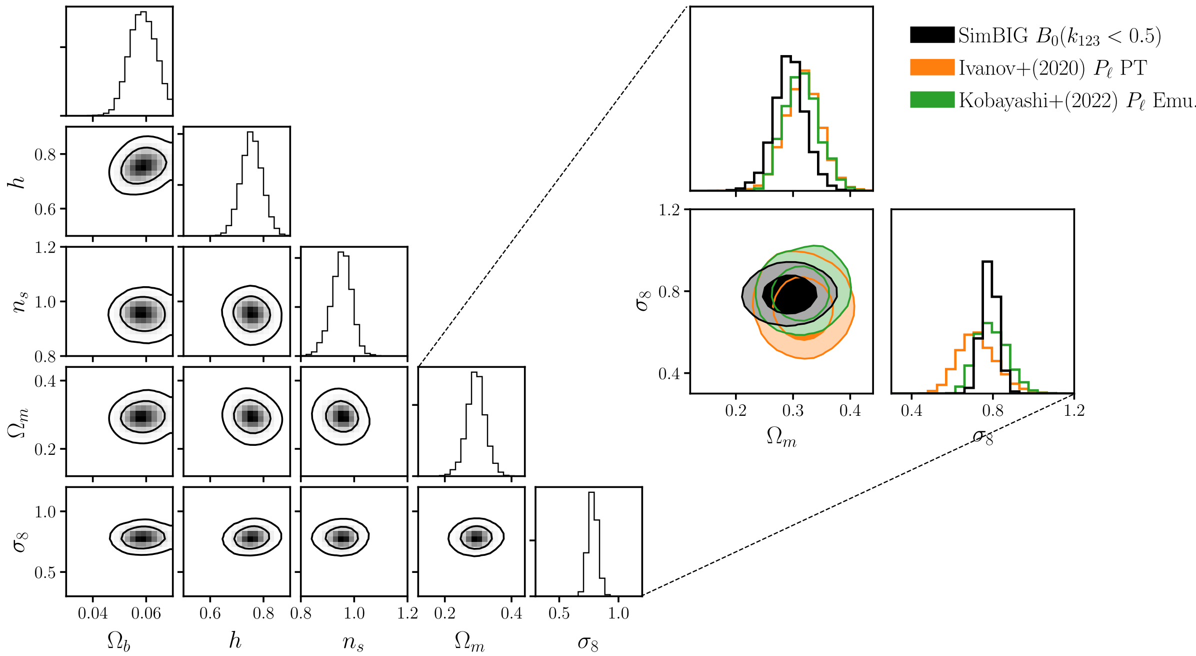

FIG. 6. Left: Posterior of cosmological parameters inferred from B_0 using SimBIG. In the diagonal panels we present the marginalized 1D posterior of each parameter. The other panels present the 2D posteriors that illustrate the degeneracies between two parameters. The contours mark the 68 and 95 percentiles. By robustly analyzing B_0 down to nonlinear regimes, $k_{\max} = 0.5\ h/\mathrm{Mpc}$, we place significant constraints on all ΛCDM parameters without any priors from BBN of CMB experiments. Right: We focus on the posteriors of Ω_m and σ_8, the parameters that can be most significantly constrained by galaxy clustering alone. We derive $\Omega_m = 0.293^{+0.027}_{-0.027}$ and $\sigma_8 = 0.783^{+0.040}_{-0.038}$. Our Ω_m and σ_8 constraints are >10 and 50% tighter than the $P_\ell(k < k_{\max} = 0.25\ h/\mathrm{Mpc})$ constraints from a PT approach [11] (orange) and an emulator approach [13] (green). This improvement comes from simultaneously exploiting higher-order and nonlinear cosmological information.

Ω_m and σ_8. We include posteriors from [11] (orange), [13] (green), and SimBIG $B_0(k_{123} < 0.5\ h/\mathrm{Mpc})$ (black). The contours represent the 68 and 95 percentiles of the posteriors. We find overall good agreement among the posteriors. Compared to the P_ℓ constraints, the SimBIG $B_0(k_{123} < 0.3)$ analyses improves σ_8 by $\sim 1.33\times$. The improvement is more modest than the improvement from SimBIG $B_0(k_{123} < 0.5)$ and is broadly consistent with the [24] constraints from analyzing the B_0 to $k_{\max} = 0.23\ h/\mathrm{Mpc}$ and bispectrum quadrupole, B_2, to $k_{\max} = 0.08\ h/\mathrm{Mpc}$. References [25,76] recently found more modest improvements from the bispectrum ($\sim 1.1\times$). They, however, only include the bispectrum monopole and multipoles, respectively, out to $k_{\max} = 0.08\ h/\mathrm{Mpc}$. We refrain from a more detailed comparison since we analyze a subsample of BOSS galaxies. Nevertheless, the comparison illustrates that the B_0 on nonlinear scales contains significant additional cosmological information.

The SimBIG $B_0(k_{123} < 0.5)$ produces significantly tighter cosmological constraints than P_ℓ analyses because we exploit both non-Gaussian and nonlinear cosmological information. For σ_8, the $2\times$ improvement in precision is roughly equivalent to analyzing a galaxy sample with $> 4\times$ the volume using the standard approach. This improvement is made possible by the SimBIG forward modeling approach that is not only able to accurately model galaxy clustering to $k_{\max} = 0.5\ h/\mathrm{Mpc}$ but also robustly account for observational systematics.

Interestingly, the improvements from the SimBIG B_0 analysis enable us to inform recent "cosmic tensions", despite only using 10% of the full BOSS volume. These tensions refer to the discrepancies between the late time and early time measurements of $S_8 = \sigma_8\sqrt{\Omega_m/0.3}$ and the Hubble constant, H_0, that have been growing in statistical significance with recent observations (for a recent review see [77]). They have increased the scrutiny on ΛCDM and have led to a slew of theoretical works to explore modifications or alternatives to ΛCDM (e.g. [78–80]).

For S_8, our SimBIG B_0 constraint $S_8 = 0.774^{+0.056}_{-0.053}$ lies slightly above the constraints from weak lensing (WL) experiments (e.g. [81–86]). We do not find significant tension with either the CMB or WL experiments. Our SimBIG B_0 analysis also places significant constraints on H_0, especially when we combine our posterior with a prior on $\omega_b = \Omega_b/h^2 = 0.02268 \pm 0.00038$ from BBN using importance sampling [87–89]: $H_0 = 67.6^{+2.2}_{-1.8}$. We find a lower value of H_0 that is in good agreement with CMB and other galaxy clustering constraints.

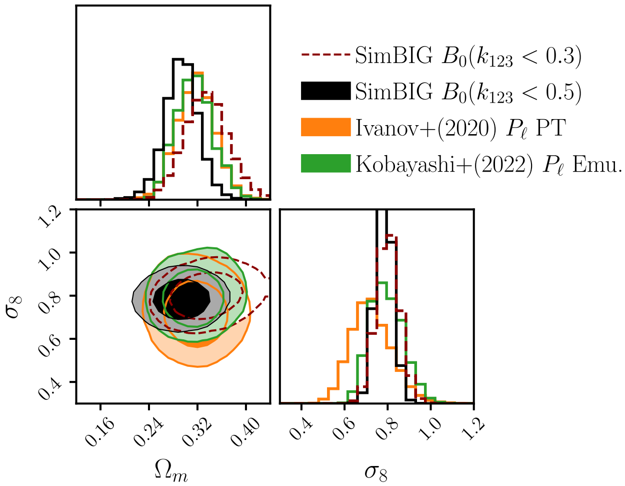

FIG. 7. (Ω_m, σ_8) posterior from theSim BIG B_0 analysis to $k_{\max} = 0.3\ h/\mathrm{Mpc}$ (red dashed). For comparison, we include posteriors from P_ℓ analyses (Ivanov *et al.* [11], orange; Kobayashi *et al.* [13], green) and the SimBIG $B_0(k_{123} < 0.5\ h/\mathrm{Mpc})$ analysis (black). The contours represent the 68 and 95 percentiles. We find overall good agreement among the posteriors. Furthermore, the improvement we find from $B_0(k_{123} < 0.3\ h/\mathrm{Mpc})$ over P_ℓ is consistent with the improvement from B_0 found in the literature (e.g. [24]). Our $B_0(k_{123} < 0.3\ h/\mathrm{Mpc})$ posterior is significantly broader than our $B_0(k_{123} < 0.5\ h/\mathrm{Mpc})$. This demonstrates that there is additional higher-order cosmological information in the nonlinear regime, $0.3 < k < 0.5\ h/\mathrm{Mpc}$, that we can robustly analyze using SIMBIG.

V. DISCUSSION

The SimBIG SBI approach relies on accurate forward modeling of the observed galaxy distributions such that the simulated and observed data are statistically indistinguishable. To achieve this, the SimBIG forward model is designed to be highly flexible and mitigate the impact of model misspecification. It uses N-body simulations that can accurately model the nonlinear matter distribution, a halo finder that robustly determines the position and velocities of dark matter halos, and a highly flexible state-of-the-art HOD.

Despite these modeling choices, the SimBIG forward model does not account for all possible effects that may impact galaxy clustering. For example, it does not include the effect of baryons on the matter clustering. Instead, since it has a subpercent effect on the matter bispectrum at $k < 0.5\ h/\mathrm{Mpc}$ (e.g. [90]), we rely on the HOD model to implicitly account for the impact. Furthermore, we do not include redshift evolution and additional observational systematics (e.g. imaging incompleteness). We refer readers to H23 for a more detailed discussion on the caveats of our forward model.

There are also caveats to our posterior validation for B_0. For instance, the comparison of the $B_0^{(c)}$ likelihoods only demonstrates the robustness near the fiducial cosmologies of the test simulations. Furthermore, some cosmological information may be lost in the q_ϕ-based compression scheme. This would then potentially underestimate the discrepancies in the full B_0 likelihood. Addressing either of these limitations, however, requires a substantially larger suite of simulations evaluated across the full prior space. We reserve developing more stringent and efficient validation of the posterior and summary statistic to future work.

Significant challenges still remain when applying forward modeling approaches to upcoming surveys. They will need to be accompanied by continual improvements to the forward model and validation. There are also challenges in extending SimBIG to the large volumes and the different galaxy samples of upcoming surveys. Nevertheless, in this work we demonstrate the clear advantages of forward modeling: by extracting cosmological information using higher-order statistics and on nonlinear scales we can *double* the precision of σ_8 constraints and significantly improve the constraints of all ΛCDM parameters. In [91], we will present forecasts SimBIG analyses applied to upcoming galaxy surveys: DESI, PFS, and *Euclid*.

VI. SUMMARY

We present the SimBIG cosmological constraints from analyzing the galaxy bispectrum monopole, $B_0(k_1, k_2, k_3)$, on nonlinear scales to $k_{\max} = 0.5\ h/\mathrm{Mpc}$. SimBIG provides a forward modeling framework that uses SBI to perform highly efficient cosmological inference using NDE with normalizing flows (H22a and H23). It enables us to leverage the predictive power of N-body simulations to accurately model higher-order clustering on small scales, which is currently inaccessible with standard PT analyses. It also allows us to more robustly include observational systematics that significantly impact galaxy clustering measurements.

After validating the accuracy and robustness of our analysis using 2,000 test simulations constructed using three different forward models, we conduct the SimBIG $B_0(k_{123} < 0.5\ h/\mathrm{Mpc})$ analysis on a subset of CMASS galaxies in the SGC of SDSS-III BOSS. We derive significant constraints on all ΛCDM parameters $(\Omega_m, \Omega_b, h, n_s, \sigma_8)$ without any external priors. Compared to standard power spectrum analyses, we infer 1.2 and 2.4× tighter constraints on $\Omega_m = 0.293^{+0.027}_{-0.027}$ and $\sigma_8 = 0.783^{+0.040}_{-0.038}$. We verify that this improvement comes from higher-order cosmological information on nonlinear scales and, when restricted to larger scales, our constraints are consistent with previous bispectrum analyses.

In this work, we apply SimBIG to ∼10% of the full BOSS volume due to the limited volume of our N-body simulations. Despite the smaller volume, we derive growth of structure, $S_8 = \sigma_8\sqrt{\Omega_m/0.3}$, constraints competitive with other cosmological probes and BOSS analyses of the full volume. Our $S_8 = 0.774^{+0.056}_{-0.053}$ constraint is statistically consistent with both CMB and weak lensing experiments.

We also derive a constraint on $H_0 = 67.6^{+2.2}_{-1.8}\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ by combining our posterior with a ω_b prior from BBN. Our H_0 constraint is consistent with early universe constraints from CMB and other LSS analyses.

Even with the limited volume of our observations, we derive competitive constraints on S_8 and H_0 by exploiting additional cosmological information in higher-order clustering on nonlinear scales. Extending SimBIG to the full BOSS volume would roughly improve the precision of our constraints by $\sim 3\times$. In an accompanying paper [91], we will present forecasts of SimBIG clustering analyses of upcoming spectroscopic galaxy surveys (e.g. DESI, PFS, *Euclid*) and demonstrate that it has to be potential to produce the leading cosmological constraints from LSS. Reference [91] will also compare the B_0 constraints from this work to SimBIG constraints derived from field-level inference using convolutional neural networks [92] and the wavelet scatter transform [93].

ACKNOWLEDGMENTS

It is a pleasure to thank Mikhail M. Ivanov and Yosuke Kobayashi for providing us with the posteriors used for comparison. We also thank Peter Melchior, Uroš Seljak, and Benjamin D. Wandelt for valuable discussions. This work was supported by the AI Accelerator program of the Schmidt Futures Foundation. J. H. has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 101025187. A. M. D. acknowledges funding from Tomalla Foundation for Research in Gravity.

[1] A. Aghamousa *et al.* (DESI Collaboration), arXiv:1611.00036.
[2] A. Aghamousa *et al.* (DESI Collaboration), arXiv:1611.00037.
[3] B. Abareshi *et al.*, Astron. J. **164**, 207 (2022).
[4] M. Takada *et al.*, Publ. Astron. Soc. Jpn. **66**, R1 (2014).
[5] N. Tamura *et al.*, in Ground-Based and Airborne Instrumentation for Astronomy VI (2016), Vol. **9908**, p. 99081M, arXiv:1608.01075.
[6] R. Laureijs *et al.*, arXiv:1110.3193.
[7] D. Spergel *et al.*, arXiv:1503.03757.
[8] Y. Wang, Z. Zhai, A. Alavi, E. Massara, A. Pisani, A. Benson, C. M. Hirata, L. Samushia, D. H. Weinberg, J. Colbert, O. Doré, T. Eifler, C. Heinrich, S. Ho, E. Krause, N. Padmanabhan, D. Spergel, and H. I. Teplitz, Astrophys. J. **928**, 1 (2022).
[9] F. Beutler *et al.*, Mon. Not. R. Astron. Soc. **466**, 2242 (2017).
[10] G. d'Amico, J. Gleyzes, N. Kokron, K. Markovic, L. Senatore, P. Zhang, F. Beutler, and H. Gil-Marín, J. Cosmol. Astropart. Phys. 05 (2020) 005.
[11] M. M. Ivanov, M. Simonović, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 05 (2020) 042.
[12] S.-F. Chen, Z. Vlah, and M. White, J. Cosmol. Astropart. Phys. 02 (2022) 008.
[13] Y. Kobayashi, T. Nishimichi, M. Takada, and H. Miyatake, Phys. Rev. D **105**, 083517 (2022).
[14] P. Zhang, G. D'Amico, L. Senatore, C. Zhao, and Y. Cai, J. Cosmol. Astropart. Phys. 02 (2022) 036.
[15] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, Phys. Rep. **367**, 1 (2002).
[16] V. Desjacques, D. Jeong, and F. Schmidt, Phys. Rep. **733**, 1 (2018).
[17] A. Banerjee and T. Abel, Mon. Not. R. Astron. Soc. **500**, 5479 (2021).
[18] M. Eickenberg, E. Allys, A. Moradinezhad Dizgah, P. Lemos, E. Massara, M. Abidi, C. Hahn, S. Hassan, B. Regaldo-Saint Blancard, S. Ho, S. Mallat, J. Andén, and F. Villaescusa-Navarro, arXiv:2204.07646.
[19] G. Valogiannis and C. Dvorkin, Phys. Rev. D **105**, 103534 (2022).
[20] K. Naidoo, E. Massara, and O. Lahav, Mon. Not. R. Astron. Soc. **513**, 3596 (2022).
[21] R. Scoccimarro, H. A. Feldman, J. N. Fry, and J. A. Frieman, Astrophys. J. **546**, 652 (2001).
[22] L. Verde *et al.*, Mon. Not. R. Astron. Soc. **335**, 432 (2002).
[23] H. Gil-Marín, W. J. Percival, L. Verde, J. R. Brownstein, C.-H. Chuang, F.-S. Kitaura, S. A. Rodríguez-Torres, and M. D. Olmstead, Mon. Not. R. Astron. Soc. **465**, 1757 (2017).
[24] G. D'Amico, Y. Donath, M. Lewandowski, L. Senatore, and P. Zhang, arXiv:2206.08327.
[25] O. H. E. Philcox and M. M. Ivanov, Phys. Rev. D **105**, 043517 (2022).
[26] E. Sefusatti and R. Scoccimarro, Phys. Rev. D **71**, 063001 (2005).
[27] C. Hahn, F. Villaescusa-Navarro, E. Castorina, and R. Scoccimarro, J. Cosmol. Astropart. Phys. 03 (2020) 040.
[28] C. Hahn and F. Villaescusa-Navarro, J. Cosmol. Astropart. Phys. 04 (2021) 029.
[29] E. Massara, F. Villaescusa-Navarro, S. Ho, N. Dalal, and D. N. Spergel, Phys. Rev. Lett. **126**, 011301 (2021).
[30] D. Gualdi, S. Novell, H. Gil-Marín, and L. Verde, J. Cosmol. Astropart. Phys. 01 (2021) 015.
[31] E. Massara, F. Villaescusa-Navarro, C. Hahn, M. M. Abidi, M. Eickenberg, S. Ho, P. Lemos, A. M. Dizgah, and B. R.-S. Blancard, Astrophys. J. **951**, 70 (2023).
[32] Y. Wang, G.-B. Zhao, K. Koyama, W. J. Percival, R. Takahashi, C. Hikage, H. Gil-Marín, C. Hahn, R. Zhao, W. Zhang, X. Mu, Y. Yu, H.-M. Zhu, and F. Ge, arXiv:2202.05248.
[33] J. Hou, A. Moradinezhad Dizgah, C. Hahn, and E. Massara, J. Cosmol. Astropart. Phys. 03 (2023) 045.
[34] L. Porth, G. M. Bernstein, R. E. Smith, and A. J. Lee, Mon. Not. R. Astron. Soc. **518**, 3344 (2023).

[35] K. Storey-Fisher, J. Tinker, Z. Zhai, J. DeRose, R. H. Wechsler, and A. Banerjee, arXiv:2210.03203.
[36] Z. Zhai, J. L. Tinker, A. Banerjee, J. DeRose, H. Guo, Y.-Y. Mao, S. McLaughlin, K. Storey-Fisher, and R. H. Wechsler, Astrophys. J. **948**, 25 (2023).
[37] A. J. Ross *et al.*, Mon. Not. R. Astron. Soc. **424**, 564 (2012).
[38] A. J. Ross *et al.*, Mon. Not. R. Astron. Soc. **464**, 1168 (2017).
[39] J. H. Yoon, K. Schawinski, Y.-K. Sheen, C. H. Ree, and S. K. Yi, Astrophys. J. Suppl. Ser. **176**, 414 (2008).
[40] H. Guo, I. Zehavi, and Z. Zheng, Astrophys. J. **756**, 127 (2012).
[41] C. Hahn, R. Scoccimarro, M. R. Blanton, J. L. Tinker, and S. A. Rodríguez-Torres, Mon. Not. R. Astron. Soc. **467**, 1940 (2017).
[42] D. Bianchi, A. Burden, W. J. Percival, D. Brooks, R. N. Cahn, J. E. Forero-Romero, M. Levi, A. J. Ross, and G. Tarle, Mon. Not. R. Astron. Soc. **481**, 2338 (2018).
[43] L. Pinol, R. N. Cahn, N. Hand, U. Seljak, and M. White, J. Cosmol. Astropart. Phys. 04 (2017) 008.
[44] A. Smith, J.-h. He, S. Cole, L. Stothert, P. Norberg, C. Baugh, D. Bianchi, M. J. Wilson, D. Brooks, J. E. Forero-Romero, J. Moustakas, W. J. Percival, G. Tarle, and R. H. Wechsler, Mon. Not. R. Astron. Soc. **484**, 1285 (2019).
[45] C. Hahn, M. Eickenberg, S. Ho, J. Hou, P. Lemos, E. Massara, C. Modi, A. M. Dizgah, B. R.-S. Blancard, and M. M. Abidi, Proc. Natl. Acad. Sci. U.S.A. **120**, e2218810120 (2023).
[46] C. Hahn, M. Eickenberg, S. Ho, J. Hou, P. Lemos, E. Massara, C. Modi, A. Moradinezhad Dizgah, B. Régaldo-Saint Blancard, and M. M. Abidi, J. Cosmol. Astropart. Phys. 04 (2023) 010.
[47] K. Cranmer, J. Brehmer, and G. Louppe, Proc. Natl. Acad. Sci. U.S.A. **117**, 30055 (2020).
[48] M. Germain, K. Gregor, I. Murray, and H. Larochelle, Proc. 32nd Int. Conf. Mach. Learn. **37**, 881 (2015).
[49] G. Papamakarios, T. Pavlakou, and I. Murray, arXiv:1705.07057.
[50] D. J. Eisenstein *et al.*, Astron. J. **142**, 72 (2011).
[51] K. S. Dawson *et al.*, Astron. J. **145**, 10 (2013).
[52] J. N. Fry, Phys. Rev. Lett. **73**, 215 (1994).
[53] S. Matarrese, L. Verde, and A. F. Heavens, Mon. Not. R. Astron. Soc. **290**, 651 (1997).
[54] R. Scoccimarro, Astrophys. J. **544**, 597 (2000).
[55] F. Villaescusa-Navarro *et al.*, Astrophys. J. Suppl. Ser. **250**, 2 (2020).
[56] P. S. Behroozi, R. H. Wechsler, and H.-Y. Wu, Astrophys. J. **762**, 109 (2013).
[57] A. A. Berlind and D. H. Weinberg, Astrophys. J. **575**, 587 (2002).
[58] Z. Zheng, A. L. Coil, and I. Zehavi, Astrophys. J. **667**, 760 (2007).
[59] E. G. Tabak and E. Vanden-Eijnden, Commun. Math. Sci. **8**, 217 (2010).
[60] E. G. Tabak and C. V. Turner, Commun. Pure Appl. Math. **66**, 145 (2013).
[61] D. Jimenez Rezende and S. Mohamed, arXiv:1505.05770.
[62] D. S. Greenberg, M. Nonnenmacher, and J. H. Macke, arXiv:1905.07488.
[63] A. Tejero-Cantero, J. Boelts, M. Deistler, J.-M. Lueckmann, C. Durkan, P. J. Gonçalves, D. S. Greenberg, and J. H. Macke, J. Open Source Softwaare **5**, 2505 (2020).
[64] C. Durkan, A. Bekasov, I. Murray, and G. Papamakarios, in *Advances in Neural Information Processing Systems*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., Vancouver, Canada, 2019), Vol. 32.
[65] C. Durkan, A. Bekasov, I. Murray, and G. Papamakarios, nflows: Normalizing flows in PyTorch (2020), 10.5281/zenodo.4296287.
[66] D. P. Kingma and J. Ba, arXiv:1412.6980.
[67] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, arXiv:1907.10902.
[68] B. Lakshminarayanan, A. Pritzel, and C. Blundell, arXiv:1612.01474.
[69] J. Alsing, T. Charnock, S. Feeney, and B. Wandelt, Mon. Not. R. Astron. Soc. **488**, 4440 (2019).
[70] R. Scoccimarro, Phys. Rev. D **92**, 083532 (2015).
[71] H. A. Feldman, N. Kaiser, and J. A. Peacock, Astrophys. J. **426**, 23 (1994).
[72] S. Talts, M. Betancourt, D. Simpson, A. Vehtari, and A. Gelman, arXiv:1804.06788.
[73] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White, Astrophys. J. **292**, 371 (1985).
[74] N. A. Maksimova, L. H. Garrison, D. J. Eisenstein, B. Hadzhiyska, S. Bose, and T. P. Satterthwaite, Mon. Not. R. Astron. Soc. **508**, 4017 (2021).
[75] B. Hadzhiyska, D. Eisenstein, S. Bose, L. H. Garrison, and N. Maksimova, Mon. Not. R. Astron. Soc. **509**, 501 (2022).
[76] M. M. Ivanov, O. H. E. Philcox, G. Cabass, T. Nishimichi, M. Simonović, and M. Zaldarriaga, Phys. Rev. D **107**, 083515 (2023).
[77] E. Abdalla *et al.*, J. High Energy Astrophys. **34**, 49 (2022).
[78] P. D. Meerburg, Phys. Rev. D **90**, 063529 (2014).
[79] A. Chudaykin, D. Gorbunov, and I. Tkachev, Phys. Rev. D **97**, 083508 (2018).
[80] E. Di Valentino, A. Melchiorri, O. Mena, and S. Vagnozzi, Phys. Rev. D **101**, 063502 (2020).
[81] M. Asgari *et al.*, Astron. Astrophys. **645**, A104 (2021).
[82] A. Amon *et al.*, Phys. Rev. D **105**, 023514 (2022).
[83] L. F. Secco *et al.* (DES Collaboration), Phys. Rev. D **105**, 023515 (2022).
[84] R. Dalal *et al.*, Phys. Rev. D **108**, 123519 (2023).
[85] S. Sugiyama *et al.*, Phys. Rev. D **108**, 123521 (2023).
[86] T. M. C. Abbott *et al.* (DES & KiDS Collaborations), Open J. Astrophys. **6**, 2305.17173 (2023).
[87] E. Aver, K. A. Olive, and E. D. Skillman, J. Cosmol. Astropart. Phys. 07 (2015) 011.
[88] R. J. Cooke, M. Pettini, and C. C. Steidel, Astrophys. J. **855**, 102 (2018).
[89] N. Schöneberg, J. Lesgourgues, and D. C. Hooper, J. Cosmol. Astropart. Phys. 10 (2019) 029.
[90] S. Foreman, W. Coulton, F. Villaescusa-Navarro, and A. Barreira, Mon. Not. R. Astron. Soc. **498**, 2887 (2020).
[91] C. Hahn, P. Lemos, R. B, and SimBIG, arXiv:2310.15246.
[92] P. Lemos, L. Parker *et al.* (SimBIG Collaboration), this issue, Phys. Rev. D **109**, 083536 (2024).
[93] B. Régaldo-Saint Blancard, C. Hahn *et al.* (SimBIG Collaboration), following paper, Phys. Rev. D **109**, 083535 (2024).