Piercing through Highly Obscured and Compton-thick AGNs in the Chandra Deep Fields. I. X-Ray Spectral and Long-term Variability Analyses

Junyao Li[1,2] , Yongquan Xue[1,2] , Mouyuan Sun[1,2] , Teng Liu[3] , Fabio Vito[4,5] , William N. Brandt[6,7,8] ,
Thomas M. Hughes[1,2,9,10] , Guang Yang[6,7] , Paolo Tozzi[11] , Shifu Zhu[6,7] , Xuechen Zheng[1,2,12] , Bin Luo[13,14,15] ,
Chien-Ting Chen[16] , Cristian Vignali[17,18] , Roberto Gilli[18] , and Xinwen Shu[19]

[1] CAS Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, Hefei 230026, People's Republic of China; lijunyao@mail.ustc.edu.cn
[2] School of Astronomy and Space Science, University of Science and Technology of China, Hefei 230026, People's Republic of China; xuey@ustc.edu.cn
[3] Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse 1, D-85748 Garching bei München, Germany
[4] Instituto de Astrofisica and Centro de Astroingenieria, Facultad de Fisica, Pontificia Universidad Catolica de Chile, Casilla 306, Santiago 22, Chile
[5] Chinese Academy of Sciences South America Center for Astronomy, National Astronomical Observatories, CAS, Beijing 100012, People's Republic of China
[6] Department of Astronomy & Astrophysics, 525 Davey Lab, The Pennsylvania State University, University Park, PA 16802, USA
[7] Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
[8] Department of Physics, The Pennsylvania State University, University Park, PA 16802, USA
[9] CAS South America Center for Astronomy, China-Chile Joint Center for Astronomy, Camino El Observatorio #1515, Las Condes, Santiago, Chile
[10] Instituto de Física y Astronomía, Universidad de Valparaíso, Avda. Gran Bretaña 1111, Valparaíso, Chile
[11] Istituto Nazionale di Astrofisica (INAF)—Osservatorio Astrofisico di Firenze, Largo Enrico Fermi 5, I-50125 Firenze Italy
[12] Leiden Observatory, Leiden University, PQ Box 9513, NL-2300 RA Leiden, The Netherlands
[13] School of Astronomy and Space Science, Nanjing University, Nanjing 210093, People's Republic of China
[14] Key Laboratory of Modern Astronomy and Astrophysics (Nanjing University), Ministry of Education, Nanjing, Jiangsu 210093, People's Republic of China
[15] Collaborative Innovation Center of Modern Astronomy and Space Exploration, Nanjing 210093, People's Republic of China
[16] Astrophysics Office, NASA Marshall Space Flight Center, ZP12, Huntsville, AL 35812, USA
[17] Dipartimento di Fisica e Astronomia, Alma Mater Studiorum, Università degli Studi di Bologna, Via Gobetti 93/2, I-40129 Bologna, Italy
[18] INAF—Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Gobetti 93/3, I-40129 Bologna, Italy
[19] Department of Physics, Anhui Normal University, Wuhu, Anhui, 241000, People's Republic of China

Abstract

We present a detailed X-ray spectral analysis of 1152 active galactic nuclei (AGNs) selected in the Chandra Deep Fields (CDFs), in order to identify highly obscured AGNs ($N_H > 10^{23}$ cm^{-2}). By fitting spectra with physical models, 436 (38%) sources with $L_X > 10^{42}$ erg s^{-1} are confirmed to be highly obscured, including 102 Compton-thick (CT) candidates. We propose a new hardness ratio measure of the obscuration level that can be used to select highly obscured AGN candidates. The completeness and accuracy of applying this method to our AGNs are 88% and 80%, respectively. The observed $\log N - \log S$ relation favors cosmic X-ray background models that predict moderate (i.e., between optimistic and pessimistic) CT number counts. Nineteen percent (6/31) of our highly obscured AGNs that have optical classifications are labeled as broad-line AGNs, suggesting that, at least for part of the AGN population, the heavy X-ray obscuration is largely a line-of-sight effect, i.e., some high column density clouds on various scales (but not necessarily a dust-enshrouded torus) along our sight line may obscure the compact X-ray emitter. After correcting for several observational biases, we obtain the intrinsic N_H distribution and its evolution. The CT/highly obscured fraction is roughly 52% and is consistent with no evident redshift evolution. We also perform long-term (≈ 17 yr in the observed frame) variability analyses for 31 sources with the largest number of counts available. Among them, 17 sources show flux variabilities: 31% (5/17) are caused by the change of N_H, 53% (9/17) are caused by the intrinsic luminosity variability, 6% (1/17) are driven by both effects, and 2 are not classified owing to large spectral fitting errors.

Key words: galaxies: active – galaxies: nuclei – quasars: general – X-rays: galaxies

Supporting material: machine-readable table

1. Introduction

Highly obscured active galactic nuclei (AGNs), which are defined as AGNs with hydrogen column density (N_H) larger than 10^{23} cm^{-2}, are believed to represent a crucial phase of active galaxies. According to our knowledge of coevolution of supermassive black holes (SMBHs) and their host galaxies (for reviews, see, e.g., Alexander & Hickox 2012; Kormendy & Ho 2013) and the hierarchical galaxy formation model, AGN activity may be triggered in a dust-enshrouded environment, into which gas inflows due to either internal (e.g., disk instabilities; e.g., Hopkins & Hernquist 2006) or external (e.g., major mergers; e.g., Di Matteo et al. 2005) processes both fuel and obscure the SMBH accretion, resulting in short-lived

heavily obscured AGNs (e.g., Fiore et al. 2012; Morganti 2017). The subsequent AGN feedback process may blow out the obscuring material and leave out an unobscured optically bright quasar. Compared with unobscured AGNs, AGNs in the highly obscured phase tend to have smaller BH masses, higher Eddington ratios (λ_{Edd}; e.g., Lanzuisi et al. 2015), and larger merger fractions (e.g., Kocevski et al. 2015; Ricci et al. 2017), which may indicate a fast growth state of central SMBHs (e.g., Goulding et al. 2011). Moreover, the cosmic X-ray background (CXB) synthesis models also require a sizable population of highly obscured AGNs, or even Compton-thick (CT) AGNs ($N_H \gtrsim 10^{24}$ cm^{-2}; see, e.g., Comastri 2004; Xue 2017; Hickox & Alexander 2018 for reviews), to reproduce the peak of CXB

at $20-30$ keV (e.g., Gilli et al. 2007; but see Treister et al. 2009). Therefore, the study of highly obscured AGNs across cosmic epochs is vital for our understanding of the AGN triggering mechanism, SMBH growth, AGN environment, and the origin of CXB.

Thanks to the powerful penetrability of high-energy X-ray photons, X-ray observations provide a great window to uncover the mysterious veil of these heavily obscured sources that are likely missed in optical surveys. In the past 20 yr or so, the deep X-ray surveys conducted by *Chandra* (e.g., Alexander et al. 2003; Xue et al. 2011, 2016; Luo et al. 2017), *XMM-Newton* (e.g., Ranalli et al. 2013), *Swift*/BAT (e.g., Baumgartner et al. 2013), and *NuSTAR* (e.g., Lansbury et al. 2017) have provided relatively unbiased AGN samples thanks to their unprecedented depths and sensitivities, which allow us to identify a significant population of heavily obscured AGNs (e.g., Risaliti et al. 1999; Brightman & Ueda 2012; Ricci et al. 2015) using X-ray color, spectral analysis, and/or the stacking technique (e.g., Alexander et al. 2011; Iwasawa et al. 2012; Georgantopoulos et al. 2013; Brightman et al. 2014; Corral et al. 2014; Del Moro et al. 2016; Koss et al. 2016). Moreover, the combination of mid-infrared (MIR), optical, and X-ray data provides additional methods to select heavily obscured systems, such as MIR excess (e.g., Daddi 2007; Alexander et al. 2008; Luo et al. 2011), high 24 μm-to-optical flux ratio (e.g., Fiore et al. 2009) and high X-ray-to-optical flux ratio (e.g., Fiore et al. 2003).

Among various methods, X-ray spectroscopy provides the most direct and unambiguous way to measure the column density of the obscuring materials. Several previous studies have focused on deriving the intrinsic N_H distribution corrected for the survey biases using deep *Chandra* survey data. Tozzi et al. (2006) presented an N_H distribution that has an approximately lognormal shape peaking at $\sim 10^{23}$ cm^{-2} and with an excess at $\sim 10^{24}$ cm^{-2}. Liu et al. (2017, hereafter L17) reported a similar N_H distribution that peaks in a higher N_H range, due to the inclusion of more sources with low X-ray luminosities (L_X) and high redshifts than that of Tozzi et al. (2006), which are expected to have relatively high N_H values. However, both works only focused on bright AGNs, and neither was dedicated to or optimized for investigating highly obscured sources. In particular, L17 excluded CT AGNs in their work and only focused on the Compton-thin (CN) population. Hence, the absorption distribution and evolution of the most deeply buried AGNs are still unclear, especially at high redshifts (Vito et al. 2018). Therefore, unveiling the apparently faint, CT regime using the deepest X-ray survey data is indispensable for fully understanding the entire AGN population.

There have also been several attempts to constrain the obscured AGN fraction and CT fraction on the basis of modeling the CXB (e.g., Gilli et al. 2007; Akylas et al. 2012), the X-ray luminosity function (LF; e.g., Aird et al. 2015; Buchner et al. 2015), or X-ray spectral analysis (e.g., Burlon et al. 2011). Although most of the studies support the picture of evolved absorption with redshift and luminosity (e.g., Hasinger 2008; Liu et al. 2017), the value of the CT fraction varies from study to study, ranging from a few percent to $\sim 50\%$, largely due to the limited sample sizes, the use of different X-ray spectral models, the unknown contribution from Compton reflection (Treister et al. 2009; Akylas et al. 2016), and the relatively poor quality of X-ray spectra at

high-N_H circumstances. Therefore, deeper X-ray observations, as well as the physically appropriate spectral models for highly obscured AGNs, are needed to robustly characterize the obscuration properties.

Among the X-ray surveys, the Chandra Deep Field (CDF) surveys (see, e.g., Brandt & Alexander 2015; Xue 2017, for reviews), which consist of the 7 Ms Chandra Deep Field-South survey (CDF-S; Luo et al. 2017) and the 2 Ms Chandra Deep Field-North survey (CDF-N; Xue et al. 2016) along with the 250 ks Extended Chandra Deep Field-South survey (E-CDF-S; Xue et al. 2016), provide us the most promising data to study highly obscured AGNs at high redshifts. In particular, the 7 Ms CDF-S, which is the deepest X-ray survey to date, significantly improves the count statistics that allow us to extract high-quality X-ray spectra, detect more faint, highly obscured sources, and perform more robust spectral analyses compared with previous 4 Ms analyses (e.g., Brightman et al. 2014). Furthermore, recent works suggest that the power spectral density (PSD) break frequency of AGN light curves might be related to N_H variability (Zhang et al. 2017; González-Martín 2018), which makes obscuration a very important factor in investigating the driving mechanism of AGN variability. Benefiting from the very long time span of the 7 Ms CDF-S data (16.4 yr in the observed frame), we are able to, for the first time, quantify the detailed variability behavior for a large, dedicated sample of highly obscured AGNs, in order to better understand the location of the obscuring materials and their contribution to AGN variability (see also Yang et al. 2016, hereafter Y16).

In this study, we construct the largest dedicated highly obscured AGN sample in the deepest Chandra surveys, which enables us to extend the studies of deeply buried sources to lower luminosities and higher redshifts in great detail, and present systematic X-ray spectral and long-term variability analyses to study their evolution and physical properties. This paper is organized as follows. We describe our data reduction procedure and sample selection in Section 2. In Section 3 we introduce our spectral fitting method. The spectral fitting result, a new hardness ratio measure of the obscuration level, the number counts of CT AGNs, and the constraint on CXB models are presented in Section 4. The reprocessed components and the covering factor (CF) of the obscuring materials are discussed in Section 5. In Section 6, by correcting for several observational biases, we constrain the intrinsic N_H distribution representative for the highly obscured AGN population and study its evolution across cosmic time. In Section 7, we select a subsample of highly obscured AGNs that have largest counts available and perform detailed long-term X-ray variability analyses, in order to find out the variable fraction, as well as the main driven mechanism of variability.

Throughout this paper, we adopt a Galactic column density of $N_H = 8.8 \times 10^{19}$ cm^{-2} for the CDF-S and $N_H = 1.6 \times 10^{20}$ cm^{-2} for the CDF-N (Stark et al. 1992). We adopt cosmological parameters of $H_0 = 70.0$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.30$, and $\Omega_\Lambda = 0.70$. All given errors are at the 1σ confidence level unless otherwise stated.

2. Data Reduction and Sample Selection

This work is based on the *Chandra* data in the 7 Ms CDF-S main-source catalog (Luo et al. 2017) and the 2 Ms CDF-N main-source catalog (Xue et al. 2016). Since the exposure in the E-CDF-S is ~ 8–28 times shallower than in the CDF-N and

CDF-S, we exclude the E-CDF-S data in this work. The redshift information for each source is adopted from the two catalogs that presented a preferred redshift value given by the ZFINAL keyword. The ZFINAL redshift values were collected and selected based on various published catalogs and followed a general preference order of spectroscopic redshift over photometric redshift (see Section 4.3 of Luo et al. 2017 and Section 2.3.4 of Xue et al. 2016 for details). The redshift (spectroscopic redshift) completenesses in the CDF-S and the CDF-N main-source catalogs are 99.4% (64.8%) and 95.2% (52.4%), respectively, and the corresponding mean 1σ photometric redshift errors are about 0.21 and 0.17, respectively.

The source spectra from individual observations were extracted using the ACIS Extract (AE) software package (Broos et al. 2010). AE generates the point-spread function (PSF) model based on the MARX ray-tracing simulator and constructs a polygonal extraction region that corresponds to an encircled-energy fraction of ∼90%. For crowded sources, AE adopts smaller extraction regions to avoid overlapping polygonal regions. The background spectra were extracted using the BETTER_BACKGROUNDS algorithm. The most significant aspect of the above photometry and spectral extraction procedure, compared to the widely used circular-aperture extraction, is that it can obtain photometry and spectra as accurate as possible and remove the contamination from neighboring sources to faint sources to the greatest extent (see Section 3.2 of Xue et al. 2011 for details). This is extremely important for our work since we are dealing with the highly obscured AGNs that are generally fainter and expected to have limited counts owing to significant obscuration.

The spectra eventually used in this work are the merged spectra for which all the individual observations were matched to a common K_s-band astrometric frame (see Section 2.2.1 in Xue et al. 2016) and stacked using the MERGE_OBSERVATIONS algorithm in AE. The corresponding response matrix files (rmf) and ancillary response files (arf) were also generated and combined during this stage.

We construct our sample by selecting sources that (1) are classified as AGN (TYPE = AGN; see Section 4.5 in Luo et al. 2017 for details), (2) have 0.5–7 keV net counts >20 (FB_COUNTS > 20) to allow basic X-ray spectral fitting, and (3) have redshift measurements (ZFINAL ! = −1) in the main catalogs. The resulting full sample consists of 1152 sources, with 660 from the CDF-S and 492 from the CDF-N, respectively. The counts and redshift distributions for the full sample are shown in Figure 1. The median counts are 112, and 53% (33%) of the sources have counts larger than 100 (200). The median redshift is 1.45, and 603 (52%) sources have spectroscopic redshifts.

3. X-Ray Spectral Analysis

3.1. Spectral Fitting Models

We use MYTorus-based models (Murphy & Yaqoob 2009) to fit the observed-frame 0.5–7 keV spectra of the full sample in order to identify heavily obscured sources. Due to limited counts, we do not bin the spectra because they may lose some key information of the sources. We use the Cash statistic (Cstat in XSPEC; Cash 1979) as our spectral fitting statistic. Cstat has a similar probability distribution to χ^2 statistics and has been proved to be more appropriate in the low-count regime. Since
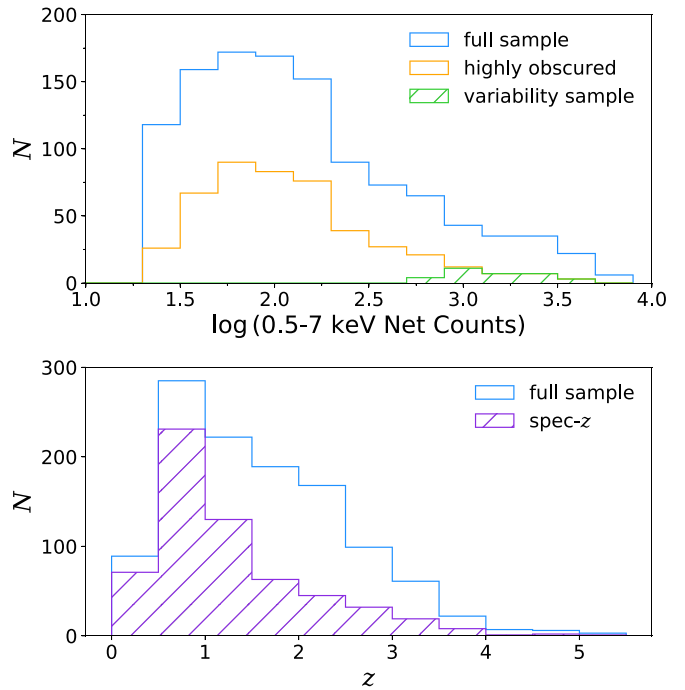
Figure 1. Top: *Chandra* net count distributions of the full sample of 1152 AGNs (blue solid histogram), the highly obscured sources confirmed with X-ray spectral fitting in Section 4 (orange dashed histogram), and the subsample selected to perform variability analyses in Section 7 (solid green histogram), respectively. Bottom: redshift distributions for the full sample (blue histogram) and the 603 sources with spectroscopic redshifts (purple histogram).
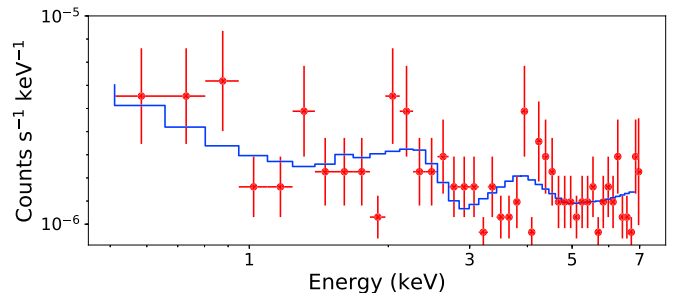
Figure 2. Example of fitting the *Chandra* background spectrum that has 332 counts using the *cplinear* model (binned only for illustration purposes).

Cstat is not appropriate for the background-subtracted spectra, we simultaneously fit the source and background spectra, with the latter (with 1642 median counts) being fit with the *cplinear* model (Broos et al. 2010) that properly describes the observed *Chandra* background by a continuous piecewise-linear function in 10 energy segments (an example is shown in Figure 2). In this way, we are able to maximize the usage of information relevant to the sources.

We adopt two models to fit the source spectra with different components and degrees of freedom (dof) in order to find a statistically robust best-fit one:

1. The MYTorus baseline model: *phabs* × (MYTZ × *zpow* + f_{ref} × MYTS + f_{ref} × *gsmooth* (*MYTL*)).
2. The soft-excess model: *phabs* × (*MYTZ* × *zpow* + f_{ref} × MYTS + f_{ref} × *gsmooth* (*MYTL*) + f_{exs} × *zpow*).
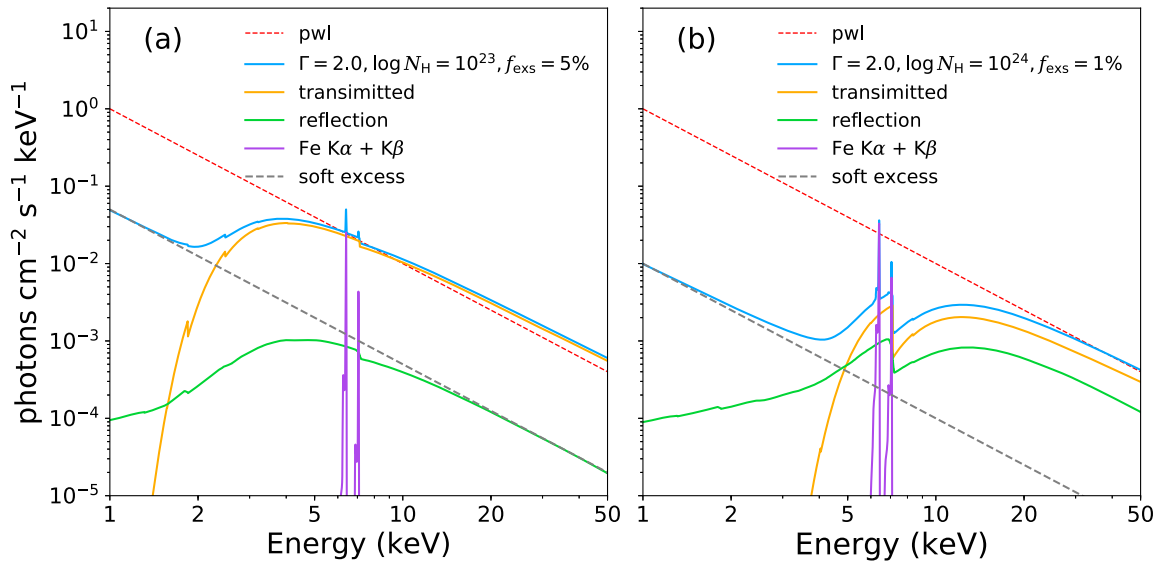
Figure 3. Examples of MYTorus-based models with different parameters used in spectral fitting. The unattenuated continuum (pwl), absorbed continuum (blue curve), transmitted zeroth-order continuum, reflection component, iron emission lines, and soft-excess component are shown in different colors. (a) Model used to generate curve A in Figure 7 with N_H, Γ, and f_{exs} fixed at 10^{23} cm^{-2}, 2.0, and 5%, respectively. (b) Model used to generate curve B in Figure 7 with N_H, Γ, and f_{exs} fixed at 10^{24} cm^{-2}, 2.0, and 1%, respectively.

Table 1
Model Definitions Used in This Work

Model	f_{ref}	f_{exs}
A	10^{-8}	0
B	10^{-8}	$0 < f_{exs} < 0.1$
C	1.0	0
D	1.0	$0 < f_{exs} < 0.1$

These models include all the typical spectral features found in highly obscured AGNs (see Figure 3). The *phabs* component models the Galactic photoelectric absorption. The MYTZ $\times zpow$ term represents the zeroth-order transmitted power-law continuum across the torus that takes into account both photoelectric absorption and Compton scattering processes. MYTS and *gsmooth* (MYTL) stand for the reflection component and the broadened Fe Kα, Kβ fluorescent emission lines, respectively. A second power law is added to represent the soft-excess component often found in AGN spectra, possibly originating from the zeroth-order continuum being scattered by the extended CN materials in obscured AGNs (Guainazzi et al. 2005; Bianchi et al. 2006; Corral et al. 2011). During spectral fitting, the inclination angle θ is fixed at 75°, which is the average value within a range of 60°–90° where the torus intercepts our line of sight (LOS). The nominal normalizations of MYTS, MYTL, and the second power law are set to be the same as that of the first power law, and we use the constants f_{ref} and f_{exs} to represent the real normalizations. f_{exs} is allowed to vary between 0 and 0.1, and we assume a 200 keV high-energy cutoff throughout. One thing to mention is that the best-fit N_H values in MYTorus represent the equatorial values, and we always use the LOS values $N_{H,LOS} = (1 - 4\cos^2\theta)^{\frac{1}{2}} N_H$ in this work.

3.2. Model Selection

Several previous works suggested that the inclusion of the soft-excess and reflection components in the spectral models is crucial for us to correctly estimate N_H of highly obscured sources (e.g., Brightman et al. 2014; Lanzuisi et al. 2015). However, the low counts of many sources do not allow us to apply complex models with free parameters since it could lead to large degeneracies. Therefore, we choose the model components and the parameter spaces according to the following criteria:

1. We fix Γ at 1.8 (e.g., Tozzi et al. 2006; Marchesi et al. 2016) for the 851 sources with counts less than 300. For the 301 sources with counts larger than 300, we set Γ free to find a best-fit value.
2. For all sources, we first fit the spectra with a free f_{ref}. If f_{ref} is less than 10^{-5}, we then fix it to 10^{-8}, which is an arbitrary value set as the lower limit for f_{ref} (i.e., indicating a negligible reflection component); otherwise, we fix f_{ref} at 1, which is the default value adopted in MYTorus.
3. To determine whether we should add a soft-excess component, we compare the Cstat between models with and without considering soft-excess emission. The best-fit model is chosen to be the one that has a statistically robust low Cstat value. More specifically, adding a new soft-excess component should improve the Cstat at least for $\Delta C = 3.84$ with 1 more dof. This criterion is based on the fact that ΔC approximately follows the χ^2 distribution; thus, $\Delta C = 3.84$ is roughly consistent with a >95% confidence level (Tozzi et al. 2006; Liu et al. 2017).
4. In order to avoid extremely untypical Γ caused by the degeneracy between Γ and N_H, we re-fit the spectra of those sources with Γ pegged at 1.4 (i.e., the lowest value permitted by MYTorus) or $\Gamma > 2.4$ (i.e., the typical maximum photon index for high-λ_{Edd} AGNs; e.g., Wang et al. 2004; Fanali et al. 2013) by fixing their Γ at 1.8. If $\Delta C = C_{fix} - C_{free} > 3.84$, we adopt the free Γ value; otherwise, we adopt $\Gamma = 1.8$.

From now on, for convenience, we refer model 1 (2) with negligible f_{ref} as model A (B) and model 1 (2) with f_{ref} fixed at 1.0 as model C (D), as summarized in Table 1. In order to ensure the consistency of the models used for subsequent bias

Table 2
X-Ray Spectral Fitting Results for Highly Obscured AGNs

XID (1)	Field (2)	R.A. (3)	Decl. (4)	z (5)	ztype (6)	HR (7)	Γ (8)	N_H (9)	L_X (10)	Counts (11)	Model (12)
8	CDF-N	188.841072	62.250256	2.794	zphot	0.17	1.80f	24.42	45.30	39	A
10	CDF-N	188.846392	62.292835	1.173	zphot	0.07	1.80f	23.88	43.15	41	C
11	CDF-N	188.847062	62.217590	1.498	zphot	−0.23	1.80f	24.63	44.90	51	D
14	CDF-N	188.853852	62.256847	1.652	zphot	0.59	1.80f	23.81	44.22	105	A
16	CDF-N	188.869802	62.240976	1.732	zphot	0.25	1.80f	23.37	43.97	197	A

Note. Column (1): source ID in the Luo et al. (2017) and Xue et al. (2016) catalogs. Column (2): field. Columns (3) and (4): R.A. and decl. for the X-ray source position. Column (5): ZFINAL redshift. Column (6): redshift type ("zspec": spectroscopic; "zphot": photometric). Column (7): hardness ratio. Column (8): photon index ("f": fixed value). Column (9): LOS column density in units of 10^{24} cm^{-2}. Column (10): logarithm of absorption-corrected rest-frame 2–10 keV luminosity in units of erg s^{-1}. Column (11): 0.5–7 keV background-subtracted net counts. Column (12): best-fit model (see Section 3.1).

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

Table 3
Best-fit Models for Highly Obscured Sources with 0.5–7 keV Net Counts < 100, 100 ⩽ Counts < 300, Counts ⩾ 300, Counts ⩾ 1000,
10^{23} cm^{-2} < N_H < 10^{24} cm^{-2}, and $N_H \geqslant 10^{24}$ cm^{-2}, Respectively

Cases (1)	N (2)	A (3)	B (4)	C (5)	D (6)	$\frac{C+D}{N}$ (7)	N_s (8)	A_s (9)	B_s (10)	C_s (11)	D_s (12)	$\frac{C_s+D_s}{N_s}$ (13)
<100 counts	210	80	21	89	20	51.9%	79	14	11	38	16	68.4%
100–300 counts	146	51	7	69	19	60.3%	67	17	5	29	16	67.2%
⩾300 counts	80	14	5	52	9	76.2%	45	4	2	31	8	86.7%
⩾1000 counts	18	3	2	9	4	72.2%	14	3	1	6	4	71.4%
10^{23} cm^{-2} < N_H < 10^{24} cm^{-2}	328	120	17	166	24	57.9%	136	27	10	80	19	72.8%
$N_H \geqslant 10^{24}$ cm^{-2}	108	25	16	44	23	62.0%	55	8	8	18	21	70.9%
$N_H \geqslant 10^{24}$ cm^{-2} and count ⩾ 150	16	3	1	6	6	75.0%	9	1	1	1	6	77.8%
Total	436	145	33	210	48	58.9%	191	35	18	98	40	72.3%

Note. Column (1): cases of four count bins and two N_H ranges corresponding to CN and CT AGNs. Column (2): total source number in each case. Columns (3)–(6): numbers of sources that are best fitted by models A−D, respectively. Note that models C and D have f_{ref} fixed at 1.0, while models A and B have negligible f_{ref} and thus weak reflection component and Fe K lines, and we show the fraction of the sources that have reflection and iron emissions lines in Column (7). Columns (8)–(13): same as Columns (2)–(7), but for the spectroscopic redshift subsample.

corrections (see Section 6) as much as possible, we have to assume fixed parameters in case of low counts. We justify our models in the Appendix by validating that the usage of fixed parameters does not significantly affect our results. We also compare our fitting results with several previous works (see Section 4.1) and those obtained from the Borus model (Baloković et al. 2018; see Section 5.3) and find good consistency.

We note that the simple absorbed power-law model (e.g., $phabs \times zwabs \times zpow$; hereafter model z) has been widely used in the literature to obtain rough estimates of AGN parameters. However, such a model is not appropriate for highly obscured AGNs since $zwabs$ only models photoelectric absorption and does not take into account the Compton scattering process, which is particularly important in the high-N_H regime. Therefore, we also fit the spectra using model z, aiming at directly testing how accurately such a simple model reproduces the main spectral parameters.

4. Spectral Fitting Results

On the basis of the spectral fitting results, we find that 39% (458/1152) of sources in the full sample are identified to be highly obscured AGNs (hereafter the highly obscured sample [HOS]). The main fitting results for the HOS using MYTorus models A–D are presented in Table 2, and the full version is available online. The best-fit models for sources in different count bins and N_H ranges are summarized in Table 3. In the

following analyses, we only consider the 436 sources that have the absorption-corrected, rest-frame 2–10 keV luminosity (calculated from the lumin command after deleting all the additional components and only keeping the $zpow$ component) $L_X > 10^{42}$ erg s^{-1} to avoid possible contamination from star-forming galaxies.

4.1. Comparisons with Previous Works and the Model of phabs × zwabs × zpow

Thanks to the increased exposure time, the improved data reduction procedure (see Table 1 in Xue et al. 2016 for a summary), the updated redshift measurements, and the usage of physically appropriate spectral fitting models for highly obscured AGNs, we are able to extract higher-quality spectra and obtain more robust parameter constraints than previous works in the same field (e.g., Tozzi et al. 2006), especially for faint sources. Before performing further analyses, we first make direct comparisons with several works that presented spectral fitting parameters in the CDF-S to understand how different methods influence the spectral fitting results. The works used for comparisons are as follows:

1. Liu et al. (2017) (L17): L17 performed an extensive X-ray spectral analysis for the bright sources (hard-band counts >80) in the 7 Ms CDF-S using the $wabs \times (plcabs \times power - law + zgauss + power - law + plcabs \times pexrav \times constant)$ model. The background spectrum
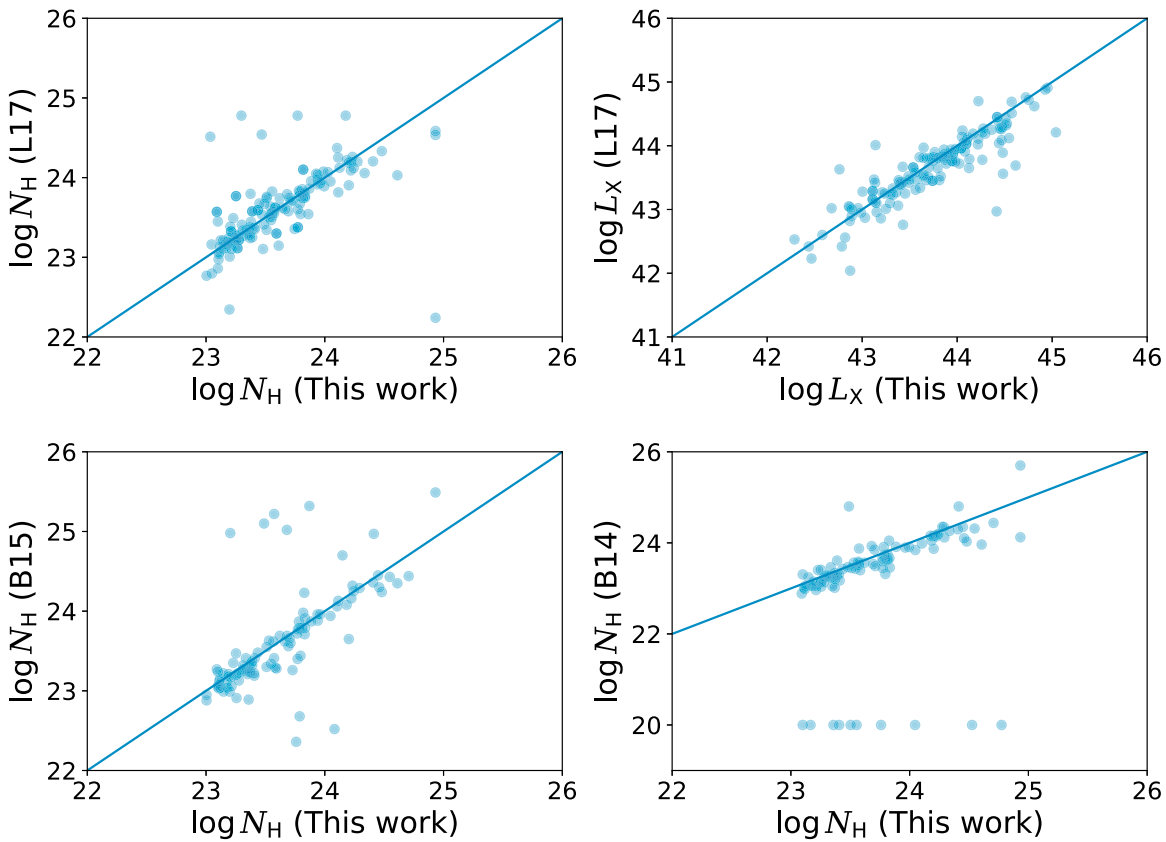
Figure 4. Comparing the spectral fitting results with Liu et al. (2017), Brightman et al. (2014), and Buchner et al. (2015) for common sources with redshift difference $\Delta z < 0.2$.

was also fitted with the *cplinear* model. The redshift information used in the two works is the same.

2. Brightman et al. (2014) (B14): B14 presented X-ray spectral fitting results in the 4 Ms CDF-S using the BNtorus model (Brightman & Nandra 2011). A total of 112 sources are found in common between the two works; 24 of them have redshifts different by more than 0.2 ($\Delta z > 0.2$) compared with the values adopted in the updated 7 Ms catalog and are neglected in the following comparison.

3. Buchner et al. (2015) (B15): The 4 Ms CDF-S spectra were analyzed using a physically motivated torus model and a Bayesian methodology to estimate spectral parameters. A total of 114 sources are found in common between the two works, and 30 of them with $\Delta z > 0.2$ are excluded.

As shown in Figure 4, our measured N_H and L_X values are in general agreement with previous works, despite that for individual sources the usage of different model configurations and data may result in large discrepancies. The largest distinction happens between B14 and our work, in that 10 highly obscured AGNs in our sample are reported to be unobscured in B14. To understand the discrepancy, we refit their spectra using the same model and method as described in B14, and the results again confirm their heavily obscured nature. Therefore, the large discrepancy could be due to the adopted data of different depths, as well as the different data

reduction and spectral extraction methods used in the two works.

The comparisons between the results obtained through MYTorus and model z are shown in Figure 5. Note that since MYTorus does not allow N_H to vary below 10^{22} cm^{-2}, a number of unobscured sources thus have best-fit $N_H = 10^{22}$ cm^{-2}. Therefore, we only consider sources with MYTorus $\log N_H > 22.2$ cm^{-2} in the comparison.

The N_H and the observed 0.5–7 keV flux derived by the two models are generally consistent. But at a given N_H, the absorption-corrected intrinsic 0.5–7 keV flux ratio and the intrinsic L_X ratio between the two models dramatically increase with N_H. It is obvious that model z underestimates the intrinsic luminosity owing to neglect of the Compton scattering process, and the discrepancy is already evident at $N_H \sim 10^{23}$ cm^{-2} (see also Burlon et al. 2011). Therefore, such models with only photoelectric absorption taken into consideration (e.g., *zwabs*, *zphabs*, *ztbabs*) must be used with caution in the highly obscured regime.

4.2. Photon Index Distribution

The photon index distribution for the 62 highly obscured sources with free Γ during the fitting process is shown in Figure 6. The mean value of the distribution is 1.82 ± 0.04, in agreement with previous X-ray spectral analyses in the CDF-S (e.g., Tozzi et al. 2006; Liu et al. 2017) and COSMOS (e.g., Lanzuisi et al. 2013). Note that our Γ distribution peaks at $\Gamma = 1.4$ instead of the mean value 1.8. This is because the
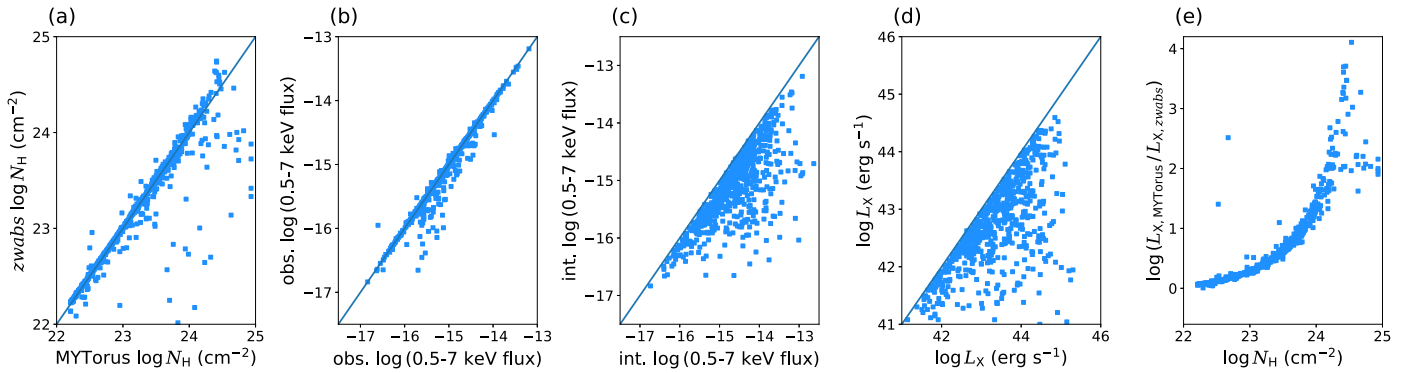
Figure 5. Comparing the spectral fitting results between using the MYTorus (x-axes) and *phabs × zwabs × zpow* (y-axes in all panels but panel (e)) models. (a–d) Comparison results of $\log N_{\rm H}$, observed 0.5–7 keV flux, intrinsic 0.5–7 keV flux, and intrinsic 2–10 keV luminosity, respectively. (e) The log-ratio of the intrinsic 2–10 keV luminosity between MYTorus and model z as a function of the MYTorus $N_{\rm H}$.
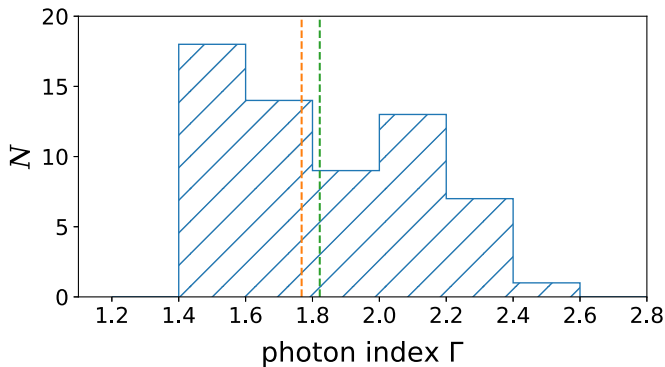

Figure 6. Distribution of the best-fit photon index for the 62 sources with free Γ. The green and red vertical dashed lines show the median ($\Gamma = 1.77$) and mean ($\Gamma = 1.82$) values of the total distribution, respectively.

Table 4
Information of Sources with Extreme Photon Indices

Field	ID	Γ	$N_{\rm H}$	Counts	ztype	zqual
CDF-S	98	1.40	$0.24^{+0.02}_{-0.01}$	3689	zphot	insecure
	135	2.53	$1.59^{+0.21}_{-0.24}$	617	zspec	insecure
	172	1.40	$0.33^{+0.07}_{-0.08}$	534	zphot	...
	243	1.40	$0.12^{+0.01}_{-0.01}$	372	zspec	secure
	249	1.40	$0.18^{+0.02}_{-0.02}$	830	zphot	...
	597	1.40	$0.24^{+0.03}_{-0.04}$	351	zspec	secure
	760	1.40	$0.72^{+0.08}_{-0.07}$	351	zspec	insecure
CDF-N	546	1.40	$0.11^{+0.01}_{-0.01}$	456	zspec	secure

Note. Columns are the same as in Table 2. The zqual column represents the quality of the spectroscopic redshift (secure or insecure). For sources with insecure spectroscopic redshifts, the X-ray source catalogs may choose to adopt the photometric redshifts as ZFINAL (see Section 4.4 in Xue et al. 2011 for details).

photon index is restricted within 1.4–2.6 in MYTorus. If we use MYTZ alone, which does not limit the Γ range to fit the spectra, most sources with Γ pegged at 1.4 will have smaller Γ, and thus the distribution will appear more symmetric with a larger dispersion.

We also try to search the potential correlations among Γ, $N_{\rm H}$, and $L_{\rm X}$ using the Spearman rank correlation test. No correlation between Γ and redshift (Spearman's $\rho = 0.00$, p-value = 0.99) is detected, indicating that the inner disk and corona structures have little evolution across cosmic time. The Spearman tests also suggest no significant correlation between Γ and $N_{\rm H}$ ($L_{\rm X}$).

It is noteworthy that there are eight sources with $\Gamma = 1.4$ or $\Gamma > 2.5$ (see Table 4). Unlike some of the high-count sources for which we artificially fix Γ at 1.8 (see Section 3.2), these sources have significantly worse fits if we do not allow Γ to vary freely (the average improvement of ΔC is 9.0 if we set Γ free). Sources with very flat photon indices are often considered to be reflection-dominated CT candidates, and their extremely obscured nature may not be revealed by our best-fit $N_{\rm H}$ (Georgantopoulos et al. 2009). Moreover, since we are dealing with the stacked spectra here, sources that possessed large spectral variability may also exhibit untypical spectral shape, as might be the case for XID 249 (see Section 7.4). Additionally, we cannot rule out the possibility that the extreme photon indices of some sources are wrongly measured owing to insecure photometric redshifts.

4.3. Hardness Ratio versus Redshift

For highly obscured AGNs, the significant absorption and scattering of soft X-ray photons may lead to large hardness ratio (HR); thus, the HR may be used as an indicator of obscuration level (e.g., Wang et al. 2004). In Figure 7 we show the observed HR (here we adopt the definition of HR as $(H-S)/(H+S)$, where H stands for the observed-frame 2–7 keV count rate and S stands for the 0.5–2 keV count rate) as a function of redshift for highly obscured sources and CT candidates confirmed by the spectral fitting, as well as less obscured AGNs (best-fit $N_{\rm H} < 10^{23}$ cm^{-2}) that contaminate our sample. Heavily obscured sources have significantly larger HR than less obscured sources as expected. We also calculate the effective photon index $\Gamma_{\rm eff}$ (obtained by fitting the spectra using XSPEC model *phabs × zpow* with $N_{\rm H}$ fixed at the Galactic value) for the HOS. Ninety percent of them have $\Gamma_{\rm eff} < 1.0$, and the median $\Gamma_{\rm eff}$ is only -0.38 for CT AGNs, which again verify their heavily obscured nature.

The anticorrelation between HR and z is due to the fact that for high-redshift sources the observed soft band actually corresponds to a much harder band in the rest frame and the photons may not be absorbed significantly. Note that the HRs of the most heavily obscured CT AGNs are not always the largest at a given redshift. This can be ascribed to the additional soft-excess component softening the CT AGN spectra, and CN
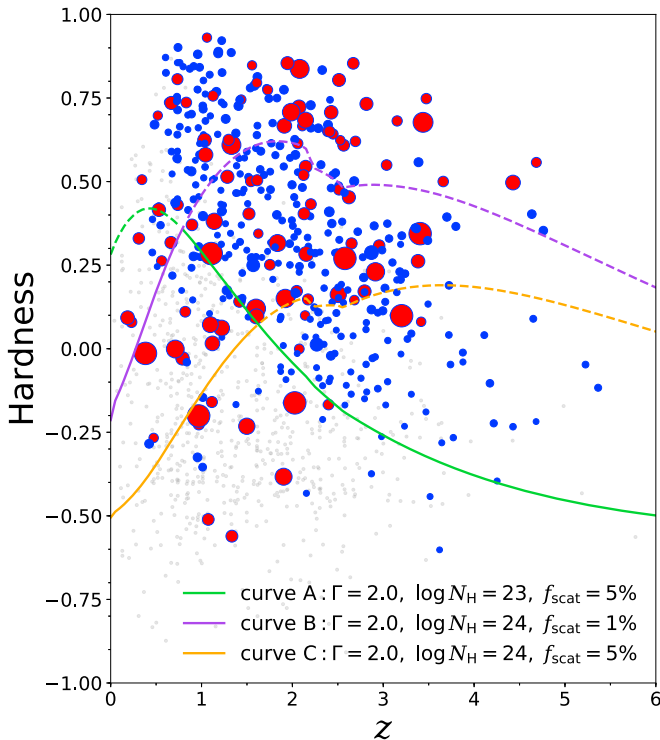
Figure 7. HR of the highly obscured sample as a function of source redshift. CT candidates selected in Section 4.5 with best-fit $N_H > 10^{24}$ cm^{-2} and the 1σ lower limit of $N_H > 5 \times 10^{23}$ cm^{-2} are shown in red, while other highly obscured AGNs are shown in blue. The size of the symbol indicates their best-fit N_H value. Less obscured sources with best-fit $N_H < 10^{23}$ cm^{-2} are shown in gray. Curves in different colors represent different selection curves presented in Section 4.3.

AGNs with intrinsically flat photon indices are also plausible to have larger HRs than CT AGNs.

To test whether a simple HR value can be used to select heavily obscured AGNs, we calculate the evolution of HR as a function of redshift for typical X-ray spectral parameters. Since our purpose is to find the critical HR for highly obscured AGNs as a function of redshift, we simply consider the threshold condition: a source with $N_H = 10^{23}$ cm^{-2}. Additionally, we set f_{ref} at 1.0, set the high-energy cutoff at 200 keV, and fix the photon index Γ of the two power laws at 2.0. The constant f_{exs} is set to 5%, which is typical for AGNs with $N_H \sim 10^{23}$ cm^{-2} (e.g., Brightman et al. 2014). The reason why we adopt $\Gamma = 2.0$ instead of the mean value 1.8 is that $\sim$96% of the sources with a well-constrained photon index in L17 have $\Gamma \lesssim 2.0$. Therefore, with this value we expect that our derived HR curve will be able to successfully select out most of the highly obscured AGNs. The model is shown in Figure 2(a), and the derived model HRs at different reshifts are shown in green in Figure 7 (curve A).

It is worth noting that not all CT AGNs have larger HRs than CN AGNs. We also show the HR curve derived for a CT AGN with $N_H = 10^{24}$ cm^{-2}. This time f_{exs} is chosen to be 1% owing to the fact that f_{exs} decreases with increasing N_H (see Section 5.2). The model is shown in Figure 2(b), and the simulated HR curve is also shown in Figure 7 (curve B). It is clear that at low redshifts, even a very small soft-excess fraction can easily dominate the soft X-ray spectra that could make CT AGNs even softer than CN AGNs. Therefore, to

avoid missing a large number of high-N_H sources at low redshifts, we use curve B at $z \leqslant 0.8$ and curve A at $z > 0.8$ as our selection curve in the following analysis (i.e., the combined curve).

For a source with given HR and redshift values, we consider it as a heavily obscured candidate if the observed HR lies above the combined HR curve. By applying this curve to our full sample, 480 sources will be selected as highly obscured candidates (hereafter the candidate sample). Eighty percent (382/480) of sources in the candidate sample are indeed highly obscured as confirmed from spectral fitting, i.e., the accuracy is 80%. The parameter distributions for the selected candidates and confirmed sources are shown in Figure 8 (top panel). Only 20% of the candidate sources are contaminated by less obscured sources, with most of them having N_H only slightly smaller than 10^{23} cm^{-2} and lying close to the boundary curve, as expected. The redshift, luminosity, and soft-excess fraction distributions for the selected candidates are very similar to sources that are confirmed to have $N_H > 10^{23}$ cm^{-2}.

In addition, we define the completeness fraction as the ratio of the number of highly obscured sources (confirmed) that lie above the selection curve to the total number of highly obscured sources (i.e., all red and blue points in Figure 7). The completeness fractions as a function of several parameters are shown in the second row in Figure 8. The average completeness fraction for the selection curve is 88% (382/436) and remains >90% for most high-redshift and high-luminosity bins, while at lower redshifts and larger f_{exs}, the completeness fraction significantly drops to <80%, since the soft excess dominates the observed spectrum, which makes highly obscured sources lie below the selection curve.

We also show the result after changing the soft-excess fraction for the model CT AGN to 5% (see curve C in Figure 7). This time the completeness fraction at low redshift is improved to 93%, but the accuracy dramatically drops to 68% since a large amount of less obscured AGNs are mistakenly included. As a trade-off, we choose to adopt the combined curve A and B as our selection curve. The functional form of this curve can be written as

$$
HR = \begin{cases}
-0.240z^3 + 0.230z^2 + 0.680z - 0.200, \\
\quad (z \leqslant 0.8); \\
-0.004z^3 + 0.077z^2 - 0.530z + 0.740, \\
\quad (z > 0.8).
\end{cases}
\tag{1}
$$

We note that the conclusions in the following sections remain largely unchanged if we use the HR-selected sample instead, suggesting that a simple HR value can be used to identify highly obscured AGNs and select a representative highly obscured sample as long as the redshift and soft-excess effects are properly taken into account while deriving the boundary curve.

4.4. Redshift and Luminosity Distributions

In Figure 9 we show the redshift distribution of the HOS in the bottom right panel. A total of 191 sources have spectroscopic redshifts, and 245 sources have photometric redshifts. The peak appears at $z \sim 1$–2.5, which is known as the peak epoch of both star formation and black hole accretion activities (e.g., Aird et al. 2015). The source amount slightly decreases down to $z \sim 0.5$ and significantly drops at $z < 0.5$. Since
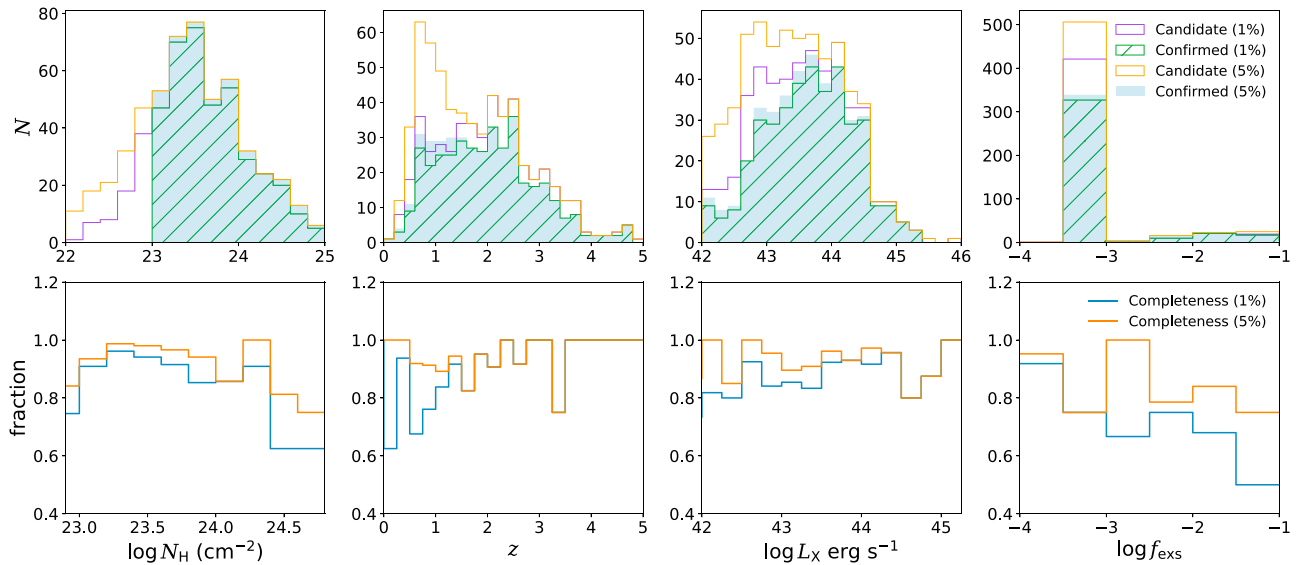
Figure 8. Top: distributions of source properties of the candidate samples selected from different HR curves (candidate) and those confirmed to be heavily obscured through spectral fitting (confirmed). Bottom: completeness fractions of the candidate samples selected by different HR curves. The 1% and 5% in the legend represent different soft-excess fractions assumed while deriving the HR curves in Figure 7. Sources without a detected soft-excess component are shown as $\log f_{\mathrm{exs}} = -4$ in the plot.

Chandra only pointed at small patches of sky in these two surveys, the small number of sources detected at $z < 0.5$ is mainly due to the small volume probed and the insufficient penetrability of the corresponding rest-frame energies at low redshifts; thus, the detection probability is severely limited in finding nearby highly obscured AGNs. The downward trend toward high redshifts is primarily due to the flux limits and may be partly due to the fact that photoelectric absorption in the soft band of mildly obscured ($N_{\mathrm{H}} \sim 10^{23}$ cm^{-2}) sources shifts out of the *Chandra* soft bandpass at $z \sim 2.5$, making it more difficult to constrain N_{H} (Vito et al. 2018).

The L_{X} distribution for the HOS is displayed in the top left panel of Figure 9. The $\log L_{\mathrm{X}}$ (L_{X} in units of erg s^{-1}) peaks in the range of 43.5–44.0, with a mean value of 43.65 ± 0.03. The luminous sources with $\log L_{\mathrm{X}} > 44.0$ (44.5) account for 32% (11%) of the whole sample. CT AGNs, in particular, have even higher luminosities, with a mean value of 44.10 ± 0.06, because the minimum detectable L_{X} for a flux-limited survey significantly increases with increasing N_{H} (see Figure 16). The correlations among L_{X}, N_{H}, and redshift are also plotted in Figure 9. The well-known Malmquist bias can be clearly seen in the lower left L_{X}–z corner, and we will discuss the influence of this bias on our results in Section 6.4.

4.5. Observed N_{H} Distribution

The observed distributions of the best-fit N_{H} in six luminosity bins of the HOS are shown in Figure 10. It appears that the amount of high-NH sources in lower-Lx bins is significantly lower than that in the higher-Lx bins, due to the fact that the minimum detectable luminosity significantly increases with NH (see Section 6.4). We show the normalized cumulative N_{H} distributions in Figure 11(a). Sources with best-fit $N_{\mathrm{H}} > 1.0 \times 10^{24}$ cm^{-2} account for ~25% (108/436) of the HOS. The distribution obtained in Brightman et al. (2014), which utilized the 4 Ms CDF-S data, is shown in the same plot for comparison. The K-S tests imply that the N_{H} distributions

from the CDF-S and the CDF-N are consistent with p-value $= 0.96$ but are different from that in Brightman et al. (2014) with p-value $\ll 0.001$, due to the fact that we identify more sources between $\log N_{\mathrm{H}} = 23.5$ and 24.0 cm^{-2}, while the distribution in Brightman et al. (2014) peaks at $\log N_{\mathrm{H}} = 23.0$–$23.5$ cm^{-2}.

Although we have significantly improved the photon statistics for each source using the deepest data, we should caution that most of the CT AGNs discovered still have counts <200 and their N_{H} might be poorly constrained. Therefore, we adopt a more conservative criterion that a source will be regarded as a CT candidate only if it has the best-fit $N_{\mathrm{H}} > 10^{24}$ cm^{-2} and the 1σ lower limit on N_{H} is constrained to be greater than 5×10^{23} cm^{-2}. This criterion selects 102 sources, with 66 from CDF-S and 36 from CDF-N, accounting for ~23% of the HOS, to be CT candidates. Note that there are some well-studied CT candidates in the literature that are not classified as CT AGNs in our sample owing to their best-fit $N_{\mathrm{H}} < 10^{24}$ cm^{-2}, such as XID 328, XID 551 (XID 153 and XID 202 in Comastri et al. 2011, respectively), XID 539 (XID 403 in Gilli et al. 2011), and XID 375 (*BzK*8608 in Feruglio et al. 2011) in the CDF-S. However, all these sources have best-fit $N_{\mathrm{H}} > 8 \times 10^{23}$ cm^{-2}, which shows good consistency with previous studies; thus, we confirm their heavily obscured nature using deeper *Chandra* data.

We note that at $f_{2-10\,\mathrm{keV}} > 10^{-15}$ erg cm^{-2} s^{-1}, the number of our CT AGNs in the CDF-N (22 sources) is higher than that found by Georgantopoulos et al. (2009), which presented 10 CT candidates in the same field. This may be because Georgantopoulos et al. (2009) mainly focused on searching reflection-dominated CT AGNs and possibly missed some transmission-dominated sources. We also fit the 22 CT candidates using the BNtorus model (Brightman & Nandra 2011) and find that 20 sources have best-fit $N_{\mathrm{H}} > 10^{24}$ cm^{-2} and $f_{2-10\,\mathrm{keV}} > 10^{-15}$ erg cm^{-2} s^{-1}, which confirms the MYTorus result.
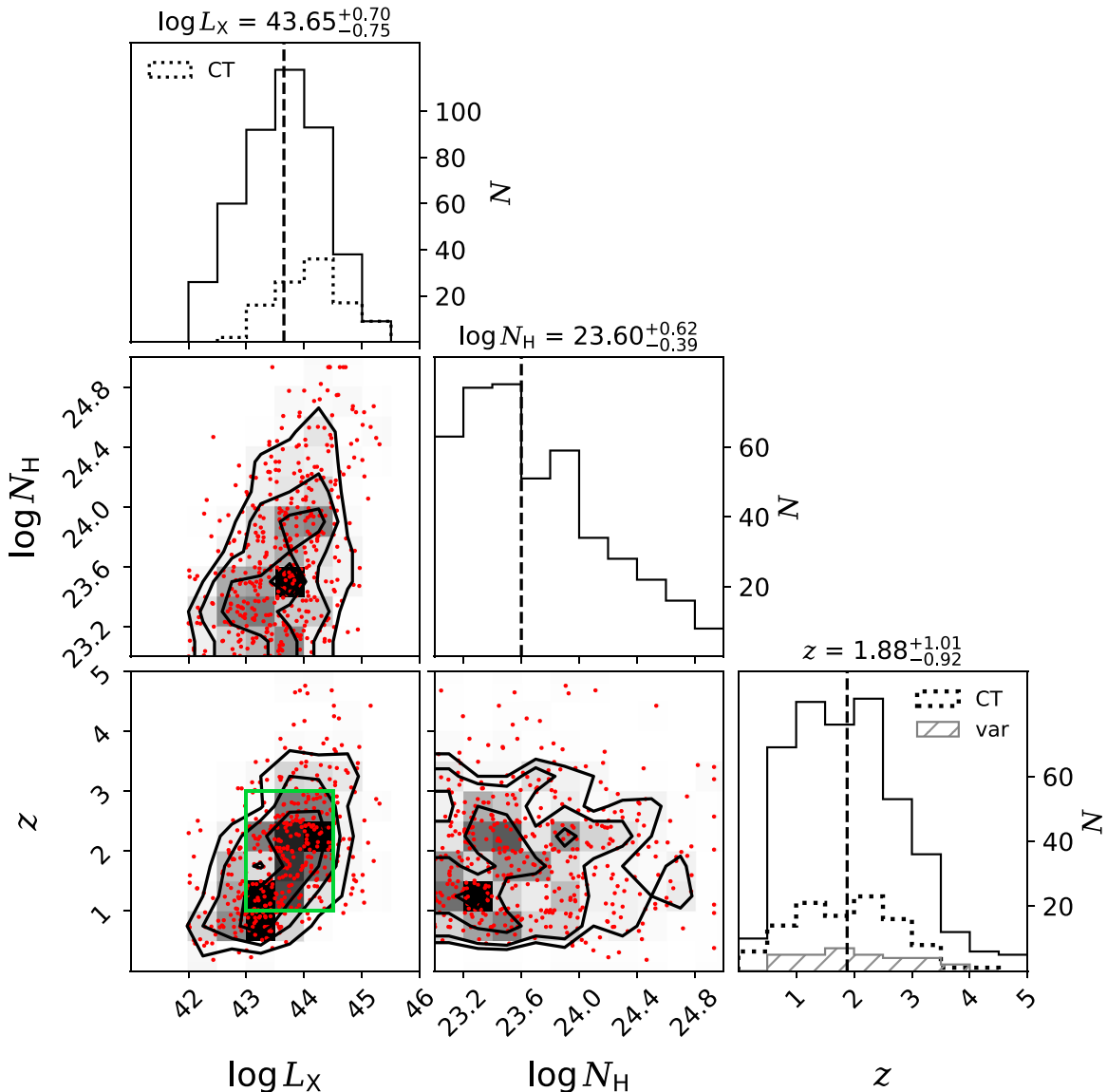
Figure 9. Triangle plot of the correlations among L_X, N_H, and redshift. The outer parts of the triangle show the distributions of L_X, N_H and redshift from the top left to the bottom right panels, respectively. The vertical dashed lines show the median value for each distribution. In the redshift and L_X distribution panels, the solid histograms represent the total sample, and the dotted ones show the distributions of CT candidates selected in Section 4.5. The shaded histogram shows the redshift distribution of our variability sample in Section 7. The inner parts of the triangle show the scatter plots and corresponding density maps overlaid with contours among L_X, N_H, and redshift, respectively. In the L_X vs. z plane, the green rectangle shows the subsample we select in Section 6.5 to investigate the evolution of the intrinsic fraction of CT AGNs among highly obscured ones.

4.6. Number Counts for CT AGNs

We calculate the observed number counts ($\log N - \log S$) for CT candidates selected in Section 4.5. The observed $\log N - \log S$ is defined as the observed source amount divided by survey area. However, as we will discuss in Section 6.2, the sky coverage is not a constant value across the whole flux and N_H ranges; hence, we assign a weighting factor $1/\omega_{area}$ (see Section 6.2 for details) for each source while calculating their cumulative number counts. The corrected results are shown in Figure 11(b), and several number count measurements presented in previous works in the 4 Ms CDF-S (Brightman & Ueda 2012), X-UDS (Kocevski et al. 2018), and COSMOS (Lanzuisi et al. 2018) fields are also displayed for comparison. Note that Brightman & Ueda (2012) reported the result in the 0.5–8 keV band and we convert the 0.5–8 keV flux into

2–10 keV by assuming a $\Gamma = -0.4$ power law, which is the median effective photon index for our CT AGNs.

The corrected number counts for the CDF-S and CDF-N are generally consistent within the error bars at higher fluxes, despite that we obtain six CT AGNs in the CDF-N but only three in the CDF-S at $f_{2-10\,keV} > 4 \times 10^{-15}$ erg cm^{-2} s^{-1}, possibly caused by the low number statistics and the cosmic variance due to the small sky coverage of CDFs (e.g., Harrison et al. 2012). At the faint end, both CDF-S and CDF-N number counts flatten owing to the limited detection ability, and the CDF-N number counts are significantly smaller than the CDF-S, because the three times shallower exposure limits its detectability of faint CT sources (see Figure 11(c)). Our results are also in agreement with the 4 Ms CDF-S result at the faint end, the COSMOS result at the bright end, and the X-UDS result at moderate fluxes.
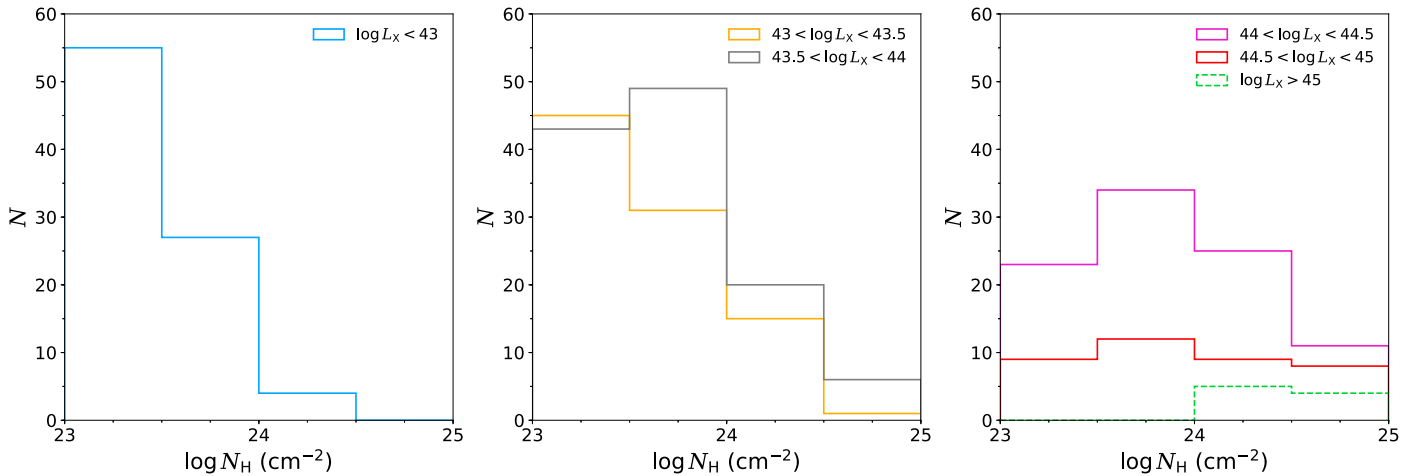
Figure 10. Observed N_H distributions in six L_X bins. The small number of high-N_H sources in low-L_X bins is due to the fact that such sources are difficult to detect. This bias will be corrected in Section 6.4.

We also compare the observed $\log N - \log S$ with several CXB model predictions in $z = 0-5$ (Gilli et al. 2007; Ballantyne et al. 2011; Akylas et al. 2012; Ueda et al. 2014). For the Gilli et al. (2007),[20] Akylas et al. (2012),[21] and Ueda et al. (2014)[22] CXB models, the luminosity range used to derive the predicted number counts is $\log L_X = 42-45.5$ erg s^{-1}, similar to our sample; for the Ballantyne et al. (2011) model, the predicted result is presented in $\log L_X = 41.5-48$ erg s^{-1}. As can be seen from Figure 11(b), our observed $\log N - \log S$ prefers the moderate CT number counts as predicted by Akylas et al. (2012) and Ueda et al. (2014), while other models more or less overestimate or underestimate the number counts.

In summary, the observed parameter distributions and relationships presented in this section provide a basic description of the highly obscured AGN population and are crucial for distinguishing various CXB models. However, these observed distributions, in particular, the observed N_H distribution, are influenced by several biases. We will discuss more details about these biases and reconstruct the intrinsic N_H distribution representative for the highly obscured AGN population in Section 6.

5. The Properties of Obscuring Materials

5.1. X-Ray Highly Obscured Broad-line AGNs

We collect optical classification results for our CDF-S sample by cross-matching their optical/near-IR/IR/radio counterpart positions presented in the X-ray source catalogs (CP_RA and CP_DEC) with the Silverman et al. (2010) E-CDF-S optical spectroscopic catalog using a 0″.5 matching radius. Among the 55 matched sources, 31 sources have been classified. To our surprise, 19% (6/31) of them are labeled as broad-line AGN (BLAGN). Detailed information on these sources is given in Table 5.

Most of the six sources have sufficient counts and reliable redshift measurements to constrain N_H. Three sources have insecure spectroscopic redshifts, although they are labeled as

BLAGN, and the 7 Ms main catalog chooses to adopt the photometric redshift as ZFINAL for two of them. This could be due to the fact that they have low-S/N optical spectra or only one emission line is available for sources within particular redshift ranges, which makes it difficult to conclusively determine the line nature, thus giving an insecure redshift. For three sources with insecure redshifts, we set redshift as a free parameter in the spectral fitting and obtain the best-fit X-ray redshift and corresponding N_H. All of them are still best fitted by $N_H > 10^{23}$ cm^{-2}. The redshift range in which the source will remain X-ray highly obscured is also listed. In general, the N_H estimates for the six sources are robust. This result suggests that the heavy X-ray absorption in a fraction of sources is largely an LOS effect caused by some compact clumpy clouds obscuring the central X-ray-emitting region, but the global CF for the high-N_H materials is limited, and thus the BLR is not blocked; alternativley, the heavy X-ray obscuration may be produced by the BLR itself.

5.2. Soft-excess Fraction Dependences

There are 80 sources in our sample that require a second power law to fit the soft-excess component. The origin of this component is still a puzzle. Many different models, e.g., warm Comptonization or blurred ionized reflection from the disk, partially ionized absorption in a wind, or power-law continuum in other directions scattered into the LOS from large-scale CN matter, are proposed to explain the diverse situations found in different sources (e.g., Boissay et al. 2016). However, for highly obscured AGNs, the scattered mechanism is preferred since either blurred reflection or warm Comptonization from the accretion disk will be significantly attenuated by the obscuring materials.

We show the f_{exs}, which represents the relative normalization of the soft component with respect to the intrinsic power law, as a function of N_H in Figure 12. We find a significant anticorrelation between the two parameters: Spearman's $\rho = -0.66$, p-value $= \ll 0.001$. The mean f_{exs} is 3.7% for the total sample. The scattered fraction decreases from 6.1% to 4.3% to 1.3% for $N_H < 5 \times 10^{23}$ cm^{-2}, 5×10^{23} cm$^{-2} < N_H < 1.5 \times 10^{24}$ cm^{-2}, and $N_H > 1.5 \times 10^{24}$ cm^{-2}, respectively.

To verify that this anticorrelation is not caused by the parameter degeneracy that sources with a large soft-excess fraction and a high N_H might be misclassified as low N_H, low

Table 5
Information of X-Ray Highly Obscured BLAGNs in the CDF-S

XID (1)	Model (2)	f_{ref} (3)	Oclass (4)	z (5)	ztype (6)	zqual (7)	counts (8)	$\log N_H$ (cm^{-2}) (9)	z-range (10)	z_x (11)	$\log N_{H,x}$ (cm^{-2}) (12)
882	A	$\ll 1$	BLAGN	3.19	zspec	Secure	80	23.1	⋯	⋯	⋯
944	C	1	BLAGN	0.96	zphot	Insecure	69	23.9	all	2.00	24.1
16	A	$\ll 1$	BLAGN	3.10	zphot	Insecure	409	23.7	$z > 1.4$	1.66	23.2
399	C	1	BLAGN	1.73	zspec	Secure	1077	23.2	⋯	⋯	⋯
968	C	1	BLAGN	2.03	zspec	Secure	500	23.8	⋯	⋯	⋯
977	C	1	BLAGN	4.64	zspec	Insecure	225	23.9	$z > 1.2$	3.25	23.9

Note. Column (4): optical classification result from Silverman et al. (2010). Thirty-one X-ray highly obscured AGNs in our CDF-S sample have optical classification results, and six of them are identified as BLAGNs. Column (10): redshift range for sources with insecure redshifts to be determined as being X-ray highly obscured. Column (11): best-fit X-ray redshift by setting redshift as a free parameter during spectral fitting. Column (12): best-fit $\log N_H$ when adopting the X-ray best-fit redshift. Other columns have the same meaning as in Table 2.
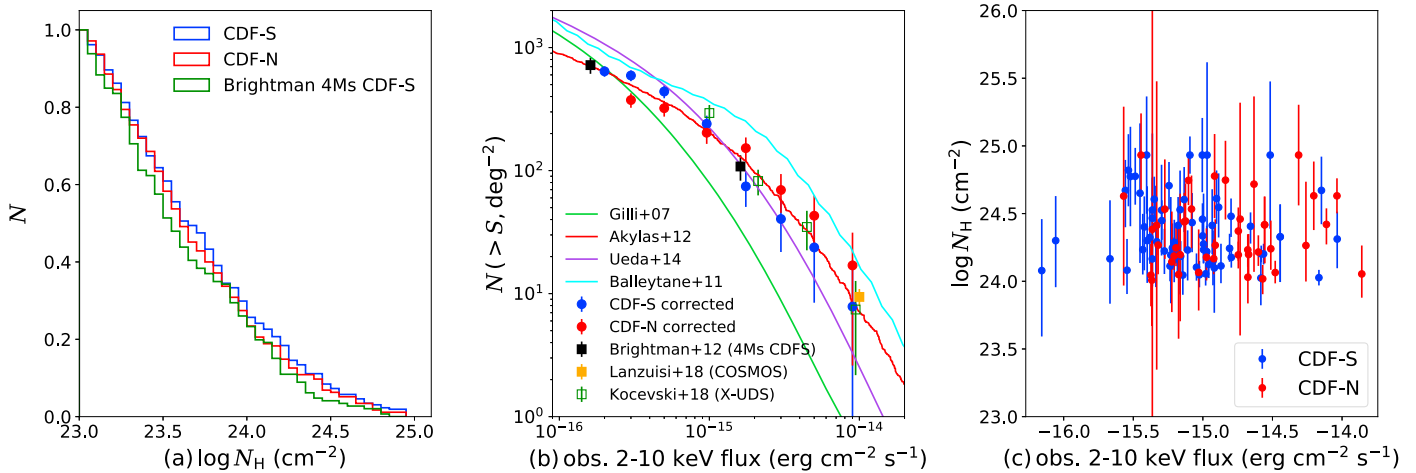

Figure 11. Left: normalized cumulative N_H distribution for our highly obscured sample. The 4 Ms CDF-S result adopted from Brightman et al. (2014) is also plotted for comparison. Middle: observed $\log N - \log S$ for CT candidates compared with previous number counts in the 4 Ms CDF-S (Brightman & Ueda 2012), X-UDS (Kocevski et al. 2018), and COSMOS (Lanzuisi et al. 2018), as well as several CXB model predictions (Gilli et al. 2007; Ballantyne et al. 2011; Akylas et al. 2012; Ueda et al. 2014). Right: column density of CT AGNs as a function of observed 2–10 keV flux in the CDF-S (blue) and CDF-N (red), respectively.
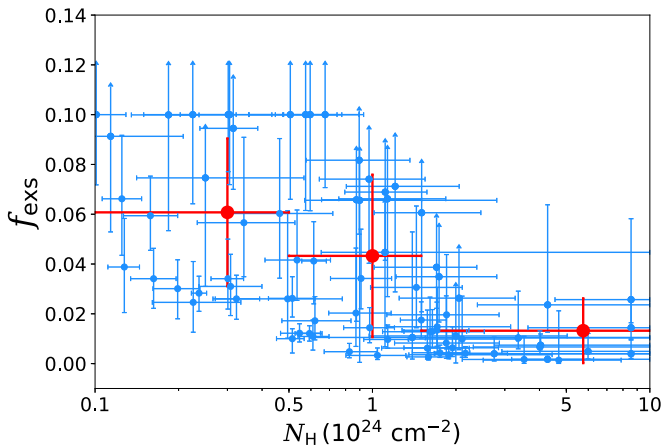

Figure 12. Soft-excess fraction f_{exs} as a function of N_H. The blue circles show the individual sources, and the red circles show the binned results. The negative correlation between f_{exs} and N_H indicates that a portion of high-N_H sources might have higher "torus" CFs, which makes it hard for the soft-excess photons to escape, under the assumption that the excess originates from the backscattered continuum in highly obscured AGNs.

f_{exs}, and low intrinsic luminosity sources owing to the low S/N, we assume that CT AGNs have exactly the same f_{exs} as CN AGNs (here we simply choose $f_{exs} = 5\%$). Then, we generate fake spectra using model D for a sample of simulated sources,

which have redshifts and N_H uniformly distributed between $z = 0$ and 4 with an interval of $\Delta z = 0.5$ and 23 cm^{-2} < $\log N_H < 25$ cm^{-2} with an interval of $\Delta \log N_H = 0.25$. For each source we simulate 100 spectra with 0.5–7 keV counts of ~50, 100, and 150, respectively. We then fit the fake spectra and perform a Spearman rank correlation test for the output N_H and f_{exs}. The Spearman's $\rho \approx 0.1$ (insensitive to the simulated spectral counts), indicating no apparent correlation; thus, the observed anticorrelation is intrinsic.

By assuming that the second power law originates from the backscattered continuum, Brightman & Ueda (2012) found that the scattered fraction depends on the opening angle, i.e., sources with small opening angles have lower f_{exs}. Moreover, Ueda et al. (2015) and Kawamuro et al. (2016) found that the ratios of [O III] and [O IV] to hard X-ray luminosity are lower in low scattering fraction sources, inferring that the low-f_{exs} AGNs are buried in small opening angle torus. Therefore, the anticorrelation between f_{exs} and N_H indicates that high-N_H sources might preferentially reside in high-CF tori. This is also consistent with the results from Brightman & Ueda (2012) and Lanzuisi et al. (2015), who found a similar anticorrelation between f_{exs} and N_H and explained it as highly obscured AGNs being also heavily geometrically obscured (but see the discussion in the next section).
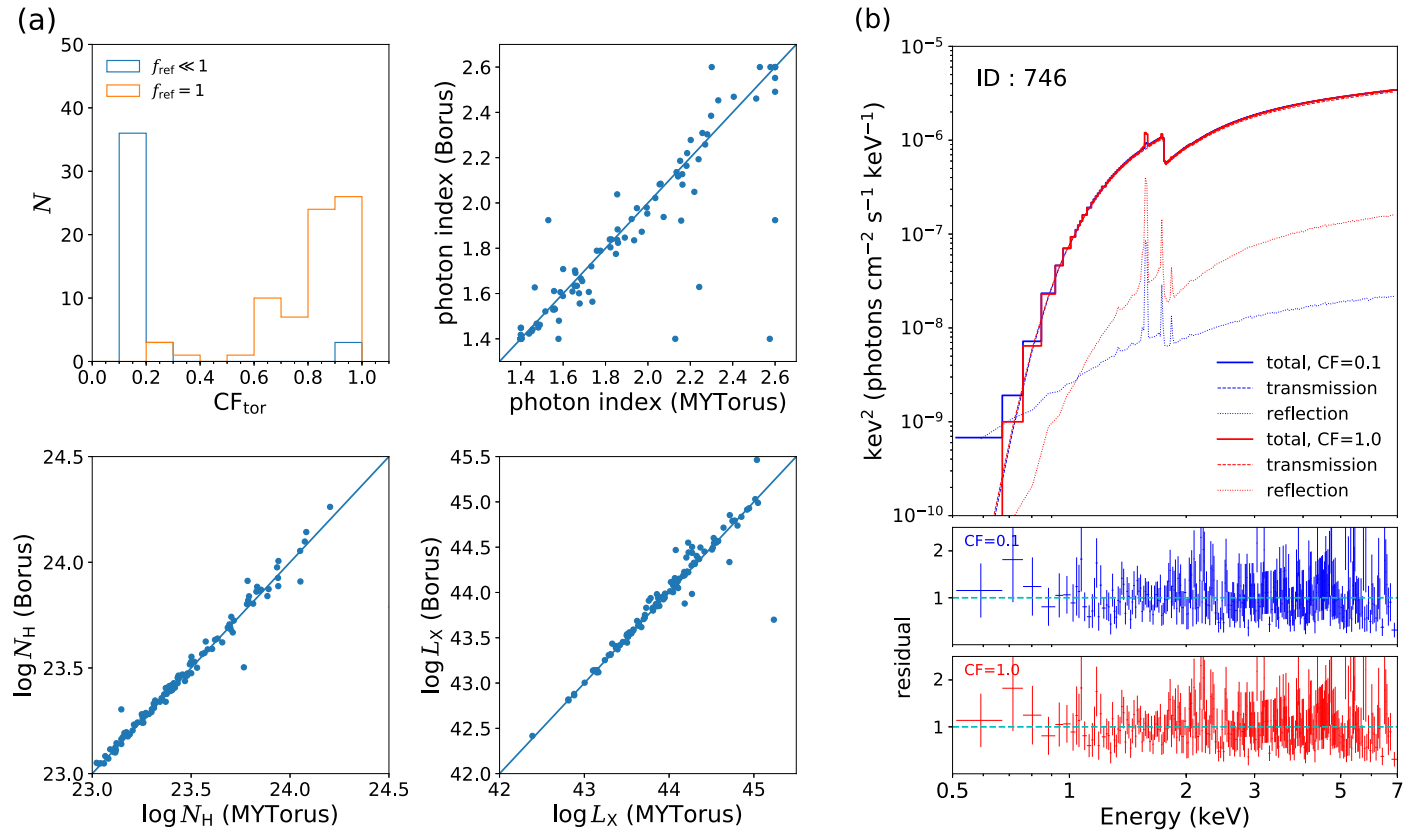
Figure 13. (a) CF derived through the Borus model (Baloković et al. 2018) and the comparisons among photon index, N_H, and L_X obtained from the Borus and MYtorus models. The photon indices in this plot are the original results without fixing Γ at 1.8 for some extreme cases. (b) Spectral fitting result for CDF-S XID 746 using the Borus model. Both the fully buried model (CF$_{tor}$ = 1.0; Cstat = 942.3) and the weak torus model (CF$_{tor}$ = 0.1; Cstat = 941.8) can well reproduce the data. The contributions from the reprocessed components to the total spectrum in both models are largely weak.

5.3. Reprocessed Components and CF

Unlike local CT AGNs in which the prominent neutral Fe lines and reflection hump (hereafter the reprocessed components) are prevalently detected in the high-quality hard X-ray spectra and can serve as unambiguous signatures of being CT (e.g., Tanimoto et al. 2018; Marchesi et al. 2019; Zhao et al. 2019), 41% of our highly obscured sources and 38% of the CT candidates have negligible f_{ref}, even for some high-count sources or CT AGNs. (see Table 3). It should be noted that the low detection rate of reprocessed components is not in conflict with their highly obscured nature. The poor S/N for low-count sources and the narrow *Chandra* spectral coverage make it challenging to detect narrow iron lines and to distinguish the transmission and reflection components, which may cause a significant overestimation of the fractions.

But we note that it is possible that highly obscured AGNs may have weak reprocessed components if the global CFs of high-N_H materials are low or the central AGN is geometrically fully buried by CT materials such that even the reflected photons cannot escape (e.g., Brightman et al. 2014). To distinguish different scenarios, we use the Borus model (Baloković et al. 2018),[23] which allows CF to vary freely, to fit the spectra for all sources with counts >200. This time we treat photon index as a free parameter in order to make a direct comparison between the two models. The distribution of the

derived CF$_{tor}$, which is defined as the cosine value of the opening angle θ_{tor} measured from the symmetry axis toward the equatorial plane, is shown in Figure 13(a). We also show the comparisons of the main spectral fitting parameters with those obtained from the MYTorus model and find good consistency.

For sources with negligible reprocessed components in the MYTorus model ($f_{ref} \ll 1$), the Borus results are consistent with the weak torus scenario (CF$_{tor} \approx 0.1$), which might suggest that their heavy obscuration is simply an LOS effect (i.e., some high column density clouds on various scales along our sight line may obscure the compact X-ray emitter), without the necessity to invoke a strongly buried nuclear environment. In contrast, sources with $f_{ref} = 1.0$ are mostly best fitted by a highly covered model (0.8 < CF$_{tor}$ < 1.0) that has torus covering angles ($90° - \theta_{tor}$) between $65°$ and $90°$. We emphasize that although the LOS and the global torus N_H are linked in the spectral fitting, those fully buried sources do not need to be covered by very high N_H materials in all directions, since the best-fit N_H will more likely converge to the values determined by the LOS component because the photoelectric absorption is a much stronger spectral feature. Moreover, we find that the average CF$_{tor}$ for $f_{exs} < 0.05$ sources is 0.60, while for $f_{exs} > 0.05$ sources the $\overline{CF_{tor}}$ is 0.32. This might provide evidence that the lower soft-excess fraction in some sources is caused by a geometrically more buried structure (see Section 5.2).

However, we caution here that the current S/N and spectral coverage do not allow us to constrain the CF from X-ray

---

[23] In the Borus model, CF$_{tor}$ = 1.0 corresponds to a fully covered torus, while CF$_{tor}$ = 0.1 represents a typical disk-like covering (see http://www.astro.caltech.edu/~mislavb/download/).
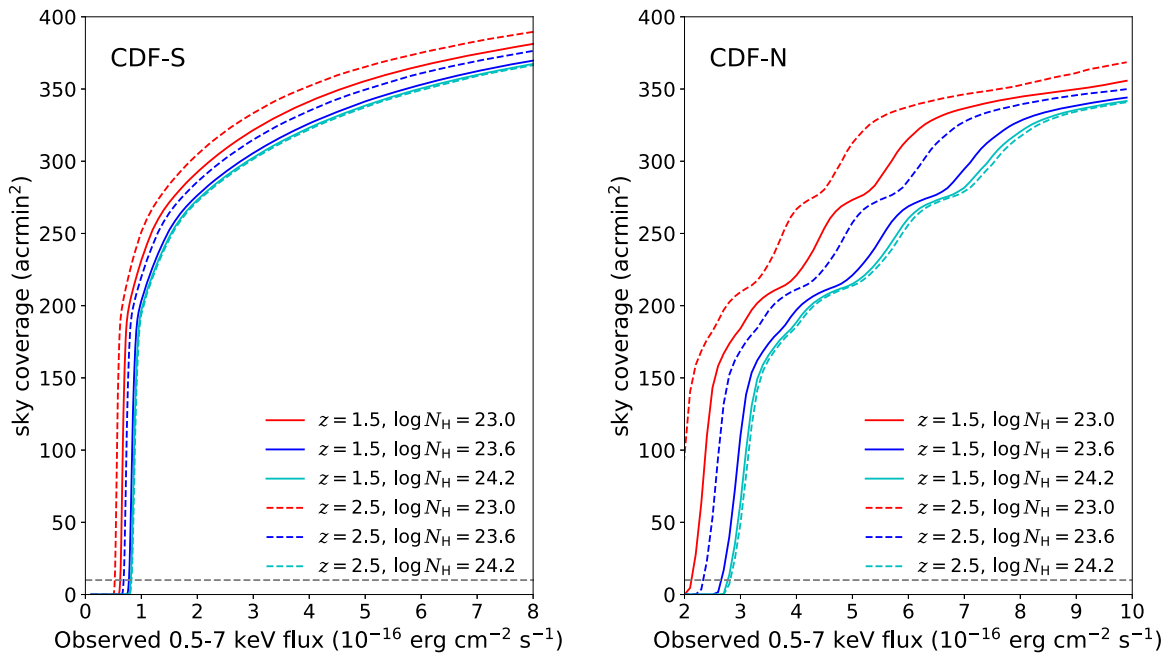
Figure 14. Sky coverage as a function of the observed 0.5–7 keV flux corresponding to the count cut 20 in the CDF-S and CDF-N, shown for three different $N_{\rm H}$ values at two different redshifts, respectively.

spectral analysis. If we fix $CF_{\rm tor}$ at 0.1 for sources that are best fitted by a fully buried model, we may still obtain acceptable fits with Cstat values only slightly increasing (an example is shown in Figure 13(b)). Conclusively disentangling several scenarios is beyond the scope of this work, and multiwavelength data are needed to further shed light on the nature of the obscuring materials (J. Y. Li et al. 2019, in preparation).

6. Intrinsic $N_{\rm H}$ Distribution for Highly Obscured AGNs

In Section 4.5, we present the observed $N_{\rm H}$ distribution, which is affected by several sample incompletenesses. In order to derive the intrinsic $N_{\rm H}$ distribution, we should take into account the errors on best-fit $N_{\rm H}$ and correct for several survey biases (Liu et al. 2017). To perform such corrections, we restrict our analysis to the 394 sources with off-axis angle $<10\!'\!.0$, for the sake of avoiding large background contamination, extremely small sky coverage (see Section 6.2), and limited detectable fraction (see Section 6.4).

6.1. Errors on Best-fit $N_{\rm H}$

To consider the uncertainty of the spectral fitting, we perform a resampling procedure to the best-fit $N_{\rm H}$. Given the asymmetric errors, we assume that the errors on $N_{\rm H}$ obey the "half-Gaussian" distribution. We first generate 1000 $N_{\rm H}$ for each source with the σ of the half-Gaussian distribution equal to the lower 1σ error, and then we generate another 1000 $N_{\rm H}$ with σ equal to the upper 1σ error. The mean value μ is set to the best-fit $N_{\rm H}$. The resampled $N_{\rm H}$ distribution (calculated by averaging the 2000 resampled distributions and also corrected for the sky coverage effect; see Section 6.2) is shown in the shaded region of Figure 18. It has an extended tail down to $N_{\rm H} < 10^{23}$ cm^{-2}, which contains only 4.0% of the resampled data, suggesting that even if considering the spectral fitting errors, most of our sources are still consistent with being highly obscured.

6.2. Sky Coverage Effect

Due to *Chandra*'s instrument features, the point-source PSF size increases and the effective exposure time dramatically decreases toward large off-axis angles. The sensitivities of detecting faint sources reduce prominently at the outskirt of the field, leading to small sky coverage while the observed flux is low. Therefore, a correction must be made.

We calculate the energy-flux-to-count-rate conversion factor (ECF) by assuming the soft-excess model (model D) with Γ, $f_{\rm ref}$, and $f_{\rm exs}$ fixed at 1.8%, 1.0%, and 1%, respectively. Combining ECF and the exposure map, we build a sensitivity map that represents the flux limits corresponding to the 20-count cut. Using this map, we calculate the sky coverage as a function of observed 0.5–7 keV flux, $N_{\rm H}$, and redshift as shown in Figure 14. Then, we measure the sky coverage for each source based on their observed flux, $N_{\rm H}$, and redshift obtained from spectral fitting. We define $\omega_{\rm area}$ as the ratio between the source sky coverage and the maximum sky coverage of the two *Chandra* surveys (484.2 and 447.5 arcmin^2 in the CDF-S and the CDF-N, respectively). To correct for the sky coverage effect, we simply weigh each source by $1/\omega_{\rm area}$ while resampling the observed $N_{\rm H}$ distribution. To avoid extremely large weights, we apply another cut for the 22 sources with sky coverage less than 100 arcmin^2, setting their weighting factors to the median factor 1.3, since these sources lie close to the count cut and we cannot rule out the possibility that their extremely small sky coverage may be a result of inappropriate spectral modeling.

6.3. Eddington Bias

Considering the measurement error of net counts, sources with intrinsically low counts may exceed our count cut (20 photons), and sources with high counts may be missed in our sample. Since the numbers of faint and bright sources are not equal, this error leads to the Eddington bias. To correct for this bias, we convolve the cumulative count distribution for AGNs
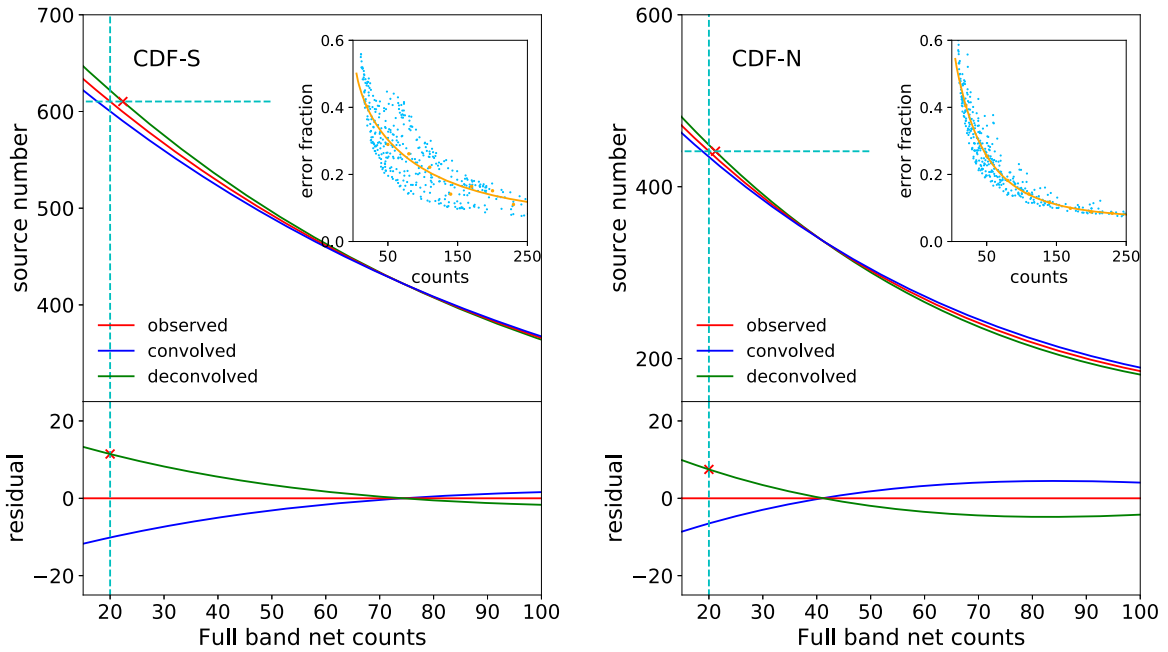
Figure 15. Illustration of the Eddington bias in the CDF-S (left) and CDF-N (right). In each panel/field, the red curve shows the smoothed cumulative count distribution for all AGNs with counts >20 and off-axis angle <10′. The blue curve is obtained by Poisson-sampling the observed count distribution by considering the measurement errors. For sources without count error measurements in the catalogs, we fit the mean error fraction f_{err} (f_{err} = count$_{error}$/count) in a given count range (shown as orange points) as a function of full-band net counts using sources with errors; the fitted curve is shown in the inset. Then, we simply assign an error to sources without errors, which is given by their total net counts times the corresponding f_{err}. We apply a pseudo-deconvolved method to obtain the deconvolved intrinsic count distribution by shifting the blue curve leftward or rightward (based on the location on the right or left side of the node of the two curves) with the value equal to the displacement between the blue and red curves, as shown by the green curve. From the residual between the observed and deconvolved curves, we can see that we missed 11.6 and 6.5 sources (shown as red crosses in the bottom panels), which correspond to the effective count cuts of 22.4 and 22.1 (shown as red crosses in the top panels) in the CDF-S and CDF-N, respectively.

in the two *Chandra* survey catalogs with the count errors using Poisson sampling in order to obtain the intrinsic count distribution. We follow the "pseudo-deconvolved" method proposed in L17 (see Section 5.1.3 of L17 for details), by shifting the observed curve leftward with the value equal to the displacement between the observed and convolved curves. We then obtain the deconvolved curve that represents the intrinsic cumulative count distribution.

The difference between the deconvolved and the observed curves can be treated as the number of sources missed or misincluded (depending on the adopted count cut) in our sample. As shown in Figure 15, at a 20-photon count cut, we miss 11.6 and 6.5 sources in the CDF-S and CDF-N, respectively. This is due to the fact that although sources with counts <20 may outnumber the bright sources, such faint sources are not detected in the source catalogs. By controlling the sample size (i.e., the cumulative source number in the y-axes) to be the same, we obtain the effective net counts of 22.4 and 21.1 for the CDF-S and CDF-N, respectively. Thus, the 20-photon count cut will be replaced by the new effective count cut while performing other corrections.

6.4. Malmquist Bias

So far we have only obtained the corrected observed N_H distribution for our sample. In order to derive the intrinsic N_H distribution for the highly obscured AGN population, we should also consider the undetectable part of the *Chandra* survey, which is known as the Malmquist bias that sources with lower luminosities will be missed by a flux-limited survey. More specifically, given the count rate limit of the survey (the

effective count cut in our situation), whether a source is detectable or not depends on its redshift, spectral shape, column density, intrinsic luminosity, and position in the field, since large off-axis angles lead to significantly lower sensitivities.

To quantify this bias, we generate fake X-ray spectra to find out the minimum L_X (L_{cut}, corresponding to our effective count cut) in different redshift and N_H bins, which can be detected at a specific position in the field. We divide the two CDFs into 10 annuli, with each corresponding to an interval of 1′. While running XSPEC simulations, we use the averaged exposure time, rmf file, and arf file calculated from all sources in each annulus. Model D is used in the simulation with the parameters set the same as in Section 6.2.

We show some of the detectable boundary curves in Figure 16. We define the detectable region as the area between the boundary curve and the maximum log $L_X = 46$ erg s^{-1}. As we can clearly see, sources with high N_H and large off-axis angles have significantly small detectable regions. We follow Equation (3) in T06:

$$F(N_H)dN_H = f(N_H)dN_H \int_0^{z_{max}} \frac{dV}{dz} dz$$
$$\times \int_{L_{cut(N_H,z)}}^{L_{max}} N(L_X, z)\, dL_X, \qquad (2)$$

where $f(N_H)$ represents the intrinsic N_H distribution, $F(N_H)$ represents the observed N_H distribution, and $N(L_X, z)$ is the X-ray LF. The detectable source fraction for each N_H bin relative to the log $N_H = 23$ cm^{-2} bin in a given redshift interval
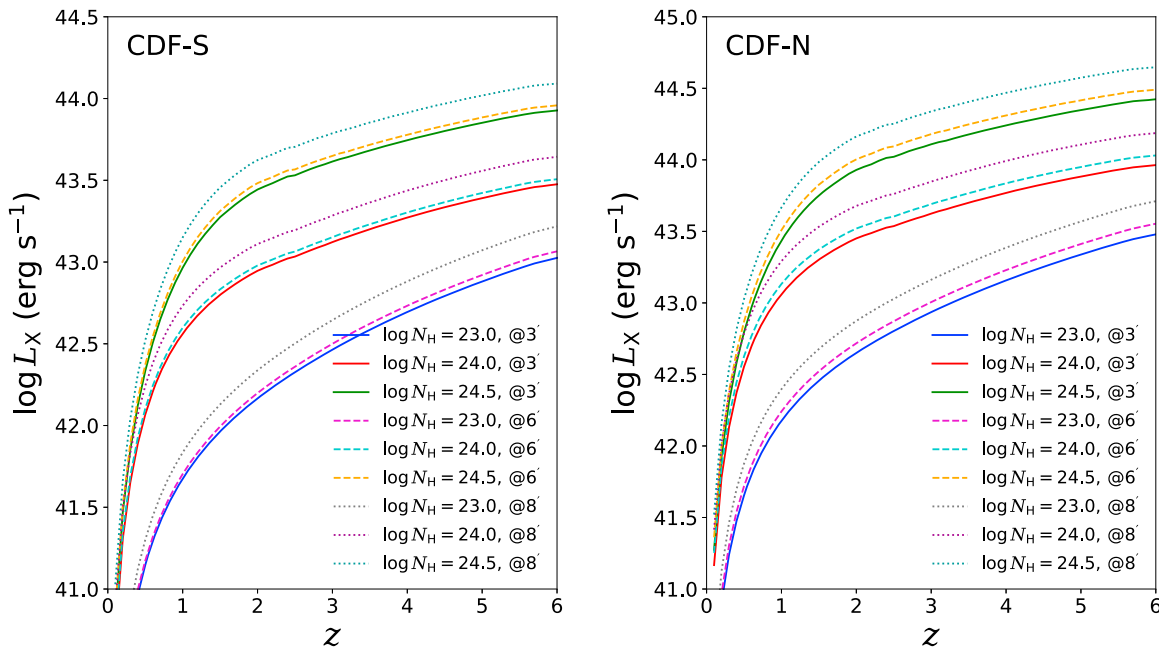
Figure 16. Detection boundary curves as a function of redshift for different N_H and off-axis angles, corresponding to the effective count cut 22.4 in the CDF-S and 21.1 in the CDF-N.

can be calculated as

$$f = \frac{\int_{L_{\text{cut}(N_H,z)}}^{L_{\text{max}}} N(L_X, z)\, dL_X}{\int_{L_{\text{cut}(23,z)}}^{L_{\text{max}}} N(L_X, z)\, dL_X}. \tag{3}$$

Thus, for the bin we choose

$$f = \begin{cases} 1, & \text{if } 23.0 \text{ cm}^{-2} \leqslant \log N_H < 23.25 \text{ cm}^{-2}, \\ <1, & \text{if } \log N_H \geqslant 23.25 \text{ cm}^{-2}. \end{cases} \tag{4}$$

Note that T06 derived the above Equation (2) by assuming that $f(N_H, L_X, z)$, which represents the possibility of detecting a source with column density in the range of N_H to $N_H + d N_H$ for a given L_X and z, varies slowly as L_X and z change. There is evidence that the obscured fraction of AGNs evolves with both luminosity and redshift. However, we do not consider such a complicated evolution and simply extract $f(N_H, L_X, z)$ from the integral. Therefore, by multiplying $1/f$ by the observed N_H distribution $F(N_H)\, dN_H$ in each bin, sources with different N_H are corrected to cover the same observable space with respect to the $\log N_H = 23$–23.25 cm^{-2} bin, and the Malmquist bias is corrected. To avoid the case of the corrected distribution being dominated by sources with extremely small detectable fractions, we apply a cut that while $f < 0.25$ the weighting factor $1/f$ will be simply set to $1/0.25$. Adopting a lower (0.2) or higher (0.3) detectable fraction cut will lead to $\sim 2\%$ systematic difference of the final intrinsic CT/highly obscured fraction.

Apparently, the correction depends on both spectral models and the detailed shape of X-ray LFs. Therefore, we compare the results using different X-ray LFs. We combine the $z < 3$ LF from Miyaji et al. (2015) and $z > 3$ LF from Georgakakis et al. (2015) as our first LF model, which is the same as in L17. As for the second model, we adopt Aird et al. (2015) LFs for obscured AGNs (22 cm^{-2} < N_H < 24 cm^{-2}). We do not use their LFs for CT AGNs, since they obtained a systematically

lower CT fraction than other works and our result in Section 4.6 prefers higher CT fractions (but see the discussion in Section 6.4 of Aird et al. 2015). The third LF we used is from Ueda et al. (2014). The three LF models are shown in Figure 17. We note that the adopted obscured AGN LFs may not be suitable for CT AGNs, but an extensive exploration of this issue is beyond the scope of this work.

6.5. Intrinsic N_H Distribution

Figure 18 shows the intrinsic N_H distributions in five redshift bins after correcting for all the aforementioned biases. We define the intrinsic CT/highly obscured source fraction f_{CH} as

$$f_{CH} = \frac{N(\log N_H \geqslant 24)}{N(\log N_H \geqslant 23)} \tag{5}$$

and list its evolution with redshift in Table 6.

As we can see, different LFs give consistent results. It is obvious that the corrected N_H distributions are quite different from the observed ones, especially in the CT regime. This highlights the importance of carefully correcting for the observational biases in order to uncover the underlying distributions. Though the intrinsic N_H distributions are different between high-redshift and low-redshift bins, the CT/highly obscured fraction is in accordance with no evident redshift evolution given the uncertainties (i.e., $f_{CH} \approx 0.52$ for all redshift bins; see Table 6). Note that Buchner et al. (2015) also claimed that the CT fraction is constant across cosmic time, but the CT fraction in their work is defined as the number of CT AGNs to the total AGN population. To avoid possible biases induced by the fact that different redshift bins are sensitive to different luminosity ranges, we also select two subsamples with 43.0 erg s^{-1} < $\log L_X$ < 44.5 erg s^{-1} at $z = 1$–2 and $z = 2$–3, respectively, which have similar L_X distributions (see Figure 9), and perform the same corrections. The results are
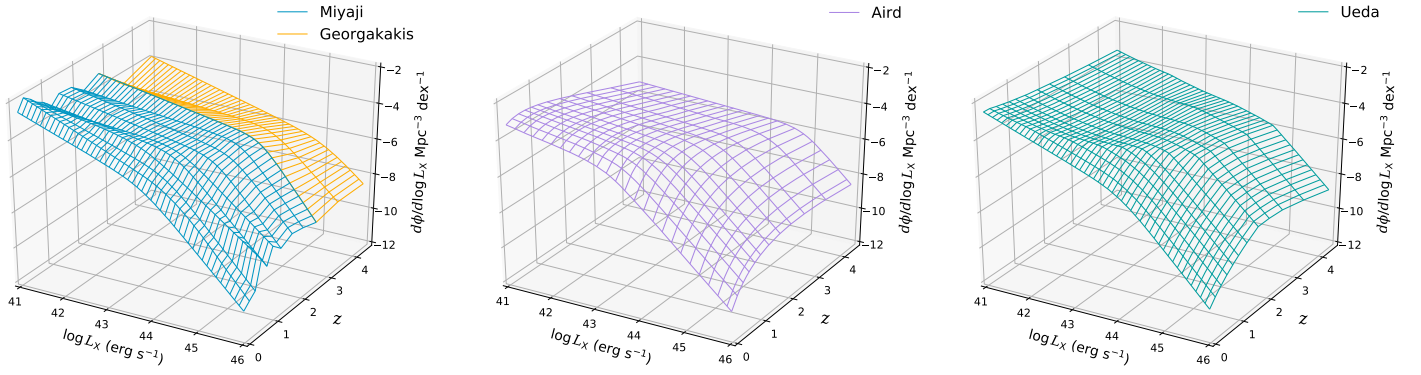
Figure 17. Left: combined X-ray LFs taken from Miyaji et al. (2015) at $z < 3$ and Georgakakis et al. (2015) at $z > 3$. Middle: obscured ($22\ \mathrm{cm}^{-2} < \log N_H < 24\ \mathrm{cm}^{-2}$) AGN LFs from Aird et al. (2015). Right: Compton-thin AGN ($\log N_H < 24\ \mathrm{cm}^{-2}$) LFs from Ueda et al. (2014).
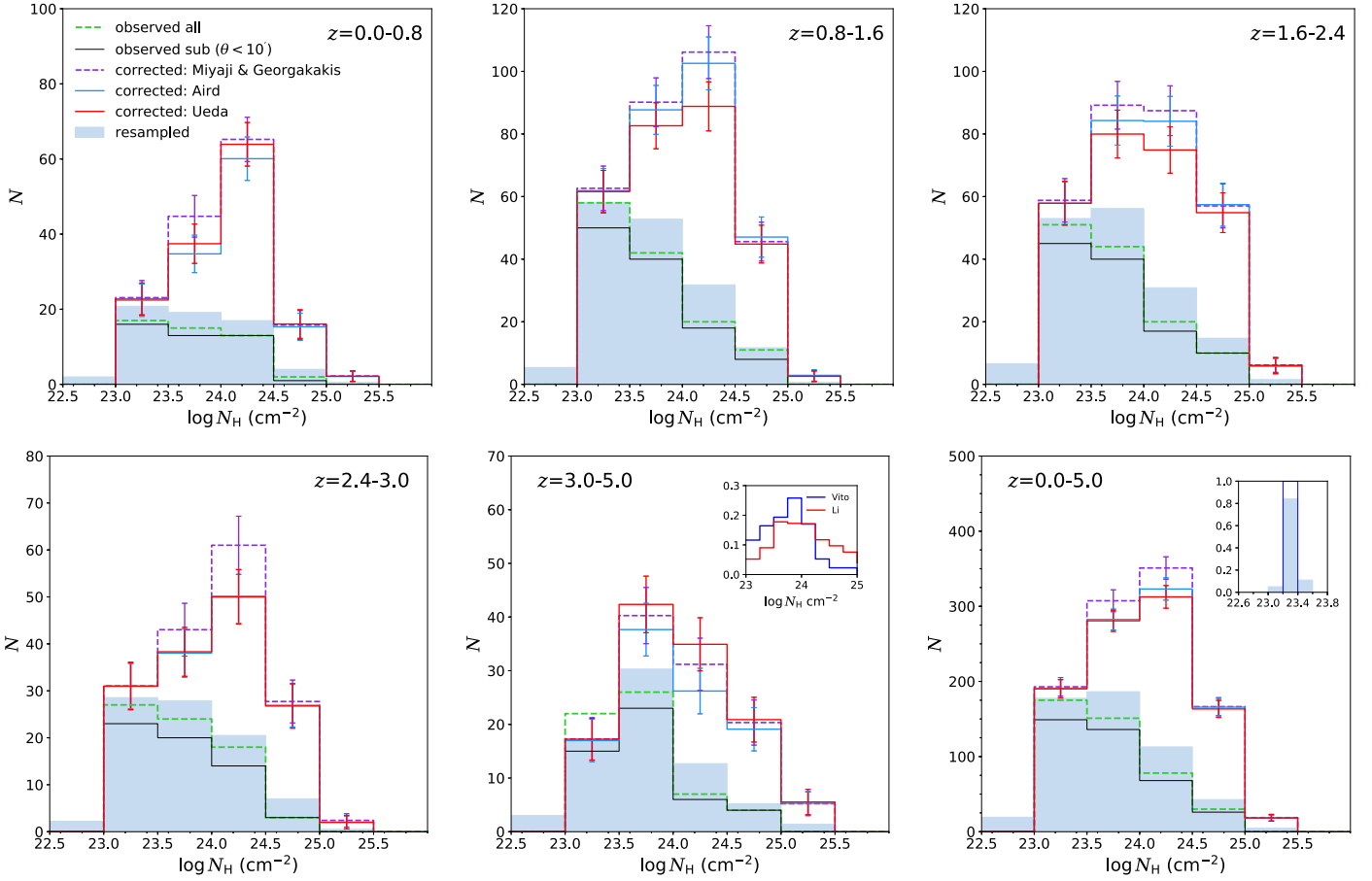
Figure 18. N_H distributions for our HOS in five redshift bins and the entire redshift range ($z = 0$–5). The error bars are estimated through bootstrapping. In each panel, the green dashed histogram shows the observed best-fit N_H distribution, the black histogram shows the observed N_H distribution for sources with off-axis angle $<10'$, the blue shaded area shows the resampled N_H distribution that takes into account the spectral fitting errors and sky coverage effect, and the red and blue solid histograms and purple dashed histogram show the intrinsic N_H distributions after correcting for the undetectable parts of the two surveys using the Ueda et al. (2014), Aird et al. (2015), and Miyaji et al. (2015) and Georgakakis et al. (2015) LFs, respectively. The comparison of the normalized N_H distribution of highly obscured AGNs at $z > 3$ presented in Vito et al. (2018) and this work is shown in the inset of the $z = 3$–5 panel; our result shows a systematically higher CT fraction. The inset in the $z = 0$–5 panel illustrates the effect of the resampling procedure, where the observed N_H distribution shown in blue is resampled into the shaded histogram by considering N_H errors.

listed in Table 6 and are still consistent with no redshift evolution.

Vito et al. (2018) presented the N_H distribution for the high-redshift ($z > 3$) AGNs also in the CDFs. Their analyses were based on modeling the X-ray spectra using the *wabs* × *zwabs* × *zpow* model for the $z > 3$ sources by carefully taking

into account the photometric redshift errors. The comparison of the two works is shown in the $z = 3.0$–5.0 bin in Figure 18. Our result shows a higher CT fraction, which might be ascribed to the different redshift values adopted in the two works, as well as the large weighting factors while correcting for the undetectable part (see Section 6.4).

Table 6
Evolution of the Observed ($f_{\rm obs}$) and Corrected Intrinsic
($f_{\rm int}$) CT/Highly Obscured Fraction with Redshift

z (1)	$f_{\rm obs,CH}$ (2)	$f_{\rm int,CH}$ (3)	$\overline{\log L_{\rm X}}$ (erg s^{-1}) (4)
Total Sample			
0.0–0.8	0.32	0.57 ± 0.07	42.9
0.8–1.6	0.24	0.50 ± 0.05	43.5
1.6–2.4	0.24	0.50 ± 0.05	43.7
2.4–3.0	0.29	0.54 ± 0.06	43.9
3.0–5.0	0.18	0.50 ± 0.09	44.1
0.0–5.0	0.25	0.52 ± 0.06	43.6
Subsample			
1.0–2.0	0.25	0.50 ± 0.05	43.7
2.0–3.0	0.26	0.51 ± 0.05	43.8

Note. The values listed are calculated by averaging the results from the three LF models. Top: results for the total sample. Bottom: results for two subsamples with 43.0 erg s^{-1} < $\log L_{\rm X}$ < 44.5 erg s^{-1} at $z = 1.0$–2.0 and at $z = 2.0$–3.0, which have similar $L_{\rm X}$ distributions (see Figure 9).

However, we note that all these aforementioned corrections are based on sources we do observe, including the correction of the undetectable part in a given $N_{\rm H}$ bin for which we have to make assumptions regarding the unknown AGN population. The series of corrections can be simply understood as multiplying an evolved weighting factor by the observed distribution, which has several limitations. First of all, as mentioned before, different redshift bins sample different luminosity ranges. Several authors have found that the obscured fraction changes with luminosity (Lusso et al. 2013; Buchner et al. 2015); thus, the $N_{\rm H}$ distribution might be luminosity dependent. However, in Equation (2) we ignore this effect and only report the result for the full luminosity range, owing to the limited CT AGN amount detected in faint-$L_{\rm X}$ bins. Second, for the extremely buried sources ($N_{\rm H} > 10^{25}$ cm^{-2}), or intrinsically very faint sources, the corrections cannot be made since there are hardly any such sources observed owing to the limit of the survey and the restrictions imposed by our sample selection count cut. Therefore, the part of the $N_{\rm H}$ distribution of sources that are under the detection limit is actually still missed in our final result. It is plausible that $N_{\rm H} > 10^{25}$ cm^{-2} sources may contribute significantly to the heavily obscured AGN population (Risaliti et al. 1999), but such sources are missed in our sample, and the corrections in this part are beyond our attainment. Therefore, the $N_{\rm H}$ distribution at the highest-$N_{\rm H}$ bin displayed in Figure 18 is highly uncertain and incomplete, so that the CT/highly obscured fractions in Table 6 should be better treated as lower limits.

7. Spectral Variability Analyses

X-ray variability is a ubiquitous feature among AGNs that can provide useful information about AGN properties and relevant underlying physical processes (e.g., Vaughan et al. 2003; McHardy et al. 2006; González-Martín & Vaughan 2012; Soldi et al. 2014; Zheng et al. 2017). In this section, we aim to study the main driving mechanism that causes the variability of

highly obscured AGNs by investigating the variability of some main spectral parameters, such as the luminosity, column density, and reflection strength.

To perform detailed long-term spectral variability analyses of highly obscured AGNs and expand the sample size, we select our sources from the HOS based on only one additional criterion, i.e., the net 0.5–7 keV counts in the total 7 Ms exposure in the CDF-S or in the total 2 Ms exposure in the CDF-N should be larger than 700 or 900, respectively. We bin data from neighboring observations as one epoch to enhance the S/N. For sources in the CDF-S that have counts > 1500, we bin the data into four epochs; for sources having 700 < counts < 1500, we bin the data into three epochs. The data for all the sources in the CDF-N are binned into three epochs. Detailed binning information can be found in Table 7. These binning strategies and the adopted count criteria can roughly ensure that the average counts in each epoch are at least ∼200. Thirty-one sources, which satisfy our selection criteria, are selected for subsequent variability analyses. The redshift distribution of this variability sample is shown in Figure 9, and the count distribution is shown in Figure 1. The mean number of counts for the variability sample is 1615. The binning process is carried out using the combine_spectra tool in CIAO to generate the source spectra, background spectra, and rmf and arf files in each epoch.

7.1. Method

We use the same spectral fitting models as described in Section 3.1 to perform spectral variability analyses, except for removing the constants $f_{\rm ref}$ and $f_{\rm exs}$ in the model and untying the normalizations of the intrinsic power law and other spectral components. We simultaneously fit the spectra in each epoch using the best-fit model for each source obtained in Section 4, but this time we allow $norm_{\rm ref}$ to vary freely to search for possible variation in the scattered continuum. Considering the relatively low counts in each epoch, the fitting strategies are adopted as follows:

1. The uncertainties on Γ can cause large degeneracies if we set it free in all epochs. Given that some sources may have large variability in Γ, we first let Γ in each epoch vary freely and obtain $C_{\rm free}$. Then, for the first two adjacent epochs (i.e., epoch1 and epoch2), we link Γ and measure the ΔC with respect to $C_{\rm free}$. If $\Delta C > 3.84$ (for $\Delta {\rm dof} = 1$), Γ is considered to be different between the two epochs at >95% confidence level, and we will set Γ free; otherwise, we will link Γ. Then for the second epoch pair (i.e., epoch2 and epoch3), we link their Γ and compare the Cstat value with the last step. After traversing all the epoch pairs, if no variability is detected, Γ is linked together.

2. The reflection and soft-excess components are often considered to be produced in large-scale clouds, e.g., torus, for highly obscured AGNs. Since the time span in the rest-frame is $(1+z)$ times less than in the observed frame and the typical redshift of our sample is relatively high, it is reasonable to assume that in such a short timescale the large-scale components may not vary dramatically. To better constrain the normalization of the intrinsic power law, we tie the normalizations of the reflection component and the soft-excess component in each epoch, respectively, unless setting them free, leading

Table 7
Observational Epochs, Variability Sample, and Binning Information

CDF-S	Start Time	End Time	CDF-N	Start Time	End Time
I	1999 Oct	2000 Dec	I	1999 Nov	2000 Nov
II	2007 Sep	2007 Nov	II	2001 Feb	2001 Nov
III	2010 Mar	2010 Jul	III	2002 Feb	2002 Feb
IV	2014 Jun	2015 Jan			

XID (1)	z (2)	N (3)	M (4)	Counts (5)	N_{obs} (6)	I (7)	II (8)	III (9)	IV (10)
				CDF-S Four Epochs					
49	2.394	4	A	2345	98	0.93	0.96	1.96	2.70
81	3.309	4	A	3579	102	0.93	0.96	1.96	2.88
98	1.412	4	B	3689	102	0.93	0.96	1.96	2.88
186	2.810	4	A	2532	102	0.93	0.96	1.96	2.88
214	3.740	4	A	2863	102	0.93	0.96	1.96	2.88
236	2.562	4	A	1578	102	0.93	0.96	1.96	2.88
308	1.097	4	B	2484	102	0.93	0.96	1.96	2.88
458	2.291	4	C	2093	102	0.93	0.96	1.96	2.88
746	3.064	4	A	2144	102	0.93	0.96	1.96	2.88
760	3.350	4	B	1974	102	0.93	0.96	1.96	2.88
876	3.470	4	A	4191	102	0.93	0.96	1.96	2.88
933	1.654	4	A	1792	102	0.93	0.96	1.96	2.88
				CDF-S Three Epochs					
63	0.737	3	B	848	29	0.45	0.56	1.04	
91	2.256	3	A	756	29	0.45	0.56	1.04	
752	0.733	3	A	866	55	0.80	1.27	1.69	
785	1.600	3	C	1345	22	0.31	0.36	0.60	
73	2.509	3	A	879	102	1.89	1.96	2.88	
240	1.185	3	C	785	102	1.89	1.96	2.88	
249	0.735	3	A	830	102	1.89	1.96	2.88	
328	1.536	3	C	1432	102	1.89	1.96	2.88	
367	0.604	3	C	948	102	1.89	1.96	2.88	
399	1.730	3	C	1077	102	1.89	1.96	2.88	
551	3.700	3	C	756	102	1.89	1.96	2.88	
621	1.213	3	A	784	102	1.89	1.96	2.88	
658	1.845	3	A	839	102	1.89	1.96	2.88	
733	2.404	3	D	973	102	1.89	1.96	2.88	
818	2.593	3	C	711	102	1.89	1.96	2.88	
840	1.220	3	C	1321	102	1.89	1.96	2.88	
846	2.483	3	A	974	102	1.89	1.96	2.88	
				CDF-N Three Epochs					
66	0.959	3	A	934	20	0.49	0.87	0.58	
143	1.727	3	A	1516	20	0.49	0.87	0.58	

Note. Top: definition of the observational epochs. Bottom: variability sample and the binning information. Column (3): number of binning epochs. Column (4): best-fit model. Column (5): total 0.5–7 keV net counts. Column (6): number of individual observations of the source during the total CDF-S 7 Ms or CDF-N 2 Ms exposure. Columns (7)–(10): exposure time in units of Ms in each epoch.

to $\Delta C > 3.84$. Finally, only one source shows a significant change in the reflection component, and the soft-excess fluxes remain constant for all sources.

3. For those sources with a weak and constant reflection component confirmed in the last step, we delete the MYTS and MYTL components from the model and only use MYTZ to fit the spectra, in order to reduce the number of free parameters. This procedure may break the self-consistency of the MYTorus model to some extent but will not influence the variability analysis, since we

only remove a small constant component in each epoch, which has very little influence on the calculated χ^2.

4. The N_H and normalization of the intrinsic power law always vary freely.

The simultaneous fitting yields the best-fit model parameters Γ, $norm_{ref}$, and $norm_{scat}$. Then, we fit the spectra in each epoch with Γ_i, $norm_{ref,i}$, and $norm_{scat,i}$ fixed at the best-fit value obtained in the simultaneous fitting to obtain the N_H, observed flux, intrinsic flux, and corresponding errors in each epoch. The spectral fitting parameters of each epoch are listed in Table 8. No significant photon index variability is detected for all sources.

7.2. L_X, N_H, and Observed Flux Variability

We identify L_X, N_H, and observed flux-variable sources based on the classical χ^2 test. For illustration, we explain how we determine L_X-variable sources. First, we use the *cflux* model in XSPEC to calculate the observed and absorption-corrected 0.5–7 keV fluxes and corresponding errors. Then, we derive the intrinsic rest-frame 0.5–7 keV luminosity using equation $L_{0.5-7\,keV} = 4\pi d_L^2 f_{0.5-7\,keV,int} (1 + z)^{\Gamma_{int}-2}$, where the "int" subscript represents the intrinsic value. Since the photon index and redshift are all fixed during spectral fitting of each single epoch, the error on $L_{0.5-7\,keV}$ is totally attributed to the error of the intrinsic flux $f_{0.5-7\,keV,int}$. Thus, the χ^2 of the rest-frame $L_{0.5-7\,keV}$ actually equals the χ^2 of the observed-frame $f_{0.5-7\,keV,int}$. If a source does not vary, i.e., the intrinsic luminosity (or N_H, flux) is constant, assuming that the errors obey the Gaussian distribution, the χ^2 value calculated from

$$\chi^2 = \sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{\sigma_{x,i}^2} \qquad (6)$$

should obey the χ^2 distribution with $(N - 1)$ dof, where x represents the tested parameter (L_X, N_H, or flux), $\sigma_{x,i}$ means the 1σ error in the ith epoch, and N is the number of epochs. The χ^2 test results are listed in Table 9. When the source with three (four) binning epochs satisfies $\chi^2 > 7.8$ (9.8), we regard it as being variable at >98% confidence level.[24] We also check the L_X- and N_H-variable identification results using the Akaike information criterion (AIC) method as described in Y16, which does not need to assume the Gaussian errors (see Section 3.2.3 of Y16 for details). The ΔAIC values for each source are also listed in Table 9. Sources with ΔAIC > 4.0 are assigned as being variable. It can be seen that 30/31 and 30/31 of sources have consistent results in identifying L_X and N_H variability while using the χ^2 test and the AIC method, respectively.[25]

Based on the χ^2 test results, sources that show observed flux variability account for $55\% \pm 13\%$ (17/31) of the entire sample. The resulting L_X and N_H variable source fractions are $29\% \pm 10\%$ (9/31) and $19\% \pm 8\%$ (6/31), respectively. For the majority of sources, we do not find any correlation between L_X and N_H variability patterns, suggesting that the main reason that causes the variation of N_H is likely not the change of

---

[24] The choice of this confidence level is because Y16 showed that it roughly corresponds to ΔAIC = 4, so that we can directly compare the χ^2 results with those obtained using the AIC method.

[25] We decide to use the χ^2 test results in the following analysis since for the two discrepant sources, CDFS XID 933 and CDFN XID 66, the AIC results show neither L_X nor N_H variability, which is inconsistent with their flux-variable nature.

Table 8
Multiepoch Spectral Fitting Results for the Variability Sample

XID	Γ	N_{H1} (a)	N_{H2}	N_{H3}	L_{X1} (b)	L_{X2}	L_{X3}	flux$_1$ (c)	flux$_2$	flux$_3$
					CDF-S					
63	1.54	$0.06^{+0.01}_{-0.01}$	$0.14^{+0.02}_{-0.03}$	$0.19^{+0.04}_{-0.03}$	0.47	0.66	0.41	$16.0^{+1.9}_{-1.8}$	$16.6^{+1.5}_{-1.5}$	$8.8^{+0.9}_{-0.9}$
73	1.81	$0.64^{+0.08}_{-0.07}$	$0.56^{+0.11}_{-0.11}$	$0.67^{+0.07}_{-0.07}$	3.37	1.87	3.78	$2.7^{+0.3}_{-0.3}$	$1.7^{+0.3}_{-0.3}$	$2.9^{+0.3}_{-0.3}$
91	2.08	$0.16^{+0.03}_{-0.03}$	$0.12^{+0.03}_{-0.03}$	$0.14^{+0.03}_{-0.03}$	1.29	1.03	1.25	$2.3^{+0.3}_{-0.3}$	$2.1^{+0.3}_{-0.3}$	$2.4^{+0.3}_{-0.3}$
240	1.40	$1.80^{+0.17}_{-0.15}$	$1.41^{+0.17}_{-0.15}$	$1.34^{+0.12}_{-0.11}$	1.76	1.21	1.04	$3.7^{+0.5}_{-0.4}$	$3.4^{+0.5}_{-0.5}$	$3.3^{+0.3}_{-0.3}$
249	1.51	$0.50^{+0.14}_{-0.12}$	$0.28^{+0.08}_{-0.08}$	$0.14^{+0.02}_{-0.01}$	0.25	0.16	0.18	$2.1^{+0.3}_{-0.3}$	$2.4^{+0.3}_{-0.3}$	$4.2^{+0.3}_{-0.3}$
328	2.12	$1.91^{-1.02}_{-0.12}$	$1.31^{+0.11}_{-0.10}$	$1.57^{+0.06}_{-0.06}$	4.37	1.69	4.94	$4.2^{+0.4}_{-0.4}$	$3.4^{+0.5}_{-0.4}$	$6.0^{+0.4}_{-0.4}$
367	2.07	$0.42^{+0.03}_{-0.02}$	$0.38^{+0.02}_{-0.02}$	$0.47^{+0.06}_{-0.04}$	0.14	0.13	0.07	$3.7^{+0.2}_{-0.3}$	$3.8^{+0.3}_{-0.3}$	$2.2^{+0.2}_{-0.2}$
399	1.40	$0.24^{+0.02}_{-0.02}$	$0.23^{+0.03}_{-0.03}$	$0.32^{+0.02}_{-0.02}$	0.38	0.13	0.32	$2.9^{+0.3}_{-0.3}$	$1.5^{+0.3}_{-0.3}$	$2.6^{+0.2}_{-0.2}$
551	1.84	$1.56^{+0.11}_{-0.10}$	$1.62^{+0.15}_{-0.14}$	$1.54^{+0.09}_{-0.09}$	4.02	2.31	3.89	$2.0^{+0.2}_{-0.2}$	$1.5^{+0.2}_{-0.2}$	$1.9^{+0.2}_{-0.2}$
621	2.11	$0.18^{+0.02}_{-0.02}$	$0.15^{+0.02}_{-0.02}$	$0.17^{+0.02}_{-0.02}$	0.40	0.32	0.30	$2.3^{+0.2}_{-0.2}$	$2.0^{+0.3}_{-0.3}$	$1.8^{+0.2}_{-0.2}$
658	1.67	$0.15^{+0.02}_{-0.02}$	$0.18^{+0.03}_{-0.03}$	$0.15^{+0.02}_{-0.02}$	0.54	0.57	0.52	$2.0^{+0.2}_{-0.2}$	$2.0^{+0.2}_{-0.2}$	$2.0^{+0.2}_{-0.2}$
733	1.84	$0.21^{+0.04}_{-0.03}$	$0.17^{+0.03}_{-0.03}$	$0.12^{+0.02}_{-0.02}$	0.94	0.96	0.92	$1.8^{+0.2}_{-0.2}$	$2.0^{+0.2}_{-0.2}$	$2.2^{+0.2}_{-0.2}$
752	2.24	$0.15^{+0.02}_{-0.02}$	$0.15^{+0.02}_{-0.02}$	$0.11^{+0.01}_{-0.01}$	0.28	0.33	0.25	$4.6^{+0.7}_{-0.7}$	$5.4^{+0.7}_{-0.7}$	$5.0^{+0.5}_{-0.5}$
785	1.97	$0.54^{+0.05}_{-0.05}$	$0.56^{+0.05}_{-0.04}$	$0.53^{+0.03}_{-0.02}$	4.73	5.97	9.26	$16.2^{+2.8}_{-2.7}$	$18.8^{+2.1}_{-2.1}$	$29.7^{+1.8}_{-1.8}$
818	1.97	$0.14^{+0.02}_{-0.02}$	$0.11^{+0.03}_{-0.03}$	$0.19^{+0.03}_{-0.02}$	0.61	0.46	0.93	$1.1^{+0.1}_{-0.1}$	$0.9^{+0.1}_{-0.1}$	$1.5^{+0.1}_{-0.1}$
840	1.60	$0.58^{+0.05}_{-0.04}$	$0.59^{+0.04}_{-0.04}$	$0.40^{+0.02}_{-0.02}$	0.53	0.49	0.59	$3.9^{+0.3}_{-0.3}$	$3.8^{+0.3}_{-0.3}$	$5.1^{+0.3}_{-0.3}$
846	1.56	$0.14^{+0.03}_{-0.02}$	$0.11^{+0.03}_{-0.03}$	$0.15^{+0.03}_{-0.03}$	0.84	0.86	0.90	$2.1^{+0.2}_{-0.2}$	$2.2^{+0.2}_{-0.2}$	$2.2^{+0.2}_{-0.2}$
					CDF-N					
66	1.92	$0.34^{+0.09}_{-0.06}$	$0.24^{+0.02}_{-0.02}$	$0.29^{+0.03}_{-0.03}$	1.19	1.49	1.10	$7.6^{+1.0}_{-1.0}$	$12.7^{+0.7}_{-0.7}$	$8.1^{+0.7}_{-0.7}$
143	1.46	$0.15^{+0.02}_{-0.02}$	$0.15^{+0.01}_{-0.01}$	$0.11^{+0.01}_{-0.01}$	2.07	2.79	2.42	$10.2^{+0.9}_{-0.9}$	$13.6^{+0.6}_{-0.7}$	$13.1^{+0.7}_{-0.7}$

XID	Γ	N_{H1}	N_{H2}	N_{H3}	N_{H4}	L_{X1}	L_{X2}	L_{X3}	L_{X4}	flux$_1$	flux$_2$	flux$_3$	flux$_4$
						CDF-S							
49	1.66	$0.19^{+0.02}_{-0.02}$	$0.15^{+0.04}_{-0.03}$	$0.30^{+0.04}_{-0.03}$	$0.16^{+0.02}_{-0.02}$	3.90	1.85	2.69	3.44	$8.3^{+0.6}_{-0.6}$	$4.4^{+0.6}_{-0.6}$	$4.7^{+0.4}_{-0.4}$	$7.8^{+0.4}_{-0.4}$
81	1.83	$0.17^{+0.02}_{-0.02}$	$0.14^{+0.02}_{-0.02}$	$0.17^{+0.02}_{-0.02}$	$0.19^{+0.02}_{-0.02}$	7.32	9.21	5.15	6.37	$7.8^{+0.5}_{-0.4}$	$10.4^{+0.6}_{-0.6}$	$5.4^{+0.4}_{-0.4}$	$6.5^{+0.3}_{-0.3}$
98	0.98	$0.18^{+0.03}_{-0.02}$	$0.17^{+0.02}_{-0.02}$	$0.18^{+0.02}_{-0.02}$	$0.17^{+0.01}_{-0.01}$	1.06	1.18	1.29	1.11	$10.3^{+0.8}_{-0.8}$	$11.4^{+1.0}_{-1.0}$	$12.4^{+0.8}_{-0.8}$	$11.1^{+0.6}_{-0.6}$
186	1.79	$0.22^{+0.02}_{-0.03}$	$0.27^{+0.03}_{-0.03}$	$0.22^{+0.02}_{-0.02}$	$0.23^{+0.02}_{-0.02}$	3.92	4.68	3.20	4.23	$5.2^{+0.4}_{-0.4}$	$5.7^{+0.4}_{-0.5}$	$4.3^{+0.3}_{-0.3}$	$5.6^{+0.2}_{-0.2}$
214	1.87	$0.26^{+0.03}_{-0.03}$	$0.35^{+0.05}_{-0.04}$	$0.30^{+0.03}_{-0.03}$	$0.30^{+0.03}_{-0.03}$	9.25	9.36	8.61	8.81	$6.3^{+0.4}_{-0.4}$	$5.6^{+0.5}_{-0.4}$	$5.5^{+0.4}_{-0.4}$	$5.7^{+0.2}_{-0.2}$
236	1.62	$0.15^{+0.03}_{-0.03}$	$0.20^{+0.03}_{-0.03}$	$0.22^{+0.04}_{-0.03}$	$0.21^{+0.03}_{-0.03}$	1.68	2.16	1.61	1.65	$3.6^{+0.3}_{-0.3}$	$4.2^{+0.5}_{-0.5}$	$3.0^{+0.3}_{-0.3}$	$3.1^{+0.2}_{-0.2}$
308	1.69	$0.21^{+0.02}_{-0.01}$	$0.24^{+0.03}_{-0.02}$	$0.21^{+0.02}_{-0.01}$	$0.22^{+0.02}_{-0.02}$	1.45	0.90	1.11	0.76	$11.9^{+0.5}_{-0.7}$	$6.8^{+0.6}_{-0.6}$	$9.2^{+0.5}_{-0.5}$	$6.2^{+0.3}_{-0.3}$
458	1.83	$0.19^{+0.02}_{-0.02}$	$0.25^{+0.02}_{-0.02}$	$0.31^{+0.02}_{-0.02}$	$0.27^{+0.02}_{-0.02}$	0.96	1.75	2.11	1.57	$3.8^{+0.4}_{-0.4}$	$5.7^{+0.5}_{-0.5}$	$4.7^{+0.3}_{-0.3}$	$3.7^{+0.2}_{-0.2}$
746	1.68	$0.41^{+0.04}_{-0.04}$	$0.46^{+0.06}_{-0.06}$	$0.45^{+0.04}_{-0.04}$	$0.55^{+0.04}_{-0.03}$	5.17	4.75	5.81	6.93	$4.9^{+0.2}_{-0.3}$	$4.2^{+0.4}_{-0.4}$	$5.2^{+0.3}_{-0.4}$	$5.4^{+0.3}_{-0.3}$
876	1.87	$0.15^{+0.02}_{-0.02}$	$0.16^{+0.02}_{-0.02}$	$0.15^{+0.02}_{-0.02}$	$0.16^{+0.02}_{-0.02}$	7.18	7.98	8.99	8.84	$6.9^{+0.4}_{-0.4}$	$7.6^{+0.5}_{-0.5}$	$8.7^{+0.4}_{-0.4}$	$8.3^{+0.3}_{-0.3}$
933	1.71	$0.31^{+0.04}_{-0.03}$	$0.23^{+0.03}_{-0.03}$	$0.25^{+0.02}_{-0.02}$	$0.23^{+0.02}_{-0.02}$	1.81	1.46	1.77	1.29	$5.6^{+0.6}_{-0.6}$	$5.3^{+0.5}_{-0.5}$	$6.2^{+0.4}_{-0.4}$	$4.7^{+0.3}_{-0.3}$
760	1.49	$0.81^{+0.12}_{-0.10}$	$0.67^{+0.10}_{-0.10}$	$0.68^{+0.06}_{-0.06}$	$0.59^{+0.05}_{-0.05}$	9.18	6.51	8.99	7.33	$6.5^{+0.9}_{-0.8}$	$5.4^{+0.9}_{-0.8}$	$7.0^{+0.6}_{-0.6}$	$6.5^{+0.4}_{-0.4}$

Note. (a) in units of 10^{24} cm^{-2}. (b) 2–10 keV intrinsic luminosity, in units of 10^{44} erg s^{-1}. (c) observed 0.5–7 keV flux, in units of 10^{-15} erg cm^{-2} s^{-1}.

ionization parameter induced by the variable L_X, but is likely the occultation of the clumpy clouds moving in/out of our LOS.

7.3. Variability Amplitude Estimation

We use the normalized excess variance σ^2_{nxv} (nxv) to estimate the variability amplitude. σ^2_{nxv} is defined as

$$\sigma^2_{\mathrm{nxv}} = \frac{1}{N\langle x\rangle^2}\sum_{i=1}^{N}[(x_i - \langle x\rangle)^2 - (\Delta x_i)^2], \quad (7)$$

where N is the number of binning epochs (three or four), x_i and Δx_i are the best-fit parameters (observed flux, L_X, N_H) and their 1σ errors, and $\langle x\rangle$ is the unweighted mean of the calculated

parameter. The error on σ^2_{nxv} is estimated as $s_D/(\langle x\rangle^2\sqrt{N})$, where

$$s_D^2 = \frac{1}{N-1}\sum_{i=1}^{N}[(x_i - \langle x\rangle)^2 - (\Delta x_i)^2 - \sigma^2_{\mathrm{nxv}}\langle x\rangle^2]^2. \quad (8)$$

We calculate σ^2_{nxv} and corresponding errors on the intrinsic rest-frame 0.5–7 keV luminosity ($\sigma^2_{\mathrm{nxv},L}$), observed 0.5–7 keV flux ($\sigma^2_{\mathrm{nxv},F}$), and N_H ($\sigma^2_{\mathrm{nxv},N}$) for each source. The results are listed in Table 9. To check whether our calculation of σ^2_{nxv} is affected by the limited counts (e.g., Zheng et al. 2017, hereafter Z17), we perform Spearman rank correlation tests and find no significant correlation between $\sigma^2_{\mathrm{nxv},F}$, $\sigma^2_{\mathrm{nxv},N}$, $\sigma^2_{\mathrm{nxv},L}$, and counts with Spearman's $\rho = 0.07$, p-value $= 0.70$;

Table 9
The Chi-Squared Value χ^2, the ΔAIC Value Δ, and the Normalized Excess Variance σ^2 of the Observed 0.5–7 keV Flux, Intrinsic Rest-frame $L_{0.5-7\,\mathrm{keV}}$, and N_H

XID	χ^2_{flux}	$\chi^2_{L_X}$	Δ_{L_X}	$\chi^2_{N_H}$	Δ_{N_H}	σ^2_{flux}	$\sigma^2_{L_X}$	$\sigma^2_{N_H}$
				CDF-S Four Epochs				
49	53.3	21.9	11.9	17.3	11.3	0.074	0.056	0.073
81	69.1	23.7	13.9	3.1	−3.2	0.058	0.037	0
98	3.6	2.9	−3.1	0.4	−5.6	0	0	0
186	13.4	9.6	2.6	2.5	−3.9	0.007	0.009	0
214	2.9	0.4	−6.0	3.0	−3.3	0	0	0
236	7.0	2.7	−3.6	4.4	−2.1	0.008	0	0
308	86.7	44.9	32.4	1.4	−4.7	0.066	0.054	0
458	21.7	28.4	13.5	7.5	2.8	0.024	0.056	0.013
746	6.8	7.7	2.0	7.8	1.9	0.004	0.010	0.003
876	12.0	5.5	−0.9	0.4	−5.6	0.005	0.002	0
933	10.3	10.3	2.2	4.9	−1.4	0.002	0.004	0.005
760	2.6	3.0	−3.1	5.6	−1.7	0	0	0
				CDF-S Three Epochs				
63	34.2	7.2	2.8	65.2	29.45	0.054	0.026	0.147
73	7.7	5.7	0.7	0.9	−3.27	0.029	0.032	0
240	0.6	1.6	−0.9	1.6	−0.52	0	0	0
249	31.0	2.2	−2.9	130.8	14.38	0.086	0	0.150
328	20.8	18.3	8.5	2.4	1.14	0.050	0.095	0
367	33.8	9.6	1.4	1.4	−0.38	0.044	0.046	0
399	13.5	14.6	8.3	2.6	−1.31	0.051	0.118	0
551	3.6	2.4	−2.2	0.1	−3.95	0.004	0	0
658	0.1	0.2	−3.8	0.8	−3.14	0	0	0
752	0.7	1.9	−2.3	6.8	1.16	0	0	0.002
785	26.0	10.6	4.3	0.2	−3.87	0.062	0.057	0
840	12.0	1.7	1.1	16.1	9.59	0.013	0	0.012
846	0.3	0.1	−3.9	1.1	−2.96	0	0	0
621	2.4	2.0	−2.0	1.3	−2.83	0	0	0
91	0.7	1.0	−3.1	1.1	−2.92	0	0	0
818	13.9	12.8	8.1	3.2	−0.7	0.035	0.067	0.004
733	2.0	0.04	−3.9	4.9	0.2	0	0	0.012
				CDF-N Three Epochs				
66	25.8	4.2	−0.4	8.6	0.29	0.050	0	0
143	11.8	6.8	2.4	8.8	4.28	0.012	0.008	0.011

Note. Negative values of σ^2 are set to 0. Sources with $\chi^2 > 7.8$ for three epochs, $\chi^2 > 9.8$ for four epochs, or ΔAIC > 4.0 will be regarded as being variable at >98% confidence level (Y16).

$\rho = -0.11$, p-value $= 0.56$; and $\rho = 0.32$, p-value $= 0.08$, respectively.

We also calculate the fractional root mean square (frms) variability amplitude, which is defined as $\sqrt{\sigma^2_{\mathrm{nxv}}}$ for sources having $\sigma^2_{\mathrm{nxv}} > 0$ and can be treated as the percentage of the variability amplitude. It is obvious that our sample lacks sources that display large variability amplitudes. The mean frms values of the total sample for the observed flux, intrinsic luminosity, and N_H are 12%, 10%, and 6%, respectively. For sources having positive nxv, the mean (maximum) frms values are 17% (29%), 19% (34%), and 16% (39%), respectively.

In addition, we do not find any correlation between $\sigma^2_{\mathrm{nxv},F}$, $\sigma^2_{\mathrm{nxv},L}$, and L_X (Figure 19) whereas other studies suggested a negative trend, i.e., luminous AGNs show relatively small variations (Y16; Z17; Paolillo et al. 2017). We perform a Spearman correlation test for the 25 common sources using the $\sigma^2_{\mathrm{nxv},F}$ derived in Z17, which is obtained from the light-curve
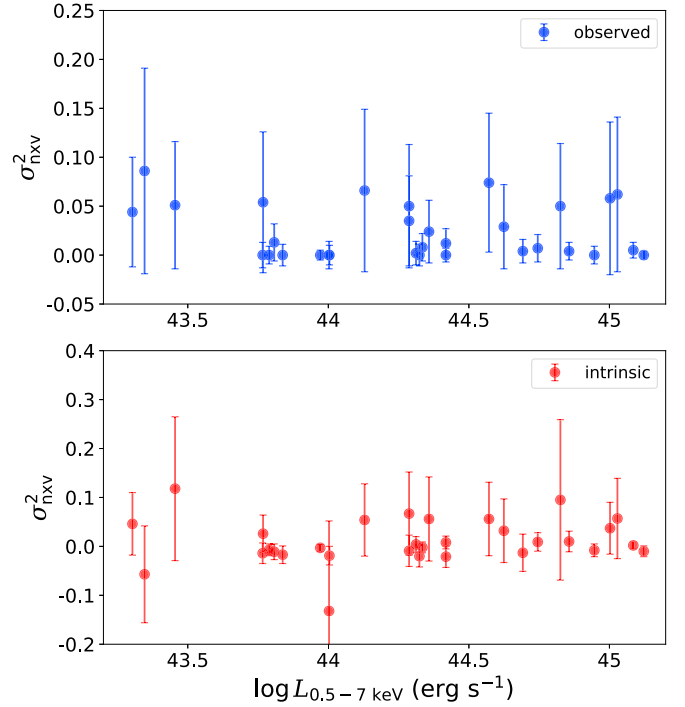

Figure 19. Normalized excess variance (σ^2_{nxv}) of the observed (top) and intrinsic (bottom) 0.5–7 keV flux vs. the intrinsic 0.5–7 keV luminosity for the variability sample. The original values without fixing negative σ^2_{nxv} at 0 are displayed. The error bars are calculated using Equation 8.

analysis; the result is still consistent with no correlation between L_X and $\sigma^2_{\mathrm{nxv},F}$ (Spearman's $\rho = -0.31$, p-value $= 0.13$). As pointed out by Allevato et al. (2013), the calculation of σ^2_{nxv} is biased at low counts and uneven cadence and should be restricted to large samples. Therefore, the limited source number and the broad redshift span may prevent us from detecting such a correlation. But beyond that, since Y16 and Z17 mainly focused on the AGNs with largest counts available in the CDF-S rather than highly obscured AGNs, the variability behavior and its underlying physical drivers of our sources and those in Y16 and Z17 may be intrinsically different.

Recently, González-Martín (2018) found a new X-ray variability plane for AGNs that links the characteristic break frequency of the PSD (f_{break}) with bolometric luminosity and N_H. This makes N_H play a non-negligible role in understanding AGN variability. As we will show in the next section, although Z17 argued that the N_H variation does not contribute significantly to the total variability in their sample, the N_H variability does play a substantial role for highly obscured AGNs. Nevertheless, the small σ^2_{nxv} values indicate that the state of highly obscured AGNs does not vary dramatically while concerning the average source properties on a 17 yr timescale in the observed frame.

7.4. Detailed Variability Analysis

There are some sources in our sample that show significant variability of the observed flux, intrinsic luminosity, or N_H. In this section, we perform detailed analyses to study their variability behavior aiming at shedding light on the leading mechanism that drives the large variabilities, and try to investigate the typical location and size of the obscuring clouds.
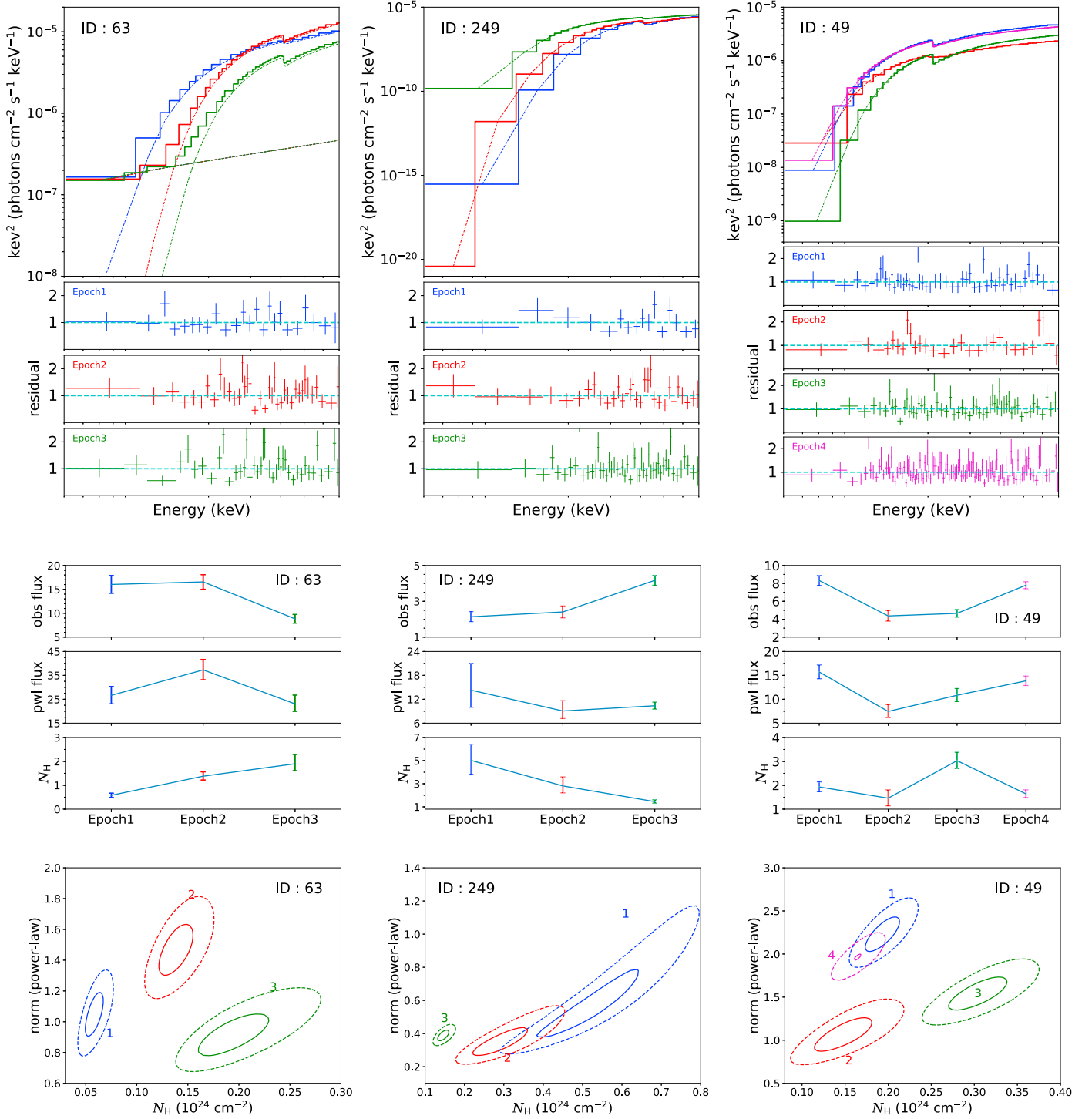
Figure 20. First row: spectra for three highly variable sources XID 63, XID 249, and XID 49 in the CDF-S. The top panels show the unfolded spectra in each epoch, and the bottom panels show the data-to-model ratios. Second row: observed 0.5–7 keV flux, intrinsic 0.5–7 keV flux (in units of 10^{-16} erg cm^{-2} s^{-1}), and $N_{\rm H}$ (in units of 10^{23} cm^{-2}) variability curves for the three sources, respectively. Third row: confidence contours of the normalizations of the intrinsic power law and $N_{\rm H}$ for the three sources. The solid and dashed curves indicate 1σ and 2σ confidence contours, respectively. The numbers annotated represent different epochs.

7.4.1. XID 63

This source has 848 counts available, and its data are binned to three epochs. It is a moderately luminous source ($\bar{L}_{\rm X} = 5.1 \times 10^{43}$ erg s^{-1}) at $z = 0.737$. The best-fit model of this source is the absorbed power law ($\Gamma = 1.54^{+0.21}_{-0.22}$) plus an additional soft-excess component. Γ and $norm_{\rm soft}$ do not show variations and are linked during all the epochs. The

unfolded spectra, light curves, and 1σ and 2σ confidence contours of $N_{\rm H}$ and normalization are displayed in the left column of Figure 20. It has $\chi^2_{N_{\rm H}} = 65.2$, $\chi^2_{f,{\rm obs}} = 34.2$, and $\chi^2_L = 7.2$, indicating significant variabilities. Although the photon index does not vary, the observed spectral shape changes prominently owing to the large variation in column density. The absorption is weak in epoch1 with

$N_{\rm H} = 5.8 \times 10^{22}$ cm^{-2}. Then, it transforms into a highly obscured state with $N_{\rm H} = 1.4 \times 10^{23}$ cm^{-2} in epoch2, and its $N_{\rm H}$ continues to rise to 1.9×10^{23} cm^{-2} in epoch3. The intrinsic flux increases by about a factor of 1.4 from epoch1 to epoch2 and decreases by about a factor of 1.6 from epoch2 to epoch3. Due to the large variability amplitude in $N_{\rm H}$, the observed flux remains constant in the first two epochs but significantly declines in epoch3 by a factor of 1.9. The $N_{\rm H}$ variability may be caused by the clumpy cloud moving into the LOS, since the variations of $N_{\rm H}$ and $L_{\rm X}$ are not correlated. By assuming this transition as an eclipse event, we can roughly estimate the distance and angular size of the cloud using the same method in Y16, thus to constrain the cloud size. We use the empirical relation between the inner torus radius r and 14-195 keV luminosity from Koshida et al. (2014) to estimate the distance. The relation can be described as $\log r = -1.04 + 0.5 \log (L_{14-195\,{\rm keV}}/10^{44})$, where r and $L_{14-195{\rm keV}}$ are in units of pc and erg s^{-1}, respectively. The $L_{14-195\,{\rm keV}} = 2.5 \times 10^{44}$ erg s^{-1} at epoch1 is obtained by extrapolating the power law to the higher-energy band. Thus, the estimated r from the inner torus to the central black hole is ≈ 0.14 pc. By assuming Keplerian motion, the orbital period of the cloud $t_{\rm orbit} = 2\pi \frac{r^{\frac{3}{2}}}{(M_{\rm BH}G)^{\frac{1}{2}}} \approx 519 \left(\frac{M_{\rm BH}}{10^8 M_\odot}\right)^{-\frac{1}{2}}$ yr. The angular size (viewed from the BH) of the cloud is estimated as $\frac{t_{\rm tran}}{t_{\rm orbit}} \times 360° < 4°.1 \left(\frac{M_{\rm BH}}{10^8 M_\odot}\right)^{\frac{1}{2}}$, where the upper limit value of $t_{\rm tran}$ is calculated as the time span (~ 6.0 yr in the rest-frame) between epoch1 and epoch2. By multiplying by r, we estimate the cloud size to be $< 0.01 \left(\frac{M_{\rm BH}}{10^8 M_\odot}\right)^{\frac{1}{2}}$ pc.

7.4.2. XID 249

This source can be well fitted by the absorbed power-law model with $\Gamma = 1.51^{+0.26}_{-0.28}$ and has the most prominent $N_{\rm H}$ variability in our sample ($\chi^2_{N_{\rm H}} = 130.8$) as shown in Figure 20. The $N_{\rm H}$ declines from 5.0×10^{23} cm^{-2} to 2.8×10^{23} cm^{-2}, and finally to 1.5×10^{23} cm^{-2} during the three epochs. The intrinsic flux of this source remains roughly constant ($\chi^2_{\rm flux,in} = 2.2$), but the strong variation in $N_{\rm H}$ causes a significant increase in the observed flux ($\chi^2_{\rm flux,obs} = 31.0$). The variability analyses of each single-epoch spectrum provide very different results compared with the 7Ms stacked spectrum ($\Gamma = 1.20$ when using MYTZ alone, which does not limit the range of Γ), also indicating strong $N_{\rm H}$ variation. Using the same method as for XID 63, we estimate the distance and the cloud size to be < 0.09 pc and $< 0.006 \left(\frac{M_{\rm BH}}{10^8 M_\odot}\right)^{\frac{1}{2}}$ pc for the occultation event from epoch2 to epoch3.

7.4.3. XID 328

This source has the highest column density in our variability sample. The epoch-mean $L_{\rm X}$ and $N_{\rm H}$ are 3.7×10^{44} erg s^{-1} and 8.5×10^{23} cm^{-2}, respectively, which are consistent with the stacked spectral analyses. The spectra can be well described by a power law ($\Gamma = 2.12^{+0.24}_{-0.24}$) plus Compton reflection continuum and strong Fe K lines. $N_{\rm H}$ and the reflection flux remain roughly constant in the time span of 5.8 yr in the rest frame, and the observed flux variations result from the intrinsic X-ray power variability. According to the stacked spectral

analyses, the reflection flux accounts for 27.5% of the observed 0.5–7 keV flux but only accounts for 3.8% of the intrinsic 0.5–7 keV flux. This source has also been reported in the literature as CT candidates (Comastri et al. 2011; Y16). Our results confirm its highly obscured nature but with higher $N_{\rm H}$, $L_{\rm X}$, and Γ.

7.4.4. XID 49

This source was observed 98 times during the total 7 Ms campaign and has 2345 counts available. The spectra can be explained by a simple absorbed power law ($\Gamma = 1.66^{+0.11}_{-0.12}$) with different column densities and normalizations (see Figure 20). The intrinsic flux decreases from 1.57×10^{-14} erg cm^{-2} s^{-1} to 7.45×10^{-15} erg cm^{-2} s^{-1} during the first two epochs and finally rises again to 1.39×10^{-14} erg cm^{-2} s^{-1} in the fourth epoch. The variability behavior of $N_{\rm H}$ does not follow the intrinsic flux variability pattern, indicating that the varying absorption is not caused by the change of ionization state, but is more likely to be caused by the transverse motion of the obscuring clouds.

7.4.5. XID 458

This source has a redshift of 2.291 and can be described by a power law ($\Gamma = 1.83^{+0.14}_{-0.13}$) with reflection hump and iron emission lines. Though the $N_{\rm H}$ is not so high during all the epochs ($\bar{N}_{\rm H} \approx 1.4 \times 10^{23}$ cm^{-2}), the reprocessed components are required to fit the 7 Ms stacked spectrum. The intrinsic flux increases from epoch1 to epoch3 and declines at epoch4. The reflection flux remains roughly constant from epoch1 to epoch2, but decreases at epoch3, and does not show any variability from epoch3 to epoch4. The different variability patterns indicate that there may be a time lag between the intrinsic continuum variability and the large-scale reflection flux variability. The decline of the reflection flux during epoch2 to epoch3 (from observed frame 2007 September to 2010 July, corresponding to rest-frame 0.76 yr) might result from significant flux decline before epoch1, and this decline has just propagated to the reflection medium (possibly the clumpy torus) and causes significant diminishment of the reflection flux. This means that the variable continuum signal needs at least 2.4 yr (rest-frame time span from 1999 October to 2007 September) to spread to the torus from the central emission region, which provides a rough lower-limit estimate, ≈ 0.7 pc, of the location of the reflecting cloud.

7.4.6. Remaining Sources

By applying similar analyses to the remaining sources, we find that 17 sources in our sample that show observed flux variabilities can be classified into three types: 29% (5/17) are caused by the change of $N_{\rm H}$, 53% (9/17) are caused by the AGN intrinsic luminosity variability, and 6% (1/17) are driven by both effects. Note that there are two sources (CDF-S XIDs 186 and 876) identified to be flux variable but showing neither $L_{\rm X}$ nor $N_{\rm H}$ variability according to the χ^2 tests. However, their best-fit intrinsic luminosities both show obvious variations, but due to large errors, the corresponding χ^2_L values are smaller than the critical value. The high $N_{\rm H}$-variable fraction among flux-variable sources confirms our previous thoughts in Section 7.3 that the $N_{\rm H}$ variability is a key ingredient for investigating the variability in highly obscured AGNs.

8. Conclusions

In this study, we present systematic X-ray spectral analyses of 436 highly obscured AGNs ($N_H > 10^{23}$ cm^{-2}) with $L_X > 10^{42}$ erg s^{-1} identified in the CDFs, which make up the largest dedicated highly obscured AGN sample to date, to explore their physical properties and evolution. We also carry out detailed long-term variability analyses for a subsample of 31 sources with largest counts available to investigate the main driver of their spectral variability and the typical variability amplitude. Below we summarize our main results.

1. We perform detailed X-ray spectral fitting for 1152 AGNs in the CDFs with observed-frame 0.5–7 keV net counts >20 using physically appropriate MYTorus-based models, in order to identify heavily obscured ones. By limiting our analyses to sources with log L_X > 42 erg s^{-1} in order to avoid possible contamination from star-forming galaxies, 436 AGNs are confirmed to be highly obscured, with $\overline{z} = 1.9$ and $\overline{\log L_X} = 43.6$ erg s^{-1}.

2. We select 102 CT candidates with best-fit $N_H > 10^{24}$ cm^{-2} and 1σ lower limit $> 5 \times 10^{23}$ cm^{-2}, accounting for ~23% of the highly obscured sample. The observed log N–log S for CT AGNs prefers the moderate CT number counts as predicted by Akylas et al. (2012) and Ueda et al. (2014), while other models (e.g., Gilli et al. 2007; Ballantyne et al. 2011) more or less overestimate or underestimate the number counts.

3. We present a new HR measure of the obscuration level as a function of redshift (the HR curve), which can be used to select heavily obscured AGN candidates without resorting to detailed spectral fitting. The completeness and accuracy by applying the HR curve on the CDF's AGN population to identify highly obscured ones are 88% and 80%, respectively.

4. We find a strong negative correlation between the soft-excess fraction f_{exs} and N_H with a Spearman rank correlation coefficient $\rho = -0.66$. By assuming that the soft excess originates from the backscattered continuum and treating the small f_{ref} as an indicator of high CF of the obscuring materials, this result indicates that a portion of the most heavily absorbed AGNs reside in an extremely geometrically buried environment.

5. Among the 31 CDF-S highly obscured sources that have optical classification results, 19% (6/31) of them are labeled as BLAGN, indicating that at least for part of the AGN population the high-level X-ray obscuration is largely an LOS effect, i.e., some high column density clouds on various scales (but not necessarily a dust-enshrouded torus) may obscure the compact X-ray emitter without blocking the entire BLR; alternatively, the heavy X-ray obscuration may be produced by the BLR itself.

6. After considering the errors on the best-fit N_H and correcting for the sky coverage effect, the Eddington bias, and the N_H-dependent Malmquist bias, we derive the intrinsic N_H distribution representative of the highly obscured AGN population, as well as its evolution across cosmic time. The intrinsic CT/highly obscured fraction is roughly 52% and is consistent with no evident redshift evolution.

7. We select 31 sources with 0.5–7 keV net counts >700 and >900 in the CDF-S and CDF-N, respectively, to perform long-term ($\approx$17 yr in the observed frame) spectral variability analyses. We find that the flux-variable, L_X-variable, and N_H-variable source fractions are 55% ± 13% (17/31), 29% ± 10% (9/31) and 19% ± 8% (6/31), respectively. The typical flux, L_X, and N_H variability amplitudes for those variable sources are 17%, 19%, and 16%, respectively.

8. We calculate the normalized excess variance (σ_{nxv}^2) of N_H, observed 0.5–7 keV flux, and L_X to investigate the intrinsic variability amplitude. No correlation between σ_{nxv}^2 and L_X is detected, possibly due to the observed flux variability for a significant fraction of highly obscured AGNs being caused by the change of N_H rather than the variation of L_X alone, as well as the limited sample size and the broad redshift span.

9. We discuss detailed variability behaviors for five sources that show significant N_H, L_X, or observed flux variability. The typical locations and sizes of the occultation and reflection clouds are estimated (see Section 7.4). The main driver for the variability of the 17 flux-variable sources can be classified into three types: 29% (5/17) are caused by the change of N_H, 53% (9/17) are caused by the AGN luminosity variability, and 6% (1/17) are driven by both effects. Two sources are not classified owing to large measurement errors.

Benefiting from the deepest X-ray surveys to date, our work provides meaningful constraints on the properties and evolution of the AGN obscuring materials over a broad redshift range and quantifies the detailed variability behaviors of these hidden sources, which are crucial for us to better understand the role that highly obscured AGNs played in galaxy evolution. However, X-ray data alone are incapable to show the overall perspective. In a subsequent paper of this series (J. Y. Li et al. 2019, in preparation), we will further explore the properties of highly obscured AGNs and their host galaxies by combining the wealth of multiwavelength data in the CDFs, aiming at testing whether highly obscured AGNs are the missing link in the merger-triggered AGN-galaxy coevolution theory (Sanders et al. 1988).
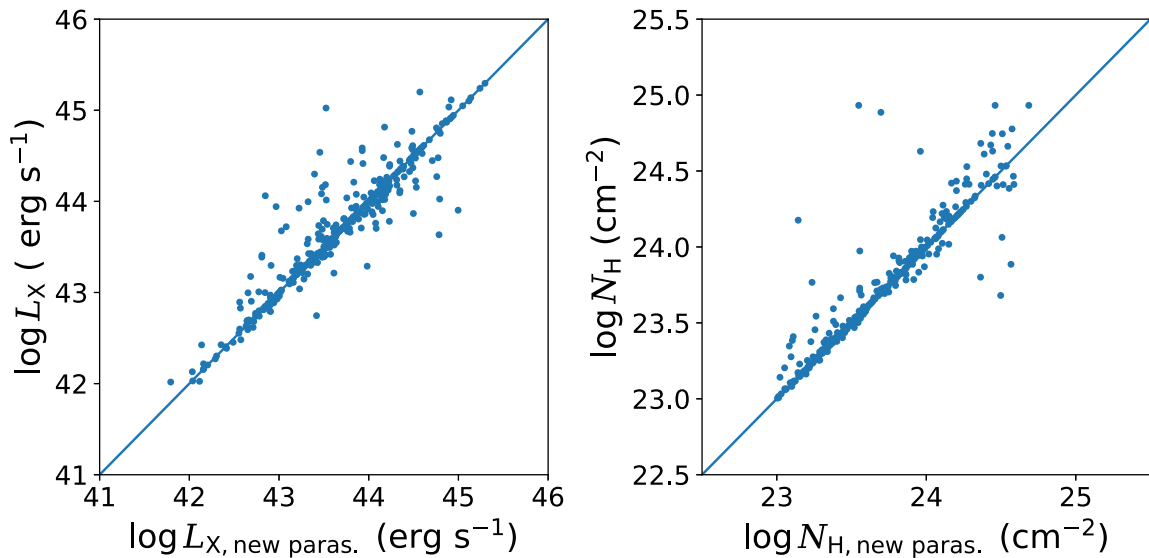
Figure 21. Comparison between the spectral fitting results obtained in Section 3 and those obtained in the Appendix using the new model parameters.

Appendix
Justification of the Adopted Models and Parameters

In the MYTorus configuration, since the reprocessed components strongly depend on the inclination angle θ, the equatorial N_H, and the assumed f_{ref}, the best-fit results will also depend on the assumed geometries and parameters. However, due to the low S/N for most of the spectra, we have to make prior assumptions about the input parameters and simply fix them to values for which their representativeness has not yet been physically validated. Here we justify that our spectral fitting results are not significantly influenced by the assumed parameters.

We first fit the spectra for the high-count sources (counts >200) with θ and f_{ref} set as free parameters. The best-fitted θ peaks at $\theta < 70°$; therefore, this time we choose to fix it at 65° for all sources. Then, we refit the spectra by keeping f_{ref} free and obtain the best-fit results. For the low-count sources, we fix f_{ref} at 2.0 for sources with detected reprocessed components, which is roughly the average value for high-count sources.

The comparison of the results with those obtained in Section 3.1 is shown in Figure 21. Despite small scatters, the results are consistent. Therefore, our simple choices of the fixed parameters are reasonable. The consistency between the MYTorus results and the Borus results obtained in Section 5.3 also validates our spectral fitting strategy.

ORCID iDs

Junyao Li ⦿ https://orcid.org/0000-0002-1605-915X
Yongquan Xue ⦿ https://orcid.org/0000-0002-1935-8104
Mouyuan Sun ⦿ https://orcid.org/0000-0002-0771-2153
Teng Liu ⦿ https://orcid.org/0000-0002-2941-6734
Fabio Vito ⦿ https://orcid.org/0000-0003-0680-9305
William N. Brandt ⦿ https://orcid.org/0000-0002-0167-2453
Guang Yang ⦿ https://orcid.org/0000-0001-8835-7722
Paolo Tozzi ⦿ https://orcid.org/0000-0003-3096-9966
Shifu Zhu ⦿ https://orcid.org/0000-0002-1653-4969
Xuechen Zheng ⦿ https://orcid.org/0000-0002-1564-0436
Bin Luo ⦿ https://orcid.org/0000-0002-9036-0063
Chien-Ting Chen ⦿ https://orcid.org/0000-0002-4945-5079
Cristian Vignali ⦿ https://orcid.org/0000-0002-8853-9611
Roberto Gilli ⦿ https://orcid.org/0000-0001-8121-6177
Xinwen Shu ⦿ https://orcid.org/0000-0002-7020-4290

References

Aird, J., Coil, A. L., Georgakakis, A., Nandra, K., & Barro, G. 2015, MNRAS, 1892, 451
Akylas, A., Georgakakis, A., Georgantopoulos, I., Brightman, M., & Nandra, K. 2012, A&A, 546, A98
Akylas, A., Georgantopoulos, I., Ranalli, P., et al. 2016, A&A, 594, A73
Alexander, D., & Hickox, R. 2012, NewAR, 56, 93
Alexander, D. M., Bauer, F. E., Brandt, W. N., et al. 2003, AJ, 126, 539
Alexander, D. M., Bauer, F. E., Brandt, W. N., et al. 2011, ApJ, 738, 44
Alexander, D. M., Chary, R., Pope, A., et al. 2008, ApJ, 687, 835
Allevato, V., Paolillo, M., Papadakis, I., & Pinto, C. 2013, ApJ, 771, 9
Ballantyne, D. R., Draper, A. R., Madsen, K. K., Rigby, J. R., & Treister, E. 2011, ApJ, 736, 56
Baloković, M., Brightman, M., Harrison, F. A., et al. 2018, ApJ, 854, 42
Baumgartner, W. H., Tueller, J., Markwardt, C. B., et al. 2013, ApJS, 207, 19
Bianchi, S., Guainazzi, M., & Chiaberge, M. 2006, A&A, 448, 499
Boissay, R., Ricci, C., & Paltani, S. 2016, A&A, 588, A70
Brandt, W. N., & Alexander, D. M. 2015, A&ARv, 23, 1
Brightman, M., & Nandra, K. 2011, MNRAS, 413, 1206
Brightman, M., Nandra, K., Salvato, M., et al. 2014, MNRAS, 443, 1999
Brightman, M., & Ueda, Y. 2012, MNRAS, 423, 702
Broos, P. S., Townsley, L. K., Feigelson, E. D., et al. 2010, ApJ, 714, 1582
Buchner, J., Georgakakis, A., Nandra, K., et al. 2015, ApJ, 802, 89
Burlon, D., Ajello, M., Greiner, J., et al. 2011, ApJ, 728, 58
Cash, W. 1979, ApJ, 228, 939
Comastri, A. 2004, in Supermassive Black Holes in the Distant Universe, 308, ed. A. J. Barger (Dordrecht: Kluwer), 245
Comastri, A., Ranalli, P., Iwasawa, K., et al. 2011, A&A, 526, L9
Corral, A., Della Ceca, R., Caccianiga, A., et al. 2011, A&A, 530, A42
Corral, A., Georgantopoulos, I., Watson, M. G., et al. 2014, A&A, 569, A71
Daddi, E. 2007, ApJ, 670, 173
Del Moro, A., Alexander, D. M., Bauer, F. E., et al. 2016, MNRAS, 456, 2105
Di Matteo, T., Springel, V., & Hernquist, L. 2005, Natur, 433, 604
Fanali, R., Caccianiga, A., Severgnini, P., et al. 2013, MNRAS, 433, 648

Feruglio, C., Daddi, E., Fiore, F., et al. 2011, ApJL, 729, L4
Fiore, F., Brusa, M., Cocchia, F., et al. 2003, A&A, 409, 79
Fiore, F., Puccetti, S., Brusa, M., et al. 2009, ApJ, 693, 447
Fiore, F., Puccetti, S., Grazian, A., et al. 2012, A&A, 537, A16
Georgakakis, A., Aird, J., Buchner, J., et al. 2015, MNRAS, 453, 1946
Georgantopoulos, I., Akylas, A., Georgakakis, A., & Rowan-Robinson, M. 2009, A&A, 507, 747
Georgantopoulos, I., Comastri, A., Vignali, C., et al. 2013, A&A, 555, A43
Gilli, R., Comastri, A., & Hasinger, G. 2007, A&A, 463, 79
Gilli, R., Su, J., Norman, C., et al. 2011, ApJL, 730, L28
González-Martín, O. 2018, ApJ, 858, 2
González-Martín, O., & Vaughan, S. 2012, A&A, 544, A80
Goulding, A. D., Alexander, D. M., Mullaney, J. R., et al. 2011, MNRAS, 411, 1231
Guainazzi, M., Matt, G., & Perola, G. C. 2005, A&A, 444, 119
Harrison, C. M., Alexander, D. M., Mullaney, J. R., et al. 2012, ApJL, 760, L15
Hasinger, G. 2008, A&A, 490, 905
Hickox, R. C., & Alexander, D. M. 2018, ARA&A, 56, 625
Hopkins, P., & Hernquist, L. 2006, ApJS, 166, 1
Iwasawa, K., Gilli, R., Vignali, C., et al. 2012, A&A, 546, A84
Kawamuro, T., Ueda, Y., Tazaki, F., Ricci, C., & Terashima, Y. 2016, ApJS, 225, 14
Kocevski, D. D., Brightman, M., Nandra, K., et al. 2015, ApJ, 814, 104
Kocevski, D. D., Hasinger, G., Brightman, M., et al. 2018, ApJS, 236, 48
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511
Koshida, S., Minezaki, T., Yoshii, Y., et al. 2014, ApJ, 788, 159
Koss, M. J., Assef, R., Baloković, M., et al. 2016, ApJ, 825, 85
Lansbury, G. B., Alexander, D. M., Aird, J., et al. 2017, ApJ, 846, 20
Lanzuisi, G., Civano, F., Elvis, M., et al. 2013, MNRAS, 431, 978
Lanzuisi, G., Civano, F., Marchesi, S., et al. 2018, MNRAS, 480, 2578
Lanzuisi, G., Ranalli, P., Georgantopoulos, I., et al. 2015, A&A, 573, 137
Liu, T., Tozzi, P., Wang, J. X., et al. 2017, ApJS, 451, 457
Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2011, ApJ, 740, 37
Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, ApJS, 228, 2
Lusso, E., Hennawi, J. F., Comastri, A., et al. 2013, ApJ, 777, 86
Marchesi, S., Ajello, M., Zhao, X., et al. 2019, ApJ, 872, 8
Marchesi, S., Lanzuisi, G., Civano, F., et al. 2016, ApJ, 830, 100
McHardy, I. M., Koerding, E., Knigge, C., Uttley, P., & Fender, R. P. 2006, Natur, 444, 730
Miyaji, T., Hasinger, G., Salvato, M., et al. 2015, ApJ, 804, 104
Morganti, R. 2017, NatAs, 1, 596
Murphy, K. D., & Yaqoob, T. 2009, MNRAS, 397, 1549
Paolillo, M., Papadakis, I., Brandt, W. N., et al. 2017, MNRAS, 471, 4398
Ranalli, P., Comastri, A., Vignali, C., et al. 2013, A&A, 555, A42
Ricci, C., Bauer, F. E., Treister, E., et al. 2017, MNRAS, 468, 1273
Ricci, C., Ueda, Y., Koss, M. J., et al. 2015, ApJL, 815, L13
Risaliti, G., Maiolino, R., & Salvati, M. 1999, ApJ, 522, 157
Sanders, D. B., Soifer, B. T., Elias, J. H., et al. 1988, ApJ, 325, 74
Silverman, J. D., Mainieri, V., Salvato, M., et al. 2010, ApJS, 191, 124
Soldi, S., Beckmann, V., Baumgartner, W. H., et al. 2014, A&A, 563, A57
Stark, A. A., Gammie, C. F., & Wilson, R. 1992, ApJS, 77, 79
Tanimoto, A., Ueda, Y., Kawamuro, T., et al. 2018, ApJ, 853, 146
Tozzi, P., Gilli, R., Mainieri, V., et al. 2006, A&A, 451, 457
Treister, E., Urry, C. M., & Virani, S. 2009, ApJ, 696, 110
Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, ApJ, 786, 104
Ueda, Y., Hashimoto, Y., Ichikawa, K., et al. 2015, ApJ, 815, 1
Vaughan, S., Edelson, R., Warwick, R. S., & Uttley, P. 2003, MNRAS, 345, 1271
Vito, F., Brandt, W. N., Yang, G., et al. 2018, MNRAS, 473, 2378
Wang, J.-M., Watarai, K.-Y., & Mineshige, S. 2004, ApJL, 607, L107
Wang, J. X., Malhotra, S., Rhoads, J. E., & Norman, C. A. 2004, ApJL, 612, L109
Xue, Y. Q. 2017, NewAR, 79, 59
Xue, Y. Q., Luo, B., Bauer, F. E., et al. 2011, ApJS, 195, 10
Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2016, ApJS, 224, 15
Yang, G., Brandt, W. N., Luo, B., et al. 2016, ApJ, 831, 145
Zhang, F., Yu, Q., & Lu, Y. 2017, ApJ, 845, 88
Zhao, X., Marchesi, S., Ajello, M., et al. 2019, ApJ, 870, 60
Zheng, X. C., Xue, Y. Q., Brandt, W. N., et al. 2017, ApJ, 849, 127