The Obscured Fraction of Quasars at Cosmic Noon

Bovornpratch Vijarnwannaluk, Masayuki Akiyama, Malte Schramm, Yoshihiro Ueda, Yoshiki Matsuoka, Yoshiak Toba, Marcin Sawicki, Stephen Gwyn, and Janek Pfugradt

1 Astronomical Institute, Tohoku University, Aramaki, Aobaku, Sendai, Miyagi 980-8587, Japan; bovornpratch.s@astro.tohoku.ac.jp
2 Graduate school of Science and Engineering, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama City, Saitama 338-8570, Japan
3 Department of Astronomy, Kyoto University, Kitashirakawa-Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan
4 Research Center for Space and Cosmic Evolution, Ehime University, 2-5 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan
5 National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
6 Institute for Computational Astrophysics and Department of Astronomy and Physics, Saint Mary’s University, Halifax, NS B3H 3C3, Canada
7 Canadian Astronomy Data Centre, NRC-Herzberg, 5071 West Saanich Road, Victoria, British Columbia, V9E 2E7, Canada
8 Academia Sinica Institute of Astronomy and Astrophysics, 11F of Astronomy-Mathematics Building, AS/NTU, No.1, Section 4, Roosevelt Road, Taipei 10617, Taiwan

Received 2022 April 15; revised 2022 September 28; accepted 2022 September 29; published 2022 December 14

Abstract

Statistical studies of X-ray selected active galactic nuclei (AGN) indicate that the fraction of obscured AGN increases with increasing redshift, and the results suggest that a significant part of the accretion growth occurs behind obscuring material in the early universe. We investigate the obscured fraction of highly accreting X-ray AGN at around the peak epoch of supermassive black hole growth utilizing the wide and deep X-ray and optical/IR imaging data sets. A unique sample of luminous X-ray selected AGNs above $z > 2$ was constructed by matching the XMM-SERVS X-ray point-source catalog with a point-spread function convolved photometric catalog covering the $u' - 4.5 \mu m$ bands. Photometric redshift, hydrogen column density, and 2–10 keV AGN luminosity of the X-ray selected AGN candidates were estimated. Using the sample of 306 2–10 keV detected AGN at above redshift 2, we estimate the fraction of AGN with $\log N_H (cm^{-2}) > 22$, assuming parametric X-ray luminosity and absorption functions. The results suggest that 76% of luminous quasars ($\log L_X (erg s^{-1}) > 44.5$) above redshift 2 are obscured. The fraction indicates an increased contribution of obscured accretion at high redshift than that in the local universe. We discuss the implications of the increasing obscured fraction with increasing redshift based on the AGN obscuration scenarios, which describe obscuration properties in the local universe. Both the obscured and unobscured $z > 2$ AGN show a broad range of SEDs and morphology, which may reflect the broad variety of host galaxy properties and physical processes associated with the obscuration.

Unified Astronomy Thesaurus concepts: Active galactic nuclei (16); Supermassive black holes (1663); Quasars (1319)

Supporting material: machine-readable table

1. Introduction

Observational results indicate that supermassive black holes (SMBHs) exist ubiquitously in all massive galaxies (see review by Kormendy & Richstone 1995; Kormendy & Ho 2013). However, how such massive black holes grow over cosmic time is not well understood. Active galactic nuclei (AGN) represent a key phase of SMBH growth during which SMBHs are actively accreting mass. The number density of AGN peaks at redshift 1–3, also known as the cosmic noon, and shows a trend in which the number density of more luminous AGN peaks earlier at redshift $\sim 3$ than that of less luminous AGN and then declines toward the local universe (Ueda et al. 2014; Delvecchio et al. 2014; Aird et al. 2015). This era represents a crucial period where the bulk of the cosmic SMBH mass density (90%) was gained through mass accretion in luminous quasars (Soltan 1982; Marconi et al. 2004; Delvecchio et al. 2014; Ueda et al. 2014).

One large uncertainty in tracing the SMBH accretion growth is the fraction of obscured accretion. Studies using hard X-rays above 10 keV (Malizia et al. 2009; Burlon et al. 2011) and cosmic X-ray background synthesis studies have shown that a non-negligible fraction of AGN at low redshift is obscured (Comastri et al. 1995; Ueda et al. 2003; Ballantyne et al. 2006; Gilli et al. 2007; Treister et al. 2009).

Although optical imaging surveys have been successful in constructing large samples of quasars at high redshifts, identification of obscured AGN using optical color selection is challenging due to their similarity in color to galaxies with no ongoing AGN activity. Several emission-line diagnostic diagrams were constructed to identify obscured AGN from galaxies (Baldwin et al. 1981; Feltre et al. 2016). However, the classification requires spectra, which can be time expensive to obtain in large numbers for faint objects at high redshifts. Alternatively, multiwavelength AGN signatures such as X-ray, mid-infrared, or radio emission can also be used to identify obscured accretion activity. Among the tracers of AGN activity at high redshift, hard X-ray data sets ($E > 2 keV$) currently provide the most reliable indicator of AGN activity. In addition, they provide the most complete view of the high-redshift AGN population compared to other AGN selection methods thanks to the strong contrast against stellar light and the lower bias against obscuration.

In studies at low redshifts ($z < 2$), the obscured fraction shows a clear anticorrelation with AGN luminosity where more
luminous quasars are less likely to be obscured. This trend has been observed ubiquitously in various AGN samples selected by optical emission lines, X-rays, and mid-infrared emission (Simpson 2005; La Franca et al. 2005; Maiolino et al. 2007; Hasinger 2008; Burlon et al. 2011; Toba et al. 2013, 2021). Following the orientation-based AGN unification scheme (Antonucci 1993; Urry & Padovani 1995), the trend can be explained by the inner torus structure receding outward due to strong illumination and sublimation of dust; thus, the opening angle within which the central engine is directly observable increases with luminosity, or, in other words, the dust covering factor decreases (Lawrence 1991; Toba et al. 2014). Another possible scenario is that the obscured fraction is controlled by radiation pressure on dust particles: AGN blow out the obscuring material after exceeding an $N_{\text{HI}}$-dependent critical Eddington ratio (Fabian et al. 2006, 2008, 2009). Evidence supporting this scenario is found in AGN in the local universe showing that the nuclear column density depends on the Eddington ratio and AGNs whose accretion activity exceeds the critical effective Eddington ratio tend to show strong blowout winds (Fabian et al. 2009; Ricci et al. 2017a; Bär et al. 2019; Yamada et al. 2021; Toba et al. 2022).

High-resolution hydrodynamic simulations demonstrate that a torus-like structure is a natural outcome of gas accretion toward the nuclear region and that the torus properties are closely related to the SMBH and nuclear interstellar matter (ISM) properties (Hopkins et al. 2012; Wada 2012; Roth et al. 2012; Wada et al. 2016; Hopkins et al. 2016). These simulations can reproduce the observed column density distribution of AGN in the local universe as well as the obscured fraction by assuming that the torus is clumpy and that AGN feedback via radiation pressure clears some portions of sight lines (Hopkins et al. 2012; Wada 2012; Roth et al. 2012; Hopkins et al. 2016; Wada et al. 2016).

There is another trend among Compton-thin AGN (CTN-AGN) with $\log N_{\text{HI}}$ (cm$^{-2}$) $< 24$ the fraction of obscured AGN with $\log N_{\text{HI}}$ (cm$^{-2}$) $> 22$ increasing from the local universe up to redshift 2 (Ueda et al. 2003; La Franca et al. 2005; Ballantyne et al. 2006; Treister & Urry 2006; Hasinger 2008; Treister et al. 2009; Ueda et al. 2014; Aird et al. 2015; Buchner et al. 2015). Since most of the obscuring material is thought to be concentrated in the nuclear region (Hickox & Alexander 2018), the redshift dependence suggests that the nuclear region contains a larger amount of gas analogous to the larger gas fraction in galaxies at high redshifts than those in the local universe (Tacconi et al. 2010; Carilli & Walter 2013). Alternatively, the evolution of the obscured fraction among CTN-AGN may be driven by the X-ray obscuration from the host galaxy (Buchner et al. 2015, 2017; Buchner & Bauer 2017).

Beyond redshift 2, the obscured fraction was estimated to be larger than in the local universe but the behavior is still unclear. Some studies suggest that the obscured fraction is constant above redshift 2 (Hasinger 2008; Kalfountzou et al. 2014; Vito et al. 2016, 2014). Also in contrast to the obscured fraction below redshift 2, it was suggested that the fraction of AGN with $\log N_{\text{HI}}$ (cm$^{-2}$) $\geq 23$ at redshift 3–5 is independent of the X-ray luminosity (Vito et al. 2014, 2018). Some studies have found that the fraction of obscured AGN decreases with decreasing X-ray luminosity (Georgakakis et al. 2015).

The large uncertainty in the fraction of obscured quasars above redshift 2 is partly due to the limited sample size associated with the limited survey area and depth of X-ray surveys. Survey depth is important for the detection of high-redshift obscured AGN due to their X-ray faintness from obscuration and large distances. However, deep X-ray data sets are often limited to less than a few degrees of the sky. Furthermore, the number density of quasars beyond redshift 2 shows a strong decline with redshift thus a large survey area with sufficient depth is needed to construct a sizable sample of high-redshift obscured quasars (McGreer et al. 2013; Kalfountzou et al. 2014; Vito et al. 2014; Georgakakis et al. 2015; Vito et al. 2016, 2018).

While X-ray emission is a reliable tracer of accretion activity, it offers limited information on the AGN properties without distance estimates. Thus, X-ray sources must be matched with an optical/IR counterpart in order to determine their distance and properties, such as the luminosity and column density of the nuclear obscuration. This means a large and deep multiwavelength data set is needed to investigate the obscured fraction of quasars in the high-redshift universe.

Recent development in large and deep X-ray surveys, which cover many legacy multiwavelength deep fields, have allowed the investigation of the obscured fraction at high redshift. Of particular interest is the XMM-Spitzer Extragalactic Representative Volume Survey (XMM-SERVS) in the XMM-LSS region (Chen et al. 2018), which is also covered by the deep Optical Imaging data set from the Hyper Suprime-Cam Subaru Strategic Survey Program (HSC-SSP; Aihara et al. 2018) and the deep $U$-band imaging data from the CFHT Large Area $U$-band Deep Survey (CLAUDS; Sawicki et al. 2019) as well as previous legacy deep IR data sets.

In this study, the obscured fraction of luminous quasars above redshift 2 during the peak epoch of the quasar accretion growth was estimated by utilizing the unique wide and deep multiwavelength data set within the XMM-SERVS region. The deep $U$-band image plays a crucial role in deriving accurate photometric redshift for objects at $z > 2$ with the Lyman-break feature. Hereafter, obscured AGN refers to X-ray obscured AGN with $\log N_{\text{HI}}$ (cm$^{-2}$) $\geq 22$ unless otherwise stated. We also investigate the correspondence between the X-ray obscuration and the rest-frame UV/optical spectral energy distribution (SED). A galactic hydrogen column density of $3.57 \times 10^{19}$ cm$^{-2}$ in the survey area (Chen et al. 2018) and a flat standard Lambda dark matter cosmology with $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$ was assumed. Magnitudes are reported in the AB magnitude system.

2. Data

2.1. XMM-SERVS

The XMM-SERVS X-ray point-source catalog in the XMM-LSS region (Chen et al. 2018) was chosen for the primary selection of AGNs. The X-ray survey observations were performed by the XMM-Newton satellite over $5.3 \text{ deg}^2$ of the XMM-LSS survey field. XMM-Newton has three detectors, MOS1, MOS2, and PN. The catalog contains the combined detection of 5242 X-ray point sources using the three detectors in the 0.5–2, 2–10, and 0.5–10 keV bands. Figure 1 shows the 0.5–10 keV flare-filtered exposure time within the survey area, which reaches $\sim 50$ ks per pointing continuously over the survey area. In some regions, deeper X-ray observations are available from other projects as listed in Table 2 in Chen et al. (2018). The survey flux limits over 90% of the total area are $1.7 \times 10^{-15}, 1.3 \times 10^{-14}$, and $6.5 \times 10^{-15}$ erg cm$^{-2}$ s$^{-1}$ in the
0.5–2, 2–10, and 0.5–10 keV bands, respectively. The flux from each detector in the catalog was derived using an energy conversion factor assuming a power-law continuum with photon index, $\Gamma = 1.7$, and galactic absorption column density. The catalog provides the combined source flux calculated using the error-weighted average of the flux estimated by all three detectors. In this paper, the count rates were converted to the expected count rates detected with the PN detector (PN-equivalent count rates) in later discussions.

### 2.2. HSC-SSP

The HSC-SSP is an optical imaging survey performed by the 8.2 m Subaru telescope using the HSC imager (Miyazaki et al. 2018). HSC is a wide-field camera with a field of view of 1.5 deg$^2$. The camera is made up of 116 2K $\times$ 4K fully depleted back-illuminated CCDs (Kamata et al. 2012) mounted at the prime focus of the telescope. Among the HSC-SSP data sets with three different depths, imaging data of deep and ultradeep layers in the XMM-LSS field are utilized from the S19A and S20A internal releases. Each HSC pointing is shown in Figure 1 in red and blue dashed circles. S19A has better seeing in the HSC $r$ band than that in the S20A data release. The depth of the HSC deep layer is 27.4, 27.1, 26.9, 26.3, and 25.3 mag in the grizy bands (Aihara et al. 2022). The depth and typical seeing size are summarized in Table 1. The depth was calculated from the median value of the 5$\sigma$ point-source limiting magnitudes of all patches. Image data from HSC were reduced using the HSC pipeline (Jurić et al. 2017; Bosch et al. 2018, 2019; Ivezic et al. 2019). Photometric and astrometric calibration was performed against the first data release of the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS1; Schlafly et al. 2012; Tonry et al. 2012; Magnier et al. 2013; Chambers et al. 2016; Magnier et al. 2020). The HSC images are warped onto predefined grids called “Tracts” and “Patchs”. Each tract is approximately $1.5\times1.5$ and divided into $9\times9$ patches.

### 2.3. CLAUDS

CLAUDS is a deep $U$-band imaging survey of the HSC deep layer performed by the 3.6 m CFHT with MegaCam (Sawicki et al. 2019). MegaCam (Boulade et al. 2003) is a wide-field
camera with 40,2048 × 4612 pixel back-illuminated CCDs, and covers an area of 1.02 deg².

The CLAUDS survey was performed using two u-band filters; u' and U. The two u-band filters are significantly different from each other. The new U-band filter has better transmission than the u''-band filter and has no red leak at 5000 Å, which is observed in the u''-band filter. For the XMM-LSS deep field, the CLAUDS survey was performed entirely using the u''-band filter. The depth of the CLAUDS in the XMM-LSS reaches 26.6 mag (5σ detection in 2″ diameter aperture), while in the ultradeep region, which corresponds to the Subaru XMM-Newton Deep Survey (SXDS; Furusawa et al. 2008) region, reaches 1 mag deeper. The survey properties are summarized in Table 1.

Basic calibration and data reduction of CLAUDS MegaCam data were performed with the Elixir software (Magnier & Cuillandre 2004) at CFHT. Elixir performs the basic data reduction before sending the data to the Canadian Astronomy Data Center to be processed with MegaPipe (Gwyn 2008). MegaPipe performs astrometric calibration against Gaia astrometry while photometry was calibrated against a Sloan Digital Sky Survey (SDSS) u-band photometry, cross checked with synthetic u-band photometry produced using a combination of Pan-STARRS (Magnier et al. 2020) g-band and GALEX NUV photometry.

2.4. VIDEO

The VISTA Deep Extragalactic Observations survey (VIDEO; Jarvis et al. 2013) is a deep near-infrared (NIR) imaging survey performed by the 4.1 m Visible and Infrared Survey Telescope for Astronomy (VISTA) at Cerro Paranal with the VISTA InfraRed CAmera (VIRCAM; Dalton et al. 2006). VIRCAM consists of 16 2K × 2K Raytheon VIRGO HgCdTe detectors.

VIDEO data release 5 mosaic images of the XMM-SERVS field are provided in the ESO Phase 3 data archive.9 In the XMM-SERVS field, the mosaic images in each band are separated into three smaller areas designated as XM1, XMM2, and XMM3. Among them, XMM1 covers the SXDS. The data were reduced at the Cambridge astronomical survey unit (CASU) using the VISTA Data Flow System (VDFS; Irwin et al. 2004). The astrometry and photometry of the survey were calibrated against the Two Micron All Sky Survey (2MASS) point-source catalog (Skrutskie et al. 2006). The final 5 σ depths in a 2″ diameter aperture are 24.51, 24.44, 24.12, and 23.77 mag in the Y, J, H, and Ks bands, respectively. The survey properties are summarized in Table 1.

2.5. SERVS

SERVS (Mauduit et al. 2012) is a deep mid-infrared imaging survey performed by the Spitzer space telescope using the Infrared Array Camera (IRAC; Fazio et al. 2004) during the post-cryogenic mission. Only the IRAC channel 1 (3.6 μm) and channel 2 (4.5 μm) are usable due to the high background in the other bands due to the outages of the cryogenic coolant.

SERVS covers five deep multiwavelength extragalactic fields (ELAIS-N1, ELAIS-S1, Lockman Hole, Chandra Deep Field South, and XMM-LSS), in total 18 deg². The mean integration time per pixel is approximately 1200 s, which is close to the confusion limit of the Spitzer IRAC data. The 5σ depths in 3.6 and 4.5 μm bands are 1.9 and 2.2 μJy in 3″ diameter aperture. They correspond to 5σ magnitudes of 23.20 and 23.04, respectively. The survey properties are summarized in Table 1.

The mosaic and uncertainty images were retrieved from the NASA/IPAC Infrared Science Archive.10 The data were processed at the Spitzer Science Center. The data reduction pipeline performs the standard image reduction and additional detector-specific processing. The images were co-added and reprojected using MOPEX to a pixel scale of 0″6 pixel⁻¹. Original photometric calibration of the IRAC data was performed using dedicated calibration observations. Crosschecks against the SWIRE survey (Lonsdale et al. 2003) suggest that a correction factor of 1.02 is needed for the 3.6 μm band but not for the 4.5 μm band. We apply the correction during catalog construction.

2.6. Spectroscopic Redshifts

Similar to other multiwavelength survey fields, there are a large number of spectroscopic redshift measurements in the XMM-LSS region. Spectroscopic redshift measurements within the survey area were compiled from various spectroscopic surveys, including those from the SDSS data releases 9 and 16 (Ahn et al. 2012), VIMOS Public Extragalactic Redshift Survey (VIPERS; Scoceddio et al. 2018), Galaxy and Mass Assembly (Liske et al. 2015), VIMOS VLT Deep Survey (VVDS; Le Fèvre et al. 2013), VANDELS (Garilli et al. 2021), Milliquas (Flesch 2019), MOSFIRE Deep Evolution Field Survey (Kriek et al. 2015), Ultradeep Survey11 (UDS; McLure et al. 2013; Bradshaw et al. 2013), 3D-Hubble Space Telescope (3D-HST; Brammer et al. 2012; Momcheva et al. 2016), and the SXDS multiwavelength catalog (Akiyama et al. 2015), as well as from individual studies in the SXDS and XMM-LSS regions (Yamada et al. 2005; Geach et al. 2007; Ouchi et al. 2008; Saito et al. 2008; Willmer et al. 2008; van Breukelen et al. 2009; Ono et al. 2010; Simpson et al. 2012; Diaz Tello et al. 2013; Melnyk et al. 2013; Yabe et al. 2014; Wang et al. 2016; Menzel et al. 2016; Ono et al. 2018). In total, 294,536 secure spectroscopic redshift records associated with 238,403 unique galaxies and AGN were compiled. The majority of the spectroscopic redshift records are from the objects in the SDSS and VIPERS catalogs, which have an i-band magnitude up to 22.5. However, the faintest magnitude of deep spectroscopic surveys such as VVDS-UDEEP, 3D-HST, and spectroscopic follow-up of X-ray sources in the SXDS from Akiyama et al. (2015) reaches an i-band magnitude of 24.75.

3. Multi-band Photometry of the Optical Counterpart

3.1. Process of Multi-band Photometry

In order to obtain the multiwavelength properties of the optical counterparts of the X-ray sources, this work uses deep imaging data in the 12 photometric filters from the data sets described above. Proper treatment of the point-spread function (PSF) shape and size differences between data sets is needed in order to obtain accurate colors for photometric redshift estimation and SED fitting. This is especially important for the

9 http://eso.org/mg/publicAccess#dataReleases

10 SERVS Team (2020).

11 https://www.nottingham.ac.uk/astronomy/UDS/data/data.html
SERVS mid-infrared data set, which suffers from severe blending due to the larger PSF size than the other data sets.

Prior-based PSF-convolved photometry was performed using T-PHOT (Merlin et al. 2015, 2016). T-PHOT uses morphological information of an object in a high-resolution image to measure its flux in images with low spatial resolution. The low-resolution image needs to have the same World Coordinate System (WCS) and the same or integers-times pixel scale as the high-resolution prior image. In this analysis, all data were resampled to the WCS defined in the HSC S19A internal data set. The process of image alignment, background subtraction, and preparation of variance images are described in the following subsections.

3.1.1. Image Alignment and Background Subtraction

First, global background subtraction was applied to the calibrated HSC image in each patch. The background levels were determined from the mean pixel value of the image after applying a $3\sigma$ clip. After the background subtraction, an additional 200 blank pixels were padded to each side of the image.

CLAUDS data used in this work were aligned to the tract-patch definition of the HSC S16A data set, a prior version to the HSC data currently used in this analysis. Therefore, there is a 1–2 pixel offset from the HSC images used in the current analysis, we match the astrometry to the HSC S19A data set and apply an additional local background subtraction using SWarp (Bertin 2010).

For the pipeline-reduced VIDEO DR5 images of the XMM1, 2, and 3, the background of each area was subtracted using SExtractor (Bertin & Arnouts 1996). The images were then resampled into the same pixel scale as in the HSC images and combined together into a single mosaic using SWarp. Resampled variance images on the same pixel scale were also produced using SWarp. These variance images are different from the weight images produced automatically by SWarp, which does not preserve the original variance. The variance images were created by converting the original weight images to variance and then passing them to SWarp as images to be combined. This method better preserves the original variance of the image than the weight images automatically generated by SWarp. Cutouts in the same tract patch as HSC images were created from the mosaics and resampled to the same WCS as in the HSC S19A images. SERVS mid-infrared images are provided as a single mosaic with the rms image. The mosaic image was resampled to the same WCS as in the HSC S19A images and the local background was subtracted using SWarp. The rms image was converted to a variance image and resampled in the same manner applied to the VIDEO images. Cutouts in the same tract patch and WCS as in the HSC S19A images were then produced from the mosaic images using SWarp.

3.1.2. PSF Modeling and Convolution Kernel Construction

The convolution kernel is a two-dimensional matrix that converts the PSF shape of the high-resolution image (HRI) to the PSF shape of the low-resolution image (LRI). It is constructed by deconvolving the LRI PSF with the HRI PSF. The PSFs were constructed locally in each patch in order to take into account the PSF variation over the survey area. However, PSF variation within each patch is ignored. The PSF of the HSC data sets were queried from the HSC database while the PSF models for CLAUDS, VIDEO, and SERVS were constructed directly from the images of stellar objects.

First, the catalogs in each patch produced using SExtractor are matched with $K_s$-band sources in the 2MASS point-source catalog (Skrutskie et al. 2006). For CLAUDS and VIDEO, 2MASS sources with $15 < K_s < 17$ were selected for the PSF modeling. Extended sources with CLASS_STAR $< 0.9$ and ELLIPTICITY $> 0.2$ in the CLAUDS and VIDEO catalogs were removed. These extended sources were identified as point sources in the 2MASS point-source catalog due to the low spatial resolution of 2MASS.

For SERVS, sources with $14 < K_s < 17$ were selected for the PSF modeling. Since the PSFs of Spitzer IRAC is asymmetric, the same criteria as the CLAUDS and VIDEO data sets cannot be used to remove extended sources. Extended sources were rejected by examining the FWHM distribution; sources whose FWHM exceeds $2\sigma$ scatter were removed.

After the removal of the extended sources, the individual images of stellar sources were recentered, background subtracted, and normalized to unity. The final PSF images were constructed using a median combination of each individual stellar source. Finally, the final PSF image is recentered and normalized once again, and the images are padded to $55 \times 55$ pixels. Local PSF images were constructed only for patches with more than seven individual stellar sources. A median PSF image of each tract is used for patches with fewer than seven individual sources. The HSC S19A $\mathcal{R}$-band PSF images were adopted as the HRI PSF and Pypher (Boucaud et al. 2016) was used to construct the convolution kernels of images in other bands.

3.1.3. Pixel–pixel Correlation Correction

It is well known that pixel–pixel correlation occurs during image resampling with SWarp. Resampling smooths the image; thus, the variance in the image becomes artificially smaller than the original image. Therefore, it is necessary to correct the underestimation of the variance.

Fixed-aperture analysis similar to Bielby et al. (2012) was performed on the background variance images to estimate the effects of pixel–pixel correlation. SExtractor was used to produce segmentation images of all CLAUDS and VIDEO images and 1000 fixed apertures of $2''$ radius were placed in regions with no source detections. The sky background was measured from the science images in the 1000 apertures while the photometric uncertainties were estimated from the weight images in the same aperture using PhotUtils (Bradley et al. 2020). A $4\sigma$ clip was used to remove outlier apertures that may have fallen on the image boundary or bad detector regions. The sky background variance ($\sigma_{\text{sky}}^2$) was estimated by fitting a Gaussian distribution to the distribution of the sky background values and compared to the variance with the median of the variance image in the same aperture ($\sigma_{\text{med}}^2$). If there is no pixel–pixel correlation, the two values are consistent with each other.

The HSC pipeline has already performed the correction for pixel–pixel correlation and no resampling was done for the reduced image. For the SERVS data set, thanks to the resampling process described in Section 3.1.1, the SERVS variance data preserved correct variance and is consistent with the variance in the sky background. On the other hand, The variances measured in CLAUDS and VIDEO were significantly smaller

---

12 Skrutskie et al. (2003).
than the variance measured in the sky background. The correction factor \( k = \frac{\sigma_{\text{sky}}^2}{\sigma_{\text{wei}}^2} \) was estimated in each patch of CLAUDS and VIDEO data.

The correction factors for the VIDEO images are constant over the survey area of the VIDEO survey. A single correction factor determined from the median of the correction factors in all of the VIDEO images was adopted for simplicity and is summarized in Table 2. On the other hand, the correction factor of CLAUDS shows significant variation due to the variation in the survey depth. The correction factor of each patch was applied individually and the median of all correction factors in each tract was used when the patch-level correction factor could not be determined. Figure 2 shows the distribution of the adopted correction factor over the CLAUDS survey area.

### 3.2. Source Detection

The primary source catalogs and segmentation images were constructed from HSC S19A r-band images using SExtractor. As summarized in Table 1, the HSC S19A r-band image has the highest resolution among the imaging data sets and is deep enough to detect a large fraction of objects in the other bands. Therefore, the r-band image is used as the high-resolution prior.

The T-PHOT fitting process fails in some regions where bright stars are present in the image. This is likely due to the fact that the SExtractor deblending algorithm divides the star and halo into many individual sources. To solve this problem, sources within the HSC S19A r-band bright star masks (Coupon et al. 2018) were removed. T-PHOT was run twice in each patch to take into account small sub-pixel astrometric offsets. The astrometric offsets were determined in the first run and applied automatically during the second run. The results include catalogs containing the fitting results, model images, residual images, and residual statistics.

Figure 3 shows optical-IR images, model images, and residual images of an optical-infrared counterpart of an X-ray source. Several sources can be seen from the u' to Ks bands but are blended together in the 3.6 and 4.5 \( \mu m \) images. The residual images in the 3.6 and 4.5 \( \mu m \) bands show systematic residuals. This is possibly due to the PSF asymmetry and its variation over the field of view of the IRAC data sets.

### 3.3. Creation of the Multi-band Photometric Catalog

The T-PHOT catalogs were combined together and duplicate objects in overlapping patch regions were removed from the catalog. The photometric magnitudes and uncertainties in each band were calculated using the zero-points and zero-point uncertainties shown in Table 3. If the signal-to-noise ratio of the measurement is below 2\( \sigma \), then 2\( \sigma \) upper limits were adopted instead.

Objects that fall into the HSC S20A bright star mask, bad detector regions affected by stray light, or detector defects, sources on the edge of the survey area, and sources with failed fitting results are flagged in the catalog. In addition, sources that are likely local galaxies and were broken up by the detection algorithm were also identified using region files created from the HyperLeda catalog (Makarov et al. 2014). Lastly, sources containing saturated pixels in the HSC images were flagged using the HSC mask images. The Galactic reddening value for each object was retrieved from the IRAS reddening map\(^3\) in Schlegel et al. (1998). The Galactic dust attenuation in each band was calculated using the Galactic dust extinction law (Fitzpatrick 1999).

### 3.4. Survey Area

Because the multiwavelength photometry sample does not cover the entire sample of the X-ray point sources in Chen et al. (2018), we redefined the survey area of the X-ray sample based on the availability of the multiwavelength photometry considering the following conditions:

1. The optical-IR counterpart is in the HSC, VIDEO \( H \) band, and SERVS-IRAC1 coverage.
2. The optical-IR counterpart is not in any bright star masks of HSC nor in the bad regions of the VIDEO \( H \) band.

The first condition was imposed to maximize the coverage of multiwavelength photometry, while the second condition was imposed to remove regions where the optical-IR images were affected by image artifacts such as stray light, bright star halos, and edges of the images. Figure 4 shows the distribution of the X-ray sources that meet the above criteria shown as orange symbols. Out of the 5237 XMM-SERVS X-ray sources with an optical-IR counterpart, 3542 X-ray sources are selected as the primary sample for the statistical discussion.

The total area within the redefined area was estimated using a Monte Carlo simulation by randomly distributing 100,000 mock data points within the original survey area presented in Chen et al. (2018) and calculating the fraction of data points that satisfy the criteria. The estimated survey area for statistical analysis is 3.52 deg\(^2\). The area curve in the 0.5–2, 2–10, and 0.5–10 keV bands was calculated by normalizing the maximum

---

**Table 2**

| Band | Correction \( k = \frac{\sigma_{\text{sky}}^2}{\sigma_{\text{wei}}^2} \) |
|------|------------------|
| VIDEO Y | 16.78 |
| VIDEO J | 11.88 |
| VIDEO H | 11.56 |
| VIDEO Ks | 8.32 |

---

**Figure 2.** Distribution of the pixel–pixel correlation correction factor for the variance in the CLAUDS data at each patch position.

---

3 https://irsa.ipac.caltech.edu/applications/DUST/
area of the area curve presented in Chen et al. (2018) to be 3.52 deg$^2$.

In order to check the updated area curve, the logN-logS of the primary sample based on the redefined survey area is compared to that of the entire XMM-SERVS sample. Figure 5 shows the logN – logS relation in the 0.5–2, 2–10, and 0.5–10 keV bands for the primary sample (orange) and the original XMM-SERVS (blue). The logN – logS based on the redefined survey area is consistent with the logN – logS of the original survey area within the Poisson uncertainty. We conclude that the normalized survey area reproduces the survey area of the primary sample well.

### 3.5. Cross Matching with the X-Ray Catalog

In order to select the optical counterpart of each X-ray source, first, the identification in Chen et al. (2018) was adopted. The original optical counterparts of 3180 X-ray sources have a corresponding object in the PSF-convolved photometric catalog within a 2" radius. In Chen et al. (2018), the majority of X-ray sources (2762 sources) were matched with their counterparts in the SERVS catalog. The SERVS-matched counterparts sometimes contain multiple counterparts in the photometric catalog due to the large PSF size of IRAC.

| Table 3 | Photometric Zero-point |
|---------|------------------------|
| Band    | Zero-point  | Uncertainty |
| CLAUDS  | 3.0         | 0.035       |
| HSC g    | 27.0        | 0.010       |
| HSC r    | 27.0        | 0.010       |
| HSC i    | 27.0        | 0.010       |
| HSC z    | 27.0        | 0.011       |
| HSC y    | 27.0        | 0.013       |
| VIDEO    | 30.0        | 0.020       |
| SERVS 3.6| 23.9        | 0.030       |

Note. Measurements were converted from megajansky per steradian to microjansky per pixel.

area of the area curve presented in Chen et al. (2018) to be 3.52 deg$^2$.

In order to check the updated area curve, the logN-logS of the primary sample based on the redefined survey area is compared to that of the entire XMM-SERVS sample. Figure 5 shows the logN – logS relation in the 0.5–2, 2–10, and 0.5–10 keV bands for the primary sample (orange) and the original XMM-SERVS (blue). The logN – logS based on the redefined survey area is consistent with the logN – logS of the original survey area within the Poisson uncertainty. We conclude that the normalized survey area reproduces the survey area of the primary sample well.
For each SERVS-matched source, we examined the number of \( r \)-band detected neighbors within a 2′″ radius of the counterpart, and 342 sources have multiple optical counterparts suggesting they are blended. The brightest source in the SERVS 3.6 \( \mu \)m band was chosen as the counterpart of the X-ray source among the blended sources. This modification changed the optical counterpart of 23 of the blended sources.

The remaining 362 sources with no corresponding counterpart within the PSF-convolved catalog show no or faint objects in the HSC S19A \( r \) band but show a significant detection in the near-IR (NIR) bands. In order to recover the optically faint counterparts, we produced PSF-matched cutouts by matching the PSF to that of the VIDEO \( H \) band using the PSF described in Section 3.1.2. Aperture photometry using a 2″ diameter aperture was performed using SExtractor in dual-image mode on the optical and NIR images with the VIDEO \( H \)-band image as the detection image. The 2″ diameter aperture contains \( \sim 80\% \) of the PSF flux. Aperture correction factors were calculated from the VIDEO \( H \)-band growth curves. The factor was calculated for each patch and applied to the patch individually in order to account for the PSF variation. For the mid-infrared data sets, prior-based PSF-convolved photometry was
performed using the VIDEO H-band images as the high-resolution images. Photometry for 282 of the 362 HSC S19A r-band non-detected sources was successfully obtained.

The images of the remaining 80 sources show neither the HSC S19A r-band nor the VIDEO H-band source but some show faint HSC S19A r-band objects close to the detection limit. We do not attempt to recover these sources because the optical identification with such faint sources is uncertain. The photometry of the 282 VIDEO H-band detected sources and 3180 r-band detected sources were combined together into a single catalog. In summary, multiwavelength photometry is obtained for 97.7%, 3462 out of the 3542 primary X-ray sources. A description of the catalog of the primary X-ray AGN sample and multiwavelength photometry in the HSC-Deep XMM-LSS region is provided in the Appendix.

4. Analysis

4.1. Photometric Redshift

Out of the 3462 primary X-ray sources, 1321 sources have prior spectroscopic redshift measurements. For the remaining X-ray sources, we calculated the photometric redshift using the photometric redshift code LePhare (Arnouts et al. 1999; Ilbert et al. 2006). LePhare estimates the photometric redshift by minimizing the $\chi^2$ between the observed and model photometry that was derived from template SEDs. For X-ray detected sources, empirical templates of galaxies, local AGN, as well as composite AGN templates constructed from AGN and galaxy templates in Salvato et al. (2011) were used. For X-ray non-detected sources, galaxy models from Ilbert et al. (2009) were used. The model magnitudes were calculated between redshift 0 and 6 with steps of 0.05. We use both SMC and Calzetti extinction laws (Prevot et al. 1984; Calzetti et al. 1994) and fit the reddening as a free parameter with $E(B-V) = 0, 0.025, 0.005, 0.075, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.125, 0.15, 0.175, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, and 0.6 mag. No additional emission lines were considered for the AGN models but are added to the galaxy models. LePhare has the capability of estimating systematic zero-point shifts and applying them to the photometry to reduce the deviation from spectroscopic redshifts. The systematic shifts are derived using the spectroscopic redshift sample. The zero-point correction determined from X-ray non-detected sources was adopted to the X-ray detected sources and is shown in Table 4. In both cases, an additional uncertainty (ERR_SCALE) of 0.05 mag was also adopted in order to take into account any additional systematic uncertainty associated with the photometry.

Photometric redshift performance was evaluated based on the median absolute deviation ($\sigma_{NMAD}$; Hoaglin et al. 1983) defined as

$$\sigma_{NMAD} = 1.48 \times \text{median} \left( \frac{|z_{\text{phz}} - z_{\text{spec}}|}{1 + z_{\text{spec}}} \right),$$

where $z_{\text{phz}}$ and $z_{\text{spec}}$ are the photometric redshift and spectroscopic redshift, respectively. The outlier fraction ($f_{\text{out}}$) is the fraction of objects whose normalized absolute deviation $(z_{\text{phz}} - z_{\text{spec}})/(1 + z_{\text{spec}})$ is larger than 0.15 in the spectroscopic redshift sample. A threshold of 0.15 was adopted following studies in the COSMOS field (Ilbert et al. 2009; Laigle et al. 2016). The photometric redshift performance was evaluated using sources that do not lie within the stellar locus in the $g - z$ against $z - 3.6 \, \mu m$ plane

$$(g - z) - 0.5937(z - [3.6]) < -1.7,$$

where $g, z,$ and $[3.6]$ are the $g, z,$ and $3.6 \, \mu m$ band magnitudes. Galaxy templates were used to evaluate the photometric redshift performance of the X-ray non-detected sources. Similarly, AGN or galaxy templates were used to evaluate the photometric redshift performance of the X-ray detected source. Figure 6 shows the comparison between photo-$z$ and spec-$z$ of X-ray non-detected (left) and X-ray detected sources (right).

A photometric redshift scatter ($\sigma_{NMAD}$) of 0.07 and an outlier fraction ($f_{\text{out}}$) of 26% was achieved for X-ray detected sources with four catastrophic failures where the photometric redshift could not be determined. For comparison, a photometric redshift scatter $\sigma_{NMAD}$ of 0.031 and outlier fraction $f_{\text{out}}$ of 3% was achieved for the X-ray non-detected galaxies. The $\sigma_{NMAD}$ and $f_{\text{out}}$ for X-ray detected sources are worse than X-ray non-detected sources. Photometric redshifts for AGN can be difficult to accurately determine due to the flat featureless UV-optical continuum of unobscured AGN.

Figure 7 shows the distribution of the spectroscopic and photometric redshift of the primary X-ray sources as a stacked histogram. Above redshift 2, most of the redshift estimates are from the photometric redshift. In summary, 3458 AGN of the primary sample have a spectroscopic redshift or 12-band photometric redshift; thus, we achieved a total of 99.8% redshift completeness thanks to the deep multiwavelength data set.

4.2. Hydrogen Column Density Estimation

The hydrogen column density ($\log N_H$) associated with the nuclear X-ray emission of the primary X-ray source was calculated using the X-ray hardness ratio (HR) and best-redshift estimation assuming an intrinsic AGN X-ray spectrum. The HR is defined as $(H - S)/(H + S)$ where $S$ is the 0.5–2 keV band count rate and $H$ is the 2–10 keV band count rate. In the XMM-SERVS catalog, the flux of each source is derived by combining measurements from multiple detectors with weighting. The PN-equivalent count rate was calculated by dividing the flux by the energy conversion factor (ECF) of the PN used in Chen et al. (2018). As a check, the PN-equivalent count rate was compared with the reported count rate of sources detected with the PN detector in Chen et al. (2018). The two count rates are consistent with each other for the 2–10 keV band but show a systematic offset in the 0.5–2 keV band. A correction factor of 1.12 was applied to the 0.5–2 keV band to make the PN-equivalent count rate consistent with that measured with the PN detector.

Table 4

| Band | LePhare Systematic Zero-point Shifts |
|------|------------------------------------|
| CLAUDS $u'$ | 0.156 | VIDEO Y | 0.002 |
| HSC $g$ | -0.027 | VIDEO J | 0.049 |
| HSC $r$ | -0.037 | VIDEO H | 0.070 |
| HSC $i$ | -0.0254 | VIDEO Ks | -0.021 |
| HSC $z$ | -0.022 | SERVS 3.6 $\mu m$ | -0.005 |
| HSC $y$ | -0.044 | SERVS 4.5 $\mu m$ | 0.003 |
A phenomenological AGN model constructed from a linear combination of a cutoff power-law, pexrav reflection component \cite{MagdziarzZdziarski1995}, and scattered AGN continuum was assumed as the AGN X-ray spectra.

\[
\text{Model} = \text{tbabs} \cdot (\text{zphabs} \cdot \text{cabs} \cdot \text{zcutoffpl}) + \text{pexrav} + \text{constant} \cdot \text{zcutoffpl},
\]

where tbabs is the galactic X-ray extinction and zphabs is the absorption associated with the AGN, cabs is the additional Compton scattering, and constant is the fraction of the scattered AGN continuum. The parameters of the phenomenological model were set according to the best-fitted values from Ricci et al. \cite{Ricci2017} based on AGN in the local universe assuming a photon index of $\Gamma = 1.8$. For the reflection component, the reflection strength was set to 1.0 and the inclination angle was set to 30°. The fraction of the scattered AGN continuum was set to 1% of the transmitted AGN continuum. The exponential cutoff was set to 381 keV for all components. The elemental abundances were set to the solar abundance in Anders & Grevesse \cite{AndersGrevesse1989}.

The HR model grid was calculated in the redshift range between 0.001 to 6 and $N_{\text{H}}$ grid between 20 and 26 using XSPEC \cite{Arnaud1996} assuming an on-axis response matrix file (RMF) and ancillary matrix files (ARF) of the XMM-Newton PN detector.\footnote{https://www.cosmos.esa.int/web/xmm-newton/epic-response-files} Figure 8 shows a comparison between the HR of the primary sample and the HR of the model grid. Spectroscopically identified type 1 broad-line AGN (BL-AGN) are marked with open circles, and their HR shows a mild increasing trend with increasing redshift. Since most of the BL-AGN should have low column density ($N_{\text{H}} < 22$), the increasing X-ray hardness is unlikely driven by increasing obscuration. The increasing X-ray hardness can be explained by the Compton reflection component at rest frame 10–30 keV shifting into the observed hard 2–10 keV band at high redshift. It should be noted that broad-line classification is only available for a limited sample.

The column density was estimated only for AGN above redshift 2 and only between $N_{\text{H}}$ and $\log N_{\text{H}}$ (cm$^{-2}$) grid between 20 and 24 where the HR can distinguish the obscuration of CTN-AGN. The redshift dependence of the HR indicates that obscured AGNs with column density less than $N_{\text{H}} < 22$ can hardly be distinguished. The column density can only be constrained for sources detected in both the 0.5–2 and 2–10 keV bands. Only upper or lower limits on the column density can be derived for X-ray sources detected only in a single band. It should be reminded that the lower limit of the column density for sources that are detected only in the 2–10 keV band is mostly larger than $N_{\text{H}} > 23$ above $z > 2$. Therefore, sources detected only in the 2–10 keV band were assumed to...
have $\log N_H$ (cm$^{-2}$) > 23 in the following discussion. Furthermore, some AGN show a smaller HR than that of an AGN model with $\Gamma = 1.8$. These sources likely have a softer AGN X-ray spectrum with a larger photon index ($\Gamma > 1.8$) and no obscuration. We assign a column density of $\log N_H$ (cm$^{-2}$) = 20 to them.

### 4.3. High-redshift AGN Sample Selection

In order to examine the obscured fraction of luminous AGNs above redshift 2, 673 AGN between redshifts 2 and 5 were selected based on the best available redshift (either spec-$z$ or photo-$z$) as the high-redshift AGN sample. Of the 672 AGN, 251 AGN were detected in both the 0.5–2 and 2–10 keV bands, while 275 and 53 were detected only in the 0.5–2 and 2–10 keV bands, respectively. Finally, 93 AGN were detected only in the 0.5–10 keV band. Within the 672 AGN, 203 (30%) have spectroscopic redshift and more than half of the spectroscopic sample is BL-AGN (71%).

The intrinsic X-ray 2–10 keV luminosity was derived using

$$L_X = \text{ext}(z, N_H)k(z)4\pi D_L^2(z)f_X$$

by assuming the best available redshift and $\log N_H$, where $D_L$ is the luminosity distance, $f_X$ is the observed X-ray flux in the 2–10 keV band, $k(z)$ is the k-correction term, and $\text{ext}(z, N_H)$ is the extinction correction in the observed frame. The observed fluxes were calculated from the PN-equivalent count rate in the 2–10 keV band assuming an ECF of 1.26 $\times 10^{11}$ counts s$^{-1}$/erg s$^{-1}$ cm$^{-2}$.$^{15}$ The ECF and extinction correction was calculated using XSPEC by assuming the phenomenological AGN model presented in Section 4.2. The ECF was from the model without absorption while the extinction correction was the ratio between the model without absorption to that with $\log N_H = 20–24$.

Figures 9 and 10 show the 2–10 keV luminosity and $\log N_H$ of the high-redshift AGN. BL-AGN are marked with a black open circle. Most of the AGN detected in both the 0.5–2 and 2–10 keV bands are those with $\log N_H$ (cm$^{-2}$) < 23.5 and possess quasar-level luminosity ($\log L_X$ (erg s$^{-1}$) > 44.5). For AGN detected only in the 0.5–2 keV band, only upper limits can be placed on $\log N_H$ and $\log L_X$ and the upper limits show a large scatter. On the other hand, the lower limits of $\log N_H$ and $\log L_X$ are derived for AGN detected only in the 2–10 keV band. Most of them are located above $\log N_H > 23$, which implies that they are heavily obscured AGN.

### 4.4. LDDE Model and Absorption Function

In order to model the intrinsic number of AGN at $z > 2$, the functional form of the luminosity and absorption functions were assumed following Ueda et al. (2014) because the size and luminosity coverage of the current sample are not large enough to determine the overall shape of the luminosity function. The hard X-ray AGN luminosity function describes the number density of CTN-AGN with $\log N_H$ (cm$^{-2}$) = 20–24. It is expressed as an evolving double power law following the luminosity-dependent density evolution (LDDE) model. The hard X-ray luminosity function of CTN-AGN in the local universe is described as follows:

$$d\Phi_{CTN}(L_X, z = 0) = \frac{d\Phi_{CTN}(L_X, z = 0)}{d \log L_X} = A \left[ \left( \frac{L_X}{L_\star} \right)^\gamma_1 + \left( \frac{L_X}{L_\star} \right)^\gamma_2 \right]^{-1},$$

where $A$ is the normalization of the luminosity function, $L_\star$ is the break luminosity, and $\gamma_1$ and $\gamma_2$ are the slopes of the connected power laws. The luminosity function outside the local universe follows a luminosity and redshift-dependent evolution:

$$d\Phi_{CTN}(L_X, z) = \frac{d\Phi_{CTN}(L_X, z = 0)}{d \log L_X} e(L_X, z),$$

where $e(L_X, z)$ is the evolutionary term, which describes the redshift dependence of the luminosity function following Equation (16) in Ueda et al. (2014).

Ueda et al. (2003) introduced the absorption function $f_{abs}$, which is the probability distribution function defined between $\log N_H$ (cm$^{-2}$) = 20 and 26. The function is normalized to 1 between $\log N_H$ (cm$^{-2}$) = 20 and 24 as follows:

$$\int_{20}^{24} f_{abs}(L_X, z; N_H) d \log N_H = 1.$$
log $N_{\rm H}$ (cm$^{-2}$) = 20 and 22 together. The definition is as follows:

$$f_{\rm abs}(L_X, z; N_{\rm H}) = \begin{cases} \frac{1 - \psi(L_X, z)}{2} & [20 \leq \log N_{\rm H} < 22] \\ \frac{1}{1 + \epsilon} \psi(L_X, z) & [22 \leq \log N_{\rm H} < 23] \\ \frac{1}{1 + \epsilon} \psi(L_X, z) & [23 \leq \log N_{\rm H} < 24] \\ \frac{1 + \epsilon}{\epsilon} \psi(L_X, z) & [24 \leq \log N_{\rm H} < 26] \\ \end{cases}$$

where $\psi(L_X, z)$ is the fraction of CTN-AGN with log $N_{\rm H}$ (cm$^{-2}$) $\geq$ 22 and $\epsilon$ is the ratio of AGN with log $N_{\rm H}$ (cm$^{-2}$) = 23–24 to those with log $N_{\rm H}$ (cm$^{-2}$) = 22–23, and $f_{\rm CTK}$ is the relative number of Compton-thick AGN (CTK-AGN) relative to obscured CTN-AGN, which is fixed to 1. The constraints on $f_{\rm CTK}$ are discussed in Section 6.3.

The redshift and luminosity dependence of the absorption function can be described with the function $\psi(L_X, z)$. The obscured fraction in the local universe shows a linearly decreasing dependence on the X-ray luminosity:

$$\psi(L_X, z) = \min[\psi_{\text{max}}, \max(\psi_{43.75}(z) - \beta(\log L_X - 43.75), \psi_{\text{min}})]$$

where $\psi_{\text{min}}$ and $\psi_{\text{max}}$ are the minimum and maximum obscured fractions of CTN-AGN, which are determined to be 0.2 and 0.84, respectively (Ueda et al. 2014). The maximum obscured fraction in Ueda et al. (2014) was set to 0.84 since larger values of $\psi$ will return a negative probability for the absorption function in the log $N_{\rm H}$ (cm$^{-2}$) = 20–21 bin. Our modification of the absorption function in log $N_{\rm H}$ (cm$^{-2}$) = 20 – 22 removes this effect; hence, larger values of $\psi_{\text{max}}$ are allowed as discussed in the later sections. $\beta$ is the slope of the decrease.

Figure 9. (Left) The intrinsic 2–10 keV luminosity of high-redshift AGN detected in both the 0.5–2 and 2–10 keV bands as a function of redshift (purple). BL-AGN are shown with black open circles. The typical uncertainty for the spec-z and photo-z samples is shown with the error bar with short and long caps, respectively. (Right) Same as the left panel but upper limits and lower limits for AGN detected only the 0.5–2 keV band (blue downward triangles) or the 2–10 keV band (orange upward triangles).

Figure 10. (Left) log $N_{\rm H}$ of high-redshift AGN detected in both the 0.5–2 and 2–10 keV bands as a function of redshift (purple). BL-AGN are shown with black open circles. The bottom panel shows the redshift distribution of AGN with HR softer than the minimum HR calculated by XSPEC set to have log $N_{\rm H}$ (cm$^{-2}$) = 20, and the data points in this panel are randomly shifted in the $y$-axis for clarity. The typical uncertainty is shown with the error bar at the bottom left corner. (Right) Same as the left panel but upper limits and lower limits for AGN detected only in the 0.5–2 keV band (blue downward triangles) or the 2–10 keV band (orange upward triangles).
which is set to 0.24 (Ueda et al. 2014), and \( \psi_{43.75}(z) \) is the obscured fraction at log \( L_X \) (erg s\(^{-1}\)) = 43.75.

The redshift evolution of the obscured fraction \( \psi(L_X, z) \) is described by \( \psi_{43.75}(z) \)

\[
\psi_{43.75}(z) = \begin{cases} 
\psi^0_{43.75}(1 + z)^{a_1} & z < 2.0 \\
\psi^0_{43.75}(1 + 2)^{a_1} = \psi^2_{43.75} & z \geq 2.0
\end{cases}
\]

where \( \psi^0_{43.75} \) is the obscured fraction at log \( L_X \) (erg s\(^{-1}\)) = 43.75 in the local universe. It was determined to be 0.43 ± 0.03 (Ueda et al. 2014). The redshift dependence parameter \( a_1 \) is determined to be 0.48 ± 0.05. We assumed a constant \( \psi_{43.75}(z = 2) = \psi^2_{43.75} \) in the redshift range above redshift 2.

It should be pointed out that \( \psi_{43.75}(z) \) is a parameter used to control the evolution of the obscured fraction \( \psi(L_X, z) \) and is not limited to between \( \psi_{\text{min}} \) and \( \psi_{\text{max}} \). As defined in Equation (2), \( \psi_{43.75}(z) \) is equal to \( \psi(43.75, z) \) only when \( \psi_{43.75}(z) \leq \psi_{\text{max}} \), and beyond that would suggest that \( \psi(43.75, z) \) has saturated at \( \psi_{\text{max}} \) with no further evolution with redshift but the obscured fraction of higher luminosity AGN can still be lower. Therefore, the obscured fraction was estimated with \( \psi(L_X, z) \) (see Section 5.1).

### 4.5. Survey Area Function

In order to estimate the obscured fraction at high redshifts, it is important to evaluate the dependence of the effective survey volume as a function of luminosity, redshift, and amount of obscuration. The survey area that is sensitive enough to detect obscured AGN can be smaller than that for unobscured AGN at the same intrinsic luminosity. The survey area function \( \Omega(L_X, z, \log N_{HI}) \) was estimated as a function of AGN 2–10 keV luminosity, redshift, and \( \log N_{HI} \) using XSPEC (Arnaud 1996).

The same phenomenological AGN model, RMF, and ARF calibration files of the PN detector used to estimate the \( N_{HI} \), and the survey area curve defined in Section 3.4 are considered.

Figure 11 shows the survey area at redshift 2.5 as a function of log \( L_X \) and \( \log N_{HI} \) in the 2–10 keV band. At a fixed X-ray luminosity, the accessible survey area decreases with increasing column density due to increasing X-ray extinction. Above log \( N_{HI} \) (cm\(^{-2}\)) = 24, the area becomes constant because the scattered and reflected components dominate the spectra.

### 4.6. Maximum-likelihood Fitting

Maximum-likelihood (ML) fitting was performed to estimate the obscured fraction of quasars using the high-redshift AGN sample. The sample for the ML fitting was constructed from the 304 \( z = 2–5 \) AGN detected in the 2–10 keV band. For each source, the survey area based on log \( L_X \), \( z \), and log \( N_{HI} \) was calculated in order to evaluate the possibility of detecting that AGN. One AGN was removed since the corresponding survey area is zero. For the AGN detected only in the 2–10 keV band, we assign the column density of log \( N_{HI} \) (cm\(^{-2}\)) = 23.5.

The ML method is a parametric fitting method, which uses the observed parameters of each object without binning. The likelihood function \( L \) is generally defined as the product of all probability densities \( P \) in the sample. The probability density is defined as the probability of finding the \( i \)th object with log \( N_{HI} \) at log \( L_X \) and \( z \) as \( P_i \):

\[
P_i = \frac{f_{\text{abs}}(L_X^i, z^i; \log N_{HI})(L_X^i, z^i, \log N_{HI})}{\int_{24}^{30} f_{\text{abs}}(L_X^i, z^i; \log N_{HI})(L_X^i, z^i, \log N_{HI})d \log N_{HI}},
\]

where \( f_{\text{abs}} \) and \( \Omega \) are the absorption function and survey area function, respectively.

The likelihood function in the logarithmic form \( M \) is then defined as the sum of the logarithmic probability density of the \( i \)th AGN within the fitting sample

\[
M(\psi^2_{43.75}, \epsilon) = -2\sum_i \ln P_i(\psi^2_{43.75}, \epsilon).
\]

The best-fit parameters are obtained by minimizing the likelihood function over the parameter space of interest. In our case, we set \( \epsilon \) and \( \psi^2_{43.75} \) as free parameters. The 1σ uncertainty of the best-fit parameters was estimated based on the range where the log-likelihood value changes from the minimum by 1.

### 5. Results

#### 5.1. Obscured Fraction in the Luminous End

The ML fitting was performed in two ways, (1) by fitting \( \psi^2_{43.75} \) with fixed \( \epsilon = 1.7 \), which is the parameter determined in the local universe (Ueda et al. 2014) (hereafter we refer to as “1D”) and (2) by fitting both \( \psi^2_{43.75} \) and \( \epsilon \) simultaneously (hereafter “2D”).

First, the fitting was performed with the maximum obscured fraction \( \psi_{\text{max}} \) set to be 0.84 following Ueda et al. (2014). The results suggested that the fitting is affected by the choice of \( \psi_{\text{max}} \). Thus, the maximum obscured fraction \( \psi_{\text{max}} \) was set to 0.99 instead. The best-fit results from the 1D and 2D cases are shown in Table 5.

Figure 12 shows the observed \( \log N_{HI} \) distribution of the 2–10 keV band detected AGN. The comparison suggests that the fitting results from the 2D case reproduce the observed distribution better than 1D. Figure 13 shows the predicted distribution of redshift, 2–10 keV luminosity, and log \( N_{HI} \) based on the best-fitted parameters from the 2D-ML-fit. The predicted distributions reproduce the observed distribution well.
The intrinsic number of AGN in each luminosity range of $L_\times$ is estimated to be $0.77 \pm 0.04$, as it corresponds approximately to the 1D and 2D-ML fitting results. For the expected number of AGN, the survey area is replaced with the survey area function $\Omega(L_\times, z, \log N_{HI})$.

Based on the best-fit parameters from the 2D-ML fit, the obscured fraction of the luminous quasars with $L_\times (\text{erg s}^{-1}) = 44.5$ is estimated to be $0.78 \pm 0.02$. The best-fit slope suggests that the obscured fraction shows almost no luminosity dependence. However, it should be noted that the luminosity coverage is limited to between $L_\times (\text{erg s}^{-1}) = 44$ and 45, and the obscured fraction at higher luminosity cannot be constrained.

### 5.3. Systematic Uncertainties in the Analysis

First, systematic uncertainties arise from the photon index and the reflection strength assumed in the phenomenological AGN model, which is used to calculate the column density using the hardness ratio.

If we assume a photon index of 1.7 with the same reflection strength, the obscured fraction for luminous quasars with $L_\times (\text{erg s}^{-1}) = 44-45$ is estimated to be $0.77 \pm 0.04$, which is approximately $9\%$ lower than when assuming a photon index of 1.8 ($0.85 \pm 0.03$). Assuming a photon index of 1.8 with a stronger reflection strength of 1.3 reduces the obscured fraction to $0.82 \pm 0.03$ for luminous quasars, corresponding to a systematic change of approximately $3\%$. The assumption of the photon index and the reflection strength does not strongly affect the estimate of the column density ratio ($\epsilon$) since the results are consistent with each other within uncertainties.

Second, the obscured fraction based on a hard X-ray selected AGN sample could suffer from a bias in which the 2–10 keV count rates close to the detection limit are larger than the true count rates due to statistical fluctuations of the photon count rates (Eddington bias). As a result, hard X-ray-selected AGN samples may have harder hardness ratios and as a result, larger column densities overall. This statistical fluctuation may also affect intrinsically unobscured AGN, which makes them show a large column density consistent with obscured AGN due to positive fluctuation in the 2–10 keV band.

### Table 5

| Parameter | 1D | 2D |
|-----------|----|----|
| $\epsilon$ | 1.7 (fixed) | $1.4 \pm 0.3$ |
| $\psi_{43,75}$ | $1.01 \pm 0.04$ | $0.99 \pm 0.04$ |

Ueda et al. (2014) suggest an obscured fraction of $0.50 \pm 0.09$ above redshift 2.

For comparison with other studies, the obscured fraction of AGN with $L_\times (\text{erg s}^{-1}) = 44-45$ based on the 2D-ML best-fit parameters is adopted. For CTN-AGN, the obscured fraction is $0.85 \pm 0.04$. If we calculate the obscured fraction with $\log N_{HI} (\text{cm}^{-2}) > 23$, the obscured fraction is $0.50 \pm 0.06$.

### 5.2. Slope of the Obscured Fraction on the Luminous End

The ML method applied in Section 5.1 assumes that the slope of the obscured fraction dependence on luminosity does not depend on redshift. Therefore, the rate at that the obscured fraction increases is the same at all luminosity as long as it has not saturated. Here, the slope of the luminosity dependence on the luminous end at high redshift was investigated by applying the ML method to determine the slope parameter $\beta$ assuming that $\psi_{43,75} = 0.73$ following Ueda et al. (2014) and $\epsilon = 1.4$ following the 2D-ML-fitting results.

The best-fit slope determined using the ML fit is $-0.09 \pm 0.03$, which suggests an obscured fraction of luminous CTN-AGN with $L_\times (\text{erg s}^{-1}) = 44-45$ is $0.78 \pm 0.02$. The best-fit slope suggests that the obscured fraction shows almost no luminosity dependence. However, it should be noted that the luminosity coverage is limited to between $L_\times (\text{erg s}^{-1}) = 44$ and 45, and the obscured fraction at higher luminosity cannot be constrained.
Last, the estimates of the column density and luminosity may also be affected by the uncertainty in the photometric redshift. Contamination from low-redshift AGNs (z < 2) can also affect the best-fit results. We estimate the contamination rate from low-redshift AGN (z_{spec} < 2) based on the fraction of outlier spectroscopically confirmed quasars above redshift 2 (AGN above the red-dashed line with z_{phz} > 2 as shown in Figure 6) over all AGN above redshift 2 (z_{phz} > 2) to be 31%.

6. Discussion

6.1. Redshift Dependence of the Obscured Fraction

Using the best-fit parameters from the 2D-ML method, we calculate the obscured fraction of CTN-AGN as a function of the 2–10 keV luminosity by assuming the luminosity-dependent obscured fraction presented in Section 4.4. Figure 14 shows the best-fit obscured fraction based on the 2D-ML best-fit parameters compared with previous measurements in the local universe (bottom panel) (Burlon et al. 2011; Ueda et al. 2014; Georgakakis et al. 2017) and at z = 2–5 (top panel) (Hasinger 2008; Iwasawa et al. 2012; Hiroi et al. 2012; Kalfoountzou et al. 2014; Ueda et al. 2014; Liu et al. 2017). Our estimate of the obscured fraction is larger than the obscured fraction in the local universe, which supports the increasing trend in the obscured fraction toward high redshift. At high redshift, our estimate of the obscured fraction at log L_X (erg s^{-1}) = 43.75 is the largest compared with studies in the same redshift range, except for Liu et al. (2017), in which the obscured fraction in the same redshift range and with log L_X (erg s^{-1}) = 43.5–44.2 is determined to be 0.91 ± 0.03. The luminosity coverage of their sample is below the luminosity range of our sample but consistent with our estimate of the obscured fraction at log L_X (erg s^{-1}) = 43.75 if we extrapolate the fraction toward lower luminosity without the maximum limit of 0.84 and compare it with the obscured fraction of AGN with log N_H (cm^{-2}) > 22.

The upper and lower panels of Figure 15 show the obscured fraction of AGN with log N_H (cm^{-2}) > 22 and >23, respectively. The obscured fraction is larger than the obscured fraction in the same luminosity range as in Ueda et al. (2014) (black solid line). The larger obscured fraction suggests that the evolution of the obscured fraction needs to be stronger at z < 2. In order to reconcile with our estimate of the obscured fraction, the evolution parameter (a1) of the obscured fraction must be 0.75^{+0.10}_{-0.08}. However, this will systematically increase the obscured fraction at z < 2. Alternatively, the obscured fraction may still evolve above redshift 2 and the obscured fraction of high-luminosity AGN saturates at a higher redshift than that of low-luminosity AGN.

Vito et al. (2014) estimated the obscured fraction of log N_H (cm^{-2}) > 23 and log L_X (erg s^{-1}) ≥ 43 AGN at redshift 3–5 to be 0.54 ± 0.05. Our sample resides at a lower redshift than that in Vito et al. (2014). However, our model assumes no evolution of the obscured fraction above redshift two. The obscured fraction of AGN based on the same definition and redshift and luminosity range was estimated to be 0.57 ± 0.05 based on the best-fit parameters of the 2D-ML fit. This is consistent with that in Vito et al. (2014) within 1σ uncertainty.

Liu et al. (2017) also examined the obscured fraction using AGN samples from the C-COSMOS legacy (Civano et al. 2011, 2016) combined with the Chandra Deep Field South 7M catalog (CDF-S; Luo et al. 2017). The obscured fraction of AGN with log N_H (cm^{-2}) = 22–23 among those with log N_H (cm^{-2}) < 23. For AGN with log L_X (erg s^{-1}) = 44.1–44.9 at redshift 2–3, the fraction is 0.67 ± 0.07. The obscured fraction estimated based on the same definition between log L_X (erg s^{-1}) = 44 and 45 using the 2D-ML best-fit parameters is 0.70^{+0.05}_{-0.01}, which is consistent with the above value within 1σ uncertainty.

Recently, Gilli et al. (2022) estimated the obscured fraction of AGNs as a function of redshift by constructing an ISM model based on ALMA data. Our results are consistent with
those of Gilli et al. (2022) for AGN with $N_H$ (cm$^{-2}$) > 23 with $L_X$ (erg s$^{-1}$) $\sim$ 44, but higher than the 1σ uncertainties for the obscured fraction of AGN with $N_H$ (cm$^{-2}$) > 22 at the same luminosity. The larger obscured fraction in our study in the case of $N_H$ (cm$^{-2}$) > 22 may be due to the difficulty in separating unobscured AGN from mildly obscured AGN through the hardness ratios.

6.2. Implications of an Increasing Obscured Fraction

The increasing trend in the obscured fraction has strong implications for the physical structure and evolution in the nuclear environment of the AGN. The larger obscured fraction at high redshift implies that there is a larger amount of obscuring material within the nuclear, circumnuclear, or host galaxy scale compared to AGN in the local universe. The trend may be a direct result of the evolution in the host galaxy-scale properties of ISM. Observations of massive galaxies at high redshift show that they are more compact (van der Wel et al. 2014) and have higher gas fractions (Tacconi et al. 2010; Carilli & Walter 2013) compared to those in the local universe with the same stellar mass. This may suggest that...
the gas density is higher on all spatial scales of the host galaxy compared to those in the local universe. As a result, the higher occurrence of obscuration in the host galaxy or circumnuclear region may explain the increasing trend of the obscured fraction of CTN-AGN (Buchner & Bauer 2017; Buchner et al. 2017; Circosta et al. 2019; Fabian et al. 2008; Gilli et al. 2022) and may explain the Compton-thick fraction (D’Amato et al. 2020). Due to the denser ISM and metal abundance of gas in the circumnuclear region and host galaxy at high redshift, AGN feedback can be less efficient in clearing sight lines since gas can easily cool and replenish the obscuring material (Trebitsch et al. 2019). The decrease in the obscured fraction from high to low redshift may then be explained due to gas consumption by star formation and AGN accretion (Hirschmann et al. 2014). Feedback by AGN-driven winds further reduces the obscuration within and possibly beyond the nuclear region, as spectroscopic observations of high-redshift AGN show that AGN can drive winds with velocities from several hundred up to a thousand kilometers per second, such outflows can reach out to several kiloparsecs beyond the nuclear region (Collet et al. 2016; Nesvadba et al. 2017; Davies et al. 2020).

Another possibility is that the trend in the obscured fraction is driven by the triggering of AGN by major mergers. Galaxy merger simulations suggest that luminous quasars may go through an evolutionary sequence. In this scenario, strong gravitational effects by a merging event funnel large amounts of gas and dust toward the nuclear region, which results in heavily obscured nuclear activity (Hopkins et al. 2006, 2008; Hickox et al. 2009). During the obscured quasar phase, obscuration with a large column density ($\log N_H > 22$) as well as near-Eddington limited accretion are induced. This phase is expected to last as long as 10 times the following blowout phase, in which strong winds driven by the AGN blows out the obscuring material leaving an unobscured quasar at the end of the sequence. It is possible that the fraction of AGN triggered by major merger is higher at higher redshifts; observations of merging galaxies in the local universe show that they are heavily obscured compared to those triggered in other processes (Ricci et al. 2017c, 2021).

Lastly, the trend in obscured fraction with redshift may be compatible in the context of radiation pressure on nuclear dust where after exceeding an $N_H$ critical effective Eddington ratio, the dust in the nuclear region is blown out (Fabian et al. 2006, 2008, 2009; Ishibashi & Fabian 2015). In contrast to the merger-driven scenario, mergers are not a prerequisite in order to drive strong outflows. In this model, nuclear obscuration larger than $N_H > 5 \times 10^{21}$ cm$^{-2}$ occurs from long-lived dust clouds near the AGN while milder obscuration is due to dust lanes outside the SMBH gravitational sphere of influence and independent of the AGN Eddington ratio. At high redshift, accretion activity occurs with higher Eddington ratios than AGN in the local universe (Nobuta et al. 2012; Schulze et al. 2015). The higher average Eddington ratio suggests more AGN are in a blowout phase and less affected by nuclear obscuration. One possibility to explain the increasing trend of the obscured fraction is that the dust abundance in the nuclear region for high-redshift AGN is lower than those in the local universe (Fabian et al. 2009).

6.3. High-redshift CTK-AGN

CTK-AGN are heavily obscured AGN with $\log N_H > 24$. Due to the heavy obscuration, the AGN X-ray continuum emission is strongly suppressed; thus, their detection requires the selection of hard X-rays above the rest frame of 10 keV.

Among the high-redshift AGN sample, 53 AGN were detected only in the 2–10 keV band and have lower limits for $\log N_H$ larger than $\log N_H > 23$. Considering the band shifting effect, non-detection in the 0.5–2 keV band suggests their AGN X-ray continuum below rest frame 6–8 keV is heavily suppressed. Therefore, the 53 AGN may be considered as CTK-AGN candidates at high redshift.

Assuming the phenomenological AGN model, the survey area function presented in Section 4.5, and using the absorption function in Section 4.4 with $f_{\text{CTK}} = 1$, the expected number of CTK-AGN detected with $\log N_H = 24–26$ is $25.61^{+1.95}_{-1.61}$. This may suggest that roughly half of the 2–10 single band detected AGN may be CTK-AGN candidates. However, the column density determination based on the 0.5–2.0 and 2.0–10.0 keV bands at HR cannot discriminate heavily obscured CTN-AGN and CTK-AGN in the sample as shown in Figure 8. In addition, the adoption of physical torus models, X-ray spectral analysis, or secondary tracers of AGN luminosity are generally required to reliably confirm the CTK nature (Ricci et al. 2017b). A full discrimination of the CTK population is beyond the scope of this paper.

6.4. Rest-frame SED

We examine the correspondence between the X-ray obscuration and the SED shapes between UV and IR bands by examining the rest-frame SED of the high-redshift quasars ($\log L_X(\text{erg s}^{-1}) > 44.5$) detected in the 2–10 keV band. For each quasar, the SEDs were constructed by interpolating between the 12-band photometry shifted to rest frame and normalized at 5500 Å. We also include additional IRAC3 (5.8 μm), IRAC4 (8 μm), and MIPS 24 μm bands photometry from the SWIRE data set (Lonsdale et al. 2003). The AGN sample was separated into X-ray unobscured ($\log N_H < 22$) and X-ray obscured ($\log N_H > 22$) AGN based on the column density. The median SED of X-ray unobscured and obscured quasars was constructed by median combining the individual SED of X-ray unobscured and obscured quasars, respectively.

Figure 16 shows the median SED of the high-redshift quasars compared with the type 1 and type 2 QSO SEDs in Polletta et al. (2007) as well as the median SED of high-redshift BL-AGN. The median SED of X-ray unobscured AGN is flat similar to the median SED of the BL-AGN. This is consistent with the expected power-law SED of unobscured AGN. On the other hand, the median SED of X-ray obscured AGN shows a redder UV, optical, and NIR continuum compared to the BL-AGN.

Both the X-ray unobscured and obscured AGN show a variety of SED shapes as shown in the 16th and 84th percentile SED distribution. More than 16% of the obscured AGN have a blue UV continuum similar to the median SED of the BL-AGN, while 16% of the unobscured AGN are redder than the median SED of the obscured AGN. This suggests that the correspondence between the UV-optical SED and X-ray obscuration is not strong.

In order to further examine the correspondence between optical properties, UV-optical-NIR color and morphology, and X-ray obscuration, the AGN were separated into four groups based on the rest frame $\mu$–$H$ color and UV morphology. The rest-frame colors were calculated from the SED fitting results.
of LePhare. The morphology was inferred from the HSC $i$-band flux ratio between the HSC S20A $i$-band PSF and cmodel flux. The PSF (cmodel) flux is derived by fitting the PSF (PSF or galaxy) model. If the AGN appears as a point source on the image, then the ratio is expected to approach 1 while extended AGN will have a smaller flux ratio. We consider AGN with a flux ratio larger than 0.95 as point sources.

The distribution of the high-redshift AGN on the rest-frame $u' - H$ color and the flux-ratio plane is shown in Figure 17. Most of the X-ray unobscured AGN have blue rest-frame color and morphology similar to a point source. For the X-ray obscured AGN, approximately 49% show extended morphology while the remaining sources possess morphology consistent with a point source. X-ray obscured AGN also show a broad color distribution where some X-ray obscured AGN have rest-frame optical NIR colors consistent with the X-ray unobscured AGN.

The high-redshift AGN sample was divided into four samples based on the color against the morphology plane, and the column density distribution of each sample is shown in Figure 18. Most of the AGN with extended morphology have log $N_\text{H}$ consistent with the X-ray obscured AGN with log $N_\text{H}$ (cm$^{-2}$) > 22. On the other hand, blue point-source AGN (case 1) has a mixture of column density log $N_\text{H}$ both consistent with X-ray obscured and unobscured AGN.

For each type of morphology, we performed a K-S test between the log $N_\text{H}$ distribution for those that are extended and those which are point source. The possibility that the blue and red point source (case 1 and case 2) were drawn from the same parent distribution is rejected at a 0.01 level of significance ($D_{\text{max}} = 0.391$, $D_{\text{crit}} = 0.277$). However, the number of samples compared between case 1 and case 2 is limited. On the other hand, the possibility that the blue and red extended sources (case 3 and case 4) were drawn from the same parent distribution cannot be rejected at a 0.01 level of significance ($D_{\text{max}} = 0.256$, $D_{\text{crit}} = 0.405$). This suggests that both red and blue extended sources may contain X-ray obscured AGN.

The largest difference between the cumulative distribution of log $N_\text{H}$ of blue extended AGN and red extended AGN comes from the log $N_\text{H}$ (cm$^{-2}$) = 22–23 bin. Obscuration of the largest column densities (log $N_\text{H}$ (cm$^{-2}$) > 23) generally occurs on the nuclear scale while moderate obscuration can occur on the kiloparsec scale (Hickox & Alexander 2018). It is possible that red extended X-ray obscured AGN have larger amounts of gas and dust in the host galaxy scale than blue extended X-ray obscured AGN. As a result, the red color can be explained by the larger dust extinction in the ISM. Another possibility is that the blue UV-optical NIR colors found in blue extended X-ray obscured AGN is scattered light from the luminous quasars (Alexander et al. 2013; Assef et al. 2016; Alexandroff et al. 2018; Assef et al. 2020) or from ongoing unobscured star formation in the host galaxy and that the AGN is obscured as suggested by the extended morphology of the host galaxy.

In order to examine the relation between the UV spectral properties and X-ray obscuration, the optical spectra of 26 AGN with SDSS spectroscopic data were examined. The morphology of 21 of these AGN is consistent with a point-source object. Three sources have flux ratios >0.90 close to the stellarity threshold. Of the 26 AGN, two AGN have flux ratios consistent with an extended source with one AGN showing a narrow C IV emission line ($\sigma < 1000$ km s$^{-1}$). Among these 26 AGN, two AGN show absorption associated with the C IV emission line. The remaining AGN have spectra consistent with optically unobscured AGN with broad UV emission lines. Broad absorption line quasars (BAL-QSO) are known to be X-ray obscured (Page et al. 2011; Maiolino et al. 2010; Page et al. 2017; Strebyanska et al. 2010); therefore, some of the X-ray obscured AGN with blue UV continuum and point-source morphology may be BAL-QSOs. X-ray obscuration may also be due to warm or ionized absorbers with no dust (Piconcelli et al. 2005; Merloni et al. 2014). This is overall consistent with the concept of the unified model (Antonucci 1993; Urry & Padovani 1995). Detection of broad emission lines and blue continuum suggests that the line of sight toward

---

**Figure 16.** (Left) Median SED of X-ray unobscured AGN with log $N_\text{H}$ (cm$^{-2}$) < 22 (black solid line). The 16th and 84th percentile SEDs are plotted with black dashed lines. (Right) Same as the left panel but for X-ray obscured AGN with log $N_\text{H}$ (cm$^{-2}$) ≥ 22 (black solid line). In both diagrams, the median SED of high-redshift BL-AGN is shown in orange while type 1 QSO and type 2 QSO SED from Polletta et al. (2007) are shown in blue and red, respectively. The number in parenthesis represents the number of objects used to produce the SED. All SEDs are normalized at 5500 Å.
the broad-line region and the accretion disk is unobscured by
dust but may contain ionized or dust-free X-ray absorbers along
the line of sight due to the strong UV radiation.

We conclude that the trend in which X-ray unobscured AGN
have a flat SED, blue UV-optical color, and point-source
morphology while X-ray obscured AGN have a reddened SED
and extended morphology is present but the correspondence is
not tight. The large variety in the SED shapes may be due to
different types of X-ray absorbers, scattered AGN emission,
and the variety of host galaxy star formation and dust content.

6.5. How Obscured Quasars are Missed by Optical Color
Selection

A large number of high-redshift unobscured quasars are
selected based on their optical color and morphology in wide
and deep optical imaging surveys. This technique enables us to
select faint unobscured AGN but can miss obscured AGN. We
examine the relationship between the high-redshift X-ray
selected AGN and those with the optical color and morphology
criteria as used in Akiyama et al. (2018) and Pouliasis
et al. (2022).

Figure 19 shows the color distribution of AGN in the entire
X-ray sample compared with those at redshift 3.5–4.1. The
color and morphology selection in Akiyama et al. (2018) was
constructed to select AGN with point-source morphology at
redshift 3.5–4.1. The color selection was designed to minimize
contamination by AGN at other redshifts, galaxies at \( z \sim 1 \),
and low-mass galactic stars, which are the major contaminants. In
this section, point-source morphology is defined based on the
adaptive moment measurement (Hirata & Seljak 2003) used in
Akiyama et al. (2018) which is available in the HSC-SSP
database. We adopt the condition

\[
\begin{align*}
\frac{i_{\text{hsmsourcemomentsround_shape11}}}{i_{\text{hsmpsfmoments_shape11}}} & < 1.1 \\
\frac{i_{\text{hsmsourcemomentsround_shape22}}}{i_{\text{hsmpsfmoments_shape22}}} & < 1.1
\end{align*}
\]

to define point-source objects. For comparison with the point-
source morphology defined based on the flux ratio in Section
6.4, the adaptive moment criteria correspond to objects with
flux ratios of \( \sim 0.97 \). Approximately 16% of AGN at redshift
3.5–4.1 were selected based on color-selection criterion and
half of them have point-source morphology. It should be noted
that the color criteria are determined for stellar objects brighter
than \( i < 24 \) mag, and most of the X-ray selected objects are
fainter than \( i > 24 \) mag.

In addition to a modified version of the color selection used
in Akiyama et al. (2018), Pouliasis et al. (2022) use the Lyman-
break criteria (Ono et al. 2018). The Lyman-break criterion
selects 65% of the X-ray selected AGN at redshift 3.5–4.1,
including both point-source and extended objects. The
remaining objects, which were not selected by the color-
selection criteria possess redder colors. We conclude that the
AGN selection based on the optical color can miss obscured
AGN, which have reddened or host-dominated colors. More-
over, the application of morphological selection excludes
obscured AGN, which resides in extended host galaxies.

7. Summary

We construct a multiwavelength PSF-convolved photometric
catalog from 12 deep and wide imaging data sets covering from
the \( u^* \) band to 4.5 \( \mu \text{m} \) in the HSC-Deep XMM-LSS survey area
using the HSC \( r \)-band as the high-resolution prior. A sample of
high-redshift AGN was constructed by matching the XMM-
SERVS X-ray point-source catalog (Chen et al. 2018) with the
multiwavelength catalog and selecting AGN above redshift 2
Figure 19. (Top) $g - r$ vs. $r - z$ color distribution. AGN at redshift 3.5–4.1 selected and not selected by the color-morphology criterion in Akiyama et al. (2018) is shown as blue round and red square symbols, respectively. (Bottom) $r - i$ vs. $g - r$ color distribution. AGN at redshift 3.5–4.1 selected and not selected by the color criterion in Ono et al. (2018) is shown as blue round and red square symbols, respectively. For both plots, the gray contours show the color distribution of AGN in the primary sample while objects with point-source morphology are shown with crossed symbols.
This opportunity of observing the universe from Mauna Kea, which has cultural, historical, and natural significance in Hawaii.

This paper makes use of software developed for the Vera C. Rubin Observatory. We thank the Rubin Observatory for making their code available as free software at http://pipelines.lsst.io/.

These data were obtained and processed as part of CLAUDS, which is a collaboration between astronomers from Canada, France, and China described in Sawicki et al. (2019). CLAUDS is based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the CFHT, which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l’Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. CLAUDS uses data obtained in part through the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories, Chinese Academy of Sciences, and the Special Fund for Astronomy from the Ministry of Finance of China. CLAUDS uses data products from TERAPIX and the Canadian Astronomy Data Centre (CADC) and was carried out using resources from Compute Canada and Canadian Advanced Network For Astrophysical Research (CANFAR).

Based on data obtained from the ESO Science Archive Facility with DOI: https://doi.org/10.18727/archive/58.

This research has made use of the NASA/IPAC Infrared Science Archive, which is funded by the National Aeronautics and Space Administration and operated by the California Institute of Technology.

This publication makes use of data products from 2MASS, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

This research made use of Photutils, an Astropy package for the detection and photometry of astronomical sources (Bradley et al. 2020).

Facilities: IRSA, CFHT, Subaru, VISTA, Spitzer, XMM-Newton.

Software: NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), AstroPy (The Astropy Collaboration et al. 2022), Matplotlib (Hunter 2007), Photutils (Bradley et al. 2020), LePhare (Arnouts et al. 1999; Ilbert et al. 2006), SExtractor (Bertin & Arnouts 1996), SWarp (Bertin 2010), T-PHOT (Merlin et al. 2015; 2016), PyPher (Boucaud et al. 2016).

### Appendix

**Catalog Description**

A description of the X-ray AGN catalog is shown in Table 6.

| Number | Column Name | Unit | Description |
|--------|-------------|------|-------------|
| 1      | NameXID     |      | X-ray source ID |
| 2      | RAxcen      | deg  | X-ray center R.A. |
| 3      | DXcen       | deg  | X-ray center decl. |
| 4      | SBml        |      | 0.5–2 keV detection likelihood |
| 5      | HBml        |      | 2–10 keV detection likelihood |
| 6      | FBml        |      | 0.5–10 keV detection likelihood |
| 7      | oircat      |      | Matched optical-IR catalog |
| 8      | RAOir       | deg  | OIR R.A. |
| 9      | DEoir       | deg  | OIR decl. |

**Additional Derived X-Ray Information**

| Number | Column Name | Unit | Description |
|--------|-------------|------|-------------|
| 10     | SBcrt       | ct ks\(^{-1}\) | 0.5–2 keV count rate\(^{a,b}\) |
| 11     | eSBcrt      | ct ks\(^{-1}\) | 0.5–2 keV count-rate uncertainty\(^{a,b}\) |
| 12     | HBcrt       | ct ks\(^{-1}\) | 2–10 keV count rate\(^{a,b}\) |
| 13     | eHBcrt      | ct ks\(^{-1}\) | 2–10 keV count-rate uncertainty\(^{a,b}\) |
| 14     | HR          |      | Hardness ratio\(^{b}\) |
| 15     | eHR         |      | Hardness ratio uncertainty\(^{b}\) |
| 16     | blagn       |      | Optical BL-AGN\(^{c}\) |

**Optical and IR PSF-convolved Photometry**

| Number | Column Name | Unit | Description |
|--------|-------------|------|-------------|
| 17     | XHSCID      |      | XHSC source identifier |
| 18     | RAxhsc      | deg  | XHSC source R.A. |
| 19     | DXhsc       | deg  | XHSC source decl. |
| 20     | tract       |      | HSC tract identifier |
| 21     | patch       |      | HSC patch identifier |
| 22     | nircat      |      | Detection flag\(^{a}\) |
| 23     | usmag       | mag  | CLAUDS ux-band magnitude |
| 24     | eusmag      | mag  | CLAUDS ux-band magnitude uncertainty |
| 25     | gmag        | mag  | HSC g-band magnitude |
| 26     | egmag       | mag  | HSC g-band magnitude uncertainty |
| 27     | rmag        | mag  | HSC r-band magnitude |
| 28     | ermag       | mag  | HSC r-band magnitude uncertainty |
| 29     | imag        | mag  | HSC i-band magnitude |

This research made use of data products from 2MASS, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

This research made use of Photutils, an Astropy package for the detection and photometry of astronomical sources (Bradley et al. 2020).

Facilities: IRSA, CFHT, Subaru, VISTA, Spitzer, XMM-Newton.

Software: NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), AstroPy (The Astropy Collaboration et al. 2022), Matplotlib (Hunter 2007), Photutils (Bradley et al. 2020), LePhare (Arnouts et al. 1999; Ilbert et al. 2006), SExtractor (Bertin & Arnouts 1996), SWarp (Bertin 2010), T-PHOT (Merlin et al. 2015; 2016), PyPher (Boucaud et al. 2016).
| Number | Column Name | Unit | Description |
|--------|-------------|------|-------------|
| 30 | e_imag | mag | HSC i-band magnitude uncertainty |
| 31 | zmag | mag | HSC z-band magnitude |
| 32 | e_zmag | mag | HSC z-band magnitude uncertainty |
| 33 | ymag | mag | HSC y-band magnitude |
| 34 | e_ymag | mag | HSC y-band magnitude uncertainty |
| 35 | Ymag | mag | VIDEO Y-band magnitude |
| 36 | e_Ymag | mag | VIDEO Y-band magnitude uncertainty |
| 37 | Jmag | mag | VIDEO J-band magnitude |
| 38 | e_Jmag | mag | VIDEO J-band magnitude uncertainty |
| 39 | Hmag | mag | VIDEO H-band magnitude |
| 40 | e_Hmag | mag | VIDEO H-band magnitude uncertainty |
| 41 | Ksmag | mag | VIDEO Ks-band magnitude |
| 42 | e_Ksmag | mag | VIDEO Ks-band magnitude uncertainty |
| 43 | ch1mag | mag | SERVS 3.6 μm magnitude |
| 44 | e_ch1mag | mag | SERVS 3.6 μm magnitude uncertainty |
| 45 | ch2mag | mag | SERVS 4.5 μm magnitude |
| 46 | e_ch2mag | mag | SERVS 4.5 μm magnitude uncertainty |
| 47 | EBV | mag | Galactic E(B – V) |
| 48 | AV | mag | Galactic V-band attenuation |
| 49 | f_usmag | … | CLAUDS 1μm-band bad photometry flag |
| 50 | f_gmag | … | HSC g-band bad photometry flag |
| 51 | f_rmag | … | HSC r-band bad photometry flag |
| 52 | f_imag | … | HSC i-band bad photometry flag |
| 53 | f_zmag | … | HSC z-band bad photometry flag |
| 54 | f_ymag | … | HSC y-band bad photometry flag |
| 55 | f_Ymag | … | VIDEO Y-band bad photometry flag |
| 56 | f_Jmag | … | VIDEO J-band bad photometry flag |
| 57 | f_Hmag | … | VIDEO H-band bad photometry flag |
| 58 | f_Ksmag | … | VIDEO Ks-band bad photometry flag |
| 59 | f_ch1mag | … | SERVS 3.6 μm bad photometry flag |
| 60 | f_ch2mag | … | SERVS 4.5 μm bad photometry flag |
| 61 | Fleda | … | LEDA association flag |

### Spectroscopic Redshift and LePhare Photo-z

| Number | Column Name | Unit | Description |
|--------|-------------|------|-------------|
| 62 | zspecid | … | Spectroscopic redshift ID |
| 63 | zspec | … | Spectroscopic redshift |
| 64 | zphot | … | Best photometric redshift |
| 65 | E_zphot | … | Lower 68% confidence photo-z |
| 66 | e_zphot | … | Upper 68% confidence photo-z |
| 67 | chibest | … | Best-fit Chi-square |
| 68 | nband | … | Number of bands used |
| 69 | zsec | … | Secondary photo-z |
| 70 | chisec | … | Secondary photo-z Chi-square |
| 71 | chistar | … | Galactic star photo-z Chi-square |
| 72 | usMag | mag | CLAUDS 1μm-band absolute magnitude |
| 73 | gMag | mag | HSC g-band absolute magnitude |
| 74 | rMag | mag | HSC r-band absolute magnitude |
| 75 | iMag | mag | HSC i-band absolute magnitude |
| 76 | zMag | mag | HSC z-band absolute magnitude |
| 77 | yMag | mag | HSC y-band absolute magnitude |
| 78 | YMag | mag | VIDEO Y-band absolute magnitude |
| 79 | JMag | mag | VIDEO J-band absolute magnitude |
| 80 | HMag | mag | VIDEO H-band absolute magnitude |
| 81 | KsMag | mag | VIDEO Ks-band absolute magnitude |
| 82 | ch1Mag | mag | SERVS 3.6 μm absolute magnitude |
| 83 | ch2Mag | mag | SERVS 4.5 μm absolute magnitude |

### Derived AGN Properties Used in the Analysis

| Number | Column Name | Unit | Description |
|--------|-------------|------|-------------|
| 84 | zprime | … | Best redshift used in the analysis |
| 85 | SBflux | 10^{-3} W m^{-2} | 0.5–2 keV band flux |
| 86 | e_SBflux | 10^{-3} W m^{-2} | 0.5–2 keV band flux uncertainty |
| 87 | HBflux | 10^{-3} W m^{-2} | 2–10 keV band flux |
| 88 | e_HBflux | 10^{-3} W m^{-2} | 2–10 keV band flux uncertainty |
| Number | Column Name | Unit | Description |
|--------|-------------|------|-------------|
| 98     | e_HBloglx   | [10^{-7} W] | 2-10 keV absorption corrected luminosity from HB^a |
| 96     | HBBloglx    | [10^{-7} W] | 2-10 keV absorption corrected luminosity from HB^a |
| 95     | e_SBloglx   | [10^{-7} W] | 0.5-10 keV absorption corrected luminosity from FB^a |
| 94     | SBloglx     | [10^{-7} W] | 0.5-10 keV absorption corrected luminosity from FB^a |
| 93     | B_logNH     | [cm^{-2}] | Column density lower limit |
| 92     | b_logNH     | [cm^{-2}] | Column density lower limit |
| 91     | logNH       | [cm^{-2}] | Hydrogen column density |
| 90     | eFBloglx    | [10^{-7} W] | 0.5-10 keV band flux uncertainty |
| 89     | FBflux      | 10^{-7} Wm^{-2} | 0.5-10 keV band flux |

Notes.

^a SB = 0.5–2 keV, HB = 2–10 keV, FB = 0.5–10 keV.

^b PN-equivalent count rates.

^c BL-AGN flag from Chen et al. (2018).

^d 0 = r-band detected, 1 = H-band detected.

^e 0 = clean photometry, 1 = bad photometry.

^f 0/1 = Not associated/associated with sources in the HyperLeda catalog.

^g Absolute magnitudes calculated from LePhare.

(This table is available in its entirety in machine-readable form.)
