CrossMark

X-Ray Spectral Analyses of AGNs from the 7Ms *Chandra* Deep Field-South Survey: The Distribution, Variability, and Evolutions of AGN Obscuration

Teng Liu (刘腾)[1,2,3], Paolo Tozzi[4],
Jun-Xian Wang (王俊贤)[1,2], William N. Brandt[5,6,7], Cristian Vignali[8,9],
Yongquan Xue (薛永泉)[1,2], Donald P. Schneider[5,6], Andrea Comastri[9], Guang Yang[5,6], Franz E. Bauer[10,11,12], Maurizio Paolillo[13],
Bin Luo[14], Roberto Gilli[9], Q. Daniel Wang[3], Mauro Giavalisco[3], Zhiyuan Ji[3], David M Alexander[15], Vincenzo Mainieri[16],
Ohad Shemmer[17], Anton Koekemoer[18], and Guido Risaliti[4]

[1] CAS Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, Hefei 230026, China
lewtonstein@gmail.com, jxw@ustc.edu.cn
[2] School of Astronomy and Space Science, University of Science and Technology of China, Hefei 230026, China
[3] Astronomy Department, University of Massachusetts, Amherst, MA 01003, USA
[4] Istituto Nazionale di Astrofisica (INAF)—Osservatorio Astrofisico di Firenze, Largo Enrico Fermi 5, I-50125 Firenze, Italy; ptozzi@arcetri.astro.it
[5] Department of Astronomy and Astrophysics, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA
[6] Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
[7] Department of Physics, 104 Davey Lab, The Pennsylvania State University, University Park, PA 16802, USA
[8] Dipartimento di Fisica e Astronomia, Alma Mater Studiorum, Università degli Studi di Bologna, Viale Berti Pichat 6/2, I-40127 Bologna, Italy
[9] INAF—Osservatorio Astronomico di Bologna, Via Gobetti 93/3, I-40129, Bologna, Italy
[10] Instituto de Astrofísica and Centro de Astroingeniería, Facultad de Física, Pontificia Universidad Católica de Chile, Casilla 306, Santiago 22, Chile
[11] Millennium Institute of Astrophysics (MAS), Nuncio Monseñor Sótero Sanz 100, Providencia, Santiago, Chile
[12] Space Science Institute, 4750 Walnut Street, Suite 205, Boulder, CO 80301, USA
[13] Dip. di Fisica, Universitá di Napoli Federico II, C.U. di Monte Sant'Angelo, Via Cintia ed. 6, I-80126 Naples, Italy
[14] School of Astronomy & Space Science, Nanjing University, Nanjing 210093, China
[15] Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham, DH1 3LE, UK
[16] European Southern Observatory, Karl-Schwarzschild-Straße 2, D-85748, Garching bei München, Germany
[17] Department of Physics, University of North Texas, Denton, TX 76203, USA
[18] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

Abstract

We present a detailed spectral analysis of the brightest active galactic nuclei (AGNs) identified in the 7Ms *Chandra* Deep Field-South (CDF-S) survey over a time span of 16 years. Using a model of an intrinsically absorbed power-law plus reflection, with possible soft excess and narrow Fe Kα line, we perform a systematic X-ray spectral analysis, both on the total 7Ms exposure and in four different periods with lengths of 2–21 months. With this approach, we not only present the power-law slopes, column densities N_H, observed fluxes, and absorption-corrected 2–10 keV luminosities L_X for our sample of AGNs, but also identify significant spectral variabilities among them on timescales of years. We find that the N_H variabilities can be ascribed to two different types of mechanisms, either flux-driven or flux-independent. We also find that the correlation between the narrow Fe line EW and N_H can be well explained by the continuum suppression with increasing N_H. Accounting for the sample incompleteness and bias, we measure the intrinsic distribution of N_H for the CDF-S AGN population and present reselected subsamples that are complete with respect to N_H. The N_H-complete subsamples enable us to decouple the dependences of N_H on L_X and on redshift. Combining our data with those from C-COSMOS, we confirm the anticorrelation between the average N_H and L_X of AGN, and find a significant increase of the AGN-obscured fraction with redshift at any luminosity. The obscured fraction can be described as $f_{obscured} \approx 0.42 \, (1 + z)^{0.60}$.

Key words: catalogs – galaxies: active – galaxies: evolution – surveys – X-rays: galaxies

Supporting material: machine-readable tables

1. Introduction

Active galactic nuclei (AGNs) are important for understanding the formation and evolution of galaxies. It is now well established that the large majority of galaxies experience periods of nuclear activity, as witnessed by the ubiquitous presence of supermassive black holes (SMBHs) in their bulges (e.g., Kormendy & Ho 2013). A privileged observational window to select and characterize AGN is the 0.5–10 keV X-ray band, which has become particularly effective thanks to the advent of revolutionary X-ray facilities in the last 16 years, such as *Chandra* and *XMM-Newton*. Although only 5%–10% of the total nuclear emission emerges in the X-ray band, the relative strength of X-ray to other band (optical, infrared, radio)

emission in AGN is much higher than that in stars. This trait allows one to identify AGN out to very high redshift in deep high-resolution surveys. At least to first order, the majority of AGN spectra can be well described by an intrinsic power-law undergoing photoelectric absorption and Compton scattering by line-of-sight obscuring material, an unabsorbed power-law produced by scattering from surrounding ionized material, and a reflection component from surrounding cold material. These features make X-ray spectral analysis a powerful tool to measure the accretion properties and the surrounding environment of SMBHs. Therefore, tracing the X-ray evolution of AGN across cosmic epochs is crucial to reconstruct the cosmic history of accretion onto the SMBH and the properties of the host galaxy at the same time.

In this framework, significant results have been obtained thanks to a number of high-sensitivity X-ray surveys with large and medium sky coverage, such as C-COSMOS (Elvis et al. 2009; Lanzuisi et al. 2013), XMM-COSMOS (Hasinger et al. 2007), COSMOS-Legacy (Civano et al. 2016; Marchesi et al. 2016), CDF-S (Giacconi et al. 2002; Luo et al. 2008; Xue et al. 2011; Luo et al. 2017), CDF-N (Brandt et al. 2001; Alexander et al. 2003; Xue et al. 2016), Extended CDF-S (Lehmer et al. 2005; Virani et al. 2006; Xue et al. 2016), AEGIS-X (Laird et al. 2009), XMM-LSS (Pierre et al. 2007, 2016), and the XMM survey of CDF-S (Comastri et al. 2011; Ranalli et al. 2013). Among this set, the CDF-S survey, which recently reached a cumulative exposure time of 7Ms, represents the deepest observation of the X-ray sky obtained as of today and in the foreseeable future (Luo et al. 2017). Despite its small solid angle (484 arcmin^2), the CDF-S is the only survey that enables the characterization of low-luminosity and high-redshift X-ray sources.

The 7Ms CDF-S data have been collected across the entire lifespan of the *Chandra* satellite (1999–2016). Several groups have already used the CDF-S data to provide systematic investigations of the X-ray properties of AGN (e.g., Rosati et al. 2002; Paolillo et al. 2004; Saez et al. 2008; Raimundo et al. 2010; Alexander et al. 2011; Comastri et al. 2011; Luo et al. 2011; Rafferty et al. 2011; Lehmer et al. 2012; Young et al. 2012; Castelló-Mor et al. 2013; Vito et al. 2013, 2016). Of particular interest to this work, Tozzi et al. (2006) presented the first systematic X-ray spectral analysis of the CDF-S sources on the basis of the first 1Ms exposure using traditional spectral fitting techniques. Based on the 4Ms CDF-S data, Buchner et al. (2014) performed spectral analysis on the AGNs with a different approach. They developed a Bayesian framework for model comparison and parameter estimation with X-ray spectra, and used it to select among several different spectral models the one that best represents the data. Other investigations focused on the spectral analysis of specific X-ray source subpopulations, such as normal galaxies (Lehmer et al. 2008; Vattakunnel et al. 2012; Lehmer et al. 2016), high-redshift AGN (Vito et al. 2013), or single sources (Norman et al. 2002). The CDF-S field has also been observed for 3Ms with *XMM-Newton* (e.g., Comastri et al. 2011; Ranalli et al. 2013). However, we limit this work to the 7Ms *Chandra* data, because the much higher spatial resolution of *Chandra* compared with *XMM-Newton* is essential in resolving high-redshift sources, identifying multiband counterparts, and eliminating contamination from nearby sources; it also brings about a high spectral signal-to-noise ratio (S/N) by minimizing noise in source-extraction regions.

Among the most relevant parameters shaping the X-ray emission from AGN, the equivalent hydrogen column density N_H represents the effect of the photoelectric absorption (mostly due to the metals present in the obscuring material, which are implicitly assumed to have solar metallicity) and Thompson scattering on the intrinsic power-law emission. Generally, the obscuring material is related to the pc-scale dusty torus, which produces the highest N_H values, or to the diffuse interstellar medium (ISM) in the host galaxy, which can also create N_H as high as $\sim 10^{22-23.5}$ cm^{-2} (e.g., Simcoe et al. 1997; Goulding et al. 2012; Buchner & Bauer 2017). It is also found that 100 pc-scale dust filaments, which might be the nuclear fueling channels, could also be responsible for the obscuration (Prieto et al. 2014). The presence of the obscuring material is likely related to both the fueling of AGN from the host galaxy and the AGN feedback to the host galaxy. The geometry of the absorbing material is another relevant factor. Particularly, the orientation along the line of sight plays a key role in the unification model of AGN (Antonucci 1993; Netzer 2015). However, the observed correlation with star formation (e.g., Page et al. 2004; Alexander et al. 2005; Stevens et al. 2005; Chen et al. 2015; Ellison et al. 2016) indicates that AGN obscuration can be related to a phase of the coevolution of galaxies and their SMBHs (Hopkins et al. 2006; Alexander & Hickox 2012), rather than just due to an orientation effect. Morphological studies of AGN host galaxies show that highly obscured AGNs tend to reside in galaxies undergoing dynamical compaction (Chang et al. 2017) or galaxies exhibiting interaction or merger signatures (Kocevski et al. 2015; Lanzuisi et al. 2015). For a particular Compton-thick QSO at redshift 4.75, Gilli et al. (2014) found that the heavy obscuration could be attributed to a compact starburst region. In brief, the intrinsic obscuration of a given AGN does not have a simple and immediate physical interpretation because of its complex origins. A thorough understanding of the distribution of AGN obscuration and its dependence on the intrinsic (absorption-corrected) luminosity and on cosmic epoch is mandatory to understanding, at least statistically, the nature and properties of the emission mechanism, the AGN environment, the coevolution of AGN and the host galaxy, and the synthesis of the cosmic X-ray background (Gilli et al. 2007).

There have been several attempts to measure the N_H distribution of AGN in the pre-*Chandra* era (e.g., Maiolino et al. 1998; Bassani et al. 1999; Risaliti et al. 1999); however, their results were severely limited by the X-ray data. Thanks to the excellent performance of *Chandra*, Tozzi et al. (2006) corrected for both incompleteness and sampling-volume effects of the 1Ms CDF-S AGN sample and recovered the intrinsic distribution of N_H of the CDF-S AGN population (log $L_X \lesssim 45$, $z \lesssim 4$) with high accuracy. They found an approximately log-normal distribution that peaks around 10^{23} cm^{-2} with a $\sigma \sim 1.1$ dex, not including the peak at low N_H (below 10^{20} cm^{-2}). Based on a local AGN sample detected by *Swift-BAT* in the 15–195 keV band, which is less biased against obscured AGN, Burlon et al. (2011) presented a similar N_H distribution that peaks between 10^{23} and $10^{23.5}$ cm^{-2}. Based on a mid-infrared 12 μm selected local AGN sample, which is even less biased against obscured AGN than the 15–195 keV hard X-ray emission, Brightman & Nandra (2011) reported a similar distribution with an obscured peak between 10^{23} and 10^{24} cm^{-2}. Using a large AGN sample selected from CDF-S, AEGIS-XD, COSMOS, and XMM-XXL surveys, Buchner et al. (2015) provided intrinsic N_H distributions in three segregated redshift intervals between redshift 0.5 and 2.1, and found a higher fraction of sources at $N_H \approx 10^{23}$ cm^{-2} when the redshift increases up to >1. There are other investigations that presented the observed N_H distribution of AGN, but without any correction for selection bias (e.g., Castelló-Mor et al. 2013; Brightman et al. 2014).

Many works have shown that the fraction of obscured AGNs declines at high X-ray luminosity (e.g., Lawrence & Elvis 1982; Treister & Urry 2006; Hasinger 2008; Brightman & Nandra 2011; Burlon et al. 2011; Lusso et al. 2013; Brightman et al. 2014). This behavior can be explained by a decreased covering factor of the obscuring material at high luminosity (Lawrence 1991; Lamastra et al. 2006; Maiolino et al. 2007), or

as a result of higher intrinsic luminosities in unobscured than in obscured AGNs (Lawrence & Elvis 2010; Burlon et al. 2011; Liu et al. 2014; Sazonov et al. 2015). However, other studies suggest that the relation between intrinsic absorption and luminosity is more complex, and may be non-monotonic. Some studies indicate that in the very low luminosity regime, the N_H distribution of AGN drops with decreasing luminosity (Elitzur & Ho 2009; Brightman & Nandra 2011; Burlon et al. 2011; Buchner et al. 2015). It has also been suggested that the obscured fraction rises again in the very high luminosity regime (Stern et al. 2014; Assef et al. 2015).

In general, the fraction of X-ray obscured AGN has been found by several studies to rise with redshift (La Franca et al. 2005; Ballantyne et al. 2006; Treister & Urry 2006; Tozzi et al. 2006; Hasinger 2008; Hiroi et al. 2012; Iwasawa et al. 2012; Brightman et al. 2014; Ueda et al. 2014; Vito et al. 2014; Buchner et al. 2015). However, no such evolution was found or attributed to biases in other investigations (Dwelly & Page 2006; Gilli et al. 2007, 2010; Lusso et al. 2013). The uncertainty is mainly caused by limited sample size and rough N_H measurement, which is often based upon the X-ray hardness ratio rather than spectral fitting. In particular, the strong dependence of the average N_H on luminosity places a large obstacle in identifying any dependence of average N_H on redshift because of the strong L–z correlation of sources in a flux-limited sample. A sizable sample with wide dynamical ranges in luminosity and redshift that can be split into narrow luminosity and redshift bins while maintaining good count statistics is essential to distinguish any redshift dependence from the luminosity dependence.

The picture outlined here points toward a significant complexity, where different fueling mechanisms need to be invoked at different luminosities and different cosmic epochs. In this paper we exploit the 7Ms CDF-S—the deepest X-ray data ever obtained—to investigate the distribution of intrinsic absorption among AGN over a wide range of redshift and luminosity. In addition to the unprecedented X-ray survey depth, continuous multiband follow-ups of this field allow us to perform excellent AGN classification and redshift measurement (Luo et al. 2017), which are essential in measuring the N_H values of AGN and their distribution across the AGN population. We apply updated data-processing techniques to these data, and provide systematic spectral analyses of the AGNs. A few analysis methods that have been widely used in the past few years are applied, including astrometry correction on the data, refined selection of source and background extraction regions, a more elaborate spectral stacking method, and a more accurate spectral fitting statistic. The lengthy time interval (16 years) of the 7Ms exposure provides us long-term-averaged properties of the sources. To measure the AGN obscuration more accurately, we include variability in the spectral fitting strategy, which includes not only the variation of the intrinsic luminosity, but also the change in the obscuration on timescales of a few years (e.g., Risaliti et al. 2002; Yang et al. 2016). On shorter timescales, most of our sources have insufficient statistics to measure N_H accurately. Our final aim is to characterize the intrinsic distribution and evolution of AGN obscuration based on the systematic spectral analyses.

Throughout this paper, we adopt the *WMAP* cosmology, with $\Omega_m = 0.272$, $\Omega_\Lambda = 0.728$ and $H_0 = 70.4$ km s^{-1} Mpc^{-1} (Komatsu et al. 2011). All of the X-ray fluxes and luminosities

Table 1
7Ms CDF-S Observations Divided into Four Periods

Period	Observation Date	Time Span	Exposure Time
I	1999.10–2000.12	14 months	1Ms
11 ObsIDs:	1431-0 1431-1 441 582 2406 2405 2312 1672 2409 2313 2239		
II	2007.09–2007.11	2 months	1Ms
12 ObsIDs:	8591 9593 9718 8593 8597 8595 8592 8596 9575 9578 8594 9596		
III	2010.03–2010.07	4 months	2Ms
31 ObsIDs:	12043 12123 12044 12128 12045 12129 12135 12046 12047 12137 12138 12055 12213 12048 12049 12050 12222 12219 12051 12218 12223 12052 12220 12053 12054 12230 12231 12227 12233 12232 12234		
IV	2014.06–2016.03	21 months	3Ms
48 ObsIDs:	16183 16180 16456 16641 16457 16644 16463 17417 17416 16454 16176 16175 16178 16177 16620 16462 17535 17542 16184 16182 16181 17546 16186 16187 16188 16450 16190 16189 17556 16179 17573 17633 17634 16453 16451 16461 16191 16460 16459 17552 16455 16458 17677 18709 18719 16452 18730 16185		

quoted throughout this paper have been corrected for the Galactic absorption, which has a column density of 8.8×10^{19} cm^{-2} (Stark et al. 1992) in the CDF-S field.

2. Data Processing

2.1. CDF-S Observations

The 7Ms CDF-S survey is comprised of observations performed between 1999 October 14, and 2016 March 24, (UTC). Excluding one observation compromised by telemetry saturation and other issues (ObsID 581), there are 102 observations (observation IDs listed in Table 1) in the data set. The exposures collected across 16 years can be grouped into four distinct periods, each spanning 2–21 months. Figure 1 displays the distribution of the exposure and the four periods we identified. Because of the decline of quantum efficiency of the CCD, the average response at the aimpoint[19] changes considerably across the 16 years of operation, whereas it can be considered fairly constant within a single period ($<6\%$ variation), as shown in Figure 2. We consider the cumulative spectra of X-ray sources in each period in order to mitigate the effects of AGN variability on timescales of years (e.g., Paolillo et al. 2004; Vagnetti et al. 2016; Yang et al. 2016) and reduce the uncertainty of combining the time-dependent instrument calibrations.

2.2. Data Processing

All the data are processed with *CIAO*4.8 using the calibration release CALDB 4.7.0. The data are reduced using the CHANDRA_REPRO tool. For each observation, the absolute astrometry is refined by matching the coordinates of sources detected using *wavdetect* to the 100 brightest sources of the 4Ms CDF-S catalog (Xue et al. 2011), which have been aligned to VLA 1.4 GHz radio astrometric frame. We use a simple

---

[19] Here the average aimpoint response for each observation is defined as the average quantum efficiency [counts photon^{-1}], which is calculated across all the 1024×1024 pixels of CCD3, multiplied by the effective area [cm^2] at the aimpoint.
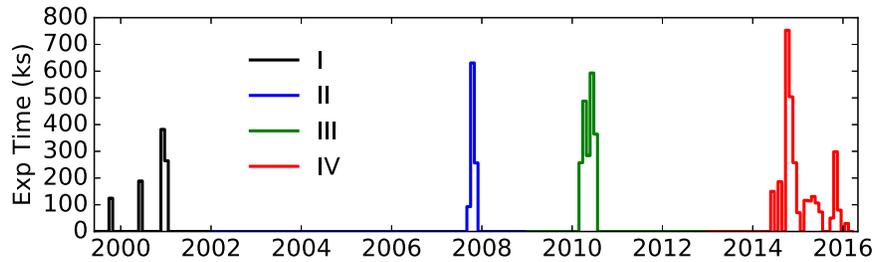
Figure 1. Color-coded histogram of the exposure time of the four periods of CDF-S observations as listed in Table 1. The bin size is 30 days. The approximate total exposure time of each period in sequence is 1Ms, 1Ms, 2Ms, and 3Ms.

iterative sigma-clipping routine (the CIAO task DEFLARE) to detect and remove background flares from the data of each CCD chip in each observation. For the VFAINT-mode exposures (92 out of 102) we apply the standard VFAINT background cleaning to remove the "bad" events, which are most likely associated with cosmic rays. This cleaning procedure could remove some real X-ray events as background in the case of bright unresolved sources. We check such an effect on the 10 brightest sources, and find that the loss of net counts is less than 2%. Finally, the exposures are combined using the FLUX_OBS task to create stacked images and exposure maps in the soft (0.5–2 keV) and hard (2–7 keV) bands, respectively. Exposure maps are computed for a monochromatic energy of 1.5 and 3.8 keV for the soft and hard bands, respectively.

2.3. Spectra Extraction

Our spectral analysis is based on the 7Ms CDF-S point-source catalog that includes 1008 X-ray sources (Luo et al. 2017). To optimize the source-extraction region, we generate an accurate point-spread function (PSF) image at the position of each source for each exposure using the ray trace simulation tool SAOTRACE, and measure the 94% energy-enclosed contour at 2.3 keV (the effective energy for the 0.5–7 keV band). In the cases where the extraction regions of two or more nearby sources overlap, the enclosed energy fraction is reduced to separate the regions. For some sources that lie inside an extended source (e.g., Finoguenov et al. 2015) or in a very crowded region, where the extraction regions significantly overlap, we reduce the source-extraction region manually. To prevent exceptionally large extraction regions at CCD gaps and borders where the PSF is distorted by the nonuniform local exposure map, the extraction region is confined to the 95% energy-enclosed circle measured with the CIAO PSF task. The loss of the source flux caused by the extraction region is recovered by applying an energy-dependent aperture correction to the spectral ancillary response files.

We define a background extraction region as an annulus around each source. To select the inner circle, which is used to mask the source signal, we measure the 97% energy-enclosed radius R_{97} at the position of each source with the PSF tool. At radii larger than R_{97} the *Chandra* PSF is highly diffused, and a negligible amount of signal from the source falls into the annular background region surrounding the source-extraction region. Only in some cases, when the source is very bright or located far from the aimpoint or in very crowded region, do we have to use a larger (a factor of 1.2–2) inner radius to make sure the background annulus is free of source signal. Moreover, we manually mask visible diffuse emission from the background measurement, consulting the extended CDF-S source catalog

presented by Finoguenov et al. (2015). After all the source signals are masked as above, we select the outer radius for each source i according to the total effective area in the source-extraction region $\int A_{i,\mathrm{src}}$. The background regions are not necessarily complete annuli; they could be broken by the mask of nearby sources. The background-region size is determined by the "backscal" parameter, which is defined as the ratio between the total effective areas in the source region and in the background region $\int A_{i,\mathrm{src}} \big/ \int A_{i,\mathrm{bkg}}$. Effective areas are computed for each exposure at 2.3 keV. We chose a backscal for each source by iteration such that the background annular region determined by this backscal includes a total of ~1000 photons in the 7Ms exposure in the total (0.5–7 keV) band. For each source, the source and background regions vary among the observations, while the backscal remains approximately constant. An upper limit of 30″ is set to the outer radius to prevent exceptionally large background regions.

Calibration files, i.e., response matrix files (RMF) and ancillary response files (ARF), are generated for each source in each exposure. It is often the case that a source in CDF-S is only visible after stacking multiple observations and may not have any photons (even background photon) within the extraction region for a given exposure. This is relevant for the majority of the sources, especially for faint source with fewer than 102 (total number of observations) net counts lying at the aimpoint of the field of view (FOV). Such an exposure is discarded when extracting spectra, background, and calibration files for this particular source; however, its exposure time is retained, see Section 2.4 for details. Although all the AGNs are expected to be unresolved in X-ray, some sources with off-axis angles >5′[20] and at least five photons in the source-extraction region in 0.5–7 keV band are treated as extended when creating the response files, by weighting the effective area for the soft photon distributions across their large extraction region. Although in most cases this treatment has a minor effect, it is technically more valid. In the case of hard sources without any soft photons in a specific ObsID, the hard-band counts are used as weights. An energy-dependent aperture correction is applied to the ARF with the *arfcorr* task.

2.4. Spectra for Combined Analysis

The net counts of the CDF-S sources span a wide range, as shown in Figure 3. Given the large number of individual exposures, the average number of net counts in one specific exposure is extremely low. It is thus meaningless to perform a joint analysis keeping spectra and calibration files from all individual exposures for the bulk of the sample. As we are interested in the spectral analysis of the largest number of

---

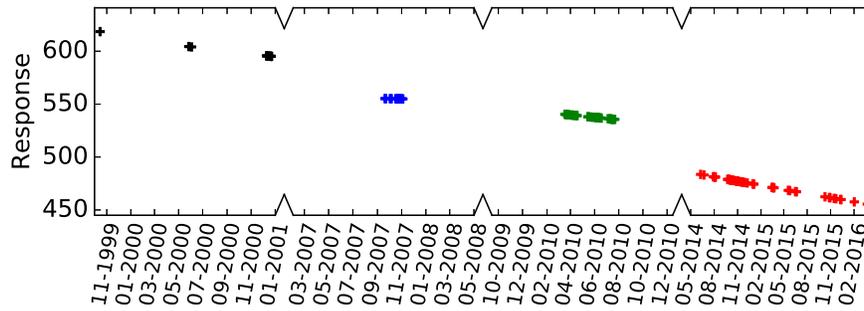[20] Sources with off-axis angles >5′ have 95% energy-enclosed PSF radii $\gtrsim 5''$.

Figure 2. Average aimpoint response [cm^2 counts/photon] of ACIS-I at 1.5 keV as a function of observation date. The quantum efficiency of the ACIS-I CCDs declined at most by 26%.
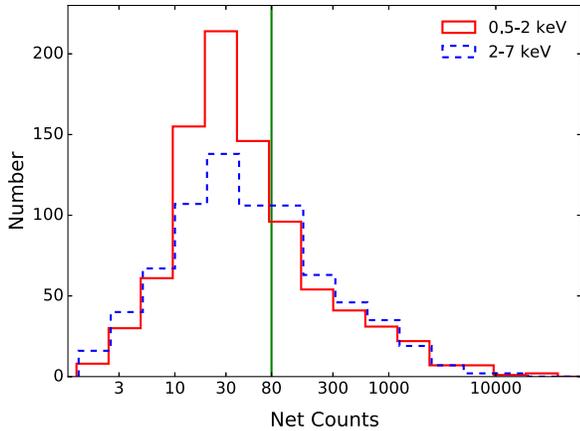

Figure 3. Distributions of *Chandra* net counts of the AGNs within the extraction radius in the soft and hard bands. The green vertical line corresponds to 80 net counts.

sources, we choose to combine the spectra within each of the four periods and within the total 7Ms exposure, generating five sets of stacked spectra and averaged response files.

The PHA spectral files of the source and background are simply stacked using the *FTOOLS* task *mathpha*. Since the backscal parameter is determined in advance and remains approximately the same in all the ObsIDs for each particular source, we set the backscal of the stacked spectra as the counts-weighted mean of the backscal of each ObsID. The exposure time of the stacked background PHA is directly calculated by summing the exposure time of each single background spectrum, while for the exposure time of the stacked source PHA, we have to account for the ObsIDs where there is no photon recorded in the source-extraction region; although such ObsIDs do not contribute any signal to the stacked spectrum. The exposure time T_j of such an ObsID j is normalized to the mean effective area of the source and then added into the total exposure time as follows: first, we measure $\bar{A}_j$, the mean effective area inside the extraction region of the source at the effective energy of the broad band, 2.3 keV, for each ObsID j. Of the ObsIDs where a signal is recorded within the source-extraction region, we calculate the mean effective area of the source $\langle \bar{A}_j \rangle$. Then, we multiply T_j by $\bar{A}_j / \langle \bar{A}_j \rangle$ and add it to the total exposure time.

To compute the averaged RMF and ARF, we consider only the ObsIDs where there is at least one photon within the source-extraction region. We simply use the broadband photon counts to weight the RMF, while for the ARF, which is more variable because of the vignetting effect and the long-term

degeneration of CCD quantum efficiency (see Figure 2), we use a weight of $C_j/\bar{A}_j$, where C_j is the broadband photon counts in ObsID j, and $\bar{A}_j$ is the mean effective area inside the extraction region at the effective energy of the broadband, 2.3 keV. This choice leads to the most accurate average flux measurement, taking the flux variation of the AGN into account, as explained in detail in the Appendix.

2.5. Sample Selection

In this work, we focus on the spectra of the AGNs of the 7Ms CDF-S main-source catalog. As reported in Luo et al. (2017), an AGN is selected if it satisfies any one of several criteria, including a high intrinsic X-ray luminosity (with $L_{0.5-7\,\mathrm{keV}} > 3 \times 10^{42}$ erg s^{-1}), a high ratio of X-ray flux to flux in other bands (optical, near-infrared, or radio), a hard X-ray spectrum (with an effective power-law slope $\Gamma < 1$, which is obtained without considering the intrinsic absorption of the AGN), and optical spectroscopic AGN features.

For each source, the net counts are measured from the 7Ms stacked source and background spectra. The distributions of the soft and hard-band net counts of AGNs in the 7Ms data are shown in Figure 3. The bulk of the spectra have fewer than 100 net counts, providing poor constraints on spectral parameters. To reach meaningful characterization of the largest number of AGN, we set a threshold on the net counts as low as possible to select a bright subsample for spectral analyses. To avoid possible bias induced by the low statistics, Tozzi et al. (2006) conservatively defined an X-ray bright sample suitable for spectral analysis by considering sources exceeding at least one of the following thresholds: 170 total counts, 120 soft counts, and 80 hard counts, based on the first 1Ms CDF-S stacked spectra. In this work, we select only sources with at least 80 net counts in the hard band, which is less affected by obscuration than the soft band. The 80 net counts threshold corresponds to a 2–7 keV flux of 2×10^{-16} erg cm^{-2} s^{-1} at the aimpoint of CDF-S (see Figure 23). This threshold appears more stringent than that in Tozzi et al. (2006); however, justified by the fact that the 7Ms background is about seven times higher, it is actually less stringent in terms of source-detection significance for the same number of net counts. In particular for faint off-axis sources with large extraction regions, the signal is background dominated. In addition, at large off-axis angles where the PSF and effective exposure change dramatically, the S/N is severely reduced. With this selection threshold, we are sampling AGNs in the luminosity range where most of the emission due to the cosmic accretion onto an SMBH is produced (see Section 5.2). To keep our sample as large as possible, we exclude only the FOV beyond a 9.′5 off-axis angle
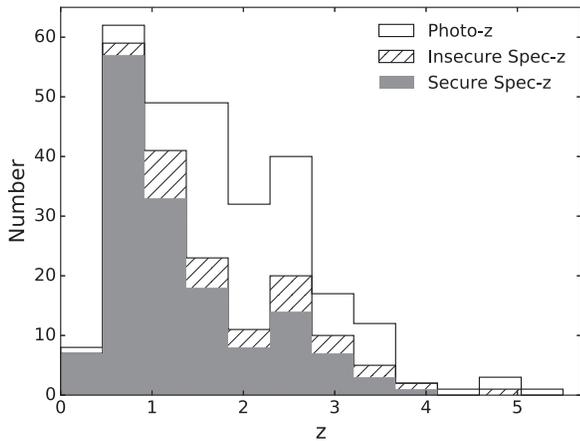
Figure 4. Redshift distribution of our sample.

from the aimpoint. Finally, we select 269 AGNs with at least 80 hard-band net counts and published redshift measurements. In addition to this main sample, we select a supplementary sample from the central region within an off-axis angle of $4'.5$ with at least 60 net counts in the hard band, in order to fully exploit the 7Ms CDF-S data. The supplementary sample, which contains seven AGNs, is only used in the reselected subsamples in Section 5.3, where the 80 net counts threshold becomes irrelevant.

Our final sample contains 276 AGNs, has a median redshift of 1.6, and a median number of 0.5–7 keV band net counts of 440. The redshift measurements are collected by Luo et al. (2017) from 25 spectroscopic-z catalogs and 5 photometric-z catalogs. They selected preferred redshifts carefully from different catalogs and demonstrated that the photometric-z measurements have a good quality by comparing the photometric-z to the available spectroscopic-z. Based on the detection of a narrow 6.4 keV Fe Kα line (see Section 3.3), we replace the photometric-z of 5 sources with our X-ray spectroscopic redshifts, which are considered insecure. Finally, of all the redshift measurements, 148 (54%) are secure spectroscopic redshifts, 31 (11%) are insecure spectroscopic redshifts, and 97 (35%) are photometric redshifts. As shown in Figure 4, photometric measurements mostly lie at relatively high redshift. We note that quite a number of the sources have their redshifts changed with respect to that used in Tozzi et al. (2006). See further comparison in Section 4.8.

3. Spectral Analysis

3.1. Spectral Fitting Method

The study of the deep X-ray sky necessarily requires the use of long exposures, often taken at different epochs, such as in the case of CDF-S. Clearly, in spectral analyses we must take significant variability in AGNs into account, which may reflect changes not only in the intrinsic luminosity, but also in the obscuration (e.g., Yang et al. 2016). In this work, which is aimed at exploiting the full statistics of the deep 7Ms exposure, we group ObsIDs which are close in time into four periods and check for significant variation between periods. To retain the energy resolution as much as possible, each stacked spectrum is grouped as mildly as possible so that each energy bin contains at least one photon (see the Appendix in Lanzuisi et al. 2013). We increase the grouping level (bin size) to speed

up the fitting only for the brightest sources in our sample. If a source has broadband total counts $N_{\rm tot} > 1000$, we group its spectrum to include at least $N_{\rm tot}/1000 + 1$ photons in each bin. The low-counts regime of our spectra requires use of the C statistic (Cash 1979; Nousek & Shue 1989) rather than χ^2.

With *Xspec* v12.9.0 (Arnaud 1996), we perform spectral analysis for each source following four different approaches:

A: fitting the background-subtracted spectrum stacked within each period independently.
B: fitting the 7Ms stacked, background-subtracted spectrum.
C: fitting the background-subtracted spectra stacked within each period simultaneously.
D: fitting the source and background spectra stacked within each period simultaneously.

For a source covered by all four periods, method B deals with one spectrum, methods A and C deal with four spectra, and method D with eight spectra. Methods A, B, and C make use of the standard C statistic. In a model comparison, the change in C statistic, ΔC, which approximately follows a χ^2 distribution, can be used as an indicator of the confidence level of the fitting improvement. Specifically, a model is providing a statistically significant improvement at a confidence level of 95% when the C statistic is reduced by $\Delta C > 3.84$ and $\Delta C > 5.99$ for one and two additional degrees of freedom (DOF), respectively. Method D, which models both the source and the background spectra, adopts a slightly different statistic, namely, the W statistic. This approach mitigates a weakness of the commonly used method of fitting a background-subtracted spectrum with the C statistic, which incorrectly assumes that the background-subtracted spectrum has a Poissonian error distribution. In this work, we use method D to obtain the final estimation of the parameters; the other three C statistic methods are used for different purposes, as described below, when the ΔC method is needed to evaluate the model improvement.

3.2. Spectral Models

3.2.1. Selection of our Models

The source spectral model is `wabs` * (`zwabs`*`powerlaw` + `zgauss` + `powerlaw` + `zwabs`*`pexrav`*`constant`); an illustration is given in the upper panel of Figure 5. The `wabs` (Wilms et al. 2000) accounts for the Galactic absorption, which is fixed at a column density of $8.8 \times 10^{19}\,{\rm cm}^{-2}$ (Stark et al. 1992). The model is composed of four additive components. The `zwabs`*`powerlaw` describes the primary power-law with intrinsic obscuration, that is, the cumulative effect of the absorbing material in the circumnuclear region and possibly in the host galaxy, expressed in equivalent hydrogen column density assuming solar metallicity. The component `zgauss` describes a Gaussian emission line with a zero width to fit an unresolved 6.4 keV Fe $K\alpha$ line when present. The second `powerlaw` is used for a soft-excess component, which is occasionally found in the soft band in addition to the primary power-law. A cold-reflection component is modeled with the `zwabs` * `pexrav` * `constant`, where the absorption is fixed to $10^{23}\,{\rm cm}^{-2}$, as discussed later. The four components do not always appear for each source; initially, only the primary power-law and the reflection are considered. The emission line and the additional power-law are included only if they are statistically required, as described in Section 3.3.
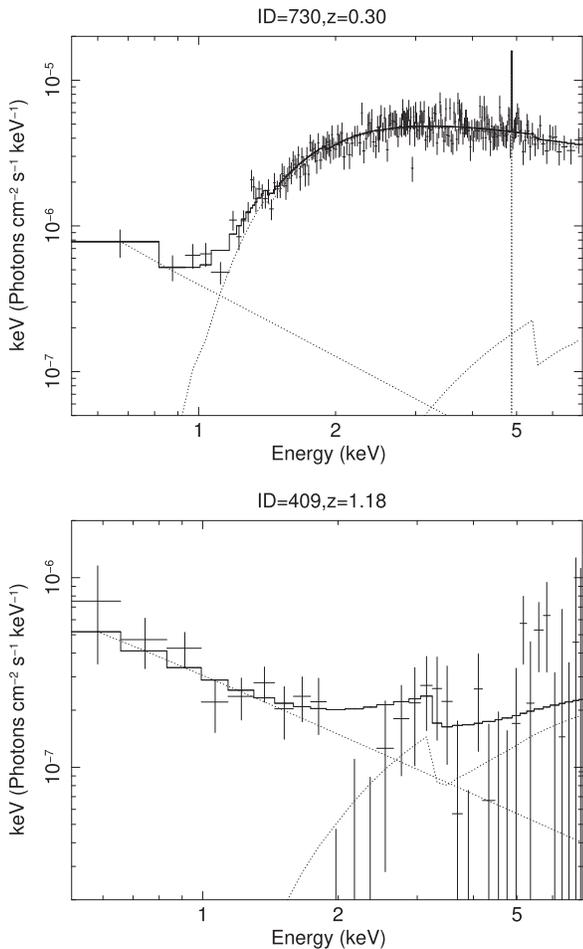
Figure 5. Two examples of the spectra. Upper panel: source 730 (with 8790 net counts in the 0.5–7 keV band) is a Compton-thin AGN fitted with the standard model, which is composed of an obscured power-law (the dominant component), a narrow Fe Kα line, a soft excess, and a `zwabs*pexrav` reflection. The reflection component only contributes a small fraction of the signal in the hard band. Lower panel: source 409 (with 440 net counts in the 0.5–7 keV band) is a Compton-thick AGN fitted with the Compton-thick model in which the relative strength of the reflection is set free. The reflection dominates the hard-band emission. Note that this Compton-thick model is only used in Section 3.4.1 to identify Compton-thick AGN.

The absorption model `zwabs` works well in the Compton-thin regime and has been widely used. However, it considers only photoelectric absorption but not Compton scattering, which starts to be relevant at $N_H >$ a few 10^{23} cm^{-2}. In order to identify Compton-thick AGNs and measure the N_H of highly obscured AGNs with more accuracy, we check how the shortage of the `zwabs` model affects the results by replacing the `zwabs*powerlaw` with `plcabs` (Yaqoob 1997), which describes X-ray transmission of an isotropic source located at the center of a uniform, spherical distribution of matter, correctly taking into account Compton scattering.

The slope and normalization of the `pexrav` component are linked to those of the primary power-law, and the cutoff energy is fixed at 300 keV. Although it has been found that the reflection strength is larger ($R \approx 2.2$) in highly obscured sources (10^{23}–10^{24} cm^{-2}) than in less obscured ones ($R \lesssim 0.5$, see Ricci et al. 2011), we fix the reflection scaling factor R at 0.5 for all the sources for simplicity. By definition, the R parameter regulates the relative strength of reflection to the primary power-law. However, we always fix R at a constant value and use the additional "constant" parameter to regulate the relative reflection strength for convenience. In the standard model, which is Compton-thin, this "constant" parameter is fixed at 1. It is only set free and used in identifying Compton-thick sources, as shown below, where the parameter can be large, indicating relatively strong reflection.

The X-ray reflected emission of AGN might arise from the accretion disk or the inner region of the dusty torus. Considering the realistic geometry of the torus and the AGN obscuration from torus-scale to galaxy-scale (Buchner & Bauer 2017), it is unlikely that all the reflected X-ray photons could leave the galaxy without any absorption, as expected by the flat-surface reflection model "pexrav." The X-ray reflection, if separated from the transmitted power-law as a stand-alone component, must be self-absorbed by the torus or obscured by material on a larger scale. In a physical torus model (e.g., Murphy & Yaqoob 2009; Brightman & Nandra 2011) that treats the absorption, scattering, and reflection self-consistently, the "self-absorption" of reflection is naturally considered. To obtain a simple rendition for a self-absorbed reflection model without adding any free parameters, we add an absorption to `pexrav`. This absorption is irrelevant to the absorption for the primary power-law, which corresponds to only the line-of-sight absorber and is highly dependent on the viewing orientation; it corresponds to a majority of the obscuring material in the galaxy, which must have a significant covering factor to the core, and is less variable among AGNs with different viewing orientations compared with the line-of-sight absorption. According to a comparison with the MYTorus model in Section 3.2.2, we set $^P N_H = 10^{23}$ cm^{-2} for the self-absorption. Therefore, even after adding this absorption, our reflection model still contains no free parameter.

In case of method D, the background is modeled with the `cplinear` model (Broos et al. 2010) plus two narrow Gaussian emission lines. The `cplinear` describes the background continuum by fluxes at 10 vertex energies. The vertex energies are selected dynamically between 0.5 and 7 keV, letting each segment contain the same number of photons.

The number of vertices is reduced in order to have at least 10 photons in each segment. The Gaussian lines describe the two most prominent instrument emission lines in the 0.5–7 keV band, one at 1.486 keV (Al Kα) and another between 2.1 and 2.2 keV (Au Mα, β). We first fit the background spectra with this model and then fix the background parameters at the best-fit values when fitting the source and background simultaneously.

3.2.2. Justification of our Models

Since the low S/N of our spectra does not allow us to constrain any parameter of the reflection component, we have to make proper assumptions about the model to describe a typical case of the reflected emission. The model we adopt, `zwabs*pexrav`, is an effective model that may not correspond to a realistic description of the torus. Here we test its validity by comparing it with a few physical models that describe the X-ray reprocessing considering more detailed torus structures (e.g., Murphy & Yaqoob 2009; Brightman & Nandra 2011).

First, we compare the spectral shape of our reflection model with that of the MYTorus model (Murphy & Yaqoob 2009), which provides the spectrum of the reflection component
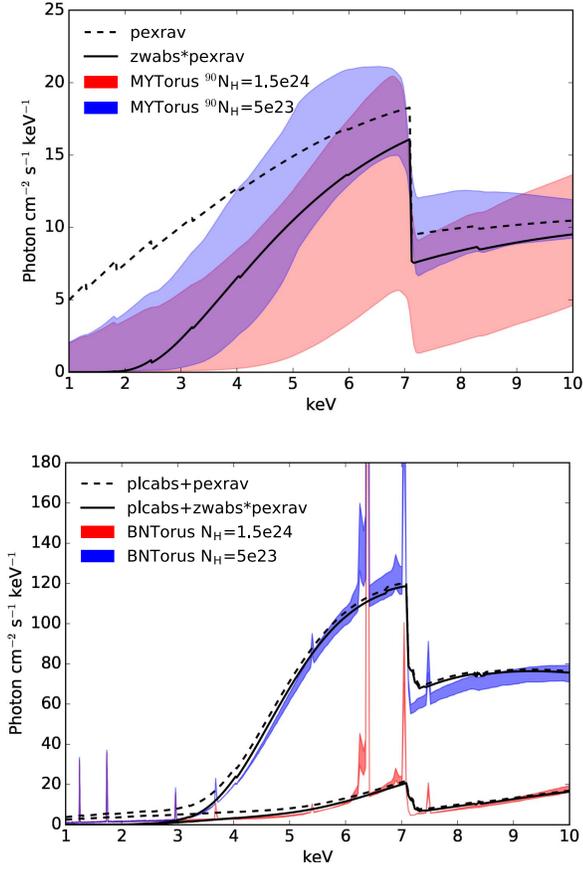
Figure 6. Upper panel: the reflection model adopted in this work (zwabs*pexrav with an $^P N_H$ of 10^{23} cm^{-2}, black solid line) compared with MYTorus reflection models (color filled regions) with diameter column densities of $^{90}N_H = 5 \times 10^{23}$ cm^{-2} (blue) and $^{90}N_H = 1.5 \times 10^{24}$ cm^{-2} (red) between inclination angles of $\theta = 0°$ (face-on) and $\theta = 90°$ (edge-on). The unabsorbed pexrav model is plotted with a dashed black line. All the models are derived from the same intrinsic primary power-law and have $z = 0$. Lower panel: the model of transmitted power-law plus reflection adopted in this work (plcabs+zwabs*pexrav, black solid lines) compared with BNTorus models (color filled regions) in two cases with $N_H = 5 \times 10^{23}$ cm^{-2} (blue) and $N_H = 1.5 \times 10^{24}$ cm^{-2} (red), respectively. The filled ranges correspond to torus opening angles between $\theta_{torus} = 30°$ and $\theta_{torus} = 60°$. Without the self-absorption, the plcabs+pexrav model is plotted with the black dashed line. All the models have $z = 0$. They are all derived from the same intrinsic primary power-law, but the normalizations of the plcabs-based models are multiplied by factors of 75% and 50% in the cases of $N_H = 5 \times 10^{23}$ cm^{-2} and $N_H = 1.5 \times 10^{24}$ cm^{-2}, respectively.

considering a toroidal torus structure. In the upper panel of Figure 6, we show the MYTorus reflection in two cases with the column densities through the diameter of the torus tube $^{90}N_H$ (not the line-of-sight N_H) of 5×10^{23} cm^{-2} (red lines) and 1.5×10^{24} cm^{-2} (blue) and with the inclination angles (between the line of sight and the symmetry axis of the torus) θ between 0° and 90°, respectively. The shape and strength of the reflection both depend significantly on $^{90}N_H$ and θ, having a large dynamical range. It is stronger when face-on than when edge-on and stronger with $^{90}N_H = 5 \times 10^{23}$ than with 1.5×10^{24} cm^{-2}. Without any self-absorption, the pexrav model is clearly softer than the MYTorus model. Adding the self-absorption of $^P N_H = 10^{23}$ cm^{-2} to pexrav, the shape of zwabs*pexrav is much more similar to that of MYTorus, and in both cases where $^{90}N_H = 5 \times 10^{23}$ cm^{-2} and $^{90}N_H = 1.5 \times 10^{24}$ cm^{-2}, it lies between the face-on ($\theta = 0°$) and edge-on ($\theta = 90°$) instances of MYTorus.

Therefore, with an order-of-magnitude estimate of $^P N_H = 10^{23}$ cm^{-2}, our reflection model can be considered as an intermediate instance of the various reflection models. According to the MYTorus model, the reflection is weaker at higher $^{90}N_H$ and larger θ, where the line-of-sight N_H would be higher. By adding the absorption to pexrav, the strength of the pexrav model is reduced by 32% in the 2–7 keV band. This weaker reflection setting in our model suggests that we are likely modeling the reflection in the high-N_H cases better than in the low-N_H cases. This is helpful to our aim of measuring N_H, since in low-N_H cases where the X-ray emission is largely dominated by the primary power-law, the reflection is not as significant as in the high-N_H cases.

Second, to check the relative strength of the reflection to the primary power-law of our model, we compare our spectral shape of primary power-law plus reflection with that of BNTorus model (Brightman & Nandra 2011), which provides the spectrum of the total transmitted and reflected emission considering a biconical torus structure. In the lower panel of Figure 6, we compare our model with the BNTorus model in two cases with N_H (independent of inclination angle) of 5×10^{23} cm^{-2} and 1.5×10^{24} cm^{-2}. The BNTorus model shown in the figure has opening angles θ_{torus} between 30° and 60°, and is stronger with smaller θ_{torus}. We note that the BNTorus spectra are weaker than plcabs-based models at the same N_H, likely because of differences in the cross-section of absorption and/or scattering and in the abundance of elements. To compare the spectral shapes in Figure 6, the normalizations of the plcabs-based models are multiplied by factors of 75% and 50% in the cases of $N_H = 5 \times 10^{23}$ cm^{-2} and 1×10^{24} cm^{-2}, respectively. In this work, we focus on how to measure N_H accurately and ignore this minor effect on the measurement of intrinsic luminosity, which only has a moderate impact on our results. In both cases where $N_H = 5 \times 10^{23}$ cm^{-2} and 1.5×10^{24} cm^{-2}, the spectral shape of our model is very similar to the continuum of the BNTorus model. If the self-absorption were not added, the plcabs+pexrav model would be softer (dashed line in the lower panel of Figure 6) than BNTorus. It is found by Liu & Li (2015) that the reflection in the BNTorus model is overestimated because of a lack of torus self-absorption on the reflected emission from the inner region of the torus. This further strengthens the necessity of adding the self-absorption to pexrav in our model.

3.3. Spectral Fitting Strategy

Before determining the final spectral fitting model for each source, we need to choose the spectral components—whether a narrow Fe Kα line or a soft-excess component is needed, and decide whether, in different periods, the power-law slope should be kept as a free parameter or fixed to a constant, and whether intrinsic absorption N_H should be linked together or left free to vary.

3.3.1. Determining Power-law Slopes

Ideally, all the parameters should be allowed to vary in each period. However, most of the CDF-S sources have low S/N, which hampers the measurement of each single spectral parameter with sufficient accuracy. In particular, there is a strong degeneracy between the power-law slope and the intrinsic absorption, such that in the low S/N regime, a very steep slope can be accommodated with a very high absorption level, and a very flat slope can be obtained with a severely

underestimated absorption. To avoid such a degeneracy, we link the power-law slopes of all the periods together and set it free only if the slope parameter Γ could be well constrained, that is, the relative error (1σ error divided by the best-fit value) of Γ is lower than 10% and the N_H is lower than 5×10^{23} cm^{-2}. The best-fit Γ values in these well-constrained cases have a median value of 1.8 (see Section 4.1)—a typical slope of AGN found or adopted in a huge number of papers. Therefore, in all the other cases, we fix Γ at 1.8.

Under this assumption, we can focus on the distribution of N_H of our sources, despite the disadvantage that the dispersion of N_H could be slightly reduced.

3.3.2. Searching for Fe Lines

A narrow Fe Kα line is commonly detected in AGNs, but its detection is limited by the quality of the spectrum. Using fitting method B, we search for the existence of a narrow Fe Kα line by comparing the fitting statistics with and without the line component in the model. For simplicity, the line energy and flux are assumed to be constant in the four periods. First, we claim a line detection when the best fit with a narrow line component at 6.4 keV has an improvement $\Delta C > 3.84$, corresponding to a more than 95% confidence level with $\Delta DOF = 1$ (degree of freedom). Although the ΔC method could not provide accurate probability in line detection (Protassov et al. 2002), we can still use it as a rough selection method. Then for the sources with photometric redshifts, we search for the narrow Fe Kα line by letting the redshift vary. In five cases when $\Delta C > 9.21$, which corresponds to 99% confidence with $\Delta DOF = 2$, we claim a line detection and replace the photometric redshift with an X-ray spectroscopic redshift, which guarantees a line energy of 6.4 keV. The ID, old redshift, and new redshift of the five sources are 98: 1.41–1.99, 646: 2.13–1.49, 733: 2.40–2.64, 940: 3.31–3.08, and 958: 0.87–0.89, respectively. For the first one (ID = 98), the same X-ray spectroscopic redshift has been reported by Del Moro et al. (2014).

Broad Fe lines are even harder to detect in the low S/N regime. For 29 sources that have an N_H below 10^{22} cm^{-2}, a spectroscopic redshift measurement, and at least 1000 net counts in the 0.5–7 keV band, we search for the broad Fe line by setting both the line energy and width free. This component is taken as detected if $\Delta C > 11.34$, which corresponds to a 99% confidence with $\Delta DOF = 3$. Then we fix the line width at 0. If $\Delta C < 2.71$ (90% with $\Delta DOF = 1$), the line width is consistent with 0. We consider such lines as narrow Fe Kα lines.

Eventually, we detect 50 narrow Fe Kα lines and 5 broad Fe lines. In the final model, the line component is adopted only if a line is detected. The line energy is set free; the line width is fixed at 0 for narrow lines and set free for broad lines.

3.3.3. Searching for the Soft Excess

A soft excess is often detected in the soft X-ray band of AGN, but its origin is uncertain. We add a secondary power-law as a phenomenological model for such a component, in order to cope with different physical origins. The normalization of the secondary power-law component is restricted to <10% of that of the primary power-law, and its slope is constrained to be equal to or steeper than that of the primary one. In the cases of obscured AGN, this secondary power-law could describe the power-law scattered back into the line of sight (e.g., Bianchi et al. 2006; Guainazzi & Bianchi 2007), which has the same slope, but <10% flux of the primary power-law (Brightman & Ueda 2012). In the cases of unobscured AGN, the soft excess could be due a blurred reflection from ionized disk (e.g., Crummy et al. 2006) or warm Comptonization emission from the disk (e.g., Mehdipour et al. 2011; Matt et al. 2014). Regardless of the physical origin, this component must have a steeper slope to rise above the primary power-law. Meanwhile, we avoid a secondary power-law flatter than that of the primary also in order to avoid a severe component degeneracy problem in the fitting. We establish the presence of a significant soft excess when the best fit improves by $\Delta C > 5.99$ after adding the secondary power-law component (95% confidence level with $\Delta DOF = 2$). Since the soft excess can be variable on timescales of years, we search for the soft excess both in each period and in the 7Ms stacked spectrum, thus fitting methods A and B are used here. If a soft excess is detected in one period for one source, the additional power-law component is activated for this period. If the excess is only detectable in the 7Ms stacked spectrum, the additional power-laws are activated for all four periods, but with the same slope and scatter fraction (ratio of the power-law normalization to that of the primary power-law).

In the cases of highly obscured sources, the soft-excess component is likely a scattered power-law. Since the primary power-law is severely reduced, the soft excess could even dominate the 0.5–7 keV band, if the scattered fraction is higher than a few percent. As described in Section 3.4.1, we look for a hard excess in order to select Compton-thick sources. For sources identified as Compton-thick in this way, for example, source 409 (lower panel of Figure 5), we consider the power-law that dominates the soft band as a scattered component, and activate the secondary power-law components in all four periods with the same scatter fraction, so that we can measure the N_H and intrinsic luminosity on the basis of the hard component. In these cases, the slope of such a secondary power-law is linked to that of the primary. In other cases when the N_H is above 5×10^{23} cm^{-2}, the slopes of the primary and the secondary power-law are also linked.

In some other cases when necessary, we add a soft-excess component even if it is not significantly detected, in order to ensure that the detected N_H variation is not due to soft excess, see Section 3.3.4 for details. Eventually, 85 sources in our sample have an activated soft-excess component in the final model, 27 of them are set with a constant scatter fraction, and the others are free in one or a few of the four periods. These sources have relatively lower redshifts, whose distribution is different from that of the whole sample at a Kolmogorov-Smirnov (K-S) test probability of 99%. This is because at high redshifts, the observed 0.5–7 keV band corresponds to a harder band where the soft excess is less prominent and the spectral S/N is relatively lower.

3.3.4. Searching for N_H Variations

When we fit the spectra of all the periods simultaneously as in methods C and D, the primary power-law has a constant slope (either free or fixed) and an independent normalization, and by default, the intrinsic absorption has a constant N_H. After the model components are selected, we check for variability of

N_H between each pair of the four periods among the 171 sources that have at least 300 broadband net counts in the 7Ms exposure. Here we use fitting method C. First, we find the best-fit after setting N_H free in all the four periods. Then, for each pair of periods, we link the N_H parameter and measure the ΔC with respect to the best fit obtained with all the four parameters set free. When $\Delta C > 6.64$ ($\Delta DOF = 1$), N_H is considered to be different between the two corresponding periods at a >99% confidence level. In some cases, since the soft-excess component is detected in one period but not the other, the apparent N_H variation might be caused by the soft-excess component. In order to guarantee that the N_H variation is independent of the soft excess in such cases, a soft-excess component, which does not improve the fit as significantly as required by our selection threshold above, is added manually into the spectral model of each period. When no significant variability is found, the N_H of all the four periods are linked together. When N_H is found to be significantly different between two periods, their N_H values are set independent of each other; for the other periods, if the best-fit N_H obtained when all the N_H were set free is between the two independent ones, it is set to the mean value of the independent ones; if the best-fit N_H is larger or smaller than both the independent ones, it is linked to the independent one that is closer. Eventually, our final fitting reports more than one N_H for a source if its N_H is significantly varying, and only one average value if not.

3.3.5. Measuring Intrinsic Luminosities

At this point, the final spectral model has been set for each source, and the best-fit parameters have been obtained with using fitting method D. In order to calculate a 7Ms averaged intrinsic luminosity, we set the flux to be constant among the four periods, and calculate the mean absorption-corrected rest-frame 2–10 keV flux of the primary power-law on the basis of the spectral modeling. To compute the error range of this flux, we add a "cflux" component in front of the "power-law" component in the spectral model. This component allows us to use the "error" task in Xspec to calculate the error of the absorption-corrected flux in the same manner as the other spectral parameters.

3.4. Identifying Compton-thick AGNs

We follow four procedures to identify Compton-thick candidates.

3.4.1. Exceptionally Strong Reflection

When $N_H > 1.5 \times 10^{24}\,\mathrm{cm}^{-2}$, the primary power-law is severely reduced by the line-of-sight obscuration, while the reflection component, determined by the intrinsic strength of the primary power-law and the material around the core, is relatively independent of the line-of-sight N_H. Therefore, Compton-thick AGN have the defining characteristic of exceptionally strong reflection compared with the primary power-law. In the soft band, the reflection component, which has an extremely hard spectral shape, can be easily swamped by other soft components, such as a scattered power-law; it is only prominent in the hard band. Our first attempt to identify Compton-thick AGN is to look for an excess in the hard band, which indicates a reflection component with an exceptionally large relative strength.

Some AGNs might change their states between Compton-thick and Compton-thin. However, hampered by the low S/N, spectroscopic identification of Compton-thick AGN would be less sensitive in each single period. Therefore, here we only use the stacked 7Ms spectra (fitting method B). We compare our standard Compton-thin model with a Compton-thick model in order to identify sources that can be better described by the latter one. In our standard Compton-thin model, the relative strength of the reflection to the unabsorbed primary power-law is small and fixed. The Compton-thick model is converted from the standard model by setting the `constant` in the "`zwabs*pexrav*constant`" component free. In this way, an exceptionally strong reflection component will manifest itself in terms of a large `constant`, as illustrated with an example (source 409) in Figure 5. Such strong reflection, which could hardly be caused by the observed weak power-law, indicates that the primary power-law is hidden (obscured) and the observed power-law that dominates the soft band should be considered as a scattered component. In order to understand the relative strength of the reflection intuitively, here we set the N_H in this component at $10^{22}\,\mathrm{cm}^{-2}$ and R at 1, which approximately describe a reflection from an infinite flat surface. To prevent too much model flexibility, which exceeds the constraining capability of the low S/N spectra, we fix the power-law slope at 1.8 and exclude the soft-excess component from the Compton-thick model. Based on such models, we find 40 sources that are better fitted with the Compton-thick model at a >95% confidence level, which corresponds to $\Delta C > 3.84$ when $\Delta DOF = 1$. Of these, we select 23 sources whose `constant` has a best-fit value >7 and a 90% lower limit >2; since the reflection component of a Compton-thick AGN must be not only significant, but also exceptionally strong.

Even though strong hard excesses are found in these sources, further checking is still needed. In highly obscured cases (with an N_H of a few $10^{23}\,\mathrm{cm}^{-2}$), especially in the low S/N regime, it is hard to distinguish between a highly obscured power-law and a reflection component; a hard excess might be explained by either of them. To check this possibility, we fit the spectra using our standard Compton-thin model with the secondary power-law activated; so that the soft X-ray emission is fitted with the unobscured secondary power-law, and the hard excess is fitted with the obscured primary power-law. This model has the same DOF as the Compton-thick model. We find that for 13 sources, the double power-law Compton-thin model fits the spectra better, with a best-fit N_H below or above $1.5 \times 10^{24}\,\mathrm{cm}^{-2}$. We add a scattered power-law, which has the same Γ as the primary power-law but at most 10% of the primary power-law's normalization, to the model of such sources, as mentioned in Section 3.3.3. Further checking will be applied to them in the next procedure. For the other 10 sources, the Compton-thick model fits the spectra better. However, this does not ensure that they are Compton-thick. We note that the reason they are better fitted with the Compton-thick model is that their power-law component, as opposed to the hard excess, is obscured with $N_H > 10^{22}\,\mathrm{cm}^{-2}$; so that this obscured power-law cannot be well fitted with the unobscured power-law model in the Compton-thin model. In such cases, although a strong hard excess is detected, our explanation of this excess as a reflection caused by a hidden primary power-law that is much stronger than the observed one seems not proper any more. Because a line-of-sight absorber has been found, to imagine another independent Compton-thick absorber, which is possible in principle, however, is

overexplaining the data. Alternatively, the hard excess could be caused by partial-covering obscuration or special geometry of reflecting material, which cannot be constrained with our low S/N spectra. Therefore, we do not consider such sources as Compton-thick. By now, we have not selected any Compton-thick sources, but have done essential preparations for next procedure.

3.4.2. Large Best-fit N_H

The secondary power-law in our spectral model is a flexible term of "soft excess." In highly obscured cases, it can be used to describe a scattered power-law that could dominate the observed soft-band flux. As mentioned above, when a strong hard excess is found, the secondary power-law is activated in the final model, so that the primary power-law appears as a hump in the hard-band spectrum. This mechanism also works in some other cases where a soft-excess component is found in Section 3.3.3. Our second Compton-thick AGN identification procedure is based on the best-fit N_H obtained with the final spectral model.

Here we replace the absorption model zwabs with the plcabs, which gives more accurate N_H measurement in the high N_H cases, see Section 4.7 for details. In the high N_H cases, with a large uncertainty, the measured N_H often has an error range crossing the Compton-thick defining threshold of $N_H = 1.5 \times 10^{24}$ cm^{-2}. For simplicity, we just select Compton-thick sources by comparing the best-fit N_H with 1.5×10^{24} cm^{-2}. Excluding the Compton-thick candidates as selected in the last section, we find 22 Compton-thick sources with a best-fit N_H $> 1.5 \times 10^{24}$ cm^{-2}.

3.4.3. Narrow Fe Kα Line

The presence of a strong narrow Fe Kα line (EW $\gtrsim 1$ keV) is also an indicator for a Compton-thick AGN (e.g., Levenson et al. 2002). The Fe line EW is positively correlated with N_H (e.g., Leahy & Creighton 1993; Bassani et al. 1999; Liu & Wang 2010), since the higher line-of-sight N_H depresses the continuum, but not the line. We show this correlation later in Section 4.2. Generally, strong Fe Kα lines with EW ~ 1 keV are detected in highly obscured AGNs; unobscured AGNs do not display such strong lines. For each source with a best-fit rest-frame Fe Kα line EW > 1 keV, if the N_H is $< 10^{21}$ cm^{-2}, we consider its slightly obscured power-law emission as a scattered component rather than the primary power-law; and thus they are likely Compton-thick sources, whose expected hard excess is not detected because of the low S/N. In addition to the Compton-thick candidates found above through the continuum fitting, we find no further candidates with this method. The low efficiency of this method is discussed in Section 4.4.

3.4.4. Mid-infrared 12μm Luminosity

Waste heat in the mid-infrared (MIR) band of AGN is an efficient tool to find deeply buried AGN (e.g., Daddi et al. 2007; Fiore et al. 2008, 2009; Gandhi et al. 2009; Georgantopoulos et al. 2009; Severgnini et al. 2012; Asmus et al. 2015; Stern 2015; Corral et al. 2016; Isobe et al. 2016). We cross-correlate our sample with the GOODS-*Herschel* catalog (Elbaz et al. 2011) with a maximum separation of 1″, and find 167 MIR counterparts. These sources are provided with 8μm, 24μm, and 100μm fluxes, although some objects
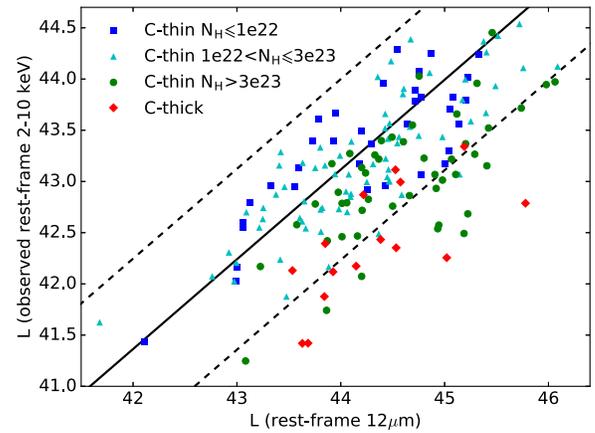

Figure 7. Scatter plot of the rest-frame 2–10 keV observed X-ray luminosity and the total rest-frame 12 μm luminosity. The sources are divided into four subsamples according to the N_H. The solid line corresponds to the best-fit $L_{2-10 \, \text{keV,intrinsic}} - \log L_{12\mu m}$ correlation, and the dashed lines represent its 2σ scatter.

are not detected in all the bands. Since all of our sources are AGN, we assume a power-law shape of the MIR spectrum and measure the rest-frame 12 μm fluxes of these sources by simple interpolation. This assumption is not necessarily true because in some cases, the MIR spectrum of AGN may deviate from an ideal power-law because of a strong star burst component or polycyclic aromatic hydrocarbon (PAH) features (e.g., Ichikawa et al. 2012, 2017; Georgantopoulos et al. 2013; Del Moro et al. 2016). Therefore, the measured 12 μm flux can be up to a few times higher than the genuine MIR emission from the AGN torus.

Based on the sources with $N_H < 10^{23}$ cm^{-2}, we perform orthogonal distance regression (ODR) between the MIR luminosity and the absorption-corrected 2–10 keV luminosity $\log L_{2-10 \, \text{keV,intrinsic}}$ (presented in Section 4.1), and find $\log L_{2-10 \, \text{keV,intrinsic}} = \log L_{12 \, \mu m} \times 0.88 + 4.56$, with a 1σ scatter on $\log L_{2-10 \, \text{keV,intrinsic}}$ of 0.44, as shown in Figure 7. From the best-fit model of the 7Ms stacked spectra, we measure the absorbed rest-frame 2–10 keV luminosity $\log L_{2-10 \, \text{keV,observed}}$, and plot it with the rest-frame 12 μm luminosity in Figure 7. The sources are divided into four subsamples according to the N_H measured with our final model. Average N_H is used in the cases of varying N_H. Despite the large scatter, it is still visible that sources with higher N_H have lower observed X-ray luminosity. If one source has $\log L_{2-10 \, \text{keV,observed}}$ below the 2σ confidence interval of the above correlation, that is, $\log L_{2-10 \, \text{keV,observed}} < \log L_{12\mu m} \times 0.88 + 4.56 - 2 \times 0.44$ (the dashed line in Figure 7), we consider it as highly obscured, but not necessarily Compton-thick. Considering the inaccuracy of our MIR flux measurements, we are adopting a conservative selection rule here. For X-ray sources satisfying this selection rule, intrinsic absorption with an N_H well above 10^{23} cm^{-2} is expected. If the measured N_H of such a source is lower than 10^{22} cm^{-2}, we consider its soft-band continuum as dominated by scattered power-law and select it as a Compton-thick candidate, similar as done in Section 3.4.3. In addition to those we identified above, no further Compton-thick candidates are found with this method either.

Table 2
Counts and Fluxes of Each Source in Each Period

ID	Period	Soft Cts	Hard Cts	Soft Flux	Hard Flux
20	I	33 ± 12	6 ± 13	$2.96^{+0.693}_{-1.04}$	$5.09^{+1.66}_{-1.7}$
20	II	102 ± 13	50 ± 14	$21.6^{+1.25}_{-3.83}$	$37.1^{+6}_{-6.07}$
20	III	153 ± 17	78 ± 23	$11.2^{+0.565}_{-1.93}$	$19.2^{+3.03}_{-3.03}$
20	IV	102 ± 15	102 ± 22	$10.1^{+0.726}_{-1.94}$	$17.3^{+2.79}_{-2.78}$
21	I	235 ± 16	60 ± 10	$25.8^{+2.02}_{-1.9}$	$35.7^{+3.09}_{-3.09}$
21	II	4 ± 3	0 ± 5	$6.46^{+3.61}_{-3.59}$	$8.96^{+5}_{-5.07}$
21	III	61 ± 11	24 ± 13	$5.47^{+1.06}_{-1.04}$	$7.59^{+1.46}_{-1.48}$
21	IV	55 ± 11	63 ± 15	$5.89^{+1.27}_{-1.25}$	$8.16^{+1.7}_{-1.7}$
22	I	78 ± 12	0 ± 13	$6.79^{+0.678}_{-1.74}$	$9.71^{+1.86}_{-1.88}$
22	II	25 ± 10	25 ± 15	$5.91^{+1.1}_{-2}$	$8.6^{+2.25}_{-2.23}$

Note. Column 1: source id (Luo et al. 2017). Column 2: period. Column 3: 0.5–2 keV net counts. Column 4: 2–7 keV net counts. Column 5: 0.5–2 keV observed flux in 10^{-16} erg cm^{-2} s^{-1}. Column 6: 2–7 keV observed flux in 10^{-16} erg cm^{-2} s^{-1}.

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

4. Results of Spectral Analyses

4.1. Basic Properties

We report the source net counts and observed fluxes in Table 2. The net counts are measured directly from the source and background spectra. The observed fluxes are calculated on the basis of our best-fit models.

We present the spectral-fitting results in Table 3, which contains the following entries:

1. Source ID in the 7Ms CDF-S catalog (Luo et al. 2017).
2. Source redshift and quality flag of the redshift (Luo et al. 2017).
3. Flag of N_H variation.
4. Intrinsic absorption N_H.
5. Primary power-law slope Γ.
6. Observed 0.5–2.0 keV net count rate.
7. Observed 2.0–7.0 keV net count rate.
8. Absorption-corrected rest-frame 2–10 keV luminosity.
9. Sample completeness flag, which is defined later in Section 5.3.1.
10. Existence of soft excess.

We are able to place good constraints on Γ for 95 sources (see Section 3.3). The distribution of the parameter Γ has a median value of 1.82 and a standard deviation of 0.15, similar to that found by Tozzi et al. (2006) and Yang et al. (2016) for CDF-S AGNs, as shown in Figure 8. With an 1σ confidence interval of (1.80,1.83), which is measured by a bootstrap method, the median Γ is consistent with 1.8. For the remaining sources, Γ is fixed at 1.8 during the spectral fitting.

We check several factors that may affect the measurement of Γ, including N_H, soft excess, and radio-loudness. First, we split these sources into two subsamples with $N_H < 10^{22}$ cm^{-2} and $N_H \geqslant 10^{22}$ cm^{-2} (see Figure 8). They show identical Γ distributions, with a K-S-test probability as low as 30% to be different. Then we compare the sources with and without a detected soft-excess component (see Figure 8) to ensure that the addition of an secondary soft power-law does not cause a flat Γ. The distributions are not significantly different (K-S test probability 80%), and the Γ of sources with soft-excess components seem slightly steeper. Therefore, we are not affected by this bias. Since we are using a secondary power-law to approximately describe a soft-excess component that might have a very different origin in different cases, we do not further investigate their properties. To check the impact of radio-loud sources, we match our sources with the CDF-S 1.4 GHz radio sources (Bonzini et al. 2013) within a separation of $2''$ and find 63 counterparts; 13 of them which have a 1.4 GHz to V-band flux ratio $>10^{1.4}$ and are classified as radio-loud by Bonzini et al. (2013). We plot 9 of them that have free Γ in Figure 8. They show an identical Γ distribution to the other sources, with a K-S-test probability as low as 20% to be different.

4.2. Fe Lines

A narrow Fe Kα line is detected in 50 sources and a broad one is detected in 5 sources, as listed Tables 4 and 5. Their line energies, widths (for broad line), and EWs are measured by fitting the 7Ms stacked spectra (method B). The low S/N of the broad lines do not allow us to fit them with a physical model. They could be relativistic broad lines from the inner region of the accretion disk, or might be a blending of multiple lines at different ionization levels. Below we discuss the properties of the narrow lines. The distributions of the line central energy and EW are shown in Figure 9. Our detection procedure ensures that each line energy is consistent with 6.4 keV, which indicates an origin from cold neutral gas. The scatter in the distribution is caused by measurement uncertainty. The EWs of the sources span a large range from a few hundred to a few thousand eV.

In Figure 10 we show the rest-frame EW of the detected iron lines versus N_H. For sources with varying N_H, an average value in log space $\langle \log N_H \rangle$ is used. Considering only the sources with secure spectroscopic redshifts, a strong correlation is shown with a Spearman's rank probability $>99.999\%$. We build a simple model to examine this correlation. As reflection emission, the narrow Fe Kα line originates from circumnuclear materials such as the outer region of the accretion disk, the broad line region, or the dusty torus (e.g., Shu et al. 2010, 2011). It is justifiable to consider the strength of this line emission as independent of N_H, which is determined by the line-of-sight material. We assume an N_H-independent line flux, and use our standard model to measure the flux density of the continuum at 6.4 keV at a specific N_H; then an EW is calculated by dividing the line flux by the continuum flux density. When we fit this model to the sources with secure redshifts, as shown in Figure 10, the line flux can be calculated. We find that this flux corresponds to an EW of 135 eV when unobscured, with a 1σ scatter of 80–230 eV. This is broadly consistent with the EW measured for a local type I AGN sample by Liu & Wang (2010). As shown in Figure 10, the data are well described by our toy model. However, we remark that this does not mean that the Fe line or continuum strength is independent of N_H. Liu & Wang (2010) find that the narrow Fe Kα line is 2.9 times weaker in Compton-thin type II AGN than in type I AGN in terms of luminosity; and Liu et al. (2014) find that the intrinsic X-ray emission (unobscured power-law) is 2.8 times weaker in Compton-thin type II AGN than in type I AGN. Here these N_H-dependent effects largely cancel each other.

Table 3
Spectral Properties

ID	z	Period	N_H	Γ	Soft Rate	Hard Rate	L	C–N_H	Model
Compton-thin									
20	1.3700i		$1.14^{+0.782}_{-0.694}$	1.8f	5.9e–07	3.2e–07	$43.31^{+0.05}_{-0.06}$	23.5	···
21	1.0650s		$0.13^{+0.298}_{-0.13}$	1.8f	8.1e–07	3.1e–07	$43.18^{+0.04}_{-0.03}$	23.5	···
22	1.9400s		$1.08^{+1.86}_{-1.08}$	1.8f	2.4e–07	1.3e–07	$43.39^{+0.08}_{-0.08}$	23.5	···
24	2.3143p		$58.2^{+14.4}_{-11.5}$	1.8f	1.2e–07	4.4e–07	$44.19^{+0.09}_{-0.10}$	24	···
26	2.3040s		$0.94^{+1.15}_{-0.94}$	1.8f	7e–07	2.7e–07	$43.89^{+0.04}_{-0.04}$	24	···
27	2.9112p		$23.4^{+8.68}_{-7.31}$	1.8f	1.6e–07	2.4e–07	$44.05^{+0.09}_{-0.10}$	24	···
32	1.3740s		$15.2^{+8.35}_{-5.92}$	1.8f	8.3e–08	2.4e–07	$43.28^{+0.14}_{-0.16}$	23.5	1
31	1.3310p		$1.97^{+0.845}_{-0.751}$	$1.88^{+0.21}_{-0.19}$	7.9e–07	5.7e–07	$43.61^{+0.04}_{-0.04}$	24	···
33	1.7846p		$16.5^{+10.3}_{-6.97}$	1.8f	8.9e–08	1.7e–07	$43.42^{+0.15}_{-0.17}$	23.5	···
34	2.9400s		$43.3^{+24.4}_{-18.7}$	1.8f	1e–07	1.6e–07	$43.95^{+0.16}_{-0.19}$	24	···

Note. Column 1: source id. Column 2: redshift. Flag "s": secure spectroscopic; "i": insecure spectroscopic; "p": photometry, "x": X-ray spectroscopic-z. Column 3: period, if N_H is found variable. This field is empty when N_H is not variable. When more than one period is given in one line, their N_H are set as the same. When one period of a source whose N_H is variable is not given in the table, its N_H is set as the average of the N_H of the other periods. Column 4: N_H in 10^{22} cm^{-2} with 90% errors. Flag "f" means fixed. Note the N_H for Compton-thick sources are considered unreliable and not used in this work. Column 5: Γ with 90% errors. Flag "f" means fixed. Column 6,7: 0.5–2 and 2–7 keV observed net count rates, averaged among the four periods, weighted by the exposure time. Column 8: log 2–10 keV unabsorbed luminosity with 90% errors, in erg s^{-1}. This value is highly uncertain in the cases of extreme Compton-thick absorption, where the transmitted power-law becomes too weak. Column 9: completeness flag C–N_H. Use the filter "C–$N_H \geqslant \widetilde{N_H}$" to select a subsample that is complete with respect to N_H at log $N_H < \widetilde{N_H}$, where $\widetilde{N_H}$ can be 23, 23.5, or 24. −1 means Compton-thick. Column 10: period number in which soft-excess component is added. 0 means soft excess is found in the 7Ms stacked spectrum. Empty means no soft excess is found.

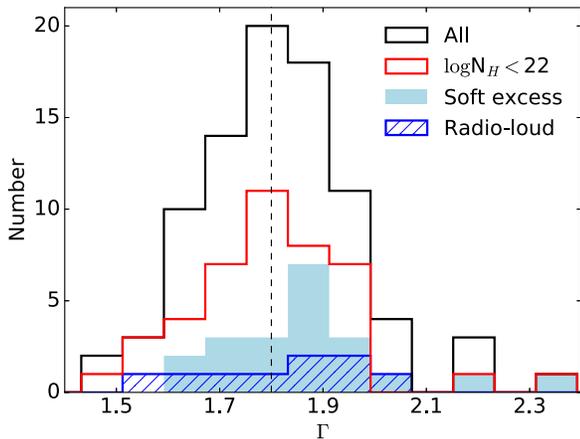(This table is available in its entirety in machine-readable form.)

Figure 8. Distribution of best-fit Γ for the 95 sources whose Γ are well constrained. The vertical line shows the median value of this distribution ($\Gamma = 1.8$). The red histogram shows the subsample with $N_H < 10^{22}$ cm^{-2}. The light-blue filled part corresponds to the sources with a detected soft-excess component. The blue dashed part shows the 9 radio-loud sources.

4.3. N_H Variation

Yang et al. (2016) studied the N_H variations of AGN using the 6Ms CDF-S data available at the time of their analysis. As we did, they divide the CDF-S observations into four periods. The first three periods are the same as ours, and the fourth period is supplemented with the last 1Ms of observations in this work. They found 11 sources with reliable N_H variations. In this work, we use the 7Ms CDF-S data, which are somewhat different, adopt a slightly different selection threshold, and search among a larger population of sources. For 39 sources, we find the N_H to be variable among 63 pairs of periods. Three sources (ID 898, 328, 252) found by Yang et al. (2016) are not considered as N_H variable in this work. The first because of

the low significance, and the other two are considered as having variable soft-excess components.

The physical origin of the obscuration variability is uncertain. The obscuration variability might be driven by the flux variability of the central engine, in the sense that a stronger illumination from the central engine could depress the line-of-sight obscuration through ionizing or blowing away the obscuring material; or the obscuration variability might be independent of the flux variability if it is caused by obscuring clouds that move across the line of sight. To study the correlation between obscuration variability and flux variability, we define a factor k between each pair of N_H variable spectra $k = (\log N_H^1 - \log N_H^2)/(\log L^1 - \log L^2)$, where L is the 2–10 keV intrinsic luminosity. A positive k indicates a positive correlation and a negative k signifies a reverse trend of variation. To take parameter uncertainties into account, we use the *Xspec* "simpars" command to generate 1000 sets of parameters for each of the concerned spectra, and calculate the k factors using them, making a smoothed k distribution. Considering the degeneracy between Γ and N_H, to avoid the possible effect of unnoticed Γ variation, we exclude the sources whose Γ are not well constrained and thus fixed at 1.8 from the analysis here; and for the remaining 42 pairs of spectra of 24 sources, we let Γ vary freely in each period when generating the random sets of parameters.

In addition to calculating the k between each pair of N_H variable spectra, we also calculate this value between the resampled values (random parameters generated with simpars) of each single spectrum, which is named k_d, in order to show the natural degeneracy between the parameters. As shown in Figure 11, the N_H–L degeneracy (blue histogram) appears as a distribution systematically biased to $k_d > 0$; the k factor calculated between our N_H-variable pairs shows a tendency to $k < 0$, in spite of the existing degeneracy effect. The k distribution has a large scatter, suggesting that the N_H variation

Table 4
Narrow Fe Kα Lines

ID	Energy (keV)	EW (eV)
26	$6.36^{+0.05}_{-0.07}$	354^{+145}_{-135}
50	$6.46^{+0.05}_{-0.04}$	444^{+162}_{-184}
58	$6.53^{+0.10}_{-0.12}$	879^{+400}_{-320}
89	$6.50^{+0.19}_{-0.08}$	132^{+59}_{-61}
98	$6.52^{+0.12}_{-0.11}$	200^{+78}_{-72}
106	$6.49^{+0.06}_{-0.06}$	241^{+86}_{-81}
119	$6.42^{+0.04}_{-0.03}$	228^{+63}_{-55}
135	$6.35^{+0.09}_{-0.09}$	411^{+183}_{-165}
174	$6.43^{+0.03}_{-0.04}$	1767^{+4322}_{-200}
208	$6.35^{+0.10}_{-0.08}$	154^{+44}_{-44}
240	$6.40^{+0.04}_{-0.05}$	411^{+144}_{-142}
242	$6.44^{+0.04}_{-0.04}$	130^{+32}_{-39}
290	$6.31^{+0.05}_{-0.04}$	1235_{-810}
328	$6.41^{+0.06}_{-0.06}$	261^{+96}_{-94}
355	$6.42^{+0.02}_{-0.03}$	1278_{-858}
357	$6.58^{+0.08}_{-0.10}$	688^{+1005}_{-320}
367	$6.33^{+0.07}_{-0.07}$	257^{+83}_{-83}
386	$6.42^{+0.10}_{-0.10}$	719^{+549}_{-463}
399	$6.27^{+0.06}_{-0.06}$	336^{+142}_{-117}
402	$6.36^{+0.07}_{-0.07}$	805^{+930}_{-452}
430	$6.41^{+0.05}_{-0.06}$	528^{+291}_{-186}
447	$6.35^{+0.09}_{-0.08}$	674^{+617}_{-358}
448	$6.16^{+0.08}_{-0.07}$	487^{+222}_{-190}
458	$6.58^{+0.10}_{-0.24}$	188^{+76}_{-72}
485	$6.39^{+0.08}_{-0.07}$	141^{+63}_{-60}
495	$6.46^{+0.03}_{-0.03}$	66^{+16}_{-17}
507	$6.36^{+0.05}_{-0.05}$	158^{+53}_{-48}
551	$6.53^{+0.11}_{-0.08}$	583^{+263}_{-179}
557	$6.27^{+0.05}_{-0.06}$	137^{+49}_{-51}
614	$6.27^{+0.11}_{-0.18}$	1441^{+517}_{-1031}
621	$6.42^{+0.34}_{-0.22}$	217^{+115}_{-113}
638	$6.36^{+0.12}_{-0.12}$	361^{+222}_{-189}
643	$6.25^{+0.04}_{-0.04}$	219^{+73}_{-68}
646	$6.39^{+0.06}_{-0.06}$	2002_{-1350}
666	$6.44^{+0.06}_{-0.06}$	2569_{-1453}
730	$6.32^{+0.09}_{-0.03}$	95^{+27}_{-26}
733	$6.37^{+0.07}_{-0.06}$	400^{+135}_{-116}
735	$6.45^{+0.10}_{-0.31}$	153^{+66}_{-71}
748	$6.30^{+0.11}_{-0.11}$	214^{+115}_{-107}
752	$6.42^{+0.05}_{-0.05}$	156^{+87}_{-83}
805	$6.35^{+0.09}_{-0.09}$	351^{+155}_{-148}
826	$6.34^{+0.05}_{-0.07}$	866^{+1529}_{-402}
840	$6.30^{+0.05}_{-0.08}$	186^{+73}_{-73}
867	$6.45^{+0.08}_{-0.07}$	451^{+340}_{-271}
868	$6.37^{+0.03}_{-0.04}$	1450^{+830}_{-500}
940	$6.42^{+0.08}_{-0.08}$	1036^{+1159}_{-408}
958	$6.40^{+0.05}_{-0.06}$	544^{+786}_{-314}
981	$6.32^{+0.06}_{-0.05}$	279^{+125}_{-111}
986	$6.37^{+0.04}_{-0.04}$	261^{+117}_{-102}
988	$6.36^{+0.04}_{-0.05}$	187^{+120}_{-101}

Note. Col 1: ID; Col 2: rest-frame central energy with 1σ error; Col 3: rest-frame EW with 1σ error.

Table 5
Broad Fe Kα Lines

ID	Energy (keV)	σ (keV)	EW (eV)
175	$6.52^{+0.05}_{-0.05}$	$0.20^{+0.09}_{-0.09}$	108^{+26}_{-25}
479	$6.61^{+0.46}_{-0.79}$	$2.60^{+1.41}_{-0.80}$	703^{+341}_{-386}
716	$6.53^{+0.16}_{-0.14}$	$0.81^{+0.34}_{-0.25}$	323^{+112}_{-110}
856	$6.59^{+0.09}_{-0.12}$	$0.36^{+0.33}_{-0.17}$	267^{+302}_{-262}
898	$6.09^{+0.08}_{-0.10}$	$0.39^{+0.36}_{-0.21}$	424^{+131}_{-129}

Note. Column 1: ID; Column 2: rest-frame central energy with 1σ error; Column 3: rest-frame width with 1σ error; Column 3: rest-frame EW with 1σ error.
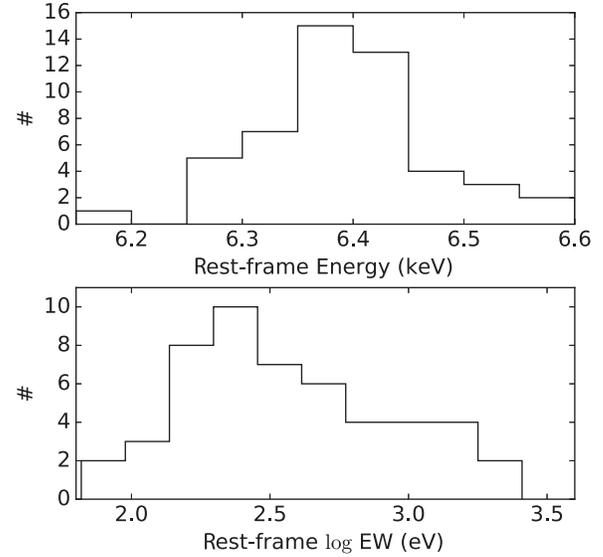

Figure 9. Distributions of the central energy and EW of the detected narrow Fe Kα lines. The line energies are consistent with 6.4 keV, and the width of the energy distribution is due to measurement uncertainties.
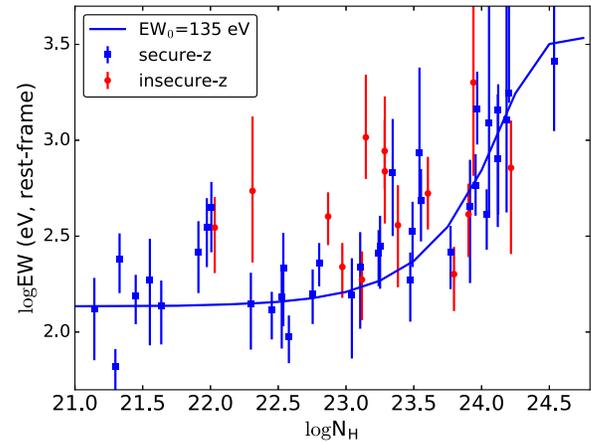

Figure 10. Scatter plot of log $N_{\rm H}$ and log EW of the detected narrow Fe Kα line. Sources with secure and insecure redshift measurements are plotted in blue and red, respectively. The blue line shows our model of the line with an $N_{\rm H}$-independent flux and EW = 135 eV when unobscured.

could be attributed to different processes in different cases. The distribution is likely composed of two components, one with a small scatter and the other very large. Clearly, the small-scatter component has $k < 0$. It is consistent with the case of

illumination-depressed obscuration, which leads to a reverse variation trend. In this case, a large flux variation amplitude is needed, while the obscuration variation cannot be very large; therefore, $|k|$ is relatively small. In the other case, an obscuring cloud moving across the line of sight could easily cause a dramatic variation of obscuration irrespective of the luminosity,
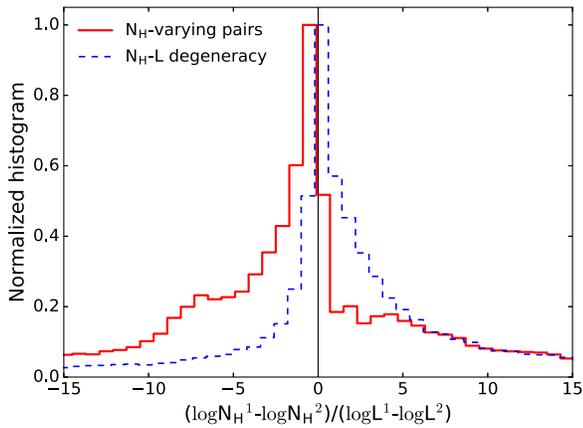
Figure 11. Resampled k distributions, calculated between each pair of N_H variable spectra (red) and each pair of random values of the same spectrum (blue). The histograms are normalized to the peak value.
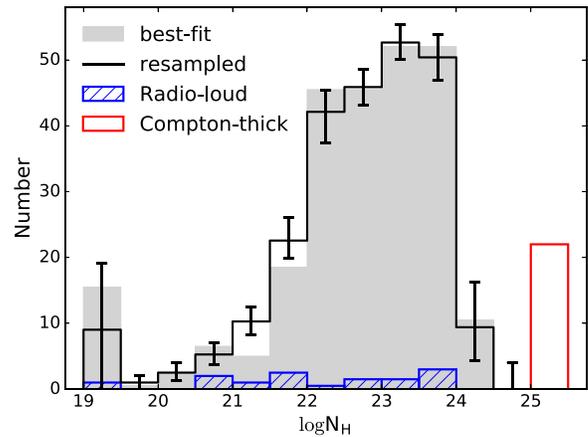

Figure 12. Observed N_H distribution. The gray filled histogram shows the best-fit N_H distribution of the Compton-thin sources. All the sources with best-fit log $N_H < = 19$ are plotted in the leftmost bin. Compton-thick sources are shown in the red histogram at log $N_H > 25$. The black empty histogram shows the resampled N_H distribution of the Compton-thin sources, using mock values generated according to the statistical error on N_H. The 1σ error bars of the histogram are obtained from the scatter of 1000 resampled histograms. The blue dashed histogram corresponds to the 13 radio-loud sources.

resulting in a largely scattered k distribution. We remark that by excluding the sources whose Γ are fixed at 1.8, we are biased against high-N_H sources (with $N_H > 10^{23}$ cm^{-2}), which tend to have a fixed Γ according to our analysis method. Obscuring material with such high N_H are less likely to be affected by the illumination of the central engine. Therefore, excluding them is helpful in revealing the reverse trend, which occurs more likely in the low-N_H cases.

4.4. Distribution of N_H Excluding Compton-thick Sources

Based on the four strategies described in Section 3.4, we classify 22 (8% out of 276) sources to be Compton-thick candidates. As shown in Figure 7, these Compton-thick candidates have significantly lower X-ray luminosities than most of the others, but are not distinctly different from some highly obscured Compton-thin sources because of the large scatter.

Although we have searched for AGN N_H variation among the four periods, we do not find any transition between Compton-thin and Compton-thick states. However, this is not evidence against the existence of such AGN in our sample. A faint Compton-thick phase might hide in one period of a source, whose N_H is linked to that of other Compton-thin periods because of the low S/N.

The identification of Compton-thick sources through X-ray spectral analyses depends on detailed selection rules. In the low S/N regime, it is hard to determine whether a highly obscured source is Compton-thick or -thin (see also the comparison of Compton-thick AGN identification results from different works by Castelló-Mor et al. 2013). We note that some Compton-thick candidates identified by previous spectral analyses are classified as Compton-thin in this work, including sources 375 (Feruglio et al. 2011, named BzK8608 therein), 551 (Comastri et al. 2011, CXOCDFSJ033229.8-275106), 328 (Norman et al. 2002; Comastri et al. 2011, CXOCDFSJ033218.3-275055), 419 (Del Moro et al. 2016, ID = 23), 240 (Georgantopoulos et al. 2013, XID = 191), and 867 (Georgantopoulos et al. 2013, XID = 634). According to our spectral analysis, these sources are found to be Compton-thin but still highly obscured, with N_H between 5.5×10^{23}–1.1×10^{24} cm^{-2}. Therefore, there is not any strong conflict. Of the 14 Compton-thick candidates selected in Tozzi et al. (2006), 3 are not included in this sample because of their low net counts, 2 are still identified

as Compton-thick, and 9 are classified as Compton-thin. For these 9 sources, we find that their low S/N spectra in Period I could indeed be well fitted with a reflection-dominated model. However, they are considered as Compton-thin according to our analysis of the 7Ms data. Apparently, we are more conservative in this work in selecting Compton-thick AGN. The differences between our results and previous classifications are caused both by different data and different spectral models, especially the latter. Instead of modeling Compton-thin and Compton-thick sources with pure transmitted power-law and pure reflection models, respectively, as done in Tozzi et al. (2006), we take into account a reflection component in addition to the transmitted power-law in the Compton-thin cases, and account for a scattered power-law component in addition to the dominating reflection emission in the Compton-thick cases. With such more realistic models, we can select Compton-thick candidates with higher reliabilities.

We detect neutral Fe Kα lines in 4 out of the 22 Compton-thick sources. Although strong Fe Kα lines are commonly detected in local bona fide Compton-thick AGNs, for some Compton-thick AGNs, the Fe Kα line might be weak because of a high ionization state or a low elemental abundance. For example, NGC 7674, which is identified as Compton-thick by *NuSTAR*, has a weak neutral Fe Kα line of EW ≈ 0.4 keV (Gandhi et al. 2017), likely because of its high bolometric luminosity (high ionization state). Our 22 Compton-thick sources have high redshifts, 20 of them with $z > 0.7$ and 6 with $z > 2$. Their bolometric luminosities might be higher than the known local Compton-thick AGNs. It is possible that they are intrinsically weak in the neutral Fe Kα line like NGC 7674. Even if not intrinsically weak, the line detection is severely hampered by the low S/N. Through simulations, Koss et al. (2015) found that the strong Fe Kα line (EW = 1.11 keV) in a Compton-thick AGN NGC 3933 ($z = 0.0125$) can be detected at an S/N = 3 by the 4 Ms CDF-S only if $z < 0.2$. This explains why the detection of neutral Fe Kα lines is inefficient in identifying Compton-thick AGNs in our high-z sample (Section 3.4.3).
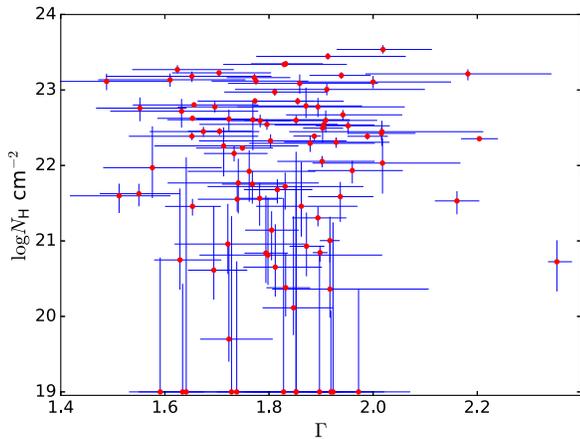
Figure 13. Scatter plot of the best-fit Γ and N_H. Only the sources with free Γ are shown here. Error bars correspond to 1σ.

After classifying Compton-thick sources, we can study the N_H distribution of our sample. Hereafter, since in the low-obscuration regime N_H cannot be well constrained, we set all the sources with best-fit $\log N_H$ below 19 at $\log N_H = 19$. Figure 12 shows the distribution of the best-fit N_H. To take into account the measurement uncertainty, which is often asymmetrical, we generate 1000 random N_H values for each Compton-thin spectrum, assuming two half-Gaussian profiles on the lower and upper side of the best-fit value using the lower and upper 1σ error as σ of the profiles. With these 1000 sets of N_H values, we obtain 1000 histograms, and we plot the median histogram and the corresponding 1σ errors in Figure 12. Although the resampling approach smoothes the N_H distribution slightly, it is still better to use the resampled values because the N_H uncertainty is nonuniform among the sources. The uncertainty is larger for sources with lower S/N (net counts), in other words, with lower luminosity or higher redshift. It is also larger for high-z sources with low N_H, since the observed-frame 0.5–7 keV band probes a harder rest-frame band, where the obscuration feature becomes less prominent. It is important to take these factors into account in order to study the intrinsic distribution of N_H and the luminosity- and redshift-dependences of N_H. By now, this distribution is simply the distribution of intrinsic absorption measured across the sample, including the statistical errors. We correct for the selection effects to derive the intrinsic absorption distribution across the AGN population in Section 5.

In Figure 12 we also show the matched radio-loud sources from Bonzini et al. (2013). They show a flatter N_H distribution as opposed to the other sources, which congregate at high N_H, with a K-S test probability of 99% to be different. These radio-loud AGNs are likely a mix of mildly absorbed and strongly absorbed galaxies, similar to the sources found by Tundo et al. (2012). Constituting only $\lesssim$5% of the sample, these radio-loud sources have no significant effect on the results. We consider all the sources as a whole population in the following analyses.

4.5. Correlation between Spectral Parameters

The scatter plot of Γ and N_H is shown in Figure 13. Only sources with measured Γ are shown. We check for possible correlation between Γ and N_H, since a degeneracy between the two parameters could show up at low S/N. In particular, high
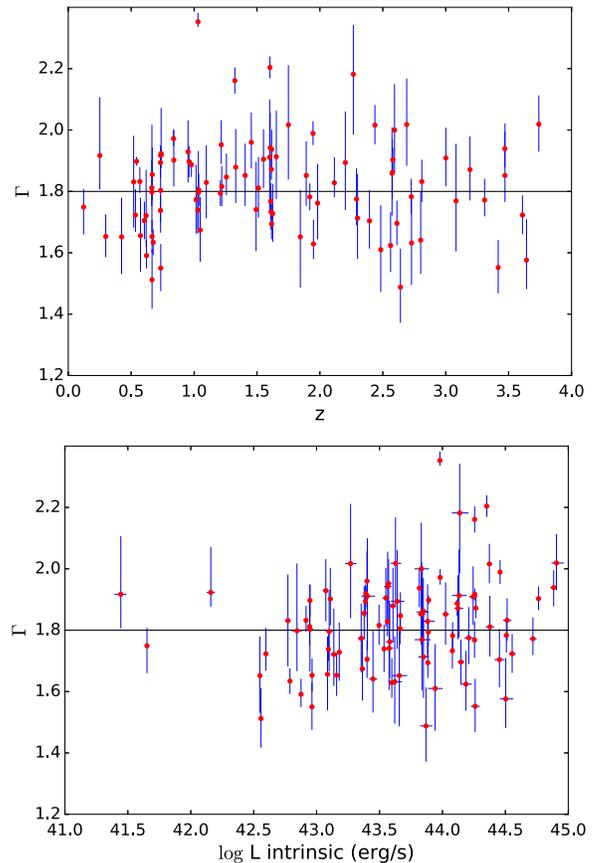

Figure 14. Upper panel: scatter plot of best-fit Γ vs. redshift. Lower panel: best-fit Γ vs. rest-frame 2–10 keV intrinsic luminosities. Only the sources with free Γ in the model are plotted. The line corresponds to the median Γ of 1.8. Error bars correspond to 1σ.

N_H can be accommodated with high Γ. We find a Spearman's rank correlation coefficient of 0.04, which corresponds to a probability of 30% to reject the null hypothesis that they are uncorrelated. Therefore no correlation is found, and our spectral fitting strategy is not significantly affected by this kind of bias.

In Figure 14 we show the scatter of Γ with redshift and intrinsic luminosity. There is no correlation between Γ and z, with a Spearman coefficient as low as 0.02. Γ and L show a slight correlation, with a Spearman coefficient of 0.20, corresponding to a probability of 95%. A simple linear fit results in a slope of 0.05 ± 0.02. The Γ–L correlation, if any, could be attributed to the positive correlation between Γ and λ_{Edd} (e.g., Shemmer et al. 2006, 2008; Risaliti et al. 2009; Brightman et al. 2013; Fanali et al. 2013), which is expected to arise from the more efficient cooling in the coronae of higher λ_{Edd} systems. However, considering the weak correlation and the large measurement uncertainties, we ignore this effect in this work.

In Figure 15 we show the correlation of N_H with redshift and intrinsic luminosity. For the Compton-thin sources, the Spearman correlation coefficient is 0.33 between the best-fit N_H and z, and 0.22 between N_H and intrinsic luminosity, both corresponding to a probability of >99.9%. We discuss these
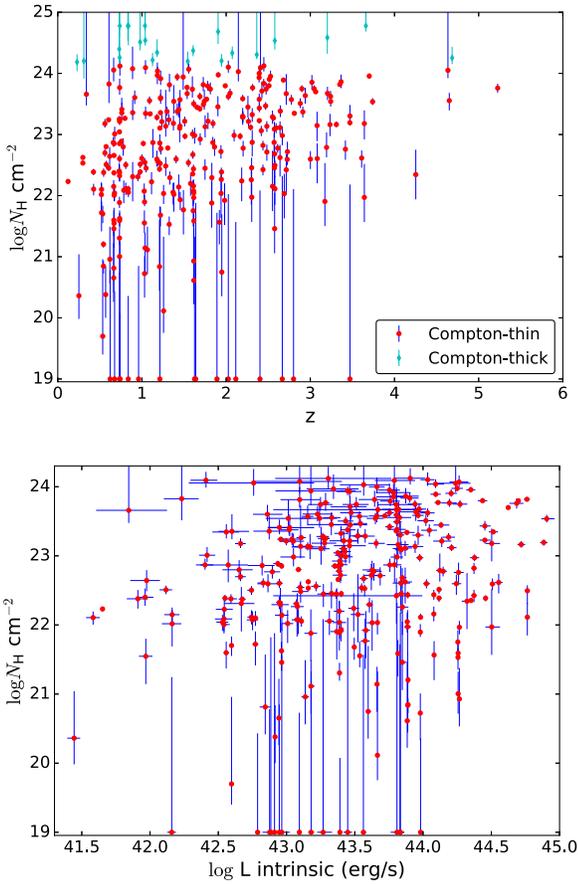
Figure 15. Best-fit N_H vs. redshift. Error bars correspond to 1σ. Some sources with very low or very high N_H have very large uncertainties. Scatter plot of the best-fit N_H and rest-frame 2–10 keV unabsorbed luminosities of the Compton-thin sources. Error bars correspond to 1σ.
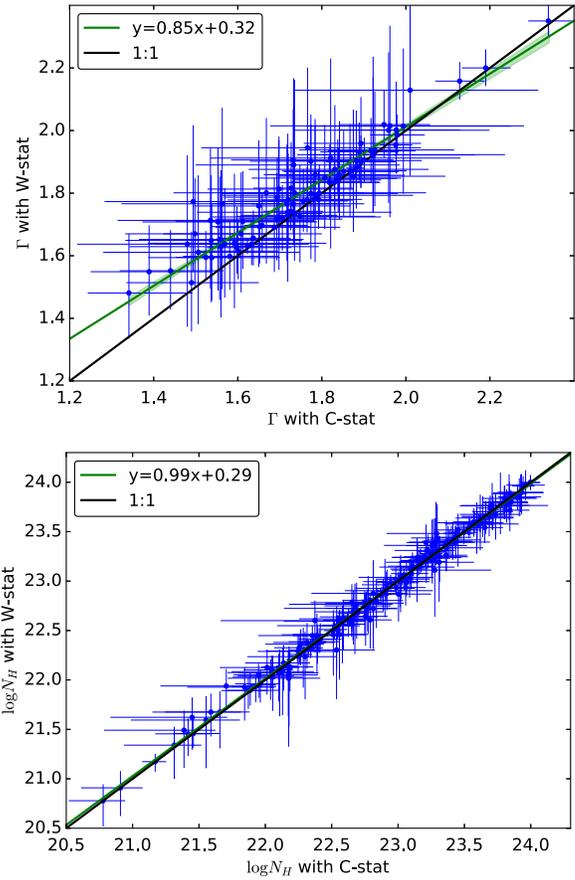

Figure 16. Power-law slope and column density as obtained using the spectral fitting methods C (standard C statistic) and D (W statistic). The best-fit lines and 1σ confidence intervals are plotted with the green lines and green shadings.

correlations after evaluating several selection biases in Section 5.3.

4.6. C-statistic versus W-statistic

As introduced in Section 3.1, both methods C (C-statistic) and D (W-statistic) can be used to obtain the final spectral fitting results. In principle, method D uses a more accurate statistical approach. Here we explore whether these two methods result in significant differences. In Figure 16 we compare the best-fit power-law slopes and the column densities obtained using the two methods. Only the sources with Γ derived from fitting are involved in the Γ comparison, and only the sources whose log N_H are below 24 according to both methods and whose 90% confidence intervals of log N_H are less than 1 are involved in the N_H comparison. Clearly, N_H is not affected by the choice of statistical method, as shown by the ODR best-fit line $y = (0.99 \pm 0.01)x + 0.29 \pm 0.14$ in Figure 16. However, at low values, Γ measured using the W statistic is typically higher ($\Gamma_W - \Gamma_C = 0.09$ at $\Gamma_C = 1.4$) than that measured using C statistic. Meanwhile, the <1 slope of the Γ_W–Γ_C best-fit line $y = (0.85 \pm 0.03)x + 0.32 \pm 0.05$ suggests that a slightly smaller scatter in the Γ distribution of the sample will be obtained using the W-statistic. However, because of the low S/N of our data, the difference caused by different statistical methods is small compared to the uncertainty. Therefore, we conclude that this effect is not significant.

Figure 17. NH measured using `zwabs` and `plcabs` with 1σ error for the Compton-thin sources, whose N_H are well constrained.

4.7. `zwabs` versus `plcabs`

As we mentioned in Section 3.2, the `zwabs` model considers only photoelectric absorption, but not Compton scattering. To check how this affects the results, we replace the `zwabs*po- werlaw` in the final model with `plcabs` (Yaqoob 1997), which approximately takes Compton scattering into account by modeling the X-ray transmission of an isotropic source located at the center of a uniform spherical distribution of matter. We set the maximum number of scattering to 3 and the high-energy cutoff e-folding energy to 300 keV. In Figure 17 we compare
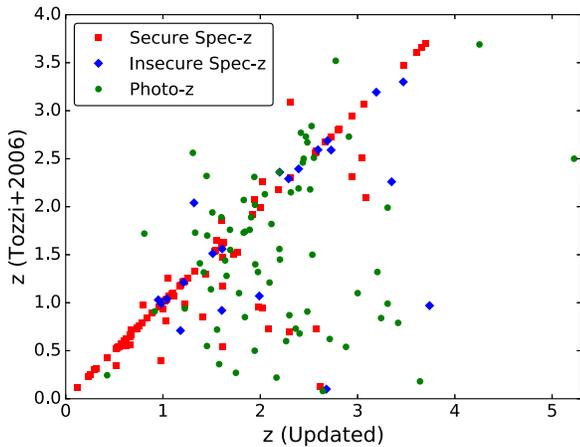
Figure 18. Three types of redshifts adopted in this work, compared with those used in Tozzi et al. (2006). Red square: secure spectroscopic; blue diamond: insecure spectroscopic; green circle: photometric.

the N_H measured with this model and with `zwabs`. The difference starts to appear above 10^{23} cm^{-2}, and is very small (by a factor of $\lesssim 0.3\%$ at log $N_H = 24$ cm^{-2}) compared to the uncertainty of our data. We conclude that this shortcoming of `zwabs` is negligible in this work. However, we still report the N_H measured with `plcabs`, and use it in identifying Compton-thick AGNs.

4.8. Comparison with Previous Works on X-Ray Spectral Analyses of CDF-S Sources

Several works have presented X-ray spectral analyses of CDF-S sources over the last 15 years (e.g., Tozzi et al. 2006; Brightman et al. 2014; Comastri & XMM-CDFS Team 2013; Buchner et al. 2014). The improvement presented in this work consists in not only the much longer exposure time, but also in the improved calibration of *Chandra* data, the updated redshift measurements, and different spectral analysis approaches. We compare our spectral fitting results with those obtained in Tozzi et al. (2006) and Buchner et al. (2014), in order to understand how these different aspects affect the spectral analyses.

As shown in Figure 18, the redshifts in the CDF-S catalog were dramatically improved in the past decade, mostly thanks to the continuous multiband follow-up and updated photometric redshift measurements, and in some cases due to the careful multiband identification in Luo et al. (2017). About 50% of the sources have a redshift that is different by more than $\Delta z = 0.05$ with respect to the values used in Tozzi et al. (2006), and the updated redshifts tend to be higher. Below, we limit our comparison with the Tozzi et al. (2006) results to the sources with $\Delta z < 0.05$. We also consider the same period, which corresponds to Period I in our analyses (the first 1 Ms). For the sources whose Γ were fixed at 1.8 in Tozzi et al. (2006), we also fix their Γ at 1.8. For the other sources, we compare the best-fit values of Γ obtained in the two works (see Figure 19), and obtain an ODR best-fit line of $y = (0.88 \pm 0.09)x + (0.27 \pm 0.16)$. According to this line, at $\Gamma = 1.4$ as obtained in Tozzi et al. (2006), we obtain $\Gamma = 1.5$ in this work. The reason for the higher Γ we obtain is that we use the W statistic instead of the C statistic, and we include a reflection component in the spectral model. In Figure 20 we show the comparison of the best-fit values of N_H in the two works. The uncertainty is very large in the cases of unobscured ($N_H < 10^{20}$ cm^{-2}) low S/N sources. For the sources that have an N_H between 10^{20} and 1.5×10^{24} cm^{-2} according to both works, we
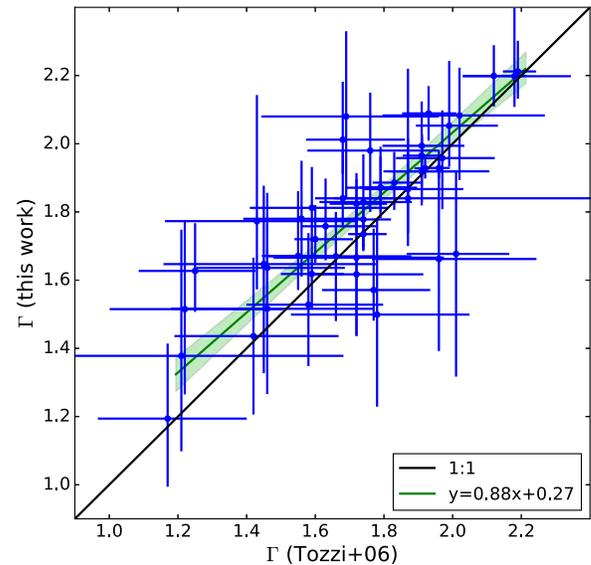

Figure 19. Photon indices measured in Tozzi et al. (2006) and in this work in Period I for the sources with consistent redshifts. Sources whose photon indices are fixed at 1.8 are excluded. Errors correspond to 1σ. The linear fitting result and 1σ confidence interval are plotted with the green line and shading.
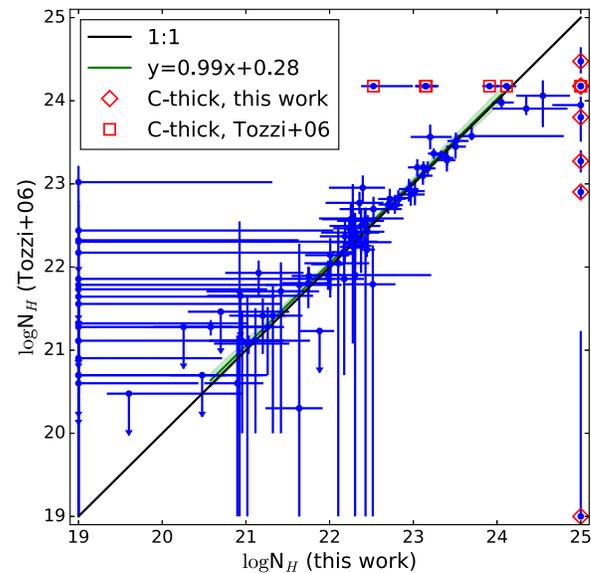

Figure 20. Column densities N_H with 1σ errors measured in Tozzi et al. (2006) and in this work in Period I for the sources with consistent redshifts. Sources identified as Compton-thick in Tozzi et al. (2006) and in this work are marked with red squares and red diamonds. We set the log N_H of our Compton-thick sources at 25. The linear fitting result and 1σ confidence interval are plotted with green line and shade.

perform an ODR fit. The best-fit line $y = (0.99 \pm 0.05)x + (0.28 \pm 1.08)$ is fully consistent with 1:1. Significant differences occur in the highly obscured regime. They are caused by the improved spectral models used in this work. In conclusion, although we used an updated calibration, new extraction procedures for the source and the background, and different spectral models, our spectral fitting results are in broad agreement with those of Tozzi et al. (2006).

Similarly, we compare our spectral fitting results with those of Buchner et al. (2014). Between the 4Ms CDF-S catalog they used and our updated 7Ms catalog, we still found $\geqslant 30\%$ of
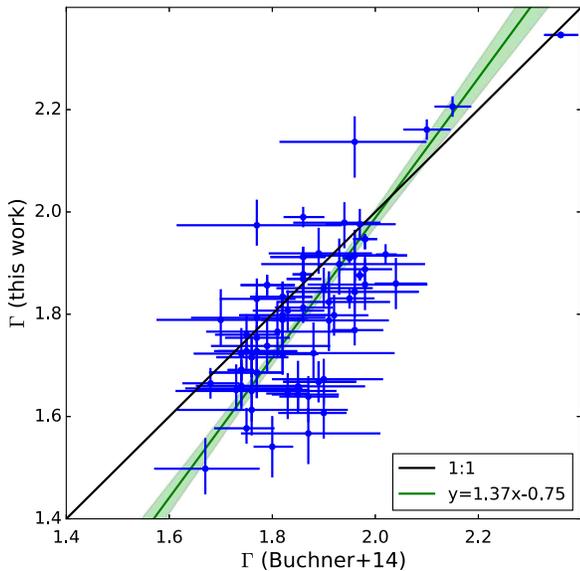
Figure 21. Photon indices Γ with 1σ errors measured in Buchner et al. (2014) on the basis of 4Ms CDF-S and in this work on the basis of 7Ms CDF-S. Sources whose photon indices are fixed at 1.8 in this work are not shown. The linear fitting result and 1σ confidence interval are plotted with the green line and shading.
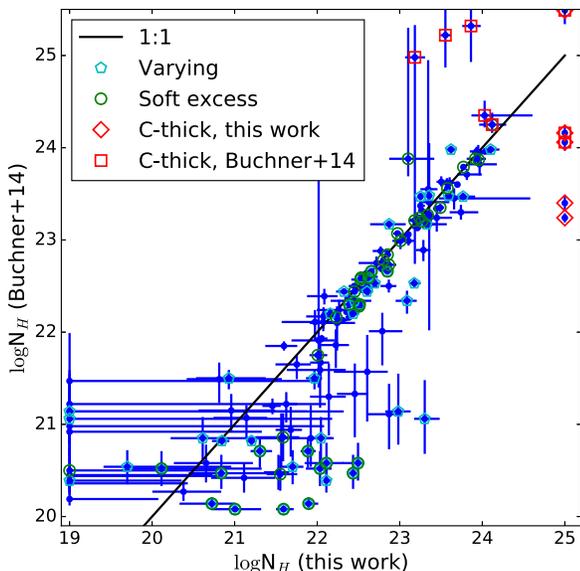

Figure 22. Column densities N_H with 1σ errors measured in Buchner et al. (2014) and in this work. Sources identified as Compton-thick in Buchner et al. (2014) and in this work are marked with red squares and red diamonds, respectively. We set the log N_H of our Compton-thick sources at 25. If N_H is found to vary among different periods in this work, all the values are plotted and marked with cyan pentagons. Sources with soft-excess components, whose spectral model contains an additional power-law, are marked with green circles.

sources (mostly photometric) with redshifts changed by more than 0.05. Again we compare only those with $\Delta z < 0.05$. In Figure 21, it is clearly shown that the Γ values obtained in Buchner et al. (2014) are higher than ours. As shown in Figure 8, we found a median $\bar{\Gamma}$ of 1.8 in our sample. The steeper slopes found by Buchner et al. (2014) might be caused by the prior distribution of Γ assumed in the Bayesian method, which has a mean value of 1.95 according to *GINGA* observations by Nandra & Pounds (1994), or by neglecting

soft excess or setting a stronger reflection component in the spectral model.

In Figure 22 we compare the N_H obtained in both works. Results from the two works are well consistent at around log $N_H = 23$, while at higher and lower N_H, there is large scatter. In the regime of higher N_H, the Compton-thick AGN identifications are different for a few sources, because, as we point out in Section 4.1, with limited information, the identification of Compton-thick AGN is highly dependent on the detailed selection method. Meanwhile, we obtained higher N_H for several heavily obscured Compton-thin sources compared with Buchner et al. (2014). In the case of low N_H, the scatter is even larger. These differences could be attributed to the different models we used and in some cases to spectral variations. Buchner et al. (2014) used a physical torus model with additional reflection and scattering, while we use a phenomenological model in which absorption, reflection, and scattering are considered but not physically connected. Buchner et al. (2014) demonstrated that in the Compton-thin regime, the two approaches can describe the observed spectra of the full sample equally well. However, various particular sources may prefer different model configurations, as also shown by Buchner et al. (2014). Particularly, we set the relative reflection strength of all the sources to a typical value, because the spectral quality does not allow us to place any constraint on it. For highly obscured sources, whose 2–7 keV fluxes could be dominated by the reflection component, the measured N_H and Γ are dependent on this configuration. As this configuration is chosen with the aim of obtaining systematic spectral properties of the sample, it is not necessarily accurate for each particular source. Another factor is the soft-excess component, the existence of which has a strong impact on the N_H measurement. We add this component to the spectral model only if it is detectable in the spectrum. Adding it to the soft band, the measured N_H on the primary power-law can be dramatically enlarged (see Figure 22). In Figure 22 we also mark sources that are found to have a varying N_H among the four observation periods. Clearly, spectral variation also plays a part in the large differences of N_H measurements in some cases. Considering the degeneracy between N_H and Γ, the difference in Γ might be another reason for the N_H differences.

In conclusion, although more differences are found between this work and Buchner et al. (2014), all these differences are understandable, which can be largely attributed to the different spectral models used in the two works. Both works apply systematic analyses to the whole sample. But for each specific source, especially low S/N ones, different spectral-analysis strategies import different assumptions into the models, thus leading to different N_H measurements.

5. Properties of the Intrinsic Obscuration

In this section we dissect the sample-selection function and draw general conclusions about the properties of the distribution of intrinsic obscuration N_H across the AGN population in two ways. First, using the sample selected with at least 80 hard-band net counts, we retrieve the intrinsic N_H distribution, which is representative of an AGN population that is well defined according to L and z, by correcting the sample-selection bias using the known luminosity function of AGN; second, instead of correcting our sample for the missing part, we trim the incomplete part without any assumption about the luminosity function and build a subsample that is complete for the
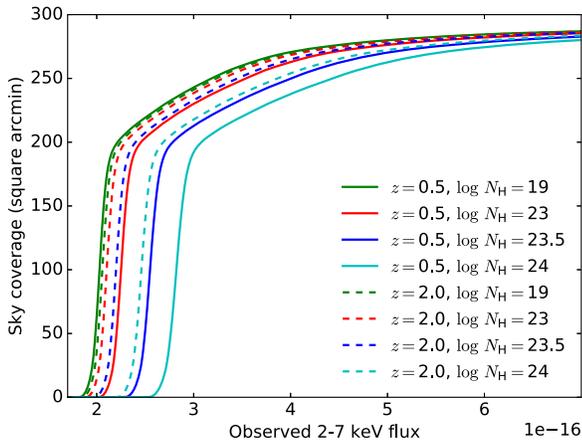
Figure 23. Sky coverage as a function of the 2–7 keV observed flux (erg cm^{-2} s^{-1}) corresponding to 80 hard-band net counts for a set of N_H values at redshifts of 0.5 (solid lines) and 2 (dashed lines).

Figure 24. Conversion surfaces from any point in the L_X–z space to hard-band observed net counts, obtained at the aimpoint with a log N_H of 22 (blue grid and line) and 24 (red grid and line). Only the parts above 80 net counts of the surfaces are shown. The lines in the L_X–z space at the base are projections of the intersections of the two surfaces with the plane where the net counts = 80.

distribution of the column density, in order to probe the dependence of N_H on the intrinsic X-ray luminosity and the cosmic epoch. The supplementary sample is included in the second part.

5.1. Sample Selection Function

5.1.1. Sky Coverage

For the combined 7Ms exposure of the CDF-S field, not only does the effective area decrease as a function of the off-axis angle, but also the exposure time at off-axis angles larger than ~8 arcmin. This is because in the outskirts, abrupt variations in exposure time are due to regions imaged only in some of the 102 exposures as a result of the different roll angles. Taking into account both effective area and exposure time, we build a map of the flux limit across the field, which obviously shows low values (maximum sensitivity) in a limited region around the center, and increasingly higher values toward the edges of the field. Using this map, we measure the sky coverage as a function of the hard flux corresponding to 80 net counts, which is the selection threshold of our sample.

The sky coverage is relevant to measuring the N_H distribution because the correspondence between detected counts and emitted flux depends on the spectral shape. In particular, intrinsic absorption can dramatically change the conversion factor between them. In Figure 23 we show the sky coverage of the CDF-S for our sample assuming our standard spectral model (absorbed power-law with $\Gamma = 1.8$ plus reflection) with a set of different N_H and redshift values. The sky coverage is biased against the detection of sources with high N_H. This bias is less severe at high redshift because of the inverse K-correction of highly obscured sources. In Section 2.5 we choose to apply the sample-selection threshold in the hard band in order to reduce the effect of sky coverage on the N_H distribution. We further correct this bias through weighting each source by the reciprocal of the sky coverage corresponding to its observed flux when computing the N_H distribution.

5.1.2. N_H-dependent Malmquist Bias

A flux-limited survey is typically biased against sources with lower luminosity at higher redshift. This effect is termed Malmquist bias. We express the Malmquist bias by drawing a
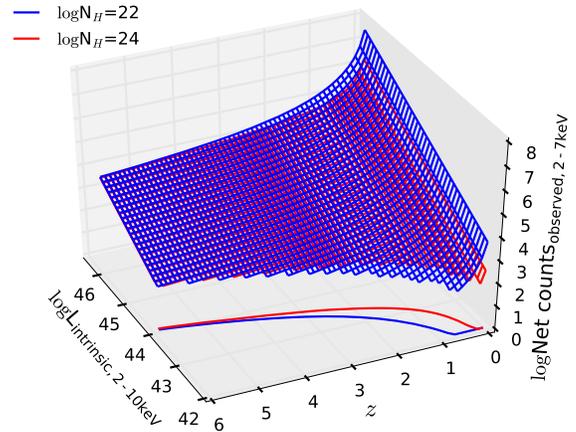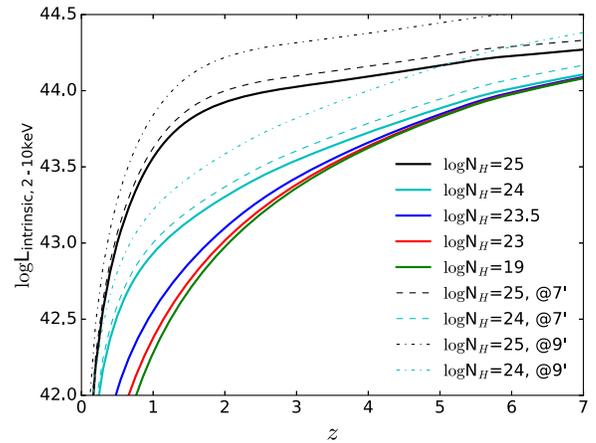

Figure 25. Boundary of detectable regions in L_X–z space with respect to different N_H. Solid lines correspond to the aimpoint region (off-axis angle <1.5). Dashed lines and dash-dotted lines correspond to off-axis angles of 7′ and 9′, respectively.

detection boundary curve in the space of luminosity and redshift, which represents the flux limit of the sample. By the general definition, the boundary corresponding to a flux limit is determined by the survey depth, which in our case decreases with off-axis angle. However, instead of using the flux-limited sample, our aim is to study the N_H distribution in a complete sample that is unbiased to the intrinsic X-ray luminosity L_X. Therefore, we have to take into account the dependence of observed flux on N_H, and express the Malmquist bias (the boundary curve) as N_H dependent.

To model the N_H-dependent Malmquist bias, we divide the CDF-S field into nine annular regions according to off-axis angle with steps of 1′ and central radii of 1′–9′. The circular region within 0.5, which is very small in area compared to the annuli, is assigned to the first annulus. In each annulus, at each specific log N_H between 19 and 25 with a step of 0.5, we convert each point in the L_X–z space into observed net counts by running *Xspec* with a model spectrum and the real response files and exposure time. A visual example of the conversion is shown in Figure 24. We take the standard model that was used in our spectral fitting, that is, an absorbed power-law plus a
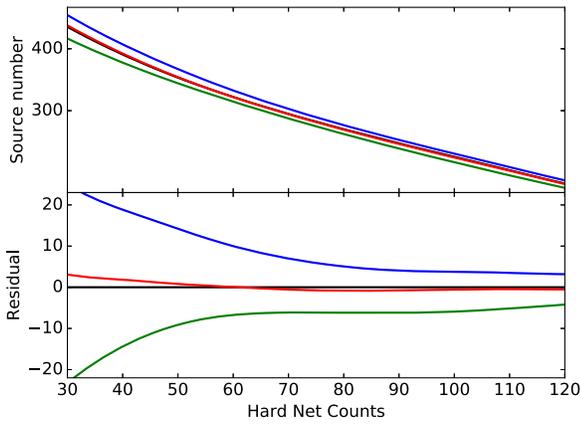
Figure 26. In the upper panel, the black line, on which an almost identical red line is superimposed, shows the smoothed cumulative number counts distribution of the CDF-S 7Ms point-source catalog. Convolving the black line with the measurement error of net counts results in the blue line. The green line is the horizontal reflection of the blue line about the black line. As a consistency check, the green line is convolved with the measurement error, giving rise to the red line, which is found to be almost identical to the black line at >50 net counts. The residuals of the lines to the black line are shown in the lower panel.

scattered power-law and a reflection, with $\Gamma = 1.8$, $R = 0.5$, $E_{cutoff} = 300$ keV. The scattered fraction is set to 1.7%, which is the average scattered fraction for 4 Ms CDF-S sources (Brightman & Ueda 2012). At each typical off-axis angle, the averaged response files and exposure time of the sources are used. In addition to repeating this in each annulus, we also carry out the whole field averaging case by taking the response files and exposure-time-averaged files from all the sources in the field.

Using the conversion surface from L_X–z into net counts, the sample-selection threshold of 80 net counts is in turn converted back into the L–z space, defining a detectable boundary curve —the luminosity limit at each redshift. In Figure 25 we show the detectable boundaries at a few specific N_H values at the aimpoint (within 1.′5). The boundary is less affected by N_H below $\log N_H = 23$, thanks to our hard-band sample-selection threshold. We also plot the boundaries corresponding to $\log N_H = 24$ and 25 at off-axis angles of 7′ (median off-axis angle of our sample) and 9′ (maximum) in Figure 25, to show the variation in survey depth.

5.1.3. Eddington Bias

As shown in Figure 3, faint sources naturally outnumber bright ones in the CDF-S point-source catalog. Considering that the net-count measurement could have large uncertainty in the low S/N regime, especially at large off-axis angles where the background is high, the sample-selection threshold on net counts introduces Eddington bias. In other words, more sources are included than expected because of net-count measurement uncertainties. We take this effect into account with an approach as follows. As shown in Figure 26, we convolve the number-count distribution of the 7Ms CDF-S point-source catalog as a function of hard-band net counts with the measurement uncertainty. For each source the net counts are replaced by thousands of Poissonian random values following its measurement uncertainty. The uncertainty-convolved number-count distribution is flatter, with relatively larger net counts. As shown in Figure 26, we take the rightward shift of the number counts caused by the convolution as an estimation of the

Eddington bias. Shifting the number counts leftward by the same amount, we take the horizontal reflection of the uncertainty-convolved distribution about the original distribution as an approximation of the deconvolved number counts. This is our pseudo-deconvolving method. To check its accuracy, we convolve the deconvolved number counts with the measurement uncertainty. As shown in Figure 26, we obtain number counts that are almost identical to the original one at >50 net counts.

Figure 26 is truncated at 30 net counts because at lower counts, a "faint-end" effect becomes significant, which is caused by the undetected sources that are not considered in the simulation. However, 80 net counts is well above the peak value of the net-count distribution (see Figure 3), the faint-end effect is negligible at 80 net counts.

Comparing the original and the deconvolved number counts, with our selection threshold of 80 net counts, we include ∼10 more sources into our sample because of Eddington bias. Allowing the same sample size, the 80 net counts threshold on the original number-count distribution corresponds to 74 net counts on the deconvolved one. To correct the Eddington bias, we use this value as an "effective" threshold. Specifically, in the correction described above in Sections 5.1.1 and 5.1.2, we use 74 instead of the nominal net-count threshold 80.

5.2. Intrinsic Obscuration Distribution

Correcting for the effects described above, we can retrieve the intrinsic N_H distribution representative of a well-defined AGN population from the observed one as follows. First, we correct for the sky coverage effect, that is, a source with higher N_H tends to have lower observed flux and thus lower sky coverage. While calculating the resampled distribution of N_H from the spectral-fitting results, we weight each source by the reciprocal of its corresponding sky coverage. As shown in Figure 27 (the "corrected.1" histogram), this correction has a minor effect.

Second, we correct for the N_H-dependent Malmquist bias. To do this, we need to know the distribution of AGN across the L–z space. Miyaji et al. (2015) modeled the AGN intrinsic X-ray luminosity function as a double power-law in a series of redshift bins from 0 to 5.8. Georgakakis et al. (2015) also provided a double power-law luminosity function, but focusing on the high-redshift section of $3 < z < 5$. We combine their works to model the AGN luminosity function in the $0 < z < 3$ and $3 < z < 5$ bins using Miyaji et al. (2015) and Georgakakis et al. (2015), respectively. The combined luminosity function is shown in Figure 28.

Concerning a specific redshift range, for each specific N_H, we define an observable space as the L–z space between $\log L = 45$ and the whole-field-averaged completeness boundary on L. As shown in Figure 25, a lower N_H corresponds to a lower detectable L boundary, and thus a larger observable space. In a set of selected $\log N_H$ between 19 and 25 with a step of 0.5, $\log N_H = 19$ corresponds to the largest observable space; all the others are sub-spaces of it. By integrating the AGN luminosity function over the observable space, we obtain the AGN number density in it. For each specific N_H, the fraction of the AGNs in its detectable space over the AGNs in the largest, $\log N_H = 19$ detectable space is shown in Figure 29 for a few redshift bins. Using this curve of fraction, we can correct the unobservable part below the completeness boundaries, so that sources with different N_H cover the same L–z
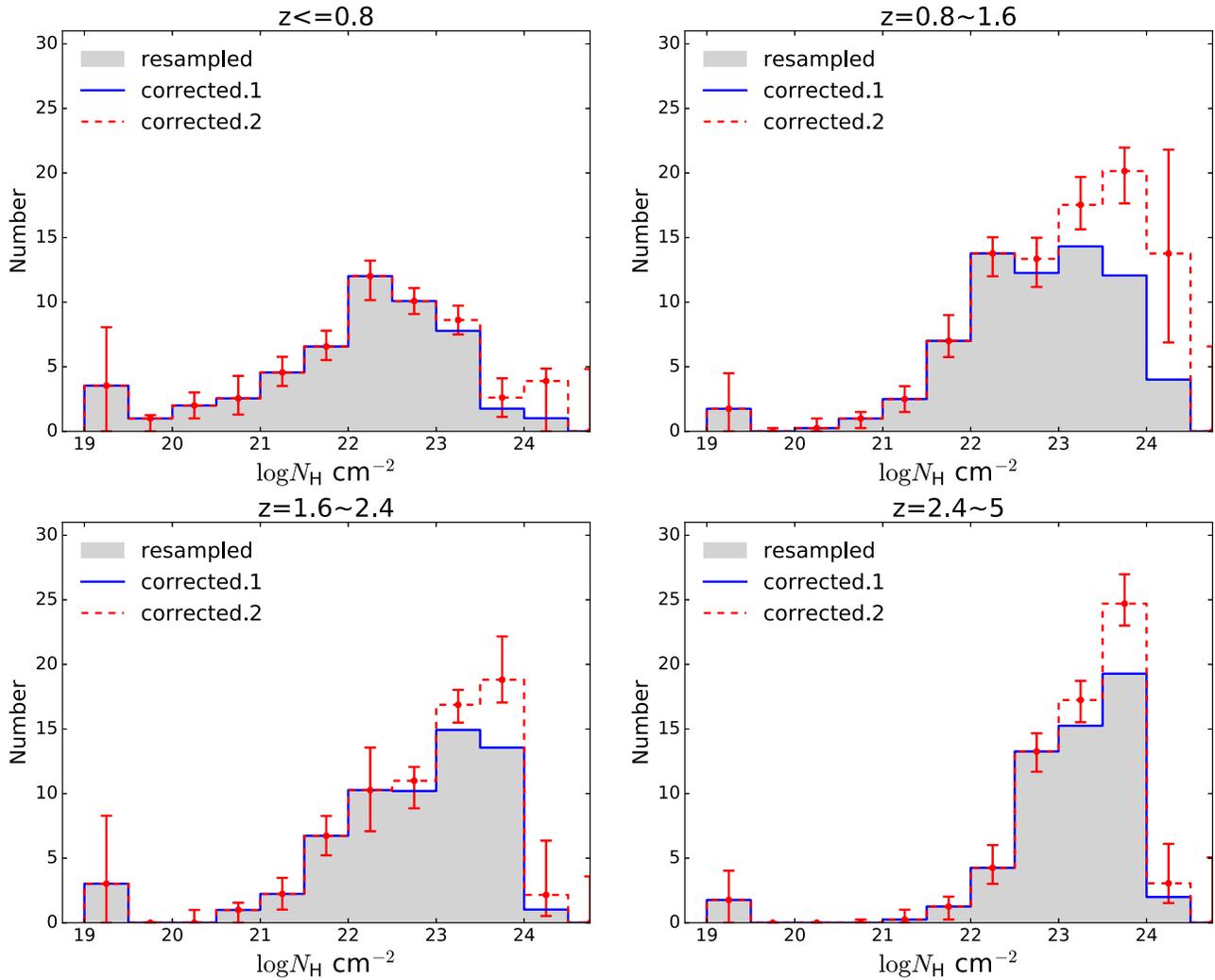
Figure 27. N_H distributions in the observable spaces corresponding to four redshift bins, as show in Figure 25. The 90% percentile luminosity range of the sources in the four redshift bins are $41.65 \sim 43.40$, $42.57 \sim 44.12$, $43.05 \sim 44.35$, and $43.46 \sim 44.76$, respectively, in the order of rising z. Gray filled histograms show the resampled N_H distribution of the Compton-thin sources, the same as plotted in Figure 12. The "corrected.1" (blue solid line) histogram shows the N_H distribution after correction for the inhomogeneous survey depth, which has a minor effect. Then after further correcting for the N_H-dependent Malmquist bias, the N_H distribution is plotted as the "corrected.2" (red dashed line) histogram. The 1σ error obtained from bootstrapping is plotted on the "corrected.2" histogram.

space—the log $N_H = 19$ observable space. This space is illustrated in Figure 28; the corrected N_H distribution, as shown in Figure 27 (the "corrected.2" histogram), is representative of the AGN population in this space. Clearly, the lower boundary of L is well below the knee of the LF, indicating that we are sampling the typical AGNs in the universe that produce a large fraction of the cosmic accretion power. For log N_H below 24, the unobservable part in the L–z space to be corrected is small compared to the observed space. Therefore, this correction is not sensitive to the selection of LF.

In the calculation, we assume that the shape of the luminosity function is independent of N_H. In other words, the luminosity functions of AGN with a different N_H have different normalizations but the same shape. It has been found that obscured AGN fraction is dependent on intrinsic luminosity, as discussed in Sections 1 and 5.3.3. Apparently, this obscuration-luminosity correlation suggests different LFs of obscured and unobscured AGNs. However, it does not necessarily indicate different LF shapes at different N_H; it could also be explained as a result of anisotropy of AGN X-ray emission (Lawrence & Elvis 2010; Burlon et al. 2011; Liu et al. 2014; Sazonov et al.

2015), that is, different LF normalizations rather than shapes at different N_H. Therefore, the assumption of an N_H-independent LF shape does not contradict our results; we do not discuss more complex corrections based on N_H-dependent LFs.

We remark that the two biases corrected above are independent of each other. Although they both relate to how the N_H affects the observed flux, the sky-coverage effect is caused by the varying depth across the whole field, while the Malmquist bias occurs at a specific depth, which in the correction above corresponds to the average depth across the field.

As shown in Figure 27, the N_H distribution of the Compton-thin sources shows large differences between redshift bins because they sample AGN populations with different L and different z. The higher N_H in high-z bins represents the combined dependence of N_H on L and z. Summing up the intrinsic N_H distributions in the four redshift bins, we plot the intrinsic N_H distribution of our whole sample in Figure 30. It is consistent with that found by Tozzi et al. (2006) below log $N_H < 23$. At log $N_H > 23$, we find a larger fraction of highly obscured sources with log $N_H > 23.5$. This is because the
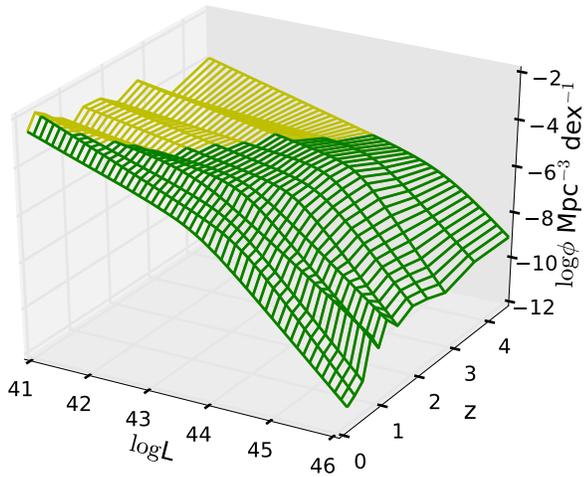
Figure 28. 2–10 keV intrinsic luminosity functions of AGNs in a series of redshift bins, obtained combining the results by Miyaji et al. (2015) and Georgakakis et al. (2015). The redshift grid values correspond to the central values of the redshift bins. The green part is covered by our sample assuming that AGN have no intrinsic obscuration. Note that this boundary corresponds to our sample-selection threshold, not the source-detection threshold. The 7Ms CDF-S catalog (Luo et al. 2017) extends well beyond this boundary in the L–z space.
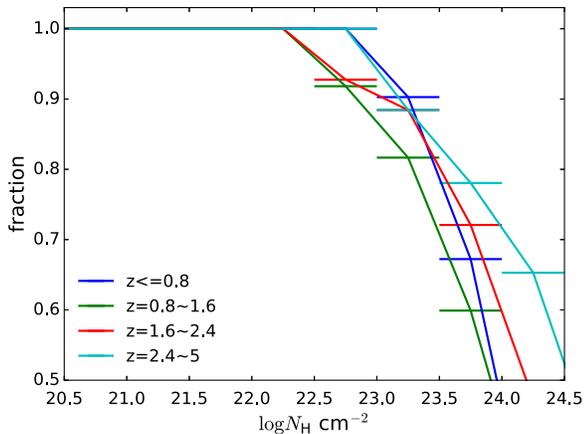
Figure 29. Fraction of sources detectable within our sample with a given N_H, computed for the luminosity range log$L$ within 42–45, in four redshift bins.

samples are different. In this work, the seven times longer exposure allows spectral analyses on more low-L and high-z sources, which tend to have higher N_H, as we show in Section 5.3. Instead of selecting the bright sample on the basis of soft, hard, and total band emission, as done in Tozzi et al. (2006), the sample selection in this work only makes use of the hard-band emission. Therefore, because we select fewer sources with very soft spectra, our sources in this work are systematically harder and thus more obscured. Additionally, we note that the updated redshift measurements are systematically higher than those used in Tozzi et al. (2006) (Section 4.8), such that the degeneracy between z and N_H in the spectral fitting also leads to higher N_H measurements in this work.

5.3. Luminosity Dependence and Evolution

5.3.1. Selection of Subsamples Unbiased with Respect to N_H

To investigate the luminosity and redshift dependence of N_H, we split the sample into 2D bins in the L–z space. This allows
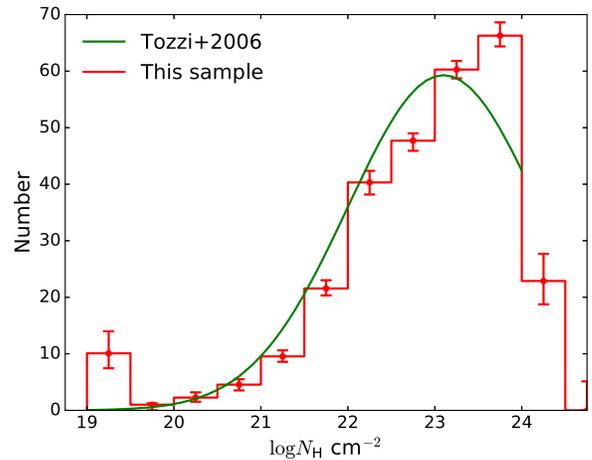

Figure 30. Intrinsic N_H distribution of our whole sample, measured by summing up the intrinsic N_H distributions in the four redshift bins ($z < 5$) as shown in Figure 27. The 90% percentile luminosity range of our sample is $42.41 \sim 44.45$. The green line that is derived from the 1Ms CDF-S by Tozzi et al. (2006), is a log-normal distribution centered at log $N_H = 23.1$ with a $\sigma = 1.1$ and normalized to the number of AGN used in this work between log N_H of 19 and 24.

us to check for a luminosity dependence of N_H at the same redshift and check for redshift dependence at the same luminosity. Clearly, two prerequisite conditions are required: first, wide ranges in the L–z space must be covered by the survey; second, the sample in each bin must be complete. As the deepest X-ray survey, the CDF-S provides an essential jigsaw piece among the X-ray surveys, which extends the coverage of X-ray surveys to the lowest luminosity and the highest redshift. However, in some bins at low luminosity and high redshift, the flux-limited sample is incomplete with respect to N_H. As shown in Figure 25, the completeness boundaries corresponding to different N_H are widely distributed in the L–z space. The sample will be biased against high N_H sources unless we apply the highest boundary, which corresponds to the highest N_H and largest off-axis angle, to the whole sample. This is clearly infeasible because it excludes most of the sources in the CDF-S. In the previous section, we handled this incompleteness with respect to N_H by correcting for the unobserved AGN population assuming a luminosity function. The disadvantages of this correction are that it is dependent on the uncertain part of the luminosity function at low L and high z, and it must be performed every time before this sample is used jointly with other X-ray surveys in the future. Therefore, in this section we select subsamples that are complete with respect to N_H by trimming the incomplete part while still maximizing the CDF-S sample. As the flux limit is irrelevant here, we include the supplementary sample prepared in Section 2.5 in selecting the complete subsamples. Such complete subsamples can be directly used in joint studies with other X-ray surveys. Note here that by "complete" we mean unbiased with respect to N_H at specific L and z values above the completeness boundaries, not with respect to L or z. The complete subsamples describe only the "observable" AGN population, leaving the low-L and high-z sources below the boundaries and the extremely obscured sources beyond a specific N_H unaccounted for.

As shown in Figure 25, as a function of the off-axis angle, we have determined a completeness boundary for each specific N_H. Based on these boundaries, we define a "completeness" flag as follows. Excluding the sources with intrinsic
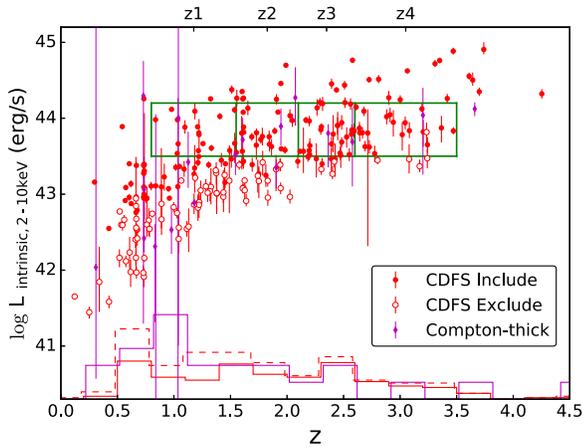
Figure 31. Scatter plot of the AGN sample in the luminosity—redshift space, with 90% errors on luminosity. The $\log N_H < 24$ N_H-complete subsample is plotted with red solid points (marked as "Include") and the other Compton-thin sources with red empty points (marked as "Exclude"). At the bottom, the normalized redshift distributions of the complete subsample ("Include") and the whole sample ("Include"+"Exclude") are shown as red solid and dashed histograms, respectively. The Compton-thick sources are also shown with the magenta points and the blue histogram. The four green boxes, which are marked as $z1$, $z2$, $z3$, and $z4$, are the regions selected to analyze the N_H dependence on redshift. Empty points are excluded from this analysis even if they fall in the selected bins.

luminosities below the $\log N_H = 24$ boundaries corresponding to their off-axis angles, we select sources that pass the filter of the $\log N_H = 24$ completeness boundaries, and flag them as "completeness" = 24. Similarly, we filter our sample with the $\log N_H = 23.5$ and $\log N_H = 23$ boundaries, and flag the selected sources with "completeness" of 23.5 and 23, respectively. A source selected by more than one filter is assigned the highest flag, a source selected by none is flagged as 0. Clearly, sources with a higher completeness flag represent a subsample of sources with a lower completeness flag. Under this definition, we can simply select an N_H-complete subsample by requiring the completeness flag $\geqslant$ a specific N_H of 23, 23.5, or 24, with the aim of studying the N_H distribution below this specific N_H. The $\log N_H < 24$ complete subsample and accordingly excluded sources are shown in Figure 31.

5.3.2. Evolution of N_H with Redshift

The unprecedented survey depth of the CDF-S provides the best opportunity to investigate the redshift dependence of AGN obscuration. As shown in Figure 31, we select a narrow luminosity band between 43.5 and 44.2, within which our sample has a large range of redshift. Then we select four redshift bins at grid points of $z = 0.8$, 1.55, 2.1, 2.6, and 3.5. For each bin that contains a subsample unbiased with respect to $\log N_H < 24$, we calculate the average $\log N_H$ and the obscured fraction—the fraction of N_H values between 10^{22}–10^{24} cm^{-2} of all those below 10^{24} cm^{-2}. We take the N_H measurement error into account through a bootstrapping procedure. For each spectrum of each Compton-thin source, we have generated 1000 random N_H values following its N_H error distribution, as done in the N_H distribution resampling. The random $\log N_H$ below 24 are used to measure the 68% confidence range of $\langle \log N_H \rangle$ and obscured fraction. When calculating $\langle \log N_H \rangle$, $\log N_H$ below 19 are set to 19. The subsample size, $\langle \log N_H \rangle$, and obscured fractions in each bin are shown in the upper panel of Table 6. Both quantities are

Table 6
Number of Sources, Median 0.5–7 keV net Counts, Average of log N_H, and Obscured Fraction in Each Bin, Obtained from the CDF-S Data

	Whole Subsample			
Bins	$z1$	$z2$	$z3$	$z4$
Number	17	21	26	20
Median Cts	2224	770	491	381
$\langle \log N_H \rangle$	$22.36^{+0.14}_{-0.14}$	$22.64^{+0.13}_{-0.13}$	$22.95^{+0.10}_{-0.10}$	$23.11^{+0.10}_{-0.10}$
ObsFrac	$0.67^{+0.06}_{-0.06}$	$0.76^{+0.05}_{-0.05}$	$0.91^{+0.03}_{-0.03}$	$0.95^{+0.03}_{-0.03}$

	Spectroscopic-z Only		
Bins	$z1$	$z2$	$z3$
Number	12	12	12
Median Cts	2934	1312	476
$\langle \log N_H \rangle$	$22.18^{+0.17}_{-0.17}$	$22.60^{+0.14}_{-0.14}$	$22.79^{+0.13}_{-0.12}$
ObsFrac	$0.58^{+0.08}_{-0.07}$	$0.74^{+0.07}_{-0.07}$	$0.90^{+0.04}_{-0.04}$

Note. The obscured fraction is the fraction of N_H values between 10^{22}–10^{24} cm^{-2} of all those below 10^{24} cm^{-2}. In the upper panel, the data are obtained using the CDF-S subsample that is unbiased with respect to log $N_H < 24$. The data in the lower panel are obtained after excluding the photometric-z sources.
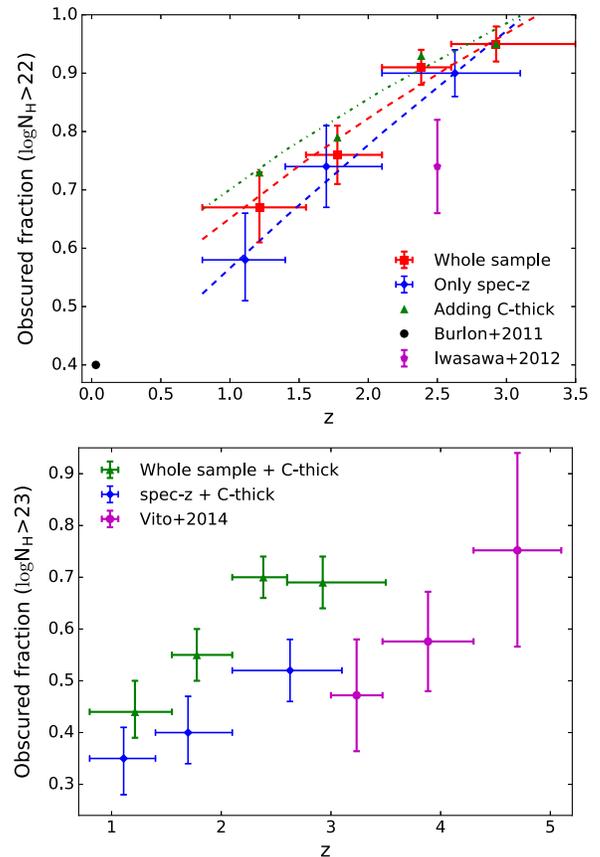

Figure 32. Upper panel: fractions of sources with log $N_H > 22$ in the four bins ($0.8 < z < 3.5$, $43.5 < \log L < 44.2$) as shown in Figure 31 and best-fit lines. Results obtained from the whole "completeness" = 24 sample are plotted in red, the spectroscopic-z subsample in blue. Results obtained after adding the Compton-thick sources to the whole sample are shown in green; the error bars are omitted for simplicity. The Burlon et al. (2011) and Iwasawa et al. (2012) results correspond to $\log L \approx 43.85$ and $\log L \approx 44.2$, respectively. Lower panel: fractions of sources with log $N_H > 23$. Results from our sample adding the Compton-thick sources are plotted in green, from the spectroscopic-z subsample adding the Compton-thick sources are plotted in blue. Results from the $42.85 < \log L < 44.5$ sample of Vito et al. (2014) are plotted in magenta.

increasing with redshift. The obscured fractions are plotted in Figure 32. We fit the correlation between obscured fraction and redshift with $f_{24} = \beta \, (1 + z)^\alpha$, and find best-fit values of $\beta = 0.42 \pm 0.09$ and $\alpha = 0.60 \pm 0.17$.

According to the result of Burlon et al. (2011), at $\log L = 43.85$, which is the average 2–10 keV luminosity of our selected luminosity range (43.5 $\sim$ 44.2), the AGN-obscured fraction at $z \sim 0.03$ is about 40%. As shown in the upper panel of Figure 32, the N_H evolution from local to $z = 2$ is very strong. However, above $z = 2$ the absorbed fraction is likely saturating and shows weak evolution. It is consistent with the result of Hasinger (2008). They found that the AGN-obscured fraction increases with z with an $\alpha = 0.62 \pm 0.11$, and becomes saturated at $z > 2$. Vito et al. (2014) also noted the weak evolution of N_H at high z.

Hasinger (2008) found an evolution slope of $\alpha = 0.48 \pm 0.08$ over a wide redshift range 0–3.2. Similarly, Ueda et al. (2014) found an $\alpha = 0.48 \pm 0.05$. To compare with these results, we take the Compton-thick sources into the account, as done in Hasinger (2008) and Ueda et al. (2014). We remark the shortcoming of including Compton-thick sources: the sampling of Compton-thick AGNs can be highly incomplete and the measurements of their intrinsic luminosity can be highly unreliable. As shown in the upper panel of Figure 32, the fraction is slightly increased, and the best-fit slope of $\alpha = 0.45^{+0.10}_{-0.09}$ is consistent with the previous results. Our Compton-thin sample shows a slightly steeper slope of 0.59 because we find fewer Compton-thick AGN at high z, as shown in Figure 31.

In the upper panel of Figure 32, we also show the result of Iwasawa et al. (2012), who found an obscured fraction among Compton-thin AGN of $74 \pm 8\%$ at $z \sim 2.5$ and $\log L \sim 44.2$. It is lower than the fraction measured in this work. To compare our results with those of Vito et al. (2014), who define an obscured fraction as the fraction of sources with $\log N_H > 23$ rather than $\log N_H > 22$, we also calculate this fraction in our selected redshift bins. As shown in the lower panel of Figure 32, compared with our results, the fractions obtained from their $42.85 < \log L < 44.5$ sample also appear to be lower. This deviation might be caused by different luminosities or incompleteness of their sample. Our conservativeness in identifying Compton-thick AGN might introduce an overestimation of the obscured fraction at high z, in the sense that a few Compton-thick AGNs might be misclassified as highly obscured Compton-thin because of the large uncertainty in the N_H measurement. However, there are only a few observable Compton-thick sources, and even if misclassified as Compton-thin, they are likely considered to be highly obscured sources with $\log N_H > 23$. Therefore, we consider this as a minor effect.

Considering the short duty cycle of AGN (10^7–10^8 years, e.g., Parma et al. 2007), the strong evolution of AGN N_H with z is obviously associated with the evolution of galaxy. One associated factor is the high merger rate at high-z, in the sense that merger triggers AGN and such AGN are more obscured (Di Matteo et al. 2005; Kocevski et al. 2015; Lanzuisi et al. 2015). Another factor is the higher gas fraction at high-z (Carilli & Walter 2013), which not only leads to a higher N_H directly, but also likely leads to a longer obscured phase and a larger covering factor.
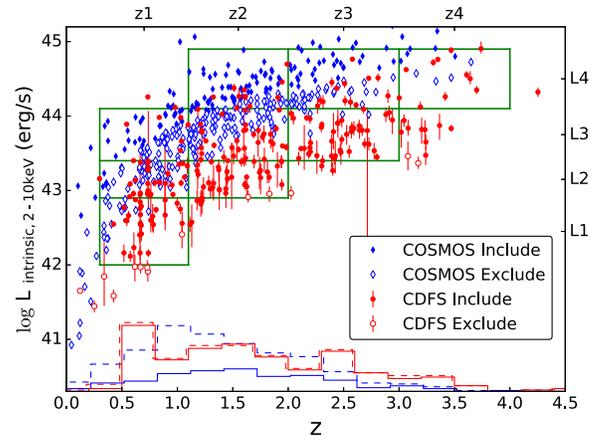

Figure 33. Same as Figure 31, but including C-COSMOS data (red for CDF-S, blue for C-COSMOS). The selected subsamples are complete for $\log N_H < 23$.

5.3.3. Decomposing the Luminosity- and Redshift-dependence of N_H

To complement the high-luminosity regime, which is not well sampled by the CDF-S, we add AGN observed from the C-COSMOS, a wider and shallower survey (Elvis et al. 2009; Lanzuisi et al. 2013; Civano et al. 2016). The 2 deg^2 COSMOS field was observed by *Chandra* for 160 ks. Lanzuisi et al. (2013) presented the properties of the AGNs in this survey. We combine this sample with ours, as shown in Figure 33. As done for our sample, similar "completeness boundaries" are calculated for C-COSMOS, considering its 160 ks exposure time and its sample-selection threshold of 70 net counts in the broad 0.5–7 keV band. Because of the shallow depth, C-COSMOS is severely incomplete for highly obscured AGN. Therefore, we only make use of the $\log N_H < 23$ part. The N_H-complete subsamples of C-COSMOS and CDF-S for $\log N_H < 23$ are shown in Figure 33.

We select a 4×4 grid for the combined sample at $0.3 < z1 < 1.1 < z2 < 2 < z3 < 3 < z4 < 4$ and $42 < L1 < 42.9 < L2 < 43.4 < L3 < 44.1 < L4 < 44.9$. The grids are selected following several criteria:

1. Exclude the high-L low-z corner because there are few sources, as limited by the total sky coverage.
2. Exclude the low-L high-z corner, which is below the sample-selection boundaries.
3. Choose a proper number of bins such that each cell contains a sufficient number of sources.
4. Distribute the sources evenly among the cells such that each cell contains approximately the same number of sources.
5. Distribute the sources evenly in each valid cell, avoiding an empty low-L high-z corner in any cell.

The number of sources in each bin is listed in Table 7. One cell at high-L low-z and six cells at low-L high-z are excluded. For the subsamples in the valid cells, the $\langle \log N_H \rangle$ below 23 are calculated similarly as above. The obscured fraction in each bin is calculated as the fraction of the N_H values between $10^{22} \sim 10^{23}$ cm^{-2} of all those below 10^{23} cm^{-2}. The results are shown in Figure 34 and Table 7, where excluded cells are left empty.

Limited by the sample size, the $\langle \log N_H \rangle$ cannot be constrained very tightly, especially at high redshift. However, trends of luminosity- and redshift-dependences are apparent. In each redshift bin, the mean N_H decreases with luminosity. In

Table 7
Number of Sources, Median 0.5–7 keV net Counts, Average of log N_H, and Obscured Fraction in each Bin, Obtained by Combining CDF-S and C-COSMOS

Whole Subsample				
Number	$z1$	$z2$	$z3$	$z4$
$L4$	7	68	39	13
$L3$	34	40	43	9
$L2$	35	33	3	0
$L1$	29	7	0	0
Median Cts	$z1$	$z2$	$z3$	$z4$
$L4$	...	284	235	1501
$L3$	494	654	335	...
$L2$	888	337	...	...
$L1$	303	...	...	...
$\langle \log N_H \rangle$	$z1$	$z2$	$z3$	$z4$
$L4$	...	$21.33^{+0.15}_{-0.15}$	$21.73^{+0.22}_{-0.21}$	$22.56^{+0.12}_{-0.14}$
$L3$	$21.32^{+0.17}_{-0.17}$	$21.89^{+0.11}_{-0.10}$	$22.06^{+0.15}_{-0.16}$	...
$L2$	$21.55^{+0.14}_{-0.13}$	$22.11^{+0.11}_{-0.10}$	...	...
$L1$	$21.81^{+0.11}_{-0.12}$	...	...	...
ObsFrac	$z1$	$z2$	$z3$	$z4$
$L4$	...	$0.40^{+0.05}_{-0.06}$	$0.67^{+0.07}_{-0.07}$	$0.96^{+0.04}_{-0.06}$
$L3$	$0.32^{+0.06}_{-0.07}$	$0.57^{+0.05}_{-0.06}$	$0.76^{+0.05}_{-0.05}$	...
$L2$	$0.50^{+0.05}_{-0.05}$	$0.75^{+0.05}_{-0.06}$	...	...
$L1$	$0.63^{+0.05}_{-0.05}$	...	...	...
Spectroscopic-z Only				
Number	$z1$	$z2$	$z3$	
$L3$	1	29	24	
$L2$	37	35	16	
$L1$	54	13	7	
Median Cts	$z1$	$z2$	$z3$	
$L3$	...	354	220	
$L2$	518	488	566	
$L1$	785	228	...	
$\langle \log N_H \rangle$	$z1$	$z2$	$z3$	
$L3$	...	$21.13^{+0.25}_{-0.26}$	$21.31^{+0.29}_{-0.29}$	
$L2$	$21.29^{+0.15}_{-0.15}$	$21.79^{+0.12}_{-0.12}$	$22.10^{+0.16}_{-0.15}$	
$L1$	$21.65^{+0.09}_{-0.09}$	...	...	
ObsFrac	$z1$	$z2$	$z3$	
$L3$	...	$0.42^{+0.08}_{-0.08}$	$0.51^{+0.09}_{-0.10}$	
$L2$	$0.28^{+0.06}_{-0.06}$	$0.51^{+0.06}_{-0.06}$	$0.72^{+0.06}_{-0.07}$	
$L1$	$0.54^{+0.04}_{-0.04}$	...	...	

Note. The median counts, $\langle \log N_H \rangle$, and obscured fractions are given only for the valid cells. The obscured fraction corresponds to the fraction of the N_H values between $10^{22} \sim 10^{23}$ cm^{-2} of all those below 10^{23} cm^{-2}. The data in the upper panel correspond to the subsample that is unbiased with respect to log $N_H < 23$. In the lower panel, the data are obtained after excluding the photometric-z sources.

each luminosity bin, the mean N_H increases with redshift. Obscured fractions show similar luminosity- and redshift-dependences. Limited by the luminosity range of our sample, we are not able to see any turnover of the N_H–luminosity anticorrelation at low luminosities.
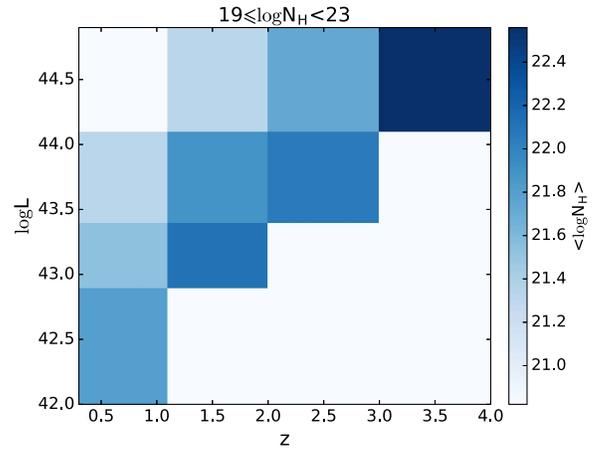

Figure 34. Mean log N_H in grid, combining CDF-S and C-COSMOS data. Only values below 23 are considered. Values below 19 are taken as 19. White cells are excluded.

5.3.4. Discussion of Other Systematic Effects

Our analysis so far aimed at minimizing the effects of selection bias on the correlation between N_H, luminosity and redshift. In this section, we discuss other systematic effects related to degeneracy between spectral parameters.

The best-fit N_H and intrinsic luminosity L are positively correlated, in the sense that higher N_H gives rise to higher absorption-correction and thus higher L. The degeneracy between them can be strong in low S/N and high N_H cases, where the uncertainty of N_H is large. In Sections 5.3.2 and 5.3.3, we have managed to attenuate this effect while comparing the N_H distribution in different bins in the L–z space. First, we consider only the uncertainty of N_H but not the uncertainty of L while dividing the sources into bins. Second, we keep the valid cells away from sample-selection boundaries (grid selection criterion 5 in Section 5.3.3) to reduce boundary effect, that is, including high-L high-N_H sources from the below-boundary region and excluding low-L low-N_H sources from the above-boundary region because of the uncertainty of L and the N_H-L degeneracy. The positive N_H–L degeneracy could weaken the negative correlation we found between average N_H and L. However, as shown in Figure 34 and the upper panel of Table 7, the average N_H clearly declines at higher L, indicating that the effect of the N_H-L degeneracy is negligible.

An accurate redshift is essential in measuring both N_H and L. In our sample, 37% of the sources have photometric redshifts, and they lie at relatively high-z. For such a photometric-z source, the redshift value, which has a large uncertainty, is positively correlated with N_H since the absorption feature in the spectrum is considered to occur at a higher energy with a larger z; and z is positively correlated with L as a larger z produces a larger luminosity distance. The N_H-z degeneracy enhances the N_H evolution we found; and the L–z degeneracy affects the thoroughness of our decomposition of the L-dependence and z-dependence of average N_H. To test the robustness of our results, we repeat the experiments without the photometric-z sources.

Since we have few sources in this case, instead of the four redshift bins show in Figure 31, we choose three redshift bins at grids of $z = 0.8$, 1.4, 2.1, and 3.1, and calculate the $\langle \log N_H \rangle$ and obscured fraction of the spectroscopic-z subsample in each bin in the same manner as described above. The results are compared with the total sample in Table 6 and Figure 32. We
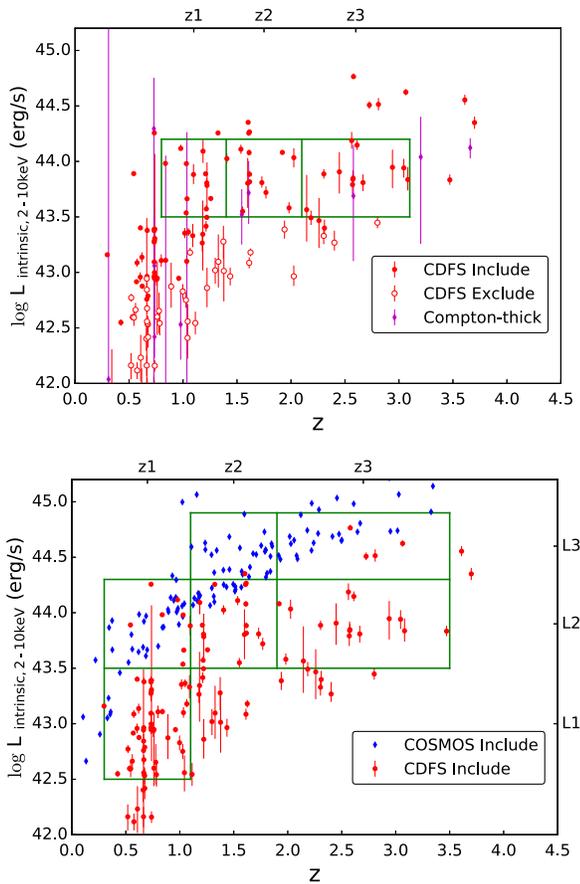
Figure 35. Upper and lower panels are the same as Figures 31 and 33, respectively, but only include spectroscopic-z sources.

find that compared with the results obtained from the whole sample, the log $N_H > 22$ obscured fractions are reduced slightly, while the log $N_H > 23$ obscured fractions are reduced significantly, especially at high z.

The 2D distribution of average N_H using the spectroscopic-z subsample also deserves consideration. We rebin this subsample to a 3×3 grid at $0.3 < z1 < 1.1 < z2 < 1.9 < z3 < 3.5$ and $42.5 < L1 < 43.5 < L2 < 44.3 < L3 < 44.9$, as shown in Figure 35. The $\langle \log N_H \rangle$ and obscured fraction are calculated in the same manner and are shown in the lower panel of Table 7. Again, the $\langle \log N_H \rangle$ and obscured fraction are reduced for the spectroscopic-z subsample, but it is still clear that the obscuration decreases with L at the same z and increases with z at the same L.

The reduced fraction of obscured AGN in the spectroscopic-z subsample is expected by the N_H-z degeneracy, which leads to higher N_H at high-z and lower N_H at low-z. However, the spectroscopic-z subsample is likely biased against obscured AGN, in the sense that X-ray obscured AGNs also suffer more extinction in the optical band and are thus less likely to be spectroscopically observed. Therefore, the difference can only be partly attributed to the N_H-z degeneracy; the results from the spectroscopic-z subsample should be considered as a conservative estimation of the N_H dependence on z.

We conclude that the luminosity- and redshift-dependences of N_H that we have found are robust, having all the systematics understood.

6. Summary

Over the past 16 years, the CDF-S has been cumulatively observed for 7Ms, making it the deepest X-ray survey to date. We present a detailed spectral analysis for a sample of the brightest AGNs in the 7Ms CDF-S. This sample includes 276 sources that (1) are classified as AGN on the basis of multiband information, (2) are detected with at least 80 hard-band net counts, (3) have a redshift measurement, and (4) lie within an off-axis angle of $9.'5$. The new 7Ms CDF-S data and catalog provide not only spectra with an improved S/N, but also updated redshift measurements compared to those used in previous works (e.g., Tozzi et al. 2006; Buchner et al. 2014). On the basis of these improvements, we perform a systematic X-ray spectral analysis for the bright AGN sample, with the emphasis on the properties of AGN intrinsic obscuration. Our standard spectral model is composed of an intrinsically absorbed power-law and a "self-absorbed" cold-reflection component that describes a typical reflection from the dusty torus. A soft excess and/or a neutral Fe Kα line are added to the model when the component improves the fitting significantly. The analysis is performed both on the total 7Ms exposures and in four different periods across the 16 years. In this way, we are able to exploit the high S/N of the cumulative spectra whenever possible, and also to identify significant spectral variabilities on timescales of a few years. The spectral fitting results are summarized as follows.

1. We present the power-law slopes, column densities, observed fluxes, and absorption-corrected luminosities for our sample of AGNs. We also present the net counts and fluxes of each AGN in each period. For sources showing N_H variability, the different N_H in different periods are given.

2. Narrow Fe Kα lines are found in 50 sources and broad Fe Kα lines are found in 5 sources. The EWs of the narrow Fe Kα lines show a clear correlation with the N_H of the sources, which is well explained by a toy model assuming a constant line flux independent of the line-of-sight N_H. This line flux corresponds to an EW of 135 eV when the power-law is unobscured.

3. We find N_H variation in 39 sources. By checking the correlation between the variations of N_H and intrinsic X-ray luminosity L, we show two types of N_H variations: one anticorrelated with L, with a relatively small amplitude, and one independent of L, with a large amplitude. They can be attributed to N_H depression by the central engine or to an obscuring cloud that moves across the line of sight, respectively.

4. On the basis of the X-ray spectral shape, Fe line EW, and relative strength of X-ray to MIR 12 μm emission, we identify 22 (8% of the sample) Compton-thick candidates. These sources show a systematically lower X-ray-to-MIR ratio.

Thanks to the well-understood selection function of our sample, we are able to measure quantitatively the sample-selection biases that impact the observed N_H distribution, including the sky-coverage effect, which biases against sources with low observed flux, the N_H-dependent Malmquist bias, which leads to different coverage in the L_X–z space of the AGN population for different N_H, and the Eddington bias. Based on thorough analyses of these effects, we are able to recover the

intrinsic distribution and evolution of AGN obscuration from the observed data. The results are summarized as follows.

1. Correcting the sample-selection biases, we recover the intrinsic distribution of N_H of an AGN population that is well defined by our sample-selection function. The intrinsic N_H distribution changes significantly between different redshift bins because of the strong dependence of N_H on L and z. Our hard-band selected sample contains more highly obscured sources than the sample used in the 1Ms CDF-S AGN spectral analysis work by Tozzi et al. (2006), thus showing a higher peak value of N_H (between $10^{23.5}$ and 10^{24} cm^{-2}) than therein.

2. As the deepest X-ray survey at present and in the foreseeable future, the 7Ms CDF-S provides the unique opportunity to study the low-L—high-z AGN population. The limitations of such a pencil-beam survey are also obvious, as it has a limited sky coverage and thus does not sample the high-L AGN population well. Therefore, it is essential to combine the 7Ms CDF-S with other wide and shallow surveys, in order to study the dependence of N_H on L and z. With this goal, we define a "completeness" flag for each source on the basis of the modeling of the sample-selection bias, with which one can easily trim the "incomplete part" from our sample and select a subsample that is unbiased with respect to N_H in the range below a specific N_H. Such N_H-complete samples can be easily used in a joint study with other or future surveys to investigate the obscuration of AGN, as the selection bias has been eliminated.

3. Based on our N_H-complete sample, we measure the average N_H and obscured AGN fraction as a function of redshift in a narrow luminosity range ($43.5 \sim 44.2$). We find a strong evolution of the obscured fraction with z, which can be expressed as $f_{24} = 0.42 \pm 0.09 (1 + z)^{0.60 \pm 0.17}$. At $z > 2$, the obscured fraction likely saturates, showing a weak evolution. The obscured fraction measured from our sample is higher than those measured in previous works.

4. Using our N_H-complete sample, we are also able to distinguish the luminosity dependence and redshift dependence of N_H. Combining our data with that from the wider and shallower survey C-COSMOS (Lanzuisi et al. 2013), we measure the average N_H and obscured fraction in 2D L_X–z bins with L_X between 10^{42} and 10^{45} erg s^{-1} and the redshift up to our limit of measurement ($z \approx 4$). We find that at any redshift, the average N_H (or obscured fraction) decreases with intrinsic X-ray luminosity L_X, and at any L_X, it increases with redshift.

Appendix
Weight Selection in Spectral Response File Combination

For a given source, the intrinsic flux S_j (in units of erg cm^{-2} s^{-1}) incident on the ACIS-I instrument during ObsID j produces photon counts $C_j(I)$ in the energy channel I:

$$C_j(I) = \int dE \, S_j(E) T_j A_j(E) R_j(E, I), \tag{1}$$

where $A_j(E)$ and $R_j(E, I)$ are the ancillary and matrix response files, specific to the ObsID and the position of the source in the field of view.

Stacking the spectra from all the ObsID in a period, we have the stacked spectrum C_{tot} and the total exposure time T_{tot}:

$$C_{tot}(E) \equiv \sum_j C_j(E)$$

$$T_{tot} \equiv \sum_j T_j.$$

To perform spectral analyses on the stacked spectrum, we use a relation of the same kind as Equation (1):

$$C_{tot}(I) = \int dE \, \tilde{S}(E) T_{tot} \tilde{A}(E) \tilde{R}(E, I), \tag{2}$$

and based on the two expressions of C_{tot},

$$\sum_j \int dE \, S_j(E) T_j A_j(E) R_j(E, I)$$
$$= \int dE \, \tilde{S}(E) T_{tot} \tilde{A}(E) \tilde{R}(E, I), \tag{3}$$

we need to find proper expressions for $\tilde{A}(E)$ and $\tilde{R}(E, I)$ so that the average flux we measured from the stacked spectrum is correct, as in this form:

$$\tilde{S}(E) = \frac{\sum_j S_j(E) T_j}{T_{tot}}.$$

The RMF $R(E, I)$ affects the spectrum C in each channel. But as a normalized response matrix, its impact on the broadband integrated flux is secondary compared to that of the ARF, which directly affects the flux measurement in terms of effective area. Moreover, the ARF is variable because of

both the significant degradation of CCD quantum efficiency over the past 16 years (as shown in Figure 2) and the vignetting effect at positions off-axis from the aimpoint, while the RMF is relatively constant. Therefore, we can consider the RMF of a source as roughly constant and simplify it from Equation (3). To average the RMF from each ObsID within a period, we simply use the broadband photon counts as weights, while for the ARF, which is of more significance in flux calibration, we derive the optimized weight from Equation (3), in order to achieve the most accurate definition of average flux. The solution is straightforward in the case of constant flux $S(E)$. But with the aim of studying AGNs that are highly variable on timescales of months, we take $S(E)$ as time-dependent. Thus in addition to time and position dependence of the ACIS-I effective area, we also keep track of source variability in the weight.

After simplifying $R(E, I)$, the left-hand side of Equation (3) can be written as

$$\sum_j \int dE \, S_j(E) T_j \, A_j(E)$$
$$= \int dE \sum_j S_j(E) T_j \, A_j(E)$$
$$= \int dE \, \frac{\sum_j S_j(E) T_j}{\sum_j T_j} \, \frac{\sum_j T_j}{1} \, \frac{\sum_j S_j(E) T_j A_j(E)}{\sum_j S_j(E) T_j}$$
$$= \int dE \, \tilde{S}(E) T_{\text{tot}} \, \frac{\sum_j S_j(E) T_j A_j(E)}{\sum_j S_j(E) T_j}.$$

By a direct comparison with the right-hand side, we find

$$\tilde{A}(E) = \frac{\sum_j S_j(E) T_j A_j(E)}{\sum_j S_j(E) T_j}.$$

Obviously, the optimized weight is $S_j(E) T_j$.

Consider the broadband photon counts C_j, the energy-averaged flux S_j, and the mean effective area $\bar{A}_j$ at the effective energy of the broad band, 2.3 keV, we have $C_j \simeq S_j \times T_j \times \bar{A}_j$. Thus we can use $C_j/\bar{A}_j$ as an approximation of $S_j(E) T_j$.

References

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., Bauer, F. E., Chapman, S. C., et al. 2005, ApJ, 632, 736
Alexander, D. M., & Hickox, R. C. 2012, NewAR, 56, 93
Antonucci, R. 1993, ARA&A, 31, 473
Arnaud, K. A. 1996, in ASP Conf. Ser. 101, Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby & J. Barnes (San Francisco, CA: ASP), 17
Asmus, D., Gandhi, P., Hönig, S. F., Smette, A., & Duschl, W. J. 2015, MNRAS, 454, 766
Assef, R. J., Eisenhardt, P. R. M., Stern, D., et al. 2015, ApJ, 804, 27
Ballantyne, D. R., Shi, Y., Rieke, G. H., et al. 2006, ApJ, 653, 1070
Bassani, L., Dadina, M., Maiolino, R., et al. 1999, ApJS, 121, 473
Bianchi, S., Guainazzi, M., & Chiaberge, M. 2006, A&A, 448, 499
Blackburn, J. K. 1995, in ASP Conf. Ser. 77, Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes (San Francisco, CA: ASP), 367
Bonzini, M., Padovani, P., Mainieri, V., et al. 2013, MNRAS, 436, 3759
Brandt, W. N., Alexander, D. M., Hornschemeier, A. E., et al. 2001, AJ, 122, 2810
Brightman, M., & Nandra, K. 2011, MNRAS, 413, 1206
Brightman, M., Nandra, K., Salvato, M., et al. 2014, MNRAS, 443, 1999
Brightman, M., Silverman, J. D., Mainieri, V., et al. 2013, MNRAS, 433, 2485
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., & Bauer, F. E. 2017, MNRAS, 465, 4348
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Buchner, J., Georgakakis, A., Nandra, K., et al. 2015, ApJ, 802, 89
Burlon, D., Ajello, M., Greiner, J., et al. 2011, ApJ, 728, 58
Carilli, C. L., & Walter, F. 2013, ARA&A, 51, 105
Cash, W. 1979, ApJ, 228, 939
Castelló-Mor, N., Carrera, F. J., Alonso-Herrero, A., et al. 2013, A&A, 556, A114
Chang, Y.-Y., Le Floc'h, E., Juneau, S., et al. 2017, MNRAS, 466, L103
Chen, C.-T. J., Hickox, R. C., Alberts, S., et al. 2015, ApJ, 802, 50
Civano, F., Marchesi, S., Comastri, A., et al. 2016, ApJ, 819, 62
Comastri, A. & XMM-CDFS Team 2013, MmSAI, 84, 661
Comastri, A., Ranalli, P., Iwasawa, K., et al. 2011, A&A, 526, L9
Corral, A., Georgantopoulos, I., Comastri, A., et al. 2016, A&A, 592, A109
Crummy, J., Fabian, A. C., Gallo, L., & Ross, R. R. 2006, MNRAS, 365, 1067
Daddi, E., Alexander, D. M., Dickinson, M., et al. 2007, ApJ, 670, 173
Del Moro, A., Alexander, D. M., Bauer, F. E., et al. 2016, MNRAS, 456, 2105
Del Moro, A., Mullaney, J. R., Alexander, D. M., et al. 2014, ApJ, 786, 16
Di Matteo, T., Springel, V., & Hernquist, L. 2005, Natur, 433, 604
Dwelly, T., & Page, M. J. 2006, MNRAS, 372, 1755
Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119
Elitzur, M., & Ho, L. C. 2009, ApJL, 701, L91
Ellison, S. L., Teimoorinia, H., Rosario, D. J., & Mendel, J. T. 2016, MNRAS, 458, L34
Elvis, M., Civano, F., Vignali, C., et al. 2009, ApJS, 184, 158
Fanali, R., Caccianiga, A., Severgnini, P., et al. 2013, MNRAS, 433, 648
Feruglio, C., Daddi, E., Fiore, F., et al. 2011, ApJL, 729, L4
Finoguenov, A., Tanaka, M., Cooper, M., et al. 2015, A&A, 576, A130
Fiore, F., Grazian, A., Santini, P., et al. 2008, ApJ, 672, 94
Fiore, F., Puccetti, S., Brusa, M., et al. 2009, ApJ, 693, 447
Fruscione, A., McDowell, J. C., Allen, G. E., et al. 2006, Proc. SPIE, 6270, 62701V
Gandhi, P., Annuar, A., Lansbury, G. B., et al. 2017, MNRAS, 467, 4606
Gandhi, P., Horst, H., Smette, A., et al. 2009, A&A, 502, 457
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
Giacconi, R., Zirm, A., Wang, J., et al. 2002, ApJS, 139, 369
Gilli, R., Comastri, A., & Hasinger, G. 2007, A&A, 463, 79
Gilli, R., Comastri, A., Vignali, C., Ranalli, P., & Iwasawa, K. 2010, in AIP Conf. Ser. 1248, X-ray Astronomy 2009; Present Status, Multi-Wavelength Approach and Future Perspectives (Melville, NY: AIP), 359
Gilli, R., Norman, C., Vignali, C., et al. 2014, A&A, 562, A67
Goulding, A. D., Alexander, D. M., Bauer, F. E., et al. 2012, ApJ, 755, 5
Guainazzi, M., & Bianchi, S. 2007, MNRAS, 374, 1290
Hasinger, G. 2008, A&A, 490, 905
Hasinger, G., Cappelluti, N., Brunner, H., et al. 2007, ApJS, 172, 29
Hiroi, K., Ueda, Y., Akiyama, M., & Watson, M. G. 2012, ApJ, 758, 49
Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, ApJS, 163, 1
Ichikawa, K., Ricci, C., Ueda, Y., et al. 2017, ApJ, 835, 74
Ichikawa, K., Ueda, Y., Terashima, Y., et al. 2012, ApJ, 754, 45
Isobe, N., Kawamuro, T., Oyabu, S., et al. 2016, PASJ, 68, 98
Iwasawa, K., Gilli, R., Vignali, C., et al. 2012, A&A, 546, A84
Kocevski, D. D., Brightman, M., Nandra, K., et al. 2015, ApJ, 814, 104
Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511
Koss, M. J., Romero-Cañizales, C., Baronchelli, L., et al. 2015, ApJ, 807, 149
La Franca, F., Fiore, F., Comastri, A., et al. 2005, ApJ, 635, 864
Laird, E. S., Nandra, K., Georgakakis, A., et al. 2009, ApJS, 180, 102
Lamastra, A., Perola, G. C., & Matt, G. 2006, A&A, 449, 551
Lanzuisi, G., Civano, F., Elvis, M., et al. 2013, MNRAS, 431, 978
Lanzuisi, G., Ranalli, P., Georgantopoulos, I., et al. 2015, A&A, 573, A137
Lawrence, A. 1991, MNRAS, 252, 586
Lawrence, A., & Elvis, M. 1982, ApJ, 256, 410
Lawrence, A., & Elvis, M. 2010, ApJ, 714, 561
Leahy, D. A., & Creighton, J. 1993, MNRAS, 263, 314
Lehmer, B. D., Basu-Zych, A. R., Mineo, S., et al. 2016, ApJ, 825, 7
Lehmer, B. D., Brandt, W. N., Alexander, D. M., et al. 2005, ApJS, 161, 21
Lehmer, B. D., Brandt, W. N., Alexander, D. M., et al. 2008, ApJ, 681, 1163
Lehmer, B. D., Xue, Y. Q., Brandt, W. N., et al. 2012, ApJ, 752, 46
Levenson, N. A., Krolik, J. H., Życki, P. T., et al. 2002, ApJL, 573, L81
Liu, T., & Wang, J.-X. 2010, ApJ, 725, 2381

Liu, T., Wang, J.-X., Yang, H., Zhu, F.-F., & Zhou, Y.-Y. 2014, ApJ, 783, 106
Liu, Y., & Li, X. 2015, MNRAS, 448, L53
Luo, B., Bauer, F. E., Brandt, W. N., et al. 2008, ApJS, 179, 19
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
Maiolino, R., Salvati, M., Bassani, L., et al. 1998, A&A, 338, 781
Maiolino, R., Shemmer, O., Imanishi, M., et al. 2007, A&A, 468, 979
Marchesi, S., Lanzuisi, G., Civano, F., et al. 2016, ApJ, 830, 100
Matt, G., Marinucci, A., Guainazzi, M., et al. 2014, MNRAS, 439, 3016
Mehdipour, M., Branduardi-Raymont, G., Kaastra, J. S., et al. 2011, A&A, 534, A39
Miyaji, T., Hasinger, G., Salvato, M., et al. 2015, ApJ, 804, 104
Murphy, K. D., & Yaqoob, T. 2009, MNRAS, 397, 1549
Nandra, K., & Pounds, K. A. 1994, MNRAS, 268, 405
Netzer, H. 2015, ARA&A, 53, 365
Norman, C., Hasinger, G., Giacconi, R., et al. 2002, ApJ, 571, 218
Nousek, J. A., & Shue, D. R. 1989, ApJ, 342, 1207
Page, M. J., Stevens, J. A., Ivison, R. J., & Carrera, F. J. 2004, ApJL, 611, L85
Paolillo, M., Schreier, E. J., Giacconi, R., Koekemoer, A. M., & Grogin, N. A. 2004, ApJ, 611, 93
Parma, P., Murgia, M., de Ruiter, H. R., et al. 2007, A&A, 470, 875
Pierre, M., Chiappetti, L., Pacaud, F., et al. 2007, MNRAS, 382, 279
Pierre, M., Pacaud, F., Adami, C., et al. 2016, A&A, 592, A1
Prieto, M. A., Mezcua, M., Fernández-Ontiveros, J. A., & Schartmann, M. 2014, MNRAS, 442, 2145
Protassov, R., van Dyk, D. A., Connors, A., Kashyap, V. L., & Siemiginowska, A. 2002, ApJ, 571, 545
Rafferty, D. A., Brandt, W. N., Alexander, D. M., et al. 2011, ApJ, 742, 3
Raimundo, S. I., Fabian, A. C., Bauer, F. E., et al. 2010, MNRAS, 408, 1714
Ranalli, P., Comastri, A., Vignali, C., et al. 2013, A&A, 555, A42
Ricci, C., Walter, R., Courvoisier, T. J.-L., & Paltani, S. 2011, A&A, 532, A102
Risaliti, G., Elvis, M., & Nicastro, F. 2002, ApJ, 571, 234
Risaliti, G., Maiolino, R., & Salvati, M. 1999, ApJ, 522, 157
Risaliti, G., Young, M., & Elvis, M. 2009, ApJL, 700, L6
Rosati, P., Tozzi, P., Giacconi, R., et al. 2002, ApJ, 566, 667
Saez, C., Chartas, G., Brandt, W. N., et al. 2008, AJ, 135, 1505
Sazonov, S., Churazov, E., & Krivonos, R. 2015, MNRAS, 454, 1202
Severgnini, P., Caccianiga, A., & Della Ceca, R. 2012, A&A, 542, A46
Shemmer, O., Brandt, W. N., Netzer, H., Maiolino, R., & Kaspi, S. 2006, ApJL, 646, L29
Shemmer, O., Brandt, W. N., Netzer, H., Maiolino, R., & Kaspi, S. 2008, ApJ, 682, 81
Shu, X. W., Yaqoob, T., & Wang, J. X. 2010, ApJS, 187, 581
Shu, X. W., Yaqoob, T., & Wang, J. X. 2011, ApJ, 738, 147
Simcoe, R., McLeod, K. K., Schachter, J., & Elvis, M. 1997, ApJ, 489, 615
Stark, A. A., Gammie, C. F., Wilson, R. W., et al. 1992, ApJS, 79, 77
Stern, D. 2015, ApJ, 807, 129
Stern, D., Lansbury, G. B., Assef, R. J., et al. 2014, ApJ, 794, 102
Stevens, J. A., Page, M. J., Ivison, R. J., et al. 2005, MNRAS, 360, 610
Tozzi, P., Gilli, R., Mainieri, V., et al. 2006, A&A, 451, 457
Treister, E., & Urry, C. M. 2006, ApJL, 652, L79
Tundo, E., Tozzi, P., & Chiaberge, M. 2012, MNRAS, 420, 187
Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, ApJ, 786, 104
Vagnetti, F., Middei, R., Antonucci, M., Paolillo, M., & Serafinelli, R. 2016, A&A, 593, A55
Vattakunnel, S., Tozzi, P., Matteucci, F., et al. 2012, MNRAS, 420, 2190
Virani, S. N., Treister, E., Urry, C. M., & Gawiser, E. 2006, AJ, 131, 2373
Vito, F., Gilli, R., Vignali, C., et al. 2014, MNRAS, 445, 3557
Vito, F., Gilli, R., Vignali, C., et al. 2016, MNRAS, 463, 348
Vito, F., Vignali, C., Gilli, R., et al. 2013, MNRAS, 428, 354
Wilms, J., Allen, A., & McCray, R. 2000, ApJ, 542, 914
Xue, Y. Q., Luo, B., Brandt, W. N., 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
Yaqoob, T. 1997, ApJ, 479, 184
Young, M., Brandt, W. N., Xue, Y. Q., et al. 2012, ApJ, 748, 124