






# First constraints from marked angular power spectra with Subaru Hyper Suprime-Cam Survey First-Year Data

Jessica A Cowell <sup>1,2,3\*</sup>, Joaquin Armijo <sup>2,3</sup>, Leander Thiele <sup>2,3</sup>, Gabriela A. Marques,<sup>4,5,6</sup> Camila P. Novaes <sup>7</sup>, Daniela Grandón,<sup>8,9</sup> Sihao Cheng <sup>10,11</sup>, Masato Shirasaki,<sup>12,13</sup> David Alonso<sup>1</sup> and Jia Liu<sup>2,3</sup>

<sup>1</sup>Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK

<sup>2</sup>Kavli IPMU (WPI), UTIAS, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8583, Japan

<sup>3</sup>Center for Data-Driven Discovery, Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan

<sup>4</sup>Centro Brasileiro de Pesquisas Físicas, R. Dr. Xavier Sigaud, 150 - Botafogo, Rio de Janeiro, RJ 22290-180, Brazil

<sup>5</sup>Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

<sup>6</sup>Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA

<sup>7</sup>Divisão de Astrofísica, Instituto Nacional de Pesquisas Espaciais, Av. dos Astronautas 1758, Jardim da Granja, São José dos Campos, SP, Brazil

<sup>8</sup>Instituto de Física y Astronomía, Facultad de Ciencias, Universidad de Valparaíso, Avenida Gran Bretaña 1111, Valparaíso, Chile

<sup>9</sup>Mathematical Institute, Leiden University, Gorlaeus Building, Einsteinweg 55, NL-2333 CA Leiden, the Netherlands

<sup>10</sup>Institute for Advanced Study, 1 Einstein Dr., Princeton, NJ 08540, USA

<sup>11</sup>Perimeter Institute for Theoretical Physics, 31 Caroline St N, Waterloo, ON N2L 2Y5, Canada

<sup>12</sup>National Astronomical Observatory of Japan (NAOJ), National Institutes of Natural Sciences, Osawa, Mitaka, Tokyo 181-8588, Japan

<sup>13</sup>The Institute of Statistical Mathematics, Tachikawa, Tokyo 190-8562, Japan

Accepted 2025 January 3. Received 2025 November 26; in original form 2025 July 22

## ABSTRACT

We present the first application of marked power spectra to weak lensing data, using maps from the Subaru Hyper Suprime-Cam Year 1 (HSC-Y1) survey. Marked convergence fields, constructed by weighting the convergence field with non-linear functions of its smoothed version, are designed to encode higher-order information while remaining computationally tractable. Using simulations tailored to the HSC-Y1 data, we test three mark functions that up- or down-weight different density environments. Our results show that combining multiple types of marked auto and cross-spectra improves constraints on the clustering amplitude parameter  $S_8 \equiv \sigma_8 \sqrt{\Omega_m/0.3}$  by  $\approx 43$  per cent compared to standard two-point power spectra. When applied to the HSC-Y1 data, this translates into a constraint on  $S_8 = 0.807 \pm 0.024$ . We assess the sensitivity of the marked power spectra to systematics, including baryonic effects, intrinsic alignment, photometric redshifts, and multiplicative shear bias. We note that some of the additional information introduced by the marked field originates from scales smaller than the scale cut, and is partly Gaussian in nature. This does not invalidate our systematic tests. These results demonstrate the promise of marked statistics as a practical and powerful tool for extracting non-Gaussian information from weak lensing surveys.

**Key words:** gravitational lensing: weak – cosmological parameters – large-scale structure.

## 1 INTRODUCTION

A key challenge faced by existing and upcoming surveys is to extract cosmological information from the complex, non-Gaussian matter field formed through non-linear gravitational evolution. One of the most promising means to study the matter inhomogeneities is weak gravitational lensing, the subtle distortion of light paths caused by the gravitational tidal field of the large-scale structure (see recent reviews by e.g. M. Kilbinger 2015; R. Mandelbaum 2018). The lensing convergence field traces the integrated mass density along the line of sight and is therefore

sensitive to the fundamental cosmological parameters encoded within it.

Most lensing analyses (T. Abbott et al. 2022; R. Dalal et al. 2023; A. H. Wright et al. 2025, for example), rely on two-point statistics, such as the two-point correlation function or the power spectrum, which measure matter clustering as a function of scale. They are optimal for Gaussian fields, whose statistical information is fully captured in the covariance matrix. However, the late-time matter density field is strongly non-Gaussian, motivating the development of new statistics that can recover the non-Gaussian (NG) information missed by two-point statistics, especially for use in future surveys with more precision (e.g. Euclid Collaboration 2023). Examples of NG statistics relevant for weak lensing include but are not limited to: higher order moments (L. F. Secco et al. 2022; M. Gatti et al. 2025), peaks

\* E-mail: [jessica.cowell@physics.ox.ac.uk](mailto:jessica.cowell@physics.ox.ac.uk)

and minima counts (J. Liu et al. 2015; N. Martinet et al. 2018; D. Zürcher et al. 2022; G. A. Marques et al. 2024a; D. Grandón & E. Sellentin 2025), Minkowski functionals (A. Petri et al. 2015), scattering transforms (S. Cheng et al. 2025b; M. Gatti et al. 2025), persistent homology (J. Calles et al. 2024), probability distribution functions (J. Liu & M. S. Madhavacheril 2019; Boyle et al. 2021; L. Thiele et al. 2023), machine learning (J. Fluri et al. 2022), field level inference (A. J. Zhou et al. 2024) among others.

Alternatively, instead of creating new statistics, one could perform a non-linear transformation on the fields before applying the usual two-point statistical machinery, such as Gaussianized fields (D. H. Weinberg 1992; M. C. Neyrinck 2011; M. C. Neyrinck, I. Szapudi & A. S. Szalay 2011), logarithmic density mappings (M. C. Neyrinck, I. Szapudi & A. S. Szalay 2009; H.-J. Seo et al. 2011), and reconstructed density fields (D. J. Eisenstein et al. 2007). By doing so, higher order information is brought into the statistic, however, it is also known that even utilizing all higher order moments of the field misses substantial information (J. Carron 2011) compared to transformed fields on small scales. Our work adopts one such field transformation, known as *marking* – also referred to as marked statistics, the marked power spectra, or the marked correlation function. The marked power spectra have the advantage of containing higher order information while keeping the speed and simplicity of the two-point framework, making their computation much quicker than explicitly calculating bispectra, trispectra, etc. The two-point nature also makes integration into current analysis pipelines easy. This method was first developed for the galaxy overdensity field, where galaxies were weighted by their intrinsic properties before their spatial correlations were computed (D. Stoyan 1984; R. K. Sheth 2005; R. Skibba et al. 2006; U. Sureshkumar et al. 2021). More recently, density-marked statistics, where the field is weighted by a non-linear function of the smoothed density field, have been introduced to enhance constraints on modified gravity (M. White 2016; J. Armijo et al. 2018; 2024a; 2024b; C. Hernández-Aguayo, C. M. Baugh & B. Li 2018; G. Valogiannis & R. Bean 2018; M. Kärcher, J. Bel & S. de la Torre 2024), neutrino masses (E. Massara et al. 2021), and other cosmological parameters (S. Satpathy et al. 2019; A. Aviles et al. 2020; O. H. Philcox, E. Massara & D. N. Spergel 2020; O. H. E. Philcox, A. Aviles & E. Massara 2021; E. Massara et al. 2023; J. A. Cowell, D. Alonso & J. Liu 2024; M. Marinucci et al. 2024; H. Ebina & M. White 2025). The marked power spectrum was recently applied to galaxy clustering data and achieved  $1.2\times$  tighter constraints on  $S_8 \equiv \sigma_8\sqrt{\Omega_m}/0.3$  (E. Massara et al. 2024). Although analytical approaches to the marked power spectrum have been developed using perturbation theory (O. H. Philcox, E. Massara & D. N. Spergel 2020; M. Kärcher et al. 2024; H. Ebina & M. White 2025), they break down at small scales. Simulations offer a natural alternative, enabling accurate modelling of non-linear regimes beyond the reach of perturbative methods. For example, J. A. Cowell et al. (2024) employed Fisher analysis to optimize information extraction using  $N$ -body simulations with and a Gaussian process-based mark function. Simulations also allow us to incorporate small-scale observational effects and survey geometry, which are difficult to model analytically.

In this work, we apply the marked power spectrum for the first time to weak lensing data, using convergence maps derived from the Hyper Suprime-Cam Year 1 (HSC-Y1) survey. This also represents the first application of marked statistics to projected fields. This paper is complemented by a series of papers applying

non-Gaussian statistics to HSC-Y1 using the same observational data and simulations that investigate counts of peaks and minima counts (G. A. Marques et al. 2024a), probability distribution functions (PDFs; L. Thiele et al. 2023), scattering transform coefficients (S. Cheng et al. 2025b), Minkowski functionals (J. Armijo et al. 2025), impact of baryonic feedback on NG statistics (D. Grandón et al. 2024), and a combination of NG statistics with a likelihood-free inference approach (C. P. Novaes et al. 2025).

The structure of this paper is as follows: in Section 2, we describe the theory of marked fields, and our choices of mark functions. In Section 3, we discuss the data, simulations, emulation methods, and systematic tests. We present our results in Section 4 and the main conclusions of this work in Section 5.

## 2 THEORY OF MARKED FIELDS

### 2.1 Mark functions

Density-weighted marks have traditionally been used for three-dimensional density fields, such as the matter or galaxy fields. In this paper, we follow the same spirit but apply marking to the two-dimensional lensing convergence field  $\kappa$ , which is proportional to the matter overdensity integrated along the line of sight. We define the convergence-weighted marked field  $\Delta(\kappa)$  as,

$$\Delta(\kappa) = m(\kappa_\theta)\kappa, \quad (1)$$

where  $m(\kappa_\theta)$  is the mark function and  $\kappa_\theta$  is the convergence field smoothed with a Gaussian filter of width  $\theta$ . In this work, we broadly refer to the angular power spectrum of the marked convergence field  $C_\ell^{\Delta\Delta}$ , the cross-spectrum with the original  $\kappa$  field  $C_\ell^{\Delta\kappa}$ , as well as their combination as the marked power spectrum, or spectra in the case of multiple mark functions.

### 2.2 Choice of mark function

The choice of the mark function remains an open question. While J. A. Cowell et al. (2024) looked to find the optimal mark for cosmological parameters  $\Omega_m$  and  $S_8$  using  $N$ -body simulations at a set redshift, the choice of the function maximizing the information gain may be vastly different depending on the targeted parameter, or even properties of the field itself, such as redshift. For example, applications to studies of modified gravity (M. Kärcher et al. 2024), neutrino mass (E. Massara et al. 2021), and primordial non-Gaussianity (M. Marinucci et al. 2024), would likely need to probe different volumes and scales, and therefore up-weight different regions. Moreover, the choice can be different if one wants to limit oneself to an analytically tractable mark (M. Marinucci et al. 2024; H. Ebina & M. White 2025), or explore options computationally. Moreover, as shown in J. A. Cowell et al. (2024), if all auto and cross-spectra are used, the mark function has an affine symmetry, which allows one to transform two seemingly different marks into one another and still get equivalent constraints.

As a first study on projected fields, we focus on the application, leaving the optimisation of the mark functions for future work. We adopt three distinct mark functions from the recent literature, each with a different weighting profile of the density field:

(A) The Gaussian process (GP) mark introduced in J. A. Cowell et al. (2024), which we denote as  $m_A$ . We use the ‘optimal’ mark shape found for a  $10 h^{-1}$  Mpc smoothing scale, rescaling it, such that the ‘nodes’ are spaced equidistantly between the maximum and minimum values of the smoothed convergence

field. This mark function is shown in yellow in the top left panel of Fig. 1. This mark function is a smooth function between four pre-defined nodes and fitted by a GP – a collection of multivariate normal random variables, fully defined by a mean function (which we set to zero) and a kernel function parametrizing the covariance between any two points. For the kernel, we use the radial basis function, also known as the ‘squared exponential’ kernel (SE), which is specified by only two parameters: an amplitude  $\sigma^2$  and length scale  $\ell$ :

$$k_{\text{SE}}(x, x') = \sigma^2 \exp\left(-\frac{(x-x')^2}{2\ell^2}\right). \quad (2)$$

Using this kernel,<sup>1</sup> the mark function  $m(\kappa_\theta)$  is defined in terms of its values at a set of four nodes at fixed  $\kappa_\theta$  values. We label the node positions  $\kappa_\theta^* \equiv (\kappa_{\theta,1}^*, \dots, \kappa_{\theta,4}^*)$  with corresponding values of mark function  $\mathbf{m}_* \equiv (m(\kappa_{\theta,1}^*), \dots, m(\kappa_{\theta,4}^*))$ . The value of the function at any other  $\kappa_\theta$  is given by the most likely GP conditioned on the known node values

$$m_{\mathcal{A}}(\kappa_\theta) = \mathbf{k}_*^T(\kappa_\theta) \mathbf{K}_*^{-1} \mathbf{m}_*, \quad (3)$$

where the elements of the vector  $\mathbf{k}_*(\kappa_\theta)$  and matrix  $\mathbf{K}_*$  are

$$k_{*,i}(\kappa_\theta) \equiv k_{\text{SE}}(\kappa_\theta, \kappa_{\theta,i}^*), \quad K_{*,ij} \equiv k_{\text{SE}}(\kappa_{\theta,i}^*, \kappa_{\theta,j}^*). \quad (4)$$

The  $\kappa$  position of the four nodes is recalibrated to be equally spaced between the maximum and minimum values of each field, incorporating the different ranges of the  $\kappa$  fields, capturing the dependence on smoothing scales, galaxy shape noise, and redshifts. By doing this, we are effectively introducing slightly different mark functions while preserving the general shape of the function.

(B) The mark given by the smoothed field itself re-weighted by its standard deviation for stability, which we denote as  $m_B$  and show in green in the top left panel of Fig. 1. i.e.

$$m_B = \frac{\kappa_\theta}{\sigma(\kappa_\theta)}. \quad (5)$$

This choice is inspired by the recent work by H. Ebina & M. White (2025), exploiting the fact that this lower-order mark is analytically tractable. It is also intuitive to see the contributions from higher-order terms arising in this mark, for example, the correlation function of the original field and marked field, i.e.  $\langle \kappa(\hat{\mathbf{n}}) \Delta(\hat{\mathbf{n}}') \rangle = \langle \kappa(\hat{\mathbf{n}}) \kappa(\hat{\mathbf{n}}') \kappa_\theta(\hat{\mathbf{n}}') \rangle / \sigma(\kappa_\theta)$ . We expect this to be similar to skew- or kurt-spectra (D. Munshi et al. 2020, 2022), with a difference introduced by the smoothing kernel, which we discuss later.

(C) A modified version of the mark function originally introduced in M. White (2016) in the context of modified gravity, which we denote as  $m_C$  and show in red in the top left panel of Fig. 1,

$$m_C(\kappa_\theta) = (1 + f(x))^{-p}, \quad (6)$$

where

$$x = \frac{\kappa_\theta}{\sigma(\kappa_\theta)(1+b)}, \quad (7)$$

<sup>1</sup>We set the amplitude to be 20 and the length scale to be 0.7. However, these choices are somewhat arbitrary – so long as the length scale is larger than the distance between nodes, one can recover the shape of the mark function.

with the safety function  $f$  defined as

$$f(x) = \begin{cases} x, & \text{if } x \geq -0.999 \\ (1-\delta) \tanh(10x), & \text{otherwise} \end{cases} \quad (8)$$

to avoid singularities. In this analysis, we choose the values  $p = 0.5$ ,  $\delta = 0.02$ , and  $b = 0.1$  for the free parameters of this mark function, closely following the parameters used by E. Massara et al. (2023) and J. Armijo et al. (2024a). This mark function up-weights underdense regions of the field. Compared to the negative power-law up-weighting adopted by M. White (2016), our function up-weights underdense regions with a constant factor, which is more suitable for two-dimensional projected fields.

We show an example of the mark functions and their fields for an example HSC-Y1 sky patch in Fig. 1. In the left column, we show the three marks as a function of the convergence value in the upper panel, the original convergence field  $\kappa$  in the middle panel, and its smoothed version with a smoothing scale  $\theta = 10$  arcmin in the lower panel. In the middle column, we show the marks  $m_i(\kappa_\theta)$  for the three mark functions. In the right column, we show the corresponding marked fields  $\Delta_i$  (equation 1), from which we compute the marked power spectrum.

We can see the intuitive motivation of each mark more clearly in these plots, by looking at the bright regions in the middle and right columns, and where they correspond to in the original  $\kappa$  and  $\kappa_\theta$  maps in the left column. The GP mark,  $m_{\mathcal{A}}$ , anticorrelates low-medium overdensities and underdensities, such that the marked field appears inverted when compared with  $\kappa_\theta$ . Meanwhile,  $m_B$  is proportional to the smoothed field, enhancing the extreme overdensities and underdensities.  $m_C$  up-weights underdense regions, but is equal to a constant for positive  $\kappa$ , such that we expect the information content from these regions of this marked field to have no additional information compared to the standard power spectra. Note that this example is for 10 arcmin smoothing, but we also include down to 2 arcmin smoothing in the baseline analysis, where the small-scale structure and scale of the bright regions are more comparable.

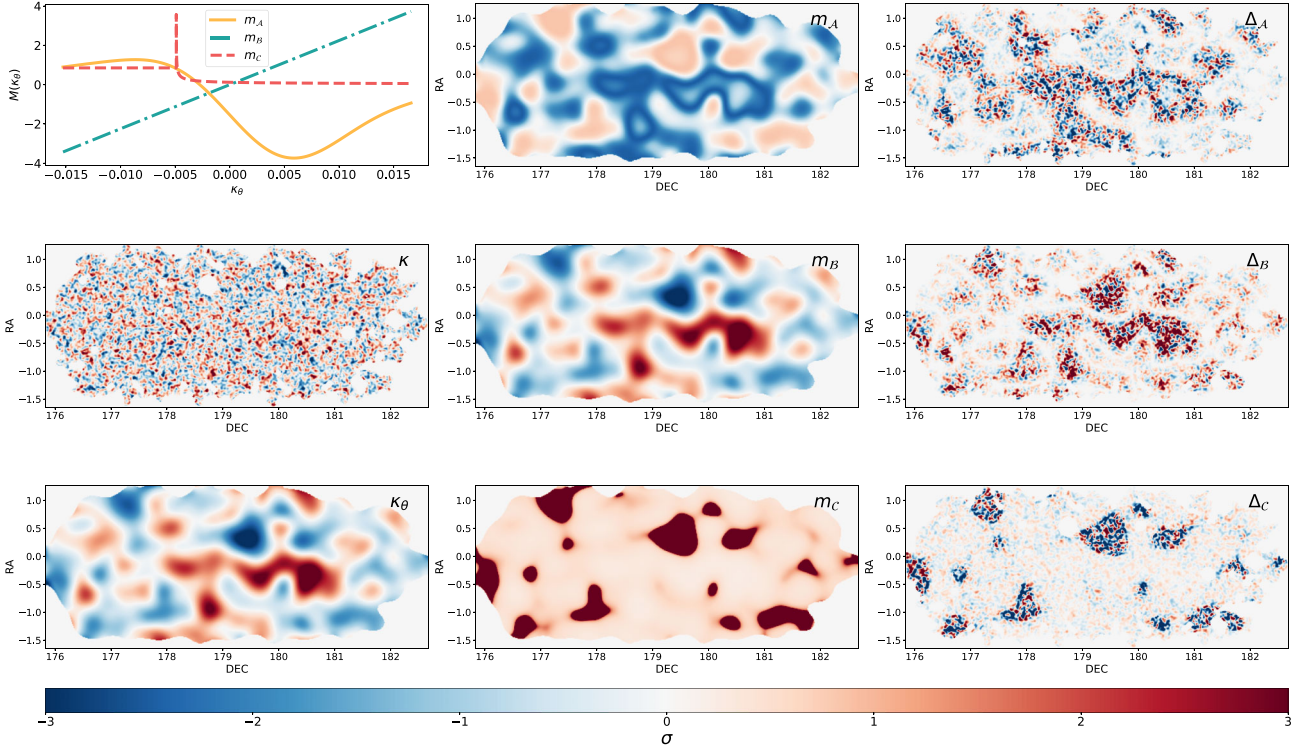
## 3 METHODS

### 3.1 HSC-Y1 data

We use lensing convergence maps estimated from the HSC-Y1 galaxy shapes catalogue (R. Mandelbaum et al. 2017). The data consist of six fields, spanning 136.9 deg<sup>2</sup>, and source redshifts  $0.3 < z < 1.5$ , calculated using the MLZ code (M. Tanaka et al. 2017). After masking, there is a total number density of 17 arcmin<sup>-2</sup> galaxies. Source galaxies are split into four tomographic redshift bins with edges [0.3, 0.6, 0.9, 1.2, 1.5]. Convergence maps are constructed from the shear maps using the Kaiser-Squires (KS) inversion (N. Kaiser & G. Squires 1993), which we repeat for smoothing scales of 1, 2, 4, and 10 arcmin. We also perform inpainting on the maps before performing the KS inversion to avoid  $E$ - $B$  leakage due to the mask. Finally, we reapply the corresponding smoothed masks to the convergence fields. In our analysis, we treat the 1 arcmin smoothed maps as the ‘original’  $\kappa$  maps, and will denote them as  $\kappa$  rather than  $\kappa_{\theta=1}$ .

### 3.2 Simulations

We use two sets of simulations, the ‘covariance set’ and the ‘cosmo-varied’ set. The covariance matrix is estimated



**Figure 1.** We illustrate the process of creating marked fields  $\Delta_i$ , using the example of the convergence field from the HSC-Y1 ‘wide12h’ field with a 10 arcsec smoothing scale in the third redshift bin ( $0.9 < z < 1.2$ ). Left column (top to bottom): example mark functions as a function of the smoothed field  $m_i(\kappa_\theta)$ ; the original convergence field  $\kappa$ , smoothed at 1 arcmin; and the smoothed field  $\kappa_\theta$  for  $\theta = 10$  arcmin. Middle column: mark functions applied to  $\kappa_\theta$ . Right column: marked fields  $\Delta_i = m_i(\kappa_\theta)\kappa$ , used for calculating the marked power spectra. All mark functions and fields are renormalised by their standard deviations for visualization. The colour thus represents fractions of the standard deviation for each field.

from a set of 2268 pseudo-independent realisations generated from 108 full-sky  $N$ -body simulations. They are run with a fiducial cosmology taken from the best-fitting result of the Wilkinson Microwave Anisotropy Probe 9-yr data (G. Hinshaw et al. 2013) ( $\Omega_b = 0.046$ ,  $\Omega_m = 0.279$ ,  $\Omega_\Lambda = 0.721$ ,  $h = 0.7$ ,  $\sigma_8 = 0.82$ ,  $n_s = 0.97$ ), generated from the simulations in R. Takahashi et al. (2017). To model the marked power spectra in lieu of an analytical prediction, we use a separate suite of 100 cosmology-varied simulations presented in M. Shirasaki et al. (2021) and G. A. Marques et al. (2024a). These simulations vary  $\Omega_m$  and  $\sigma_8$ , with 50 quasi-independent realizations for each cosmological model.

The produced mocks incorporate observational effects of the HSC-Y1 catalogue, such as multiplicative bias, inhomogeneity of source galaxies, survey geometry, variations in the lensing weight, redshift, and galaxy distribution. Full details of the method for tailoring the simulations to the HSC-Y1 data as well as adding systematic effects to the simulations, can be found in M. Shirasaki & N. Yoshida (2014), M. Shirasaki et al. (2019), and G. A. Marques et al. (2024a).

### 3.3 Power spectra

The calculation of the marked or standard power spectra is identical, with the only difference being the field itself. We use the flat sky approximation to calculate the power spectra of the fields using the pseudo- $C_\ell$  approach (E. Hivon et al. 2002) as implemented in NAMASTER (D. Alonso, J. Sanchez & A. Slosar 2019), which corrects for the coupling of different Fourier modes due to the mask. We compute power spectra in 14 logarithmically spaced

bins in the range  $80 < \ell < 6500$ , following C. Hikage et al. (2019) and G. A. Marques et al. (2024a). We take the weighted average of the power spectra calculated for each of the six HSC-Y1 fields, where the weights are equal to the number of galaxies per field.

### 3.4 Covariance

Our full data vector for the baseline analysis consists of the  $C_\ell$  measurement of the original field  $\kappa$  map smoothed to 1 arcmin, and the cross- and autospectra of nine marked fields, consisting of three different mark functions and three smoothing choices,  $\theta = [2', 4', 10']$ .<sup>2</sup> Our data vector will then be as follows,

$$\mathbf{D} = \left\{ C_\ell^{\kappa, \kappa}, C_\ell^{\kappa, \Delta_{\mathcal{X}, \theta}}, C_\ell^{\Delta_{\mathcal{X}, \theta}, \Delta_{\mathcal{X}, \theta}} \right\}, \quad (9)$$

where  $\mathcal{X} \in \{A, B, C\}$ ,  $\theta \in \{2', 4', 10'\}$ , and  $\Delta_{\mathcal{X}, \theta} \equiv m_{\mathcal{X}}(\kappa_\theta)\kappa$ . Note that we do not include cross-correlations between separate marks, smoothing scales, or tomographic bins. We use three redshift bins, omitting the fourth bin due to potential issues with the photometric redshift calibration (R. Dalal et al. 2023; X. Li et al. 2023; G. A. Marques et al. 2024b), and 14  $\ell$  bins, leading to a data vector of length 1064 before scale cuts, or 380 after scale cuts (discussed in Section 3.7).

It is crucial to verify that the covariance matrix is accurately estimated, especially for such a large data vector. When inverting

<sup>2</sup>Note, this is the smoothing scale of the *mark function*, such that a smoothing scale of 10 arcmin will still include information from the 1 arcmin field due to the definition of the mark function in equation (1).

the covariance matrix, we apply the Anderson–Hartlap factor (J. Hartlap, P. Simon & P. Schneider 2007) to correct for the finite number of simulations used to calculate the covariance, such that the inverted covariance is  $C_{\text{corrected}}^{-1} = \alpha C^{-1}$ , where  $\alpha = \frac{N_s - N_d - 2}{N_s - 1}$ ,  $N_d = 380$  is the number of data points, and  $N_s = 2268$  is the number of simulations.

We examine the correlation matrix across marks and smoothing scale, and show an example sub-set for cross-spectra in Fig. 2. The first three panels show correlations across smoothing scales for individual marks. The fourth panel shows correlations across the three mark functions. We only show data using the third redshift bin ( $0.9 < z < 1.2$ ) as an example.  $m_B$  displays strong correlations across scales due to its construction (second panel). Negligible correlations between different marks are seen (fourth panel), demonstrating the relatively independent information they carry. This motivates our combined analysis of all three marks.

### 3.5 Emulator

We use a GP to emulate our summary statistics, i.e. the power spectrum and various marked power spectra. We use the GP Regression implemented in SCIKIT-LEARN<sup>3</sup> using a radial basis function kernel with a fixed length scale of 5. We found the emulator accuracy independent of the length scale. Each bandpower is emulated with its own GP Regression, such that points at different bandpowers are effectively treated as independent. The emulator is trained on 100 different cosmologies, using the average (marked) power spectrum over 50 realizations for each cosmology. We evaluate the emulator’s accuracy using cosmology varied simulations in two stages, both employing the ‘leave-one-out’ approach, where the emulator is trained on 99 out of 100 cosmological models, and then used to predict the power spectrum and corresponding marked power spectra for the excluded model.

First, we test the accuracy of the predicted power spectra and marked power spectra for the test model. For each cosmology, we calculate the difference between the predicted power spectra and the true power spectrum, averaging over 50 realizations. For all scales, the emulator is accurate to within  $\sim 10$  per cent standard deviation, where the standard deviation is evaluated across realizations. We also visually inspect these deviations to check that there is no trend with cosmology or scale.

Next, we investigate the accuracy of the inferred cosmology when sampling with an emulated data vector. We show deviations in cosmological parameters  $S_8$  and  $\Omega_m$  using the emulated data vector in comparison to the value inferred using the data vector calculated from the simulation data Fig. 3. In both cases, we use the mean of the posterior. Here, we use the combination of all three marks. The deviations are quantified with respect to the standard deviation of the parameter posterior as calculated using the GETDIST package (A. Lewis 2019). The purple star corresponds to the cosmology of the covariance-training simulations, while the blue box indicates the prior regions used. The points outside the prior are still used to train the emulator. Interestingly, we identify one point located relatively close to the fiducial cosmology used for the covariance matrix estimation that shows a  $1.2\sigma$  bias in  $S_8$ . Although the origin of this bias remains unclear, we verified that including or excluding this point from the emulator has no impact on our final constraint. Our analysis shows a

large deviation on the  $\Omega_m$  value, which is largely due to the fact that  $\Omega_m$  is prior-dominated, giving an artificially smaller standard deviation. Therefore, in our final results, we focus on  $S_8$  instead of  $\Omega_m$ . We refer the reader to Appendix A for more details.

### 3.6 Likelihood estimation

To reduce noise in the covariance matrix, we compress our data vector and covariance matrix using the `moped` compression scheme (A. F. Heavens, R. Jimenez & O. Lahav 2000). This reduces our data vector to only two elements, corresponding to the number of parameters to be inferred. We assume a Gaussian likelihood,

$$\mathcal{L}(\mathbf{D}|\theta_p) = -\frac{1}{2} [\mathbf{D} - \mathbf{x}(\theta_p)]^T \mathbf{C}^{-1} [\mathbf{D} - \mathbf{x}(\theta_p)] + \text{const.}, \quad (10)$$

where  $\mathbf{D}$  is the summary statistic of the data,  $\mathbf{C}^{-1}$  is the inverse of the covariance matrix with the Anderson–Hartlap correction, and the theoretical prediction  $\mathbf{x}(\theta_p)$  is produced by the emulator described in Section 3.5, as a function of the parameters  $\theta_p = [\Omega_m, S_8]$ . We use the `cobaya` Markov chain Monte Carlo (MCMC) sampler (A. Lewis & S. Bridle 2002; J. Torrado & A. Lewis 2021) to sample the posterior, using 10 chains with uniform priors on  $\Omega_m$  of  $[0.1, 0.4]$  and  $S_8$  of  $[0.5, 0.1]$ .

### 3.7 Scale cuts and systematics

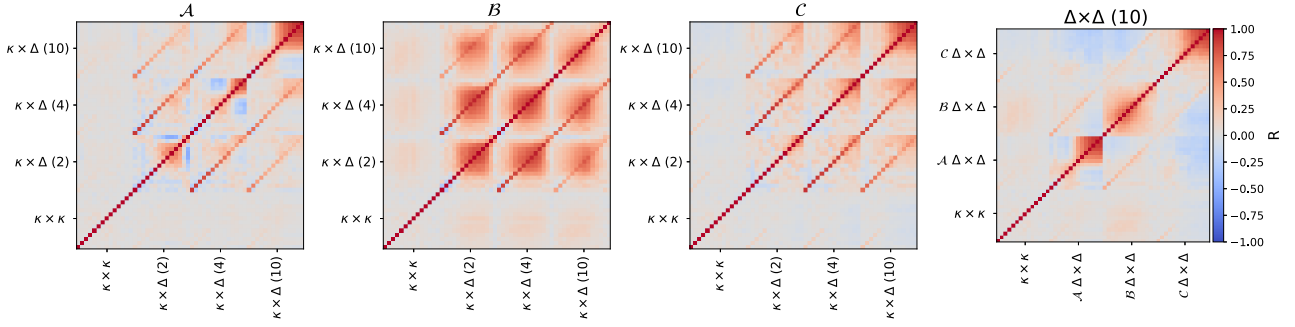
The marked power spectrum is constructed by weighting the local density field with a function of its smoothed counterpart in real space, which inherently mixes different scales in Fourier space. As a result, the scale cuts typically applied to the two-point correlation function may not be appropriate for marked power spectra. Similarly, various systematics may exhibit different scale dependencies compared to the standard power spectrum. We explore the full range of scales in the presence of systematics and provide a more detailed discussion in Section 3.8.

For our baseline analysis, we include all three mark functions and three smoothing scales for the mark function,  $\theta = [2', 4', 10']$ . We select scale cuts such that the induced shift in  $S_8$  due to systematics satisfies  $\Delta S_8 \lesssim 0.4\sigma(S_8)$  across all tested systematics. We find that adopting a maximum multipole of  $\ell_{\text{max}} = 1500$  meets this criterion. We fix  $\ell_{\text{min}} = 300$  based on the unmodelled systematics found in M. Oguri et al. (2017). We present the corresponding inferred values from simulations with systematic contamination in Fig. 4. The top data point of the plot shows the ‘Fiducial’ case, in which no systematics are present. The shaded bands represent the  $\pm 0.4\sigma_{\text{Fid}}$  thresholds for both  $S_8$  and  $\Omega_m$ . Below this, we show the shifts in inferred parameter values resulting from systematic-contaminated data vectors. We consider the analysis robust if these shifted values remain within the shaded bands. We adopt the same systematic tests employed in previous papers in this series, with a detailed description of the contamination procedures provided in G. A. Marques et al. (2024a). For completeness, we also provide a summary of these procedures below.

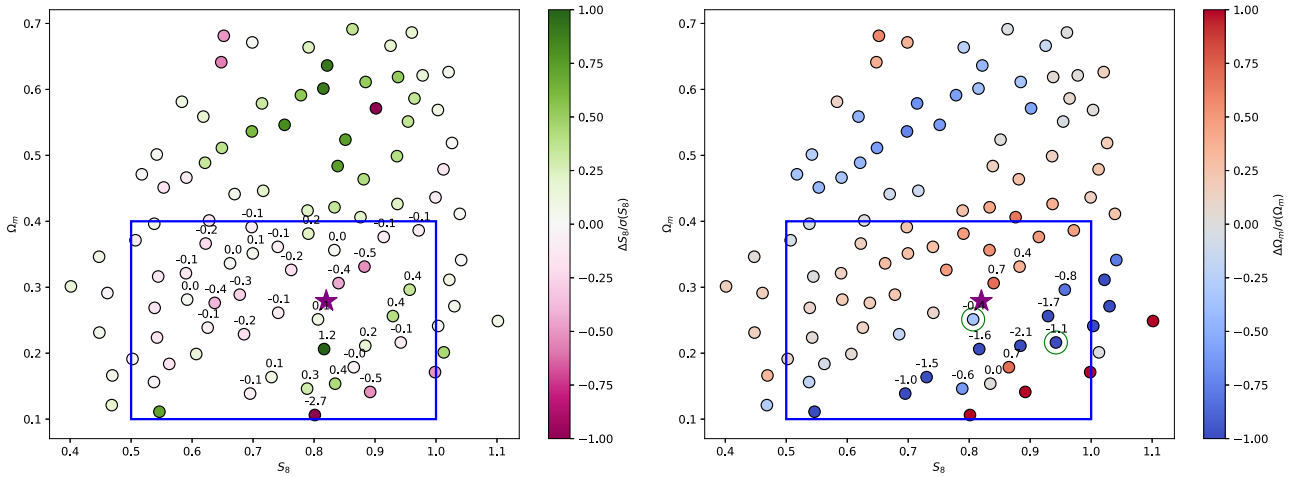
#### 3.7.1 Baryonic effects

Baryonic effects are known to suppress the matter power spectrum on small scales. Since marked power spectra can transfer information between small and large scales, it is essential to test their sensitivity to such effects. To do this, we compute the

<sup>3</sup><https://scikit-learn.org/>



**Figure 2.** Example correlation matrices for sections of the full covariance matrix. In the first three columns, we show the blocks corresponding to the cross spectra  $[C_{\ell}^{\kappa \Delta}]$  of the three marks,  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{C}$ , defined in Section 2.2 for smoothing scales of 2, 4, and 10 arcmin. The standard power spectrum is also shown in the left corner for comparison, labelled  $\kappa \times \kappa$ . On the far right, we show the correlation between these marks for the example of 10 arcmin. All blocks are before scale cuts. All are shown using the third redshift bin ( $0.9 < z < 1.2$ ) as an example.].



**Figure 3.** Accuracy validation using the leave-one-out method. For each of the 100 cosmologies (shown as points), we remove that simulation from the emulator and compare inference between using an emulated data vector or using it as an input data vector to evaluate the shift in the inferred values of  $S_8$  (left) and  $\Omega_m$  (right) relative to the true values. Offsets are expressed in units of the posterior standard deviation. The analysis uses the baseline set-up with all three mark functions and  $\ell_{\max} = 1500$ . The box indicates the prior range; numerical offsets are displayed only for posteriors that remain within this range. Contours shown in green correspond to cases with numerical biases exceeding  $1\sigma$  but that do not appear concerning upon visual inspection.

marked power spectra using both hydrodynamic and dark matter-only simulations from the  $\kappa$ TNG data set (K. Osato, J. Liu & Z. Haiman 2021). We then contaminate the data vector, generated from the ‘covariance training set’ simulations, using the ratio of the hydrodynamic to dark matter-only statistics, and examine the resulting parameter shifts. Overall, we find that baryonic effects introduce minimal biases in the  $S_8$  inference, even when small scales are included, as further discussed in Section 3.8. For our baseline scale cut choice, we see a shift of only  $-0.09\sigma_{\text{Fid}}(S_8)$ . We find the bias on  $\Omega_m$  to be much higher, although still within the statistical uncertainties (approximately  $-0.35\sigma_{\text{Fid}}(\Omega_m)$ ). A related study by D. Grandón et al. (2024) analysed the influence of baryonic feedback on peak counts, minimum counts, the probability distribution function, and the scattering transform in the HSC-Y1 data set. They similarly reported negligible shifts in  $S_8$  across all statistics.

### 3.7.2 Intrinsic alignment

Intrinsic alignments (IA) are a known source of systematic error in weak lensing measurements, arising when galaxy shapes are

influenced by local interactions and become intrinsically correlated to the surrounding large-scale structure. Although many models exist to describe this effect, in this work we focus solely on the non-linear Tidal Alignment (NLA) model (S. Bridle & L. King 2007). In the NLA framework, the strength of intrinsic alignment is controlled by the coupling parameter  $A_{\text{IA}}$ , such that IA-matter and IA-IA power spectra are

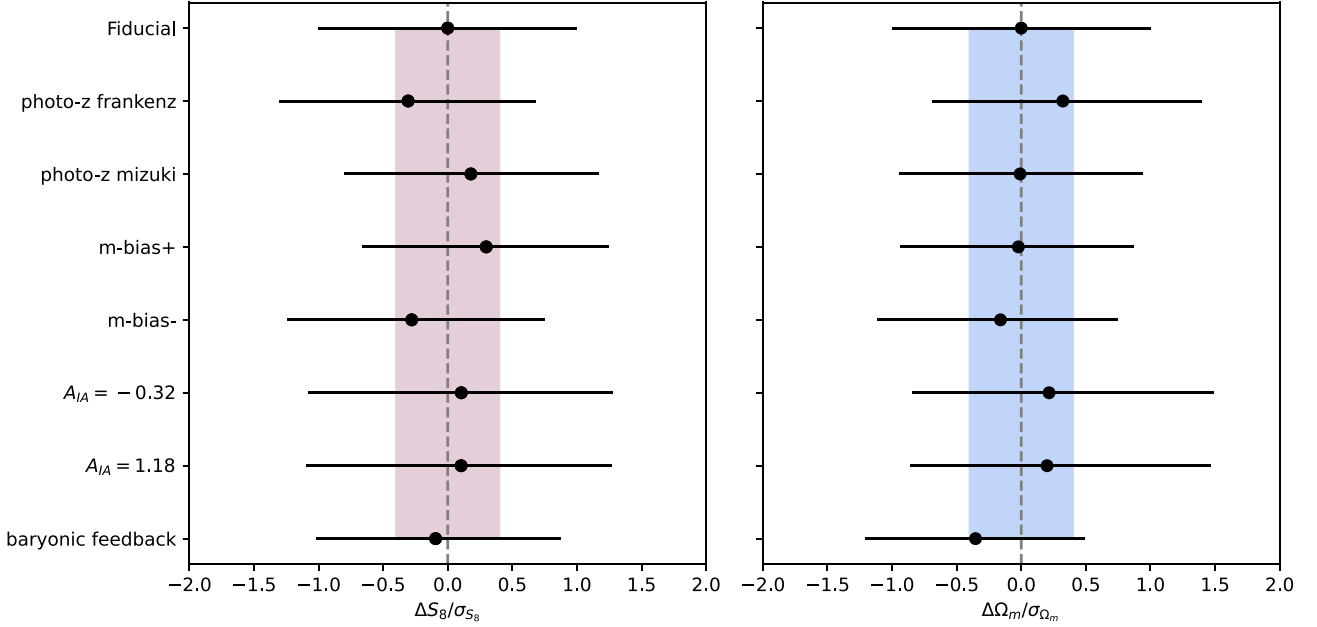
$$P_{\text{Im}}(k, z) = A_{\text{IA}}(z) P_{\text{NL}}(k, z) \quad P_{\text{II}}(k, z) = A_{\text{IA}}^2(z) P_{\text{NL}}(k, z), \quad (11)$$

where  $P_{\text{NL}}(k, z)$  is the non-linear matter power spectrum.

Following previous studies on the HSC-Y1 data, values were reported as  $A_{\text{IA}} = 0.38 \pm 0.7$  (C. Hikage et al. 2019), and  $A_{\text{IA}} = 0.91^{+0.27}_{-0.32}$  (T. Hamana et al. 2020). For this analysis, we choose to test two values of  $A_{\text{IA}} = \{-0.32, 1.18\}$ , corresponding to the extremes of these cases within  $1\sigma$ . We find the intrinsic alignment to give biases of up to  $0.12\sigma_{\text{Fid}}(S_8)$ .

### 3.7.3 Photometric redshifts

Photometric redshift (photo- $z$ ) estimation is an important challenge in cosmological weak lensing, and a potential source of



**Figure 4.** Analysis of the impact of systematic effects on  $S_8$  (left) and  $\Omega_m$  (right). For each contaminated data vector we repeat the inference and recover the error bars shown. The coloured regions show  $1/3\sigma$  confidence interval of the underlying cosmology with no systematics added. We check to see the systematics do not bias the inference outside of this range.

the  $S_8$  tension (A. H. Wright et al. 2025). In HSC, the redshifts are determined using several independent codes, described in M. Tanaka et al. (2017). Our baseline analysis uses source redshifts ranging to  $0.3 < z < 1.5$ , and uses the best-fitting photo-z determined by a machine-learning code based on self-organizing maps (MLZ). We test the effect of using two alternative redshift estimate techniques, using an extra 100 simulations generated with the Flexible Regression over Associated Neighbours with Kernel dEnSity estimation for Redshifts (FRANKEN-Z) and the Mizuki method (M. Tanaka et al. 2017) which utilized theoretical stellar population synthesis models, and using the power spectra from these simulations as input data vector. For the Mizuki method, we find a shift of  $+0.2\sigma_{\text{Fid}}(S_8)$  for the baseline analysis choices. However, based on the scale-cut analysis we find an inconsistent pattern when using the FRANKEN-Z method, for  $\ell < 1000$ , we find a shift of  $-0.75\sigma_{\text{Fid}}(S_8)$ . When including smaller scales for  $\ell < 2500$ , we find up to  $-0.85\sigma_{\text{Fid}}(S_8)$  shifts. However, at the baseline scale choice of  $\ell_{\text{max}} = 1500$ , we observe a shift of only  $-0.31\sigma_{\text{Fid}}(S_8)$ , well within our error budget. This could be an indication of noise from lack of sufficient photo-z simulations.

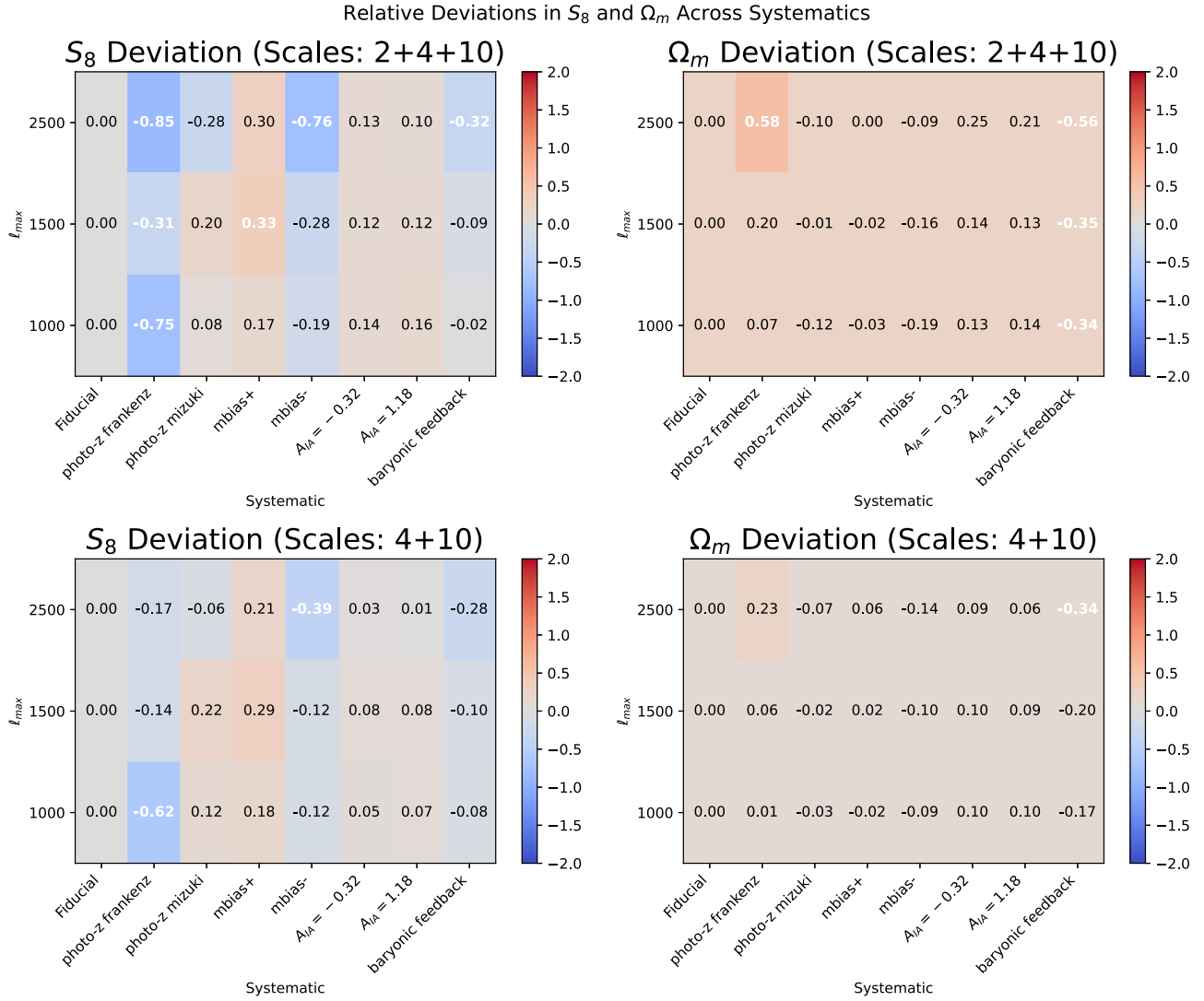
### 3.7.4 Multiplicative bias

Multiplicative bias (*m-bias*) refers to a potential systematic in weak lensing that scales the shear field inferred from galaxy shapes measured in real images with respect to the true underlying lensing field. This is caused by the miscalibration of the galaxy shapes and shape noise bias (L. Miller et al. 2013). The simulation mocks are generated with one multiplicative bias value for each field; however, there are per cent-level uncertainties on the multiplicative bias for HSC-Y1. We account for this by testing the pipeline with a ‘true’ data set a set of realizations generated with *m-bias* purposefully miscalibrated with  $\Delta m = \pm 0.01$ , represented by *m-bias+* and *m-bias-* in Fig. 4. We find shifts of up to  $0.33\sigma_{\text{Fid}}(S_8)$  for our baseline scale cuts. It is worth noting

that, in the case of the  $\mathcal{B}$  mark function, we expect the impact of multiplicative bias can be predicted analytically and easily marginalized over (although we do not do so in this work).

### 3.8 Impact of smoothing and scale cuts on systematic bias

One might expect that including smaller physical scales would amplify the bias of some systematic effects, such as baryonic feedback not included in the main simulations, while others should remain relatively constant. However, as discussed in Section 4.4, applying scale cuts to the marked power spectra does not fully isolate the information from these scales, due to the intrinsic scale mixing introduced by the mark function. More specifically, the smoothing scale used in the mark determines the degree of ‘leakage’ from smaller physical scales, such that information from the smoothing scale size could always be present in the data. In Fig. 5, we illustrate how the inferred cosmological parameters shift in response to different systematic effects as a function of  $\ell_{\text{max}}$ . We examine three smoothing scales used in the baseline analysis, 2, 4, and 10 arcmin, and also compare to the case where the 2 arcmin smoothing is excluded. Each column corresponds to a different systematic, while the y-axis denotes increasing  $\ell_{\text{max}}$  (with  $\ell_{\text{min}} = 300$  fixed). The colour indicates the deviation in the inferred  $S_8$  and  $\Omega_m$  from their fiducial values, in units of their standard deviation  $\Delta\sigma$ . Red denotes an overestimation, blue an underestimation, and white text flags deviations greater than  $0.3\sigma$ . For baryonic effects, we observe the expected behaviour increasing suppression of  $S_8$  as smaller scales are included. However, we also observe some unexpected scale dependence of systematics, such as photo-z. Specifically, on the scale-cut analysis we find an inconsistent pattern when using the FRANKEN-Z method, for  $\ell < 1000$ , we find a shift of  $-0.75\sigma_{\text{Fid}}(S_8)$ . When including smaller scales for  $\ell < 2500$ , we find up to  $-0.85\sigma_{\text{Fid}}(S_8)$  shifts. However, at the baseline scale choice of  $\ell_{\text{max}} = 1500$ , we observe a shift of only  $-0.31\sigma_{\text{Fid}}(S_8)$ , well within our error budget. This



**Figure 5.** xHeatmap displaying the bias introduced by performing inference on contaminated data vectors, compared to the true cosmology. The colour indicates the deviation in the inferred parameters  $S_8$  and  $\Omega_m$ , expressed in units of their standard deviation  $\Delta\sigma$ . Red denotes overestimation, blue underestimation, and white text highlights shifts larger than  $0.3\sigma$ . The top row corresponds to the baseline analysis choices of smoothing scales, while the bottom row removes the 2 arcmin mark functions.

could be an indication of noise from lack of simulations, and will be investigated in future work.

Also, the model with strong intrinsic alignment ( $A_{IA} = 1.18$ ), shows increasing bias with *larger* scale cuts. An important caveat to this figure is that this illustrates the behaviour from **all three marks**, corresponding to our baseline analysis. The behaviour of systematics for each mark also varies, such that statements true of systematic bias for this analysis may not be true in the general case of ‘marked statistics’.

These trends highlight the complex interactions between mark choice, smoothing scale, and systematics. Since the marked statistics mix scales and amplify non-Gaussian information, they may respond to systematics in less predictable ways than standard two-point functions. A detailed analysis of these interactions is left to future work.

## 4 RESULTS

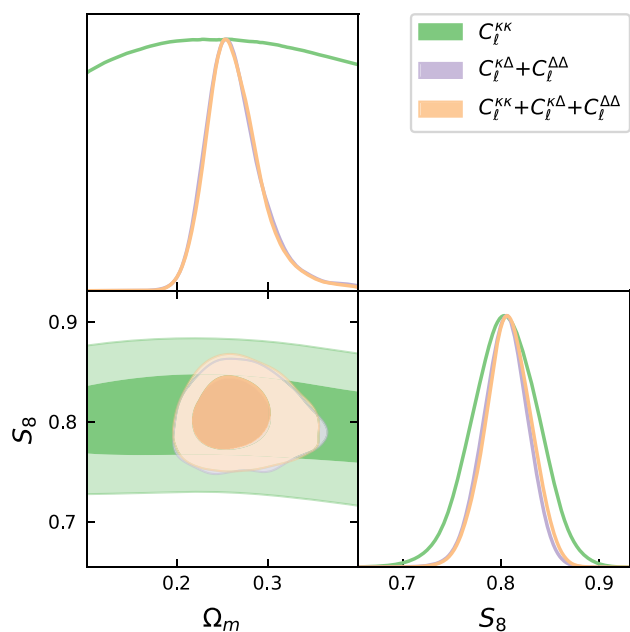
We first present the constraints on  $S_8$  found using our baseline choice inference (Section 4.1). We then discuss the impact of MNRAS **546**, 1–17 (2026)

the choice of mark on these results (Section 4.3). Finally, in Section 4.4 we discuss the information in the cross- and autospectra of the marked field in relation to higher order correlators.

### 4.1 Baseline analysis

For our baseline analysis, we combine all three mark functions and configurations with the scale cut  $300 < \ell < 1500$  (Section 3.7). Although the marked power spectra also contain information about the field on scales outside this range, which we discuss more in Section 4.4. We show results in Fig. 6, where the constraints found using the standard power spectrum,  $C_\ell^{KK}$ , are shown in green, the constraints obtained from the marked spectra ( $C_\ell^{K\Delta} + C_\ell^{\Delta\Delta}$ ) are shown in purple, and the combination of all spectra is shown in orange. The resulting constraints on  $S_8$  are  $S_8 = 0.803 \pm 0.034$ ,  $S_8 = 0.804 \pm 0.023$ , and  $S_8 = 0.807 \pm 0.024$ , respectively.

Our fiducial constraint using HSC-Y1 data is  $S_8 = 0.807 \pm 0.024$ , showing a  $1.4\times$  smaller error compared to that from  $C_\ell^{KK}$  alone using the same scale cuts. We find adding  $C_\ell^{KK}$  does not



**Figure 6.** Constraints on  $S_8$  and  $\Omega_m$  using marked power spectra and HSC-Y1 data, using our baseline analysis choice of combining three mark functions ( $A$ ,  $B$ ,  $C$ ) over three smoothing scales of 2, 4, 10 arcmin, using scale cuts of  $\ell < 1500$ , and no cross-correlations between separate tomographic bins, different mark functions, or smoothing scales. We colour the 68 per cent (inner line) and 95 per cent (outer line) confidence interval contours.  $\kappa$  represents the original convergence field smoothed by 1 arcmin, while  $\Delta$  corresponds to the marked field, such that  $C_\ell^{\kappa\kappa}$  corresponds to the typical power spectrum.

improve our constraint, likely due to the fact that the information in  $C_\ell^{\kappa\kappa}$  is already encoded in  $C_\ell^{\Delta\Delta}$  (see discussions in Section 4.4). We also obtain  $\Omega_m = 0.265^{+0.022}_{-0.035}$ . However, the inferred value of  $\Omega_m$  is affected by imperfect emulator accuracy (Section 3.5). Therefore, caution must be exercised before interpreting this result.

Fig. 7 compares our result to other measurements of  $S_8$ . Cosmic microwave background (CMB)-derived constraints appear at the top of the figure, followed by weak lensing results from two-point statistics, including both power spectra and correlation functions, shown in dark blue. Our result from marked spectra is shown in pink, while in purple we show results from other non-Gaussian analysis of HSC-Y1 in this series of papers. Finally, results from other studies using higher-order statistics are shown in light blue at the bottom. Overall, our constraint lies between previous weak lensing and CMB results, positioned near the centre of the so-called  $S_8$  tension. The constraints are also in agreement with the most recent measurement from the Kilo-Degree Survey (KiDS) Legacy data (A. H. Wright et al. 2025).

#### 4.2 Comparison to other NG stats

The HSC Y1 NG statistics considered here include the Minkowski functionals (MFs; J. Armijo et al. 2025) scattering transform (ST; S. Cheng et al. 2025b), peak and minima counts (G. A. Marques et al. 2024a), the probability density function (PDF; L. Thiele et al. 2023), and an simulation-based inference (SBI) combination of the above, labelled as ‘NG’ (C. P. Novaes et al. 2025). Each statistic probes distinct aspects of the convergence field: the PDF is a one-

point statistic; as are peak counts, which should be sensitive to the **highly non-linear regime**, while minima trace **underdense regions**, similar to the sensitivity of the marked statistic. Meanwhile, the MFs quantify topological features of the field. The ST captures multiscale correlations and is related to higher-order moments of the field. For detailed discussion we refer the readers to the respective papers. Direct comparison between the HSC-Y1 NG statistic constraints is nuanced, as each analysis adopts different scale cuts and analysis choices based on the systematic tests discussed in Section 3.8. For example, the implementation of the ST in S. Cheng et al. (2025b), while yielding the tightest  $S_8$  constraints to date, employs a substantially larger scale cut (up to  $\ell_{\max} = 3500$ ) and includes cross-tomographic correlations, which are not used in this work. Conversely, the other non-Gaussian statistics considered here effectively adopt smaller  $\ell$  cuts due to failing systematic tests. We can, however, contrast their relative performance with respect to the power spectrum  $C_\ell$  under similar scale selections, although this is not a truly direct comparison due to potential smaller scales leaking into the NG statistics from smoothing kernels. Among the NG statistics, the *marked power spectrum* shows the strongest performance of 43 per cent, compared to the two-point power spectra at  $\ell = 1500$ , while ST is only 20 per cent when compared to  $C_\ell$  down to  $\ell = 3500$ . A thorough, controlled comparison of non-Gaussian statistics under identical analysis and map level scale cuts is left for future work.

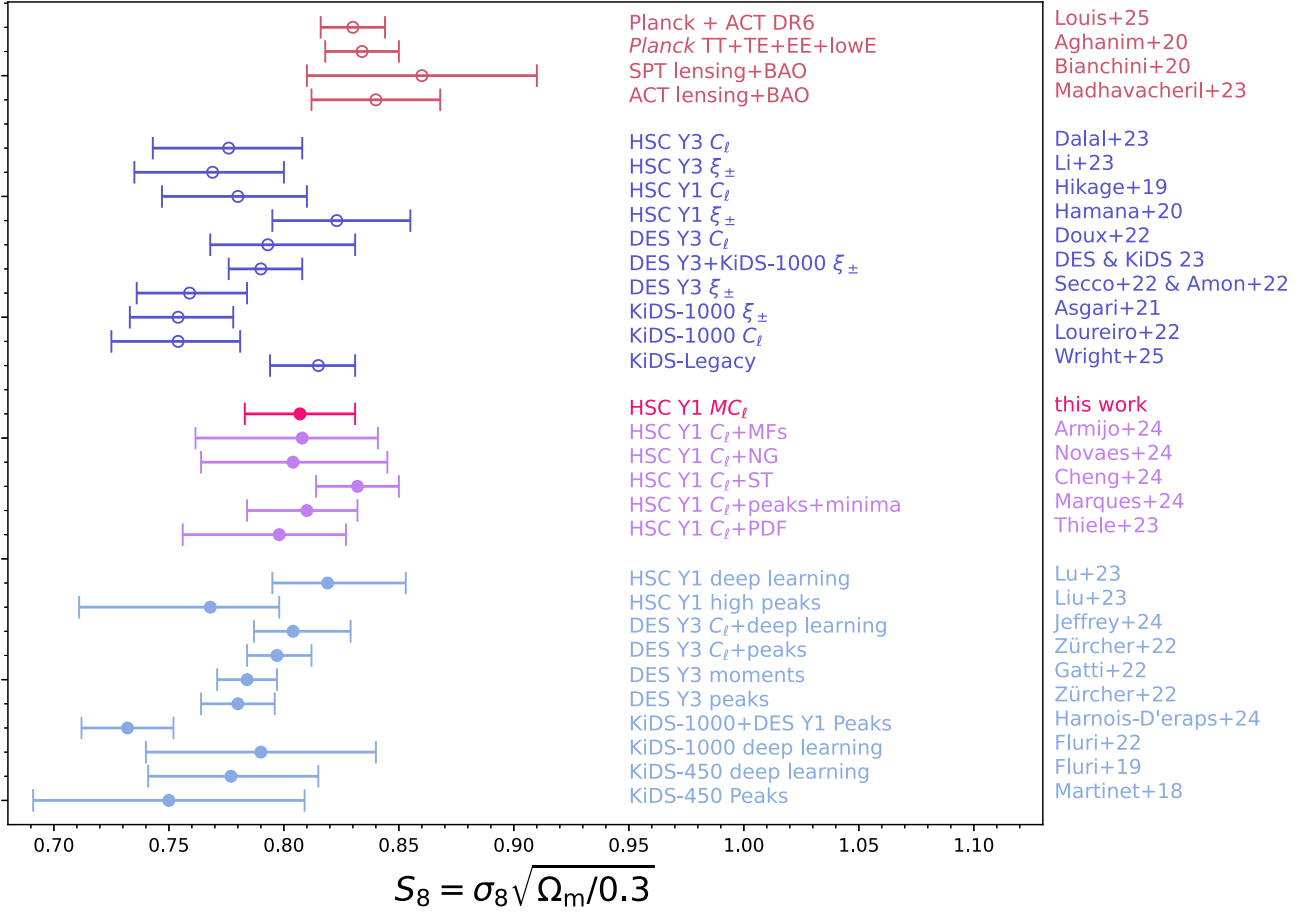
#### 4.3 Comparison of mark functions

We next investigate the constraining power of the individual mark functions and its auto or cross-spectra with the original field. As described in Section 2.2, the three marks have different properties, with  $A$  antiweighting mild over and underdensities,  $B$  multiplying by the smoothed field, up-weighting the more extreme values of the field, and  $C$  up-weighting underdensities.

We show the constraints obtained from HSC-Y1 data for each mark function, including all three smoothing scales, in Fig. 8, showing the combination of all spectra in panel (a), constraints without the power spectrum in (b), and the auto and cross-spectra in (c) and (d), respectively. Mark  $A$  is shown in yellow,  $B$  in red,  $C$  in green, and the combination, representing our baseline analysis choice, in blue. We show the corresponding  $S_8$  errors in Fig. 9, where the improvement in the value of  $\sigma(S_8)$  found from  $C_\ell^{\kappa\kappa}$  in each case is displayed in text above each bar. The horizontal red line shows the constraint from  $C_\ell^{\kappa\kappa}$  at the same scale cuts, such that any result below it corresponds to an improvement.

The results in this figure can be summarized as follows:

- (i) From all three panels, we see the combination of three marks significantly shrinks the contours, and no single mark function can obtain comparable constraints to the combination  $A + B + C$  shown in blue, especially with regard to  $\Omega_m$ .
- (ii) Although in panel (a), which corresponds to the solid bars in Fig. 9, the constraints on  $S_8$  appear visually similar for each mark,  $A$ , in yellow, has the best individual improvement factor of 1.12, compared to 1.04 for the others. Combining all three marks, however, leads to a factor of 1.43 improvement.
- (iii) For  $\Omega_m$  it is clear that  $B$  has the most constraining power, with an error of  $\sigma(\Omega_m) = 0.033$ . Since  $C_\ell^{\kappa\kappa}$  cannot constrain  $\Omega_m$  within our prior range, it is difficult to quantify the improvement factor in this case. In any case, the recovered error is around three times smaller than the width of the prior. We expect the leading



**Figure 7.**  $S_8$  constraints from various surveys and methods. The top group in red corresponds to results from CMB surveys, meanwhile, the blue/purple are all from weak lensing. In purple, we show the results of the other non-Gaussian (NG) statistics using HSC-Y1, with our result in pink. The bottom lighter blue shows constraints from other studies using NG statistics.

information to come from some configurations of the bispectrum and trispectrum, which we discuss further in Section 4.4.

(iv) In panels (c) and (d), corresponding to the hashed bars, we can see the three marks have different degeneracy directions in the  $S_8 - \Omega_m$  plane. Most notably,  $\mathcal{B}$  appears almost perpendicular to  $\mathcal{C}$ . This degeneracy-breaking could explain the constraining power of  $\mathcal{B}$  in  $\Omega_m$ , as well as why its combination with the other marks increases the overall constraining power.

(v) Even without including the standard power spectrum (i.e.,  $C_\ell^{\kappa\kappa}$ ), tight constraints on  $S_8$  can still be achieved, as illustrated in Fig. 9. Notably, only the solid bar includes  $\kappa\kappa$ . However, significant improvement beyond the standard power spectrum is only seen when combining multiple mark functions, as shown by the blue curves. Additionally, the combination of the marked cross- and autospectra for  $\mathcal{B}$  and  $\mathcal{C}$  (dotted lines) yields constraints nearly as tight (an improvement factor of  $\sim 0.99$  when compared with  $C_\ell^{\kappa\kappa}$  alone).

Since each mark weights different regions of the convergence field, it is therefore somewhat intuitive that the constraints from combining these three mark functions are significantly enhanced, similar to how combining power spectra from different cosmological environments can yield tighter constraints (T. Bonnaire et al. 2022). This complementarity, particularly in breaking parameter degeneracies, highlights the strength of the multimark

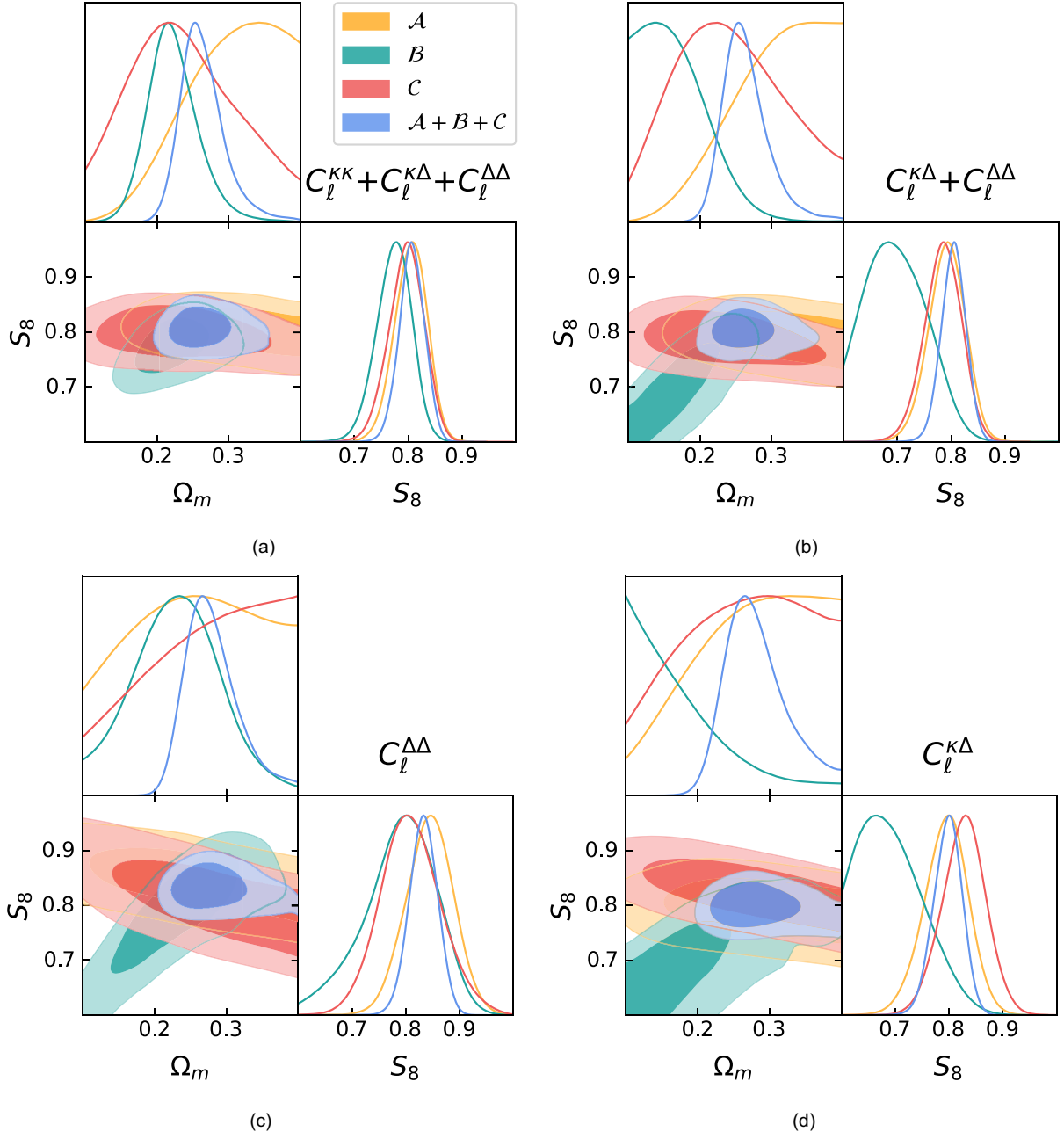
approach and motivates its use in future analyses aimed at extracting maximal information from weak lensing data.

#### 4.4 Relationship to $n$ -point correlation functions

The weak lensing bispectrum contains additional information compared to the power spectrum (M. Takada & B. Jain 2004), and success in tightening constraints of  $S_8$  and  $\Omega_m$  has been found using third-order shear correlation functions or moments from other weak lensing surveys such as KiDS (P. A. Burger et al. 2024) and DES (L. F. Secco et al. 2022; R. C. H. Gomes et al. 2025).

For the  $m_A$  and  $m_C$  marks the direct relation to  $n$ -point correlators is somewhat contrived, however in the case of the simplest mark,  $m_B = \kappa_\theta / \sigma(\kappa_\theta)$ , the cross-spectra  $C_\ell^{\kappa\Delta_B}$  can be expressed purely in terms of the projected bispectrum, and the autocorrelation  $C_\ell^{\Delta_B\Delta_B}$  can be written as a four-point function. In this sense, the mark choice of  $m_B$  is similar in spirit, to skew spectra (D. Munshi et al. 2020) or kurt spectra (D. Munshi et al. 2022), where one takes the correlation function of a polynomial of the field. The marked spectra differ in the introduction of the smoothing kernel, similar to the filtering step applied in L. Harscouet et al. (2025).

We show the parameter constraints derived from each of the spectra in Fig. 10. Panel (a) focuses on the individual components, while panel (b) shows the constraining power of different combi-

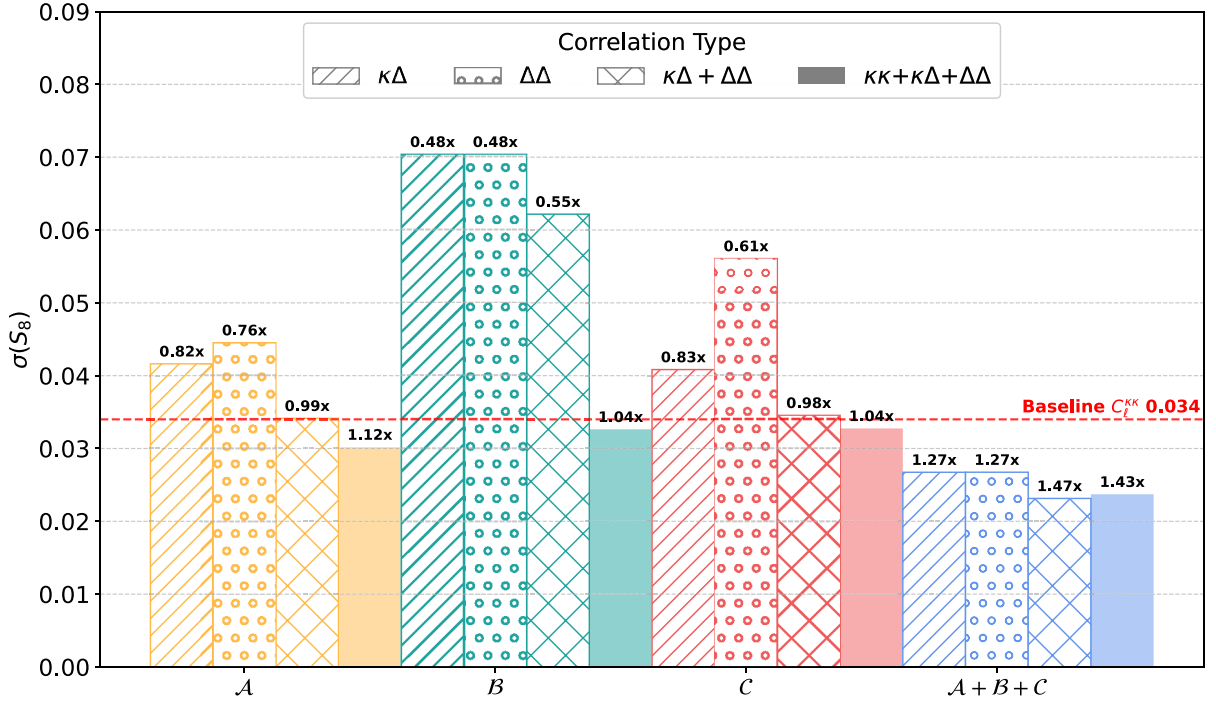


**Figure 8.** Constraints on  $\Omega_m$  and  $S_8 = \sigma_8 \sqrt{\Omega_m}/0.3$  from three different marked power spectra using HSC-Y1 data.  $\Delta$  denotes the marked field, and  $\kappa$  the original convergence field. The three marks emphasise different regions:  $A$  anticorrelates and up-weights medium over and underdensities,  $B$  up-weights overdense regions, particularly the extremes of the field, and  $C$  up-weights only underdense regions. Contours show 68 per cent (inner) and 95 per cent (outer) confidence intervals. The blue contour represents the combined constraints from all three marks. Panel (a) includes all auto and cross-spectra of both the original and marked fields. Panel (b) shows the combination of the auto and cross-marked power spectra *without* the traditional power spectrum, while panels (c) and (d) isolate contributions from the auto and cross-spectra, respectively. The uncertainty constraints from all displayed posteriors are displayed in Fig. 9.

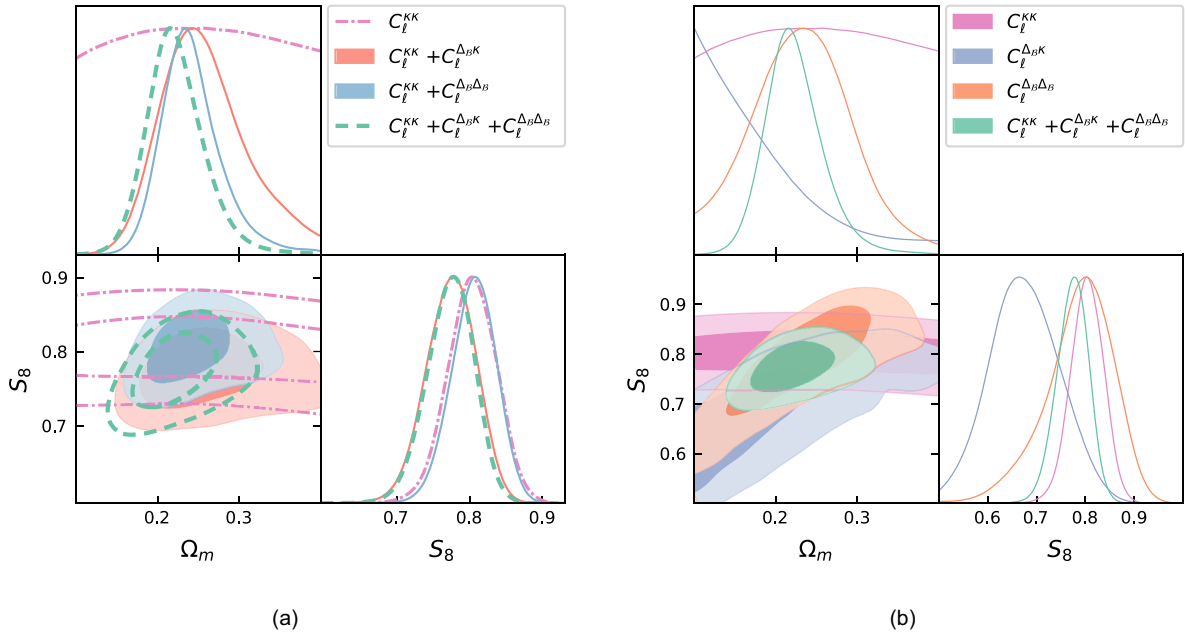
nations. Focussing first on panel (a), the ‘bispectrum-like’ term,  $C_\ell^{\kappa\Delta B}$ , is shown in the blue contour, while the autocorrelation  $C_\ell^{\Delta B\Delta B}$ , is shown in orange. The results obtained from the standard power spectrum  $C_\ell^{\kappa\kappa}$  are shown in pink, and the combination of all three spectra is shown as a green contour. While each term has a slightly different  $S_8 - \Omega_m$  degeneracy direction, we also see that the  $C_\ell^{\kappa\Delta B}$  contour hits the lower bound of the  $\Omega_m$  prior, and

favours lower values of  $S_8$ , leading the autocorrelation to appear more constraining than the cross-correlation.

In panel (b), we look at the combinations of the three spectra. We do not plot the  $C_\ell^{\kappa\Delta B} + C_\ell^{\Delta B\Delta B}$  combination, as it is not particularly constraining for this mark (see Fig. 9). It is interesting to note that the combination  $C_\ell^{\kappa\kappa} + C_\ell^{\Delta\Delta}$ , plotted in purple, leads to better constraints than  $C_\ell^{\kappa\kappa} + C_\ell^{\kappa\Delta}$ , plotted in orange.



**Figure 9.** Posterior uncertainties on  $S_8$  for different analysis configurations. Results are shown for individual mark choices ( $A, B, C$ ; in distinct colours) as well as their combination (blue). Within each configuration, we compare the constraints obtained from the correlation functions of the original convergence field  $\kappa$  and the marked field  $\Delta$ . The baseline  $S_8$  uncertainty from the autopower spectrum of the unmarked field,  $\kappa\kappa$ , is indicated by the horizontal red line. Numbers above each bar denote the relative improvement with respect to this baseline.



**Figure 10.** Posteriors of the different combinations of  $B$  marked power spectra, using the mark function  $m_B = \kappa_\theta$  with the HSC-Y1 data.  $\Delta$  denotes the marked field, while  $\kappa$  is the original convergence field. We colour the 68 per cent (inner line) and 95 per cent (outer line) confidence intervals contours. On the left, we show the individual components, which are the power spectrum  $C_\ell^{\kappa\kappa}$ , the cross-spectra  $C_\ell^{\Delta_B\kappa}$  (bispectrum-like), and autospectra  $C_\ell^{\Delta_B\Delta_B}$  (trispectrum-like) of the marked fields. On the right, we show the constraining power from different combinations of these terms.

One might expect that, since  $C_\ell^{\kappa\Delta}$  is sourced by the bispectrum, it should dominate the improvement on cosmological constraints

with respect to the power spectrum alone. This is because the bispectrum captures the leading-order non-Gaussian

signal, has a higher signal-to-noise ratio than higher-order connected correlators (e.g. the trispectrum), and is less affected by observational noise.

To understand this better, it is therefore useful to expand the marked power spectra in terms of  $n$ -point statistics. For simplicity, we work in the flat sky approximation. Let  $\Delta(\mathbf{x}) \equiv \kappa(\mathbf{x})\kappa_\theta(\mathbf{x})$ , where  $\theta$  donates the smoothing scale.<sup>4</sup> From the convolution theorem, in Fourier space

$$\Delta_\ell = \int \frac{d^2L}{(2\pi)^2} \kappa_{\theta, \mathbf{L}} \kappa_{\ell - \mathbf{L}} = \int \frac{d^2L}{(2\pi)^2} W(L\theta) \kappa_{\mathbf{L}} \kappa_{\ell - \mathbf{L}}, \quad (12)$$

where  $W(L\theta)$  is the smoothing kernel. The cross-spectrum of the original and marked fields is

$$C_\ell^{\kappa\Delta} = \int \frac{d^2L}{(2\pi)^2} W(L\theta) B_\kappa(\ell, L, |\mathbf{l} - \mathbf{L}|). \quad (13)$$

The window function restricts the effective contributions to modes with wavenumber  $L \lesssim 1/\theta$ , such that we get squeezed-like contributions of the form  $B(\ell, 1/\theta, \ell - 1/\theta)$ .

Similarly, the  $\Delta$ - $\Delta$  correlator can be expanded as

$$C_\ell^{\Delta\Delta} = \left[ \underbrace{\int \frac{d^2L}{(2\pi)^2} W^2(L\theta) 2C_{\mathbf{L}}^{\kappa\kappa} C_{|\mathbf{l}-\mathbf{L}|}^{\kappa\kappa}}_{\text{Gaussian part}} + \underbrace{\int \frac{d^2L d^2L'}{(2\pi)^4} W(L\theta) W(L'\theta) T^\kappa(\mathbf{L}, \mathbf{l} - \mathbf{L}, -\mathbf{L}', -\mathbf{l} + \mathbf{L}')}_{\text{non-Gaussian part}} \right], \quad (14)$$

where  $T^\kappa(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \mathbf{l}_4)$  is the connected (non-Gaussian) trispectrum of the convergence field. We can see that, while the second term above is purely non-Gaussian, the first contribution contains only the Gaussian information found in the disconnected trispectrum. As in the case of  $C_\ell^{\kappa\Delta}$ , the range of scales the measurements are sensitive to depend on the smoothing scale  $\theta$ . Specifically, the Gaussian contribution above is sensitive to the power spectrum on scales  $\ell \lesssim \ell_{\text{max}}^{\text{2pt}} + 1/\theta$ , where  $\ell_{\text{max}}^{\text{2pt}} = 1500$  is the scale cut used in our analysis. For a smoothing scale of e.g.  $\theta = 2$  arcmin, using  $\ell \sim 1/\theta \sim 1700$ , this means that the Gaussian contribution to  $C_\ell^{\Delta\Delta}$  is sensitive to the power spectrum on multipoles  $\ell \lesssim 3200$ . This explains the apparent additional constraining power from the marked power spectrum, likely coming from Gaussian fluctuations on smaller scales than available to the measured power spectrum, and not from the field's intrinsic non-Gaussianity. As a consistency test, we verified that the constraints obtained from  $C_\ell^{\Delta\Delta}$  for larger smoothing scales were consistently less tight than those found for  $\theta = 2$  arcmin (e.g. the error on  $S_8$  grows by 121 per cent when using  $\theta = 10$  arcmin).

## 5 CONCLUSIONS

In this work, we present the first application of marked angular power spectra to weak lensing data, using simulated convergence maps tailored to the Subaru Hyper Suprime-Cam Year 1 (HSC-Y1) survey, which is the current survey with the most similar number density of galaxies to Stage-IV surveys such as the Rubin Observatory's Legacy Survey of Space and Time (LSST; The LSST Dark Energy Science Collaboration 2018). We investigated the

performance of three different mark functions, each designed to emphasise different density environments, by evaluating their statistical power in constraining cosmological parameters.

To model the marked statistics as a function of cosmology, we used a Gaussian process emulator trained on a suite of 100 cosmological simulations varying  $\Omega_m$  and  $\sigma_8$ , each with 50 pseudo-independent realisations. The covariance matrix was estimated from 2268 pseudo-independent maps derived from 108 full-sky simulations at a fiducial cosmology, cut and masked to match the HSC-Y1 footprint and galaxy properties. We note that, for simplicity, we did not exploit cross-correlations between fields in different redshift bins, which could significantly tighten the constraints found using both power spectra and marked statistics.

Our key findings are as follows:

(i) **The combination of different marked statistics enhances cosmological constraints (Fig. 8):** the combination of all three mark functions yields significantly tighter constraints on cosmological parameters compared to using individual marks or the standard power spectrum alone. This is particularly true for  $S_8$  and  $\Omega_m$ , with each mark probing different regions of the density field and contributing complementary information.

(ii) **Improved  $S_8$  constraints (Figs 6 and 7):** our fiducial analysis, which includes cross- and autocorrelations of marked fields with smoothing scales of 2, 4, and 10 arcmin, and angular scales in the range  $\ell \in [300, 1500]$ , yields a constraint of  $S_8 \equiv \sigma_8 \sqrt{\Omega_m/0.3} = 0.804 \pm 0.023$ , corresponding to an improvement factor of  $\sim 1.43$  (i.e. 32 per cent smaller uncertainties) over the constraint from  $C_\ell^{\kappa\kappa}$  alone, under identical scale cuts.

(iii) **Information leakage from small scales (Section 4.4):** due to the non-linear nature of the mark construction, marked power spectra exhibit a non-trivial mixing of scales, allowing small-scale information to leak into large-scale modes. We demonstrate that this effect can be mitigated through the choice of smoothing scale and validate our scale cuts ( $\ell_{\text{max}} = 1500$ ) against known systematics. Nevertheless, this implies that a significant fraction of the improvement to the constraining power can be brought about by Gaussian information on smaller scales, rather than purely non-Gaussian data combinations.

(iv) **Robustness to systematics (Fig. 4):** we assess the impact of key observational systematics, including baryonic effects, intrinsic alignments, multiplicative shear bias, and photometric redshift uncertainties, on the inferred cosmology. We find that, with appropriate smoothing choices and scale cuts, the bias in  $S_8$  remains within  $\sim 0.4\sigma$  for all tested systematics.

(v) **Theoretical implications (Section 4.4):** we show that the marked angular power spectra contain contributions analogous to bispectrum and trispectrum configurations, providing a practical and interpretable method of capturing non-Gaussian information without directly computing higher-order correlators.

These results demonstrate that marked power spectra, when applied to weak lensing data, are able to extract significant non-Gaussian information. Combined with the ease with which these statistics can be deployed in real data, this makes them a promising observable to exploit in ongoing and future experiments.

## ACKNOWLEDGEMENTS

JAC was funded by a Kavli/IPMU PhD Studentship. JA was supported by JSPS (Japan Society for the Promotion of Science) KAKENHI Grant Number JP23K19064. DA acknowledges support from the Beecroft trust. MS was supported by JSPS KAKENHI

<sup>4</sup>For simplicity, we leave out the normalization factor  $\frac{1}{\sigma(\kappa_\theta)}$ .

Grant Numbers JP24H00215 and JP24H00221. LT is supported by JSPS KAKENHI Grant 24K22878. CPN thanks Instituto Serapilheira for financial support. JL is supported by JSPS KAKENHI Grant Numbers 23H00107 and 25H00403. The Kavli IPMU was supported by the WPI (World Premier International Research Center) Initiative of the MEXT (Japanese Ministry of Education, Culture, Sports, Science and Technology). D.G. is supported by ANID FONDECYT Iniciación grant 11261860. Numerical computations were (in part) carried out on the analysis servers and the general-purpose PC cluster at the Center for Computational Astrophysics, National Astronomical Observatory of Japan.

## DATA AVAILABILITY

The code and data used for this analysis can be made available upon request to the authors.

## REFERENCES

- Abbott T. et al., 2022, *Phys. Rev. D*, 105, 023520
- Alonso D., Sanchez J., Slosar A., 2019, *MNRAS*, 484, 4127
- Armijo J., Cai Y.-C., Padilla N., Li B., Peacock J. A., 2018, *MNRAS*, 478, 3627
- Armijo J., Baugh C. M., Norberg P., Padilla N. D., 2024a, *MNRAS*, 528, 6631
- Armijo J., Baugh C. M., Norberg P., Padilla N. D., 2024b, *MNRAS*, 529, 2866
- Armijo J., Marques G. A., Novaes C. P., Thiele L., Cowell J. A., Grandón D., Shirasaki M., Liu J., 2025, *MNRAS*, 537, 3553
- Aviles A., Koyama K., Cervantes-Cota J. L., Winther H. A., Li B., 2020, *J. Cosmol. Astropart. Phys.*, 2020, 006
- Bonnaire T., Aghanim N., Kuruvilla J., Decelle A., 2022, *A&A*, 661, A146
- Boyle A., Uhlemann C., Friedrich O., Barthelemy A., Codis S., Bernardeau F., Giocoli C., Baldi M., 2021, *MNRAS*, 505, 2886
- Bridle S., King L., 2007, *New J. Phys.*, 9, 444
- Burger P. A. et al., 2024, *A&A*, 683, A103
- Calles J., Yip J. H. T., Contardo G., Noreña J., Rouhiainen A., Shiu G., 2025, *JCAP*, 2025, 064
- Carron J., 2011, *ApJ*, 738, 86
- Cheng S., Marques G. A., Grandón D., Thiele L., Shirasaki M., Ménard B., Liu J., 2025b, *J. Cosmol. Astropart. Phys.*, 2025, 006
- Cowell J. A., Alonso D., Liu J., 2024, *MNRAS*, 535, 3129
- Dalal R. et al., 2023, *Phys. Rev. D*, 108, 123519
- Ebina H., White M., 2025, *J. Cosmol. Astropart. Phys.*, 2025, 150
- Eisenstein D. J., Seo H.-J., Sirko E., Spergel D. N., 2007, *ApJ*, 664, 675
- Euclid Collaboration, 2023, *A&A*, 675, A120
- Fluri J., Kacprzak T., Lucchi A., Schneider A., Refregier A., Hofmann T., 2022, *Phys. Rev. D*, 105, 083518
- Gatti M. et al., 2025, *Phys. Rev. D*, 111, 063504
- Gomes R. C. H. et al., 2025, *Phys. Rev. D*, 112, 123514
- Grandón D., Sellentin E., 2025, *MNRAS*, 536, 2064
- Grandón D., Marques G. A., Thiele L., Cheng S., Shirasaki M., Liu J., 2024, *Phys. Rev. D*, 110, 103539
- Hamana T. et al., 2020, *PASJ*, 72, 16
- Harscouet L., Cowell J. A., Ereza J., Alonso D., Camacho H., Nicola A., Slosar A., 2025, *Open J. Astrophys.*, 8, 6
- Hartlap J., Simon P., Schneider P., 2007, *A&A*, 464, 399
- Heavens A. F., Jimenez R., Lahav O., 2000, *MNRAS*, 317, 965
- Hernández-Aguayo C., Baugh C. M., Li B., 2018, *MNRAS*, 479, 4824
- Hikage C., Oguri M., Hamana T., More S., Mandelbaum R., Takada e. a., 2019, *PASJ*, 71, 43
- Hinshaw G. et al., 2013, *ApJS*, 208, 19
- Hivon E., Gorski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, *ApJ*, 567, 2
- Kaiser N., Squires G., 1993, *ApJ*, 404, 441
- Kärcher M., Bel J., de la Torre S., 2024, *A&A*, 694, A253
- Kilbinger M., 2015, *Rep. Prog. Phys.*, 78, 086901
- Lewis A., 2019, *J. Cosmol. Astropart. Phys.*, 2025, 025
- Lewis A., Bridle S., 2002, *Phys. Rev. D*, 66, 103511
- Li X. et al., 2023, *Phys. Rev. D*, 108, 123518
- Liu J., Madhavacheril M. S., 2019, *Phys. Rev. D*, 99, 083508
- Liu J., Petri A., Haiman Z., Hui L., Kratochvil J. M., May M., 2015, *Phys. Rev. D*, 91, 063507
- Mandelbaum R., 2018, *ARA&A*, 56, 393
- Mandelbaum R. et al., 2017, *PASJ*, SP1, S25
- Marinucci M. et al., 2024, *J. Cosmol. Astropart. Phys.*, 2025, 036
- Marques G. A. et al., 2024a, *MNRAS*, 528, 4513
- Marques G. A. et al., 2024b, *J. Cosmol. Astropart. Phys.*, 2024, 033
- Martinet N. et al., 2018, *MNRAS*, 474, 712
- Massara E., Villaescusa-Navarro F., Ho S., Dalal N., Spergel D. N., 2021, *Phys. Rev. Lett.*, 126, 011301
- Massara E. et al., 2023, *ApJ*, 951, 70
- Massara E. et al., 2024, *Phys. Rev. D*, 112, 083507
- Miller L. et al., 2013, *MNRAS*, 429, 2858
- Munshi D., Namikawa T., Kitching T. D., McEwen J. D., Bouchet F. R., 2020, *MNRAS*, 498, 6057
- Munshi D., Lee H., Dvorkin C., McEwen J. D., 2022, *J. Cosmol. Astropart. Phys.*, 2022, 020
- Neyrinck M. C., 2011, *ApJ*, 742, 91
- Neyrinck M. C., Szapudi I., Szalay A. S., 2009, *ApJ*, 698, L90
- Neyrinck M. C., Szapudi I., Szalay A. S., 2011, *ApJ*, 731, 116
- Novaes C. P. et al., 2025, *Phys. Rev. D*, 111, 083510
- Oguri M. et al., 2017, *PASJ*, 70, S26
- Osato K., Liu J., Haiman Z., 2021, *MNRAS*, 502, 5593
- Petri A., Liu J., Haiman Z., May M., Hui L., Kratochvil J. M., 2015, *Phys. Rev. D*, 91, 103511
- Philcox O. H., Massara E., Spergel D. N., 2020, *Phys. Rev. D*, 102, 043516
- Philcox O. H. E., Aviles A., Massara E., 2021, *J. Cosmol. Astropart. Phys.*, 2021, 038
- Satpathy S., A C Croft R., Ho S., Li B., 2019, *MNRAS*, 484, 2148
- Secco L. F. et al., 2022, *Phys. Rev. D*, 105, 103537
- Seo H.-J., Sato M., Dodelson S., Jain B., Takada M., 2011, *ApJ*, 729, L11
- Sheth R. K., 2005, *MNRAS*, 364, 796
- Shirasaki M., Yoshida N., 2014, *ApJ*, 786, 43
- Shirasaki M., Hamana T., Takada M., Takahashi R., Miyatake H., 2019, *MNRAS*, 486, 52
- Shirasaki M., Moriwaki K., Oogi T., Yoshida N., Ikeda S., Nishimichi T., 2021, *MNRAS*, 504, 1825
- Skibba R., Sheth R. K., Connolly A. J., Scranton R., 2006, *MNRAS*, 369, 68
- Smith R. E. et al., 2003, *MNRAS*, 341, 1311
- Stoyan D., 1984, *Mathematische Nachrichten*, 116, 197
- Sureshkumar U. et al., 2021, *A&A*, 653, A35
- Takada M., Jain B., 2004, *MNRAS*, 348, 897
- Takahashi R., Hamana T., Shirasaki M., Namikawa T., Nishimichi T., Osato K., Shiroshima K., 2017, *ApJ*, 850, 24
- Tanaka M. et al., 2018, *PASJ*, 70, S9
- The LSST Dark Energy Science Collaboration, 2018, preprint (arXiv:1809.01669)
- Thiele L., Marques G. A., Liu J., Shirasaki M., 2023, *Phys. Rev. D*, 108, 123526
- Torrado J., Lewis A., 2021, *J. Cosmol. Astropart. Phys.*, 2021, 057
- Valogiannis G., Bean R., 2018, *Phys. Rev. D*, 97, 023535
- Weinberg D. H., 1992, *MNRAS*, 254, 315
- White M., 2016, *J. Cosmol. Astropart. Phys.*, 2016, 057
- Wright A. H. et al., 2025, *A&A*, 703, A158
- Zhou A. J., Li X., Dodelson S., Mandelbaum R., 2024, *Phys. Rev. D*, 110, 023539
- Zürcher D. et al., 2022, *MNRAS*, 511, 2075

## APPENDIX A: EMULATOR BIAS

As briefly mentioned in Section 3.5 where we discuss the tests performed on our emulator, when testing the full pipeline by

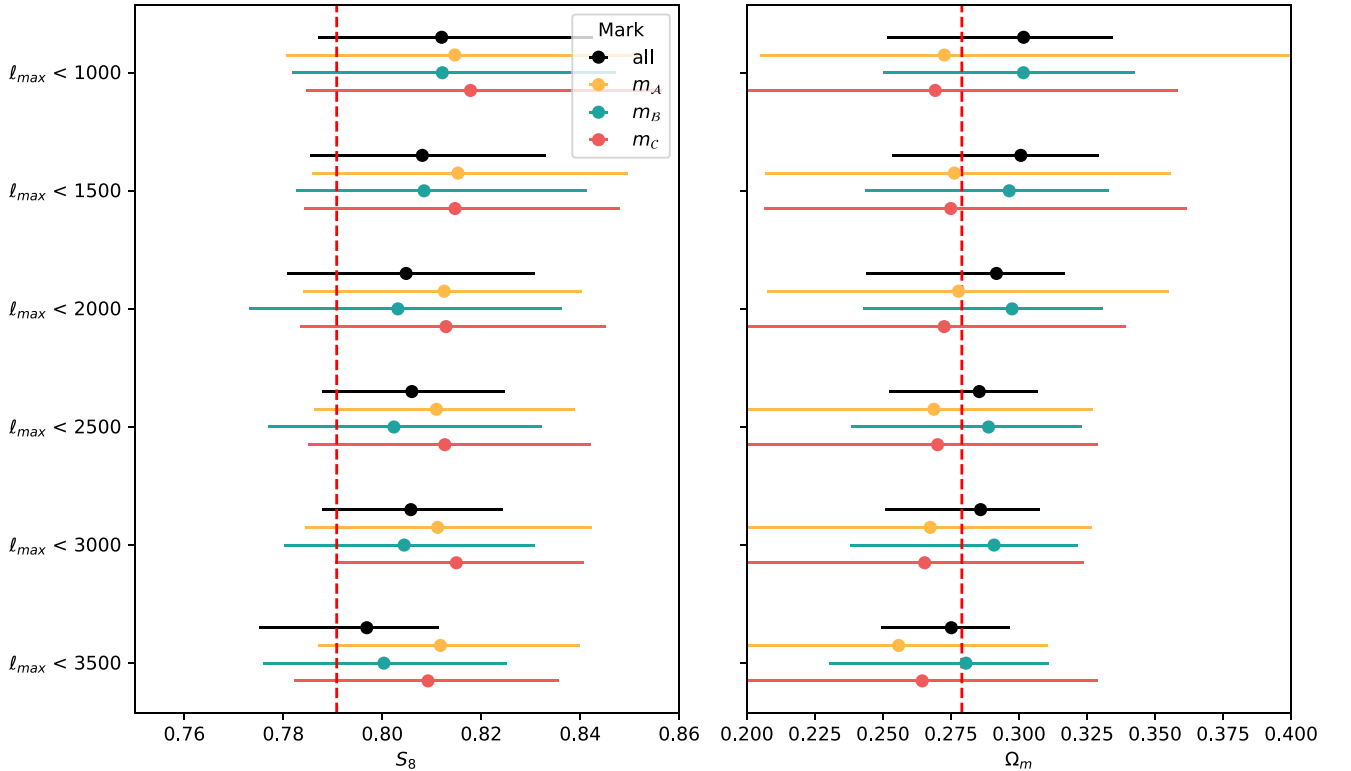
inputting data vectors calculated from simulations with varied cosmology, our analysis shows a shift trend on the  $\Omega_m$  value recovered by the inference, visible in the right panel of Fig. 3. It is worth noting that, in most of our analysis setups,  $\Omega_m$  is prior-dominated, and thus the metric used in this figure does not represent the emulator bias as a fraction of the real statistical uncertainties; a visual inspection of the contours indeed revealed that many of them are indeed prior-dominated, giving an artificially smaller standard deviation. Therefore, for posteriors that hit the prior bounds, we do not display any numerical offset. Based on this visual check, we also mark cosmologies whose posteriors visually contain the true value within  $1\sigma$  with a green circle. For those where the posterior appears truly biased, we show the numerical offset in terms of standard deviation beside the point.

The offsets are significantly higher than those found for  $S_8$ , for which constraints are data-dominated. We find  $m_B$  to give the largest numerical contribution to this bias; however, it also drives the constraining power in  $\Omega_m$ . The cosmology models with the worst bias also seem to be consistent across marks  $m_A$  and  $m_B$ , but  $m_C$  has fewer trends toward overestimating  $\Omega_m$ . We expect this mark to be no more sensitive to small-scale overdensities than the usual power spectrum, which suggests that the source of the discrepancy could be related to small scales; however, this trend is not seen clearly in the residuals of the power spectra themselves. We leave the investigation of the source of such bias to future work. For this reason, we place emphasis on the improved constraining power, rather than the value of  $\Omega_m$  inferred from our analysis.

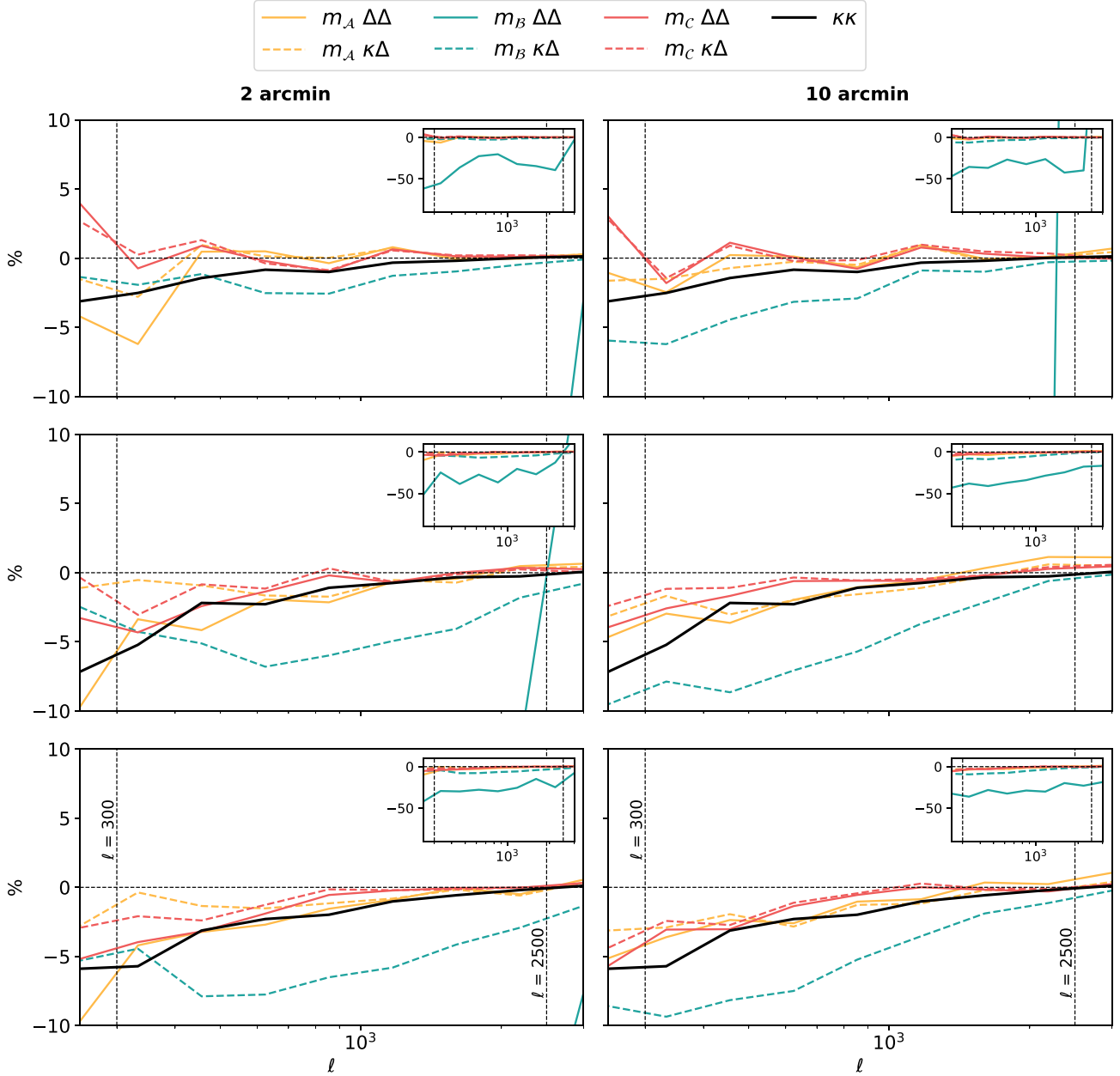
## APPENDIX B: SIMULATION DISCREPANCIES

We observe a systematic offset when inferring the cosmology of the covariance set simulations from (R. Takahashi et al. 2017). Inference is performed using an emulator trained on the cosmo-varied suite, presented in G. A. Marques et al. (2024a), based on the simulations of M. Shirasaki et al. (2021). This is a known phenomenon observed in the other NG HSC-Y1 papers using this analysis technique (L. Thiele et al. 2023; D. Grandón et al. 2024; G. A. Marques et al. 2024a; J. Armijo et al. 2025; S. Cheng et al. 2025b; C. P. Novaes et al. 2025).

Fig. B1 shows the combined posterior constraints for all three mark functions across the three baseline smoothing scales, when inputting the mean of the covariance training simulations as a data vector. In black, we plot the combination of all marks, as used in the baseline analysis. In colour, we plot the results using only one of the mark fields such that the data vector is  $[C_\ell^{KK}, C_\ell^{K\Delta}, C_\ell^{\Delta\Delta}]$ . For all analysis choices,  $S_8$  is overestimated, regardless of the choice of  $\ell_{\max}$ , while  $\Omega_m$  trend varies depending on mark choice. The combined inference remains largely consistent with expectations, though a systematic shift in  $S_8$  persists. This bias is likely caused by a mismatch in the predicted and true power spectra. It is known that for the unmarked power spectrum  $C_\ell^{KK}$ , the cosmo-varied suite has less power than the covariance suite on large physical scales. This effect has been confirmed in prior analyses, and is discussed in the appendix of S. Cheng et al. (2025b). We explore the difference at the power spectrum level in Fig. B2, which shows the residuals between the emulated and true marked power spectra, broken down by mark type, smoothing scale, and redshift bin.



**Figure B1.** Posterior constraints for all three mark functions ( $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{C}$ ) using smoothing scales of 2, 4, and 10 arcmin when inputting a data vector calculated from the covariance-training simulation set, and evaluated using the emulator trained on the cosmo-varied simulations. The true cosmology of the covariance training set is marked in red. The three marks are shown in different colours, with the combination of the three shown in black.



**Figure B2.** Percentage residuals between the power spectra constructed by the emulator trained on the varied cosmology simulations, and true power spectra measured from the covariance training simulations directly. The residuals are defined as  $100 \times (C_\ell^{\text{true}} - C_\ell^{\text{emu}})/C_\ell^{\text{true}}$ , for the covariance training set. We show the residuals for all three mark functions, both cross- and autospectra, and two smoothing scales: 2 and 10 arcmin. Results are shown across all redshift bins. The standard  $C_\ell^{\kappa\kappa}$  spectrum is shown in black.

In black, labelled  $\kappa\kappa$ , we show the standard power spectrum  $C_\ell^{\kappa\kappa}$ , where the lack of power on large scales is visible. The three mark functions exhibit distinct behaviours:

(i) The  $m_A$  mark (yellow), which up-weights and anticorrelates under and overdense regions, shows moderate, scale-dependent deviations, which appear to worsen with increasing redshift.

(ii) The  $m_B$  mark (blue), which overweights overdensities, displays the largest residuals, particularly in the cross-spectrum component shown in the inset, with deviations an order of magnitude larger than other marks. This term is sensitive to some configurations of the squeezed bispectrum. This spectrum, in particular, has very low signal-to-noise, with a large variance

across the covariance simulations. It is possible that the 50 realizations of the cosmo-varied simulations are not enough to accurately predict this spectrum. Alternatively, unlike the autospectrum,  $C_\ell^{\kappa\Delta}$  does not contain information proportional to the two-point function, and at all scales captures small-scale information on the scale of the smoothing scale. However, we see a similar behaviour for 10 and 2 arcmin, suggesting this behaviour is more complicated than just a small-scale effect. A future investigation using the bispectrum applied to this data set could help illuminate this discrepancy.

(iii) The  $m_C$  mark (red), which up-weights underdense regions, shows smoother residuals and is less strongly biased. Interest-

ingly, at lower redshifts it appears to have positive residuals, which switch with increasing redshift.

Interestingly, while  $m_{\mathcal{B}}$  shows the largest spectral residuals, this does not translate into a proportionally large bias in parameter inference (see Fig. B1). This suggests that some of the deviation may cancel out in the compressed data vector, or that the source of the discrepancy could be the low signal to noise, which is accounted for by the covariance matrix.

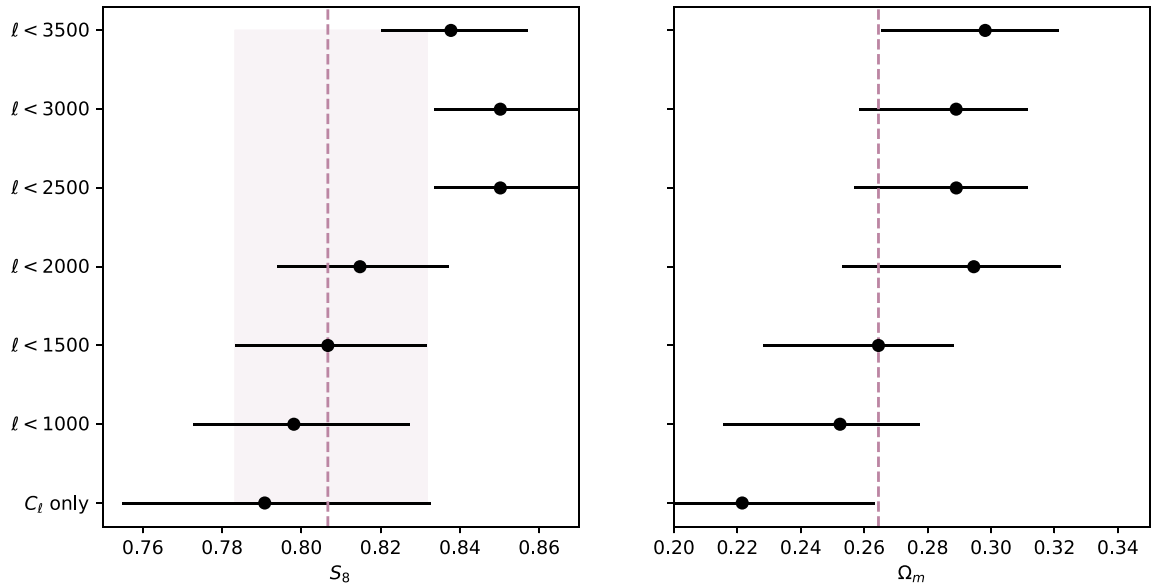
The cause of the discrepancy between the two simulations is not fully understood. The covariance simulations have been validated against HALOFIT (R. E. Smith et al. 2003) predictions and reproduce the convergence power spectrum within 5 per cent up to  $\ell \sim 3000$ , however they are almost affected by finite thickness effects, making their lensing power spectra very large scales smaller than the popular HALOFIT predictions (R. Takahashi et al. 2017). The finite thickness effects are totally simulation-dependent and cannot be corrected with any post processing. The cosmo-varied simulations from M. Shirasaki et al. (2021) have higher mass resolution and are expected to contain more small-scale power. However, the observed deficit in large-scale power is harder to explain. One possibility is the use of adaptive box sizes along the light cone in the cosmo-varied suite, which may introduce subtle inconsistencies across scales.

These findings highlight the importance of consistency between simulation suites when building emulators for non-Gaussian statistics, especially when scale mixing is involved. Further investigation into the discrepancy between the simulation sets is needed to fully interpret these biases, and it may be difficult to find perfect consistency of  $S_8$  inferences among different mocks at this stage.

## APPENDIX C: SCALE DEPENDENCE ON HSC-Y1 DATA

In the analysis of the real HSC-Y1 data, we observe a consistent increase in the inferred value of  $S_8$  as smaller angular scales are included in the data vector. We show the constraints on HSC-Y1 data in Fig. C1, where  $S_8$  rises as higher  $\ell$  modes are added. This trend is not seen when repeating the same analysis on simulated data vectors, as shown in Fig. B1, but is true for both the marked and standard power spectra. In simulations,  $S_8$  estimates remain either consistent or exhibit a weak decrease with increasing  $\ell_{\max}$ . This is true whether the data vector is generated from the covariance suite or from the cosmology-varied simulations.

This discrepancy raises several possible explanations. One is that residual systematics in the HSC-Y1 data, such as unaccounted baryonic effects, intrinsic alignment, or photo- $z$  shifts, could contribute to this scale-dependent bias. However, extensive systematic testing (see Section 3.8) suggests that such effects alone cannot fully account for the observed trend. Another possibility is that the real Universe contains small-scale structures or physics not captured by the current simulations, which rely on gravity-only (or fixed feedback) models. More likely is that the simulations do not have high enough resolution at small scales to accurately model the HSC-Y1 data, leading to a lack of power on small scales. This would cause the real universe to appear to have a higher clustering at smaller scales, boosting the inferred value of  $S_8$ . It is also known that the high convergence regions, which are highly up-weighted by mark  $\mathcal{B}$ , contain rich information about cluster-sized haloes, however the simulations we use likely cannot accurately predict this regimem leading to a significant bias. Nevertheless, our fiducial scale cuts ( $\ell_{\max} = 1500$ ) ensure that this potential bias is kept within the statistical uncertainties. We leave a detailed investigation of this behaviour to future work.



**Figure C1.** Constraints on  $S_8$  and  $\Omega_m$  from HSC-Y1 data using the baseline analysis choices of three marks and three smoothing scales, but varying the scale cuts with  $\ell_{\min} = 300$  fixed. We do not take cross-correlations between different mark models or tomographic bins. We show the constraint from our baseline analysis ( $\ell < 1500$ ), which passed systematic validation for  $S_8$  in pink.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.

© The Author(s) 2026.

Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.