1. Imperativeness of Improving Source Catalogs

The *Chandra* Deep Fields (CDFs) have critically contributed to the characterization of the 0.5–8 keV cosmic X-ray background (CXRB) sources. The 2 Ms CDF-North (CDF-N; Alexander at al. 2003, hereafter A03) and the 4 Ms CDF-South (CDF-S; Xue et al. 2011, hereafter X11) are the two deepest *Chandra* surveys, and the latter is complemented by the 1 Ms Extended-CDF-S (E-CDF-S, which consists of four flanking, contiguous 250 ks *Chandra* observations; e.g., Lehmer et al. 2005, hereafter L05).[1] The CDFs have enormous supporting multiwavelength investments that are key to source identification and characterization, and will remain a crucial resource in interpreting the nature of extragalactic populations identified using the superb multiwavelength surveys (e.g., *JWST*, ALMA, and EVLA) over the coming decades. *The legacy value of the CDFs is reflected in part by the* ≈ 1600 *citations to past CDF catalog papers alone* (e.g., Brandt et al. 2001; Giacconi et al. 2002; A03; L05; Luo et al. 2008; X11).

Over the last decade there have been major improvements in the methodology of producing *Chandra* source catalogs. In the recent production of the 4 Ms CDF-S point-source catalogs (X11), we have successfully implemented a number of improvements in methodology (see Table 1) compared to that used for producing earlier CDF catalogs. *Given the parallel importance of the CDF-N and E-CDF-S to the CDF-S, we propose to create improved 2 Ms CDF-N and 1 Ms E-CDF-S source catalogs implementing and further extending the methodology adopted by X11.* The E-CDF-S, though not as deep as the CDF-N and CDF-S, is also a premiere deep-survey field, and its data help significantly with measurements of sources located at large off-axis angles in the CDF-S proper.

Table 1: Improvements in the 4 Ms CDF-S Catalogs over Existing CDF Catalogs

	2 Ms CDF-N/1 Ms E-CDF-S	4 Ms CDF-S	Section
Source detection	WAVDETECT-only	WAVDETECT + ACIS Extract (AE) no-source probability	2.1
Extraction region	Circular aperture	AE polygonal regions	2.3
Crowded sources	Manual extraction	AE extraction by automatically shrinking regions	2.3
Background estimate	Source-masking approach	AE BETTER_BACKGROUNDS algorithm	2.3
X-ray photometry	Cumulative images	AE merging of extractions on individual images	2.3
Spectra and light curves	Not provided	Provided	2.4
Source identification	Not provided	Provided	3.1
Photometric redshifts	Not provided	Provided	3.2
Source classification	Not provided	Provided	3.3

In addition to adopting the improved methodology, we will also provide spectra and light curves, multiwavelength identifications, photometric redshifts, and source classifications for the ≈ 1600 sources expected in the improved catalogs (see Table 1). *The proposed work will bring the 2 Ms CDF-N and 1 Ms E-CDF-S catalogs to the same high-quality level as the 4 Ms CDF-S catalogs, thus enabling uniform and reliable field-to-field comparisons.* Compared to earlier catalogs, the improved catalogs will probe fainter and more obscured sources (i.e., of the expected total ≈ 1600 sources, $\gtrsim 220$ will be new and generally fainter and more obscured) with high confidence in their validity, due to the combination of the flexible and reliable two-stage source-detection approach (§ 2.1), the best possible X-ray photometry (§ 2.3), and the secure multiwavelength matching (§ 3.1). *As such, the improved catalogs will enable better characterization of all the* ≈ 1600 *sources, thus improving our understanding of the physics and evolution of obscured AGNs and distant galaxies.* Numerous studies can be carried out *more reliably* with the improved catalogs, e.g.: **(1)** number counts apportioned by source class across multiple fields that allow for examining cosmic variance, **(2)** deriving the best possible X-ray spectral and variability constraints upon obscured AGNs and distant galaxies, and **(3)** comparative studies between fields of resolving the 0.5–8 keV CXRB.

[1] There are alternative CDF catalogs, e.g., those produced by the Imperial College group (see Laird et al. 2009 for their methodologies) and the Yale group (Virani et al. 2005). Similar to the A03 and L05 catalogs, the production of those alternative catalogs did not use the ACIS Extract point-source analysis software and/or did not take a two-stage source-detection approach (see § 2 for details). For simplicity, here we use the A03 and L05 catalogs for discussion.

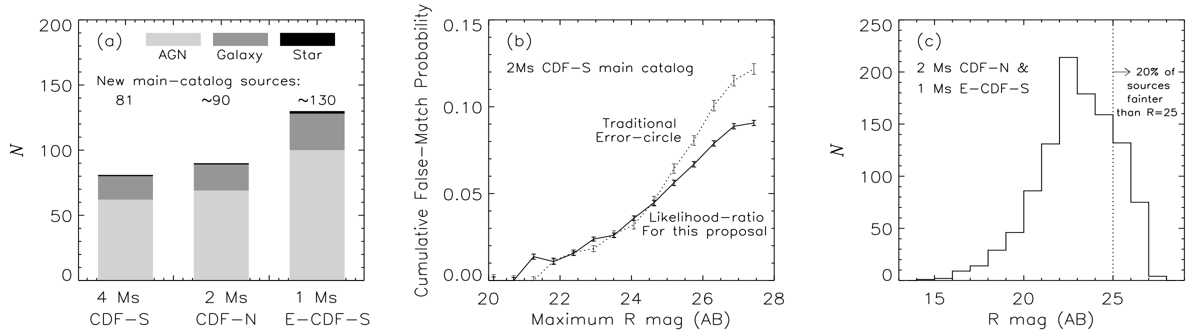

Figure 1: (a) Numbers of (expected) new main-catalog sources, broken down as AGNs (*light grey block*), galaxies (*dark gray block*), and stars (*black block*), due to the use of the WAVDETECT plus AE binomial no-source probability criterion source-detection method, for the 4 Ms CDF-S (X11), the 2 Ms CDF-N, and the 1 Ms E-CDF-S, respectively. The numbers of (expected) new sources are annotated accordingly. (b) 2 Ms CDF-S: Cumulative false-match probability as a function of the R-band limiting magnitude for the traditional error-circle matching method (*dotted curve*) and the superior likelihood-ratio matching method to be used in the proposed work (*solid curve*; Luo et al. 2010). (c) Histogram of R-band magnitude for the CDF-N and E-CDF-S sources (A03; L05). The error-circle method ($r_{\rm match}$=2″) was used to make this plot; $\approx 20\%$ of the sources have $R \geq 25$ (*dotted line*), for which the likelihood-ratio matching becomes critical.

2. Improved X-ray Source Catalogs and Data Products

In X11, we have successfully implemented improved methodology (see Table 1) by making extensive use of the ACIS Extract (AE; Broos et al. 2010) point-source analysis software that accurately computes source X-ray properties (importantly point spread function; PSF), when combining multiple observations with different roll angles and/or aim points. We will adopt and refine further this improved methodology in the production of the new CDF-N and E-CDF-S source catalogs.

2.1 New source identifications via flexible and reliable detection approach: As opposed to the WAVDETECT-only source-detection approach used in previous CDF catalog production (e.g., A03; L05), we will adopt a two-stage source-detection approach as in X11: WAVDETECT plus the AE binomial no-source probability criterion. We will first generate a candidate-list catalog of sources detected by WAVDETECT at a false-positive probability threshold of 10^{-5} or 10^{-4}. We will then obtain a more conservative *main catalog* by removing low-significance source candidates, according to the AE-computed binomial no-source probabilities (P; probabilities of sources not being real considering their local backgrounds). This approach not only produces source catalogs that are of similar reliability to those produced by running WAVDETECT at the more typical false-positive probability threshold of 10^{-6} or 10^{-7} (e.g., A03; L05), but also allows for flexibility in including additional legitimate sources that fall below the 10^{-6} or 10^{-7} threshold. In the 4 Ms CDF-S catalogs (X11), such a two-stage source-detection approach, WAVDETECT at 10^{-5} plus $P < 0.004$, maximized the number of reliable sources detected, with a "net gain" of 81 reliable sources that have secure multiwavelength counterparts, compared to the WAVDETECT-only (at 10^{-6}) approach (see Fig. 1a). The combination of the flexible and reliable two-stage source-detection approach, the best possible X-ray photometry (§ 2.3), and the secure multiwavelength identifications (§ 3.1) will probe *fainter and more obscured* sources with high confidence in their validity. For the new catalogs, we will also experiment with other combinations of WAVDETECT significance level and P value to make the optimal choice that maximizes the number of reliable sources detected. We expect to reveal $\gtrsim 220$ new sources in the new catalogs (see Fig. 1a), with a total of ≈ 1600 sources (including those in the supplementary catalogs).

We will also produce *supplementary catalogs* consisting of X-ray sources lying below the formal detection limit that correlate spatially with bright optical, IR, and radio sources. For example, 36 additional sources with $0.004 < P < 0.1$ and $R \leq 24$ made up the supplementary catalog of X11.

2.2 Improved positions and absolute astrometry: In A03 and L05, X-ray source positions were determined using the matched-filter and centroid techniques, and the absolute X-ray source positions were refined by matching the X-ray sources to the VLA 1.4 GHz radio sources (Richards 2000) and the MPG/ESO WFI R-band optical sources (Giavalisco et al. 2004), respectively. We will improve X-ray source positions by using the centroid and *improved* matched-filter techniques (Broos et al. 2010) and improve also the absolute astrometry of the new catalogs by utilizing a roughly 4 times deeper VLA 1.4 GHz radio catalog (Morrison et al. 2010) for the CDF-N and a VLA 1.4 GHz radio catalog (Miller et al. 2008, 2012) for the E-CDF-S, respectively. The better X-ray positions and astrometry will enable better measurements of X-ray photometry and better source identifications (§ 3.1).

2.3 Improved photometry: Compared to "traditional" circular-aperture photometry (e.g., A03; L05), the most important difference in the AE-computed photometry is the use of polygonal source-extraction regions. AE models the *Chandra* High Resolution Mirror Assembly and detector using the MARX ray-tracing simulator to obtain PSF models. It then constructs polygonal extraction regions that approximate specified encircled-energy fractions (EEFs) of local PSFs measured at multiple energies. Rather than adopting a typical 90% EEF for all sources (e.g., X11), we will perform simulations to determine the best EEF for a source, which likely depends on off-axis angle and/or source counts. The more accurate treatment of complex source extraction regions and more appropriate EEF choices will maximize the signal-to-noise ratio of sources. Moreover, for crowded sources in overlapping polygonal extraction regions, AE automatically utilizes smaller extraction regions that are chosen to be as large as possible without overlapping, in contrast to manual extractions that were required in previous CDF catalog production.

AE uses an algorithm for estimating backgrounds that utilizes unmasked data to model the spatial flux distributions for the point source of interest and the neighboring sources. The algorithm then computes local background counts within the background regions, subtracting contributions from the point source and its neighboring sources. This algorithm takes into account the effects of neighboring sources, CCD gaps, and exposure-time variations, thus producing accurate background extractions, which are critical for crowded sources, sources located on/near CCD gaps, and in particular *faint sources* [$\approx$ 41% ($\approx$ 44%) of the CDF-N (E-CDF-S) sources have < 50 (< 25) counts in the 0.5–8 keV band; A03; L05].

In earlier CDF catalog production (e.g., A03; L05), X-ray photometry was performed on merged images. In contrast, we will use AE to analyze individual observations independently (e.g., PSF modeling and source and background extractions) and merge the data to produce the *best possible* X-ray photometry and related products (e.g., merged X-ray spectra) for each source.

2.4 X-ray spectra and light curves: We will publicly provide merged X-ray spectra and light curves of the detected sources, which will be obtained by AE through the above sophisticated extraction of photometry, thereby enabling the *best possible* X-ray spectral and variability constraints upon the detected sources.

2.5 Summary of advantages: Compared to earlier catalogs (e.g., A03; L05), our improved source catalogs will have a number of advantages, including, **(1)** detecting $\gtrsim 220$ new and reliable sources, **(2)** providing the best possible X-ray positions and photometry for all sources, **(3)** enabling uniform and reliable field-to-field comparisons (e.g., studying cosmic variance), and **(4)** obtaining better source characterization by using the improved X-ray photometry that allows better spectral and variability constraints, leading to better understanding of the source nature. As we outline below, these improvements are in addition to improved multiwavelength counterpart matching.

2.6 Data products release plan: We will publish the improved source catalogs and associated data products (e.g., images, exposure maps, background maps, sensitivity maps, and X-ray spectra and light curves), which will be highly beneficial to the community, *thereby contributing to the most effective exploitation of the large investments in the CDF surveys.*

3. Multiwavelength Identifications, Photometric Redshifts, and Classifications

We will also provide the *most up-to-date multiwavelength identification catalog* that includes the deepest multiband data, latest redshift measurements/derivations, and most reliable source classifications, which will significantly *push forward CDF studies on obscured AGNs and distant galaxies.*

3.1 Multiwavelength identifications: We will utilize the substantial deep multi-band data and the likelihood-ratio technique to perform the best possible identification of the detected X-ray sources, noting that the multi-band data have grown significantly and the latter technique has not yet been used on earlier CDF-N and E-CDF-S catalogs. Our detailed identification analyses for the 2 Ms CDF-S (Luo et al. 2010), implementing likelihood-ratio matching across five bands (R, z, K, 3.6 μm, 1.4 GHz), have shown the power of this approach while also quantifying the significant challenges in source identification at faint magnitudes. As shown in Fig. 1b, likelihood-ratio matching significantly outperforms error-circle matching at $R \gtrsim 25$, which is critical for such *faint sources* that account for $\approx 20\%$ of all the earlier CDF-N and E-CDF-S sources (see Fig. 1c; for the improved catalogs, this fraction will increase to $\approx 30\%$). We will perform simulations to obtain the most reliable estimate of false-match rate for likelihood-ratio matching, which will be reduced significantly because most of our detected sources have true counterparts and thus have smaller susceptibilities to spurious matches (Broos et al. 2011).

3.2 Photometric/spectroscopic redshifts ($z_{\rm phot}$'s/$z_{\rm spec}$'s): We will compile all the available $z_{\rm spec}$'s, including the latest measurements, e.g., the DEEP3/GOODS-N redshifts (Cooper et al. 2011). For sources without $z_{\rm spec}$'s ($\approx 58\%$ of earlier sources), we will derive high-quality $z_{\rm phot}$'s using the superb and growing multi-band photometry (≥ 18 and ≥ 35 distinct mid-IR-to-UV bands for the CDF-N and E-CDF-S, respectively), following the procedure used for the 2 Ms CDF-S main-catalog sources, with a treatment of photometry that will include likelihood-matching, manual source deblending, and appropriate upper limits (Luo et al. 2010). Our assembled broadband photometry (from radio, IR, optical, UV to X-ray) will also be a great resource that allows for detailed SED studies and disentangling AGN/star-formation contributions to the bolometric output.

3.3 Source classifications: Following X11, we will categorize the X-ray sources as follows: **(1)** identify AGNs based on distinct AGN physical properties and optical spectroscopic information (additionally, identify low-luminosity AGNs using long-term X-ray variability; Young et al. 2012); **(2)** identify stars by cross-matching the X-ray sources with the spectroscopically identified stars and the likely stars indicated by *HST* stellarity indices or best-fit stellar templates, in conjunction with inspection of the *HST* images; and **(3)** classify the non-AGN and non-star sources as galaxies.

4. Use of Funding

To execute the proposed work, we request funds of \$58,000 for 9-month salary support, publication charges, and overheads. The 9-month support period is necessary because this proposed work requires substantial efforts including, e.g., **(1)** performing numerous simulations to determine optimal EEFs and experimenting with various combinations of WAVDETECT significance level and P value, **(2)** working with > 10 multiwavelength data sets (e.g., for likelihood-ratio matching, manual source deblending, and deriving high-quality $z_{\rm phot}$'s), and **(3)** writing a catalog paper with lasting value. The paper will provide high-quality uniform X-ray catalogs and multiwavelength matching for the CDFs, which is likely to be highly cited based on previous experience (e.g., the A03 CDF-N catalog paper alone has > 500 citations). Given our $\approx$ 6-month experience/effort cataloging the 4 Ms CDF-S (X11), we expect the proposed work, which involves a total solid angle that is ≈ 3 times that of the 4 Ms CDF-S field, to be finished within 9 months.

• Alexander et al. 03, AJ, 126, 539 (A03) • Brandt et al. 01, AJ, 122, 2810 • Broos et al. 10, ApJ, 714, 1582 • Broos et al. 11, ApJS, 194, 2 • Cooper et al. 11, ApJS, 193, 14 • Giacconi et al. 02, ApJS, 139, 369 • Giavalisco et al. 04, ApJL, 600, L93 • Laird et al. 09, ApJS, 180, 102 • Lehmer et al. 05, ApJS, 161, 21 (L05) • Luo et al. 08, ApJS, 179, 19 • Luo et al. 10, ApJS, 187, 560 • Miller et al. 08, ApJS, 179, 114 • Miller et al. 12, in prep. • Morrison et al. 10, ApJS, 188, 178 • Richards 00, ApJ, 533, 611 • Virani et al. 06, AJ, 131, 2373 • Xue et al. 11, ApJS, 195, 10 (X11) • Young et al. 12, ApJ (arXiv:1201.4391)