A Morphological Study of Galaxies Hosting Optical Variability-selected AGNs in the COSMOS Field

Yuxing Zhong1, Akio K. Inoue1,2, Satoshi Yamanaka2,3,4, and Toru Yamada5

1 Department of Pure and Applied Physics, Graduate School of Advanced Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku, Tokyo 169-8555, Japan
2 Waseda Research Institute of Science and Engineering, Faculty of Science and Engineering, Waseda University, 3-4-1, Okubo, Shinjuku, Tokyo 169-8555, Japan
3 General Education Department, National Institute of Technology, Toba College, 1-1, Ikegami-cho, Toba, Mie 517-8501, Japan
4 Research Center for Space and Cosmic Evolution, Ehime University, 2-5, Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan
5 ISAS/JAXA, 3-1-1, Yoshinodai, Chuo-ku, Sagamihara, Kanagawa, 252-5210, Japan

Received 2021 September 29; revised 2021 November 19; accepted 2021 November 29; published 2022 February 2

Abstract

Morphological studies are crucial to investigate the connections between active galactic nucleus (AGN) activities and the evolution of galaxies. Substantial studies have found that radiative-mode AGNs primarily reside in disk galaxies, questioning the merger-driven mechanism of AGN activities. In this study, through Sérsic profile fitting and nonparametric morphological parameter measurements, we investigated the morphology of host galaxies of 485 optical variability-selected low-luminosity AGNs at $z \leq 4.26$ in the COSMOS field. We analyzed high-resolution images of the Hubble Space Telescope to measure these morphological parameters. We only successfully measured the morphological parameters for 76 objects and most AGN hosts ($\sim 70\%$) were visually compact point-like sources. We examined the obtained morphological information as a function of redshift and compared them with literature data. We found that these AGN host galaxies showed no clear morphological preference. However, the merger rate increased with higher host star formation rate and AGN luminosity. Interestingly, we found ongoing star formation consistent with the typical star-forming populations in both elliptical and spiral galaxies, while these two types of galaxies were more symmetric than normal star-forming galaxies. These results suggest that optical variability-selected AGNs have higher probabilities to reside in elliptical galaxies than infrared-selected AGNs, whose host galaxies have a strong disk dominance, and support recent findings that the AGN feedback can enhance star-forming activities in host galaxies.

Unified Astronomy Thesaurus concepts: Active galaxies (17); AGN host galaxies (2017); Galaxy classification systems (582)

1. Introduction

Galaxy morphology is a direct means to reveal the interaction of galaxies with their environments and the impact of the internal perturbation. Based on galaxies’ visual appearance, morphological studies provide a unique method to investigate the evolution of galaxies. In 1926, Edwin Hubble proposed a galaxy morphology classification scheme, the so-called Hubble sequence, which divided galaxies into four classes: ellipticals, lenticulars, spirals, and irregulars (Hubble 1926). In addition, in 1959, de Vaucouleurs extended Hubble’s scheme, taking rings into consideration (de Vaucouleurs 1959). Recently, a well-known project, Galaxy Zoo, proposed more detailed morphological classifications, including disturbed features and cigar-shape classifications for edge-on galaxies (Lintott et al. 2008; Willett et al. 2013). Along with a purely visual investigation, parametric (Sérsic index $n$) and nonparametric methods ($G$-values, $M_{20}$, Concentration (C), Asymmetry (A), Smoothness ($S$), and Ellipticity) have been proposed to quantitatively describe the light distribution within a galaxy. Galaxies difficult to visually classify at high redshifts or due to edge-on structures can be more accurately distinguished through these parameters. For the Sérsic index, Ravindranath et al. (2004) found $n = 2$ efficient enough to separate early- and late-type galaxies, and Cassata et al. (2011) successfully applied this value to high-$z$ Hubble Space Telescope (HST) galaxies. Conselice (2003) presented the values of $G$, $A$, and $S$ for nearby galaxies, and $G$ decreased as the galaxies varied from ellipticals to disks and irregulars, whereas $A$ and $S$ increased.

Black holes (BHs) are ubiquitously found in the centers of galaxies. The co-evolution of a galaxy and its central BH is one of the most attractive issues in modern astronomy. Unfortunately, owing to the limitation of observational techniques, we can only investigate such connections via a proxy, which is known as the active galactic nucleus (AGN). In most cases, an AGN resides in the center of its host galaxy on a very small scale and is fueled by the gas inflow—an accretion onto the BH. Generally, AGNs that have dust-obscured structures, along with broad-/narrow-line regions (BLRs/NLRs), are rich in emission lines. These AGNs are called radiative-mode. They are more likely to be held by moderately massive ($M_\odot \sim 10^{10}$ to a few times $10^{11} M_\odot$) disk systems undergoing star-forming activities with star formation rates (SFRs) corresponding to those of the typical star-forming populations at epochs of up to $z \sim 2$ (Bournaud et al. 2012; Harrison et al. 2012). By contrast, jet-mode AGNs that are radiatively inefficient with advection-dominated accretion flows tend to reside in massive ellipticals, or spheroid systems (Heckman & Best 2014). Such co-evolution studies help us understand the growth mechanism of BHs, how AGNs influence the SFR, and the global structure of the host galaxy.
Prior to such detailed studies, AGNs must be selected from samples of astronomical objects such as normal galaxies. One primary approach to select AGNs is based on X-ray observations. Mainieri et al. (2011) selected 142 Type 2 QSOs in the COSMOS field via XMM-Newton observatory data with \(L_X[0.5–10\,\text{keV}] = 10^{44} - 10^{45}\,\text{erg} \, \text{s}^{-1} \) and \(\sim 0.8 < z < 2.0\). They found that the majority of their objects reside in early-type galaxies, and the minority belongs to prominent disk systems or mergers. Meanwhile, at \(z \sim 1\), about 62% \(\pm 7\%\) Type 2 QSO hosts are actively forming stars and this tendency becomes more apparent at higher redshifts, suggesting an evolutionary effect. The evolution of the specific SFR (sSFR) of these QSO hosts with redshift excellently agrees with that of normal star-forming galaxies (SFGs) at \(1 < z < 3\). In addition, using Herschel PACS observations, Santini et al. (2012) found that the hosts of their X-ray-selected radiative-mode AGNs in the GOODS-S and -N fields at \(<0.5 < z < 2.5\) exhibited enhanced SFRs and sSFRs in comparison with mass-matched inactive galaxies, indicating that SFGs are more likely to host AGNs.

Infrared (IR) observations are paramount for explorations of obscured AGNs and act as complements to X-ray data. This method can unveil AGNs that even the deepest X-ray observations cannot detect. Schawinski et al. (2012) studied the nature of quasar host galaxies based on mid-infrared (MIR, \(24\,\mu m\)) selected dust-obscured galaxies (DOGs) at \(z \sim 2\) using HST WFC3/IR imaging data. Contrary to X-ray-selected Type 2 AGN hosts analyzed by Mainieri et al. (2011), most of their DOGs are disk systems. Further, they argued that only a minority of these objects are mergers. This merger rate of AGN hosts is also supported by the study of IR-AGN hosts up to \(z \sim 2.5\) (Chang et al. 2017); they suggested that the merger rate of most luminous AGNs with \(\log(L_{IR}/L_{B}) \sim 12.5\) could be as large as 50%.

According to the sample selections, redshifts, and techniques used in different studies, the relations between the host galaxy’s morphology and AGN activities are still under debate. Pierce et al. (2007) selected 94 AGNs based on X-ray observations at \(0.2 < z < 1.2\) and found most of their host galaxies were classified as E/S0/Sa; nonparametric techniques revealed no evident morphological preference in the hosts of IR-AGNs. Gabor et al. (2009) performed a two-component decomposition for \(\sim 400\) X-ray-selected AGN hosts at \(0.3 < z < 1.0\) and found that the morphology of host galaxies spanned a wide range from bulge-dominated (1.5 < Sérsic index \(n < 10\)) to disk-dominated \((n \sim 1)\) systems and peaks between these two systems. Villforth et al. (2014) simulated AGN components and co-added them with stellar mass–matched control samples at \(0.2 < z < 0.8\), finding similar distributions of asymmetries, Sérsic indices, and ellipticities between AGN hosts and control sample galaxies. At \(1.25 < z < 2.67\), Schawinski et al. (2011) considered point-source components and suggested that over half of their 57 X-ray AGN hosts resided in disk-dominated systems.

Low-luminosity AGNs (LLAGNs) suffer contamination by light from their host galaxies, where usual color selection techniques do no work. Therefore, the flux variability, possibly arising from the instability of the accretion disk, in multi-epoch observations can be employed as a new selection technique. By studying the host galaxies of MIR variability-selected AGNs, Villforth et al. (2012) found common disturbed morphological features, and that it was secular processes–tidal processes and minor mergers, in particular–that triggered LLAGN activities rather than major mergers.

In this study, we explore the morphology of host galaxies whose central AGNs are selected based on their optical variabilities up to \(z \sim 3\). Such a method provides a more efficient selection of unobscured AGNs with low bolometric luminosities compared with that of X-ray or IR observations. Accompanied by this low obscuration is the high contamination of their light to the host galaxies. Further, high redshifts beget a cosmological surface brightness dimming of the galaxies, making it more difficult to visually determine the morphological classification. Considering these, we employ a 2D surface brightness fitting (Sérsic index), which includes a correction for central brightness excess, and nonparametric methods less affected by an AGN component but also corrected for it. To obtain detailed information about hosts’ SFRs, masses, and AGN fractions to investigate AGN-host connections, spectral energy distribution (SED) fitting is performed using IR to X-ray photometric data.

The rest of this article is organized as follows. In Section 2, we describe the samples as well as imaging and photometric data. Section 3 introduces both parametric and nonparametric methods, along with how we perform the SED fitting. The results of morphological measurements are presented in Section 4. Indirect results derived from different methods, as well as implications, are presented in Section 5. Finally, we summarize in Section 6. We use the following ΛCDM cosmological parameters to calculate the radius in the unit of kpc: \(H_0 = 70\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}, \Omega_M = 0.27\), and \(\Omega_\Lambda = 0.73\). All magnitudes in this article follow the AB system (Oke & Gunn 1983).

2. Sample and Data

2.1. Optical Variability-selected AGNs

The parent catalog of the AGNs examined in this study is based on optical variability-selection (Subaru Hyper-Suprime-Cam g, r, i, z; Furusawa et al. 2018; Kawanomoto et al. 2018; Komiyama et al. 2018; Miyazaki et al. 2018) by Kimura et al. (2020), consisting of 491 variability-selected AGNs in the Cosmic Evolution Survey (COSMOS, Scoville et al. 2007) field. The brightnesses of the AGNs in this catalog range from \(17.56\) to \(25.87\) in \(i\)-band magnitude, and the redshifts are spanned up to \(z = 4.26\).

The redshift distribution of these variability-selected AGNs is plotted in Figure 1. Of all of these objects, \(441\) (\(\sim 90\%\)) were detected in X-ray observations with Chandra, and \(337\) objects (\(\sim 69\%\)) had spectroscopic redshifts. The available spectroscopic redshifts are provided by \(z_{\text{spec}}\) in the Subaru Hyper-Suprime-Cam (HSC) PDR2 catalog, and are collected from zCOSMOS DR3 (Lilly et al. 2009), PRIMUS DR1 (Coil et al. 2011; Cool et al. 2013), VVDS (Le Fèvre et al. 2013), Sloan Digital Sky Survey (SDSS) DR12 (Alam et al. 2015), FMOS-COSMOS (Silverman et al. 2015), 3D-HST (Momcheva et al. 2016), and the DEIMOS 10 Spectroscopic Survey Catalog (DEIMOS catalog; Hasinger et al. 2018). If no spectroscopic redshift is available, \(z_{\text{best}}\) (best photometric redshift) in the Chandra catalog (Marchesi et al. 2016) is used for X-ray-detected objects, and \(z_{\text{PDF}}\) (photometric redshift measured using the galaxy templates) in the COSMOS2015 catalog (Laigle et al. 2016) is used for X-ray-undetected objects.

---

6 https://hsc-release.mtk.nao.ac.jp/doc/index.php/database-2/
To visually inspect the morphology of host galaxies of the optical variability-selected AGNs, we choose the I-band (F814W) imaging data obtained with the HST Advanced Camera for Surveys (ACS) from the COSMOS-HST Treasury project (Koekemoer et al. 2007; Massey et al. 2010). We searched around the coordinate of each object recorded in the Subaru HSC catalog offered by Kimura et al. (2020) within a radius of 1.5 in the COSMOS-HST catalog to obtain the HST cutout for each AGN host galaxy, and 485 of 491 objects were found in the COSMOS-HST Treasury project (Koekemoer et al. 2007; Massey et al. 2010). The cutout of each object had 202 pixels in height and width with a resolution per pixel of 0.03, corresponding to 6' × 6' in size.

Checking the imaging data, we found that many host galaxies were contaminated by their central AGNs, leading to bright point-spread function (PSF)–like components in their centers, or they were completely dominated by the PSF. Therefore, we considered the PSF component a necessary term to fit the 2D Sérsic brightness profile.

We used the Tiny Tim HST PSF Modeling Tool (Krist et al. 2011) to generate the PSF of the I-band (F814W) images. The final modeled PSF was the average of the stacking of 17 PSFs modeled with 17 types (O, B, A, F, G, K, and M, including intermediate types) of stars, which were provided by Tiny Tim. The PSFs modeled with some types of stars showed unexpected gaps or bumps in their radial light profiles. We performed stacking to avoid the possible influences of such gaps on successive measurements.

A real point source convolved with the real PSF should always be larger than the size of the PSF. To ensure the reliability of the modeled PSF, its FWHM was measured via the IRAF-psfmeasure command and compared with a few PSF-dominated objects (the entire object is visually in the shape of the PSF) that could be considered real PSFs. The modeled one had an FWHM of 2.853 pixels, whereas two objects at z > 4 both had an FWHM larger than 3 pixels, and an object at z = 3.599 had FWHM = 2.75 pixels. Although the FWHM of this object at z = 3.599 was smaller than that of the modeled PSF, we considered that such a 0.1 pixel difference was insignificant. Thus, we believe this modeling is sufficient for our purpose of involving it as a necessary term for the 2D Sérsic fitting. Figure 2 depicts our modeled PSF and the z = 3.599 object as an example of PSF-dominated host galaxies.

2.3. Photometric Data

We obtained multiwavelength photometry from the COSMOS2015 catalog (Laigle et al. 2016). The data include near-ultraviolet (NUV, 200–300 nm) and far-ultraviolet (FUV, 100–200 nm) observations from the Galaxy Evolution Explorer (GALEX) satellite (Zamojski et al. 2007; Capak et al. 2007b), UV observations (300–400 nm) from the Canada–France–Hawaii Telescope (u band, CFHT/MegaCam), and optical data from the COSMOS-20 survey taken with Subaru Suprime-Cam in five broad bands (B, V, i, r, and z), six intermediate bands (IB427, IB464, IB505, IB574, IB709, and IB827), and two narrow bands (NB711 and NB816; Taniguchi et al. 2007, 2015). In the near-infrared (NIR) range, Y-, J-, H-, and Ks-band data taken with WIRCam and Ultra-VISTA (McCracken et al. 2010, 2012) were used. In the MIR range, the SPLASH-COSMOS survey (Spitzer Large Area Survey with HSC; PI: P. Capak), S-COSMOS (Sanders et al. 2007), and the Spitzer Extended Mission Deep Survey (Ashby et al. 2013) provide data at [3.6], [4.5], [5.8], and [8.0] μm. The units of these photometric data were converted from Absolute Magnitude (AB mag) to flux densities (Fν) with the unit mJy (Fν = 1016.90 mJy/cm2 × 1000). The fractional errors in AB mag were used as the uncertainties of Fν.

Many objects were detected in X-ray observations, and we selected Chandra 0.5–10 keV band data originally from the “Chandra COSMOS Legacy” survey (Civano et al. 2016) for our X-ray-detected objects and converted their flux units of erg s−1 cm−2 to the units of flux density mJy (corresponding to the frequency of 2.297 × 1015 Hz). We assumed a 10% fractional difference of the observed flux as the uncertainty, which was typical for our objects based on the catalog by Civano et al. (2016).

3. Methods

3.1. Parametric Method

The basic parametric method includes measurement of the Sérsic index (n) and the effective radius (r_e, the radius that encloses 50% of the galaxy’s total flux). We performed 2D Sérsic + PSF profile fitting by employing Astropy (Astropy Collaboration et al. 2013, 2018) and its affiliated package statmorph (Rodriguez-Gomez et al. 2019). This package was also used for the nonparametric measurements in the next subsection.
To perform the fitting, the segmentation map, a 2D array with the same size as the cutout, that labels the pixel belonging to the source should be created in advance. We used photutils (Bradley et al. 2020) to create this image segmentation. For the first step, a smoothing box, which was defined as a 2D circular Gaussian kernel with an FWHM of 3 pixels, was used to smooth the cutout image and increase the object’s detectability. Next, we estimated the background noise of the cutout and detected the target object according to the threshold defined as $2\sigma$ per pixel above the smoothed background noise. The segmentation maps were generated based on this threshold and the FWHM size of the smoothing box. For several objects too faint or showing irregular structures, it might fail in basic morphological measurements or 2D Sérsic fitting. To correctly perform the fitting we adjusted the threshold (1.5–3$\sigma$ range) and the FWHM size of the smoothing box. We compared the measured values before and after adjusting the parameters for some successfully measured objects and found only insignificant changes for both parametric and nonparametric measurements.

As described in Section 2.2, an AGN and its host galaxy was a composite of two components, which meant we needed a decomposition to perform a more accurate fitting. Otherwise, the bright AGN in the center will enhance the central brightness, resulting in biased Sérsic indices and nonparametric parameters (see Section 3.3). Although statmorph could perform a de-convolution of the galaxy profile convolved with the PSF, it could not decompose the host galaxy and its central AGN. We describe how we correct for this systematic error in Section 3.3.

### 3.2. Nonparametric Method

The statmorph package also provides a set of nonparametric parameter measurements (Conselice 2003; Lotz et al. 2004). This method was used to measure the light distribution of a galaxy. Nonparametric measurements performed in this study include five parameters as described below.

The first is Gini ($G$) defined by the Lorentz curve of the galaxy’s light distribution. This parameter describes whether the light is evenly distributed among all pixels or concentrates on several pixels. The value of $G$ ranges from 0 to 1; 1 indicates that a single pixel encloses all brightness and 0 indicates all pixels have the same brightness. Notably, $G$ does not consider spatial positions.

The second parameter is $M_{20}$, the second-order moment of the galaxy’s brightest region enclosing 20% of the total flux of the galaxy (Lotz et al. 2004). It is a tracer of any bright nucleus, bar, spiral arm, and off-center star cluster within a galaxy. The value of $M_{20}$ is always negative; the lower it is, the higher the concentration anywhere within the galaxy.

The third parameter is Concentration ($C$), simply defined as follows (Conselice 2003):

$$C = 5 \times \log \left( \frac{r_{80}}{r_{20}} \right).$$

where $r_{80}$ and $r_{20}$ are radii enclosing 80% and 20% of the galaxy’s light, respectively. Concentration is directly related to the galaxy’s central brightness; a larger value indicates a brighter central region and a larger Sérsic index.

The fourth parameter is Asymmetry ($A$), which is an indicator of whether there exist asymmetric components within the galaxy. Spiral galaxies usually have large Asymmetry due to their spiral arms. In addition, inhomogeneous star-forming activities will also result in asymmetric structures. Although the value of $A$ is usually positive, because such a measurement depends on background noise, negative values may appear in low signal-to-noise ratio (SNR) cases.

The last parameter is Smoothness ($S$, or Clumpiness), which describes whether there is any clumpy structure within the galaxy. Elliptical galaxies are typical smooth systems holding small values of Smoothness, and SFGs tend to have more clumpy regions.

To calculate these five parameters, statmorph would create a Gini segmentation based on the Petrosian radius ($r_{\text{petro}}$, Petrosian 1976), which is given by

$$\eta(R) = \frac{I(R)}{\langle I(<R) \rangle},$$

where $I(R)$ is the surface brightness at a radius $R$, and $\langle I(<R) \rangle$ is the mean surface brightness within the radius $R$. We choose $\eta(R = r_{\text{petro}}) = 0.2$ (Lotz et al. 2004) to perform calculations. In fact, we tested measurements with different values for $\eta$ and found that the Gini_segmentation would ignore substructures, such as spiral arms, as $\eta$ increased and the measured values only varied slightly. Figure 3 gives an example output plot of the measurement, including Sérsic fitting and basic parameter measurements.

After checking the results, we found that over 70% of our objects had $S = 0$. We investigated the outputs and found all of these objects were PSF-like, indicating a significant optically unobscured AGN fraction. For all of these objects, there was no reliable Gini segmentation as well because the $r_{\text{petro}}$ always merely enclosed the bright core (at almost the FWHM size) of the PSF-like object, making it impossible to measure the nonparametric morphological parameters of the host galaxies. Then, we further checked the Gini_segmentation of the remaining objects so that the nonparametric parameters can be measured on the basis of a reliable enclosure of the galaxy’s light. As a result, we found many cases where the objects had no reliable Gini_segmentation with nonparametric parameters, concentrating around the combination of $G = 0.5 \pm 0.02$, $M_{20} \sim -1.74$, and $C = 3.0 \pm 0.2$. In $G-M_{20}$ diagnostics, these objects were all classified as Sb/Sc/Irr, leading to an overestimation of the fraction of disk systems (see Section 5.1). Real objects with this combination of $G$, $M_{20}$, and $C$ were also excluded.

### 3.3. Correction for Effects of the Central PSF Components

For the remaining objects, there were central bright PSF-like components in many cases. To correct for any systematic effect due to the PSF-like components, in particular, to investigate how the Sérsic index was affected, we modeled AGN host galaxies. The modeling includes several preset parameters: Sérsic index ($n$), effective radius ($r_e$), light intensity $I_e$ at the effective radius, ellipticity, and ratio of the total flux of the host galaxy and the AGN $(f_{\text{host}}/f_{\text{AGN}})$.

As the first step, we modeled 300 galaxies whose intrinsic Sérsic index ranged from 0.1 to 5.0 (50 samples) and ellipticities ranged from 0 to 0.5 (six groups), both with a 0.1 increase between two consecutive setups. Then, we considered three different effective radii ($r_e = 5, 10, 15$ pixels) and obtained $300 \times 3 = 900$ galaxies. The intensity at the effective
The detection thresholds and sizes of smoothing boxes were adjusted to ensure an enclosure of the entire galaxy including substructures as much as possible. The nonparametric parameters defined by the Petrosian radius were assumed to be \( I_e = 10 \) electrons s\(^{-1}\) pix\(^{-1}\), which was arbitrarily chosen, but scaled by an SNR, as described below. The unit is the same as for the actual HST cutout images. To model an AGN component, the brightest pixel of the modeled galaxy made in the first step was co-added with the value of the total flux multiplied by a factor. This factor was given by the following combinations of \( f_{\text{host}}/f_{\text{AGN}} = 1:0.01, 1:0.1, 1:0.2, ..., 1:0.9, 1:1.0, 1:1.5, 1:2, 1:2.5, 1:3, 1:4, \) and 1:5. Hence, we had 900 galaxies harboring AGNs modeled in total. These model galaxies, including the AGN component, were convolved with the F814W PSF. Then, we added noise in each pixel of the images following an error function whose standard deviation \( \sigma = 0.4 \) electrons s\(^{-1}\) pix\(^{-1}\), corresponding to an S/N = 25 at the effective radius. The SNR at the effective radius was chosen as a typical value observed in our actual AGN host galaxies.\(^7\)

By comparing the assumed values of Sérsic index and \( r_e \) with the values measured by statmorph, we found many cases where the differences were unacceptably large (measured \( n > 10 \)) and the host galaxy was heavily contaminated by the AGN. To make these models good matches to real objects, we adopted the following criteria to exclude invalid results:

1. Either \( \text{flag\_sersic} == 0 \) or \( \text{flag} == 0 \) (any error flag returned by statmorph)
2. Either \( r_e \) or \( \text{sersic\_rhalf} \)\(^8\) is smaller than that of the F814W PSF (2.19 and 1.52 pixels, respectively).
3. The size of the Gini\_segmentation is smaller than the effective radius as well as 50% of the size of the original segmentation.
4. The measured Smoothness is equal to 0.
5. Cases without reliable a Gini\_segmentation, for which \( G = 0.5 \pm 0.02 \), \( M_{50} \sim -1.74 \), and \( C = 3.0 \pm 0.2 \).
6. Objects with a failure in the calculation of the total flux, because of which an AGN component cannot be modeled.

By adopting these criteria, 7697 modeled AGN hosts were excluded and there were 7603 successfully modeled AGN host galaxies. Except for the last criterion, the others were applied to real objects to exclude unsuccessful measurements, ensuring a fair comparison between real and simulated AGN+host galaxies. As mentioned in Section 3.2, over 70% objects had \( S = 0 \), and so the fourth criterion produced the greatest rejection of valid objects. These host galaxies suffered from both cosmological surface brightness dimming and heavy contamination from their central AGNs, resulting in a complete PSF-like morphology, of which object ID002 in the right panel

---

\(^7\) In this procedure, the total flux of model galaxies without the AGN component, \( f_{\text{host}} \), was calculated by the method as follows. We convolved the model galaxies with the F814W PSF and co-added with the pixel noise following an error function with \( \sigma = 0.1 \) electrons s\(^{-1}\) pix\(^{-1}\). These high S/N = 100 model galaxies were used only to make a reliable segmentation map, wherein we calculated the total flux, \( f_{\text{host}} \), as the sum of the pixel counts.

\(^8\) Radius \( r_e \) indicates the effective radius of the object in the original image, and \( \text{sersic\_rhalf} \) indicates the effective radius of the model fitted by 2D Sérsic fitting.
of Figure 2 is representative. The correction method for the Sérsic index is described in the Appendix in detail. In addition, if the fraction of \( r_e \) to \( r_{\text{sersic rhalf}} \) of the real object was out of the range of the modeled hosts within the corresponding effective radius, there was no available correction for the object, and it was also excluded (see the Appendix).

Corrections for \( G \), \( M_{20} \), and \( C \) were also made through simulations. Unlike the Sérsic index that depends on a certain mathematical form, nonparametric parameters are much less affected by an AGN component only if Gini segmentation has a reliable enclosure of the host galaxy’s light and the fractional differences of values \( (\text{value}_{\text{galaxy+AGN}} - \text{value}_{\text{galaxy}}) / \text{value}_{\text{galaxy+AGN}} \) for a galaxy with and without an AGN can be applied for the correction.

The basic parameters of the simulation are the same as those for the correction for the Sérsic index other than the amplitude of the point source (or the total flux of the AGN) and the definition of \( f_{\text{host}}: f_{\text{AGN}} \). Apart from the simulation parameters, the nonparametric parameters were measured for the host galaxies without AGN components and then measured after the two components were co-added so that the fractional differences could be calculated. In this simulation, the total flux is set to 10 and \( f_{\text{host}} \) and \( f_{\text{AGN}} \) were the pixel counts of the central pixel (which is the brightest one) of the galaxy and the AGN, respectively. Before co-adding the two components, both of them were convolved with the PSF and then normalized so that the brightest pixels had pixel counts equal to 1; thus, \( f_{\text{host}}: f_{\text{AGN}} \) could be well controlled. There were four groups of the flux ratio between the host and AGN: \( f_{\text{host}}: f_{\text{AGN}} = 3:1, 1:1, 1:4, \) and \( 1:6 \). For \( f_{\text{host}}: f_{\text{AGN}} = 3:1 \), there was no visible PSF, and the other three had visible PSF components. Thus, the modeled host galaxies were further divided into two groups: with or without visible AGN components. Modeled host galaxies without visible AGN components corresponded to real objects whose central AGNs could not be visually confirmed or those with \( r_e > 16 \) pixels, whereas modeled hosts with visible AGN components were matched to real objects that had apparent AGN components. The fractional differences are listed in Table 1. To correct for the measurement biases, the following equation was adopted:

\[
N_{\text{P corrected}} = N_{\text{P measured}} \times (1 - f_{\text{NP}}),
\]

where \( N_P \) represents the nonparametric parameter, and \( f_{\text{NP}} \) is the fractional difference of the corresponding parameter.

Through modeling, it was found that galaxies with a small \( r_e \) and a large Sérsic index (\( n \geq 3 \)) were convolved to be the PSF shape even if they did not have AGN components. By applying the criteria, along with those failing in the measurement, 349 objects were excluded, including many spheroid systems (\( n > 2 \)), before they could be further discussed. This might result in a potential biased dominant morphology of the whole sample with 485 AGN host galaxies.

### 3.4. SED Fitting

Our samples were recorded in the Laigle et al. (2016) catalog, which provides MASS_BEST (the stellar mass of the entire galaxy) and SFR_BEST computed by the SED fitting. However, the SED fitting performed in this catalog did not include IR data from the Spitzer MIR channels, leading to large uncertainties of the SFR. Further, they did not take AGN contributions into consideration. Thus, to obtain results not only from the host galaxies but also from the AGNs such as the AGN luminosity and to update the catalog, we decided to perform our SED fitting. We chose X-CIGALE (Yang et al. 2020), the X-ray version of CIGALE (Burgarella et al. 2005; Noll et al. 2009; Boquien et al. 2019) to perform SED fitting of the photometric data retrieved from the COSMOS2015 catalog (see Section 2.3).

As the first step, a high-dimensional parameter grid of SED models consisting of all components that contribute to the emission was populated by X-CIGALE. Then, the goodness of the fit for each model was computed, and the best-fit SED model for each sample galaxy was identified via the reduced-\( \chi^2 \) statistic (Ramos Padilla et al. 2020).

The components used in the first step were required to be specified; we list these parameters and values in Table 2 to define the X-CIGALE grid of AGN hosts SEDs. For any other components not mentioned in the Table, default settings provided by Yang et al. (2020) for AKARI and COSMOS AGN SEDs were used.

The star formation history (SFH) of the host was treated with a delayed SFH model (Ciesla et al. 2015): \( SFR \propto \tau_{\text{main}} \exp(-t/\tau_{\text{main}}) \). Since the onset of the star-forming activity, the increase in the SFR is almost linear rather than a sudden burst; after a peak, the SFR decreases exponentially. There are two parameters controlling this model: the age from the onset of star formation \( (t_0) \), and the e-folding time \( (\tau_{\text{main}}) \) that also determines the timing of the SFR peak. Through adjusting these two parameters, ongoing or recent starburst events as well as a quenched phase could be simulated.

To model the emission from the stars, we adopted the Bruzual & Charlot (2003) stellar population synthesis library. The metallicity set by default was 0.02. However, considering that we were handling AGN samples, we increased the value to 0.05 because AGNs would enrich their environments, resulting in a higher chemical abundance (Zubovas et al. 2013; Taylor & Kobayashi 2015).

Owing to the existence of the obscuring structure, emissions from the AGN will be re-emitted at longer wavelengths. To simulate this reddening effect, we adopted SKIRTOR, a clumpy two-phase torus model derived from a modern radiative-transfer method (Stalevski et al. 2012, 2016). This model depends on several parameters such as the average edge-on optical depth at 9.7 \( \mu m \), the ratio of outer to inner radii, and the inclination. Because Kimura et al. (2020) suggested there

---

**Table 1**

| Criterion          | \( r_e \) |
|--------------------|----------|
|                    | 5        | 10       | 15       |
| Gini: \( G \)      |          |          |          |
| PSF visible        | 0.059    | 0.087    | 0.067    |
| PSF not visible    | 0.016    | 0.013    | 0.008    |
| \( M_{20} \)       |          |          |          |
| PSF visible        | 0.078    | 0.111    | 0.124    |
| PSF not visible    | 0.026    | 0.025    | 0.020    |
| Concentration: \( C \) |          |          |          |
| PSF visible        | 0.115    | 0.183    | 0.124    |
| PSF not visible    | 0.040    | 0.024    | 0.019    |

**Note.**

\(^{a}\) The effective radii are in units of pixels with a resolution of 0\(^{\prime}\)03 per pixel.
First, we describe our studies on AGN hosts and normal galaxies in the literature. morphological measurements along with those of previous 0.05 Metallicity R 0, 20, 50, 70 The inclination of the line of sight i 20, 40, 60 The angle measured between the equatorial plane and the edge of the torus 5, 7 The average edge-on optical depth at 9.7 μm 0.05 Additional power-law index modifying the attenuation curve denoted in Boquien et al. (2019) Clumpy two-phase torus model: SKIRTOR2016 (Stalevski et al. 2012, 2016) τ 5, 7 The average edge-on optical depth at 9.7 μm θ 20, 40, 60 The angle measured between the equatorial plane and the edge of the torus δ 0, 20, 50, 70 The inclination of the line of sight R_out/R_in 20 The ratio of the outer radius to the inner radius fracAGN 0.01, 0.1, 0.2, ..., 0.99 The fraction of AGN IR luminosity in the total IR luminosity may be a significant fraction of optically unobscured Type 1 AGNs among our AGN host galaxies, we allowed these parameters to vary within wider ranges.

With the settings in Table 2, we computed over 500 million models. In Table 3, we compare our median stellar masses and SFRs of G − M_20 + V classification samples (see Section 5.1) with those in COSMOS2015 at three redshift bins, and Figure 4 shows the corresponding values in the COSMOS2015 catalog as a function of our computed values. SFRs in COSMOS2015 show large scattering and are 2 dex higher than ours in the extreme case, which can be a consequence of the absence of IR data and the ignorance of AGN components. The medians of the stellar masses are closer and most of the new estimates are 0.05–0.5-dex smaller than those in COSMOS2015. Figure 5 depicts the SED fitting results of the object ID003 at z = 0.977 with a visible merger-like morphology. The host galaxy is rich in emission lines, suggesting ongoing star formation. The solid orange line indicates the emission from the AGN, and the AGN contributes a significant fraction to the observed fluxes, from which we may infer that this is a Type 1 AGN with a visible nucleus. The computed SFR and inclination are 22.97 M_⊙ yr^{-1} and 21.11 ± 21.96, respectively.

### 4. Results

In this section, we will present the results of our morphological measurements along with those of previous studies on AGN hosts and normal galaxies in the literature. First, we describe our final sample to be discussed further.

There were 491 objects in Kimura et al. (2020), and 485 of these were found in the HST-COSMOS database. By applying the selection criteria described in Section 3.3, only 76 objects remained. There were another 26 objects that failed to pass the criteria due to invalid Gini segmentation or outlying r_e/sersic_rhalf, but their morphology could be confidently determined by visual classification. Therefore, these 26 objects were kept, while the measurements of both parametric and nonparametric parameters were not discussed other than the SED fitting results. Unfortunately, all AGNs at z > 3 failed to pass the criteria because the PSF-like component dominated the entire structure of their host galaxies. Finally, there were only five valid objects at 1.5 < z < 3.0.

For the 76 remaining objects found in the COSMOS2015 catalog, we computed the host galaxies’ stellar masses and SFRs by using X-CIGALE as described in the previous section. The redshift−mass and redshift−SFR distributions are plotted in the upper and lower panels of Figure 6, respectively. There are 72 (≈94.7%) host galaxies that lie within the star formation main sequence (see Figure 15), which can be considered as star-forming galaxies (SFGs), with a median mass of 2.81 × 10^{10} M_☉. The central AGNs of these 76 objects have L_{AGN,median} ~ 1.11 × 10^{47} W, which is more than 1 dex fainter than the typical bolometric luminosity of samples in SDSS Quasar DR12 (Kozłowski 2017). Therefore, we believe that these valid objects are good samples of low-luminosity populations. In the following subsections of morphological measurements, we focus only on these 76 valid objects.
4.1. Visual Classifications

Visual classification is the most original and explicit method to determine the morphology of a galaxy. However, owing to the difficulties in the classification of edge-on galaxies and those at high redshifts, where galaxies are more compact, we decided to follow the Galaxy Zoo 2 (GZ2; Lintott et al. 2011; Willett et al. 2013) field guide to classify the total 102 objects into four groups: spirals, ellipticals, irregulars/mergers, and point sources. Even after the rejection of point-like sources for which we could not obtain reliable morphological measurements as discussed in Section 3.3, we still found many small and compact sources among the valid objects, forcing us to

Figure 4. Star formation rates (SFRs, left panel) and stellar masses (right panel) in the COSMOS2015 catalog as a function of the values computed in this work.

Figure 5. An example of the best-fit models obtained from our SED fitting performed by X-CIGALE. This is the SED of the object ID003, a galaxy with visible merger-like morphology at $z = 0.977$. The reduced-$\chi^2$ indicates the goodness of the fit, and 0.56 ensures a high-quality and realistic fitting. The nebular emission contributes several emission lines in the Model spectrum including the Ly$\alpha$ and H$\alpha$ lines at $\lambda \sim 0.24$ $\mu$m and 1.3 $\mu$m, respectively, suggesting that the galaxy is undergoing star-forming activity.

Figure 6. Top: stellar mass of the entire galaxy sample computed by X-CIGALE as a function of the redshift. The dashed lines divide the sample into three bins in stellar mass. Bottom: SFR averaged over the recent 100 Myr computed by X-CIGALE as a function of the redshift. The solid and dashed lines divide the sample into three bins in SFR and redshift, respectively. The blue crosses represent host galaxies with AGNs detected by Chandra, while coral triangles are AGNs undetected in X-rays.
compact ellipticals lie on the boundary of ellipticals and point sources. Patchy star-forming regions may be classified as irregular/merger, and point sources. We show examples of the four classes of galaxies in Figure 7. The result of the visual classification is listed in Table 4. Notably, there exist some objects that could not be confidently classified via a purely visual inspection. These objects include point sources with spatial extensions of several pixels and that are classified as ellipticals and galaxies with asymmetric arms that cannot be distinguished between spiral galaxies or late-stage mergers. Tiny arm-like substructures are also difficult to distinguish between spiral arms or streams.

By visual investigation, many spiral galaxies showed some disturbed features such as asymmetric spiral arms and streams, which suggested that these galaxies were undergoing strong star formation, merger activities, or interactions. Owing to such difficulties, some visual classifications might be suspicious; we will discuss morphological classifications based on their Sérsic index, Gini−M20 and log(Gini)−log(Asymmetry) diagnostics in Section 5.1.

This visual method was also used as a criterion to decide the correction factor for the Sérsic index and nonparametric parameters, as introduced in Section 3.3. Among the 76 valid objects, 26 had visible PSF components in their centers, corresponding to the modeled host galaxies that had their central regions not three times brighter than AGNs, and 50 did not have such visible PSF-like components.

### 4.2. Sérsic Index and Effective Radius

In this section, we compare the results of 2D Sérsic fitting with some other AGN host galaxy samples in the literature. First, we investigated the influence of the correction on distributions of the Sérsic index.

The values of the Sérsic index were corrected for bias due to the central bright AGN (Section 3.3). To clearly show this effect, we compare the values before and after the correction in Figure 8. The corrected histogram significantly shifted to smaller Sérsic index values compared to the histogram before correction. Indeed, the Sérsic index measurements without correction were biased toward higher values, with some other AGN host galaxy samples in the literature. This led to a significant shift in the Sérsic index values, with a correction to correct the bias.

In the upper panel of Figure 9, our results are compared with the host galaxies of IR-selected AGNs within 0.5 < z < 1.5 that have log(M⋆/M⊙) > 10.5 (Chang et al. 2017), whereas ours range from ∼109 M⊙ to ∼1011 M⊙. The IR-selected AGNs were significantly obscured and the host galaxies could be considered as a degradation of, not a contamination from, AGNs. Considering n = 2 (Ravindranath et al. 2004; Cassata et al. 2011) as the boundary that divided disk and spheroid systems, most host galaxies of IR-AGNs were disk systems in the top-left panel of Figure 9. However, the Sérsic index of the host galaxies of our optical variability-selected AGNs had a concentration at n ∼ 3 along with a higher cumulative probability (∼60%) at n > 2. This led to a significant difference in the n distribution of these two samples in a Kolmogorov

---

**Table 4**

| Classification       | Number |
|----------------------|--------|
| Spiral               | 32     |
| Elliptical           | 30     |
| Irregular/merger     | 17     |
| Point-source         | 23     |

**Note.**

4 The purely visual inspection has uncertainties due to the fact that spirals with patchy star-forming regions may be classified as irregular/merger and compact ellipticals lie on the boundary of ellipticals and point sources.
−Smirnov test (KS test, \( p\)-value \( \ll 0.01 \)). The distributions of \( r_e \) of these two samples were similar, but we found an absence of our objects in the most compact hosts (\( r_e \leq 0.42 \) kpc) and extended (\( r_e > 10 \) kpc) cases compared to the distributions of IR-AGN hosts. This led to a less significant (\( p\)-value \( \approx 0.04 \)), but still possible, difference between the two samples. We explain this as being due to the fact that our most compact hosts were dominated by the AGNs because of unobscuration and they failed to pass the selection criteria. In addition, the average size of the IR-AGN host galaxies was about 3.56 kpc (\( r_e = 10^{0.552} \) kpc, measured error considered), larger than ours (\( r_e = 10^{0.438} \approx 2.74 \) kpc). However, considering the significant fraction of extremely extended IR-AGN host galaxies, we found no evidence of more compact sizes of hosts for the optical variability-selected AGNs.

We also compare with the host galaxies of \( \sim 5000 \) quasars from SDSS DR14 at \( 0 < z < 1.0 \), \( L_{bol} = 10^{44.0 \pm 0.5} \text{ erg s}^{-1} \), and \( 9.5 < \log(M_*/M_\odot) < 11.5 \) (Li et al. 2021) in the middle panel of Figure 9. The morphology of these quasar host galaxies was studied using HSC five-band (grizy) optical imaging data. Considering light contributions of quasars to the hosts, Li et al. (2021) also made corrections for the Sérsic index by adding AGNs to the centers of galaxies with different magnitudes. Apparently, the difference of \( n \) between the two samples was highly significant (\( p\)-value \( \ll 0.01 \)). However, in this comparison, our host galaxies showed a spheroid dominance while the majority of quasar host galaxies were still disk systems. In addition, Li et al. (2021) performed corrections for \( n \) as well; for some quasar-dominated hosts fainter than HSC \( i \) band \( \sim 23 \) mag, their intrinsic parameters could not be recovered. Such a

---

**Figure 9.** Comparison of Sérsic index and \( r_e \) in the literature. The redshift ranges of our samples match with the comparisons. Top: comparison with IR-selected AGNs at \( 0.5 < z < 1.5 \) (Chang et al. 2017). Middle: comparison with quasars observed with Subaru HSC at \( 0.2 < z < 1.0 \) (Li et al. 2021). Bottom: comparison with normal galaxies in the COSMOS field at up to \( z \sim 1 \) (Sargent et al. 2007; Scarlata et al. 2007). Variables \( \bar{n} \) and \( \bar{r_e} \) are average values of the Sérsic index and the effective radius, respectively.
difficulty in the measurements of low-luminosity hosts was also seen in our results before the valid sample selection; over \(\sim 60\%\) host galaxies were fainter than \(I_{\text{mag}} \sim 22.5\) mag and were dominated by AGN components, and these objects had both \(S=0\) and \(flag_{\text{sersic}}=1\). We removed all such objects and many of them even had a Sérsic index \(n > 10\). Significantly different \((p\text{-value} < 0.01)\) distributions of \(r_e\) could also be seen, such that our sample was much more compact than the SDSS quasars. We attributed this difference to the fact that Li et al. (2021) used HSC imaging data. Because of atmospheric seeing \((O''7)\), objects in HSC images were more extended, resulting in larger \(r_e\) values.

It is also interesting to compare with 16,538 normal galaxies selected with the HST \(I\)-band magnitude \((I_{\text{F814W}} < 22.5)\) from the Zurich Structure & Morphology catalog (Sargent et al. 2007; Scarlata et al. 2007), shown in the bottom panel of Figure 9. Among these normal galaxies, \(\sim 12,000\) were disk systems. The distributions of \(n\) significantly differed \((p\text{-value} \ll 0.01)\), which was attributed to both the spheroid dominance in the AGN host galaxies and the fact that disk systems that host AGNs had higher \(n\) than normal disk systems. For \(r_e\), our host galaxies were smaller in size \((p\text{-value} \ll 0.01)\) and showed clearly separated concentrated regions. As indicated by the average values, the size of our AGN hosts was roughly \(70\%\) of that of normal galaxies.

As suggested in Kimura et al. (2020), optical variability-selection is more efficient at selecting LLAGNs. This was supported by the result of SED fitting, in that our AGNs had a median luminosity of \(L_{\text{bol}} = 1.11 \times 10^{44.6}\) erg s\(^{-1}\), while SDSS quasars spanned \(L_{\text{bol}} = 10^{44.0-46.5}\) erg s\(^{-1}\) with a median at \(L_{\text{bol}} \sim 10^{45.2}\) erg s\(^{-1}\), which was larger than ours by 1 dex (Li et al. 2021). LLAGNs are likely to be hosted by super-massive black holes (SMBHs) that have lower accretion rates, which are more likely to reside in the centers of elliptical galaxies (Heckman & Best 2014). This explains the different morphological dominance seen in the three comparisons as well as the more compact sizes when they are compared with the sample of normal galaxies, \(70\%\) of which are spiral galaxies (Sargent et al. 2007). The smaller sizes may also imply that our AGN host SFGs are undergoing a process of dynamical compaction, probably arising from gas inflow. Cosmological hydrodynamical simulations of galaxy formation suggest that highly perturbed wet disks fed by cold streams may experience a dissipative contraction phase (Dekel & Burkert 2014). In this scenario, the gas inflow, which is often associated with disk instability, toward the central region of the galaxy triggers the initial contraction and acts as the energy provider to maintain substantial high-level turbulence, leading to a massive core and enhanced SFRs, along with the triggering of accretion onto the SMBH and induced AGN activity (Bournaud et al. 2011; Zolotov et al. 2015).

4.3. Nonparametric Parameters

Nonparametric parameters provide more detailed investigations of the light distributions of our host galaxies. There are four parameters presented in the results: Gini \((G)\), \(M_{20}\), Concentration \((C)\), and Asymmetry \((A)\).

In the upper panel of Figure 10, we plot the distributions of the GMCs of our AGN host galaxies and those of normal galaxies from Sargent et al. (2007). The correction described in Section 3.3 was also applied to our measurements, although the AGN effects were small. Apparently, normal galaxies have more even light distributions, as seen from their smaller \(G\) values. Combined with \(M_{20}\) and \(C\), it implies there are brighter central regions in the host galaxies of our AGN sample. Given that we corrected for the bias due to the central AGN component in the measurements, the brighter central parts...
might serve as evidence of central star formation induced by the AGN. As shown by Asymmetry, there was only a minute increase compared with normal galaxies, suggesting insignificant impacts of AGNs on the global structures of the host galaxies.

In the bottom panel of Figure 10, we compare the results with the COSMOS Zamojski Morphology catalog (Zamojski et al. 2007). This catalog includes 8146 SFGs selected based on HST ACS I-band photometry with $I_{F814W} < 23$ mag at $0.2 < z < 1.0$ and with $9.5 < \log(M_*/M_\odot) < 11.6$. The distributions of $G$, $M_{20}$, and $C$ show the same pattern as for the normal galaxies of Sargent et al. (2007). Meanwhile, the distributions of $A$ differ. Their SFGs show an extended tail to high $A$ values in the distributions, indicating intense and inhomogeneous star formation in their disks. Our AGN host galaxies are comparatively much more symmetric, which is also shown by the comparison in the upper panel. The significant fraction of spheroid systems implied by the Sérsic index might be an explanation of the difference in the distributions of $A$.

We also plot the GMCA measurements of our AGN samples within three redshift, stellar mass, and SFR bins in Figure 11. The average values of $G$ and $A$ at different redshift bins are similar, whereas $M_{20}$ and $C$ show a slight evolution pattern. This could be explained as being the result of the size evolution with cosmic time, i.e., at higher redshifts, a galaxy is more compact, and spheroid systems show a steeper decrease than disk systems (Conselice 2014), along with surface brightness dimming with redshift (Lotz et al. 2006). Therefore, with a smaller spatial extension, the contribution of the brightest pixel for high-$z$ objects does not influence the measurement as significantly as for low-$z$ objects. The rest-frame UV light would be redshifted to the wavelengths at which we search for high-$z$ objects. Therefore, we expect more UV light for higher-$z$ galaxies. Because star-forming regions are the main contributor in UV, more disturbed and asymmetric structures can be expected as seen in Zamojski et al. (2007). However, such an asymmetry evolution with redshift is not clearly seen in $A$.

Comparisons of different stellar mass and SFR bins are shown in the middle and lower panels of Figure 11, respectively. We can barely see any differences between $G$, $M_{20}$, and $C$ for both comparisons, indicating that the stellar masses and star formation have no clear dependencies on the global morphology. In addition, as seen in the distribution of $A$ for different SFR bins, although there is an increase in $A$ toward higher SFRs, the evolution is quite insignificant compared with that of normal SFGs, whereas these hosts of optical variability-selected AGNs are also SFGs. Such results again require a
morphological classification to explain them and suggest that the star-forming activity in these hosts might more likely be related to the galaxies’ central regions.

4.4. AGN Fraction Contribution

Through SED fitting, we also computed the AGN fraction defined as the AGN luminosity fraction in the total (AGN + galaxy) IR luminosity (Boquien et al. 2019; Yang et al. 2020). In this section, we investigate the relations of the AGN fraction with the Sérsic index, $n$, and the effective radius, $r_e$, to discover whether these parameters correlate with the AGN fraction.

Figure 12 shows the relations between the AGN fraction and the Sérsic index, $n$, as well as the effective radius, $r_e$. The effective radius $r_e$ and Sérsic index $n$ show negative correlations with the AGN fraction. Their linear fittings with the AGN fraction are given by: $r = -(1.61 \pm 0.91)f + (3.38 \pm 0.37)$ and $n = -(0.41 \pm 0.85)f + (2.52 \pm 0.35)$, where $r$ is the effective radius, $n$ is the Sérsic index, and $f$ is the AGN fraction. According to these equations, a 20% AGN contribution leads to a 3.3% decrease of the Sérsic index and a 9.5% decrease in the $r_e$ compared with those at a zero AGN fraction, respectively. A 50% AGN contribution corresponds to an increased Sérsic index by 8.1% and a decreased $r_e$ by 23.8% compared with the case without any AGN effect.

To measure the significance of these correlations, we calculated Pearson’s correlation coefficients. For $r_e$, it had a $p$-value of 0.08 > 0.05 and a correlation coefficient of $r = -0.2$, and for the Sérsic index, the $p$-value and correlation coefficient were 0.63 > 0.05 and $r = -0.06$, respectively. It suggests that $r_e$ is weakly correlated with the AGN fraction and the correlation is not statistically significant. This weak correlation between $r_e$ and the AGN fraction agrees with the process of dynamical contraction that arises from the central gas inflow. Meanwhile, the Sérsic index, in principle, does not correlate with the AGN fraction at all, which might be because we corrected for the influence of AGN components and removed extremely compact objects. We calculated the $p$-values of the AGN fraction between nonparametric parameters. However, all of them had a $p$-value $\geq 0.05$, indicating no arguable correlations between the AGN fraction and nonparametric parameters.

5. Discussion

5.1. Nonparametric Parameter Diagnostics

Other than visual classification, a method to determine which class the galaxy belongs to uses diagnostics based on nonparametric parameters, including $G$–$M_{20}$ ($G$–$M_{20}$) and $\log$(Gini)–$\log$(Asymmetry); ($\log$(G)–$\log$(A)) diagnostics.

In the left panel of Figure 13, we plot the $G$–$M_{20}$ diagnostics that divide the region into three parts: E/S0/Sa, Sb/Sc/Irr, and Mergers. The division lines are obtained from Lotz et al. (2008) with the following definitions for Extended Groth Strip (EGS) galaxies at $0.2 < z < 1.2$: Mergers: $G > -0.14M_{20} + 0.33$, E/S0/Sa: $G \leq -0.14M_{20} + 0.33$, and $G > 0.14M_{20} + 0.80$; and Sb/Sc/Irr: $G \leq -0.14M_{20} + 0.33$ and $G \leq 0.14M_{20} + 0.80$. We show the fractions of different identifications at different redshifts in Table 5. Clearly, most AGN host galaxies at low redshifts reside in the E/S0/Sa category, whereas at higher redshifts, the probability of them being found in the Sb/Sc/Irr region increases. In addition, the fraction of mergers is small within the full redshift range, suggesting that mergers are not the major mechanism triggering AGN activity, consistent with the findings of Chang et al. (2017). However, as the redshift increases, the merger fractions increases as well, which may cause a sudden strong gas inflow toward the galaxy center and eventually trigger AGN activity. In our visual classification, many spiral galaxies show disturbed features, which might support this implication.

The $\log$(G)–$\log$(A) diagnostics are plotted in the right panel of Figure 13. The total sample had a size of 71 objects because five objects had negative $A$ values due to high background noise levels. The division lines are defined by Capak et al. (2007a), who studied the morphology of the galaxies in the COSMOS field at $0 < z < 1.2$: the division line between the irregulars and spirals is $\log_{10}A = 2.353 \times \log_{10}G + 0.353$, and the division line between the spirals and ellipticals is: $\log_{10}A = 5.50 \times \log_{10}G + 0.825$. In the entire redshift range, 65 ($\sim$91.5%) are

| Redshift | $G$–$M_{20}$ diagnostics$^a$ | $G$–$M_{20}$ diagnostics + Visual classification$^b$ |
|----------|-------------------------------|-----------------------------------------------|
| $0 < z < 1.0$ (59) | 66.1% 20.0% 0% | 56.1% 36.8% 0% |
| $1.0 < z < 1.5$ (12) | 50.0% 41.7% 8.3% | 52.8% 38.6% 8.6% |
| $1.5 < z < 3.0$ (5) | 20.0% 60.0% 20.0% | 20.0% 60.0% 20.0% |
| $0 < z < 3.0$ (102) | G–$M_{20}$ diagnostics | 46.1% 44.1% 9.8% |

Note.
$^a$ Numbers of objects within each redshift bin are shown in parentheses.
$^b$ Twenty-six objects with a visually confirmed morphology that did not pass the selection criteria are included in the $G$–$M_{20}$ + V classification.
adapted the formula in Capak et al. (2007a) to divide the regions. Gray histograms in the distribution simply show the overlapping objects.

classified as elliptical galaxies, and only six objects (~8.5%) are spiral galaxies. Further, there are no irregular or merging galaxies. The number of spiral galaxies in this diagnostics is even smaller than that of the visually confirmed spirals. We infer that such distributions can be attributed to the low A and larger G at the same time. The A values of these AGN hosts have only a slight increase compared with normal galaxies, whereas they have much more uneven light distributions.

Although according to the $G - M_{20}$ diagnostics 46 objects (~60.5%) are E/S0/Sa, because this class includes S0 and Sa, we expected a smaller fraction of ellipticals. Further, the 2D Sérsic fitting depends on the underlying mathematical form, which means it cannot be used as an indicator of merging/irregular galaxies. In addition, when we analyze these systems, errors may occur and reduce the valid sample size. Therefore, we combine the $G - M_{20}$ diagnostics and visual classifications ($G - M_{20} + V$) to determine the final morphology result for valid objects, which includes objects that are successfully measured and objects with visually recognizable morphological structures that failed to pass the selection criteria. In $G - M_{20}$ diagnostics, 17 of 22 visually classified ellipticals are classified as E/S0/Sa, whereas only 10 of 20 visually classified spirals are classified as Sb/Sc/Irr. Among the other 10 visually classified spiral galaxies, six had two visible spiral arms, by which they were classified as Sa. We took the results of the $G - M_{20}$ diagnostics as the priority and visually re-investigated the objects with inconsistent classifications between the two methods to avoid misclassifications of $G - M_{20}$ diagnostics because of unenclosed faint substructures. The visually confirmed S0/Sa galaxies were all classified as disk systems, whereas the remaining ones were classified as spheroid systems. The final classifications of $G - M_{20} + V$ are listed in Table 5.

As seen in the $G - M_{20}$ diagnostics, the dominant system is a spheroid system at $z < 1.5$. Although at $z > 1.5$, disk galaxies have a higher fraction, because of the limited sample size, this dominance is unarguable. At $z \lesssim 1$, the ellipticals occupy the absolute majority and the number of ellipticals is even over half of all ellipticals within the entire redshift range. One possible explanation for this result could be due to less UV light, which primarily arises from star-forming activity, being received by the F814W filter at lower redshifts. In addition, spheroid systems experience a size evolution that drops more steeply with cosmic time than for disk systems, resulting in a smaller fraction of detectable ellipticals at a higher redshift. This is clearly shown in the valid sample selection, such that ~70% of host galaxies are excluded because of a PSF dominance, caused by small $r_e$, large $n$, and bright AGN components.

5.2. Implications for Galaxy–AGN Co-evolution

As shown in the upper panel of Figure 14, the majority of the hosts of our variability-selected AGNs are undergoing star-forming activities. To investigate if they have a stronger SFR than do normal SFGs that lay around the star formation main sequence at the corresponding epochs, we plot the sSFR against stellar mass in Figure 15. These star-forming AGN hosts lie within the main sequence if a scatter of ±1 dex is considered. This suggests that there is no intensively enhanced SFR in the spiral galaxies that are intrinsically SFGs. But interestingly, for ellipticals, which are representative of quiescent galaxies with low SFRs, we have found that they around the star formation main sequence, which suggests that their star-forming activities might be triggered by AGNs. Such star-forming elliptical galaxies hosting AGNs might explain a strange pattern shown in the Asymmetry comparison in the lower panel of Figure 10, i.e., that star-forming AGN host galaxies show more symmetric structures than do normal SFGs. The Asymmetry within different SFR bins in the lower panel of Figure 11 is a complement to this pattern: higher SFRs only have negligible impacts on the global structures of the host galaxies. However, there is an almost equal fraction of disk systems, which should be more asymmetric than ellipticals.
Figure 14. Top: AGN luminosity as a function of the SFR. Bottom: AGN luminosity as a function of redshift. The meaning of the symbols is indicated in the panel. For both panels the classifications are based on the combination of the $G-M_{20}$ diagnostics and visual classification. The dashed lines in the upper and lower panels indicate the median values of log(SFR) and log($L_{AGN}$), respectively.

Figure 15. Specific star formation rate (sSFR) as a function of galaxy mass. Different types of galaxies represented by different shapes are based on $G-M_{20}$ diagnostics + visual classification. The different lines show the SFR main sequences at different redshifts adapted from Chang et al. (2017).

In this study, we studied the morphology of host galaxies with optical variability-selected AGNs at $0 < z < 3$ in the COSMOS field. The host morphology was evaluated using parametric ($n$) and nonparametric ($G$, $M_{20}$, $C$, and $A$) morphological parameters and investigated with SFRs using HST imaging ($\sim 0.03$) and SED fitting. Our main conclusions are as follows.

1. The sizes of host galaxies of the optical variability-selected AGNs up to $z \sim 1$ are more compact than for normal galaxies at the same redshift and stellar mass range by 35.7%.
2. The host galaxies of these optical variability-selected AGNs have no clear morphological preference, as seen in attributed to disk instabilities or that they are late-type mergers. Except for the disturbed galaxies, we also have found 10 mergers or irregular galaxies in the imaging data. The basic conclusion for the small merger rate at $0 < z < 3.0$ agrees with many previous studies, arguing that mergers are not the primary triggering mechanism of AGN activities (Elmegreen et al. 2008; Bournaud et al. 2012). However, in the lower panel of Figure 14, we do find that mergers are more likely to be found at higher redshifts, which agrees with the predictions of cosmological hydrodynamic simulations (Rosas-Guevara et al. 2016). We also have checked the AGN fractions, AGN luminosities, SFRs, and stellar masses of host galaxies, and we have found that, with a higher SFR and AGN luminosity, a galaxy has a larger possibility to be a merger-like object, as shown in the bottom panel of Figure 14. Such a finding agrees with the X-ray-selected and HST/WFC3 imaged heavily obscured AGNs at $z \sim 1$ (Kocevski et al. 2015). This can be understood as an aspect of the evolution of the massive galaxy, i.e., the most massive systems grow mainly through hierarchical merging activities. This is more ubiquitous at younger cosmic times relative to the local universe. Then, the merging galaxies induce a sudden gas inflow into their central regions that feeds strong AGN activities compared with a steady gas inflow. Intensive SFRs and luminous AGNs are expected to be the products of these major mergers.

6. Conclusions

Our visual classification reveals that a relatively larger fraction of spiral galaxies have disturbed features. This indicates the existence of possible interactions or merger activities and conflicts with the results of Cisternas et al. (2011), who studied the hosts of X-ray-selected Type 1 AGNs at $z \sim 0.3$–1.0 and found that 85% of their objects showed normal undisturbed morphological patterns. Owing to the limitation of observations, we cannot say with certainty that these features are
the number fractions of disk (~44.1%) and spheroid (~46.1%) systems.

3. Almost all AGNs (~94.6%) reside in SFGs (log(SFR)$_{\text{med}}$ ~ 0.7 $M_\odot$ yr$^{-1}$) that have a very similar Asymmetry distribution compared with that of normal galaxies, and much more symmetric structures compared with normal SFGs. This can be explained by the fraction of elliptical galaxies (44.9%), suggesting that AGN feedback enhances the star formation of spheroid systems and the star-forming activity influenced by AGN feedback only varies within a small central scale rather than across the entire system.

4. The fraction of major mergers in the variability-selected AGNs is as small as ~9.8%, which suggests that major mergers are not the main triggering mechanism of AGN activities; however, the merger rate increases with increasing redshift, AGN luminosity, and SFR values.

We thank Yu-yen Chang, Mark Sargent, Michel Zamojski, and Junyao Li for providing their data in the literature and the referee for the useful comments. This work would have not be possible without Vicente Rodriguez-Gomez’s technical support. This work used the NASA/IPAC Infrared Science Archive (IRSA) at the NASA Infrared Processing and Analysis Center (IPAC), located on the campus of the California Institute of Technology (Caltech). We gratefully acknowledge the contributions of the entire COSMOS collaboration. The COSMOS team in France acknowledges support from the Centre National d’Études Spatiales.

Software: statmorph (Rodríguez-Gomez et al. 2019); Photutils (Bradley et al. 2020); Tiny Tim (Krist et al. 2011); Astropy (Astropy Collaboration et al. 2013, 2018); X-CIGALE (Boquien et al. 2019; Yang et al. 2020).

Appendix

Correction Formulae for the Sérsic Index

In this Appendix we describe how we correct the measured Sérsic index $n$ via the modeling results in Section 3.3. In Figure 16, we plot the fraction of the intrinsic Sérsic index $n_{\text{intrinsic}}$ relative to the measured value $n_{\text{measure}}$ against the ratio of the effective radius $r_e$ to the effective radius of the 2D Sérsic profile. We then performed polynomial fitting to the fifth degree to find how $n_{\text{intrinsic}}/n_{\text{measure}}$ and $r_e/sersic$_{rhalf} are correlated for different radii.

For real objects with $r_e < 8$, the polynomial fitting for models with $r_e = 5$ is applied:

$$
\frac{n_{\text{intrinsic}}}{n_{\text{measure}}} = -384.7x^5 + 2729x^4 - 7679x^3 + 10710x^2 - 7408x + 2033, \quad (A1)
$$

where $x = r_e/sersic$_{rhalf} and lies within 1.05 $\leq x \leq$ 1.70. For real objects with 8 $\leq r_e < 12$, the polynomial fitting for models with $r_e = 10$ is applied:

$$
\frac{n_{\text{intrinsic}}}{n_{\text{measure}}} = 893.1x^5 - 4989x^4 + 11020x^3 - 12030x^2 - 6507x - 1394, \quad (A2)
$$

where $x = r_e/sersic$_{rhalf} and lies within 0.82 $\leq x \leq$ 1.27. For real objects with 12 $\leq r_e$, the polynomial fitting for models with $r_e = 10$ is applied:

$$
\frac{n_{\text{intrinsic}}}{n_{\text{measure}}} = 2225x^5 - 11200x^4 + 22450x^3 - 22420x^2 + 11150x - 2208, \quad (A3)
$$

where $x = r_e/sersic$_{rhalf} and lies within 0.78 $\leq x \leq$ 1.18.

As described in Section 3.3, the $r_e$ and $sersic$_{rhalf} are calculated for each real object and, if the fraction of these two parameters is out of the range of $x$ for each equation, there will be no correction available. For valid objects, the Sérsic index can be corrected via simply multiplying the polynomial as a function of $r_e/sersic$_{rhalf} by the measured Sérsic index.

![Figure 16](image.png)

Figure 16. The fraction of the intrinsic Sérsic index $n_{\text{intrinsic}}$ to the measured value $n_{\text{measure}}$ as a function of the ratio of the effective radius $r_e$ to the effective radius of the 2D Sérsic profile $sersic$_{rhalf}. From left to right, the preset effective radii are 5, 10, and 15 pixels, respectively.
