Decoding Cosmological Information in Weak-Lensing Mass Maps with Generative Adversarial Networks

Masato Shirasaki,[1, *] Naoki Yoshida,[2, 3, 4, 5] Shiro Ikeda,[6, 7] Taira Oogi,[3] and Takahiro Nishimichi[8, 3]

[1]*National Astronomical Observatory of Japan (NAOJ), Mitaka, Tokyo 181-8588, Japan*
[2]*Department of Physics, University of Tokyo, Tokyo 113-0033, Japan*
[3]*Kavli Institute for the Physics and Mathematics of the Universe (WPI),*
University of Tokyo, Kashiwa, Chiba 277-8583, Japan
[4]*Institute for Physics of Intelligence, University of Tokyo, Tokyo 113-0033, Japan*
[5]*Research Center for the Early Universe, Faculty of Science, University of Tokyo, Tokyo 113-0033, Japan*
[6]*The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan*
[7]*Department of Statistical Science, Graduate University for Advanced Studies,*
10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
[8]*Center for Gravitational Physics, Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan*
(Dated: December 2, 2019)

Galaxy imaging surveys enable us to map the cosmic matter density field through weak gravitational lensing analysis. The density reconstruction is compromised by a variety of noise originating from observational conditions, galaxy number density fluctuations, and intrinsic galaxy properties. We propose a deep-learning approach based on generative adversarial networks (GANs) to reduce the noise in the weak lensing map under realistic conditions. We perform image-to-image translation using conditional GANs in order to produce noiseless lensing maps using the first-year data of the Subaru Hyper Suprime-Cam (HSC) survey. We train the conditional GANs by using 30000 sets of mock HSC catalogs that directly incorporate observational effects. We show that an ensemble learning method with GANs can reproduce the one-point probability distribution function (PDF) of the lensing convergence map within a $0.5 - 1\sigma$ level. We use the reconstructed PDFs to estimate a cosmological parameter $S_8 = \sigma_8\sqrt{\Omega_{m0}/0.3}$, where Ω_{m0} and σ_8 represent the mean and the scatter in the cosmic matter density. The reconstructed PDFs place tighter constraint, with the statistical uncertainty in S_8 reduced by a factor of 2 compared to the noisy PDF. This is equivalent to increasing the survey area by 4 without denoising by GANs. Finally, we apply our denoising method to the first-year HSC data, to place 2σ-level cosmological constraints of $S_8 < 0.777\,(\mathrm{stat}) + 0.105\,(\mathrm{sys})$ and $S_8 < 0.633\,(\mathrm{stat}) + 0.114\,(\mathrm{sys})$ for the noisy and denoised data, respectively.

I. INTRODUCTION

Impressive progress has been made in observational cosmology in the past decades. An array of multi-wavelength astronomical data have established the standard cosmological model of our universe, referred to as ΛCDM model, with precise determination of major cosmological parameters. The nature of the main energy contents in our universe remains unknown, however. Invisible mass component called dark matter is needed to explain the formation of large-scale structures in the universe [1, 2], and an exotic form in energy appears to be responsible for the accelerating expansion of the late-time universe [3].

In order to reveal the nature of dark matter and the late-time cosmic acceleration, a number of astronomical surveys for detailed observations of large-scale structures are ongoing. Accurate measurement of cosmic lensing shear signals is one of the primary goals aimed at in the ongoing galaxy imaging surveys such as the Kilo-Degree Survey (KiDS[1]), the Dark Energy Survey (DES[2]), and the Subaru Hyper Suprime-Cam Survey (HSC[3]), even for upcoming projects including the Wide Field Infrared Survey Telescope (WFIRST[4]), the Large Synoptic Survey Telescope (LSST[5]), and Euclid[6].

The large-scale matter distribution can be reconstructed through measurements of lensing shear signals by collecting a large set of galaxy images [4–6]. Although the image distortion of individual galaxy shapes is usually tiny, averaging over many galaxies makes it possible to infer the distribution of underlying dark matter in an unbiased way in principle. However, there are well known challenges when extracting the rich cosmological information from the reconstructed weak lens-

* masato.shirasaki@nao.ac.jp

[1] http://kids.strw.leidenuniv.nl/index.php
[2] https://www.darkenergysurvey.org/
[3] http://hsc.mtk.nao.ac.jp/ssp/
[4] https://wfirst.gsfc.nasa.gov/
[5] https://www.lsst.org/
[6] http://sci.esa.int/euclid/

ing maps. The non-linear gravitational growth of large-scale structures makes statistical properties of the weak lensing signals complicated. Numerical simulations have shown that simple statistics for random Gaussian fields are not sufficient to fully describe the statistical information imprinted in weak lensing maps [7–9]. To utilize the so-called non-Gaussian information, various approaches have been proposed [10–26], but no single statistic can capture the full information, unfortunately.

In practice, the observed weak lensing maps are largely contaminated by noises arising from intrinsic galaxy properties and observational conditions. The former is commonly called as shape noise, which contaminates the original physical information. It is known that the overall effect of the noise contamination can be estimated and mitigated for some Gaussian statistics [27, 28], but it is not understood quantitatively how the shape noise affects non-Gaussian statistics. Non-Gaussian information can potentially be a powerful probe to validate the ΛCDM model and test a variety of extended cosmological models even in the presence of shape noises [29–31]. Clearly, it is necessary to devise a noise reduction method in order to maximize the science returns from wide-field lensing surveys.

A straightforward method to reducing the shape noise is to smooth weak lensing maps over a large angular scale (e.g. $\sim 20-30$ arcmins in Refs. [32, 33]), but the smoothing also erases the non-Gaussian information in the map [7, 34]. A novel approach has been proposed to keep a high angular resolution of $\sim$ 1arcmin and yet preserves non-Gaussian information [35]. The method is based on a deep-learning framework called conditional generative adversarial networks (GANs) [36]. Thanks to the rich expressive power of deep neural networks, conditional GANs enable to denoise the weak lensing map on a pixel-by-pixel basis (see also Ref. [37] for a similar attempt with a deep learning method).

In this paper, we construct and test conditional GANs that can be applied to the real galaxy imaging data taken by the Subaru HSC survey [38]. Using a large set of mock HSC catalogs [39], we train the GANs in a realistic situation and test the denoising capability by our GANs using 1000 test data. We assess several possible systematic errors in the denoising process, and investigate the cosmological dependence on the denoised map by making the best use of numerical simulations of gravitational lensing. We then show how well our GANs denoise the real HSC data and improve cosmological parameter constraints.

The rest of the present paper is organized as follows. In Section II, we summarize the basics of gravitational lensing. Section III describes the HSC data as well as our numerical simulations used for training and testing GANs. In Section IV, we provide the analysis methods including the details of our training strategy of GANs and the likelihood analysis for testing the cosmological model. In Section V, we show the results of our denoised map for the HSC data and the gain of our GAN-based denoising for the cosmological constraints. Concluding remarks and discussions are given in Section VI.

II. WEAK GRAVITATIONAL LENSING

A. Basics

We first summarize the basics of gravitational lensing induced by the large-scale structure. Weak gravitational lensing effect is characterized by the distortion of the image of a source object by the following 2D matrix:

$$A_{ij} = \frac{\partial \beta^i}{\partial \theta^j} \equiv \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}, \quad (1)$$

where $\boldsymbol{\theta}$ represents the observed position of a source object, $\boldsymbol{\beta}$ is the true position, κ is the convergence, and γ is the shear. In the weak lensing regime (i.e., $\kappa, \gamma \ll 1$), each component of A_{ij} can be related to the second derivative of the gravitational potential Φ [40]. Using the Poisson equation and the Born approximation, one can express the weak lensing convergence field as the weighted integral of matter overdensity field $\delta_{\mathrm{m}}(\boldsymbol{x})$:

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_H} \mathrm{d}\chi \, q(\chi) \delta_{\mathrm{m}}(\chi, r(\chi)\boldsymbol{\theta}), \quad (2)$$

where χ is the comoving distance, χ_H is the comoving distance up to $z \to \infty$ and $q(\chi)$ is called lensing kernel. For a given redshift distribution of source galaxies, the lensing kernel is expressed as

$$q(\chi) = \frac{3}{2} \left(\frac{H_0}{c} \right)^2 \Omega_{\mathrm{m}0} \frac{r(\chi)}{a(\chi)} \int_\chi^{\chi_H} \mathrm{d}\chi' p(\chi') \frac{r(\chi' - \chi)}{r(\chi')}, \quad (3)$$

where $r(\chi)$ is the angular diameter distance and $p(\chi)$ represents the redshift distribution of source galaxies normalized to $\int_0^{\chi_H} \mathrm{d}\chi \, p(\chi) = 1$.

B. Smoothed lensing convergence map

In optical imaging surveys, galaxies' shapes (ellipticities) are commonly used to estimate the shear component γ in Eq. (1). Since each component in the tensor A_{ij} is given by the second derivative of the gravitational potential, one can reconstruct the convergence field from the observed shear, in Fourier space, as

$$\hat{\kappa}(\boldsymbol{\ell}) = \frac{\ell_1^2 - \ell_2^2}{\ell_1^2 + \ell_2^2} \hat{\gamma}_1(\boldsymbol{\ell}) + \frac{2\ell_1 \ell_2}{\ell_1^2 + \ell_2^2} \hat{\gamma}_2(\boldsymbol{\ell}), \quad (4)$$

where $\hat{\kappa}$ and $\hat{\gamma}$ are the convergence and shear in Fourier space, and $\boldsymbol{\ell}$ is the wave vector with components ℓ_1 and ℓ_2 [5].

For a given source galaxy, one considers the relation between the observed ellipticity $\epsilon_{\mathrm{obs},\alpha}$ and the expected

shear $\tilde{\gamma}_\alpha$,

$$\tilde{\gamma}_\alpha = \frac{\epsilon_{\text{obs},\alpha}}{2\mathcal{R}}, \qquad (5)$$

$$\tilde{\gamma}_\alpha = (1 + m_\text{b})\gamma_{\text{true},\alpha} + c_\alpha, \qquad (6)$$

where $\mathcal{R}$ is the conversion factor to represent the response of the distortion of the galaxy image to a small shear [41], $\gamma_{\text{true},\alpha}$ is the true value of cosmic shear, and m_b and c_α are the multiplicative and additive biases to assess possible systematic uncertainty in galaxy shape measurements. In practice, before employing the conversion in Eq. (4), one must first construct the smoothed shear field on grids [42],

$$\gamma_{\text{grid},\alpha}(\boldsymbol{\theta}) = \frac{\sum_{i\in\boldsymbol{\theta}} w_i \left(\epsilon_{i,\text{obs},\alpha}/2\mathcal{R} - c_{i,\alpha}\right)}{(1 + \langle m_\text{b}\rangle) \sum_{i\in\boldsymbol{\theta}} w_i}, \qquad (7)$$

$$\langle m_\text{b}\rangle = \frac{\sum_{i\in\text{all}} w_i m_{\text{b},i}}{\sum_{i\in\text{all}} w_i}, \qquad (8)$$

$$\gamma_{\text{sm},\alpha}(\boldsymbol{\theta}) = \int \mathrm{d}^2\phi \, \gamma_{\text{grid},\alpha}(\boldsymbol{\phi})W(\boldsymbol{\phi} - \boldsymbol{\theta}) \qquad (9)$$

where $\boldsymbol{\theta}_i$ is the position of the i-th source galaxy, w_i represents the inverse variance weight, and $W(\boldsymbol{\theta})$ is a smoothing filer. In the above, $\sum_{i\in\boldsymbol{\theta}}$ represents the summation over the galaxies in the pixel at the angular coordinate $\boldsymbol{\theta}$, while $\sum_{i\in\text{all}}$ is the sum over all the galaxies in our survey window. In this paper, we assume the functional form for W as

$$W(\boldsymbol{\theta}) = \frac{1}{\pi\theta^2}\left[1 - \left(1 + \frac{\theta^2}{\theta_s^2}\right)\exp\left(-\frac{\theta^2}{\theta_s^2}\right)\right], \quad (10)$$

for $\theta \leq 10\,\theta_s$ and $W(\boldsymbol{\theta}) = 0$ otherwise. We set $\theta_s = 3\,\text{arcmins}$ throughout this paper[7]. Using Eqs. (4) and (9), one can derive the smoothed convergence field from the observed imaging data through Fast Fourier Transform (FFT).

Apart from the systematic uncertainty by shape measurements, the observed ellipticity can be expressed as a sum of two term in practice:

$$\epsilon_{\text{obs}} = 2\mathcal{R}\,\gamma + \epsilon_\text{N}, \qquad (11)$$

where γ is the lensing shear of interest and ϵ_N represents noise that originates from the intrinsic galaxy shape and from observational conditions, referred to as shape noise. Accordingly, we have two components in the observed lensing map as

$$\kappa_{\text{obs}} = \kappa_{\text{WL}} + \kappa_\text{N}. \qquad (12)$$

---

[7] The smoothing scale θ_s is commonly adopted to search for massive galaxy clusters in a smoothed lensing map [43]. Using numerical simulations, Ref. [44] has found a one-to-one correspondence between the peaks on a smoothed map by the filter in Eq. (10) and massive galaxy clusters at $z = 0.1 - 0.3$ when imposing the peak height to be larger than $\sim 5\sigma$.

The shape noise is much larger than the lensing shear term for individual objects in typical galaxy imaging surveys. Hence, the observed map κ_{obs} is significantly contaminated by the shape noise on a pixel-by-pixel basis. This fact makes it challenging to extract the cosmological information contained in the map. Our objective in this paper is to estimate the noiseless field κ_{WL} from the observed (noisy) map κ_{obs}. For this purpose, we adopt a framework of conditional generative adversarial networks (GANs).

III. DATA

A. Subaru Hyper Suprime-Cam Survey

Hyper Suprime-Cam (HSC) is a wide-field imaging camera installed at the prime focus of the 8.2-meter Subaru telescope [38, 45–48]. The Wide Layer in the HSC survey will cover 1400 deg^2 in five broad photometric bands ($grizy$) in its 5-year operation, with superb image quality of sub-arcsec seeing. In this paper, we use a galaxy shape catalog that has been produced for cosmological weak lensing analysis in the first year data release (HSC S16A hereafter). Details of the galaxy shape measurements and catalog information are found in Ref. [49].

In brief, the HSC S16A galaxy shape catalog is based on the HSC Wide-Layer data taken from March 2014 to April 2016 over 90 nights. We apply the same set of galaxies as in Ref. [49] to construct a secure shape catalog for weak lensing analysis. The sky areas around bright stars are masked [50]. The HSC S16A weak lensing shear catalog covers 136.9 deg^2 that consists of 6 disjoint patches: XMM, GAMA09H, GAMA15H, HECTOMAP, VVDS, and WIDE12H. Among these 6 patches, we pay special attention to the XMM in this paper because there exist publicly available catalogs of galaxy clusters in optical [51] and X-ray bands [52] in the XMM patch. We use these cluster catalogs to examine the reliability of our denoising process, i.e., we test the accuracy by performing object-by-object matching.

In the HSC S16A shape catalog, the shapes of galaxies are estimated using the re-Gaussianization PSF correction method applied to the i-band coadded images [53]. In the XMM region, the survey window is defined such that 1) the number of visits within `HEALPix` pixels with `NSIDE=1024` to be $(g, r, i, z, y) \geq (4, 4, 4, 6, 6)$ and the i-band limiting magnitude to be greater than 25.6, 2) the PSF modeling is sufficiently good to meet our requirements on PSF model size residuals and residual shear correlation functions, 3) there are no disconnected `HEALPix` pixels after the cut 1) and 2), and 4) the galaxies do not lie within the bright object masks. For details of defining these masks, see Ref. [49].

The redshift distribution of source galaxies is estimated from the HSC five broadband photometry. Ref. [54] measured photometric redshifts (photo-z's) of galaxies in the HSC survey by using several different methods. Among
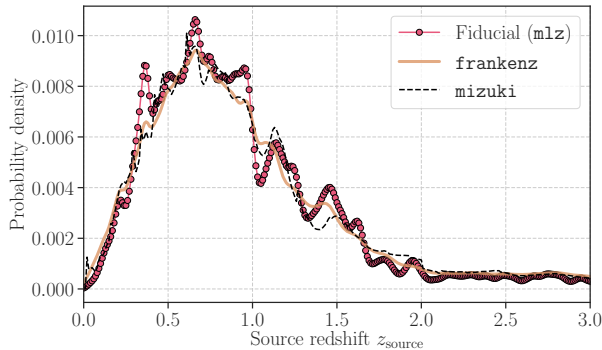
FIG. 1. The stacked photometric redshift distribution for the galaxies in the XMM field. The line with points shows the estimate by our baseline method, while the yellow solid and black dashed lines stand for the results by `frankenz` and `mizuki`, respectively.

them, we choose the photo-z with a machine-learning code based on self-organizing map (`mlz`) as a baseline. To study the impact of photo-z estimation with different methods, we consider two additional photo-z's estimated from a classical template-fitting code (`mizuki`) and a hybrid code combining machine learning with template fitting (`frankenz`). For our analysis, we select the source galaxies by their *best* estimates (see Ref. [54]) of the photo-z's (z^{best}) in the redshift range from 0.3 to 1.5 as done in the main cosmological analyses for the HSC S16A data [55]. For a given method of the photo-z estimation, individual HSC galaxies are assigned a posterior probability distribution function (PDF) of redshift. Figure 1 shows the stacked PDFs for the source galaxies in the XMM. The mean source redshifts are found to be 0.96, 1.01, and 1.01 for the estimates by `mlz`, `mizuki`, and `frankenz` , respectively.

We then reconstruct the smoothed convergence field from the HSC S16A data as described in Section II B. Adopting a flat-sky approximation, we first create a pixelized shear map for the XMM on regular grids with a grid size of 1.5 arcmins. We then apply FFT and perform convolution in Fourier space to obtain the smoothed convergence field. Note that we limit the maximum number of grids on a side to be 256 in our analysis. Currently, it is still computationally expensive to train GANs with large-size images with decent computer resources (see Ref. [56] for a recent attempt). Since our aim here is to analyze lensing convergence maps with an arcmin resolution, the pixel size is set to be $\sim 1\,\text{arcmin}$. We will analyze observational data for a larger region in our future work. Our survey window covers the range of $[30.9, 37.3]\,\text{deg}$ and $[-7.29, -0.89]\,\text{deg}$ in right ascension (RA) and declination (dec), respectively. There are 1345810 source galaxies available with photo-z estimates.

In actual observations, there are missing galaxy shear

data due to bright star masks. The observed regions have also complex geometry. Applying our method directly to such regions likely generates additional noises [57]. We determine the mask regions for each convergence map by using the smoothed number density map of the input galaxies with the same smoothing kernel as in Eq. (10). Then we mask all the pixels with the smoothed galaxy number density less than 0.5 times the mean number density. After masking, the data region is found to cover 21.4 deg^2.

B. Mock HSC observations

We use a large set of simulation data for training our conditional GANs. Table I summarizes our mock simulations.

1. Fiducial simulations

We first describe the mock shape catalogs for HSC S16A. The mock catalogs are generated from 108 full-sky lensing simulations presented in Ref. [59][8]. In Ref. [59], the authors perform a suite of cosmological N-body simulations with 2048^3 particles and generate lensing convergence maps and halo catalogs. The N-body simulations assume the standard ΛCDM cosmology consistent with the 9-year WMAP cosmology [58] with the CDM density parameter $\Omega_{\text{cdm}} = 0.233$, the baryon density $\Omega_{\text{b0}} = 0.046$, the matter density $\Omega_{\text{m0}} = \Omega_{\text{cdm}} + \Omega_{\text{b0}} = 0.279$, the cosmological constant $\Omega_{\Lambda} = 0.721$, the Hubble parameter $h = 0.7$, the amplitude of density fluctuations $\sigma_8 = 0.82$, and the spectral index $n_s = 0.97$. The gravitational lensing effect is simulated with the multiple lens-plane algorithm on a curved sky [44, 60]. Light-ray deflection is directly followed by using the projected matter density field produced by the outputs from the N-body simulations. Each lensing simulation data consists of 38 different source planes at redshift less than 5.3. Realistic source redshift distributions are implemented following the curves in Figure 1.

To generate mock shape catalogs, we employ essentially the same method as developed in Refs. [20, 61]. We use the full-sky simulations combined with the observed photometric redshifts and angular positions of real galaxies. Provided the real catalog of source galaxies, where each galaxy contains information on the position (RA and Dec), shape, redshift, and the lensing weight, we perform the following four-steps:

(i): Set the RA and Dec of the survey window in the full-sky realization.

Name	# of realizations	Cosmology	Note	Reference
Fiducial	2268	WMAP9 cosmology [58]	Photo-z info by `mlz`	Sec. III B 1
Photo-z run 1	100	-	Photo-z info by `mizuki`	Sec. III B 2
Photo-z run 2	100	-	Photo-z info by `frankenz`	-
multiplicative-bias run 1	100	-	Change $\langle m_{\rm b} \rangle$ by $+0.01$	Sec. III B 3
multiplicative-bias run 2	100	-	Change $\langle m_{\rm b} \rangle$ by -0.01	-
Cosmology-varied run	50×100	100 different models (Fig 2)	Photo-z info by `mlz`	Sec. III B 4

TABLE I. Summary of our mock catalogs for Subaru Hyper-Suprime Cam Survey first-year data. For each of 100 cosmological models (parameter sets), we have 50 realizations of mock catalogs.

(ii): Populate source galaxies on the light-cone using original angular positions and redshifts of the observed galaxies.

(iii): Rotate the shape of each source galaxy at random to erase the real lensing signal.

(iv): Add the lensing shear on each source galaxy using the lensing simulations

Note that our method maintains the observed properties of the source galaxies on the sky. We increase the number of realizations of the mock catalogs by extracting multiple separate regions from a single full-sky simulation. Finally we obtain 2268 mock catalogs in total.

2. Photometric redshift uncertainties

In the fiducial mock catalogs, we utilize the photo-z information estimated by `mlz`. To examine possible systematic effects owing to photo-z uncertainties, we generate additional mock realizations adopting the two other redshift estimates by `mizuki` or `frankenz`. We produce 100 mock realizations of the HSC S16A catalogs for each model, and use them to evaluate the impact of photo-z uncertainty in our denoising process, as well as of cosmological parameter inference.

3. Image calibration uncertainties

We use a single value of multiplicative bias $\langle m_{\rm b} \rangle$ (defined in Eq. [8]) when generating our fiducial mock catalogs. Estimating $\langle m_{\rm b} \rangle$ is based on image simulations, and thus there remains a 1%-level uncertainty [62]. To account for possible systematic effects by the misestimation of the multiplicative bias, we make additional mock realizations by changing $\langle m_{\rm b} \rangle \rightarrow \langle m_{\rm b} \rangle + \Delta m_{\rm b}$ in the production process. We assume two values of $\Delta m_{\rm b} = \pm 0.01$. For each value of $\Delta m_{\rm b}$, we produce 100 mock realizations of the HSC S16A.

4. Varying cosmological models

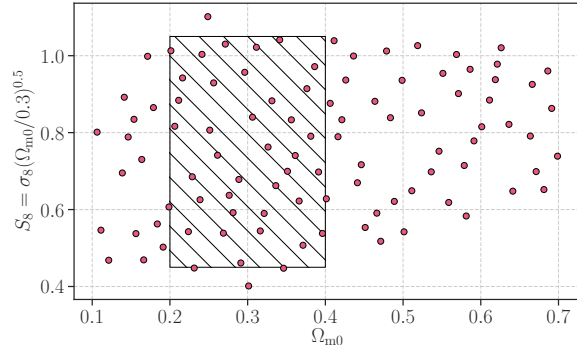To study the cosmological dependence on weak lensing maps, we also generate mock catalogs of the HSC S16A

FIG. 2. The 100 different cosmological models to study the cosmological dependence on weak lensing maps. At each point, we generate 50 mock realizations of the HSC S16A data. The hatched region represents the prior information adopted in our cosmological parameter inference (see, Sec. IV B).

data by varying cosmological models. We design the cosmological models for simulations so as to cover a much wider area in the two-parameter space $(\Omega_{\rm m0}, \sigma_8)$ than the constraints by the current galaxy imaging surveys [55, 63–65]. We choose a sample of cosmological models in the $\Omega_{\rm m0} - \sigma_8$ plane by using a public R package to generate the maximum-distance sliced Latin Hypercube Designs (LHDs) [66]. The code allows flexible experimental designs by subdividing the whole designs into different "slices". The distance metric is maximized both for the individual slices and the whole design. We first generate 240 designs in two slices (120 designs each) in a two dimensional rectangle specified by $0.1 \le \Omega_{\rm m0} \le 0.7$ and $0.4 \le \sigma_8(\Omega_{\rm m0}/0.3)^{0.6} \le 1.1$ using the codes. We then restrict the designs to those with $0.4 \le \sigma_8 \le 1.4$. This leaves 200 designs (100 per slice). We use the 100 designs only in the first slice for this study. Figure 2 shows the resultant 100 cosmological models adopted in our simulations. Note that we set $\Omega_\Lambda = 1 - \Omega_{\rm m0}$ assuming a spatially flat universe. For other parameters, we adopt $\Omega_{\rm b0}h^2 = 0.02225$, $h = 0.6727$ and $n_s = 0.9645$. These parameters are consistent with the results from Planck 2015 [2].
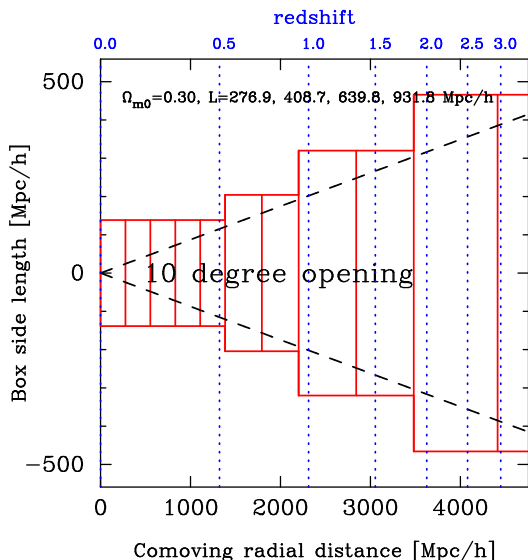
FIG. 3. The configuration of N-body boxes in our ray-tracing simulation for the cosmology with $\Omega_{\mathrm{m}0} = 0.3$.

For each cosmological model, we perform ray-tracing simulations under a flat-sky approximation. We adopt the multiple lens-plane algorithm [7, 8] to simulate the gravitational lensing effects on a light cone of angular size $10° \times 10°$. We place a set of N-body simulations with different volumes to cover a wide redshift range as well as have higher mass and spatial resolutions at lower redshifts (e.g. see, Ref. [9]). We consider four different box sizes on a side and each box size is varied as a function of the cosmological model. The box size L_{box} of the N-body simulations for our ray-tracing simulations is set by the following criteria:

$$L_{\mathrm{box},1} = \chi(z = 0.5) \times (\theta_{\mathrm{sim}} + \Delta\theta), \tag{13}$$

$$L_{\mathrm{box},2} = \chi(z = 0.8) \times (\theta_{\mathrm{sim}} + \Delta\theta), \tag{14}$$

$$L_{\mathrm{box},3} = \chi(z = 1.5) \times (\theta_{\mathrm{sim}} + \Delta\theta), \tag{15}$$

$$L_{\mathrm{box},4} = \chi(z = 3.0) \times (\theta_{\mathrm{sim}} + \Delta\theta), \tag{16}$$

where $L_{\mathrm{box},i}$ is the box size for the i-th smallest-volume simulation and $\theta_{\mathrm{sim}} = 10\,\mathrm{deg}$. We introduce the buffer in opening angle to compute L_{box} and set $\Delta\theta = 2\,\mathrm{deg}$. We then place the N-body simulations with the box size of $L_{\mathrm{box},1}, L_{\mathrm{box},2}, L_{\mathrm{box},3}$ and $L_{\mathrm{box},4}$ to cover the light cone in the redshift range of $0 < z < 0.5$, $0.5 < z < 0.8$, $0.8 < z < 1.5$, and $1.5 < z < 3.0$, respectively. Figure 3 shows an example of the configuration of N-body boxes in our ray-tracing simulation in the case of $\Omega_{\mathrm{m}0} = 0.3$.

For a given single N-body simulation volume, we produce two sets of the projected density fields with a projection depth of $L_{\mathrm{box}}/2$ on 9600^2 grids by using the triangular-shaped cloud assignment scheme [67]. By solving the discretized lens equation numerically, we obtain the lensing convergence κ and shear γ on 4096^2 grids with a grid size of 0.15 arcmin. A single realization of our ray-tracing data consists of 22 source planes in the range of $z \lesssim 3$. We perform 50 ray-tracing realizations of the underlying density field by randomly shifting the simulation boxes assuming periodic boundary conditions. We finally produce the mock catalog of the HSC S16A as described in Sec III B 1.

When running cosmological N-body simulations, we use the parallel Tree-Particle Mesh code GADGET2 [68]. We generate the initial conditions using a parallel code developed by Refs. [69, 70], which employ the second-order Lagrangian perturbation theory [71]. The number of N-body particles is set to 512^3. We set the initial redshift by $1 + z_{\mathrm{init}} = 36\,(512/L_{\mathrm{box}})$, where we compute the linear matter transfer function using CAMB [72]. Note that our choice of the initial redshift is motivated by the detailed study of Ref. [73].

IV. ANALYSIS

A. Denoising by deep-learning networks

1. Conditional generative adversarial networks

To perform mapping from a noisy lensing field κ_{obs} to a noiseless counterpart κ_{WL}, we use a model of conditional generative adversarial networks developed in Ref [36]. The networks have two main components, a generator and a discriminator. We train the networks so that the generator applies some transformation to the input noisy field κ_{obs} to output a *noise* field κ_{N}. The discriminator compares the input image to an unknown image (either a target image from the data set or an output image from the generator) and tries to judge if it is produced by the generator. To be specific, the input image for the discriminator is set to the noisy field κ_{obs}, while the target image is either of the noise counterpart of κ_{obs} or an output from the generator.

The structure of the generator and the discriminator in our networks is essentially the same as in Ref. [35], except for minor parameter tuning. The generator uses a U-Net structure [74] with an eight set of convolution and deconvolution layers. Each convolution layer consists of convolution with a kernel size of 5×5, the batch normalization, and the application of the activation function of leaky ReLU with a leak slope of 0.2. The deconvolution layer does the inverse operation of the convolution layer. The generator also has additional skip connections between mirror layers to propagate the small-scale information that would be lost as the size of the images decreases through the convolution process. The discriminator produces a single value from a given input image for the decision whether the input is real or a fake. The final output of the discriminator is made after the image reduction through 4 convolution layers and after averaging all the responses from the convolution layers. In the convolution layers in the discriminator, we remove the batch normalization to balance the losses of the generator and the discriminator in a stable way. The resulting

number of parameters in our networks is close to 400000.

2. Training the networks

The objective of our networks is to solve an optimization problem with a cost function expressed as a combination of loss functions as

$$\min_G \max_D \left\{ \mathcal{L}_{\mathrm{cGAN}}(G, D) + \lambda \mathcal{L}_{\mathrm{L1}} \right\}, \qquad (17)$$

where G indicates the generator and D is the discriminator. We here introduce two loss functions as

$$\mathcal{L}_{\mathrm{cGAN}}(G, D) = \mathbb{E}_{x,y} \log D(x, y)$$
$$+ \mathbb{E}_{x,z} \log \left\{ 1 - D(x, G(x, z)) \right\}, (18)$$
$$\mathcal{L}_{\mathrm{L1}}(G) = \mathbb{E}_{x,y,z} \sum_{\mathrm{map}} |y - G(x, z)|, \qquad (19)$$

where x is the input noisy field, y is the true noise field, and z is a random noise vector at the bottom layer of the generator. The function $D(X_1, X_2)$ returns the score in the range of zero to unity to evaluate if the noise counterpart of X_1 and a noise field X_2 is identical or not. In Eq. (19), the summation runs over all the pixels in a map but with the masked region excluded. In the training, we alternate between one gradient descent step on D, then one step on G. As suggested in Ref. [75], we train to maximize the term of $\log D(x, G(x, z))$. Also, we divide the objective by 2 while optimizing D, which slows down the learning rate of D relative to G.

When training the networks, we use the minibatch Stochastic Gradient Descent (SGD) method and apply the Adam solver [76], with learning rate 0.0002, momentum parameters $\beta_1 = 0.5$ and $\beta_2 = 0.9999$. We also set $\lambda = 100$ in Eq. (17). The latter controls the strength of the regularization given by the L1 norm. All the networks in this paper are trained with a batch size of 1. We initialize the model parameters in the networks from a Gaussian distribution with a mean 0 and a standard deviation of 0.02. We train our networks using the TensorFlow implementation[9] on a single NVIDIA Quadro P5000 GPU. While processing, we randomly select training and validation data from the input data sets. Each network is validated every time it learns 100 image pairs.

To prepare the training data set, we use 400 realizations of our mock HSC S16A catalogs (Sec. III B 1). Using the information of noiseless lensing maps κ and γ in our survey window, we generate 60000 noisy maps by injecting independent noise realizations at random. From the 60000 image pairs of the noisy field κ_{obs} and the underlying noise κ_{N}, we select 30000 image pairs by bootstrap sampling so that each bootstrap realization can

_____
[9] We use the modified version of `https://github.com/yenchenlin/pix2pix-tensorflow`

contain 200 realizations of noiseless lensing fields. In our previous study, we find that it is near-optimal to use 200 realizations of noiseless lensing fields and set the number of training sets to 30000 for our networks [35].

3. Production of the final denoised image

As reported in Ref. [35], a single set of our networks trained by 30000 image pairs has a large scatter in the image-to-image translation. To reduce this dependence on training data sets, we generate 100 bootstrap sampling of 30000 training data and obtain a total 100 networks for denoising. Namely, we obtain 100 candidates of the underlying noise field κ_{N} for a given noisy field κ_{obs}. To evaluate the best estimate of κ_{N}, we take the median over the 100 candidates on a pixel-by-pixel basis. Once the averaged estimate of κ_{N} is determined in this manner, we evaluate the underlying noiseless field κ_{WL} by subtracting the best noise model from the observed one κ_{obs}.

The denoising process by our networks is tested by 1000 noisy data from the fiducial mock catalogs. These test data are not used in the training process. We investigate the properties of the denoised fields in Appendix A. In short, our denoising method with the conditional GANs reproduces the one-point PDF for the noiseless field within a $0.5 - 1\sigma$ level. However, the denoised field is less accurate on a pixel-by-pixel basis with respect to the noiseless (true) counterpart. It is worth noting that the denoised field shows a tight correlation with the noiseless one at large angular scales. The cross-correlation coefficient in the two-point correlation function is close to unity for the separation angle of > 30 arcmins. It is remarkable that 44.3% of the positive peaks with their peak heights larger than 5σ on a denoised map are found to have single counterparts of massive-cluster-sized halos at $z < 1$. These results suggest that the denoised lensing map with our networks retains at least a part of the original cosmological information.

B. Likelihood analysis

To explore the possibility to infer the cosmological model with our GAN-based denoising, we perform a likelihood analysis for a summary statistic of a lensing map. To this end, we use the one-point PDF $\mathcal{P}$ of the lensing map in this paper. For a given lensing map, we measure the one-point PDF as a function of $(\kappa - \mu)/\sigma$ where μ and σ are the spatial average and root-mean-square. We perform a linear-spaced binning in the range of $-15 < (\kappa - \mu)/\sigma < 15$ with a step of 0.3. If the observed PDF follows the multivariate Gaussian distribution with covariance matrix $\boldsymbol{C}$, the log-likelihood func-

tion is expressed as

$$-2\log L(\boldsymbol{p}) = \sum_{i,j} C_{ij}^{-1} \left[\mathcal{P}_{\mathrm{obs}}(x_i) - \mathcal{P}_{\mathrm{model}}(x_i, \boldsymbol{p})\right]$$
$$\times \left[\mathcal{P}_{\mathrm{obs}}(x_j) - \mathcal{P}_{\mathrm{model}}(x_j, \boldsymbol{p})\right], \quad (20)$$

where x is the normalized lensing field with the zero mean and unit variance, $\mathcal{P}_{\mathrm{obs}}$ is the observed PDF, $\mathcal{P}_{\mathrm{model}}$ is the model template of the PDF as a function of x as well as the parameters of interest $\boldsymbol{p}$. We then minimize Eq. (20) by varying the parameters $\boldsymbol{p}$ and define the best-fit parameters at the minimum in $-2\log L$. In this paper, we consider the two parameters $(\Omega_{\mathrm{m0}}, S_8)$ to be varied, where $S_8 = \sigma_8 \sqrt{\Omega_{\mathrm{m0}}/0.3}$ is a primary parameter in the cosmological inference based on weak lensing analyses. We limit the range of the lensing PDF $\mathcal{P}$ to be larger than 0.03 to remove the high-σ information[10]. For a noisy PDF, the responsible range of pixel values is found to be $-2.1 < (\kappa - \mu)/\sigma < 2.4$, while the range is $-1.8 < (\kappa - \mu)/\sigma < 2.4$ for a denoised PDF.

Because of the limited constraining power due to the large sample variance, we resort to introducing prior information in the two-parameter space to obtain reasonable cosmological constraints. We set $0.2 \leq \Omega_{\mathrm{m0}} \leq 0.4$ and $0.45 \leq S_8 \leq 1.05$ for the prior. We adopt a uniform prior in this paper. The prior range in the parameter space is shown in the hatched region in Figure 2.

For the model template, we use the cosmology-dependent mock data sets as described in Sec. III B 4. For each of the 100 cosmological models, we have 50 realizations of the PDFs for the noisy lensing fields, and also the denoised fields by our GANs. We then use the averaged PDFs over 50 realizations for the model templates in Eq. (20). The covariance matrix is evaluated with 1000 realizations from the fiducial mock catalogs as in Sec. III B 1. Note that these 1000 realizations are not used in the training process for our GANs.

To estimate the confidence interval from the likelihood analysis, we need to compute the log-likelihood function at a designated parameter point (model), but our sampling in the parameter space is limited (see Figure 2). In this paper, we first calculate the log-likelihood values at the 100 cosmological parameter sets (cosmological models) as shown in Figure 2. We then evaluate $-2\log L$ at arbitrary points in the prior range by using a Gaussian process (GP) regression. In practice, we input 100 different sets of $-2\log L$ as a function of Ω_{m0} and S_8 and then calculate the log-likelihood for a given set of the cosmological parameters. We use a publicly available code GPy [82] and adopt a stationary Matern 3/2 kernel. To choose

the kernel function in the GP regression, we first make a training data of $-2\log L$ with 99 elements by reserving a single data of the log-likelihood from the parent 100 data points for validation. We then optimize the hyperparameters of the GP kernel with the 99 data points, do the GP interpolation to the reserved single data of $-2\log L$ (not used in the training). We repeat the test 100 times by changing the validation data point. For the validation, we use a chi-square metric as

$$\chi^2 = \sum_{i=1}^{100} \frac{[\log L_{\mathrm{test}}(\boldsymbol{p}_i) - \log L_{\mathrm{GP}}(\boldsymbol{p}_i)]^2}{\sigma_{\log L}^2(\boldsymbol{p}_i)}, \quad (21)$$

where $\log L_{\mathrm{test}}(\boldsymbol{p}_i)$ is the test log-likelihood at the i-th cosmological model, $\log L_{\mathrm{GP}}(\boldsymbol{p}_i)$ is the prediction by the GP, and $\sigma_{\log L}(\boldsymbol{p}_i)$ is defined by a 68% confidence level of the GP prediction. Among Matern 5/2, Matern 3/2, Gaussian, and Exponential kernels, the Matern 3/2 kernel shows the second smallest χ^2 with small bias. The exponential kernel provides the smallest χ^2 but this is due to the large prediction uncertainty $\sigma_{\log L}$ for different cosmological models.

After the validation, we construct the GP model of $-2\log L$ from the 100 data points and use it in the cosmological parameter inference below. The interpolation error by our GP regression is presented and discussed in Sec. V C 1 and we consider the error properly in the final constraints on the cosmological parameters.

V. RESULTS

A. Visual impression

We apply our GAN-based denoising to the real weak lensing map obtained from the HSC S16A data. Figure 4 shows a comparison of the denoised images between mock and real data set. In the figure, we normalize the lensing field so that it has zero mean and unit variance. The top left panel shows a noisy lensing field in a mock observation taken from 1000 realizations of the fiducial catalogs (Sec. III B 1). The top right panel represents the denoised weak lensing fields for the mock data. In the bottom, the left and right panels are similar to the top, but for the real HSC S16A data. On the denoised field, we mark the position of the matched galaxy cluster to the local maximum with its peak height greater than 5σ. For the mock data, we define the galaxy clusters by the dark matter halos with their masses greater than $10^{14} h^{-1} M_\odot$ and their redshifts $z < 1$ in N-body simulations. On the other hand, for the real data, we select the optically selected galaxy clusters [51] in the HSC S16A with their richness of > 15 and the X-ray selected clusters [52] in our survey window by their X-ray temperature being $> 2.14\,\mathrm{keV}$. Ref. [51] has shown that our selection of the optical richness and X-ray temperature roughly corresponds to the selection of the cluster mass by $> 10^{14} h^{-1} M_\odot$. According to the results in Appendix A 4, we expect $\sim 44.3\%$ of

---

[10] The high-σ information should be related to the abundance of massive galaxy clusters [77] and contain rich cosmological information [16, 78]. Nevertheless, such regions are likely affected by various systematic effects as the baryonic physics [79, 80] and the intrinsic alignments [81], not included in our analysis pipeline. To be conservative, we decide to remove this information in the likelihood analysis.
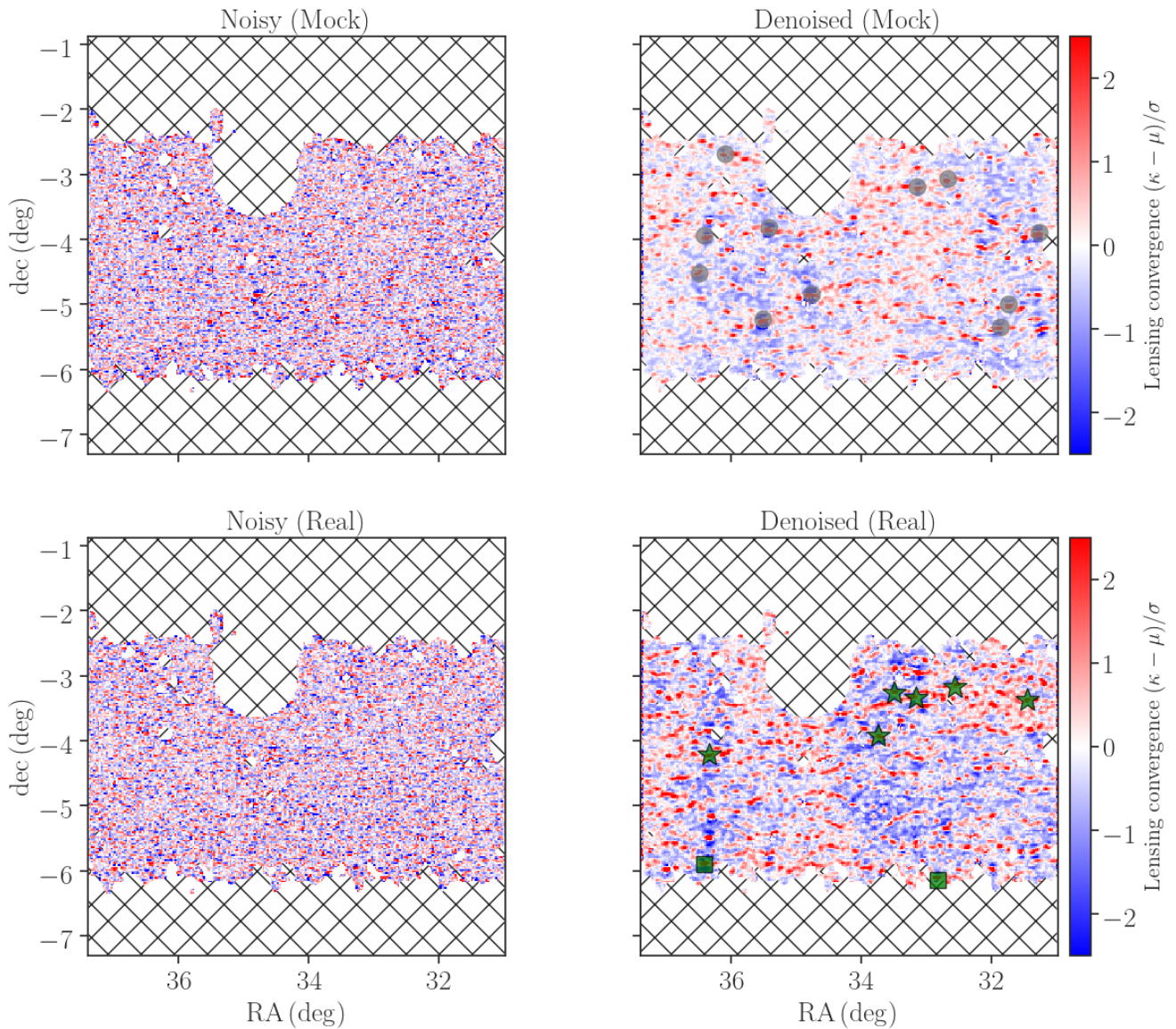
FIG. 4. Performance of the denoising of the *observed* weak lensing map in the Subaru HSC first-year data. In upper panels, the left and right panels show a noisy input map and the denoised counterpart for a mock observation among 1000 realizations, respectively. The bottom panels show the results similar to the upper ones, but for the real observational data. In the top right panel, the gray points show the matched dark matter halos to the local maxima on the denoised map. In the bottom left panel, the star and square symbols show the matched galaxy clusters selected in optical and X-ray bands, respectively. Note that the hatched region represents the masked area because of missing the data.

the peaks in a denoised field with their peak heights $> 5\sigma$ will find their counterparts of galaxy clusters. In our denoised map for the real HSC S16A, we find 30 peaks and 8 peaks have the counterparts. This matching rate is a bit smaller than the expectation from our experiments with 1000 mock observations, but it is still consistent within a 2σ level.

B. Cosmological dependence on lensing PDFs

Our primary interest is the cosmological information contents in the denoised maps. To study the cosmological dependence on the denoised lensing map, we use the mock catalogs as in Sec. III B 4. We have 50 mock re-
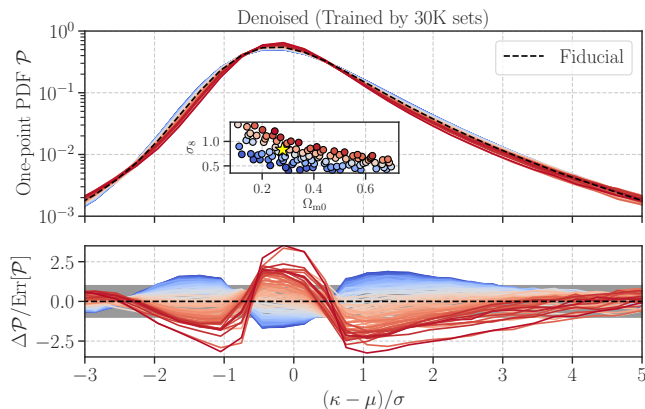
FIG. 5. The cosmological dependence of the one-point PDF of the denoised maps. The top panel shows the PDF as a function of the pixel value $(\kappa - \mu)/\sigma$, where μ and σ is the spatial average and root-mean-square for a lensing field κ, respectively. In the top, the inset figure represents the 100 cosmological models considered in the present study. The dependence of cosmological models is highlighted by the color difference. In the bottom, we show the difference of the PDF from our fiducial cosmological model (the WMAP9 cosmology [58]) normalized by the statistical uncertainty. For reference, the gray filled region in the bottom shows a $\pm 1\sigma$-level difference.

FIG. 7. The log-likelihood function by the lensing PDF with the marginalization over Ω_{m0}. The figure shows the difference of $-2 \log L$ from the minimum value for the mock input data. The black line shows the result by the noisy PDF, while the red line stands for that by the denoised PDF. The vertical dashed line shows the best-fitted value given by our likelihood analysis. Note that the true input S_8 is 0.801 for our mock data.

FIG. 6. Similar to Figure 5, but for the lensing PDF without our denoising process.

alizations of noisy lensing maps for each of 100 different cosmological models. For a given cosmology, we input a noisy map to our GANs, obtain the denoised map, and then compute the one-point PDF from the denoised map. We repeat this process for 50 realizations per each cosmological model and estimate the average PDF.

Figure 5 summarizes the cosmological dependence on

the denoised PDF for the HSC S16A. We find a clear dependence of the cosmological model and expect some of the models can be ruled out with the denoised PDF with a high confidence level (e.g. for a high S_8). The results in Figure 5 can be compared with the PDFs for noisy input maps. The cosmological dependence of the lensing PDFs without our denoising is shown in Figure 6, illustrating that the cosmological dependence is significantly weak compared to the statistical error of the PDF when one works on the original noisy lensing map. Note that the denoised PDFs are less sensitive to the cosmological parameters than the noiseless counterparts (see Appendix B).

To demonstrate the possibility of improving the cosmological constraints by the lensing PDF with our denoising method, we perform likelihood analysis as outlined in Sec. IV B. We assume that the true PDF in the observed region is identical to the average PDF for our fiducial cosmological model and adopt the cosmology-varied simulations as the model prediction. The values of $-2 \log L$ at the 100 cosmological parameter sets are then interpolated by the GP regression as described in Sec. IV B. This analysis allows us to set the expected confidence level of the cosmological constraint with the observed noisy PDF as well as the denoised counterpart for our observational setup. Figure 7 shows the log-likelihood marginalized over Ω_{m0} when the input PDF is set to the average PDF for our fiducial cosmology. We can set the expected 1σ constraint by the difference between the best-fitted parameter and its 1σ upper limit. We find the statisti-

cal uncertainty of S_8 to be 0.104 for the original noisy PDF, while it decreases to 0.051 for the denoised PDF. The improvement of the statistical error by our denoising method is a factor of $0.104/0.051 \simeq 2$. It is worth noting that our denoising plays a critical role in testing models with low S_8 with the lensing PDF. In models with lower-S_8, the overall density fluctuations and high peaks in the lensing mass map are less prominent. Since small fluctuations in a lensing map tend to be buried in the observational noises, then denoising is important to improve the sensitivity to a wider range of S_8.

C. Accounting for systematic uncertainties

1. Interpolation errors by Gaussian process

Our likelihood analysis relies on the GP regression based on 100 training data sets. The interpolation error by the GP can be estimated by the difference between the best-fit parameter inferred from the log-likelihood and its true counterpart when we input the lensing PDF with a known cosmological model. The vertical dashed lines in Figure 7 show the best-fit parameters when we input the average PDF for our fiducial cosmological model (with $S_8 = 0.801$). Note that our fiducial cosmology is not included in the training sets for the GP regression. We find the best-fit S_8 by our likelihood analysis to be 0.794 and 0.819 for the noisy and denoised PDFs, respectively. Hence, we evaluate the systematic uncertainty by the interpolation error by the GP to be 0.007 and 0.018 for the noisy and the denoised PDF.

It is worth noting that the above systematic uncertainties will be affected by the sample variance effect in the training data for the GP model. The data points of $-2 \log L$ in the construction of the GP will have the sample variance because those are evaluated by a finite number of mock realizations. In the fiducial analysis, we use 50 mock realizations to evaluate $-2 \log L$ for a given cosmological model. If we change the number of mocks to be 40, the interpolation errors are found to be 0.017 and 0.013 for the noisy and the denoised PDF, respectively. Hence, the sample variance in the training sets in our GP can be a source of the systematic uncertainties in our cosmological inference when we work on the noisy PDF. To be conservative, we set the interpolation error to be 0.017 for the noisy PDF.

2. Photometric redshifts

In our likelihood analysis, we assume the source redshift estimation by the specific method `mlz`, while other methods predict the different redshift distributions accordingly (see Figure 1). To assess the systematic uncertainty due to imperfect photo-z estimates, we perform a likelihood analysis when setting the input PDF to be the averaged PDF over 100 realizations of the mock catalogs
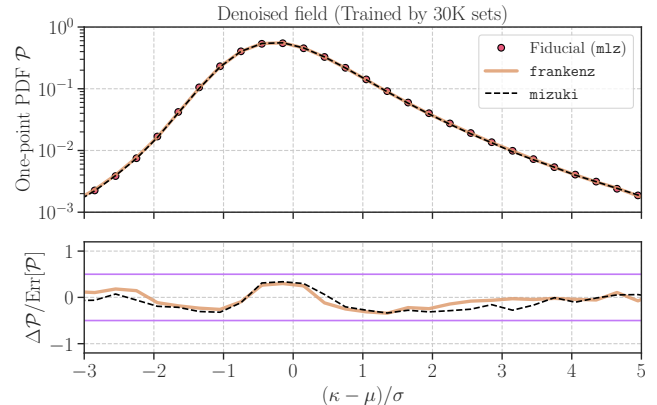

FIG. 8. The impact of photo-z uncertainties on the denoised lensing PDF. The upper panel shows the PDFs assuming the photo-z estimate by three different methods. In the bottom, we show the difference of the lensing PDF from our baseline case divided by the statistical uncertainty. For a reference, the magenta lines in the bottom represent $\pm 0.5\sigma$ levels.

with different photo-z information (Sec III B 2). Figure 8 highlights the impact of photo-z uncertainties on the denoised PDF. We find the photo-z estimate by different methods can induce the bias in the lensing PDFs with a $\sim 0.3 - 0.4\sigma$ level over a wide range of pixel values. These differences can induce the systematic uncertainty of the 1σ-level upper limit of S_8 by 0.02 and 0.036 for the noisy and the denoised PDFs, respectively.

3. Multiplicative biases

Besides, we assume the multiplicative bias defined by Eq. (8) is perfectly calibrated, but it can be misestimated with a level of 0.01. To test this systematic effect, we input the average PDF obtained from the mock catalogs as described in Sec III B 3 in our likelihood analysis. We find that a 1%-level error in the multiplicative bias can induce the shift of the 1σ upper limit of S_8 by 0.024 and 0.012 for the noisy and the denoised PDFs, respectively.

4. Baryonic physics and Intrinsic alignments

All the analyses in this paper assume the baryonic effects on the cosmic mass density can be negligible. Refs. [79, 80] examined the baryonic effects on the lensing PDF with hydrodynamical simulations and found the most prominent effect would appear in high-σ tails in the PDF. This is because the baryonic effects such as cooling, star formation, and feedback from active galactic nuclei commonly play a critical role in high-mass-density environments in the universe. In our likelihood analysis, we

remove such high-density regions by setting the range of pixel values to be $|\kappa - \mu|/\sigma \lesssim 2$.

Besides, we ignore possible correlations between the lensing shear and the shape noises. An example causing such correlations is the intrinsic alignment (IA) [81]. Although this IA effect can potentially cause the biased parameter estimation in future surveys [83], we expect it would be less important for our analysis because we do not employ clustering analyses of galaxy shapes. According to the observational facts, the IA effect is expected to be more prominent for redder galaxies (e.g. see Ref. [84]). Since redder galaxies preferentially reside in denser environments such as galaxy clusters, we would mitigate the impact of the IA effect on our analysis when removing the high-σ information from our analysis.

D. Cosmological constraint by the denoised map

We then perform the likelihood analysis with the observed PDFs in the real HSC S16A and provide a constraint on the cosmological parameter S_8. Figure 9 shows the log-likelihood functions marginalized over Ω_{m0} for both of the noisy and denoised PDFs. We find the best-fit parameter is found to be $S_8 = 0.492$ for the noisy PDF, while the denoised PDF prefers a larger value of $S_8 = 0.52$. In Appendix C, we infer the best model within our 100 different models (see Figure 2) by finding the smallest log-likelihood value. We confirm that the best model can fit the observed PDF for the noisy map as well as for the denoised counterpart. The goodness-of-fit is found to be 1.2-1.4. In Appendix D, we also study the probability distribution function of the best-fit S_8 with 1000 mock realizations assuming the true S_8 is set to 0.801. We find there are 21.8% and 5.7% chances to find the best-fit S_8 to be smaller than 0.5 for the noisy and denoised PDFs, respectively. This indicates that the small best-fit value of S_8 from our likelihood analysis may be explained by 2σ-level statistical fluctuations.

We provide the 2σ upper limits of S_8 from the observed lensing PDFs. The constraint for the noisy PDF is given by

$$
\begin{aligned}
S_8 < \ &0.777\,(\text{stat.}) + 0.017\,(\text{sys, GP}) \\
&+ 0.040\,(\text{sys, photo}{-}z) + 0.048\,(\text{sys}, m_b) \\
&= 0.882, \qquad\qquad\qquad\qquad\qquad\qquad (22)
\end{aligned}
$$

while this reduces to

$$
\begin{aligned}
S_8 < \ &0.663\,(\text{stat.}) + 0.018\,(\text{sys, GP}) \\
&+ 0.072\,(\text{sys, photo}{-}z) + 0.024\,(\text{sys}, m_b) \\
&= 0.777, \qquad\qquad\qquad\qquad\qquad\qquad (23)
\end{aligned}
$$

for the denoised PDF. Even including the systematic uncertainties, we can improve the upper limit of S_8 with a level of ~ 0.10 by our denoising without increasing the survey area. Our limits are consistent with the current constraints by other weak-lensing analyses [55, 63–65].
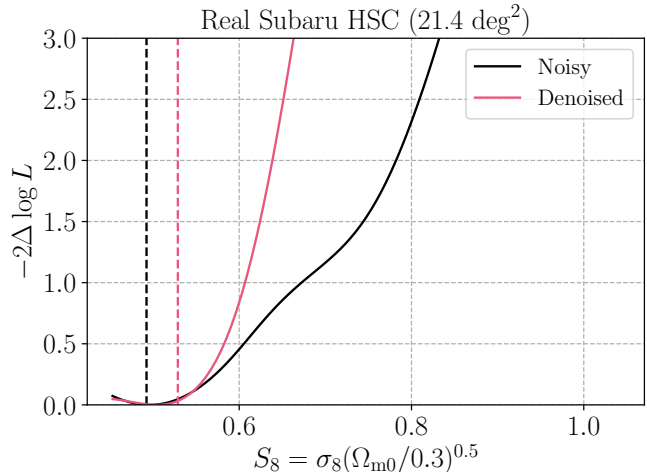

FIG. 9. Similar to Figure 7, but for the real data.

VI. CONCLUSION AND DISCUSSION

We have developed a novel method for noise reduction in cosmic mass density maps obtained from weak lensing surveys. We improve over our previous analysis [35] by incorporating more realistic properties in the training data sets and by performing an ensemble learning for conditional GANs. Our denoising method with GANs produces 100 estimates of the underlying noise fields for a given input (observation), allowing us to reduce generalization errors in the denoising process by taking the median value over 100 predictions by our networks.

We have applied our method to the real observational data by using a part of the Subaru Hyper-Suprime Cam (HSC) first-year shape catalog [49]. We confirm that the one-point PDF for the denoised map shows a stronger cosmological dependence than the original noisy counterpart. The likelihood analysis with mock observational data shows that the denoising can tighten the statistical error of a primary cosmological parameter $S_8 = \sigma_8\sqrt{\Omega_{m0}/0.3}$ by a factor of ~ 2 with the lensing PDF. The reduction of the sample covariance of the PDF measurements by ~ 4, or the increase in the survey area by the same factor, would be required to realize this improvement without denoising. We then performed the likelihood analyses with the noisy and denoised PDFs from the HSC data and provided the 2σ upper limits of $S_8 < 0.777\,(\text{stat}) + 0.105\,(\text{sys})$ for the noisy PDF and $S_8 < 0.663\,(\text{stat}) + 0.114\,(\text{sys})$ for the denoised PDF. We confirmed the small best-fit S_8 in our analysis can happen by a $\sim 5 - 20\%$ chance when the underlying true S_8 is 0.801 due to the limited sky coverage and the large sample covariance. Accounting for the systematic uncertainties and possible statistical fluctuations, we concluded that our limits are consistent with the current constraints by other lensing analyses as in Refs [55, 63–

For the first time, we show that generative adversarial networks (GANs) can be applied to precision cosmology. The framework developed in this paper can be easily generalized to other large-scale cosmological data sets. Since the observable information is limited by the cosmic variance, future cosmological analyses would need to extract some hidden information behind observational noises within a limited data size. Sophisticated modeling of cosmic large-scale structures with supercomputers can open a new window to produce mock observable "universes" as many as possible, leading to redesign cosmological analyses to defy conventional wisdom. Conditional GANs can be one of the most innovative approaches for next-generation cosmological analyses. To gain full benefits from machine learning methods in the future cosmological analyses, we have to solve some problems of large-scale computing for deep learning networks and fast and accurate modeling of the cosmic structures in multi-dimensional parameter space. Tight collaborations of astrophysics with machine learning are required to confront these challenges. Our results in this paper would be the first step toward the enhancement of the science returns by machine learning methods in future astronomical surveys.

ACKNOWLEDGMENTS

Appendix A: Properties of denoised maps

In this appendix, we summarize the statistical properties of weak lensing maps denoised by our conditional generative adversarial networks (GANs). Our training strategy for conditional GANs is provided in Sec. II B. Here, we show the validation results of the outputs from our networks by using 1000 test data sets. These test sets are based on the fiducial mock catalogs as in Sec. III B 1, while we do not use them in the training process. In the following, the lensing map is normalized so as to have zero mean and unit variance.

1. Visual comparisons

First of all, we make a visual comparison of weak lensing maps and highlight how our GAN-based denoising can work for noisy input images. Figure 10 compares the maps for a given realization in 1000 test data. In the figure, the left and right panels show an input noisy and the true noiseless counterpart, respectively. The medium represents the denoised map by our conditional GANs. In each panel, redder regions show a higher mass density, while bluer one is for a lower density. As can be seen in the figure, the image after denoising remains a similar pattern in the density contrast over a few degrees to the ground truth. Note that Figure 10 concentrates on the pixel values in the range of -2.5σ to $+2.5\sigma$, where the noises are usually dominated. Although it is not perfect, the denoised image contains some small-scale information (e.g., positive peaks) correlated with the noiseless counterparts.

2. One-point probability distributions

The one-point probability distribution function (PDF) is a simple summary statistic of the weak lensing map. Our previous study has shown that the denoised image follows a similar PDF to the noiseless true counterpart if the lensing field is properly normalized [35]. We update the previous analysis by including various observational effects such as a complex survey geometry, inhomogeneous galaxy distributions on a sky, wide redshift distributions of source galaxies, and variations of the weights in the analysis. Figure 11 shows the comparisons of PDFs averaged over 1000 realizations of lensing fields. The noiseless PDFs are significantly skewed compared to the observed noisy counterparts. Our denoising is efficient to reproduce a large skewness in the noiseless PDFs from the noisy input images. The typical bias in the reconstruction is found to range from a $0.5-1\sigma$ level over a wide range of pixel values.

3. Clustering amplitudes

To study the correlation between the denoised and the noiseless (true) fields, we perform a two-point correlation analysis. For a given set of two random fields on a sky, we define the correlation function as

$$\xi_{XY}(\theta) = \langle X(\boldsymbol{\phi})Y(\boldsymbol{\phi}+\boldsymbol{\theta})\rangle, \qquad (A1)$$

where X and Y are the two-dimensional random fields of interest. We evaluate the two-point correlation functions for the noiseless and denoised fields by using a public code `TreeCorr` [15]. We perform the linear-spaced binning in the angular separations from 3 to 60 arcmins. Figure 12 shows the averages of the two-point correlation functions over 1000 realizations.
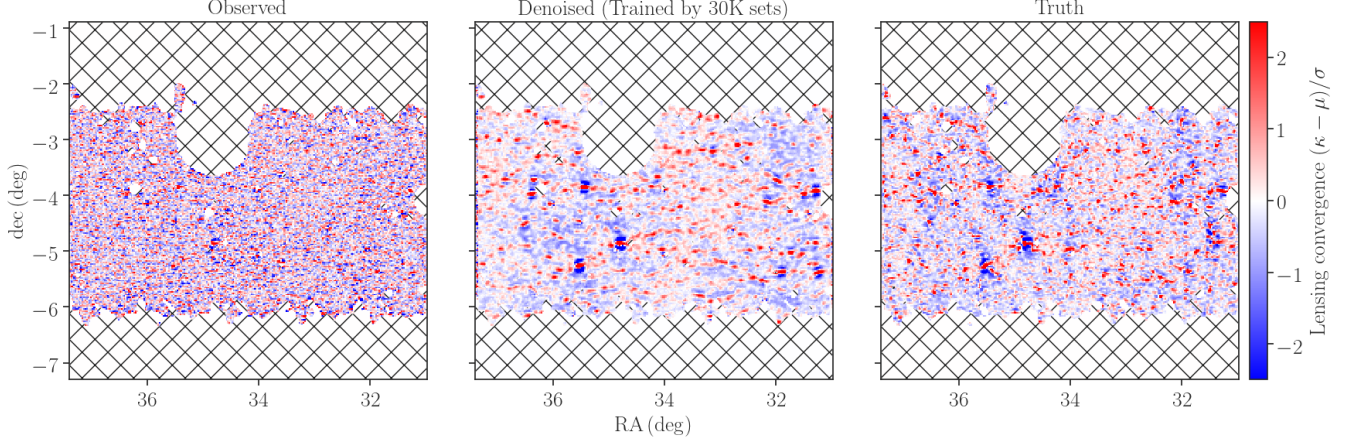
FIG. 10. An example of image-to-image translation by our networks. The left panel shows an input noisy lensing map, while the right stands for the true (noiseless) counterpart. The medium represents the reconstructed map by our conditional GANs. For the reconstructed map, we first obtain the underlying noise field from 100 bootstrap realizations of the generators in our GANs and then derive the convergence map by the residual between the input noisy map and the predicted noise. In this figure, the hatched region shows the masked area due to the presence of bright stars and inhomogeneous angular distributions of galaxies in our survey window. In the legend, μ and σ denote the spatial average and the root-mean-square of lensing fields, respectively.
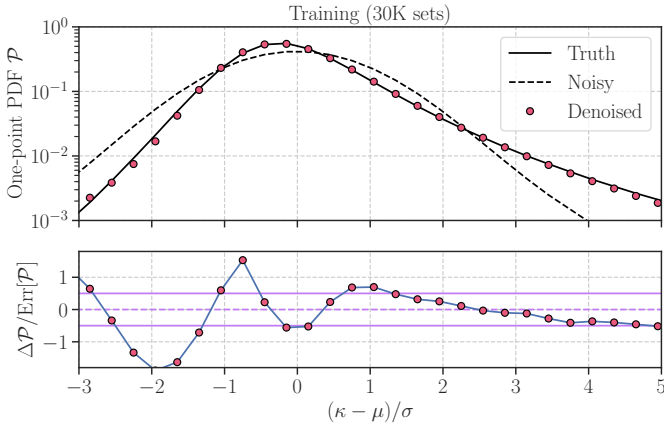
FIG. 11. The comparison of the lensing PDF among noisy, noiseless and denoised fields. The solid line in the top panel shows the averaged PDF over 1000 realizations of the noiseless lensing maps, while the dashed line is for the noisy (observed) one. The red points in the top panel show the averaged PDF for the denoised maps by our GANs. In the bottom, we show the difference between the noiseless and denoised PDFs in the unit of the sample variance of the noiseless PDFs. For a reference, we highlight $\pm 0.5\sigma$-level differences by the magenta lines at the bottom.
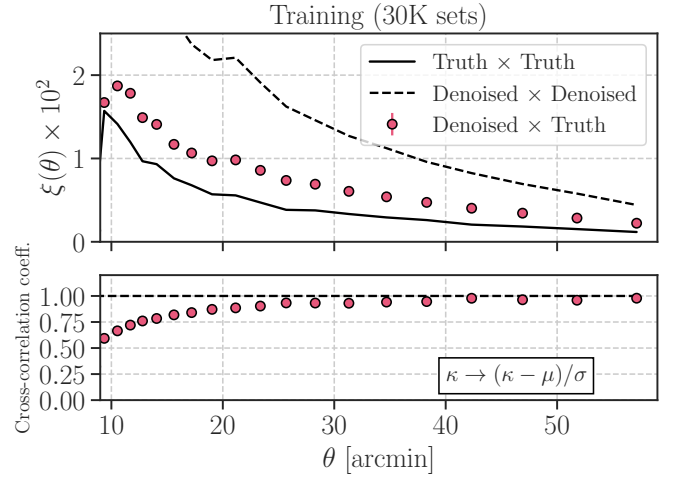
FIG. 12. The two-point correlation analysis of noiseless and denoised fields. In the top panel, the points show the cross correlation between the noiseless and denoised fields, while the solid and dashed lines are for the autocorrelation of the noiseless and denoised fields. The bottom panel represents the cross-correlation coefficient in the two-point clustering. At the angular scales larger than ~ 30 arcmins, the denoised fields are almost perfectly correlated with the noiseless counterparts.
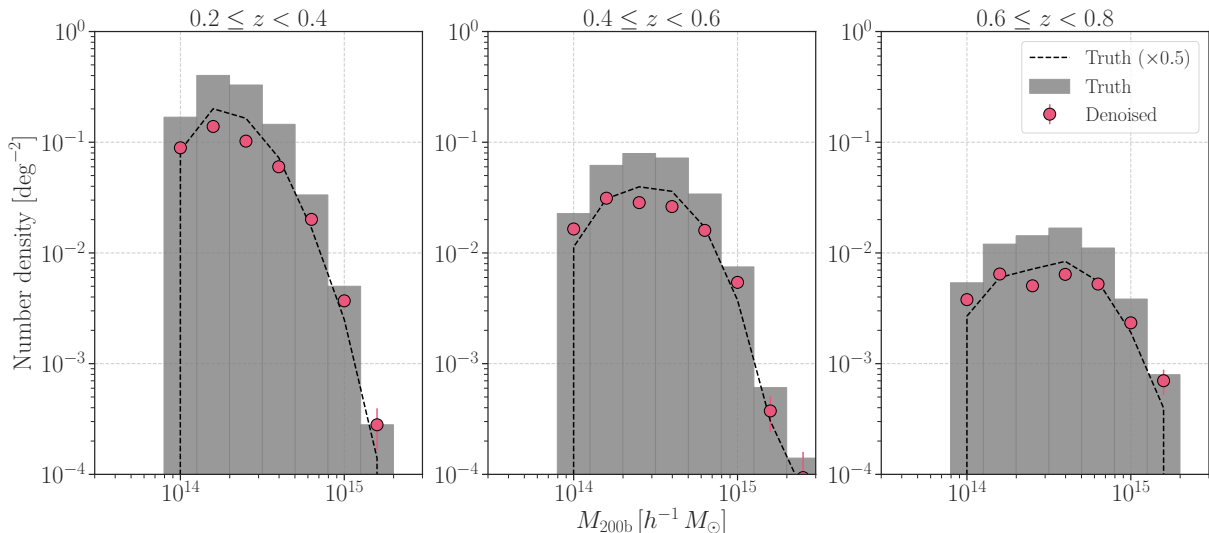
FIG. 13. The number density of the matched dark matter halos to the peaks on the lensing peaks. From the left to the right, we show the number density of the dark matter halos as a function of halo masses M_{200b} at three different redshift ranges, $0.2 \leq z < 0.4$, $0.4 \leq z < 0.6$, and $0.6 \leq z < 0.8$. The gray histogram shows the results for the true (noiseless) lensing fields, while the red points are for the denoised fields. The red points broadly follow the gray histogram except for the difference in amplitudes. For a reference, the dashed line in each panel represents the noiseless results with a multiplicative factor of 0.5.

We find that the cross-correlation function between the noiseless and denoised fields is offset from the true autocorrelation function, showing the denoised fields are biased for the ground truth. Apart from the bias, the cross-correlation coefficient (CCC) is a measure of the correlation degree in the two-point correlation analysis. In our case, the CCC is defined as

$$ r = \frac{\xi_{\text{true,denoised}}}{\sqrt{\xi_{\text{true}}\,\xi_{\text{denoised}}}}, \tag{A2} $$

where $\xi_{\text{true,denoised}}$ is the cross-correlation function between the noiseless and denoised fields, while ξ_{true} is the autocorrelation function of the noiseless field and so on. The bottom panel in Figure 12 shows that the CCC approaches unity as the angular separation increases. On the large scales ($\theta \gtrsim 30$ arcmins), the denoised fields are found to tightly correlate with the noiseless counterparts.

4. Peak-halo matching

To study small-scale information on a denoised lensing field, we examine the correspondence between dark matter halos and the local maxima in lensing maps. Since our mock HSC catalogs are originally based on cosmological N-body simulations, we can generate the light-cone halo catalogs with the same sky coverage as the mock HSC catalogs. The light-cone catalogs are produced from the inherent full-sky halo catalogs in Ref. [59]. The dark matter halos in the full-sky catalogs are identified by a phase-space temporal halo finder `Rockstar` [85]. In the

following analysis, we use the dark matter halos with their masses[11] greater than $10^{14}\,h^{-1}M_{\odot}$ at their redshift less than 1. This mass and redshift selection roughly corresponds to the real galaxy cluster catalog based on the photometric data in the HSC S16A [51].

On a given lensing map, we first identify local maximum with their peak height greater than 5σ. We then search for matched dark matter halos around the peak with a search radius of 5 arcmins [44]. When we find several halos in the search radius, we regard the matched halo as the closest halo from the position of the peak. Over 1000 realizations in our survey window, we find 41292 peaks and identify 77.9% of the peaks have the matched dark matter halos for the noiseless fields. After denoising, the number of peaks is found to be 15593 and the matching rate changes to 44.3%.

Furthermore, to validate the halo-peak matching, we study the number density of the matched dark matter halos as a function of halo masses and redshifts. Figure 13 shows the number density of the matched halos to the noiseless and the denoised peaks. As shown in the figure, the shape of the number density looks similar between two cases, indicating that the peak-halo matching for the denoised fields is not a coincidence.

———

[11] We define the halo mass as the spherical-overdensity mass with respect to 200 times mean overdensity.
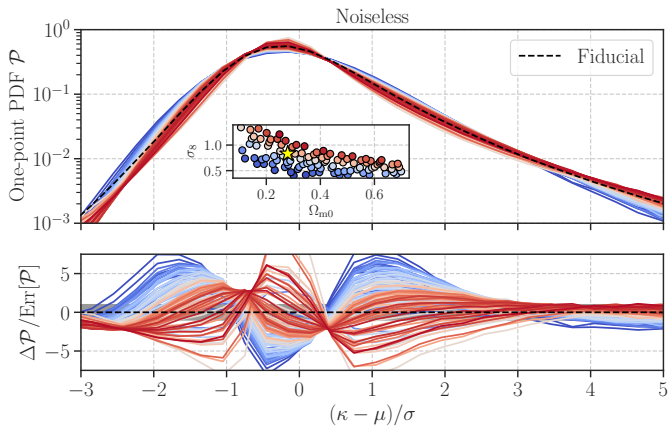
FIG. 14. Similar to Figure 5, but for the lensing PDF in the absence of shape noises.

Appendix B: Cosmological dependence of noiseless lensing PDFs

For a reference, we here show the cosmological dependence of the noiseless lensing PDFs. Figure 14 shows the dependence and illustrates that our denoising can not recover the expected cosmological dependence perfectly. Nevertheless, a trend of the cosmological dependence is similar between the noiseless and the denoised PDFs.

Appendix C: Comparison with observed lensing PDFs and simulated counterparts

Here we compare the observed lensing PDF with our model predictions based on numerical simulations. Since we have the model predictions of lensing PDFs for 100 different cosmological models as in Figure 2, we can infer the best cosmological model over the 100 models to find the smallest value of the log-likelihood (see Eq. [20]). Figure 15 and 16 summarize the comparison and show that the model based on our numerical simulations can give an acceptable fit to the observed lensing PDF for either of the noisy and denoised maps.

Appendix D: Probability distribution of the best-fitted cosmological parameter

In Sec. V D, the best-fit cosmological parameter $S_8 = \sigma_8 \sqrt{\Omega_{\rm rm0}/0.3}$ by our likelihood analysis is found to be ~ 0.5 and this is significantly small compared to the expectation from other cosmological measurements, e.g., cosmic microwave background analysis [2, 58]. To evaluate the statistical significance of our results, we study the probability distribution of the best-fit value of S_8 using
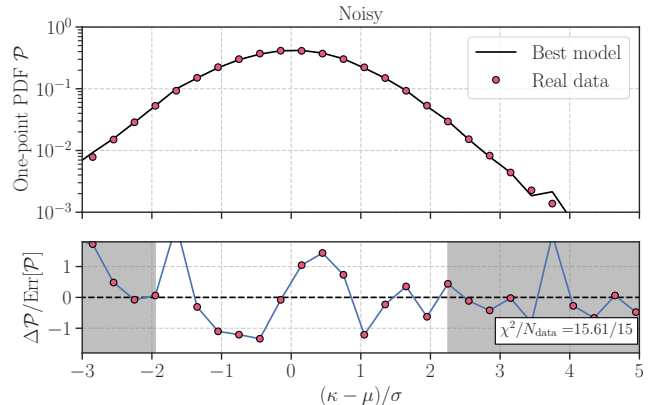

FIG. 15. The comparison of the observed lensing PDF with our simulation-based prediction. In the upper panel, the red points show the observed PDF without denoising and the black line represents the best model over 100 cosmological models in Figure 2. In the bottom, we show the residual between the observed PDF and the best model in the unit of the statistical error. The gray shaded region in the bottom infers the data points not used in the comparison. The smallest log-likelihood value $(-2 \log L)$ is found to be 15.61 at the cosmological model with $(\Omega_{\rm m0}, S_8) = (0.281, 0.591)$ for 15 data points.
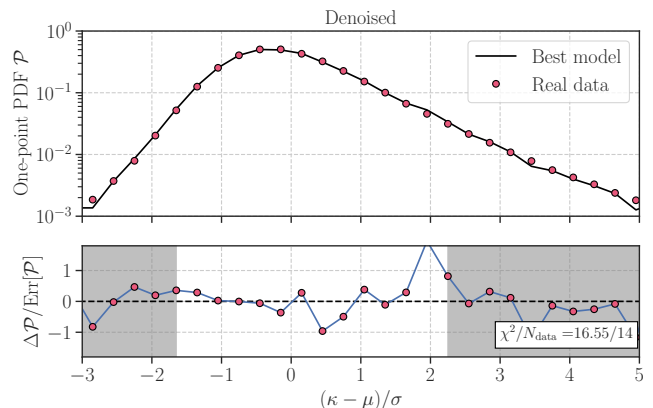

FIG. 16. Similar to Figure 15, but for the denoised PDF. The smallest log-likelihood value $(-2 \log L)$ is found to be 16.55 at the cosmological model with $(\Omega_{\rm m0}, S_8) = (0.269, 0.538)$ for 14 data points.

measurements of mock lensing PDFs. For each of 1000 mock catalogs in Sec. III B 1, we produce a noisy lensing map (Sec. II B), and then denoise the map with our conditional GANs (Sec. IV A 3). Then, we compute the lensing PDFs for both of the noisy and denoised maps. We perform likelihood analysis to find the best-fit param-

eter S_8 (Sec. IV B). We repeat this process 1000 times and estimate the probability distribution of S_8 with our observational setup.
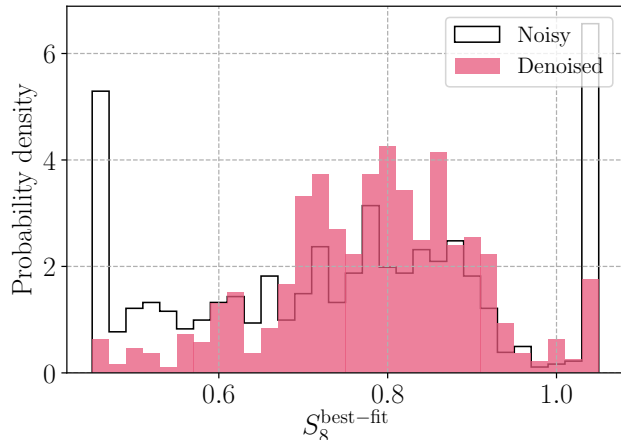

FIG. 17. The probability distribution function of the best-fit parameter $S_8 = \sigma_8\sqrt{\Omega_{m0}/0.3}$ assuming the true value is $S_8 = 0.801$. For this figure, we perform the likelihood analysis using the lensing PDFs taken from 1000 mock realizations as described in Sec. III B 1. The black open histogram shows the results for the noisy PDF, while the red filled one is for the denoised PDF. Note that we adopt the prior information of $0.45 \leq S_8 \leq 1.05$ in the likelihood analysis.

Figure 17 shows the probability distribution from our 1000 mock likelihood analyses. Note that the mock catalogs are based on the cosmological model with $S_8 = 0.801$. Because the noisy PDF has a weak sensitivity to S_8 compared to its statistical uncertainty, the best-fit parameter would follow a wide distribution around the input value. After denoising, we expect the width of the distribution can become narrower, but a long tail toward higher and lower S_8 exists. For instance, we find that 21.8% of the best-fit values in our mock analyses are smaller than 0.5 for the noisy PDFs, while this fraction reduces to 5.7% after we denoise the data. Hence, we conclude the small best-fit values of S_8 in our analysis with real HSC data can be caused by statistical fluctuations within a $\sim 2\sigma$ level, even if the true S_8 is 0.801.

[1] D. Clowe, A. Gonzalez, and M. Markevitch, Astrophys. J. **604**, 596 (2004), arXiv:astro-ph/0312273 [astro-ph].
[2] P. A. R. Ade et al. (Planck), Astron. Astrophys. **594**, A13 (2016), arXiv:1502.01589 [astro-ph.CO].
[3] D. Huterer and D. L. Shafer, Rept. Prog. Phys. **81**, 016901 (2018), arXiv:1709.01091 [astro-ph.CO].
[4] J. A. Tyson, R. A. Wenk, and F. Valdes, Astrophys. J. **349**, L1 (1990).
[5] N. Kaiser and G. Squires, Astrophys. J. **404**, 441 (1993).
[6] P. Schneider, Mon. Not. Roy. Astron. Soc. **283**, 837 (1996), arXiv:astro-ph/9601039 [astro-ph].
[7] B. Jain, U. Seljak, and S. D. M. White, Astrophys. J. **530**, 547 (2000), arXiv:astro-ph/9901191 [astro-ph].
[8] T. Hamana and Y. Mellier, Mon. Not. Roy. Astron. Soc. **327**, 169 (2001), arXiv:astro-ph/0101333 [astro-ph].
[9] M. Sato, T. Hamana, R. Takahashi, M. Takada, N. Yoshida, T. Matsubara, and N. Sugiyama, Astrophys. J. **701**, 945 (2009), arXiv:0906.2237 [astro-ph.CO].
[10] T. Matsubara and B. Jain, Astrophys. J. **552**, L89 (2001), arXiv:astro-ph/0009402 [astro-ph].
[11] J. Sato, M. Takada, Y. P. Jing, and T. Futamase, Astrophys. J. **551**, L5 (2001), arXiv:astro-ph/0104015 [astro-ph].
[12] M. Zaldarriaga and R. Scoccimarro, Astrophys. J. **584**, 559 (2003), arXiv:astro-ph/0208075 [astro-ph].
[13] M. Takada and B. Jain, Astrophys. J. **583**, L49 (2003), arXiv:astro-ph/0210261 [astro-ph].
[14] U.-L. Pen, T.-J. Zhang, L. van Waerbeke, Y. Mellier, P.-J. Zhang, and J. Dubinski, Astrophys. J. **592**, 664 (2003), arXiv:astro-ph/0302031 [astro-ph].
[15] M. Jarvis, G. Bernstein, and B. Jain, Mon. Not. Roy. Astron. Soc. **352**, 338 (2004), arXiv:astro-ph/0307393 [astro-ph].
[16] S. Wang, Z. Haiman, and M. May, Astrophys. J. **691**, 547 (2009), arXiv:0809.4052 [astro-ph].
[17] J. P. Dietrich and J. Hartlap, Mon. Not. Roy. Astron. Soc. **402**, 1049 (2010), arXiv:0906.3512.
[18] J. M. Kratochvil, Z. Haiman, and M. May, Phys. Rev. D **81**, 043519 (2010), arXiv:0907.0486 [astro-ph.CO].
[19] Z. Fan, H. Shan, and J. Liu, ApJ **719**, 1408 (2010), arXiv:1006.5121 [astro-ph.CO].
[20] M. Shirasaki and N. Yoshida, Astrophys.J. **786**, 43 (2014), arXiv:1312.5032 [astro-ph.CO].
[21] C.-A. Lin and M. Kilbinger, A&A **576**, A24 (2015), arXiv:1410.6955 [astro-ph.CO].
[22] A. Petri, J. Liu, Z. Haiman, M. May, L. Hui, and J. M. Kratochvil, Phys. Rev. D **91**, 103511 (2015), arXiv:1503.06214 [astro-ph.CO].
[23] W. R. Coulton, J. Liu, M. S. Madhavacheril, V. Böhm, and D. N. Spergel, arXiv e-prints , arXiv:1810.02374 (2018), arXiv:1810.02374 [astro-ph.CO].
[24] J. Schmelzle, A. Lucchi, T. Kacprzak, A. Amara, R. Sgier, A. Rfrgier, and T. Hofmann, (2017), arXiv:1707.05167 [astro-ph.CO].
[25] A. Gupta, J. M. Z. Matilla, D. Hsu, and Z. Haiman, Phys. Rev. **D97**, 103515 (2018), arXiv:1802.01212 [astro-ph.CO].
[26] D. Ribli, B. . Pataki, and I. Csabai, (2018), arXiv:1806.05995 [astro-ph.CO].
[27] W. Hu and M. J. White, Astrophys. J. **554**, 67 (2001),

arXiv:astro-ph/0010352 [astro-ph].

[28] P. Schneider, L. van Waerbeke, M. Kilbinger, and Y. Mellier, Astron. Astrophys. **396**, 1 (2002), arXiv:astro-ph/0206182 [astro-ph].

[29] M. Shirasaki, T. Nishimichi, B. Li, and Y. Higuchi, Mon. Not. Roy. Astron. Soc. **466**, 2402 (2017), arXiv:1610.03600 [astro-ph.CO].

[30] J. Liu and M. S. Madhavacheril, Phys. Rev. **D99**, 083508 (2019), arXiv:1809.10747 [astro-ph.CO].

[31] G. A. Marques, J. Liu, J. M. Z. Matilla, Z. Haiman, A. Bernui, and C. P. Novaes, JCAP **1906**, 019 (2019), arXiv:1812.08206 [astro-ph.CO].

[32] V. Vikram et al. (DES), Phys. Rev. **D92**, 022006 (2015), arXiv:1504.03002 [astro-ph.CO].

[33] C. Chang et al. (DES), Mon. Not. Roy. Astron. Soc. **475**, 3165 (2018), arXiv:1708.01535 [astro-ph.CO].

[34] A. Taruya, M. Takada, T. Hamana, I. Kayo, and T. Futamase, Astrophys. J. **571**, 638 (2002), arXiv:astro-ph/0202090 [astro-ph].

[35] M. Shirasaki, N. Yoshida, and S. Ikeda, Phys. Rev. **D100**, 043527 (2019), arXiv:1812.05781 [astro-ph.CO].

[36] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, ArXiv e-prints (2016), arXiv:1611.07004 [cs.CV].

[37] N. Jeffrey, F. Lanusse, O. Lahav, and J.-L. Starck, (2019), arXiv:1908.00543 [astro-ph.CO].

[38] H. Aihara et al. (HSC), PASJ **70**, S4 (2018), arXiv:1704.05858 [astro-ph.IM].

[39] M. Shirasaki, T. Hamana, M. Takada, R. Takahashi, and H. Miyatake, Mon. Not. Roy. Astron. Soc. **486**, 52 (2019), arXiv:1901.09488 [astro-ph.CO].

[40] M. Bartelmann and P. Schneider, Phys. Rep. **340**, 291 (2001), arXiv:astro-ph/9912508 [astro-ph].

[41] G. M. Bernstein and M. Jarvis, Astron. J. **123**, 583 (2002), arXiv:astro-ph/0107431 [astro-ph].

[42] C. Seitz and P. Schneider, Astron. Astrophys. **297**, 287 (1995), arXiv:astro-ph/9408050 [astro-ph].

[43] T. Hamana, M. Takada, and N. Yoshida, Mon. Not. Roy. Astron. Soc. **350**, 893 (2004), arXiv:astro-ph/0310607 [astro-ph].

[44] M. Shirasaki, T. Hamana, and N. Yoshida, Mon. Not. Roy. Astron. Soc. **453**, 3043 (2015), arXiv:1504.05672 [astro-ph.CO].

[45] S. Miyazaki et al., Astrophys. J. **807**, 22 (2015), arXiv:1504.06974 [astro-ph.CO].

[46] Y. Komiyama et al. (HSC), PASJ **70**, S2 (2018).

[47] H. Furusawa et al. (HSC), PASJ **70**, S3 (2018).

[48] S. Miyazaki et al. (HSC), PASJ **70**, S1 (2018).

[49] R. Mandelbaum et al., (2017), 10.1093/pasj/psx130, arXiv:1705.06745 [astro-ph.CO].

[50] J. Coupon, N. Czakon, J. Bosch, Y. Komiyama, E. Medezinski, S. Miyazaki, and M. Oguri, PASJ **70**, S7 (2018), arXiv:1705.00622 [astro-ph.IM].

[51] M. Oguri et al., Publ. Astron. Soc. Jap. **70**, S20 (2018), arXiv:1701.00818 [astro-ph.CO].

[52] C. Adami et al. (XXL Consortium), Astron. Astrophys. **620**, A5 (2018), arXiv:1810.03849 [astro-ph.CO].

[53] C. M. Hirata and U. Seljak, Mon. Not. Roy. Astron. Soc. **343**, 459 (2003), arXiv:astro-ph/0301054 [astro-ph].

[54] M. Tanaka, J. Coupon, B.-C. Hsieh, S. Mineo, A. J. Nishizawa, J. Speagle, H. Furusawa, S. Miyazaki, and H. Murayama, PASJ **70**, S9 (2018), arXiv:1704.05988.

[55] C. Hikage et al. (HSC), Publ. Astron. Soc. Jap. **71**, Publications of the Astronomical Society of Japan, Volume 71, Issue 2, April 2019,

[56] A. Brock, J. Donahue, and K. Simonyan, arXiv e-prints , arXiv:1809.11096 (2018), arXiv:1809.11096 [cs.LG].

[57] M. Shirasaki, N. Yoshida, and T. Hamana, ApJ **774**, 111 (2013), arXiv:1304.2164 [astro-ph.CO].

[58] G. Hinshaw et al. (WMAP), Astrophys. J. Suppl. **208**, 19 (2013), arXiv:1212.5226 [astro-ph.CO].

[59] R. Takahashi, T. Hamana, M. Shirasaki, T. Namikawa, T. Nishimichi, K. Osato, and K. Shiroyama, (2017), arXiv:1706.01472 [astro-ph.CO].

[60] M. R. Becker, Mon. Not. Roy. Astron. Soc. **435**, 115 (2013).

[61] M. Shirasaki, M. Takada, H. Miyatake, R. Takahashi, T. Hamana, T. Nishimichi, and R. Murata, Mon. Not. Roy. Astron. Soc. **470**, 3476 (2017), arXiv:1607.08679 [astro-ph.CO].

[62] R. Mandelbaum et al., Mon. Not. Roy. Astron. Soc. **481**, 3170 (2018), arXiv:1710.00885 [astro-ph.CO].

[63] H. Hildebrandt et al., Mon. Not. Roy. Astron. Soc. **465**, 1454 (2017), arXiv:1606.05338 [astro-ph.CO].

[64] M. A. Troxel et al. (DES), Phys. Rev. **D98**, 043528 (2018), arXiv:1708.01538 [astro-ph.CO].

[65] M. A. Troxel et al. (DES), Mon. Not. Roy. Astron. Soc. **479**, 4998 (2018), arXiv:1804.10663 [astro-ph.CO].

[66] S. Ba, R. W. Myers, and A. W. Brenneman, Technometrics **57**, 479 (2015).

[67] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Cambridge: University Press, —c1992, 2nd ed. (1992).

[68] V. Springel, Mon. Not. Roy. Astron. Soc. **364**, 1105 (2005), arXiv:astro-ph/0505010 [astro-ph].

[69] T. Nishimichi et al., Publ. Astron. Soc. Jap. **61**, 321 (2009), arXiv:0810.0813 [astro-ph].

[70] P. Valageas and T. Nishimichi, A&A **527**, A87 (2011), arXiv:1009.0597.

[71] M. Crocce, S. Pueblas, and R. Scoccimarro, Mon. Not. Roy. Astron. Soc. **373**, 369 (2006), arXiv:astro-ph/0606505 [astro-ph].

[72] A. Lewis, A. Challinor, and A. Lasenby, ApJ **538**, 473 (2000), arXiv:astro-ph/9911177 [astro-ph].

[73] T. Nishimichi et al., (2018), arXiv:1811.09504 [astro-ph.CO].

[74] O. Ronneberger, P. Fischer, and T. Brox, ArXiv e-prints (2015), arXiv:1505.04597 [cs.CV].

[75] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, ArXiv e-prints (2014), arXiv:1406.2661 [stat.ML].

[76] D. P. Kingma and J. Ba, ArXiv e-prints (2014), arXiv:1412.6980.

[77] G. Kruse and P. Schneider, Mon. Not. Roy. Astron. Soc. **318**, 321 (2000), arXiv:astro-ph/9904192 [astro-ph].

[78] S. Das and J. P. Ostriker, Astrophys. J. **645**, 1 (2006), arXiv:astro-ph/0512644 [astro-ph].

[79] K. Osato, M. Shirasaki, and N. Yoshida, Astrophys. J. **806**, 186 (2015), arXiv:1501.02055 [astro-ph.CO].

[80] T. Castro, M. Quartin, C. Giocoli, S. Borgani, and K. Dolag, Mon. Not. Roy. Astron. Soc. **478**, 1305 (2018), arXiv:1711.10017 [astro-ph.CO].

[81] M. A. Troxel and M. Ishak, Phys. Rept. **558**, 1 (2014), arXiv:1407.6990 [astro-ph.CO].

[82] GPy, "GPy: A gaussian process framework in python," http://github.com/SheffieldML/GPy (since 2012).

[83] E. Krause, T. Eifler, and J. Blazek, Mon. Not. Roy.

Astron. Soc. **456**, 207 (2016), arXiv:1506.08730 [astro-ph.CO].

[84] C. M. Hirata, R. Mandelbaum, M. Ishak, U. Seljak, R. Nichol, K. A. Pimbblet, N. P. Ross, and

D. Wake, Mon. Not. Roy. Astron. Soc. **381**, 1197 (2007), arXiv:astro-ph/0701671 [astro-ph].

[85] P. S. Behroozi, R. H. Wechsler, and H.-Y. Wu, ApJ **762**, 109 (2013), arXiv:1110.4372 [astro-ph.CO].