The *Chandra* Deep Fields: Lifting the Veil on Distant Active Galactic Nuclei and X-Ray Emitting Galaxies

Y. Q. Xue\textsuperscript{a,b}

\textsuperscript{a}CAS Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, Hefei 230026, China
\textsuperscript{b}School of Astronomy and Space Science, University of Science and Technology of China, Hefei 230026, China

Abstract

The *Chandra* Deep Fields (CDFs), being a major thrust among extragalactic X-ray surveys and complemented effectively by multiwavelength observations, have critically contributed to our dramatically improved characterization of the 0.5–8 keV cosmic X-ray background sources, the vast majority of which are distant active galactic nuclei (AGNs) and starburst and normal galaxies. In this review, I highlight some recent key observational results, mostly from the CDFs, on the AGN demography, the interactions between AGNs and their host galaxies, the evolution of non-active galaxy X-ray emission, and the census of X-ray galaxy groups and clusters through cosmic time, after providing the necessary background information. I then conclude by summarizing some significant open questions and discussing future prospects for moving forward.

Keywords: black holes, galaxies, active galactic nuclei, quasars, accretion

1. Introduction

1.1. Effectiveness of extragalactic X-ray surveys

Since the discovery of the cosmic X-ray background (CXRB; e.g., Giacconi et al. 1962), various major X-ray observatories have joined the efforts of resolving it into discrete cosmic sources as well as characterizing such sources, by carrying out different tiers of extragalactic X-ray surveys that range from shallow all-sky surveys to ultradeep pencil-beam surveys. Together, these surveys, being highly complementary to each other, effectively occupy the practically-accessible half of the so-called X-ray survey discovery space (i.e., X-ray flux limits achieved vs. solid angles covered; see Fig. 3 and Table 1 of Brandt & Alexander (2015) for a recent demonstration and the information of additional X-ray surveys, respectively).

X-ray AGN surveys are arguably the most effective method of identifying highly reliable and fairly complete samples of AGNs, due to several reasons (see, e.g., Section 1.1 of Brandt & Alexander 2015 for detailed reasoning and caveats): (1) Observationally, X-ray emission is nearly a universal feature of optically, infrared (IR), radio-selected AGNs that are neither highly Compton-thick (CT, i.e., with neutral hydrogen column densities of $N_H \gtrsim 1\times10^{24}$ cm$^{-2}$; e.g., Lanzuisi et al. 2015b) nor intrinsically X-ray weak (such sources are very rare; see, e.g., Wu et al. 2011; Luo et al. 2014a).

Theoretically, X-ray emission can be produced in various accretion disk models for AGNs that are applicable for a wide range of mass accretion rates (from sub-Eddington to super-Eddington accretion), disk temperatures (cold vs. hot accretion flows), gas opacities (optically thick vs. thin), and geometric structures (thin vs. thick); these models invoke a corona or corona-like component to Compton up-scatter soft photons into hard X-rays when necessary (see, e.g., Yuan...
Figure 1: Flux limits achieved versus solid angles covered by some selected X-ray surveys in the 0.5–2 keV band from *Chandra*, *XMM-Newton*, *ROSAT*, and *eROSITA*. The vertical dotted line indicates the solid angle for the whole sky. The surveys plotted are listed below, with their corresponding references shown in the parentheses: (1) for the *Chandra* Deep Fields (CDFs; red stars): the 7 Ms CDF-S survey (Luo et al. 2017), the 2 Ms CDF-N survey (Xue et al. 2016), and the 250 ks E-CDF-S survey (Xue et al. 2016), with the numbers annotated after the survey names indicating the observed X-ray source densities in their respective central 3-arcmin areas (these numbers have not been corrected for detection incompleteness or Eddington bias; see Table 1; cf. Section 2.1); (2) for the other *Chandra* surveys (blue bullets): the *Chandra* Deep Survey of the Extended Groth Strip (AEGIS-X; Laird et al. 2009), the AEGIS-X Deep survey (Nandra et al. 2015), the SSA22 protocluster survey (Lehmer et al. 2009), the Lynx survey (Stern et al. 2002), the *Chandra* COSMOS Legacy survey (Civano et al. 2016), the ELAIS N1+N2 deep X-ray survey (Manners et al. 2003), the *Chandra* Lockman Area North Survey (CLANS; Trouille et al. 2008), and the X-ray survey of the NDWFS Bootes field (XBootes; Murray et al. 2005); (3) for the *XMM-Newton* surveys (green crosses): the CDF-S survey (Ranalli et al. 2013), the CDF-N survey (Miyaji et al. 2003), the ELAIS-S1 field survey (Puccetti et al. 2006), the Subaru/XMM-Newton Deep Survey (SXDS; Ueda et al. 2008), the XMMLarge Scale Structure survey (XLSS; Chiappetti et al. 2013), the Stripe 82X survey (LaMassa et al. 2016), the XMM Medium Deep Survey (XMDS; Chiappetti et al. 2005), the COSMOS survey (Cappelluti et al. 2009), and the XXL survey (Pierre et al. 2017); (4) for the *ROSAT* surveys (purple squares): the *ROSAT* Ultra Deep Survey (UDS; Lehmann et al. 2001); the *ROSAT* North Ecliptic Pole survey (NEP; Henry et al. 2006), and the Second *ROSAT* all-sky survey source catalog (2RXS; Boller et al. 2016); and (5) the proposed final (4 years) *eROSITA* All-Sky Survey (eRASS:8; the orange dashed line with two arrow heads; Merloni et al. 2012).
& Narayan 2014 for a review). (2) X-rays can penetrate through non-highly CT columns that are common among the majority AGN populations, and become even more penetrating at high redshifts due to positive \( K \)-correction, thereby reducing significantly absorption biases, probing immediate vicinities of SMBHs, and allowing for reliable \( N_H \) measurements to uncover intrinsic (i.e., absorption-corrected) AGN luminosities. (3) X-ray emission is subject to low dilution by host-galaxy stellar emission. An X-ray point source sitting right at the center of a galaxy is very likely to be an AGN; this serves as an effective way to identify distant AGNs when it is typically unfeasible to resolve spatially AGN light from host starlight. (4) An AGN X-ray spectrum is produced through numerous line and continuum emission processes subject to obscuration, and can therefore be utilized to infer physical conditions close to the SMBH, provided that the spectrum is of sufficient signal-to-noise ratio and energy resolution.

1.2. The Chandra Deep Fields

The Chandra Deep Fields (CDFs; see Fig. 2 and Table 1) consist of the Chandra Deep Field-South (CDF-S), the Chandra Deep Field-North (CDF-N), and the Extended-Chandra Deep Field-South (E-CDF-S). The CDF-S survey was originally led by R. Giacconi during 1999–2000 (1 Ms CDF-S; Giacconi et al. 2002), extended to 2 Ms through the Director’s Discretionary Time (DDT) by the CXC director H. Tananbaum in 2007 (2 Ms CDF-S; Luo et al. 2008), awarded an additional 2 Ms DDT by H. Tananbaum in 2010 (4 Ms CDF-S; Xue et al. 2011), and eventually pushed to 7 Ms by W. N. Brandt during 2014–2016 (7 Ms CDF-S; Luo et al. 2017).

The CDF-S patch of sky, lying in the Fornax constellation, was chosen because of very low foreground Galactic \( N_H = (8.8 \times 10^{19}) \text{ cm}^{-2} \); e.g., Stark et al. 1992), no bright \((mv \leq 14)\) Galactic stars, and optimal visibility from large ground-based telescopes in Chile. The CDF-N project was initiated by G. Garmire (the first \( \approx 0.5 \) Ms) and W. N. Brandt (the second \( \approx 0.5 \) Ms) during 1999–2001 (1 Ms CDF-N; Brandt et al. 2001), and subsequently enlarged by W. N. Brandt during 2001–2002 (2 Ms CDF-N; Alexander et al. 2003; Xue et al. 2016).

The CDF-N lies in the Ursa Major constellation, and was chosen for largely similar considerations as the CDF-S (e.g., \( N_H = 1.6 \times 10^{20} \text{ cm}^{-2} \); Stark et al. 1992), in addition to the desire to cover the HDF-N that was unique back then. Furthermore, as a parallel field to the CDF-S, the CDF-N doubles the number of such deep surveys, thus controlling for influence of cosmic variance and allowing for direct comparative studies between fields. The E-CDF-S survey was carried out (PI: W. N. Brandt) in 2004 (250 ks E-CDF-S; Lehmer et al. 2005; Xue et al. 2016), which significantly expands the sky coverage of the CDF-S proper with four distinct, contiguous, and flanking pointings (thus totaling \( \approx 1 \) Ms exposure).

All the CDF observations were performed with the Advanced CCD Imaging Spectrometer (ACIS; Garmire et al. 2003) onboard the Chandra X-ray observatory (Weisskopf et al. 2000), whose sharp point-spread function (PSF) and low background make higher sensitivities achievable with longer exposures in an efficient manner. Indeed, the 7 Ms CDF-S and 2 Ms CDF-N images are the two deepest X-ray images ever taken, thus being able to explore parameter space that has never been probed by any other X-ray surveys and fulfill one of Chandra’s central design goals — revealing and characterizing the sources that constitute the CXRB. Amazingly, the faintest sources detected in the 7 Ms CDF-S have only \( \approx 1 \) count per 10 days; and the 7 Ms CDF-S will serve as a multi-decade Chandra legacy for advancing deep-survey science projects, owing to its unique combination of great depth and high angular resolution.

The motivation of going ultradeep is at least two-fold: on one hand, ultradeep exposures accumulate sufficient photons for known (faint) sources that allow for better characterization of the AGNs producing most of cosmic accretion power via X-ray spectral and variability analyses; on the other hand, ultra-sensitivities reveal a substantial amount of new sources (typically fainter and more obscured), thereby facilitating better understanding of obscured growing SMBHs through the \( z \approx 1–4 \) era of massive galaxy assembly, majority AGN populations in the first galaxies, and X-ray binary (XRB) populations in cosmologically distant starburst and normal galaxies.

1.3. Importance of multiwavelength observations

The ultradeep multiwavelength coverage, well matched with the CDF X-ray data, is critically important for many aspects, including source identification, source classification, measurements of host-galaxy

\[ 2 \text{In addition to the 7 Ms Chandra } \approx 0.3–8 \text{ keV CDF-S coverage, there are } \approx 3 \text{ Ms of CDF-S coverage with XMM-Newton at } 0.2–12 \text{ keV} \text{ (Comastri et al. 2011; Ranalli et al. 2013; see Fig. 1) and 200 ks of E-CDF-S coverage with NuSTAR at } 3–24 \text{ keV (Mullaney et al. 2015a).} \]

\[ 3 \text{In addition to the 2 Ms Chandra CDF-N coverage, there are 180 ks of CDF-N coverage with XMM-Newton (Miyaji et al. 2003; see Fig. 1) and 200 ks of CDF-N coverage with NuSTAR (completed in February 2016).} \]
Figure 2: The CDF trio — the 7 Ms CDF-S, 2 Ms CDF-N, and 250 ks E-CDF-S (see Table 1 for more information). (Top left) The locations of the CDFs in the sky, with the separation between the CDF-N and CDF-S/E-CDF-S annotated. (Others) Spatial distributions of all point sources detected in the 2 Ms CDF-N (Xue et al. 2016), 250 ks E-CDF-S (Xue et al. 2016), and 7 Ms CDF-S (Luo et al. 2017), respectively, as well as the false-color image of the central 7 Ms CDF-S (16 arcmin across; courtesy of B. Luo). The filled circles, triangles, and five-pointed stars represent AGNs, starburst and normal galaxies, and Galactic stars, respectively. Sources detected neither in the 0.5–2 keV band (soft band; SB) nor in the 2–7 keV band (hard band; HB) are coded in black (i.e., detected only in the 0.5–7 keV band; full band; FB); sources detected in the SB but not in the HB are coded in red; sources detected in the HB but not in the SB are coded in blue; and sources detected both in the SB and HB are coded in between red and blue based on ratios between the SB and SB+HB (net counts) (see the color bar at the bottom-left corner). Bold crosses roughly indicate exposure-weighted average aim points.

Physical properties (e.g., redshift, stellar mass — $M_\star$, star formation rate — SFR, absolute rest-frame magnitudes, morphology), and providing useful comparison samples of non-active galaxies. The CDFs are among the most intensively observed sky patches with enormous observational investments at many wavelengths from both ground and space. For instance, the CDF-S area has been covered by $\approx 50$ different photometric bands that span from extreme ultraviolet (UV) to far-IR (FIR) at great depths, in conjunction with numerous deep spectroscopic and imaging observations, thereby resulting in a very-high redshift success rate (97.8%
of the 7 Ms CDF-S main-catalog sources have redshift measurements with 67.2% being spectroscopic ones.\(^4\)

Luo et al. 2017; see Table [1]. Table [2] lists some of the most notable multiwavelength surveys and photometric-redshift catalogs in the CDF areas that were completed; and Figure [3] presents an example of calculating accurate \(z_{\text{phot}}\) using rich multiwavelength photometric data. By combining the X-ray and multiwavelength coverage, the CDF datasets provide a unique opportunity to study both statistical and source-by-source properties of detected sources over a large range of \(L_X - z\) parameter space (see Table [1]). Furthermore, the CDF regions will be a testbed for future large observatories such as JWST and ELT, and the ultradeep multiwavelength coverage will continue to improve persistently, thereby keeping the science exciting.

---

| Table 1: Properties of the CDFs\(^a\) |
|-------------------------------------|
| Galactic \(N_H\) (cm\(^{-2}\)) | \(8.8 \times 10^{20}\) | \(1.6 \times 10^{21}\) | \(8.8 \times 10^{20}\) |
| Observational timespan | 1999/10 – 2016/03 (16.4 yrs) | 1999/11 – 2002/02 (2.3 yrs) | 2004/02 – 2004/11 (0.8 yrs) |
| Total number of observations | 102 | 20 | 9 |
| Effective exposure (ks) | 6727 | 1896 | 235/209/240/241\(^b\) |
| Solid angle covered (arcmin\(^2\)) | 484.2 | 447.5 | 1128.6 |
| Source detection criteria | WAVDETECT at \(10^{-5}\) and \(P < 0.007\) | WAVDETECT at \(10^{-5}\) and \(P < 0.004\) | WAVDETECT at \(10^{-5}\) and \(P < 0.002\) |
| Number of sources detected\(^d\) | 1008 | 683 | 1003 |
| FB (0.5–7 keV) detected counts | (11.2, 98.9, 56916.2)\(^f\) | (8.1, 66.2, 19748.4) | (3.3, 27.1, 4010.6) |
| SB (0.5–2 keV) detected counts | (6.1, 47.4, 38817.0) | (5.4, 35.0, 14227.3) | (2.2, 18.9, 2802.6) |
| HB (2–7 keV) detected counts | (9.2, 94.6, 18137.8) | (7.7, 57.5, 5540.6) | (3.4, 20.4, 1210.8) |
| 1σ X-ray positional uncertainty (") | (0.11, 0.47, 1.28) | (0.10, 0.47, 2.02) | (0.10, 0.63, 1.30) |
| Logarithm of FB flux (erg cm\(^{-2}\) s\(^{-1}\)) | \((-16.76, -15.50, -12.96)\) | \((-16.35, -15.09, -12.70)\) | \((-15.73, -14.79, -12.88)\) |
| Logarithm of SB flux (erg cm\(^{-2}\) s\(^{-1}\)) | \((-17.11, -16.19, -13.29)\) | \((-16.83, -15.79, -13.07)\) | \((-16.13, -15.27, -13.26)\) |
| Logarithm of HB flux (erg cm\(^{-2}\) s\(^{-1}\)) | \((-16.46, -15.25, -13.13)\) | \((-16.15, -14.95, -12.95)\) | \((-15.73, -14.70, -13.02)\) |
| Faintest sources detected | 1 count per \(\approx 10\) days | 1 count per \(\approx 4\) days | 1 count per \(\approx 1\) day |
| Logarithm of \(L_X\) (erg s\(^{-1}\)) | \((39.01, 42.48, 45.05)\) | \((39.28, 42.94, 45.07)\) | \((39.89, 43.34, 45.50)\) |
| % of multiwavelength identifications | 98.4% | 98.1% | 95.5% |
| % of \(z_{\text{spec}}\) (adopted)\(^f\) | 67.2% (97.8%) | 51.4% (93.4%) | 47.5% (80.8%) |
| \(z_{\text{adopted}}\) \(^g\) | (0.000, 1.156, 5.776) | (0.000, 1.130, 5.365) | (0.000, 1.193, 7.203) |
| % of AGNs/galaxies/stars | 70.5%/28.3%/1.2% | 86.5%/11.0%/2.5% | 90.6%/6.7%/2.7% |
| AGN/galaxy/star density (deg\(^{-2}\)) \(^i\) | 13600/12100/250 | 12400/4200/100 | 5200/500/100 |

\(^a\) For source properties, here I refer only to the sources from the main catalogs of the 7 Ms CDF-S (Luo et al. 2017), 2 Ms CDF-N (Xue et al. 2016), and 250 ks E-CDF-S (Xue et al. 2016). These three main catalogs were produced using essentially the same approach (critically aided by the use of the ACIS Extract package; AE; Broos et al. 2010), which incorporates a number of recent improvements in Chandra source-cataloging methodology, and therefore maximizes the number of reliable sources detected and allows for best possible X-ray characterization of source properties (see, e.g., Table 1 of Xue et al. 2016 for details).

\(^b\) The E-CDF-S consists of four distinct, contiguous pointings that flank the CDF-S proper (see Fig. [2]).

\(^c\) \(P\) indicates the probability of a source not being real (i.e., due to background fluctuations).

\(^d\) Among the 1008+683+1003=2694 CDF sources, 298 were detected both in the CDF-S and E-CDF-S, which results in a total of 2396 unique sources.

\(^e\) The three numbers in parentheses denote the minimum, median, and maximum values.

\(^f\) This is the absorption-corrected rest-frame 0.5–7 keV luminosity (\(L_X\) hereafter).

\(^g\) \(z_{\text{adopted}}\) denotes the adopted redshifts, with secure spectroscopic redshifts (\(z_{\text{spec}}\)) preferred over photometric redshifts (\(z_{\text{phot}}\)).

\(^h\) The maximum secure \(z_{\text{spec}}\)’s are 4.762 and 5.186 for the CDF-S/E-CDF-S and CDF-N, respectively. The \(z_{\text{phot}}\) estimates above these values are subject to large uncertainties.

\(^i\) These are observed source densities calculated within the respective central \(r \leq 3\) arcmin areas.

---

\(4\) For \(z_{\text{spec}}\) references, see Section 4.3 of Luo et al. 2017 and Section 3.3.4 of Xue et al. 2016 for the CDF-S/E-CDF-S, and Section 2.3.4 of Xue et al. 2016 for the CDF-N.
1.4. Identification of X-ray AGNs in CDFs

Over past few decades, a series of observational and mostly empirical methods have been developed and refined to identify AGN candidates from the general X-ray source population detected in extragalactic X-ray deep surveys, including AGNs, starburst and normal galaxies, galaxy groups and clusters, and Galactic stars. Table 3 outlines the AGN identification criteria that have been routinely used in the CDF catalog papers and some relevant works (see, e.g., Bauer et al. 2004, Alexander et al. 2005a, Lehmer et al. 2008 and 2012, Xue et al. 2010, 2011, and 2016, and Luo et al. 2017 for details and caveats), which rely upon either direct use of the X-ray data or combined use of both the X-ray and multiwavelength data. These methods are effective in revealing distinct AGN signatures owing to three major factors: accurate X-ray source positions (thanks to Chandra’s sharp PSF and thus superb sub-arcsecond angular resolution; see Table 1) that allow for reliable multiwavelength identifications with a sophisticated likelihood-ratio matching procedure (e.g., Brusa et al. 2005, 2007; Luo et al. 2010, 2017; Xue et al. 2016), robust X-ray photometric measurements that well describe source X-ray properties (e.g., Xue et al. 2011, 2016; Luo et al. 2017), and rich multiwavelength data that critically complement the X-ray observations (see Section 1.3 and Table 2). It is worth noting that the vast majority of the CDF AGNs satisfy multiple criteria listed in Table 3 such cross-validations of AGN candidates therefore result in highly complete, reliable, and pure samples of distant AGNs for statistically meaningful investigations. In addition to AGN candidates, Galactic stars are also identified based on spectroscopic and/or imaging observations (e.g., Feigelson et al. 2004; Xue et al. 2011, 2016), and the remaining non-AGN and non-star CDF sources are regarded as galaxies, whose measured X-ray emission can be produced by a mixture of XRBs and low-rate SMBH accretion.
Figure 3: Reliable $z_{phot}$ calculation for the source with XID=355 in the 7 Ms CDF-S main catalog, which is enabled by using the rich and high-quality broadband photometric data ($\approx 40$ filters; see Table 2 for the detail of photometric data used) and the spectral energy distribution (SED) template fitting approach.

Table 3: Frequently-Used Criteria for X-ray AGN Identification in CDFs

| AGN criterion | Basic reasoning (targeted AGN subpopulation) |
|---------------|---------------------------------------------|
| $L_{X,intrinsic} \geq 3 \times 10^{42}$ erg s$^{-1}$ | Typical starburst and normal galaxies: $L_{X,intrinsic} \approx 10^{42}$ erg s$^{-1}$ (X-ray luminous AGNs) |
| $\Gamma_{effective} \leq 1$ | Obscured ($N_H \geq 10^{22}$ cm$^{-2}$) AGNs: hard X-ray spectra observed (X-ray obscured AGNs) |
| $\log(f_X/f_k) > -1$ | Elevated X-ray emission compared to host optical emission (majority AGN populations) |
| $\log(f_X/f_k) > -1.2$ | Elevated X-ray emission compared to host near-IR (NIR) emission (majority AGN populations) |
| $\log(f_X/f_\text{Ks,radio}) > -1.2$ | Elevated X-ray emission compared to host NIR emission (majority AGN populations) |
| $L_{X,intrinsic}/L_{1.4\text{GHz}} \geq 2.4 \times 10^{18}$ | Excess X-ray emission over that expected from pure star formation (AGNs with radio data) |
| Spectroscopic AGN features | Broad emission/absorption lines; high-excitation emission lines (AGNs with quality spectra) |
| Significant X-ray variability | Large-amplitude X-ray variability commonly seen in AGNs (AGNs with $L_{X,intrinsic} \geq 10^{44}$ erg s$^{-1}$) |

1.5. Scope of this review

This paper is meant to be a relatively focused and compact one, which highlights briefly some of the recent key observational results (my apology in advance for unavoidably embedding personal flavor as well as not possibly managing to cover all relevant works), mostly from the CDFs, mainly on distant AGNs, X-ray emitting galaxies, as well as galaxy groups and clusters, thus demonstrating the beauty and power of the CDF treasure trove in these relevant fields of study and hopefully arousing interest of the even broader community, in light of the recent release of the most up-to-date CDF catalogs and products (Xue et al. 2016; Luo et al. 2017) yet to be fully exploited scientifically. For additional relevant in-depth reviews of results mainly from extragalactic X-ray surveys in general, I refer readers to, e.g., Comastri (2004), Gilli (2004, 2013), Brandt and Hasinger (2005), Shankar (2009), Brandt and Alexander (2010, 2015), Alexander and Hickox (2012), Fabian (2012), Treister & Urry (2012), Korndey & Ho (2013), Heckman & Best (2014), Vignali (2014), Reines & Comastri (2016), Padovani et al. (2017), and the references therein.

The remainder of this paper is structured as follows: Section 2 describes the AGN demography results (including AGN number counts and CXRB, high-redshift AGNs and AGN X-ray luminosity function — XLF, highly obscured and CT AGNs, low-mass black
holes/AGNs, and significantly variable AGNs); Section 3 highlights results on the interactions between AGNs and their host galaxies (including AGN X-ray luminosity versus galaxy SFR, conducive host galaxy properties for AGN activity, Eddington ratio distribution and correlation between intrinsic X-ray photon index and Eddington ratio, and coeval growth of SMBHs and their hosts); Section 4 introduces results on the evolution of starburst and normal galaxy X-ray emission; Section 5 presents the census of X-ray galaxy groups and clusters; Section 6 gleams some additional results not formally fit into the above sections; and finally, Section 7 summarizes this paper and discusses future prospects. Apparently, many of the above subtopics are intertwined by nature and should be understood collectively.

2. AGN demography

2.1. Measuring AGN number counts and resolving the CXRB

The cumulative X-ray number counts (i.e., logN−logS), quantifying the increase of the cumulative number of X-ray sources per unit area (N) as a function of decreasing flux (S), has been routinely used to characterize the extragalactic X-ray source populations (e.g., Brandt et al. 2001; Rosati et al. 2002; Bauer et al. 2004; Kim et al. 2007; Georgakakis et al. 2008; Luo et al. 2008; Lehmer et al. 2012; Ehler et al. 2013; Ranalli et al. 2013). In particular, Lehmer et al. (2012) have pushed such number-counts studies to a new level (also see, e.g., Georgakakis et al. 2008) by presenting a state-of-the-art procedure for deriving number counts in the 4 Ms CDF-S, which performs simulations to obtain source recovery functions that account for detection incompleteness, and implements a new Bayesian approach to obtain flux-probability distributions that account for Eddington bias, thereby enabling reliable computation of number counts down to flux limits that are typically a factor of ≈1.5 lower than nominal survey sensitivities.

Using the 7 Ms CDF-S main catalog and following the Lehmer et al. (2012) procedure, Luo et al. (2017; see their Fig. 31) and this paper (see Figs. 4, 5, and 6 courtesy of B. D. Lehmer) present together the unprecedentedly sensitive X-ray number-counts measurements, down to 4.2 × 10^{18} erg cm^{-2} s^{-1} in the 0.5–2 keV band (soft band; SB) and 2.0 × 10^{17} erg cm^{-2} s^{-1} in the 2–7 keV band (hard band; HB; cf. the sensitivity limits shown in Table 1), for the overall source population and AGNs, respectively, confirming and/or extending significantly many previous number-counts results (e.g., Bauer et al. 2004; Lehmer et al. 2012): (1) AGNs dominate the overall number counts, and the number counts of AGNs, starburst and normal galaxies, and Galactic stars in various Chandra passbands can be adequately described by either double (for AGNs) or single power-law (for galaxies and stars) functions. (2) The galaxy power-law slope is steeper than the AGN faint-end power-law slope, indicating the rapid rise of galaxies toward faint fluxes; and, as predicted, at the faintest fluxes (f_{SB} ≤ 6.0 × 10^{-18} erg cm^{-2} s^{-1}) that are uniquely accessible to the 7 Ms CDF-S, the galaxy number counts are observed for the first time to overtake the AGN number counts. (3) At the SB flux limit, the AGN and galaxy densities reach record highs of ≈23900 and ≈26600 deg^{-2} (the former represents the highest AGN sky density measured reliably at any wavelength; cf. the observed source densities in Table 1), respectively, which correspond to ≈1.0 billion AGNs and ≈1.1 billion X-ray galaxies in the entire sky. (4) The measurements of AGN number counts, apportioned by various source properties such as z, N_H, and L_X, are enabled by virtue of the exquisite CDF-S X-ray and multiwavelength data (see Table 2), which reveal that, in both SB and HB, a) AGNs with z < 1.5 dominate the number counts while the remainder gradually catch up toward faint fluxes (see Fig. 4), and b) AGNs with N_H < 10^{22} cm^{-2} or L_X > 10^{43} erg s^{-1} dominate at the bright end while the others dominate or become increasingly important toward the faint end (see Figs. 5 and 6). These results are broadly consistent with the 4 Ms CDF-S number-counts results (Lehmer et al. 2012) and the expectations of the Gilli et al. (2007) phenomenological AGN population-synthesis models. One point worth noting is that the fraction of sources with L_X < 10^{42} erg s^{-1} that are classified as AGNs versus galaxies (see Table 3 for the AGN identification criteria) is higher in the 7 Ms CDF-S compared to the 4 Ms CDF-S, and this is reflected in the faint-end number counts with an upturn and also causes the galaxy number counts to not increase as fast as expected for the faintest new sources (also see, e.g., Smolčić et al. 2017a, 2017b for similar faint-end behaviors of radio AGNs).

With the number-counts estimates in hand, it is relatively straightforward to resolve the CXRB into individual X-ray sources by integrating their fluxes. Luo et al. (2017) and Lehmer et al. (2012) find that AGNs dominate the 0.5–7 keV CXRB, and the resolved CXRB fractions are ≈57 ± 4% and 82 ± 13% with the 4 Ms CDF-S in SB and the 2–8 keV band (both being ≈1–2% higher than using the 1–2 Ms CDFs), rising to 81 ± 4%
Figure 4: Top panels: cumulative AGN number counts (filled black circles) for the 7 Ms CDF-S in the 0.5–2 keV (left) and 2–7 keV (right) bands, apportioned by AGN subsamples with $z < 1.5$ (open red circles), $1.5 \leq z \leq 3.0$ (open green triangles), $z > 3.0$ (open blue squares), and unknown redshifts (filled gray circles). Also shown are the Gilli et al. (2007) AGN population-synthesis model predictions (in corresponding colors) either with (dotted curves) or without (solid curves) an exponential decline in the XLF at $z > 2.7$. Middle panels: fractional contribution of each AGN subsample to the overall AGN number counts. Bottom panels: the ratio between the AGN number-count data and the Gilli et al. (2007) model without a declining XLF at $z > 2.7$ (courtesy of B. D. Lehmer; cf. Fig. 9 of Lehmer et al. 2012 for the 4 Ms CDF-S case).

and $93 \pm 13\%$ with the 7 Ms CDF-S in SB and HB, when adopting total CXRB intensities of $(8.15 \pm 0.58) \times 10^{-12}$ and $(1.49 \pm 0.20) \times 10^{-11}$ erg cm$^{-2}$ s$^{-1}$ deg$^{-2}$ in SB and HB, respectively (Hickox & Markevitch 2006; Kim et al. 2007). Although it is still debated whether the resolved CXRB fraction is increasing, constant, or even decreasing toward higher bandpasses (e.g., Moretti et al. 2003; Bauer et al. 2004; Worsley et al. 2005; Georgakakis et al. 2008; Lehmer et al. 2012), the above resolved CXRB fractions are in broad agreement with many works (e.g., Bauer et al. 2004; Hickox & Markevitch 2006; Luo et al. 2011; Xue et al. 2012), given the relatively large uncertainties on resolved fractions that primarily arise from the use of different methodologies and datasets (e.g., various AGN selection techniques, data depths, and redshift completeness levels) as well as CXRB intensity measurements. Despite of many efforts, the exact normalization of the CXRB spectrum is still uncertain by $\approx 20–30\%$, due to a combination of several complicated factors that include spectral cross-calibrations, instrumental background modeling, foreground contamination by Galactic collisional thermal plasmas, stray light contamination, bright-end corrections, and cosmic variance (e.g., Bauer et al. 2004; De Luca & Molendi 2004; Hickox & Markevitch 2006; Kim et al. 2007; Luo et al. 2008; Moretti et al. 2009; Ishida et al. 2011; Tsujimoto et al. 2011; Lehmer et al. 2012; Cappelluti et al. 2017a; Madsen et al. 2017).

There have been some emerging works that study the CXRB from somewhat different viewpoints, either by pushing the studies of resolving the CXRB to higher energy bands due to the advent of NuSTAR (e.g., Ballan-
Figure 5: Same as Fig. 4 (i.e., cumulative AGN number counts for the 7 Ms CDF-S), but apportioned by AGN subsamples with different $N_H$ values (courtesy of B. D. Lehmer; cf. Fig. 11 of Lehmer et al. 2012 for the 4 Ms CDF-S case).

These studies, in conjunction with other studies, help constrain the likely existence, properties, and evolution of a missing yet potentially large population of highly obscured (i.e., $N_H \gtrsim 3 \times 10^{23}$ cm$^{-2}$) or even CT AGNs (CTAGNs hereafter; see Section 2.3), facilitating the refinement of AGN population-synthesis models (e.g., Gilli et al. 2007; Treister et al. 2009a), and revealing the various components constituting the unresolved CXRB that include likely contributions from low-luminosity AGNs (LLAGNs), highly obscured AGNs, CTAGNs, non-active galaxies, and hot gas in groups and clusters (see Section 5).

2.2. Constraining the high-redshift AGN subpopulation and AGN XLF

AGNs within the first cosmic structures at $z \approx 4–8$ and beyond are of perennial interest because their observations have serious implications for, e.g., the primordial BH seeds in the early Universe, the overall SMBH growth history (including AGN mass function and luminosity function), the origin and evolution of correlations between SMBH mass ($M_{BH}$) and galaxy properties, and the role played by AGNs during intergalactic medium (IGM) heating and reionization (see, e.g., Reines & Comastri 2016; Brandt & Vito 2017 for a review). Given that X-rays are much more penetrating and much less affected by galaxy dilution (see Section 1.1) than other wavelengths, X-ray surveys are therefore critical for assessing such high-redshift SMBH growth.

Indeed, many recent XLF works have been carried out utilizing the CDF data (note that the 7 Ms CDF-S is sufficiently sensitive to reach about 20–3 times below

---

7NuSTAR resolves $\approx 33$–39% of the CXRB in the 8–24 keV band (depending on the adopted 20–50 keV CXRB intensity, and with an additional $\approx 5$% statistical uncertainty), directly identifying CT AGNs with $N_H \lesssim 10^{25}$ cm$^{-2}$ (Harrison et al. 2016). Although NuSTAR can probe much harder (typically a factor of $\approx 3$–4) X-rays than Chandra and XMM-Newton, it is still not properly reaching the CXRB peak.
the knee of AGN XLF at \( z = 0.5–4 \) and other X-ray observations, finding that there is an exponential decline in the number density of luminous (\( L_X \gtrsim 10^{44} \text{ erg s}^{-1} \)) AGNs with \( z \gtrsim 3 \) (e.g., Barger et al. 2003a; Cristiani et al. 2004; Silverman et al. 2008a; Brusa et al. 2009a; Civano et al. 2011; Fiore et al. 2012a; Vito et al. 2013, 2014a, 2017; Kalfountzou et al. 2014; Ueda et al. 2014; Georgakakis et al. 2015; Marchesi et al. 2016). However, the situation is not as clear for lower-luminosity AGNs (\( L_X \approx 10^{43–44} \text{ erg s}^{-1} \)), largely owing to the faintness, likely heavy obscuration, survey incompleteness, limited sample size, and challenging multiwavelength follow-up of such sources. There generally also appears to be a decline in the space density of \( z > 3 \) moderate-luminosity AGNs, albeit with large uncertainties (in particular, the constraints on \( z > 4 \) AGNs are even looser due to poorer statistics; e.g., Fiore et al. 2012a; Vito et al. 2013, 2014a, 2017; Kalfountzou et al. 2014; Ueda et al. 2014; Georgakakis et al. 2015; Marchesi et al. 2016). The space density of such AGNs remains uncertain by more than an order of magnitude (e.g., Gilli et al. 2011a), effectively having no much discriminating power between the predictions of the theoretical models that satisfy other existing observational constraints (e.g., Salvaterra et al. 2007; Marulli et al. 2008; Volonteri & Stark 2011; Habouzit et al. 2016; Volonteri & Reines 2016). However, even a small number of such objects, if detected, can remedy this situation and provide critical leverage in modeling early SMBH growth. In the long run, the robust detection of a decent number (i.e., a few tens or even more) of AGNs down to \( L_X \approx 10^{43} \text{ erg s}^{-1} \) beyond \( z \approx 3–4 \) would provide invaluable insight on the formation mechanisms responsible for the first quasars and the early co-evolution history, which has to call for future large X-ray observatories that are capable of efficiently carrying out ultra-sensitive surveys over wide sky areas. Nevertheless, stacking analyses, instead of direct detections, have already proved to be a beneficial attempt along this direction.

Recently, Vito et al. (2016) utilize both the 7 Ms CDF-S and ultradeep CANDELS/GOODS-S data to sensitively measure the total X-ray emission from a sample of 2076 \( 3.5 < z < 6.5 \) optically selected and individually X-ray-undetected galaxies (see Table 1 for

![Figure 6: Same as Fig. 4 (i.e., cumulative AGN number counts for the 7 Ms CDF-S), but apportioned by AGN subsamples with different \( L_X \) values (courtesy of B. D. Lehmer; cf. Fig. 12 of Lehmer et al. 2012 for the 4 Ms CDF-S case).](image)
the redshift-range information of individually X-ray-detected sources, with a sophisticated stacking procedure that is validated by simulations. They detect high-significance (> 3.7σ) stacked X-ray emission from massive galaxies at 3.5 < z < 4.5, 2.7σ emission from those at 4.5 < z < 5.5 (99.7% confidence level; highest significance ever in such a redshift range), and no significant signal from those at 5.5 < z < 6.5 (see the top panel of Fig. 7; also see, e.g., Willott 2011; Cowie, Barger & Hasinger 2012; Fiore et al. 2012b; Basu-Zych et al. 2013). They find that the detected X-ray emission is likely dominated by high-redshift XRB populations (also see Cowie et al. 2012), by comparing it with the expected high-redshift XRB emission that is extrapolated from lower-redshift results assuming a range of XRB model prescriptions (e.g., Ranalli et al. 2003; Fragos et al. 2013a; Lehmer et al. 2016; see Section 4), and also by comparing both the total SFRs and SFR densities of the stacked galaxies with previous UV-based and SED fitting results of galaxies. Therefore, they conclude that (1) the continuous low-rate SMBH growth in individually X-ray-detected galaxies makes negligible contribution to cosmic SMBH mass assembly at 3.5 < z < 6.5, when compared to that in AGNs detected by deep X-ray surveys; and (2) the observational constraints on the faint-end ($L_X \approx 10^{42}$ erg s$^{-1}$) AGN XLF at 3.5 < z < 6.5 are achieved for the first time, which indicate a fairly flat faint-end XLF slope (see the bottom panel of Fig. 7), confirming previous results (e.g., Barger et al. 2003a; Ueda et al. 2014) and extending them down to lower X-ray luminosities.

Collectively, the AGN XLF studies over a broad stretch of redshift, including those at high redshifts that are based on direct detections and aided by the stacking-based AGN XLF constraints such as the aforementioned Vito et al. (2016) work, reveal that (1) there is an anti-hierarchical "cosmic downsizing" behavior of AGNs, i.e., the number density of powerful quasars peaks at an earlier cosmic time than that of lower-luminosity AGNs (e.g., Cowie et al. 2003; Barger et al. 2005; Hasinger, Miyaji, & Schmidt 2005; Silverman et al. 2008a; Yencho et al. 2009; Aird et al. 2010; Ueda et al. 2014; Miyaji et al. 2015); (2) the peak AGN emissivity (i.e., the AGN comoving bolometric luminosity density that is in units of erg s$^{-1}$ Mpc$^{-3}$) occurs at $z \approx 1.8$ for AGNs with a broad range of bolometric luminosity $L_{bol} = 10^{43-48}$ erg s$^{-1}$ (see, e.g., Fig. 20 of Ueda et al. 2014); (3) AGNs are unlikely to dominate reionization at $z \geq 6$, with stars playing a leading role instead (e.g., Robertson et al. 2010, 2013; Haardt & Madau 2012; Grissom, Ballantyne & Wise 2014; Cappelluti et al. 2016; Vito et al. 2016; Ricci et al. 2017; also see Section 4) but see Giallongo et al. 2012, 2015); and (4) the luminosity-density evolution (LDDE) and luminosity and density evolution (LADDE) XLF models describe the observational data well, and the LDDE models appear to be further favored by some high-redshift constraints, yet with further testing needed due to insufficient statistics at high redshifts (e.g., Miyaji et al. 2000, 2001, 2015; Aird et al. 2010; Ueda et al. 2014; Vito et al. 2014a; Georgakakis et al. 2015; Fotopoulou et al. 2016). The basic nature of the above AGN XLF related results (e.g., cosmic downsizing) appears robust, despite of many relevant details yet to be further worked out.

2.3. Unveiling the highly obscured and CT AGN sub-population

Hunting for the highly obscured AGNs, CTAGNs in particular, has been a longstanding and challenging quest for the AGN demography work. CTAGNs are of special interest given a number of arguments (see, e.g., Comastri 2004; Georgantopoulos 2013; Gilli 2013; Vignali 2014 for a review): (1) There has been growing ob-
observational evidence that a substantial fraction of AGNs are obscured by CT gas both locally (e.g., Maiolino et al. 1998; Risaliti, Maiolino, & Salvati 1999; Matt et al. 2000) and in the distant universe (e.g., Dwelly & Page 2006; Tozzi et al. 2006; Georgantopoulos et al. 2009; Merloni et al. 2014; Buchner et al. 2014, 2015; Lanzuisi et al. 2015a; Liu et al. 2017; also see Fig. 8).

2. The existence of the CTAGN subpopulation is required by the AGN population-synthesis models for the CXRB in order to account for the intensity peak of the CXRB spectrum at \( \approx 20-30 \) keV, and CTAGNs are predicted to be as abundant as moderately obscured AGNs (e.g., Gilli et al. 2007, 2013).

3. The absorbed energy of CTAGNs at short wavelengths (including optical, UV, and X-rays) is thermally reprocessed by, e.g., the torus, and eventually reemitted in FIR, making them a potential contributor to the cosmic IR background (e.g., Shi et al. 2013a, 2013b).

4. Finally, distant CTAGNs are believed to represent a crucial phase in SMBH/galaxy co-evolution models, during which large amounts of gas is funneled to the center and induces both intensive obscured accretion and powerful star formation as a result of the merger process, and then feedback processes likely take over by self-regulating SMBH growth and quenching star formation (e.g., Page et al. 2004; Granato et al. 2006; Hopkins et al. 2006a; Menci et al. 2008; Alexander & Hickox 2012). Despite their importance, the majority of distant CTAGNs escape even from the deep X-ray surveys and still remain largely elusive due to their very nature of extreme obscuration (see Fig. 8 for a demonstration); therefore, there is no good knowledge even about their space density and cosmological evolution (e.g., Aird et al. 2015a; Buchner et al. 2015). Fortunately, the situation is gradually improving as many CTAGN identification methods have been developed and many ultradeep X-ray (e.g., the CDFs) and multiwavelength surveys have been performed along the quest of CTAGNs.

The signatures of CT emission can be revealed in many ways (see Table 3 for a brief summary; some of these techniques are closely related but with different emphases), e.g.: (1) X-ray spectroscopy is the only unambiguous way to identify bona-fide CTAGNs where deep X-ray surveys play a key role. With high-quality X-ray spectra, CT \( N_H \) values can be determined reliably through careful spectral modeling, immediately unveiling the CT nature of sources (e.g., Comastri 2004; Comastri et al. 2011). (2) The presence of a strong Fe K\(_\alpha\) line complex (typically with a \( \gtrsim 1 \) keV line equivalent width) around rest-frame 6.4–7 keV in an AGN X-ray spectrum is highly indicative of the source likely being CT (see Fig. 8), e.g., Comastri 2004; Comastri et al. 2011; Georgantopoulos et al. 2013).

3. An AGN with a characteristic reflection-dominated X-ray spectrum is almost guaranteed to be a CTAGN; such a spectrum is featured by a broad Compton hump peaking at \( \approx 20-30 \) keV, rapid declines toward both low energies (due to absorption) and high energies (due to Compton down-scattering), and a power-law shape in the \( \approx 2-10 \) keV band with a strong Fe line complex atop (typically having equivalent width \( > 1 \) keV; e.g., Comastri 2004; Comastri et al. 2011). (4) In case of an X-ray spectrum having limited counts or not resorting to detailed spectral fitting, a (very) large X-ray hardness ratio or a (very) small \( \Gamma_{\text{effective}} \) (see Table 3) can be crudely used to indicate the likely CT nature of the source (e.g., Alexander et al. 2011; Gilli et al. 2011); furthermore, using X-ray colors or a combination of X-ray hardness and X-ray-to-MIR flux ratio can select CTAGNs more efficiently than using a single hardness ratio (e.g., Iwasawa et al. 2012; Severgnini et al. 2012). (5) An AGN having excess IR emission (owing to the reemission of the absorbed energy) compared to the typical IR/optical emission level of galaxies/non-CT AGNs are likely CTAGN candidates; the key to this category of “IR-excess” methods is to reduce contamination from starburst and normal galaxies (see, e.g., Daddi et al. 2007, Alexander et al. 2011, and Luo et al.

Figure 8: Model spectra of AGNs with different obscuration levels. The black, blue, green, and red curves correspond to unobscured \((N_H = 0)\), moderately obscured \((N_H = 5 \times 10^{22} \text{ cm}^{-2})\), highly obscured \((N_H = 5 \times 10^{23} \text{ cm}^{-2})\), and CT \((N_H = 5 \times 10^{24} \text{ cm}^{-2})\) AGNs, respectively. The spectra are generated using the MYTorus model with \( \Gamma = 1.9\), inclination angle=90 deg, and the same normalization (Murphy & Yaqoob 2009; courtesy of G. Yang).
Table 4: Techniques to Identify CTAGN Candidates

| (1) | CT N$_{H_1}$ values derived from X-ray spectral fitting |
| (2) | Strong Fe K$\alpha$ line complex |
| (3) | Characteristic X-ray reflection spectrum/component |
| (4) | Large X-ray hardness ratio + X-ray colors |
| (5) | IR-excess emission |
| (6) | X-ray stacking |
| (7) | Some well-calibrated luminosity ratios |
| (8) | Deep 9.7 $\mu$m Si feature with $>$ 1 optical depth |
| (9) | Spectral curvature above 10 keV |
| (10) | Broadband SED decomposition |

See main text for explanations and references therein for caveats.

There have been many works that involve the use of one or more of the above methods and prove their utility. For instance, Gilli et al. (2011) discover a $z_{\text{spec}}=4.762$ CTAGN in the 4 Ms CDF-S that is confirmed by X-ray spectral analysis, thus being the most distant bona-fide CTAGN so far. This source appears to be caught during a major coeval episode of SMBH accretion ($L_{2-10 \text{ keV}} \approx 2.5 \times 10^{44}$ erg s$^{-1}$) and stellar mass assembly (being a submillimeter galaxy (SMG) with SFR$\approx 1000 M_{\odot}$ yr$^{-1}$ that is later confirmed by ALMA observations; Coppin et al. 2009; Gilli et al. 2014) at early times (see Fig. 9). More interestingly, several lines of evidence point to the existence of outflowing winds from the central SMBH, hinting at AGN feedback at work (Gilli et al. 2014). Finding more such AGNs and constraining their number density would be crucial to reconstruct the early co-evolution history. Alexander et al. (2011) identify 11 highly obscured AGNs among a sample of $z \approx 2$ K$<22$ mag $Bz_{\text{K}}$-selected galaxies through X-ray spectral analysis in the 4 Ms CDF-S. They find that, among these 11 sources, some display excess IR emission and some prefer a pure reflection model; furthermore, the stacked X-ray spectrum of all the sources has a reflection-dominated shape and a strong Fe line feature, indicating CT absorption in some sources (see the left panel of Fig. 10); and the estimated lower-limit space-density constraints of highly-obscured AGNs/CTAGNs (i.e., $f>10^{-5}$ Mpc$^{-3}$ for CTAGNs with $L_{2-10 \text{ keV}} \gtrsim 10^{43}$ erg s$^{-1}$ and $z \approx 2$, where $f \approx 10-50\%$ is the likely CTAGN fraction within the studied sample) are already comparable to the results from some other works (e.g., Tozzi et al. 2006; Alexander et al. 2008; Fiore et al. 2009; Brightman & Ueda 2012; Del Moro et al. 2016) and the predictions from AGN population-synthesis models, albeit with large uncertainties on both observations (see the right panel of Fig. 10) and models (e.g., Gilli et al. 2007; Treister et al. 2009a). It appears that the space density of $z \approx 2$ CTAGNs is comparable to that of $z \approx 2$ unobscured AGNs, suggesting a non-negligible contribution of such CTAGNs to the overall SMBH growth at $z \approx 2$.

Most recently, Liu et al. (2017) present a systematic spectral analysis for a sample of 276 HB-selected brightest AGNs in the 7 Ms CDF-S. After correcting for sample selection biases (e.g., incompleteness and Eddington bias), they find that: (1) The intrinsic $N_{\text{H}}$ distribution varies significantly across different redshift ranges (up to $z \approx 5$) as a result of strong dependence of $N_{\text{H}}$ on both $L_{\text{X}}$ and $z$; the overall $N_{\text{H}}$ distribution peaks at $N_{\text{H}} \approx 10^{23.5-24}$ cm$^{-2}$, higher than the peak value of $N_{\text{H}} \approx 10^{23.1}$ cm$^{-2}$ derived with 1 Ms
CDF-S (Tozzi et al. 2006; also see, e.g., Burlon et al. 2011; Brightman & Nandra 2011; Buchner et al. 2015), mainly due to more highly obscured sources being revealed by the 7 Ms CDF-S. (2) The strong evolution of obscured AGN fraction with $z$ can be formulated as $f_{\text{obs}cured} = (0.43 \pm 0.07)(1 + z)^{0.59 \pm 0.12}$; at $z > 2$, $f_{\text{obs}cured}$ shows only a weak evolution, likely being saturated (see also, e.g., Hasinger 2008; Vito et al. 2014; Ueda et al. 2014); and the $f_{\text{obs}cured}$ measures are higher than previous ones (e.g., Iwasawa et al. 2012; Vito et al. 2014). (3) After combining the 7 Ms CDF-S data with the wider and shallower C-COSMOS data (Lanzuisi et al. 2013), it is clear that, at any given (small) redshift bin within the range of $0.3 < z < 4.0$, the average $N_{\text{H}}$ (or $f_{\text{obs}cured}$) decreases with $L_X$; and at any given (small) $L_X$ bin within the range of $42 < \log(L_X) < 45$, the average $N_{\text{H}}$ (or $f_{\text{obs}cured}$) increases with $z$. (4) A total of 22 CTAGNs are identified, corresponding to $\approx 8\%$ of the studied sample; such a CTAGN fraction appears comparable to or slightly lower than CTAGN fractions of $\approx 7$–$20\%$ among hard X-ray selected samples (e.g., Malizia et al. 2009; Burlon et al. 2011; Vavasudevan et al. 2013; Akylas et al. 2016). The above dependence of $f_{\text{obs}cured}$ on $L_X$ result presented by Liu et al. (2017) confirms many previous works (Treister and Urry 2006; Hasinger 2008; Brightman & Nandra 2011; Burlon et al. 2011; Lusso et al. 2013; Brightman et al. 2014; Merloni et al. 2014; Ueda et al. 2014) and extends reliably to lower $L_X$. Surprisingly, wide-field IR surveys find that $f_{\text{obs}cured}$ might rise substantially again at $L_{bol} \approx 10^{47}$ erg s$^{-1}$ (e.g., Stern et al. 2014; Assef et al. 2015). Moreover, the dependence of $f_{\text{obs}cured}$ on $z$ result given by Liu et al. (2017) also confirms many previous works (e.g., La Franca et al. 2005; Ballantyne et al. 2006; Tozzi et al. 2006; Treister and Urry 2006; Hasinger 2008; Treister et al. 2009; Hiroi et al. 2012; Iwasawa et al. 2012; Vito et al. 2013, 2014a; Brightman et al. 2014; Merloni et al. 2014; Ueda et al. 2014; Buchner et al. 2015) and extends reliably to lower $L_X$.

An improved demography picture for the highly obscured and CT AGNs that better reveals their space density and cosmological evolution can be obtained by combining and reconciling all the bits and pieces of works already present in the literature, which are based on direct X-ray spectroscopy, reliable X-ray stacking analysis, and/or various multiwavelength diagnostics (e.g., Daddi et al. 2007; Alexander et al. 2008a, 2011; Fiore et al. 2008, 2009; Treister et al. 2009b; Gilli et al. 2011; Luo et al. 2011; Brightman & Ueda 2012; Vignali et al. 2014; Del Moro et al. 2016), as well as NuSTAR’s $> 10$ keV constraints (e.g., Alexander et al. 2013; Del Moro et al. 2014; Brightman et al. 2015; Koss et al. 2016). However, there is still a long way to a complete understanding of the CTAGN demography given the currently small sample of distant bona-fide CTAGNs; and the radical change of such a situation would have to rely on new powerful and preferably hard X-ray facilities.

Last but not least, accumulating evidence shows that the AGN obscuring material may not be necessarily related to the pc-scale dusty torus as commonly accepted; instead, it may be attributed to a compact starburst region (e.g., Gilli et al. 2014), the diffuse interstellar medium in the host galaxy (with $N_{\text{H}}$ likely up to $10^{22}$–$23.5$ cm$^{-2}$; e.g., Simcoe et al. 1997; Goulding et al. 2012; Buchner & Bauer 2017; Lanzuisi et al. 2017), or 100 pc-scale dust filaments (e.g., Prieto et al. 2014). Such obscuring material existing at relatively large scales likely has multiple functions, e.g., fueling AGN accretion, igniting nuclear star formation, and exerting AGN feedback (e.g., in the form of AGN-driven outflows) to influence the host galaxy, which adds an additional complexity to the understanding of

---

*Most recently, Maiolino et al. (2017) find observational evidence for star formation ($\text{SFR} = 15 M_{\odot}$ yr$^{-1}$) occurring within a prominent galactic outflow, which is $\approx 7$–$9$ kpc away from and perhaps driven by the southern obscured AGN in a $z = 0.0448$ merging system, likely playing some role in obscuring the AGN.*
the AGN obscuration phenomenon and makes the quest for highly obscured/CT AGNs even more challenging yet more intriguing.

2.4. Searching for the low-mass BH/AGN subpopulation

Low-mass BHs (or intermediate-mass BHs; IMBHs\(^9\) with \(\approx 10^{3-6} \, M_\odot\)) are thought to naturally bridge the evolutionary gap between stellar-mass BHs (with a few tens of solar masses) and SMBHs, whose studies have important implications for the origin of SMBHs (i.e., the birth and growth of early BH seeds), the overall SMBH growth history, the common characteristics of BH phenomena across all mass scales, and many other aspects (see, e.g., Miller & Colbert 2004; Volonteri 2010; Greene 2012; Reines & Comastri 2016 for a review). However, such a population of low-mass BHs has been largely elusive, because they are difficult to find, mainly due to their low masses and thus small gravitational influences and low luminosities.

Nevertheless, the situation of searching for low-mass BHs has been improving continuously, thanks to the advent of numerous high-quality multiwavelength observations and various search techniques. Within the Local Group, it is possible to utilize stellar or gas dynamics to search for low-mass BHs residing in dwarf galaxies (less than a few \(10^9 \, M_\odot\)), which represent a plausible place to host such BHs according to, e.g., the \(M_{\text{BH}}-M_\star\) scaling relation (see, e.g., Kormendy & Ho 2013 for a review). A number of such attempts made in some local dwarf galaxies either place upper limits of \(M_{\text{BH}} \approx 10^4 \, M_\odot\) (e.g., Gebhardt et al. 2001; Valluri et al. 2005; Lora et al. 2009; Jardel & Gebhardt 2012) or find some positive measurements of \(M_{\text{BH}} \approx 10^5-6 \, M_\odot\) (e.g., Seth et al. 2010; van den Bosch & de Zeeuw 2010; den Brok et al. 2015). Going beyond the Local Group, multiwavelength techniques have to be invoked in order to reveal the signatures of low-mass AGNs typically located in dwarf galaxies. For example, optical spectroscopic features such as high-ionization narrow emission lines, broad emission lines, and narrow emission-line diagnostics reveal many samples of low-mass BHs or candidates with \(M_{\text{BH}} \approx 10^5-6 \, M_\odot\) (e.g., Greene & Ho 2004; Dong et al. 2012; Reines et al. 2013; Moran et al. 2014; Sartori et al. 2015); and some studies attempt to use MIR color diagnostics to search for low-mass AGNs in small-bulge and low-mass galaxies (e.g.,

---

\(^9\)A most recent work, which discovers an IMBH with \(2200\pm1300 \, M_\odot\) in the center of the globular cluster 47 Tucanae, has become a real eye catcher (Kiziltan, Baumgardt, & Loeb 2017).
Satyapal et al. 2014; Hainline et al. 2016). It should be noted that the above optical techniques tend to be biased toward AGNs with relatively high Eddington ratios ($\lambda_{\text{Edd}} = L_\text{bol}/L_\text{Edd}$, where $L_\text{Edd}$ is Eddington luminosity) and MIR approaches are often subject to severe contamination because of the confusion between MIR colors of dwarf starburst galaxies and AGNs. In contrast, X-ray observations do not have such limitations and can probe AGNs down to very low levels with least bias (see Section 1.1).

Indeed, many X-ray studies have accumulated growing samples of low-mass AGNs among a wide variety of low-mass host galaxies mostly in the local universe (e.g., Ghosh et al. 2008; Jia et al. 2011; Reines et al. 2011, 2014; Kamizasa, Terashima, & Awaki 2012; Lemons et al. 2015; Miller et al. 2015; Baldassare et al. 2015, 2017; Plotkin et al. 2016). The search for low-mass AGNs at moderate redshifts calls for deep X-ray surveys with direct detections or stacking analyses. Schramm et al. (2013) discover three low-mass AGNs hosted in $M_* < 3 \times 10^9 M_\odot$ galaxies at $z < 0.3$, which are individually detected in the 4 Ms CDF-S and have $L_X \approx 10^{40}$ erg s$^{-1}$ (with one being variable). Particularly, one of these sources has its broad H$\alpha$ line detected, providing a virial mass estimate of $\approx 2 \times 10^5 M_\odot$ that is consistent with the estimate derived using scaling relations between $M_{\text{BH}}$ and host galaxy properties (see Fig. 11). Pardo et al. (2016) identify 10 low-mass X-ray AGNs among a sample of $M_* < 3 \times 10^9 M_\odot$ dwarf galaxies at $z < 1$, which are estimated to typically have $\lambda_{\text{Edd}} \approx 5\%$. Most recently, Chen et al. (2017) study a sample of 10 low-mass $z < 0.3$ AGNs detected in the 3–24 keV band by the NuSTAR serendipitous survey, demonstrating the great potential and advantage of using hard X-ray observations to recover low-mass (obscured) AGNs missed by optical spectroscopic surveys and $< 10$ keV X-ray surveys. Furthermore, the use of sensitive stacking can significantly push the constraints upon low-mass AGNs out to higher redshifts. By stacking the 4 Ms CDF-S data, Xue et al. (2012) track down a population of highly obscured AGNs that are not individually detected in X-rays, which make the majority contribution to the unresolved 6–8 keV CXRB, and are hosted in faint $1 \leq z \leq 3$ galaxies that are located on the top of the blue cloud and have $M_* \approx 2 \times 10^8 M_\odot$. Mezcua et al. (2016) stack a large sample of X-ray-undetected dwarf starburst and late-type galaxies in five redshift bins up to $z = 1.5$ using the Chandra COSMOS-Legacy survey data, and find AGN X-ray emission (with mean $L_{\text{X}} \approx 10^{39–40}$ erg s$^{-1}$) in each redshift bin after removing contributions of XRB and hot gas to the stacked signal. It seems promising for subsequent studies to obtain a key advance by finding better multiwavelength tracers of which galaxies are those likely to host AGNs, which can allow much improved X-ray stacking and diagnosis of physical drivers.

Combining the efforts of searching for low-mass AGNs with optical, MIR, and particularly X-ray observations, it appears evident that low-mass AGNs are quite common from the local universe at least out to moderate redshifts. Therefore, it is imperative to further build up statistically meaningful samples of low-mass AGNs across a reasonably broad redshift range and with a variety of host properties, which will certainly help us piece together a scenario for their evolution, lend lessons to studies of high-redshift SMBH seeds, make a critical step forward in understanding the AGN demography in the low-BH mass regime, and constrain further their behavior in the scaling relations between $M_{\text{BH}}$ and host galaxy properties (e.g., Greene et al. 2010; Jiang, Greene, & Ho 2011; Graham & Scott 2013, 2015; Reines & Volonteri 2015; Baldassare et al. 2015; Ho & Kim 2016).

2.5. Examining the significantly variable AGN subpopulation

Temporal and spectral variability is a defining and ubiquitous feature of AGNs, and variability studies are valuable for probing AGN physical properties (e.g., Mushotzky et al. 1993; Ulrich et al. 1997; Peterson 2001; Vaughan et al. 2003). The versatile utility of AGN variability analyses is rendered in many aspects,
e.g.: (1) A first-order estimation of the physical size of the emission region can be obtained by searching for the minimum variability timescale according to simple light-crossing time arguments, which is often employed in high-energy (X-ray and gamma-ray) observations of blazars (e.g., Xue & Cui 2005; Abdo et al. 2011). (2) Detailed reverberation-mapping based studies allow for size estimates of various AGN components such as accretion disk, broad-line region, and inner part of dusty torus (e.g., Koshida et al. 2014; Peterson 2014; Faustnaugh et al. 2016). (3) The properties of the absorbing matter such as obscuring wind and gas can be revealed by examining the changes in absorption (e.g., Filiz Ak et al. 2013; Netzer 2015). (4) The AGN X-ray power spectral density (PSD) is often modeled with a broken power-law of red-noise nature (e.g., Uttley, McHardy, & Papadakis 2002; also see Fig. 1 of Zhu & Xue 2016 for a demonstration). The break frequency of the PSD, the amplitude of the high-frequency PSD, and the integral of that part are found to be closely related to BH properties such as \( M_{\text{BH}} \) and mass accretion rate, thereby enabling measurements of such BH properties (e.g., McHardy et al. 2004, 2013; Zhou et al. 2010; González-Martín & Vaughan 2012; Kelly et al. 2013). (5) Significant variability has often been used as an effective AGN selection technique in X-ray, UV, and optical bands (e.g., Maoz et al. 2005; Trevese et al. 2008; Boutsia et al. 2009; Villforth et al. 2010; MacLeod et al. 2011; Young et al. 2012; Falocco et al. 2015; also see Table 3). (6) AGN variability observations can place constraints on the underlying emission processes as well as the models intended to explain the physical origins of AGN variability (e.g., Xue, Yuan, & Cui 2006; Kelly et al. 2009, 2014; Cai et al. 2016; Liu et al. 2016a).

AGN X-ray variability is of particular interest, because it is generally more rapid and has larger amplitudes than variability at longer wavelengths; moreover, X-ray emission is nearly universal among AGNs (see Section 4). It and it originates from the innermost part of the system, thus being able to probe the immediate vicinity of the SMBH (e.g., Ulrich et al. 1997). Therefore, many deep X-ray surveys have been conducted to study AGN variability down to low and moderate luminosities and up to high redshifts (e.g., Almaini et al. 2000; Paolillo et al. 2004, 2017; Mateos et al. 2007; Papadakis et al. 2008; Vagnetti et al. 2011, 2016; Lanzausisi et al. 2014; Shemmer et al. 2014; Middei et al. 2017; Zheng et al. 2017; Li et al. in prep.). The observations of the CDFS, particularly the CDF-S and CDF-N, cover a timespan up to \( \sim 16 \) years (see Table 1), resulting in the longest rest-frame timescales that can possibly be probed for X-ray variability analyses of a large sample of distant AGNs, and thus enabling a range of exciting variability science.

Young et al. (2012) identify 20 AGNs out of a sample of 92 galaxies with \( z \approx 0.08-1.02 \) in the 4 Ms CDF-S, solely based on their significant long-term X-ray variability. These 20 newly-identified AGNs fail all the other non-variability AGN identification criteria adopted by the 4 Ms CDF-S cataloging work (Xue et al. 2011; also see Table 3) and were then regarded as galaxies. These 20 AGNs have observed variability behavior that cannot be explained by XRB populations, and appear to be low-luminosity unobscured AGNs given their stacked \( L_{\text{X}} \) (accretion) \( \approx 1.93 \pm 0.13 \), with estimated \( M_{\text{BH}} \) (accretion) a factor of 2.4 (22.5) lower than variable luminous AGNs at the same redshift. This study underscores the advantage of using X-ray variability analysis to find low-luminosity unobscured AGNs that would otherwise be missed. Yang et al. (2016) conduct systematic long-term X-ray variability analyses of the 68 brightest radio-quiet AGNs at \( z \approx 0.6-3.1 \) in the 6 Ms CDF-S and find: (1) Among these sources, \( \approx 90\% \) are variable, indicating widespread photon flux variability; \( \approx 74\% \) display \( L_{\text{X}} \) variability, with quasars having smaller variability amplitudes; and \( \approx 16\% \) show \( N_{\text{H}} \) variability, with variability amplitudes becoming larger for longer time separation. (2) There are a few sources possessing very interesting variability patterns: a CTAGN candidate has variable high-energy X-ray flux, implying the existence of \( \leq 0.3 \) pc reflecting matter; a broad absorption line quasar is \( L_{\text{X}} \) variable; and a \( \approx 1.21 \) source appears to be a “semi-changing-look” AGN, transiting from an X-ray unobscured to obscured state (also see, e.g., Matt, Guainazzi, & Maiolino 2003; Ricci et al. 2016 for examples of X-ray changing-look AGNs) but always remaining as optically type I (see Fig. 12; also see, e.g., LaMassa et al. 2015; McElroy et al. 2016; Runnoe et al. 2016 for examples of optical-changing-look AGNs). This work showcases the uniqueness of the CDF-S dataset in revealing long-term AGN variability and some unusual variable AGNs. Paolillo et al. (2017) apply various PSD models to study the ensemble X-ray timing properties of AGNs during a timespan of \( \geq 16 \) years up to \( z \approx 4 \) in the 7 Ms CDF-S, and find that the average \( \Delta t_{\text{edd}} \) appears to be largely constant across this redshift range given the large statistical uncertainties, with only marginal evidence for a possible increase of \( \Delta t_{\text{edd}} \) that peaks at \( z \approx 2-3 \). This study shows the potential of using X-ray variability analysis to trace the overall SMBH accretion history, which can be fully developed when future large X-ray missions enlarge AGN samples and thus reduce statistical uncertainties associated with such measurements.
3. Interactions between AGNs and their host galaxies

3.1. AGN X-ray luminosity versus galaxy SFR

The two relatively independent major astrophysical research fields, AGNs and galaxies, have begun to be intimately connected with each other since a series of important observational discoveries of tight relationships between $M_{\text{BH}}$ and galaxy bulge properties (e.g., stellar velocity dispersion, luminosity, and mass; e.g., Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002; Ferrarese & Ford 2005; Gültekin et al. 2009; Kormendy & Ho 2013). Subsequent results show that the volume density of SMBH accretion rate (scaled up by a factor of several thousand) closely tracks the cosmic SFR density out to at least $z \approx 2$ (e.g., Heckman et al. 2004; Merloni et al. 2004; Silverman et al. 2008a; Zheng et al. 2009; Aird et al. 2010; Mullaney et al. 2012b; Kormendy & Ho 2013; Aird et al. 2015b), in support of the scenario of a closely related overall SMBH-galaxy growth. As a result, the co-evolution studies of galaxies and their central SMBHs have ever since aroused great interest and become a central research theme in modern astrophysics (see, e.g., Cattaneo et al. 2009; Kormendy & Ho 2013 for a review).

The AGN X-ray luminosity is related to SMBH accretion and the galaxy SFR is related to star-formation process; as such, $L_X$ and SFR are often taken as the observational indicators of respective growth of SMBHs and galaxies, with the relation between them providing hints for interactions between AGNs and their hosts, BH accretion modes, and likely effects of AGN feedback (see, e.g., Fabian 2012 for a review). Many studies have explored the relations between $L_X$ and SFR, providing a wide range of results, e.g.: (1) For $z \gtrsim 1$ moderate-luminosity ($L_X \approx 10^{42–44}$ erg s$^{-1}$) AGNs, there is generally no strong evidence for any $L_X$ dependence of SFR, with SFR being broadly constant over this $L_X$ range and AGN host specific SFR (sSFR) usually being roughly consistent with main-sequence galaxies (e.g., Lutz et al. 2010; Shao et al. 2010; Mullaney et al. 2012a; Rosario et al. 2012, 2013a; Azadi et al. 2015; Stanley et al. 2015; see Fig. 13; but see Barger et al. 2015); while at $z \lesssim 1$, there appears to be a positive correlation between SFR and $L_X$ with such a trend being most prominent in the local universe (e.g., Shao et al. 2010; Rosario et al. 2012, 2013a; Azadi et al. 2015; Stanley et al. 2015; see Fig. 13; but see Barger et al. 2015); while at $z \leq 1$, there appears to be a positive correlation between SFR and $L_X$ with such a trend being most prominent in the local universe (e.g., Shao et al. 2010; Rosario et al. 2012; Azadi et al. 2015; Lanzuisi et al. 2017; also see, e.g., Rowan-Robinson 1995; Netzer et al. 2007; Netzer 2009; Diamond-Stanic & Rieke 2012; Matsuoka & Woo 2015 for similar positive SFR-AGN luminosity trends; but see, e.g., Shimizu et al. 2017). (2) For $z \gtrsim 1$ high-luminosity ($L_X \gtrsim 10^{44}$ erg s$^{-1}$) AGNs, some studies suggest an increase of SFR or a roughly constant SFR toward large $L_X$ (e.g., Hatziminaoglou et al. 2010; Lutz
et al. 2010; Harrison et al. 2012; Rovilos et al. 2012; Santini et al. 2012; Stanley et al. 2015; Lanzuisi et al.
2017; also see, e.g., Lutz et al. 2008; Bonfield et al. 2011; Dong & Wu 2016; Harris et al. 2016 for similar
positive SFR-AGN luminosity trends), while some sug-
gest a decrease of SFR with increasing $L_X$, possibly due
to AGN feedback (e.g., Page et al. 2012; Barger et al.
2015).

There are several possible causes that can account, at
least partially, for some of the variations between the
above different observational results about moderate-
luminosity and high-luminosity AGNs. For example, it is
always challenging to obtain robust SFR measure-
ments, especially for high-luminosity AGNs that can
produce significant FIR emission and thus contami-
nate that of host galaxies; limited sample sizes (i.e.,
poor source statistics) appear to be an unavoidable is-
ssue most of the time; and, sometimes, cosmic varia-
tion may come into play. Additionally, some physi-
cal considerations may be the more fundamental causes,
e.g.: (1) The relative timescales of several processes
can largely determine whether the observational sig-
atures of suppressed star formation (due to, e.g., AGN
feedback) can be maintained and thus clearly revealed;
such processes include, e.g., star-formation activity, lu-
minous AGN activity (and its variability), the delay
between the onsets of star formation and AGN activ-
ity, and quenching of star formation (e.g., Di Mat-
teo et al. 2008; Hopkins & Quataert 2010; Gabor &
Bournaud 2013; Hickox et al. 2014; Neistein & Netzer
2014). (2) There are likely two modes in the concomi-
tant AGN activity and star formation that can transit
from each other when conditions allow, with one mode
primarily associated with low- and moderate-luminosity
AGNs where the evolution is secular, with no close cou-
pling between instantaneous AGN luminosity and cur-
cent galaxy-integrated SFR, and another mode mainly
associated with high-luminosity AGNs where the evo-
lution is intense and rapid likely through (major) merg-
ers, with tight coupling between AGN growth and host
star formation (e.g., Lutz et al. 2010; Shao et al. 2010;
Santini et al. 2012). (3) The exact relation between $L_X$
and SFR likely depends on whether the star formation
is dominated by a nuclear (i.e., sub-pc scales) or an ex-
tended ($\approx$ kpc) component, with the former case result-
ing in tighter coupling than the latter (e.g., Hopkins &
Quataert 2010; Diamond-Stanic & Rieke 2012).

Alternatively, the connection between AGN activity
and star formation can be examined using a different
approach from the above studies, which finds tight cor-
relations between SFR and AGN luminosity by aver-
ing AGN and star-formation activities over all galax-
ies (irrespective of hosting an AGN or not) in a cos-
ological volume (e.g., Rafferty et al. 2011; Syme-
onidis et al. 2011; Mullaley et al. 2012b; Chen et al.
2013; Delvecchio et al. 2015). Such an approach fo-
cuses on investigating long-term average/overall behav-
iors, thus being immune to many detailed factors men-
tioned above. As a most recent example, Yang et al.
(2017) study the dependence of sample-mean BH accre-
tion rate ($<$BHAR$>$) on host SFR and $M_*$ with a sam-
pale of $\approx$ 18000 galaxies at $0.5 \leq z < 2.0$, using the
7 Ms CDF-S and CANDELS/GOODS-S observations.

They compare $<$BHAR$>$ for samples with different SFR
and/or $M_*$, down to levels as low as $\text{SFR} < 0.1 \ M_\odot \ yr^{-1}$
and $M_* \approx 10^8 \ M_\odot$ with reasonable completeness for the
first time, finding that (see Fig. 11): (1) $<$BHAR$>$ is
positively correlated with both SFR and $M_*$, and both
the $<$BHAR$>$-SFR and $<$BHAR$>$-$M_*$ relations can be
adequately fit using linear models with slope$\approx 1$. (2) The
$<$BHAR$>$-$M_*$ relation is tighter than the $<$BHAR$>$-
SFR relation (confirmed by partial-correlation analy-
ses), indicating that BH growth is mainly linked to
$M_*$ rather than SFR, and the $<$BHAR$>$-SFR relation is
largely secondary and results from the well-known star-
formation main sequence (e.g., Elbaz et al. 2011). (3)
Massive galaxies (i.e., $M_* \geq 10^{10} \ M_\odot$) have
larger $<$BHAR$/> <$SFR$>$ ratios than less-massive galax-
ies, suggesting that the former have higher AGN frac-
tions and/or higher SMBH fueling efficiencies than the
latter (also see, e.g., Rodighiero et al. 2015). The Yang
et al. (2017) results are in accordance to the SMBH-
galaxy co-evolution scenario, have important implica-
tions for the $M_{\text{BH}}$-$M_*$ scaling (see Section 3.4),
and make one worry about likely mass effects in at least
some of the aforementioned $L_X$-SFR studies that are
based on direct detections.

The current research status of the relation between
AGN activity and star formation is somewhat perplex-
ing, as both the observational and theoretical pieces
don’t all fit together (see, e.g., Harrison 2017 for a re-
view). Future observational studies can be improved in
many aspects, e.g., more reliable SFR measurements for
individual sources, broader parameter spaces (e.g., SFR,
$M_*$, $L_X$, $z$) to be probed, better source statistics, and
higher sample completeness. These improved studies
will provide more sensitive tests of how AGN activity
impacts star formation and make a step forward in un-
derstanding AGN triggering mechanisms.

3.2. Conducive host galaxy properties for AGN activity

What kinds of host galaxy environments are most
AGN-friendly? This is a question that invites numer-
ous observational efforts, which are made over a broad
redshift range (up to $z \approx 3–4$) and obtain many interesting results, e.g.: (1) Distant X-ray AGN host galaxies are typically more massive than non-AGN galaxies, with the AGN fraction (above a given AGN cutoff luminosity) increasing strongly toward larger stellar masses (e.g., Akiyama 2005; Papovich et al. 2006; Alonso-Herrero et al. 2008; Brusa et al. 2009b; Xue et al. 2010; Aird et al. 2012; Mullaney et al. 2012a; Wang et al. 2012b).
There is a color bimodality among galaxy populations, with the majority being red-sequence and blue-cloud galaxies and the minority green-valley galaxies lying in between (e.g., Strateva et al. 2001; Bell et al. 2004; Brammer et al. 2009; Xue et al. 2010; see Row (a) of Fig. 15). X-ray AGN host galaxies are generally found to be optically more luminous and redder than non-AGN galaxies (e.g., Barger et al. 2003b; Nandra et al. 2007; Silverman et al. 2008b; Xue et al. 2010; Bongiorno et al. 2012). However, when stellar mass-matched samples are considered, both AGN hosts and non-AGN galaxies are comparably luminous with similar colors, and the AGN fraction is largely constant or even slightly decreasing toward red host colors (e.g., Silverman et al. 2009; Xue et al. 2010; Rosario et al. 2013a; Hernán-Caballero et al. 2014; see Row (b) of Fig. 15). (3) Most X-ray AGNs reside in star-forming and starburst hosts, with the AGN fraction increasing significantly toward higher SFRs (e.g., Silverman et al. 2009; Xue et al. 2010; Rafferty et al. 2011; Rosario et al. 2013a; Wang et al. 2017). Nevertheless, when mass-matched samples are examined, once again, the above trend of rising AGN fraction upon larger SFRs becomes less prominent or even non-detectable, and there is no apparent difference in terms of star-formation properties between AGN hosts and star-forming main-sequence galaxies (e.g., Xue et al. 2010; Mullaney et al. 2012a; see Row (c) of Fig. 15). (4) The first morphological studies of X-ray AGN hosts indicate that these AGNs preferentially reside in bulge-dominated systems (e.g., Grogin et al. 2005; Pierce et al. 2007). However, the mass-matched technique does the magic once more: clear differences between morphological types of AGN hosts and non-AGN galaxies mostly disappear when the AGN and galaxy samples are matched in mass, i.e., X-ray AGNs reside in a broad range of host-galaxy types that include, e.g., disk-dominated, bulge-dominated, irregular, and point-like morphology classes (or simply categorized as undisturbed and disturbed classes; e.g., Kocevski et al. 2012; Fan et al. 2014; Villforth et al. 2014; see Fig. 16).

Based on the above X-ray AGN results, it appears clear that, among a variety of host galaxy properties (e.g., stellar mass, optical luminosity and colors, SFR, and morphology) likely conducive for AGN activity, stellar mass plays the fundamental driving role in AGN triggering, while the other properties play a secondary role; this is hardly surprising, given the mass-light/luminosity (e.g., Zibetti, Charlot, & Rix 2009), mass-color (i.e., massive galaxies generally tend to be redder; e.g., Xue et al. 2010), mass-SFR (i.e., the star-formation main sequence; e.g., Elbaz et al. 2011), and mass-morphology (i.e., E/S0 galaxies generally dominate the higher-mass population; e.g., Bundy, Ellis, & Conselice 2005) correlations. There are several physically plausible causes for the dominant role of stellar mass in triggering AGN activity, e.g., massive galaxies (1) have stronger gravitational pull to make gas fall into galaxy centers and fuel SMBHs eventually, (2) are more likely to have nuclear bars to induce gas inflow efficiently, and (3) tend to have larger SMBHs that are more capable of accreting gas from their vicinity (see, e.g., Section 4.2 of Yang et al. 2017 for details). Therefore, the conditions that are most conducive for AGN activity appear to be a massive host galaxy and a large reservoir.
Gas-rich major mergers and triggering of distant AGNs are often closely related in theoretical considerations (e.g., Sanders et al. 1988; Di Matteo et al. 2005; Hopkins et al. 2008). Observationally, however, only \( \lesssim 20\% \) of the AGN population at \( z \approx 0-2.5 \) have clear signatures of major mergers (e.g., Koss et al. 2010; Silverman et al. 2011; Kocevski et al. 2012; Cotini et al. 2013; Villforth et al. 2014), which is much below theoretical expectations. Interestingly, recent morphological studies find that major mergers only trigger the most luminous AGNs (e.g., Treister et al. 2012; Rumbaugh et al. 2017; but see, e.g., Villforth et al. 2017); and highly obscured AGN hosts tend to undergo dynamical compaction (e.g., Chang et al. 2017), interactions, or mergers (e.g., Kocevski et al. 2015; Lanzuisi et al. 2015a; Del Moro et al. 2016). In particular, Del Moro et al. (2016) study a sample of \( z \approx 2 \) MIR-luminous quasars, and find that the highly obscured quasars tend to reside in galaxies with disturbed morphologies while the unobscured/moderately obscured quasars preferentially lie in undisturbed hosts, but the disturbed quasar hosts only constitute the minority (\( \approx 40\% \)) of the entire sample (see Fig. 16), which is consistent with other findings (e.g., Kocevski et al. 2015; Lanzuisi et al. 2015a).

The above results indicate that secular processes such as galaxy bars and disk instabilities, as well as minor mergers, may be the main fueling mechanisms for majority AGN populations. Therefore, it remains appealing for future morphological studies to quantify reliably and accurately the respective contribution of each likely AGN fueling mechanism as a function of, e.g., \( L_X \) and redshift.

3.3. SMBH growth behavior revealed by \( \Gamma-\lambda_{\text{Edd}} \) relation and \( \lambda_{\text{Edd}} \) distribution

\( \lambda_{\text{Edd}} \) is an important parameter in accretion disk theories, which can be used as a primary indicator of SMBH growth rate and accretion mode.\(^{11}\) Recent observational results about the AGN \( \Gamma-\lambda_{\text{Edd}} \) relation reveal a V-shape correlation: as \( \lambda_{\text{Edd}} \) decreases continuously, the \( \Gamma-\lambda_{\text{Edd}} \) correlation changes from being positive (e.g., Wang et al. 2004; Shemmer et al. 2006, 2008; Risaliti et al. 2009; Zhou & Zhao 2010; Brightman et al. 2013; Fanali et al. 2013; Yang et al. 2015) into being moderate-luminosity AGNs, whose feedback may be either important (e.g., Fabian et al. 2008; Raimundo et al. 2010) or not (e.g., Hopkins & Hernquist 2006).

\(^{10}\)Interestingly, theoretical studies predict diverse roles played by gas (e.g., Silverman et al. 2009; Vito et al. 2014b). In combination with the fact that galaxy stellar masses can be estimated more reliably than colors and luminosities (e.g., van Dokkum et al. 2006), consequently, evolutionary studies of galaxies and SMBHs are best probed using stellar-mass selected samples (e.g., van Dokkum et al. 2006; Kriek et al. 2008) and/or mass-matched samples (e.g., Silverman et al. 2009; Xue et al. 2010).

As such, the digesting of the above results derived with mass-matched samples plainly indicates at least two points: (1) there is essentially no substantial difference between many physical properties (e.g., optical luminosity and colors, SFR, and morphology) of X-ray AGN hosts and coeval non-AGN galaxies over a broad range of redshift (but see, e.g., Mullaney et al. 2015b; Wang et al. 2017); and (2) currently there appears no strong direct observational evidence for feedback effects from moderate-luminosity X-ray AGNs (e.g., Xue et al. 2010; Bongiorno et al. 2012; Rosario et al. 2013b; Azadi et al. 2017; but see, e.g., Wang et al. 2017).\(^{10}\)
negative (e.g., Constantin et al. 2009; Gu & Cao 2009; Younes et al. 2011; Jang et al. 2014; Yang et al. 2015; Kawamuro et al. 2016), with the transition occurring at \( \lambda_{Edd} \approx 1 \).\(^{12}\) It is generally believed that these observational results reflect that, as the accretion rate decreases, the SMBH accretion mode changes accordingly, from the original standard thin accretion disk into radiatively inefficient accretion flow (e.g., Gu & Cao 2009; Brightman et al. 2013; Jang et al. 2014; Yuan & Narayan 2014). Recently, Yang et al. (2015) propose a coupled hot accretion flow-jet model, which can well explain the overall \( \Gamma - \lambda_{Edd} \) correlation and even predicts that, toward the very low \( \lambda_{Edd} \) regime (e.g., \( \lambda_{Edd} \lesssim 10^{-5} \)), this correlation will display a small bump and then level off eventually. Interestingly, Liu et al. (2016b) find that, toward the very high \( \lambda_{Edd} \) regime (e.g., \( \lambda_{Edd} \approx 0.1 \)), the \( \Gamma - \lambda_{Edd} \) correlation appears to be largely constant or even negative, which is different from previous findings. Therefore, it would be useful to build a large, uniform, and (relatively) complete AGN sample with, e.g., the CDF observations in conjunction with other data, to study in detail the \( \Gamma - \lambda_{Edd} \) correlation and its likely dependence on redshift and host galaxy properties (Sun et al. in prep.), in order to thoroughly compare observations with various SMBH accretion models. If reliably calibrated (see, e.g., Section 4.3 of Brandt & Alexander 2015 for the challenges in obtaining reliable estimates of \( \Gamma \) and \( \lambda_{Edd} \)), the \( \Gamma - \lambda_{Edd} \) correlation may be used to conveniently estimate some important AGN parameters (e.g., \( \lambda_{Edd} \), \( M_{BH} \), and mass accretion rate), which will facilitate better understanding accretion processes and evolution of AGNs.

Observational \( \lambda_{Edd} \) estimates for \( z \approx 0-4 \) X-ray AGNs span a broad range of \( \lambda_{Edd} \approx 10^{-5} - 1 \), with the majority having \( 10^{-4} - 0.1 \) (e.g., Babić et al. 2007; Ballo et al. 2007; Alonso-Herrero et al. 2008; Brusa et al. 2009b; Hickox et al. 2009; Raimundo et al. 2010; Trump et al. 2011; Lusso et al. 2012; Matsuoka et al. 2013; Azadi et al. 2015; Suh et al. 2015; Bernhard et al. 2016). On average, \( \lambda_{Edd} \) seems to increase with redshift at any given \( M_{BH} \), but has no clear evolution with redshift at any given \( L_{bol} \) (e.g., Lusso et al. 2012). Further careful analyses reveal that the intrinsic \( \lambda_{Edd} \) distribution for X-ray AGNs appears to follow a “universal” power-law with a slope independent of both \( M_* \) and redshift out to \( z \approx 2.5 \) (Aird et al. 2012, 2013; Bongiorno et al. 2012; Azadi et al. 2015; Wang et al. 2017), implying that the basic physical processes responsible for triggering and fueling the overall AGN population might be essentially the same; such a power-law distribution is in contrast to a lognormal distribution, with the former also seen for, e.g., AGNs in quiescent galaxies (e.g., Kauffmann & Heckman 2009; but see, e.g., Jones et al. 2016) and the latter seen for, e.g., the observed \( \lambda_{Edd} \) distribution for X-ray AGNs (e.g., Aird et al. 2012) as well as the \( \lambda_{Edd} \) distribution for optically luminous AGNs (e.g., Kollmeier et al. 2006) and AGNs in star-forming galaxies (e.g., Kauffmann & Heckman 2009).

Specifically, Wang et al. (2017) study the \( \lambda_{Edd} \) distribution (\( p(\lambda_{Edd}) \)) for a mass–complete sample of \( 0.5 < z < 2.5 \) moderate-luminosity X-ray AGNs in the two GOODS fields, and find some interesting results (see Fig. 17): (1) \( p(\lambda_{Edd}) \) for the overall galaxy population appears to be a power-law with a slope of \(-0.4\), being consistent with the slope obtained by Jones et al. (2016) but slightly shallower than the slope of \(-0.6\) measured by, e.g., Aird et al. (2012). (2) However, in terms of redshift evolution, \( p(\lambda_{Edd}) \) is different for galaxies with different intrinsic colors, such that red galaxies having more rapid redshift evolution (i.e., \( p(\lambda_{Edd}) \propto (1+z)^{3/2} \)), agreeing with previous results for the overall galaxy population; e.g., Aird et al. 2012; Bongiorno et al. 2012) than both blue and green galaxies (note that \( p(\lambda_{Edd}) \propto (1+z)^{1.8} \) for the overall galaxy population, which includes red, green, and blue galaxies). (3) There is marginal evidence for red galaxies having a steeper power-law slope (\(-0.6\)) of \( p(\lambda_{Edd}) \) than both blue and green galaxies, with the former slope in agreement with previous measurements (e.g., Kauffmann & Heckman 2009; Aird et al. 2012). The Wang et al. (2017) results reveal the strong dependence of SMBH accretion on their host colors, and thus caution that ambiguous conclusions may be drawn without taking such a color dependence into account.

All the above works on X-ray AGN \( \lambda_{Edd} \) distribution suffer, to various degrees, from issues such as limited sample sizes, in particular toward higher redshifts and lower X-ray luminosities, thus resulting in conclusions associated with large uncertainties. Therefore, future studies using larger and relatively complete samples, more uniform AGN identification criteria, and more reliable \( \lambda_{Edd} \) estimates would be critical to nail down the detailed behavior of intrinsic X-ray AGN \( \lambda_{Edd} \) distribution as a function of, e.g., redshift, host stellar mass, and intrinsic colors, thereby providing sharp insights into SMBH growth processes across cosmic time.

\(^{12}\)See, e.g., Fig. 1 of Gu & Cao (2009) and Fig. 2 of Yang et al. (2015) for a demonstration of the V-shape \( \Gamma - \lambda_{Edd} \) correlation, which is also seen in some well-observed individual AGNs (e.g., Sobolewska & Papadakis 2009; Emmanoulopoulos et al. 2012; Connolly et al. 2016) as well as in XRB populations (e.g., Yang et al. 2015).
3.4. Coeval growth of SMBHs and their hosts

The coeval growth of SMBHs and their host galaxies is vividly reflected (in part) in the observed tight $M_{\text{BH}}$-$M_{\text{bulge}}$ correlation (e.g., Kormendy & Gebhardt 2001; McLure & Dunlop 2002; Marconi & Hunt 2003; Ferrarese & Ford 2005; Lauer et al. 2007; Gültekin et al. 2009; Sani et al. 2011; Vika et al. 2012), which is largely in the form of $M_{\text{BH}} \sim M_{\text{bulge}}^{1.2}$ (e.g., Wandel 1999; Häring & Rix 2004; Kormendy & Ho 2013; McConnell & Ma 2013) with an intrinsic scatter of $\approx 0.3$ dex (e.g., Kormendy & Ho 2013). The tight $M_{\text{BH}}$-$M_{\text{bulge}}$ correlation can be explained either by AGN feedback where the coupling between AGN triggering and feedback-regulated star formation plays a key role (e.g., Di Matteo et al. 2005; Hopkins et al. 2006b; Sijacki et al. 2007; Booth & Schaye 2009; Fabian 2012), or by galaxy mergers based on the central limit theorem (e.g., Peng 2007; Hirschmann et al. 2010; Jahnke & Macciò 2011). In addition to the $M_{\text{BH}}$-$M_{\text{bulge}}$ relation, many works study the $M_{\text{BH}}$-$M_*$ relation instead and find similar tight correlations (e.g., Kormendy & Richstone 1995; Magorrian et al. 1998; Gültekin et al. 2009; Merloni et al. 2010; Schulze & Gebhardt 2011; McConnell & Ma 2013; Reines & Volonteri 2015; Sun et al. 2015), given the facts that $M_*$ is an important physical property (see Section 3.2 being relatively easy to measure, $M_{\text{bulge}}$ is approximately equal to $M_*$ for bulge-dominated systems, and it is challenging to separating $M_{\text{bulge}}$ from $M_*$ at $z \gtrsim 1$. Together, scaling relations between SMBHs and their hosts (including,
e.g., $M_{\text{BH}}-M_{\text{bulge}}$ and $M_{\text{BH}}-M_{\star}$) are fundamentally important to understanding SMBH and galaxy evolution across cosmic time.

Obviously, observational studies of the evolution of the $M_{\text{BH}}-M_{\star}$ ($M_{\text{bulge}}$) relation are of great interest yet with diverse results. For example, many works find that the $M_{\text{BH}}/M_{\star}$ ($M_{\text{BH}}/M_{\text{bulge}}$) ratio evolves positively with redshift, with high-redshift SMBHs being a factor of several overmassive relative to their host galaxies compared to their local counterpart systems (e.g., Peng et al. 2006a, 2006b; Ho 2007; Merloni et al. 2010; Bennert et al. 2011; Bongiorno et al. 2014; Shankar et al. 2016). However, some other works find no difference in the $M_{\text{BH}}/M_{\star}$ ratio between high-redshift and local measurements (e.g., Jahnke et al. 2009; Schramm & Silverman 2013; Sun et al. 2015). Additionally, there are works finding high-redshift undermassive SMBHs that are located below the local scaling relation (e.g., Borys et al. 2005; Alexander et al. 2008b; Urrutia et al. 2012). They find: (1) There is no evolution in the $M_{\text{BH}}/M_{\star}$ relation (i.e., "flat function") due to both sample selection limits and the steep slope of the $M_{\text{BH}}-M_{\star}$ relation. (2) The $M_{\text{BH}}-M_{\star}$ relation is of great interest yet have important implications for mass assembly histories of SMBHs and their host galaxies.

Figure 18: (Top) $M_{\text{BH}}/M_{\star}$ as a function of $z$. The three red squares indicate average $M_{\text{BH}}/M_{\star}$ for sources with $z < 1.0$, $1.0 \leq z < 1.5$, and $z \geq 1.5$. The red solid line indicates the local Haring & Rix (2004) $M_{\text{BH}}-M_{\text{bulge}}$ relation ($M_{\text{bulge}} \approx M_{\star}$ here as these local host galaxies are bulge-dominated; dashed line; the shaded region indicates its uncertainty) as it would be observed under the biases of the sample presented here. No evolution of the $M_{\text{BH}}-M_{\star}$ relation is observed at $0.2 \leq z < 2.1$. (Bottom) “Flow patterns” of SMBHs and their host galaxies in the $M_{\text{BH}}-M_{\star}$ plane, with arrows showing the evolutionary directions and colors indicating the absolute value of the specific BH mass accretion rate. The large cross indicates $1 \sigma$ uncertainties of $M_{\text{BH}}$ and $M_{\star}$. The self-maintenance behavior of the $M_{\text{BH}}-M_{\star}$ relation is evident. Adapted from Shen et al. (2015).
then examining likely dependence of the $M_{\text{BH}}$-$M_*$ scaling relation on them, thereby further unraveling their co-evolutionary details.

4. Evolution of starburst and normal galaxy X-ray emission

Remarkable insights into the formation and evolution of populations of XRBs have been gained through *Chandra* studies of local galaxies (see, e.g., Fabian 2006 for a review); only ultradeep *Chandra* observations have enough sensitivity to study the X-ray properties of starburst and normal galaxies at large cosmological distances either by individually detecting them or by sensitively stacking their signal, thus providing a direct view of the cosmic history of XRB production and evolution. Indeed, the CDFs have detected a significant X-ray galaxy population, e.g., 28.3% of the 1008 7 Ms CDF-S main-catalog sources are galaxies with $0.038 \leq z \leq 2.636$ (see Table I). As predicted by previous X-ray number-counts studies, the contribution of galaxies to the cumulative X-ray number counts is rapidly rising toward low fluxes, and galaxies outnumber AGNs in the SB for the first time around the 7 Ms CDF-S flux limits (see Fig. 31 of Luo et al. 2017 and Section 2.1). These detections will enable substantial improvement of insight into how AGN activity keeps the long-term balance between gas heating and cooling (e.g., Danielson et al. 2005; Lehmer et al. 2010; Boroson et al. 2011; Mineo et al. 2012, 2014; Vattakunnel et al. 2012; Zinn et al. 2012; Basu-Zych et al. 2013; Symeonidis et al. 2014). However, the large scatters in these X-ray scaling relations cannot be accounted for by measurement errors and/or statistical fluctuations (e.g., Hornschemeier et al. 2005; Mineo et al. 2012), and hence point to some real physical variations of, e.g., stellar ages, metallicities, and star formation histories that likely influence XRB formation and evolution significantly (see, e.g., Madau & Dickinson 2014 for a review). Fragos et al. (2013a) present a much improved framework of theoretical XRB population-synthesis models that is supported by many subsequent observational tests, being able to track XRB population evolution throughout cosmic history, make predictions for redshift evolution of the scaling relations, and identify a “best-fit” theoretical model for the local scaling relations. With the CDF and new deep multiwavelength data available, it is now plausible to use stacking techniques to isolate large populations of galaxies, obtain their global physical properties, investigate their population-averaged X-ray emission, and compare with XRB population-synthesis models in great detail (e.g., Laird et al. 2006; Lehmer et al. 2007, 2008; Cowie et al. 2007; Tremmel et al. 2013; Lehmer et al. 2014). Furthermore, *Chandra* stacking analyses have revealed evolution of the high-mass X-ray binary (HMXB) populations in late-type galaxies as a response to the rapidly increasing cosmic SFR with redshift (e.g., Reddy et al. 2005; Laird et al. 2006; Lehmer et al. 2008; Cowie et al. 2012); $z \approx 0.5$–1.4 late-type galaxies are 5–15 times more X-ray luminous (per unit stellar mass) than their local counterparts, showing evidence for “downsizing” of XRB populations. In contrast, stacking results of luminous early-type galaxies indicate little evolution of their hot interstellar gas up to $z \approx 1.2$, providing insights into how AGN activity keeps the long-term balance between gas heating and cooling (e.g., Danielson et al. 2012).

A number of studies based on *Chandra* observations have constrained scaling relations between $L_X$ and SFR for HMXB populations (i.e., $L_X$(HMXB)/SFR) and between $L_X$ and $M_*$ for low-mass X-ray binary (LMXB) populations (i.e., $L_X$(LMXB)/$M_*$; e.g., Bauer et al. 2002a; Ranalli et al. 2003; Colbert et al. 2004; Hornschemeier et al. 2005; Lehmer et al. 2010; Boroson et al. 2011; Mineo et al. 2012, 2014; Vattakunnel et al. 2012). These detections will enable substantial improvement of insight into how AGN activity keeps the long-term balance between gas heating and cooling (e.g., Danielson et al. 2005; Lehmer et al. 2010; Boroson et al. 2011; Mineo et al. 2012, 2014; Vattakunnel et al. 2012; Zinn et al. 2012; Basu-Zych et al. 2013; Symeonidis et al. 2014). However, the large scatters in these X-ray scaling relations cannot be accounted for by measurement errors and/or statistical fluctuations (e.g., Hornschemeier et al. 2005; Mineo et al. 2012), and hence point to some real physical variations of, e.g., stellar ages, metallicities, and star formation histories that likely influence XRB formation and evolution significantly (see, e.g., Madau & Dickinson 2014 for a review). Fragos et al. (2013a) present a much improved framework of theoretical XRB population-synthesis models that is supported by many subsequent observational tests, being able to track XRB population evolution throughout cosmic history, make predictions for redshift evolution of the scaling relations, and identify a “best-fit” theoretical model for the local scaling relations. With the CDF and new deep multiwavelength data available, it is now plausible to use stacking techniques to isolate large populations of galaxies, obtain their global physical properties, investigate their population-averaged X-ray emission, and compare with XRB population-synthesis models in great detail (e.g., Laird et al. 2006; Lehmer et al. 2007, 2008; Cowie et al. 2012; Basu-Zych et al. 2013).

Recently, Lehmer et al. (2016) make use of the 6 Ms CDF-S data through reliable X-ray stacking analyses to examine the dependence of galaxy XRB emission on SFR, $M_*$, and redshift (in the range of $z \approx 0$–7), thereby conducting the most powerful and robust tests to date of the Fragos et al. (2013) model predictions. Their findings are as follows: (1) Scaling relations involving SFR, $M_*$, and redshift simultaneously best characterize global galaxy X-ray emission ($L_X$), in stark contrast to a widely assumed “universal” $L_X$/SFR relation (see Fig. 19). (2) HMXB and LMXB populations appear to evolve as $L_{2-10\text{ keV}}$(HMXB)/SFR $\propto (1 + z)$ and $L_{2-10\text{ keV}}$(LMXB)/$M_*$ $\propto (1 + z)^{3/3}$, respectively, at least up to $z \approx 2.5$, which is consistent with basic XRB population-synthesis model predictions that attribute the increase in the HMXB/LMXB relation with redshift primarily to effects related to the decrease in metallicities/stellar ages (Fragos et al. 2013a; also see Aird, Coil, & Georgakakis 2017). However, the marginal agreement between the observational data and the Fragos et al. (2013) best-fit model necessitates minor revisions of such XRB population-synthesis models (see Fig. 19). (3) LMXBs likely dominate galaxy X-ray emissivity (i.e., $L_X$ per volume) out to $z \approx 1$–2, while HMXBs take over at higher redshifts. (4) The overall galaxy X-ray emissivity peaks around $z \approx 1.5$–3, mimicking the cosmic SFR density, but declines more slowly at $z \gtrsim 3$ than the latter owing to the rising $L_X$/SFR.
scaling with redshift; extrapolation of these results indicates that galaxies provide a larger X-ray emissivity than AGNs at $z \geq 6-8$, thereby dominating the reionization process, as expected by XRB population-synthesis models (see, e.g., Fragos et al. 2013a, 2013b; also see, e.g., Vito et al. 2016 mentioned in Section 2.2).

This Lehmer et al. (2016) work, together with many previous relevant studies, can be significantly improved, e.g., by utilizing statistically significant samples of local (through direct detections) and distant galaxies (mainly relying on stacking techniques) with a variety of physical characteristics (e.g., $M_*$, SFR, stellar age, metallicity, and morphology), from an optimal combination of current and scheduled, wide and deep Chandra surveys. Without question, these works will be greatly boosted even further, thanks to the employment of future powerful X-ray observatories (see Section 7).

5. Census of X-ray galaxy groups and clusters

Theoretically, there is almost no doubt that the growth and evolution of SMBHs and galaxies are environmentally dependent (e.g., Kauffmann 1996; De Lucia et al. 2006). However, in terms of detailed observational manifestations, it is still not so clear about the exact role that large-scale structures (LSSs; including galaxy groups, clusters, and superclusters) play regarding AGN triggering and star-formation activities (see, e.g., McNamara & Nulsen 2012 for a review on AGN feedback in LSSs), which is often complicated by additional likely dependencies such as host stellar mass, AGN luminosity, and redshift. Galaxy groups and clusters, lying at the high end of the cosmic density spectrum, have been intensively observed by X-ray observations that are subject to the least observational biases (e.g., Brandt et al. 2001; Bauer et al. 2002b; Giacconi et al. 2002; Finoguenov et al. 2006, 2007, 2009, 2010; Bielby et al. 2010). One of the current focuses is to accumulate a statistical sample of high-redshift (i.e., $z \geq 1.5$) groups, as they are likely progenitors of the local clusters in a statistical sense, thus being key to better understand the origin of the environmental dependence seen locally. In this regard, a number of such systems up to $z \approx 2$ have been discovered, although some of these systems need further spectroscopic observations for confirmation, e.g.: Povich et al. (2010) and Tanaka et al. (2010) independently confirm a $z = 1.62$ group in the Subaru/XMM-Newton Deep Field; Henry et al. (2010) report a possible $z = 1.75$ X-ray group; Gobat et al. (2011) find a $z = 2.07$ group by color selection; and Andreon et al. (2009) and Spitler et al. (2012) present a few $z \approx 2$ groups based on $z_{\text{phot}}$ selection. These high-redshift groups, combined with those at lower redshifts, provide an ideal resource to understand the evolution of groups and examine galaxy formation and evolution across the environments and cosmic time, and is also a powerful probe of cosmology (e.g., Finoguenov et al. 2010, 2015).

Recently, Finoguenov et al. (2015) combine the ultradeep Chandra and XMM-Newton observations in the CDF-S/E-CDF-S to carry out a systematic search for X-ray groups down to an unprecedented flux level of $2 \times 10^{-16}$ erg cm$^{-2}$ s$^{-1}$, with extensive spectroscopic observations providing identifications of galaxy groups out to high redshifts. They produce an X-ray group catalog that consists of a total of 46 spectroscopically identified systems up to $z = 1.61$ and down to low masses (see the left panel of Fig. 20), and investigate their properties, finding that: (1) The number counts and XLF of the E-CDF-S X-ray groups are in broad agreement with expectations of the $\Lambda$CDM cosmological model (also see, e.g., Liu et al. 2015). (2) The low-luminosity X-ray groups are confirmed to be low-mass systems through one- and two-point statistics as well as weak-lensing analysis. (3) The scaling relations between the X-ray luminosity and total group mass is verified and extended to regimes of lower masses and higher redshifts, based
on stacked weak lensing and clustering analyses. These groups, with broad ranges of mass, luminosity, and redshift, make a good and representative case of groups constituting the most common environment for galaxy evolution. Moreover, there are a few notable individual structures in the E-CDF-S. For instance, the two most prominent and populated LSSs in the E-CDF-S are located at $z = 0.67$ and $z = 0.73$, each traced by over 60 X-ray AGNs and many more galaxies (see the middle panel and Right-bottom panel of Fig. 20); and interestingly, the former LSS is in a shape of thick sheet with a radial size of 67.7 Mpc extending over the full E-CDF-S field, and in contrast, the latter LSS is thin (18.8 Mpc) and filamentary (e.g., Silverman et al. 2010). Another eye-catching system is the $z = 1.61$ group that is the most distant X-ray group identified in the E-CDF-S and examined in detail by Tanaka et al. (2013). They find that this group is actually the lowest-mass one (with $(3.2 \pm 0.8) \times 10^{13} M_\odot$) ever confirmed at $z > 1.5$, and exhibits a surprisingly prominent red sequence of quiescent early-type galaxies whose star formation is likely shut down by the bright AGN group members.

The above example works demonstrate that the CDFs are good sky patches for the census of X-ray groups and clusters, as further elucidated in the right panel of Fig. 20 that shows a number of CDF $z_{\text{spec}}$ spikes indicating (likely) LSSs, where the most up-to-date secure $z_{\text{spec}}$ compilations from Xue et al. (2016) and Luo et al. (2017) are adopted. Indeed, the CDFs are rich in LSSs: in the CDF-N, the previously known X-ray-source-traced LSSs located at $z = 0.12$, 0.46, 0.63, 0.843, 1.02, 1.15, 2.0, and 2.2 are all recovered by this
Figure 20: (Left) XMM-Newton detection of the extended emission on a 32 arcsec scale in the full E-CDF-S field, overlaid with the contours that show the extended emission detected in the combined Chandra and XMM-Newton images on the 32 and 64 arcsec scales. (Middle) The most prominent spectroscopically identified $z \approx 0.73$ group/cluster in the E-CDF-S (i.e., the highest redshift spike in the Right-bottom panel), with the contours overlaid onto the WIJ $R$-band image showing the extended X-ray emission detected by XMM-Newton and Chandra. X-ray sources with $|z_{\text{spec}}-0.73|<0.01$ are labeled as circles (GOODS spectroscopy) and squares (VLT/VIMOS and Keck/DEIMOS spectroscopy). (Right) Secure $z_{\text{spec}}$ distributions of the CDF X-ray sources in bins of $\Delta z_{\text{spec}}=0.02$ (the CDF-S $z_{\text{spec}}$’s are compiled by Luo et al. 2017, and the CDF-N and E-CDF-S $z_{\text{spec}}$’s by Xue et al. 2016). The vertical dotted lines indicate the previously known redshift spikes (i.e., associated with LSSs) identified by X-ray sources, and the downward arrows indicate additional likely redshift spikes identified in this paper. (Left) adapted from Finoguenov et al. (2015) and (Middle) adapted from Silverman et al. (2010).

In addition to the aforementioned exciting science that can be done with the CDFs, there are furthermore a wide variety of interesting topics that can be tackled using the CDFs, e.g.: (1) Bauer et al. (2017) report a remarkable new fast X-ray transient discovered by Luo et al. (2014b) in course of the extension of CDF-S observations from 4 Ms to 7 Ms, whose intriguing X-ray and multiwavelength properties effectively rule out the vast majority of previously known high-energy transients, leaving out only a few exotic theoretical possibilities that still cannot completely explain all observed properties. The inferred rate of such events is crudely comparable to that of gamma-ray bursts, which indicates the discovery of an untapped regime for a known transient type or a new class of transient phenomenon with its nature to be determined. (2) Cappelluti et al. (2017b) report a 3$\sigma$ detection of an $\approx 3.5$ keV emission line in the CXRB spectrum derived with a combined $\approx 10$ Ms Chandra exposure of the CDF-S and COSMOS Legacy surveys, and discuss the likely origins of this observed line that include the iron line background, S XVI charge exchange, and sterile neutrino decay. (3) Using the 4 Ms CDF-S, HST, and Spitzer data, Mitchell-Wynne et al. (2016) perform and compare the cross-correlation analyses between X-ray and optical/NIR cosmic background intensity fluctuations and find that the sources responsible for the cosmic IR background at 3.6 and 4.5 $\mu$m are at least partly dissimilar to those at $\leq 1.6 \mu$m (also see, e.g., Cappelluti et al. 2013; Helgason et al. 2014). (4) There are a significant number of SMGs identified in the CDFs, which are of great interest given their important role in the SMBH/galaxy co-evolution picture (e.g., Alexander & Hickox 2012; also see Section 3). Using the CDF and multiwavelength data, many topics can be studied in detail, including, e.g., the incidence of AGNs in SMGs, the location of SMGs in color-magnitude/mass...
diagrams, and coeval growth and evolutionary track of SMBHs and SMGs (e.g., Alexander et al. 2005b; Xue et al. 2010; Wardlow et al. 2011; Wang et al. 2013). (5) There is an extended X-ray emission identified in the CDF-N that is best explained as an inverse Compton ghost of a giant radio source (e.g., Fabian et al. 2009), calling for a systematic search for such sources in the CDFs. (6) There are a few tens of individual off-nuclear XRBs at $z \approx 0.05–0.3$ discovered in the CDFs, allowing for investigating their properties and redshift evolution as a population (e.g., Hornschemeier et al. 2004; Lehmer et al. 2006; also see Section 3). (7) There is a minority population of “interloping” Galactic stars in the CDFs, whose long-term evolution of magnetic activity can be studied by examining their X-ray emission (e.g., Feigelson et al. 2004).

7. Summary and prospects

The CDFs represent the amazing outcome of the multi-decade efforts of numerous people (counting from the proposal of Chandra submitted to NASA by R. Giacconi and H. Tananbaum in 1976 to the last CDF-S observation taken in March 2016), act as a major thrust among extragalactic X-ray surveys that are complemented effectively by deep multiwavelength observations, contribute critically to our dramatically improved characterization of the 0.5–8 keV CXRB sources, enable a wide range of scientific topics (including AGNs, starburst and normal galaxies, groups and clusters of galaxies, LSSs, etc.), and launch literally hundreds of exciting research papers.

In this paper, I have highlighted some recent key observational results that are mostly from the CDFs and enabled by the revolutionary and versatile scientific capabilities of Chandra, including the AGN demography (see Section 2), the interactions between AGNs and their host galaxies (see Section 3), the evolution of starburst and normal galaxy X-ray emission (see Section 4), and the census of X-ray galaxy groups (see Section 5) through cosmic time. The beauty and power of the CDF treasure trove will surely be further augmented given the consistent and full exploitation of the latest CDF catalogs and products (Xue et al. 2016; Luo et al. 2017) in conjunction with the ever improving multiwavelength observations.

Despite the great advances in the above research areas, it has almost always been challenging to tell a consistent and credible story in each of these areas or regarding some of the areas as a whole, which requires reconciling/distinguishing inconsistent results (owing to, e.g., differences in sample selection criteria, analysis techniques, data qualities and depths, incompleteness levels, and/or sample sizes, as well as likely effects of cosmic variance) and then stringing up all the reasonable pieces. Furthermore, there are many significant open questions that should persist for a foreseeable future, e.g.: (1) How to ultimately resolve the CXRB? (2) What processes dictate the formation and growth of the first SMBHs? (3) What drives the AGN downsizing behavior? (4) What are the origins of AGN variability? (5) How to obtain a census of highly obscured AGNs, CTAGNs, LLAGNs, and low-mass BHs as completely as possible? (6) What is the exact role of AGN feedback and what is the link between SMBH accretion and star formation? (7) How do accreting XRB populations evolve over most of cosmic time? (8) How do LSSs affect AGN activity? To address these questions, there are at least three aspects to move forward with each having both short-term and long-term goals associated with current and future facilities, respectively.

- Going wider. Ultradeep pencil-beam surveys such as the CDFs (see Table 1) are inevitably subject to the effects of cosmic variance. To remedy this situation and facilitate the science that requires larger solid-angle coverages, a straightforward solution is to widen the spatial coverages of the small-area surveys at moderate depths (thus being relatively easily achievable). Indeed, the CDF team has been in the process of proposing a $\approx 12 \text{ deg}^2$ Chandra/XMM-Newton survey at 30/50 ks depth of the SERVS areas of Wide-CDF-S (W-CDF-S, centered at CDF-S), ELAIS-S1, and XMM-LSS, with 1.3 Ms of XMM-Newton observing time awarded already (PI: W. N. Brandt; see Fig. 21). These three fields all have multiple intensive radio-to-UV observations performed or scheduled but lack the critical X-ray coverage. Therefore, the addition of the proposed X-ray observations will powerfully leverage those multiwavelength surveys by detecting thousands of new X-ray AGNs and hundreds of new X-ray groups/clusters, thereby dramatically advancing studies of SMBH growth across the full range of cosmic environments (from voids, groups, clusters, to the largest structures found in cold dark matter simulations), links between SMBH accretion and star formation, exceptional AGNs and protoclusters at high redshifts, and other topics. In the long run, the funded ESA-led mission Athena (band: 0.3–12 keV; large collecting area:

13Perhaps these different and even inconsistent observational manifestations highlight, to some degree, the vast complexities and personalities of galaxies and AGNs that essentially preclude any simple “universal” scenarios regarding many questions about them.
≈ 2.0–2.5 m²; large field of view: ≈ 40 × 40 arcmin; good angular resolution: ≈ 3–5 arcsec; scheduled launch: 2028; Nandra et al. 2013) will be ≈ 100 times more efficient in carrying out deep and wide surveys than Chandra and XMM-Newton for a given combination of solid-angle coverage and flux limit (above$f_{SB} \geq (1 - 2) \times 10^{-17}$ erg cm⁻² s⁻¹), which corresponds to Athena’s source confusion limit and sensitivity limits of ≈ 2 Ms Chandra exposures; see Table 1, thus substantially revolutionizing extragalactic X-ray studies. A large strategic mission concept, Lynx (called X-ray Surveyor previously; e.g., Weisskopf et al. 2015), is essentially to build a super Chandra with significantly enlarged collecting area (thus tremendously increased sensitivities) and still sub-arcsecond angular resolution; if selected and funded, it will take the sharpest X-ray vision even fainter and farther, thereby prodigiously transforming the field as Athena would do.

• Going harder. Sensitive hard X-ray ($\geq 10$ keV) surveys can open up a large volume of discovery space uncharted by Chandra and XMM-Newton, deciphering the even more energetic universe. Deep hard X-ray observations that fully encompass the CXRB peak can provide additional critical insights into the origin of the CXRB and unveil the highly obscured and CT AGNs. In this regard, NuSTAR, as the first focusing hard X-ray observatory in orbit, has been a pathbreaker that is significantly influencing the field. NuSTAR has already observed the CDFs with 200 ks depth (see Footnotes 2 and 3), being able to detect Seyfert-like sources ($L_X < 10^{44}$ erg s⁻¹) up to $z \approx 0.5$. Future more sensitive hard X-ray observations would be expected from HEX-P (PI: F. Harrison) if funded, which has been proposed to be a natural successor to NuSTAR.

• Going more. Additional sky coverages at the CDF flux levels are critically important as they substantially improve the statistical sample sizes of the faintest X-ray sources and also allow for a basic assessment of the effects of cosmic variance. It is well expected that the CDF-like surveys will become commonplace once the next-generation large X-ray observatories such as Athena and Lynx are put into operation. Specifically, Lynx will have the capability of efficiently performing even deeper surveys than the CDFs, thus effectively going deeper and wider simultaneously (see the above Going wider point).

Acknowledgements:

I thank the referee for careful reading and constructive comments that helped improve this review. I thank W.N. Brandt “who led the effort to obtain the deepest Chandra fields” (excerpted from the citation of his 2016 Bruno Rossi Prize), thank B.D. Lehmer for producing Figures 3, 5 and 6 exclusively for this paper, thank B. Luo for providing Figure 21 and thank G. Yang for making Figure 8. I thank D.M. Alexander, F.E. Bauer, W.N. Brandt, B.D. Lehmer, B. Luo, M.Y. Sun, F. Vito, and G. Yang for helpful feedback, discussions, and information sharing, as well as consistent collaboration. I also thank all my other collaborators for working on various projects together. The following funding support is gratefully acknowledged: the 973 Program (2015CB857004), the CAS Frontier Science Key Research Program (QYZDJ-SSW-SLH006), the National Natural Science Foundation of China (NSFC-11473026, 11421303), the Fundamental Research Funds for the Central Universities, the Thousand Young Talents Plan of China, and the CAS Strategic Priority Research Program “The Emergence of Cosmological Structures” (XDB09000000).

References

Abdo, A. A. et al., 2011. ApJL 733, 26.
Akiyama, M., 2005. ApJ 629, 72.
Alexander, D. M. et al., 2003. AJ 126, 539.
Alexander, D. M., Bauer, F. E., Chapman, S. C., Smail, I., Blain, A. W., Brandt, W. N., Ivison, R. J., 2005a. ApJ 632, 736.
Alexander, D. M., Smail, I., Bauer, F. E., Chapman, S. C., Blain, A. W., Brandt, W. N., Ivison, R. J., 2005b. N. Nature 434, 738.
Alexander, D. M. et al., 2008a. ApJ 687, 835.
Alexander, D. M. et al., 2008b. AJ 135, 1968.
Alexander, D. M. et al., 2011. ApJ 738, 44.
Alexander, D. M., Hickox, R. C., 2012. NewAR 56, 93.
Alexander, D. M. et al., 2013. ApJ 773, 125.
Allen, S. W., Evrard, A. E., Mantz, A. B., 2011. A&A 49, 409.
Almaini, O. et al., 2000. MNRAS 315, 325.
Alonso-Herrero, A. et al., 2008. ApJ 677, 127.
Aird, J. et al., 2010. MNRAS 401, 2531.
Aird, J. et al., 2012. ApJ 746, 90.
Aird, J. et al., 2013. ApJ 775, 41.
Aird, J. et al., 2015a. ApJ 815, 66.
Aird, J. et al., 2015b. MNRAS 451, 1892.
Aird, J., Coil, A. L., Georgakakis, A., 2017. MNRAS 465, 3390.
Akylas, A., Georgantopoulos, I., Ranalli, P., Oikonomou, E., Corral, A., Lanzuisi, G., 2016. A&A 594, 73.
Andreou, S., Maughan, B., Trinchieri, G., Kurk, J., 2013. A&A 507, 147.
Ashby, M. L. N. et al., 2013. ApJ 769, 80.
Assef, R. J. et al., 2015. ApJ 804, 27.
Azadi, M. et al., 2015. ApJ 806, 187.
Azadi, M. et al., 2017. ApJ 835, 27.
Babić, A., Miller, L., Jarvis, M. J., Turner, T. J., Alexander, D. M., Croom, S. M., 2007. A&A 474, 755.
Baldassare, V. F., Reines, A. E., Gallo, E., Greene, J. E., 2015. ApJ 809, 14.
Baldassare, V. F., Reines, A. E., Gallo, E., Greene, J. E., 2017. ApJ 836, 20.
Ballantyne, D. R. et al., 2006. ApJ 653, 1070.
Figure 21: (Left) Projected density map at $z = 1.4$ with comoving size $350 \times 450$ Mpc$^2$ from the Millennium Simulation (Springel et al. 2005), with the sizes of COSMOS and the proposed X-SERVS fields as well as their respective separations annotated. (Right) Median SED of AGNs with $L_X \approx 3 \times 10^{43}$ erg s$^{-1}$ in the E-CDF-S combined with a composite FIR starburst galaxy SED (Sajina et al. 2012) at $z = 1$. The data points and upward arrows indicate the current/scheduled X-SERVS multiwavelength coverage and their $z = 1$ limiting luminosities; and the rightmost blue arrow indicates the proposed critical X-ray coverage that is currently missing (courtesy of B. Luo).

Ballantyne, D. R., Draper, A. R., Madsen, K. K., Rigby, J. R., Treister, E., 2011. ApJ 736, 56.
Ballo, L. et al., 2007. ApJ 667, 97.
Barger, A. J. et al., 2002. AJ 124, 1839.
Barger, A. J., Cowie, L. L., Mushotzky, R. F., Yang, Y ., Wang, W. H., Steffen, A. T., Capak, P. 2005. AJ 129, 578.
Barger, A. J., Cowie, L. L., Chen, C.-C., Hasinger, G., Hsu, L.-Y ., Li, Y ., 2015. ApJ 801, 87.
Basu-Zych, A. R. et al., 2013. ApJ 762, 45.
Bauer, F. E., Alexander, D. M., Brandt, W. N., Hornschemeier, A. E., Vignali, C., Garmire, G. P., Schneider, D. P., 2002a. AJ 124, 2351.
Bauer, F. E. et al., 2002b. AJ 123, 765.
Bauer, F. E., Alexander, D. M., Brandt, W. N., Schneider, D. P., Treister, E., Hornschemeier, A. E., Garmire, G. P., 2004. AJ 128, 2048.
Bauer, F. E. et al., 2017. MNRAS 467, 4841.
Beckwith, S. V . W. et al., 2006. AJ 132, 1729.
Bell, E. F. et al., 2004. ApJ 608, 752.
Bennert, V . N., Auger, M. W., Treu, T., Woo, J.-H., Malkan, M. A., 2011. ApJ 742, 107.
Bernhard, E., Mullaney, J. R., Daddi, E., Ciesla, L., Schreiber, C., 2016. MNRAS 460, 902.
Broos, P. S. et al., 2010. ApJ 714, 1582.
Buchner, J. et al., 2014. A&A 564, 125.
Buchner, J., Bauer, F. E., 2017. MNRAS 465, 4348.
