

# The X-SERVS-LSS Survey : XMM source catalog

C.-T. J. Chen<sup>1</sup><sup>★</sup>, W. N. Brandt, B. Luo, P. Ranalli, and Friends

<sup>1</sup>*PSU*

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

We present source catalogs from the *XMM*-Spitzer Extragalactic Representative Volume Survey in the *XMM*-Large Scale Structure survey region (X-SERVS-LSS). The survey is comprised of 1.1 Ms new *XMM* AO15 data and 1.6 Ms *XMM* archival data, totaling 2.7 Ms flare-filtered exposure over a 5.3 deg<sup>2</sup> area. We provide X-ray catalogs of 5218 sources detected in the full (0.5–10 keV), soft (0.5–2 keV), and hard (2–10 keV) bands. Sources are first detected using *XMM* SAS task EWAVELET with THRESHOLD>4 and further verified by SAS task EMLDETECT with DET\_ML > 6, which is equivalent to a false-detection probability of  $P = 0.0025$ . A total of 2375 new X-ray sources are detected compared to the *XMM*-LSS-North survey in the same area. Our flux limits and distributions are comparable to those of the *XMM*-COSMOS survey. The median fluxes at the full, soft, and hard bands (in erg cm<sup>-2</sup> s<sup>-1</sup>) are  $8.78 \times 10^{-15}$ ,  $2.65 \times 10^{-15}$ , and  $1.48 \times 10^{-14}$ , respectively. As confirmed using available *Chandra* sources over  $\approx 5\%$  of XMM-LSS, we identified multiwavelength counterparts that are 97% reliable in the central 4.5 deg<sup>2</sup> region covered by *Spitzer* SERVS. For the full 5.3 deg<sup>2</sup> region the reliability is > 94%. Results of multiwavelength identifications are also included in the source catalog, along with basic optical-to-infrared photometry as well as spectroscopic redshifts from the publicly available surveys.

**Key words:** keyword1 – keyword2 – keyword3

## 1 INTRODUCTION (TO BE COMPLETED, PLEASE SKIP THIS SECTION)

Extragalactic X-ray surveys have played an integral role in studying the cosmic evolution of galaxies and its relation to the accreting supermassive black holes (SMBHs) in galactic centers (the active galactic nuclei, or AGNs). For the past two decades, the design of cosmic X-ray surveys has followed the “wedding cake” strategy. Namely, some surveys have ultra-deep X-ray coverages in a pencil-beam survey area while other surveys have shallower X-ray coverage but wider survey area. The wealth of data from cosmic X-ray surveys and ancillary multiwavelength surveys have provided the primary source of information that shaped our understanding of how galaxies and supermassive black holes evolve through cosmic time, where deep surveys sample high-redshift moderate-to-luminous AGNs and wide-fide surveys probe the more luminous, rare objects that are missed by surveys with smaller volumes. However, limited by the survey designs, narrow-field surveys lacks the contiguous volume to include a wide range of large-scale structures, and the wide-field surveys lack the survey sensitivity

to track the bulk of AGN population at the cosmic epoch during which the AGN activity and galaxy star formation peak (see [Brandt & Alexander 2015](#), for a recent review).

Among the cosmic X-ray surveys, the medium-depth COSMOS survey has provided indispensable data that are both sufficiently large and deep enough to allow for the first steps of understanding the connections between AGN and galaxy in different cosmic environments ([some citations here](#)). However, COSMOS alone cannot sample a full range of cosmic environments. For instance, the largest structures found in cold dark matter simulations are already as large as the angular size of COSMOS at  $z \approx 1$  (e.g., 80–100 Mpc in comoving size, which covers 2–3 deg<sup>2</sup>, see [Klypin et al. 2016](#)). Clustering analyses also demonstrate that COSMOS-sized fields are still subject to significant cosmic variance (e.g., [Meneux et al. 2009](#); [de la Torre et al. 2010](#); [Skibba et al. 2014](#)). In this work, we present an X-ray source catalog obtained with *XMM*-Newton in the X-SERVS-LSS survey region, which is the first part of the X-SERVS survey which aims to expand the parameter space of current X-ray surveys with multiple > 3 deg<sup>2</sup> surveys that reach COSMOS-like survey depth. This will enable studies of SMBH growth across the full range of cosmic environments and minimizing cosmic variance.

\* E-mail: ctchen@psu.edu (KTS)

The X-SERVS-LSS field is among the most extensively observed sky regions. We also present multiwavelength identifications, basic photometric properties, and spectroscopic redshifts obtained from the literature. These data, combined with the existing data from COSMOS, will enable a wide range of science on AGN evolution, including studies on the relation between large-scale structure and AGN triggering/feedback that can account for the cosmic variance.

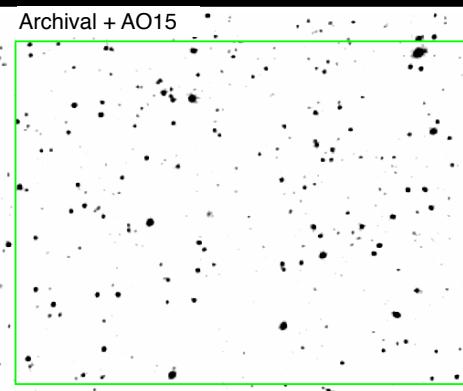
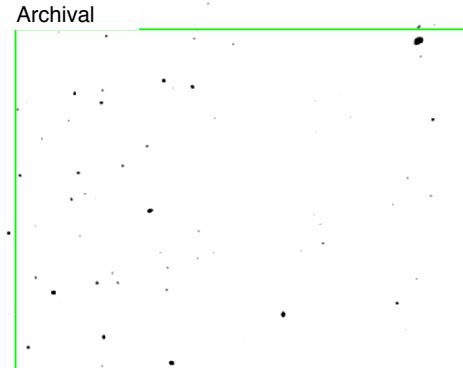
This paper is organized as follows: in Sec. 2 we present the details of the archival and new observations, and the procedures for data reduction. In Sec. 3 we describe the X-ray source searching strategies and the details of the production of the X-ray point-source catalog. Sec. 4 outlines the reliability assessment of the X-ray catalog using simulated X-ray observations. The survey sensitivity and the number counts are also presented here. In Sec. 5, we describe the multiwavelength identification methods and reliability assessments, as well as basic multiwavelength properties. In Sec. 6 we list the columns included in the multiwavelength catalog. A summary is given in Sec. 7. Throughout the paper, we assume a  $\Lambda$ CDM cosmology with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.3$ , and  $\Omega_\Lambda = 0.7$ , and AB magnitudes unless noted otherwise.

## 2 XMM OBSERVATIONS IN THE X-SERVS-LSS REGION AND DATA REDUCTION

### 2.1 XMM-Newton data in the XMM-LSS region

The XSERVS-LSS data are primarily comprised of the 1.3 Ms AO-15 *XMM-Newton* observations taken between July 2016 to February 2017. Our AO-15 data include 67 *XMM* observations covering the majority of the *XMM*-LSS survey region. For this work, we make use of all available *XMM-Newton* observations that overlap with the 1.3 Ms observations and the adjacent Subaru *XMM* Deep Survey (SXDS, Ueda et al. 2008) to create a deep *XMM* catalog contiguously covering most of the *Spitzer* SERVS-LSS field for its superior multiwavelength coverage (Mauduit et al. 2012). After excluding observations that were completely lost to flaring backgrounds (see Sec. 2.2), the archival data used here include the 51 observations culled from the *XMM*-LSS survey typically with  $\sim 10$  ks exposure time per observation, 18 observations from the *XMM-Newton* Medium Deep Survey ((XMDS, 20–25 ks exposure depth, Chiappetti et al. 2005), 4 mosaic-mode observations<sup>1</sup> as part of the *XMM*-XXL survey (Pierre et al. 2016), 4 archival *XMM* observations targeting galaxy clusters identified in the XXL and LSS surveys ( $\sim 30$ –100 ks), and the 10 observations of the SXDS ( $\sim 50$  ks exposure per observation, Ueda et al. 2008). We present the details of each observation in Table 1 (the full table is available separately).

Our 1.3 Ms observations were separated into two epochs to minimize the effects of background flaring. We first observed the SERVS-LSS sky-region with  $\sim 1$  Ms *XMM* exposure time during July–August 2016. We then screened the first-batch observations for flaring backgrounds (Sec. 2.2)



**Figure 1.** Top – Background-subtracted, smoothed image at the 0.5–10 keV band for a  $0.8 \times 0.6 \text{ deg}^2$  region centered at  $\text{RA}=35.58^\circ$ ,  $\text{DEC}=-4.965^\circ$ . This image is created using only the archival data. Bottom – Same as the top image, but all archival data and the new AO15 data are included.

and re-visit the background contaminated sky-regions using the remaining 0.3 Ms. Notably, we also observed the SXDS region in which one of the SXDS observations carried out in 2002 was severely affected by background flares. In this work, we present an X-ray catalog obtained from a  $5.3 \text{ deg}^2$  sky-region with  $34.2^\circ < \text{RA} < 37.125^\circ$  and  $-5.72 < \text{Dec} < -3.87$ . The sky-region is primarily selected by the footprint of our AO15 observations, with additional SXDS data within the *Spitzer* SERVS footprint in the corner. A total of  $\sim 3$  Ms *XMM-Newton* observations is used for generating the X-ray catalog. In Fig. 1, we show the background-subtracted, 0.5–10 keV image (see Sec. 3 for the details of data analysis) from a  $\sim 0.5 \text{ deg}^2$  region in X-SERVS-LSS generated using the AO15 + archival data. An image generated using only the archival data is also shown for comparison.

### 2.2 Data Preparation and Background flare filtering

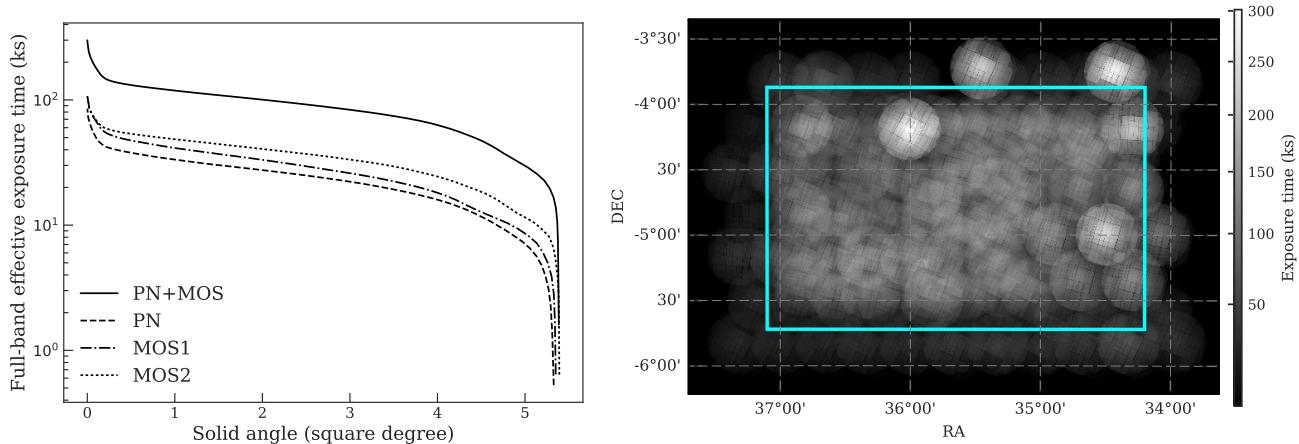
We use the *XMM* Science Analysis System (SAS) 16.1.0 and heasoft 6.21 for our data analysis. The *XMM* Observation Data Files (ODFs) are processed with SAS tasks EPICPROC (EPPROC and EMPROC for PN and MOS, respectively) to create MOS1, MOS2, PN, and PN out-of-time (OOT) event files for each ObsID. For the mosaic-mode observations and observations with unexpected interruptions due to strong background flares, we use SAS task EMOSAIC\_PREP to sep-

<sup>1</sup> Each mosaic-mode observation is comprised of  $\sim$  ten 10 ks exposures.

**Table 1.** *XMM-Newton* Observations

Field	Revolution	ObsID	Date (UT)	R.A.	Decl.	GTI (PN) (ks)	GTI (MOS1) (ks)	GTI (MOS2) (ks)
X-SERVS-LSS (AO-15)	3054	0780450101	2016-08-13T01:34:06	35.81072	-5.15989	20.91	23.61	23.61
<i>XMM</i> -LSS	1205	0404965101	2006-07-09T08:08:08	35.80953	-5.48532	3.44	10.36	9.91
XMDS	287	0111110401	2001-07-03T14:01:54	35.97582	-5.15253	21.40	27.20	27.40
SXDS	118	0112370101	2000-07-31T21:57:54	34.47819	-4.98115	39.13	42.70	42.83
<i>XMM</i> -XXL-N	2137	0677580101	2011-08-10T01:53:35	38.00217	-4.49993	4.94	5.93	5.52
XLSSJ022404.0-041328	0928	0210490101	2005-01-01T19:08:30	36.03267	-4.20230	80.28	87.98	87.98

The X-SERVS-LSS catalog presented in this work consists of 153 observations with a total of  $\sim 2.7$  Ms background-screened exposure time (MOS1), of which  $\sim 1.1$  Ms is from the new AO15 observations. The total PN+MOS GTI is  $\sim 3$  Ms. Columns from left to right: target field, *XMM* revolution, *XMM* ObsID, observation starting time, Right Ascension and Declination (J2000), background-screened good-time-interval (GTI) exposure time for PN MOS1, and MOS2.



**Figure 2.** *Left* – Cumulative survey solid angle as a function of full-band effective exposure for the X-SERVS-LSS survey. *Right* – Full-band survey exposure map (PN + MOS), the survey region from which the X-ray source catalog is constructed are marked as the rectangular box. Except for several regions with ultra-deep *XMM* follow-up observations of galaxy clusters, the *XMM* coverage in our X-SERVS-LSS region is generally homogeneous.

arate the event files into individual pseudo-exposures and assign pseudo exposure IDs<sup>2</sup>. For the mosaic mode observations, we also determine the RA/DEC of each pseudo-exposure using the AHFRA and AHFDEC values in the attitude files created using the SAS task ATTHKGEND.

For each event file, we create single event (PATTERN = 0) light curves in time bins of 100s for high (10–12 keV) and low energies (0.3–10 keV) using EVSELECT to search for events without significant background flares (the “good time intervals”, GTIs). We first remove time intervals with 10–12 keV count rates exceeding  $3\sigma$  above the mean, and then repeat the  $3\sigma$  clipping procedure for the low energy light curves. Since the background flares usually manifest as the high-count-rate tail in addition to the Gaussian-shape count-rate histogram, adopting the  $3\sigma$  clipping rule can effectively remove the high-count-rate tail while retaining useful scientific data. For a small number of event files with intense

<sup>2</sup> For mosaic-mode observations, multiple exposures are stored in a single event file for each EPIC camera after ODF processing. EMOSAIC\_PREP separates the single event file into multiple pseudo event files with the same ObsID but different EXP\_ID and EXPIDSTR, which are informative fits file keywords required by SAS tasks. For observations with unexpected interruptions, the exposure is separated into multiple event files with the same ObsID and EXP\_ID by EPICPROC, but with different EXPIDSTRs. EMOSAIC\_PREP assigns a separate EXP\_ID for each event file, which is necessary for SAS source detection tasks.

background flares, we instead filter the event files using the nominal count rate thresholds suggested by the *XMM* Science Operations Centre webpage<sup>3</sup>. Observations with GTI  $< 2$  ks are discarded. A total of 2.7 Ms exposure remains after flare-filtering, including the 1.1 Ms from AO 15 and 1.6 Ms from archival data. After screening for background flares, we further exclude events in energy ranges that overlap with the instrumental background lines (Al K $\alpha$  lines at 1.45–1.54 keV for MOS and PN; Cu lines at 7.2–7.6 keV and 7.8–8.2 keV for PN).

From the background-screened, instrumental-line-removed event files, we extract images with a 4'' pixel size using EVSELECT<sup>4</sup> in the following bands: 0.5–2 keV and 2–10 keV (soft and hard bands), and 0.5–10 keV (full band). For each image, we generate exposure maps with and without vignetting-corrections using SAS task EEXPMAP. We set USEFASTPIXELIZATION=0 and ATTREBIN=0.5 in order to obtain more accurate exposure maps. Detector masks are also generated using SAS task EMASK. We show the distribution of exposure values across the X-SERVS-LSS field and the PN+MOS exposure map in Fig. 2

<sup>3</sup> <https://www.cosmos.esa.int/web/xmm-newton/sas-thread-epic-filterbackground>

<sup>4</sup> The 4'' pixel size is set by the parameter IMAGEBINSIZE= 80 in the SAS task EVSELECT.

### 3 THE MAIN X-SERVS-LSS X-RAY CATALOG

#### 3.1 First-pass source detection and astrometric correction

The astrometric accuracy of *XMM* observations could be affected by the pointing uncertainties of the *XMM* telescope. This uncertainty is usually smaller than a few arcsecs, but could be as large as  $\sim 10''$  (e.g., Cappelluti et al. 2007; Watson et al. 2008; Rosen et al. 2016). To achieve better astrometric accuracy and to avoid any systematic offsets between different *XMM* observations, we run an initial pass of source detection for each observation and then use the first-pass source list to register the *XMM* observations onto a common WCS frame. The first-pass source detection methods are outlined below:

- (i) For the exposure taken by each instrument in each observation, we generate a temporary source list using the SAS task Ewavelet with a low likelihood threshold (THRESHOLD=4). Ewavelet is a wavelet-based algorithm running on the count-rate image generated using the image and exposure map extracted as described in Sec. 2.2.
- (ii) We use the temporary source list as an input to generate background images using the SAS task Esplinemap with METHOD=MODEL. Esplinemap fits the source-excised image with two templates: the vignetted exposure map, and the un-vignetted exposure map. The former represents the cosmic X-ray background with an astrophysical origin, while the latter represents the intrinsic instrumental noise. Esplinemap then finds the best-fit linear combination of the two templates and generate the background map. The detail of this method is described in Cappelluti et al. (2007). The background maps are used for the PSF-fitting based source detection task described in Step (iv).
- (iii) We run Ewavelet again for each observation. This time the source list is generated by running Ewavelet on the exposure map and image coadded among the PN, MOS1, and MOS2 exposures (when available) with the default likelihood threshold (THRESHOLD=5).
- (iv) For each Ewavelet source list, we use the SAS task Emldetect to re-assess the detection likelihood and determine the best-fit X-ray positions. Emldetect is a PSF-fitting software, which performs maximum likelihood fits to the input source considering the *XMM* PSF, exposure values, and background levels of the input source on each image. Emldetect also convolves the PSF with a  $\beta$ -model brightness profile for clusters and use the result to determine if the input source is extended. We use a stringent likelihood threshold (LIKMIN= 10.8) to ensure that astrometric corrections are calculated based on real detections.
- (v) For the mosaic-mode observations, the multiple pointings under the same ObsID were already registered on the same WCS frame. Therefore we do not correct for astrometry for each pseudo-exposure but only consider the astrometric offsets on an ObsID-by-ObsID basis. The source lists for the mosaic-mode observations were generated using the SAS task Emosaic\_Proc, which is a mosaic-mode wrapper for procedures similar to (i)-(iv) described above.

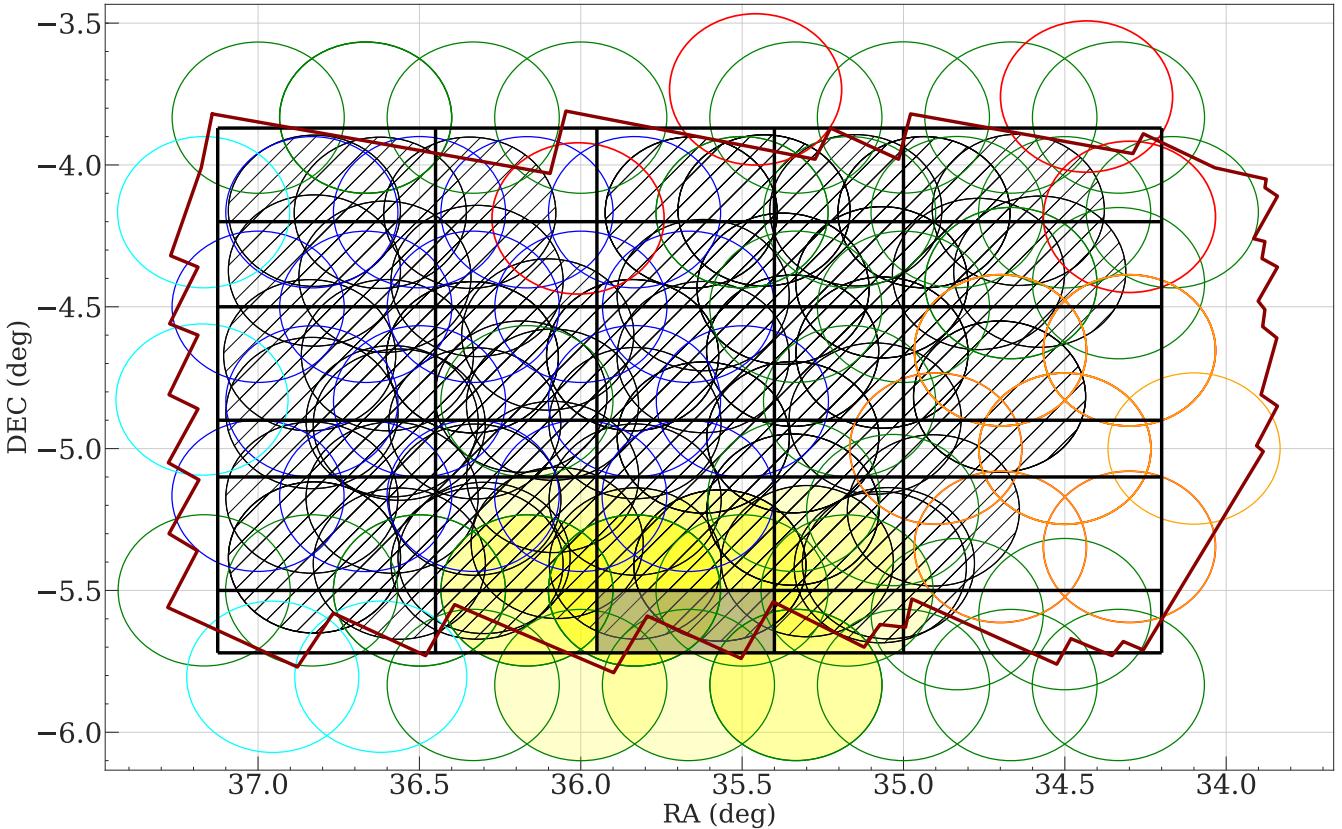
For steps (iv) and (v), the source-searching was conducted simultaneously on the images of the three EPIC-

cameras as the astrometric offsets between PN, MOS1, and MOS2 are negligible. For each ObsID, we cross-correlate the high-confidence EMLDETECT list of point sources (with the EMLDETECT flag EXT= 0) with the optical source catalog culled from the Subaru Hyper Suprime Cam Survey Public Data Release 1 (HSC-PDR1 Aihara et al. 2017), which is an ultra-deep optical photometric catalog with sub-arcsec angular resolution. The astrometry of HSC-PDR1 is calibrated to the PAN-STARRS1  $3\pi$  survey and has a  $< 0.05''$  astrometric uncertainty. More details of the HSC-PDR1 catalog can be found in Aihara et al. (2017), and is also briefly discussed in Sec. 5. For astrometric corrections, we limit the optical catalog to HSC sources with  $23 < i < 18$  to exclude possible spurious matches and matches to bright stars that might have parallaxes. The offset between each ObsID and the HSC catalog is calculated based on the maximum likelihood algorithm similar to the SAS task Eposcorr. The major difference between our approach and Eposcorr is that we use an iterative optimization approach compared to the grid-searching algorithm adopted by Eposcorr. During each iteration, we cross-correlate the optical catalog with the X-ray catalog using a  $3''$  search radius. We exclude all duplicated matches (typically only  $< 5\%$  of the X-ray sources have more than one optical counterparts in the bright HSC-PDR1 catalog). For the purpose of frame-correction, we adopt the X-ray positional uncertainties calculated based on the PSF-fitting likelihood ratios provided by EMLDETECT ( $\sigma_{eml}$  hereafter). To avoid over-weighing sources with extremely small  $\sigma_{eml}$ , we also include an irreducible systematic uncertainty when finding the best-fit values for frame-correction<sup>5</sup>. For all ObsIDs, the median number of X-ray sources with only one HSC counterpart within  $3''$  is 32. The required frame-correction offsets calculated using our approach are less than  $3''$  in both RA and DEC directions and are generally consistent with the results calculated using Eposcorr, with a median difference of  $0.1''$ . For several obsIDs the difference between our offsets and the Eposcorr offsets are indeed non-negligible ( $> 0.5''$ ). We visually inspect these ObsIDs and find that our approach does improve the alignments between the optical and corrected X-ray images. The event files and the attitude file for each ObsID are then projected onto the common WCS frame of the HSC catalog by updating the relevant keywords using a modified version of AlignEvt (Ranalli et al. 2013). Since the sky coordinates for the event files of the mosaic-mode pseudo-pointings are derived based on the reference point centered at the nominal RA and DEC positions of the mosaic-mode ObsIDs, we also recalculate the sky coordinates for these event files with the SAS task Attcalc using the true pointing positions as the reference point, which is necessary for using regular SAS tasks for mosaic-mode pseudo-exposures.

#### 3.2 Second-pass Source Detection

We re-create images, exposure maps, detector masks, background maps using the frame-corrected event files and attitude files. We then run source detection tasks for the second time considering all *XMM* observations listed in Table 1.

<sup>5</sup> We assume the systematic uncertainties to be  $0.5''$  as suggested by Watson et al. (2008)



**Figure 3.** Locations of the *XMM* observations used in this work. The AO-15 observations are marked as the hatched black circles. The green, orange, blue, and cyan circles are for *XMM*-LSS, SXDS, XMDS, and *XMM*-XXL observations, respectively. We also show the detection grids used for running EMLDETECT. The gray, shaded box demonstrates one of the grids. All of the observations with non-zero exposure time in the grid (the yellow-filled circles) are considered for source detection, and only the sources with coordinates inside the grid are kept. The final source catalog is the union of the source lists from all grids. We also plot the RA/DEC range of our catalog selection area as the solid rectangle, and the *Spitzer* SERVS footprint is marked as the red polygon.

Similar to the approach used by the *XMM*-H-ATLAS survey (Ranalli et al. 2015), we divide the X-SERVS-LSS field into multiple grids when running the second-pass source detection (see Fig. 3). We then use a custom-built wrapper of relevant *XMM*-SAS tasks to carry out the second-pass source detection, which is similar to the GRIDDETECT<sup>6</sup> tool built for the *XMM*-H-ATLAS survey (Ranalli et al. 2015). For each grid, we co-add the images and exposure maps for all observations with footprint inside the grid and run EWAVELET with a low detection threshold<sup>7</sup> on the co-added image and exposure map. For each grid, we only keep EWAVELET sources within the RA/DEC range of the grid (with a 1' padding on each side of the grid). The grid sizes are determined by the number of EWAVELET sources and the number of ObsIDs encompassed because the number of images that could be processed by a single EMLDETECT thread is limited. We then use the EWAVELET list as an input for EMLDETECT to validate the detection likelihood. Instead of running on the co-added image, EMLDETECT takes the image, exposure map, background map, detector mask of each input

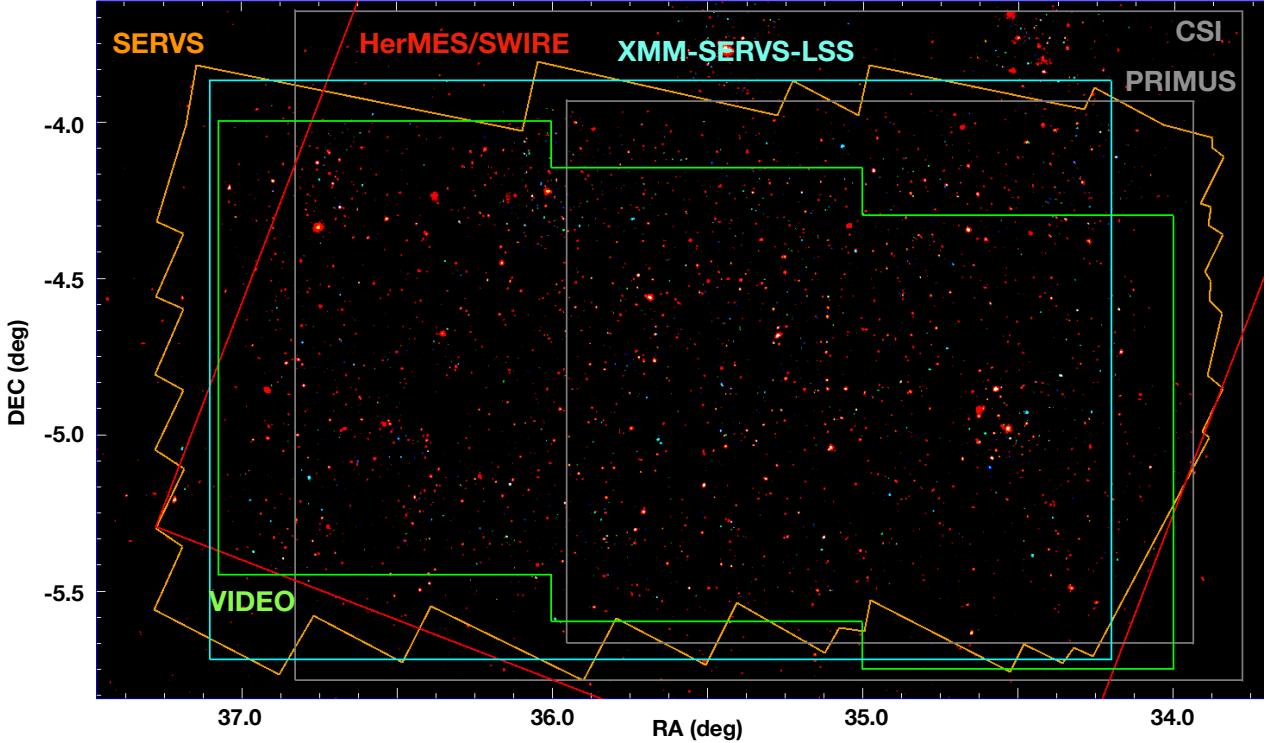
observations into account. We consider sources with EMLDETECT likelihood larger than DET\_ML=6.0 to be reliably detected. For readers interest in an even more reliable catalog, we suggest using the DET\_ML=10.8 detection threshold for the final X-ray catalog. The detection thresholds are determined by extensive simulations that are discussed in Sec. 4. The EMLDETECT list of the full X-SERVS-LSS region is constructed from the union of the sources from all grids. We search for sources in three different bands: 0.5–2 keV (soft), 2–10 keV (hard), and 0.5–10 keV (full). We show the 3-color X-ray image of the full X-SERVS-LSS survey region in Fig. 4.

### 3.3 Astrometric accuracy

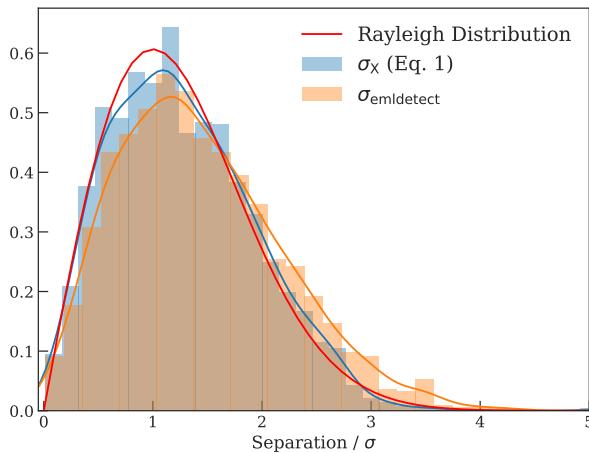
We investigate the positional accuracy of the *XMM* sources by comparing the second-pass X-ray catalog with the HSC-PDR1 catalog. Similar to the frame-correction procedures described in Sec. 3.1, we search for unique optical counterparts around the X-ray positions using a 3'' search radius. For the 5136 X-ray sources detected at the full-band during the second-pass source-searching process, a total of 2085 X-ray sources are found to have only one  $i < 23$  HSC counterpart within 3''. We use the separation between the optical

<sup>6</sup> <https://github.com/piero-ranalli/griddetect>.

<sup>7</sup> THRESHOLD=4.



**Figure 4.** Three-color (red: 0.5–2 keV, green: 2.0–4.5 keV, blue: 4.5–10 keV) image of the full X-SERVS-LSS field. Coverages of a selection of ancillary multiwavelength data are also shown. The cyan box represents the region from which the X-ray source catalog is constructed.



**Figure 5.** Histograms of the “normalized positional uncertainties” and comparison with the expected Rayleigh distribution, the solid red curve. The distribution for the empirically derived positional uncertainty,  $\sigma_x$ , for the sources matched to bright optical counterparts is marked as the blue histograms along with the kernel density estimation. The positional uncertainties calculated by EMILDETECT,  $\sigma_{\text{emil}}$ , are also shown.

and X-ray positions of this subsample as a means to determine an empirical X-ray positional uncertainties, which is a commonly adopted practice in deep X-ray surveys (e.g. Luo et al. 2010; Xue et al. 2011; Luo et al. 2016).

The X-ray positional accuracy is determined by how well can the PSF-centroid location be measured, which usually

depends on the number of counts of the detected source and the PSF size of the instrument. For the vast majority of the X-ray sources presented in this work, the detected photons come from at least three different observations, and thus the positional uncertainty dependence to the PSF size (i.e. off-axis angles) is weak. On the other hand, the X-ray positional uncertainty is strongly dependent on the number of counts of the detected sources. Using the angular separation between the 2085 X-ray sources and their unique optical counterparts, we derive an empirical relation between the number of X-ray counts,  $C^8$ , and the 68% real-position radius ( $r_{68\%}$ ) for the full-band-detected X-ray sources,  $\log_{10} r_{68\%} = -0.31 \times \log_{10} C + 0.85$ . The parameters are chosen such that 68% of the sources have positional offsets smaller than the empirical relation.

The X-ray positional uncertainty,  $\sigma_x$ , is  $r_{68\%}$  divided by a factor of 1.51517. Because the separations in both RA and DEC directions behave as a univariate normal distribution with  $\sigma_{\text{RA}}$  and  $\sigma_{\text{DEC}}$ , respectively<sup>9</sup>, the angular separation should therefore follow the joint probability distribution function of the uncertainties in the RA and DEC directions. Considering the case where  $\sigma_{\text{RA}} = \sigma_{\text{DEC}}$ , the angular separation between an optical source and an X-ray source should follow

<sup>8</sup> An upper limit of 2000 is set on  $C$  because the improvement of positional accuracy is not significant for larger source counts (e.g., Luo et al. 2016).

<sup>9</sup> Here we consider the positional uncertainties of the HSC-PDR1 catalog to be negligible compared to the *XMM* positional uncertainties.

the univariate Rayleigh distribution with the scaling parameter,  $\sigma_x$ , where  $\sigma_x = \sigma_{RA} = \sigma_{DEC}$  (see Sec. 4 of Pineau et al. 2017, for details). Since  $\sigma_x$  is not the standard deviation of a normal distribution, the 68% uncertainty of the angular separation is represented as  $r_{68\%} = 1.51517\sigma_x$ .<sup>10</sup>

For each energy band, we repeat the same process to find the best-fit relation for  $\sigma_x$  using the following equation:

$$\log_{10} \sigma_x = \alpha \times \log_{10} C + \beta. \quad (1)$$

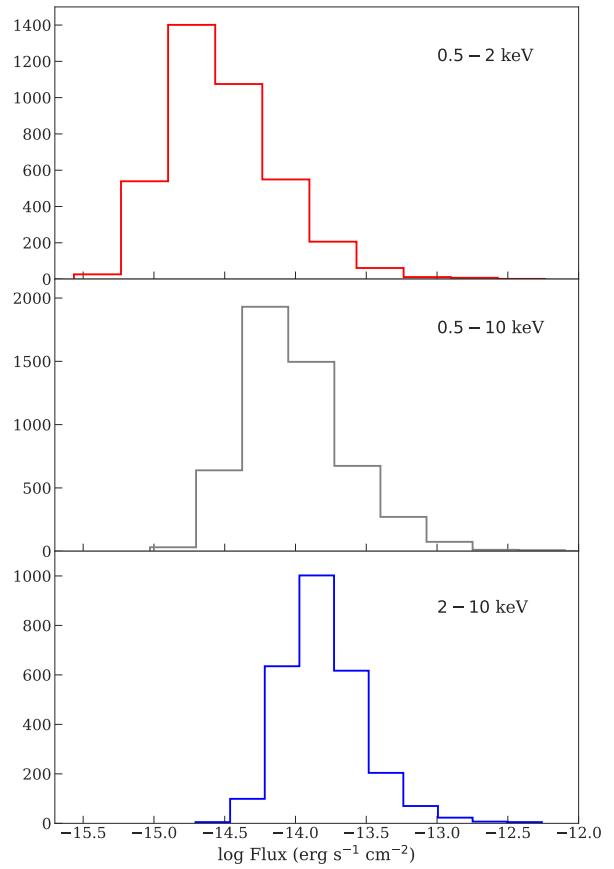
Due to the large PSF size of *XMM*, it is possible that the positional uncertainties were underestimated when deriving Equation 1 using only the optical sources within the 3'' search radius. Therefore, we adopt an iterative process. The parameters for Equation 1 are updated using the reliable matches found with the likelihood-ratio matching method described in the Sec. 5.1, and the updated astrometric uncertainties are used for running likelihood-ratio matching again. This is a stable process, as the parameters converge after two to three iterations. The average positional uncertainties ( $\sigma_x$ ) for our full-band, soft-band, and hard-band X-ray catalogs are 1.31'', 1.35'', and 1.37'', respectively. In Fig. 5, we show the comparison of the normalized separation (Separation/ $\sigma_x$ ) between the full-band X-ray sources and their bright optical counterparts with  $\sigma$  derived using Eq. 1,  $\sigma_x$ , and the original positional uncertainties calculated by EMLDETECT,  $\sigma_{eml}$ . The agreement between the Rayleigh distribution and the Separation/ $\sigma_x$  distribution of our sample demonstrates that our empirically derived  $\sigma_x$  are reliable indicators of the true positional uncertainties. As for  $\sigma_{eml}$ , previous studies have identified that an irreducible systematic uncertainties should be added to  $\sigma_{eml}$  for the normalized separation to follow Rayleigh distribution (e.g., Watson et al. 2008), but the nature of this systematic uncertainties remains unclear. We use  $\sigma_x$  as the positional uncertainties of our X-ray catalog, but  $\sigma_{eml}$  is also included in the final catalog for completeness.

### 3.4 The main X-ray source catalog

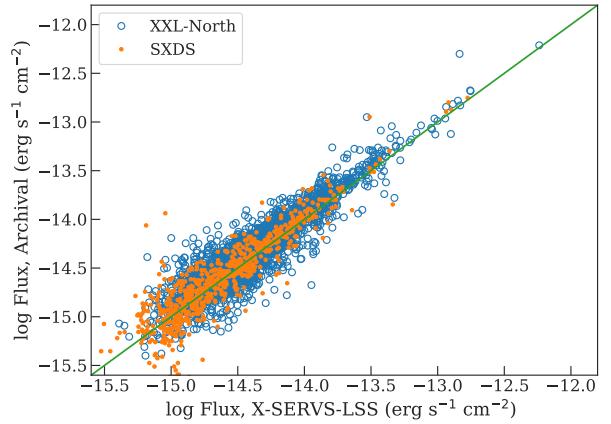
At the 0.5–10 keV, 0.5–2 keV, and 2–10 keV bands, we find 5136, 3878, 2655 point sources with DET\_ML > 6.0, respectively. We report the details of the main X-ray source catalog in Table A. We show the flux distributions of the sources detected in the three energy bands in Fig. 6. The extended sources (identified by the EXT > 0 flag of EMLDETECT) are not included as the properties of the extended X-ray emission are beyond the scope of this work.<sup>11</sup> We associate catalogs from the three energy bands using the similar approach adopted by the *XMM* Serendipitous Source Catalogue. If two sources from different catalogs have their angular separation smaller than the minimum of the following quantities, we consider them to be the same source: 1. 10'' 2. distance to the nearest-neighbor in each catalog 3. 99% positional uncertainty. The final source catalog is the union of the sources

<sup>10</sup> The factor 1.51517 is determined by integrating the Rayleigh distribution until the cumulative probability reaches 0.68. For reference, 90%, 95%, and 99.9% uncertainties corresponds to  $2.3\sigma_x$ ,  $2.7\sigma_x$ ,  $4.3\sigma_x$ , respectively.

<sup>11</sup> There are 68, 77, 11 sources identified as EXT = 0 by EMLDETECT at the 0.5–10 keV, 0.5–2 keV, and 2–10 keV bands, respectively.



**Figure 6.** Flux distributions of the sources detected at soft (top panel), full (middle panel), and hard bands.



**Figure 7.** Comparison of the soft-band X-ray fluxes of our X-ray sources and those of the *XMM* counterparts identified in the archival source catalogs within a 10'' radius, including XXL-N (blue open circles) and SXDS (orange dots).

detected at the three energy bands. We check for potential duplicated sources by visually inspecting all sources with distance to the nearest-neighbor (DIST\_NN) less than 10'', and only one set of sources are found to be duplicated, resulting in a total of 5218 unique sources. A unique X-ray source ID is assigned to each of the 5218 sources at this stage. We also derive the count-rate to flux energy conversion fac-

tors (ECFs) assuming a power-law spectrum with photon index  $\Gamma = 1.7$  and Galactic absorption<sup>12</sup>,  $N_{\text{H}} = 3.57 \times 10^{20} \text{ cm}^{-2}$ . The conversion factors for PN and MOS are (7.06, 1.99), (1.24, 0.44), and (3.17, 0.96) counts  $\text{s}^{-1} / 10^{-11} \text{ erg cm}^{-2} \text{ s}^{-1}$ , at the 0.5–2 keV, 2–10 keV, and 0.5–10 keV, respectively. For each source detected by EMLDETECT, flux from each EPIC camera is calculated separately using the corresponding ECF. The final flux of the source is the error-weighted mean of the fluxes from the three EPIC cameras, when available.

For sources that are not detected in all three bands, we calculate the source count upper limits using the mosaicked background map of the band at which the source is not detected. The mosaicked background map of each band is generated by summing up the background maps from all individual observations with bad pixels and chip gaps filled using the median values of the surrounding  $5 \times 5$  pixels. According the Poisson probability set by the EMLDETECT detection likelihood threshold ( $P_{\text{Random}}$ , the probability of the detected source to be a random Poisson fluctuation due the background), we can calculate the the minimum required source counts ( $m$  in the following equation) to exceed the random background fluctuation,  $B$ , using the regularized upper incomplete  $\Gamma$  function:

$$P_{\text{Random}} = \frac{1}{\Gamma(m)} \int_B^{\infty} t^{m-1} e^{-t} dt \quad (2)$$

Note that  $\Gamma(m)$  is the  $\Gamma$  function, which is equivalent to  $(m-1)!$ . The upper limits are set at  $P_{\text{Random}} = 2 \times 10^{-5}$  (consistent with the DET\_ML= 10.8). For each non-detected source in each band, we determine the background count by summing up the background map within the circle with 70% energy encircled fraction (EEF). We then calculate  $m$  using the SCIPY function SCIPY.SPECIAL.GAMMAINCCINV<sup>13</sup>. The flux upper limit is calculated based on the following equation:

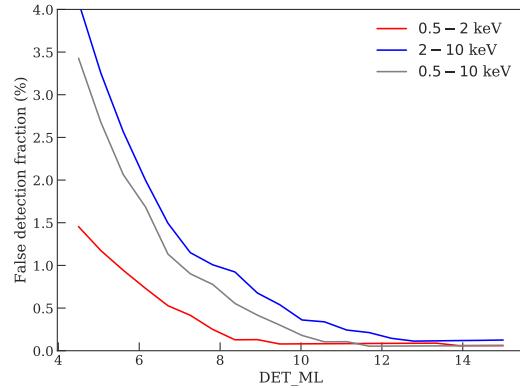
$$S = \frac{m - B}{\exp \times \text{EEF} \times \text{ECF}} \quad (3)$$

Here  $\text{EEF}=0.7$ , and  $\exp$  is the median exposure time within the 70% EEF circle. The upper limits of the source count and source count rate are also corrected for PSF loss. The flux upper limits are calculated as the exposure-time-weighed mean of the three EPIC detectors.

For each source detected in either the soft or the hard bands (or both), we calculate their hardness ratios (HR) defined as  $(H - S)/(H + S)$ , where  $H$  and  $S$  are the source counts weighted by the vignetting-corrected exposure times at the hard and the soft bands, respectively. Note that the source counts here are the default output of EMLDETECT, which is the sum of the counts from all three EPIC detectors. We report this value in our catalog for direct comparison with previous studies. The uncertainties of HR are calculated based on the count uncertainties from the output of EMLDETECT using the error propagation method described in Sec. 1.7.3 of Lyons (1991).

<sup>12</sup> Derived using the COLDEN task included in the CIAO software package. The Galactic column density is calculated at the center of the X-SERVS-LSS region at RA= 35.6625°, DEC= −4.795°.

<sup>13</sup> This is the inverse function of Equation 2.



**Figure 8.** The fraction of spurious sources detected at different DET\_ML based on simulations.

As a comparison, we also match our X-ray catalog with two archival *XMM* catalogs in the X-SERVS-LSS region, the XXL-North catalog (Liu et al. 2016) and the Subaru X-ray Deep Field (Ueda et al. 2008). We match these catalogs to our X-ray sources by searching for counterparts within a  $10''$  radius. A total of 2843 X-ray sources from XXL-North and 714 from SXDS are found to have a counterpart within the  $10''$  radius in our X-ray catalog. For these matched sources, we show the comparison between the soft-band X-ray fluxes reported in these archival catalogs to those in our catalog in Fig. 7. As expected, the majority of the archival sources detected in our catalog have consistent soft-band fluxes. Since the SXDS observations were also used for constructing XXL-North (Liu et al. 2016) catalog, we conclude that at we have detected 2375 new X-ray sources in our survey.

## 4 SURVEY SENSITIVITY AND $\log N - \log S$

### 4.1 Monte Carlo Simulation

To assess our survey sensitivity, we perform Monte Carlo simulations of mock X-ray observations. For each simulation, we generate a list of mock X-ray sources by sampling from the  $\log N - \log S$  relations reported in the XMM-COSMOS survey (Cappelluti et al. 2009, for the 0.5–2 keV and 2–10 keV bands) and the Chandra Multiwavelength Project survey (Kim et al. 2007, ChaMP, for the 0.5–10 keV band). The maximum flux of the mock X-ray catalogs is set at  $10^{-11} \text{ erg s}^{-1} \text{ cm}^{-2}$ . The minimum flux of the mock X-ray sources at each energy band is set as 0.5 dex lower than the minimum detected flux (e.g. LaMassa et al. 2016). We randomly place the mock X-ray sources in the RA/DEC range covered by the *XMM* observations used in this work. We then use a modified version of the simulator written for the *XMM*-CDFS survey (Ranalli et al. 2013), CDFS-SIM<sup>14</sup>, to create mock event files. CDFS-SIM converts X-ray fluxes to PN and MOS count rates with the same model used for deriving the ECFs, then randomly place X-ray events around the source location according to the count rates, the *XMM* PSFs

<sup>14</sup> <https://github.com/piero-ranalli/cdfs-sim>

at the given off-axis angle, and the real exposure maps. We extract images from the simulated event files using the same methods described in Sec. 3. For each observation, the simulated image is then combined with a simulated background, which is created by re-sampling the original background map according to Poisson distributions to create simulated images that mimic real observations. For each energy band, a total of 20 simulations are created.

We run the same two-stage source-detection procedures described in Sec. 3 on the simulated data products. For each simulation, we match the detected sources to the input sources within a  $10''$  search radius by minimizing the quantity  $R^2$ :

$$R^2 = \left( \frac{\Delta \text{RA}}{\sigma_{\text{RA}}} \right)^2 + \left( \frac{\Delta \text{DEC}}{\sigma_{\text{DEC}}} \right)^2 + \left( \frac{\Delta \text{RATE}}{\sigma_{\text{RATE}}} \right)^2 \quad (4)$$

This takes into account the flux and positional differences between the input catalog and the sources detected on the simulated images (e.g., Cappelluti et al. 2007; Ranalli et al. 2015). For detected sources without any input sources within the  $10''$  radius, we consider them to be spurious detections. In Fig. 6, we show the spurious fraction as a function of DET\_ML for the full, soft, and hard bands. For our catalog, we consider sources with DET\_ML > 6 to be reliably detected. At this threshold, the spurious fractions are 0.73%, 2.01%, and 1.68% for the soft, hard, and full-bands, respectively. As for the threshold at DET\_ML = 10.8, the spurious fractions become 0.08%, 0.34%, and 0.10%, respectively. For the full XMM-SERVS-LSS catalog of 5218 sources, these spurious fractions translate to  $\sim 41$  spurious detections for DET\_ML > 6, and 12 spurious detections for DET\_ML > 10.8.

## 4.2 Survey sensitivity, Sky Coverage, and $\log N - \log S$

We create sensitivity maps of our survey region at different bands using the background maps and exposure maps generated as described in Sec. 2.2. The mosaicked background and exposure maps are binned in  $5 \times 5$  pixels ( $\sim 20'' \times 20''$ ). For each pixel on the binned, mosaicked background map, the minimum required source counts to exceed the random background fluctuation are calculated using Equation 2. The sensitivity is then calculated using Equation 3 with the corresponding EEF and ECF values. We show the sky coverage in Fig. 9-left, and the soft-band sensitivity map in Fig. 9-right.

We calculate the  $\log N - \log S$  relations of our survey using the sky coverage curves described above and the following equation:

$$N(>S) = \sum_{i=1}^{N_s} \frac{1}{\Omega_i} \quad (\text{deg}^{-2}) \quad (5)$$

The  $\log N - \log S$  relations of our survey are shown in Fig. 10, along with a selection of the relations at comparable X-ray bands from the literature. Considering factors such as different spectral models and/or methods of generating survey sensitivity curves, our  $\log N - \log S$  relations are consistent with the relations at similar bands reported in the literature.

## 5 MULTIWAVELENGTH DATA IN X-SERVS-LSS

The X-SERVS-LSS region is one of the most extensively observed extragalactic fields. The publicly available multiwavelength observations in the X-SERVS-LSS region included in this work are: the CFHTLS-wide survey (Hudelot et al. 2012), the Subaru HSC Survey (Aihara et al. 2017), the VISTA Deep Extragalactic Observations (VIDEO) survey (Jarvis et al. 2013), *Spitzer* SERVS (Mauduit et al. 2012), and *Spitzer* SWIRE survey (Lonsdale et al. 2003).

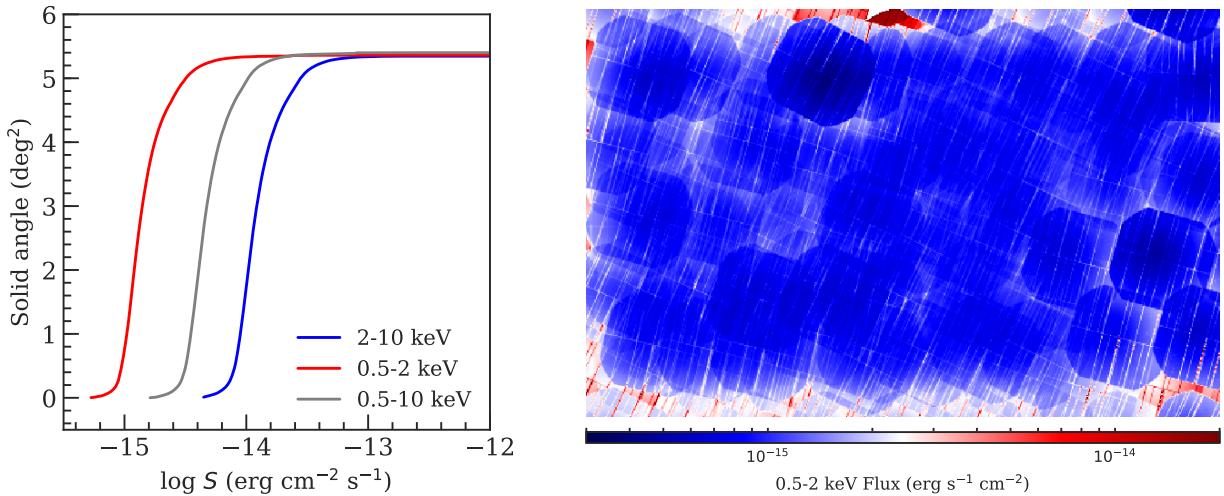
We focus on finding the correct counterparts for our X-ray sources in four deep optical-to-near-IR (OIR) catalogs, CFHTLS, HSC-PDR1, VIDEO, and SERVS. The X-SERVS-LSS field is entirely encompassed by the HSC-PDR1-wide survey. Inside the X-SERVS-LSS field, HSC-PDR1 also has the “ultra-deep” and “deep” surveys, which cover the SXDS and XMDS regions, respectively. The limiting magnitudes in  $i$ -band for the ultra-deep, deep, and wide HSC-PDR1 surveys are 27.2, 26.5, and 26.4 respectively. We select the HSC-PDR1 sources in the RA/DEC ranges marginally larger ( $1'$ ) than the X-SERVS-LSS RA/DEC range based on their  $i$ -band detections. The HSC-PDR1 sources used for our source-matching are the union of the ultra-deep, deep, and wide surveys with the DETECT\_IS\_PRIMARY, CENTROID\_SDSS\_FLAGS, and IDETECTED\_NOTJUNK flags. According to the HSC-PDR1 example script for selecting “clean objects”, we also exclude the HSC sources with FLAGS\_PIXEL\_EDGE, FLAGS\_PIXEL\_SATURATED\_CENTER, FLAGS\_PIXEL\_CR\_CENTER, FLAGS\_PIXEL\_BAD flags in the  $i$ -band to avoid unreliable  $i$ -band sources. The HSC catalogs from three different depth layers were merged according to their OBJECT\_ID. To avoid duplications, sources from the HSC wide survey within  $0.5''$  of any sources from either deep or ultra-deep survey are thrown away. Similarly, sources from the deep survey with distance less than  $0.5''$  from any ultra-deep sources are also discarded. The total number of HSC-PDR1 sources in the X-SERVS-LSS region is  $\sim 3.2 \times 10^6$ , and  $\sim 96\%$  of the X-ray sources in our main catalog have at least one HSC-PDR1 counterpart within their 99% positional uncertainty radius ( $r_{99\%}$  hereafter).

The CFHTLS-W1 survey also covers the entirety of our X-ray data, with 80% completeness limit  $i' = 24.8$ . Similar to HSC-PDR1, we select CFHTLS sources from a region slightly larger than our X-ray survey field. We limit the CFHTLS sources to those with  $SNR > 5$  at  $i'$ -band. The total source number of the  $i'$ -band selected catalog is  $\sim 8.1 \times 10^5$ . 96% of the X-ray sources in our catalog have at least one CFHTLS counterparts within  $r_{99\%}$ .

SERVS is a post-cryogenic *Spitzer* IRAC survey at the near-IR 3.6 and 4.5  $\mu\text{m}$  bands with  $\sim 2\mu\text{Jy}$  survey sensitivity limits and  $\sim 4.5 \text{ deg}^2$  area in the XMM-LSS region. We make use of the highly reliable two-band SERVS catalog built using SExtractor, obtained from the *Spitzer* Data Fusion website<sup>15</sup>, which has  $\sim 4 \times 10^5$  sources. 87% of the X-ray sources have at least one SERVS counterparts within their  $r_{99\%}$ .

VIDEO is a deep survey at near-infrared  $Z$ ,  $Y$ ,  $J$ ,  $H$  and  $K_s$  bands with  $\sim 80\%$  completeness at  $K_s < 23.8$  and a total of

<sup>15</sup> <http://www.mattiacavaccari.net/df/>.



**Figure 9.** *Left* – Sky coverage at soft, full, and hard-band of the X-SERVS-LSS survey. *Right* – Soft-band sensitivity map of the source-extraction region (the same as the cyan box shown in Fig. 2 and Fig 3).

~ $5.7 \times 10^5$  sources. 84% of the X-ray sources have at least one VIDEO counterparts within their 99% error radii.

### 5.1 The likelihood ratio method

To reliably match the X-ray sources to the OIR catalogs with much higher source densities, we employ the likelihood-ratio based matching approach similar to previous deep X-ray surveys, (e.g., Brusa et al. 2007; Luo et al. 2016). The likelihood ratio (*LR* hereafter) is defined as the ratio between the probability that the source is the correct counterpart, and the probability that the source is an unrelated background object (Sutherland & Saunders 1992):

$$LR = \frac{q(m)f(r)}{n(m)}. \quad (6)$$

Here  $q(m)$  is the magnitude distribution of the expected counterparts,  $f(r)$  is the probability distribution function of the angular separation between X-ray and OIR sources, and  $n(m)$  is the magnitude distribution of the background sources.

We calculate the background source magnitude distributions using OIR sources between  $10''$  to  $50''$  from any X-ray sources in our X-ray catalog.

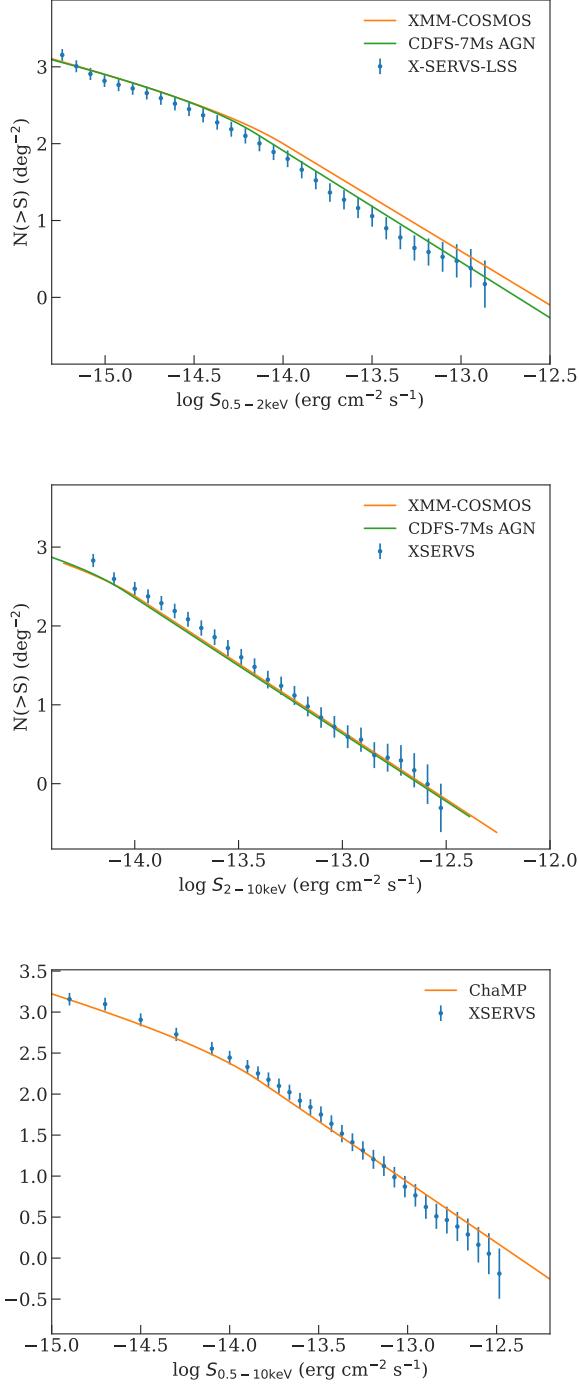
In Sutherland & Saunders (1992), the separation between the two catalogs is described as the joint probability distribution function of the uncertainties in RA and DEC directions. As discussed in Sec. 3.3, the probability density function of the angular separation should follow the Rayleigh distribution:

$$f(r) = \frac{r}{\sigma_x^2} \exp^{-\frac{r^2}{2\sigma_x^2}}. \quad (7)$$

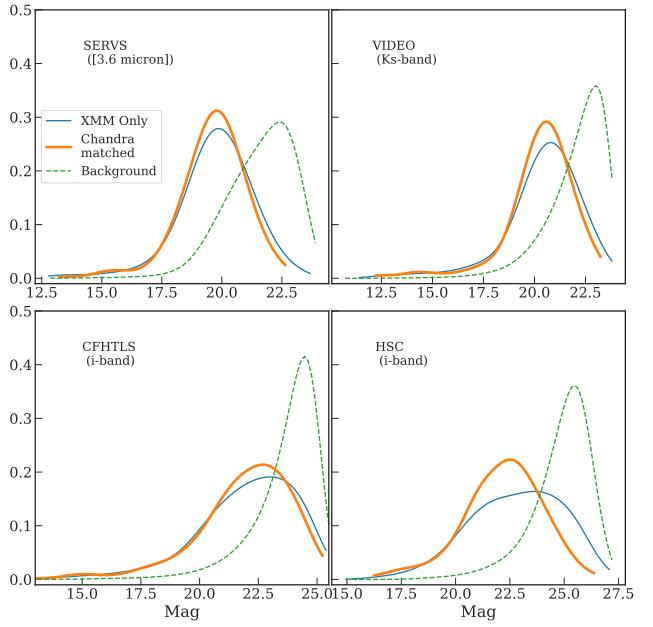
Note that the probability density function is not the two-dimensional Gaussian distribution function which is maximized at  $r=0$ .

The magnitude distribution of the expected counter parts,  $q(m)$ , is often approximated by taking all counterparts

within a small search radius and then derive  $q(m)$  by subtracting the expected number of background sources within the search radius (e.g., Brusa et al. 2007; Marchesi et al. 2016). Some studies also adopt an iterative approach to find  $q(m)$  that optimizes the matching reliability and completeness (e.g., Luo et al. 2010). These approaches can be quite uncertain for our survey because of the large positional uncertainties of *XMM* and the high source densities of the OIR catalogs, particularly for the deep optical surveys. Therefore, we obtain  $q(m)$  for our X-ray sources by first matching our *XMM* catalog to the *Chandra* Source Catalog 2.0 (CSC 2.0 Evans et al. 2010) to take advantage the higher angular resolution and positional accuracy of *Chandra*. We derive the positional uncertainties of the *Chandra* sources in our survey region using the same empirical approach described in Xue et al. (2011) by selecting CSC sources in the RA/DEC range of our catalog, and match them onto HSC-PDR1 using a  $1.5''$  radius. We select CSC sources that are uniquely matched to our X-ray catalogs within the 95% uncertainties (*Chandra* and *XMM* positional uncertainties are added in quadrature). A total of 241 *Chandra* sources are matched to our *XMM* catalog. We match these *Chandra* sources to the four OIR catalogs using Equation 6, with the  $q(m)$  derived using the iterative approach described in Luo et al. (2010). The  $q(m)$  derived from the CSC sources,  $q(m)_{Chandra}$ , is then used as an input for matching our *XMM* sources to OIR catalogs. We note that the flux distributions at soft, hard, and full bands of the *Chandra*-matched subsample are very similar to those of our entire *XMM* catalog, therefore  $q(m)_{Chandra}$  should be consistent with the intrinsic magnitude distributions of the real OIR counterparts of the full X-SERVS-LSS catalog. The counterpart matching processes are ran on four different OIR catalogs: HSC-PDR1, CFHTLS, *Spitzer*-SERVS, and VIDEO. For illustration, we show the magnitude distributions of the background sources and the magnitude distributions of the “reliable matches” derived using CSC sources in Fig. 11. For comparison, we also use the iterative method of Luo et al. (2010) to obtain  $q(m)$  for the full *XMM* catalog,  $q(m)_{XMM}$ , using a  $3''$  initial search radius, which is also plotted on Fig. 11. It is evident that



**Figure 10.** The  $\log N - \log S$  relations for our catalog in the soft-band (top), hard-band (middle), and full-band (bottom). For comparison, we also show the  $\log N - \log S$  relations from the literature. The  $\log N - \log S$  relations of our survey are generally consistent with previous studies.



**Figure 11.** Magnitude distributions (solid lines) for the expected counterparts in SERVS (top-left), VIDEO (top-right), CFTHLS (bottom-left), and HSC-PDR1 (bottom-right). We show the distributions obtained using the full *XMM-Newton* catalog ( $q(m)_{XMM}$ ), and the distributions obtained using the *Chandra* sources in the X-SERVS-LSS field ( $q(m)_{Chandra}$ ). The magnitude distributions of the background, unrelated sources are also shown in each panel as the dashed curves. This figure demonstrates that  $q(m)_{Chandra}$  significantly improves the background dominated  $q(m)_{XMM}$ .

for ultra-deep OIR catalogs such as HSC-PDR1,  $q(m)_{XMM}$  would be skewed towards the faint background sources compared to the *Chandra*-matched subsample. For the other catalogs, we find no qualitative difference between  $q(m)_{Chandra}$  and  $q(m)_{XMM}$ , but we still use  $q(m)_{Chandra}$  for consistency. We next compute the  $LR$  values for all OIR sources within a  $10''$  radius of X-ray sources using Equation 6. For each OIR catalog, we choose the  $LR$  thresholds ( $LR_{th}$ ) such that the reliability and completeness parameters are maximized (see Equation 5 of Luo et al. 2010, for details). Counterparts with  $LR > LR_{th}$  are considered to be reliably matched. The results are reported in Table 2. For each OIR catalog, we show the number of X-ray sources with at least one OIR counterpart within the 99% positional uncertainties of the X-ray sources,  $N_{99}$ ; and the number of X-ray sources with at least one reliably matched source with  $LR > LR_{th}$ ,  $N_{\text{Reliable}}$ . For an X-ray source with only one OIR counterpart within the search radius but its matching  $LR$  is smaller than  $LR_{th}$ , we consider the counterpart to be an acceptable match if the separation from the OIR position to the X-ray position is less than  $r_{99\%}$ . The total number of such cases for each OIR catalog are listed as  $N_{\text{Acceptable}}$  in Table 2. Motivated by the spurious matching rates (see Sec. 5.2 for the cross-matching reliability analysis), we first select a “primary” counterpart for each X-ray source, in the or-

**Table 2.** Counterpart Matching Results

Catalog	Limiting Magnitude	Area	Density ( $10^5/\text{deg}^2$ )	$LR_{\text{th}}$	$N_{99\%}$ (all)	$N_{99\%}$ (mean)	$N_{\text{Reliable}}$	$N_{\text{Acceptable}}$	False Rate (Simulation)	False Rate ( <i>Chandra</i> )
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
SERVS	$[3.6\mu\text{m}] < 23.1$	5.0	9.6	0.34	4530	1.46	3898	259	7.4%	2.7%
VIDEO	$K_s < 23.8$	4.5	11.3	0.24	4285	1.88	3804	90	8.0%	3.1%
CFHTLS-wide	$i < 24.8$	5.4	12.3	0.27	5022	2.37	4138	129	14.4%	5.7%
HSC-PDR1	$i < 26.5$	5.4	43.9	0.13	5002	4.39	4192	52	21.8%	6.4%
Summary	N/A	N/A	N/A	N/A	5218	N/A	4854	64	N/A	

Column 1: Catalog name. Column 2: Survey limits for each catalog in AB. Column 3: Survey area. Column 4: Source density. Column 5:  $LR$  threshold. Column 6: Total number of X-ray sources with at least one counterpart within 99% positional uncertainty in a catalog. Column 7: Average number of sources from the catalog for the X-ray sources. Column 8: Total number of X-ray sources with at least one counterpart within  $LR > LR_{\text{th}}$ . Column 9: Total number of X-ray sources without any reliable counterparts, but has only one counterpart within the 99% positional uncertainty of the X-ray source. Column 10: False matching rates determined using Monte Carlo simulations. See Sec. 5.2. Column 11: The percentage of an X-ray source to have different counterparts obtained using *XMM* and *Chandra* positions. See Sec. 5.2.

der of, SERVS, VIDEO, CFHTLS, and HSC-PDR1<sup>16</sup>. After selecting the primary OIR counterpart, we then associate different OIR catalogs with each other using a simple nearest-neighbor algorithm. Thanks to the much smaller positional uncertainties of OIR catalogs, we adopt a homogeneous search radius of  $1''$  for the OIR catalog associations, which is the approach used by the *Spitzer* Data Fusion database (Vaccari 2016).

As a result, 4854 ( $\sim 93\%$ ) X-ray sources have at least one reliable matches. Of these sources, 3898 are matched to SERVS, 371 are matched to VIDEO, 382 are from CFHTLS, and 174 are from HSC. There are also 959 sources with multiple counterparts having  $LR > LR_{\text{th}}$  in various OIR catalogs. For these sources, we select a “secondary” counterpart based on the strategies described below. First, if there are more than one reliable counterparts from SERVS or VIDEO, we select these second-best matches as the secondary counterparts. We find 62 secondary counterparts from SERVS and 29 from VIDEO. If the reliable counterparts with the highest- $LR$  from HSC and CFHTLS are still not matched to any of the primary and secondary counterparts identified so far, we choose the secondary matches as the highest- $LR$  CFHTLS counterparts (a total of 458 sources), then the highest  $LR$  HSC counterparts (a total of 255 sources). There are also a small number of cases where the highest- $LR$  counterparts from HSC and CFHTLS coincide with the SERVS/VIDEO primary counterparts, but still have a different counterpart with  $LR > LR_{\text{th}}$  in CFHTLS or HSC. We then select these OIR sources as the secondary counterparts, where 96 of them are from CFHTLS and 100 of them are from HSC. Finally, there are 22 X-ray sources with three reliable counterparts, all of the tertiary counterparts are from CFHTLS (6) and HSC (16), which are also reported in our final catalog.

For the remaining X-ray sources without any of the counterparts with  $LR > LR_{\text{th}}$ , 61 ( $\sim 1\%$ ) of them are considered to have “acceptable” matches, where 48, 7, 6 of them are matched to SERVS, VIDEO, and CFHTLS, respectively. For the remaining sources, the vast majority of them still have at least one OIR counterpart within the relevant 99% positional uncertainty circle, only 23 X-ray sources are completely isolated. Visual inspection of these sources shows that most of

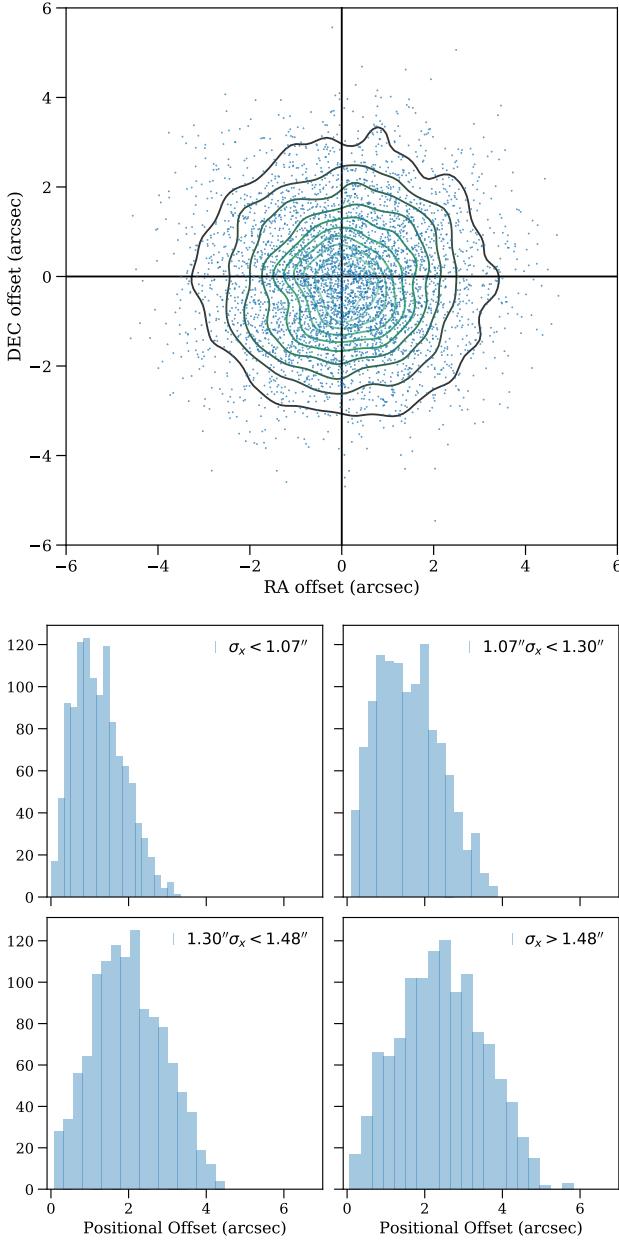
the 23 sources are in the vicinity of a bright star, thus making the pipeline OIR photometry unreliable. **Discuss these sources or show images in appendix?** We plot the positional offsets between the X-ray sources and the reliably matched sources in Fig. 12.

## 5.2 Counterpart identification reliability

We assess the reliability of the  $LR$  matching results using the Monte Carlo simulation approach described in Broos et al. (2007); Xue et al. (2011). A total of 100 simulations are carried out for each OIR catalog and the X-ray catalog. As described in Broos et al. (2007, 2011), we divide the X-ray sources into two different intrinsic populations when running simulations: the “associated population” of which the X-ray sources indeed have a counterpart in the OIR catalogs considered, and the “isolated population” of which the X-ray sources are not associated with any OIR sources. This method allows us to estimate the number of correctly matched X-ray sources ( $N_{\text{Correct}}$ ), and the number of X-ray sources having a falsely matched counterpart,  $N_{\text{False}}$  (which is comprised of the incorrect matches for the associated population and the false matches of the isolated population) for the given  $LR_{\text{th}}$ . The details of the simulation setups can be found in the appendix of Broos et al. (2007) and Sec. 5 of Broos et al. (2011). The false matching rate for each OIR catalog are calculated as  $N_{\text{False}}/(N_{\text{Correct}} + N_{\text{False}})$ . Notably, adopting the *Chandra*-matched counterpart magnitude density,  $q(m)_{\text{Chandra}}$ , does reduce the false-matching rates derived using  $q(m)_{\text{XMM}}$ . For SERVS and VIDEO catalogs, the improvements are marginal ( $< 0.5\%$ ), while the improvements for CFHTLS and HSC-PDR1 are more significant ( $\sim 2\%$  and  $6\%$ , respectively.)

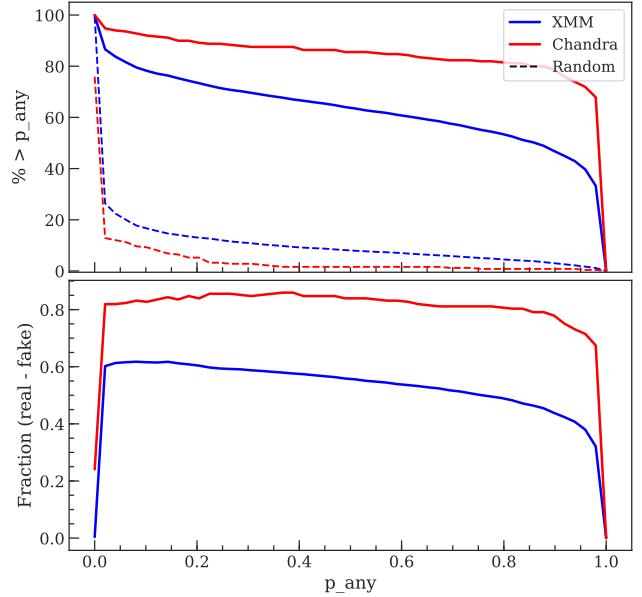
We further scrutinize the  $LR$  matching reliabilities by making use of the 241 CSC sources and their multiwavelength matching results described in Sec. 5.1. We assess the reliability of the matching results of these *Chandra* Sources using the aforementioned Monte Carlo method, and find that the false-match fractions are 2.8%, 3.3%, 0.9%, and 1.4% for CFHTLS, HSC-PDR1, SERVS, and VIDEO surveys, respectively. When directly comparing the matching results obtained with *XMM* and *Chandra* positions, 97%, 94%, 97%, and 94% of the *Chandra* matching results and the *XMM* results are the same for the SERVS, VIDEO, CFHTLS, and HSC catalogs, respectively. The excellent agreement between the matching results of the *Chandra* sources and those of

<sup>16</sup> Note that 85%, 10%, 1%, and 4% of the *Chandra* sources have their “primary” counterparts from SERVS, VIDEO, CFHTLS, and HSC, respectively.



**Figure 12.** *Top:* Distribution of the positional offsets in the RA vs. DEC plane for the 4854 reliably matched sources. Note that the residual positional uncertainties are  $< 0.1''$  in both the RA and DEC directions. *Bottom:* Histograms of positional offsets for the 4854 reliably matched sources, divided into four bins based on their positional uncertainties.

corresponding *XMM* sources suggests that the counterpart ‘‘confusion’’ should be minimal, and the false-matching rates suggested by the Monte Carlo simulation is a conservative estimation (e.g., 97% of the *Chandra* sources have the same SERVS counterparts as the *XMM-Newton* sources, which is significantly better than the 7–8% spurious rates estimated using Monte Carlo simulations.)



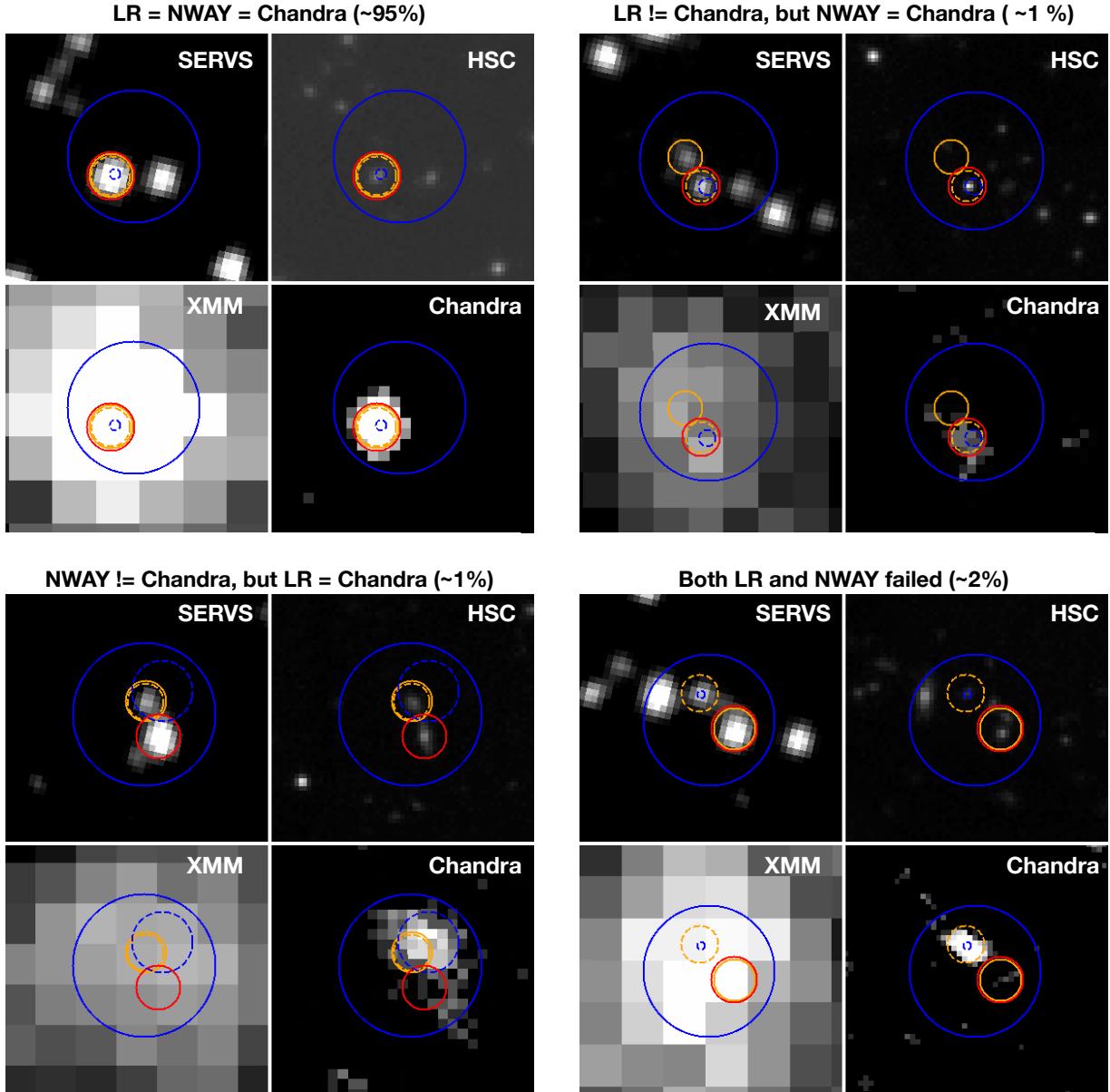
**Figure 13.** *Left:* Magnitude distribution for all sources within the search radius around *XMM* ( $3''$ ), or around the *Chandra* ( $1.5''$ ) sources that are matched to our *XMM* catalog. *Right:* Diagnostic plot for the Bayesian matching results. The y-axis is the fraction of X-ray sources with *LR* values larger than the given *LR* values at the x-axis. The y-axis is the fraction of X-ray sources with *p*\_any values larger than the given *p*\_any values at the x-axis.

### 5.3 The NWAY Bayesian catalog matching method

We supplement the *LR* matching results with the Bayesian catalog matching tool, NWAY<sup>17</sup> (Salvato et al. 2017). The fundamental difference between the Bayesian approach and the likelihood-ratio approach is that the Bayesian approach makes use of the distance and magnitude priors from multiple catalogs simultaneously to select the most probable match. The details of the NWAY matching methodology is described in Appendix B of Salvato et al. (2017).

In short, NWAY computes three informative quantities for deciding the most probable match, *p*\_single, *p*\_any and *p*\_i, where each possible counterpart has a different *p*\_single value based on their distance from the *XMM* position. This value could be weighted by the priors supplied (e.g, *q*(*m*) and *n*(*m*) in Equation 6 are similar to a magnitude prior). In our case, *p*\_single is the posterior probability for a counterpart that is correctly associated with the X-ray source based on the angular separation from the X-ray position weighted by the magnitude distribution prior, and the surface densities of the X-ray and OIR catalogs. For each X-ray source, *p*\_single of all possible counterparts are considered to compute a single *p*\_any value, which stands for the posterior probability of the X-ray source having any correct counterparts (i.e., *p*\_any = 0 if there are no OIR counterparts within the search radius of the X-ray source.) The last quantity, *p*\_i, is the relative probability of a possible counterpart being the correct match. For an X-ray source with multiple possible counterparts, the counterpart with the highest *p*\_i (*p*\_i (Best)) is considered to be the most probable match and is assigned

<sup>17</sup> <https://github.com/JohannesBuchner/nway>.



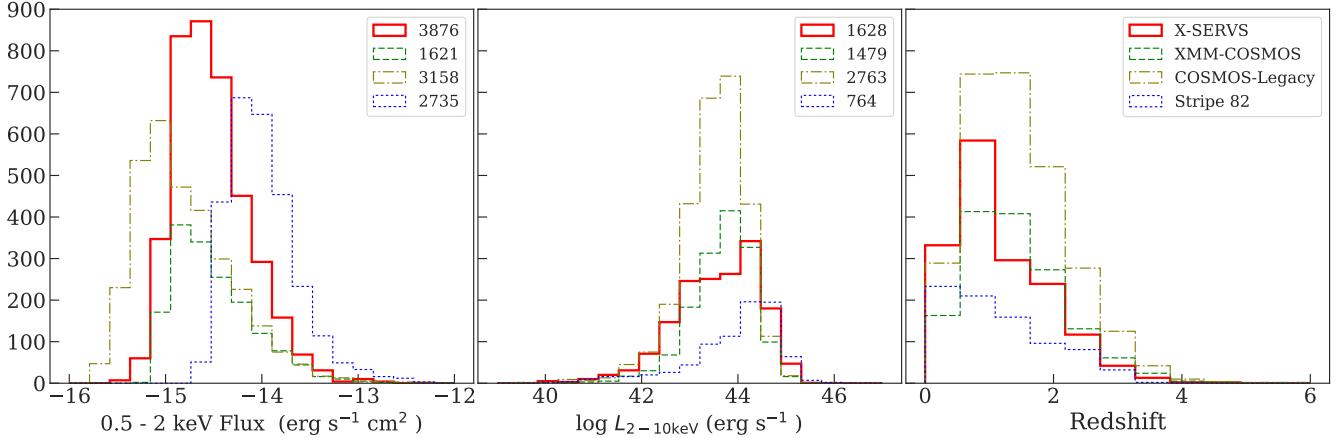
**Figure 14.** Illustrations of the comparison between the matching results using *XMM* positions or *Chandra* positions for four X-ray sources in our sample. Within the panel for each source, we show images from SERVS [3.6 $\mu$ m](top-left), HSC *i*-band (top-right), *XMM* 0.5–10 keV (bottom-left), and *Chandra* 0.5–7 keV (bottom-right). X-ray positions are marked as blue circles with a 99% error radius, with the *XMM* positions marked using solid lines and the *Chandra* positions are marked using dashed lines. SERVS counterparts identified with the *LR* method are marked as orange circles with a  $\sim 2''$  radius, solid lines mark the counterparts of the *XMM* positions and dashed lines are the counterparts of the *Chandra* positions. SERVS counterparts of the *XMM* positions identified using NWAY are shown as the red circles. For the vast majority of *XMM* sources with *Chandra* counterparts from CSC, our counterpart-matching results are identical to the results obtained using *Chandra* coordinates and positional uncertainties.

the “MATCH\_FLAG = 1” flag by NWAY. Counterparts with almost as good  $p_i$  ( $p_i > 0.5 \times p_{i\text{ (Best)}}$ ) are also flagged by NWAY as (“MATCH\_FLAG = 2”).

Similar to our *LR*-based approach, we make use of the *Chandra* sources in the X-SERVS-LSS field to compute the priors of the expected counterparts. We use the “auto” functionality of NWAY with a 1.5'' search radius for defining the “real” counterparts. In addition to the magnitude priors, we also include an additional prior based on the *Spitzer*

IRAC color from SERVS, [3.6 $\mu$ m]/[4.5 $\mu$ m]. We use the magnitude and color priors to run NWAY source matching on the full X-SERVS-LSS sample. We also run NWAY on the CSC sources with *XMM* counterparts in our survey region. For a small number of sources, this additional prior is useful for discerning two adjacent SERVS sources with comparable magnitudes (see the top-right panel of Fig. 14 for illustration).

After computing the magnitude and IRAC color priors using



**Figure 15.** A comparison between the X-SERVS-LSS (solid red line), COSMOS-Legacy (brown dash-dotted line), XMM-COSMOS (blue dotted line), and Stripe 82-X (green dashed line) surveys. Distributions shown in panels from left to right are: 0.5–2 keV flux,  $\log L_{2-10\text{keV}}$ , and redshift, respectively.

the *Chandra* sources, we run NWAY on the full X-ray catalog with a search radius of  $10''$ . All four OIR catalogs are considered simultaneously. We report the multiwavelength matches with `MATCH_FLAG=1,2` in a separate table (Table B) supplementary to the likelihood-ratio based results. As described in Salvato et al. (2017), the matching reliability can be determined by a  $p_{\text{any}}$  threshold, which is chosen based on re-running NWAY on randomly shifted “fake” X-ray catalogs. We test the fraction of spurious detections by randomly shifting the *XMM* positions, and run the Bayesian matching again. In practice, we choose the  $p_{\text{any}}$  threshold that maximizes the completeness and reliability of the matching results:

$$\text{Fraction}_{>p_{\text{any}}}(\text{Real}) - \text{Fraction}_{>p_{\text{any}}}(\text{Fake}). \quad (8)$$

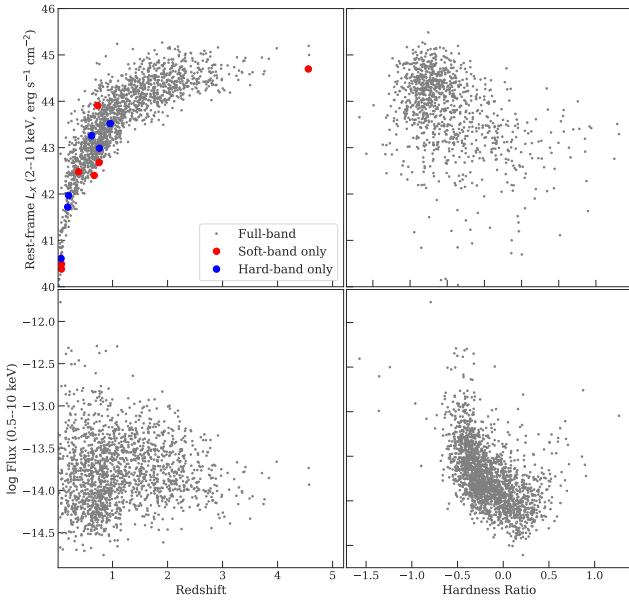
Here we define the “completeness” using  $\text{Fraction}_{>p_{\text{any}}}(\text{Real})$ , which is the fraction of real X-ray sources with  $p_{\text{any}}$  larger than the given  $p_{\text{any}}$  values. The reliability is defined as  $\text{Fraction}_{>p_{\text{any}}}(\text{Fake})$ , which is the fraction of simulated, “fake” X-ray sources with  $p_{\text{any}}$  larger than the given  $p_{\text{any}}$  values. For the full *XMM* catalog, Equation 8 is maximized at  $p_{\text{any}} = 0.08$ . Only  $\sim 80\%$  of the real X-ray sources satisfy this criterion. Meanwhile,  $\sim 20\%$  of the fake X-ray sources also satisfy this criterion.

Similar shift-and-rematch steps are also carried out for the *Chandra* sources in the X-SERVS-LSS region. For the *Chandra* matching results, Equation 8 is maximized at  $p_{\text{any}} = 0.25$ , where  $\text{Fraction}_{>p_{\text{any}}}(\text{Real}) \sim 88\%$ , and  $\text{Fraction}_{>p_{\text{any}}}(\text{Fake}) \sim 3\%$ . We show the fractions of X-ray sources with  $p_{\text{any}}$  above different  $p_{\text{any}}$  thresholds ( $\text{Fraction}_{>p_{\text{any}}}(\text{Real})$ ) in Fig. 13. The fractions for the fake X-ray catalogs are also shown as the dashed lines. The values of Equation 8 as a function of  $p_{\text{any}}$  are also shown in the bottom panel.

The “spurious fraction”,  $\text{Fraction}_{>p_{\text{any}}}(\text{Fake})$ , at our chosen  $p_{\text{any}}$  threshold is  $\sim 20\%$ . This spurious fraction is not comparable to the false-matching rates calculated based on the Monte Carlo simulations for the *LR* matching results. The high  $\text{Fraction}_{>p_{\text{any}}}(\text{Fake})$  shown here does not neces-

sarily mean that more than 20% of the multiwavelength matches are “spurious”. Given the depth of the OIR catalogs in X-SERVS, we expect that most of our X-ray sources to have a true counterpart in each OIR catalog (e.g., the deeper *Chandra* COSMOS-legacy survey has a near  $\sim 100\%$  counterpart recovery fraction given the similar OIR survey depths), and the high fraction of fake X-ray sources with  $p_{\text{any}} > p_{\text{any}}(\text{threshold})$  simply means that it is highly likely for a random background OIR source to fall into the error circle of the X-ray sources. As shown in Fig. 13, the “spurious fraction” of the *Chandra* sources is much lower than that of the full *XMM* catalog at any given  $p_{\text{any}}$  values, but the completeness is still somewhat limited ( $\sim 88\%$ ). In practice, the true spurious matching rates of NWAY should be similar to those from the Monte Carlo simulation of *LR* matching results, as the mock X-ray catalog approach only accounts for the spurious rates of the “isolated population” (see Broos et al., 2007, and Xue et al., 2011 for details), which is usually higher than the results of simulation. Therefore, we do not include any  $p_{\text{any}}$  cuts as doing so would significantly affect the completeness of the matching results.

Similar to Sec. 5.2, the matching results can also be assessed by investigating the CSC-matched subsample of 241 X-ray sources. We focus only on comparing the SERVS counterparts, as the majority of *LR* matching results are decided based on the primary counterparts from SERVS. We confirm that all *Chandra* sources have the same SERVS matching results using *LR* and NWAY. Therefore, we can use the *Chandra* results obtained with *LR* to assess the matching reliability of both *LR* and NWAY matching results for the *XMM* sources. We show examples of such comparisons in Fig. 14. We find 95% the matching results of *LR*, NWAY, and *Chandra* are identical. A small fraction ( $\approx 1\%$ ) of *LR* matching results do not agree with those of *Chandra* but could be recovered by NWAY. On the other hand, a similar fraction ( $\approx 1\%$ ) of the NWAY matching results do not agree with the *Chandra* results but could be identified by *LR*. About 2% of the *Chandra* sources have different SERVS counterparts than both *LR* and NWAY results. This sug-



**Figure 16.** Properties of the 1762 X-ray sources with spectroscopic redshift measurements. Top-left panel shows the  $L_{2-10\text{keV}} - z$  distribution, top-right panel shows  $L_{2-10\text{keV}}$  vs. hardness ratio. Bottom-left panel shows the distribution of 0.5–10 keV flux vs. redshift. Bottom-right panel shows the distribution of 0.5–10 keV flux vs. hardness ratio.

gests *LR* and *NWAY* perform similarly for finding SERVS counterparts.

However, when further scrutinizing the 95% sources with identical SERVS counterparts from *LR*, *NWAY*, and *Chandra*, we find that *NWAY* only identifies  $\sim 90\%$  of the counterparts from the other three OIR catalogs when compared to the *Chandra* matching results. These mismatches between *XMM* and *Chandra* results are likely due to how *NWAY* computes  $p_i$ . When multiple OIR catalogs are taken into account simultaneously,  $p_i$  represents the relative probability of counterparts from *all* OIR catalogs being the correct match. It is possible that sometimes  $p_i$  would be higher when the source from one of the OIR catalogs is “absent”, especially when multiple magnitude priors are involved. For instance, in the case where the X-ray source is matched to a SERVS counterpart that only has one VIDEO source within the error circle of SERVS and VIDEO catalogs, *NWAY* sometimes reports that the X-ray source has no VIDEO counterparts. This is likely due to the magnitude and color priors down-weighting the posterior probabilities of the VIDEO counterpart.

Similar mismatches are also found when comparing the *NWAY* and *LR* matching results for the full *XMM* catalog. For the 3898 sources with primary *LR*-matching counterparts from SERVS,  $\sim 95\%$  of them are matched to the same SERVS counterparts by *NWAY*. However, only 94%, 96%, and 71% of the 3898 sources have identical counterparts from VIDEO, CFHTLS, and HSC-PDR1, respectively. Further inspecting these sources suggests that most of these mismatches are due to *NWAY* not treating OIR counterparts from all catalogs to be reliable, even though they are the only source within the error circle of the reliable counter-

**Table 3.** Redshift catalogs

Catalog	Instrument	Survey sensitivity	$N_X$
3D-HST	WFCS G141 Grism	$JH_{\text{IR}} \lesssim 24$	5
CSI	IMACS (Uniform Dispersion Prism)	$[3.6\mu\text{m}]_{\text{AB}} \lesssim 21$	57
PRIMUS	IMACS (Low-Dispersion Prism)	$i \lesssim 23.5$	364
SDSS	BOSS	$r \lesssim 22.5$	1104
UDSz	VIMOS/FORS2	$K < 23$	17
VVDS	VIMOS	$17.5 \lesssim i \lesssim 24.75$	43
VIPERS	VIMOS	$i \lesssim 22.5$	172

Redshift catalogs used in this work. Column 1: Redshift survey name. Column 2: Survey instrument. Column 3: Survey sensitivity. Column 4: Total number of redshifts assigned to the X-ray sources in the main catalog

parts from other OIR catalogs. We note that *NWAY* do not have this behavior when no magnitude or color priors are used, but doing so would lose the discerning power of magnitude or color priors. Further corroborating the Bayesian method’s effectiveness of counterpart-matching with multiple OIR catalogs is beyond the scope of this work. Therefore, we list the *NWAY* matching results as-is in Table B and we consider only the *LR*-based matching results listed in Table A when exploring the multiwavelength properties in this work.

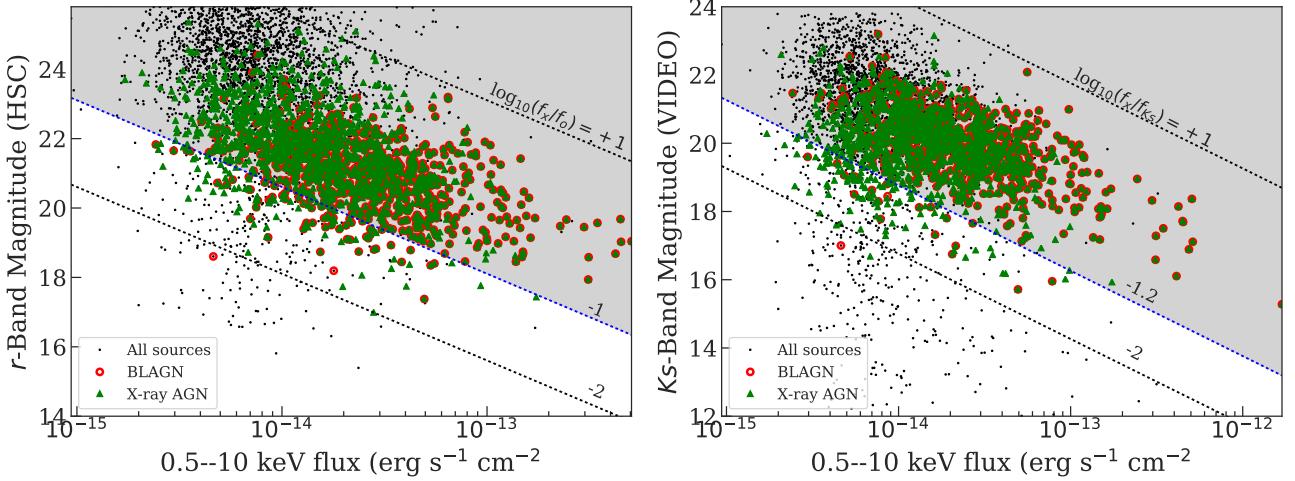
#### 5.4 Redshifts

The X-SERVS-LSS region is covered by a number of redshift surveys that target galaxies with various optical magnitude constraints, including the PRISM Multi-Object Survey (PRIMUS, Coil et al. 2010), the VIMOS Public Extragalactic Redshift Survey (VIPERS, Garilli et al. 2014), the VIMOS VLT Deep Survey (VVDS, Fèvre et al. 2013). As part of the SDSS-BOSS program, 3042 X-ray sources found in the XXL-N field ( $25 \text{ deg}^2$ ) with  $r < 22.5$  were all followed-up by SDSS (Menzel et al. 2016). Also, there are three other redshift surveys in the X-SERVS-LSS region that target near-IR selected galaxies, including the spectroscopic follow-up observations of the UKIDSS Ultra-Deep Survey (UDSz, Bradshaw et al. 2013; McLure et al. 2013), the 3D-HST Survey (Skelton et al. 2014; Momcheva et al. 2016) in the UDS region, and the Carnegie-Spitzer-IMACS Redshift Survey (CSI, Kelson et al. (2014)).

We adopt the same nearest-neighbor matching criterion with a  $1''$  matching radius to associate these redshifts to each OIR catalog. The redshift for each X-ray source is determined by the coordinates of its primary OIR counterpart. In cases where redshifts from different catalogs do not agree with each other, we choose redshifts using the following order (ranked by spectral resolution): SDSS, VVDS, VIPERS, UDSz, PRIMUS (reliable), CSI (reliable), 3D-HST, PRIMUS (acceptable), CSI (acceptable). Of the 5218 sources in our main X-ray source catalogs, 1762 of them have spectroscopic redshifts ranges from  $0.002 < z < 7.01$ . We list the properties and the relevant numbers of each redshift catalogs in Table 3.

#### 5.5 Source properties and classification

In this section we discuss the properties of the 4854 sources with reliable or acceptable counterparts. For the X-ray



**Figure 17.** Right – Distributions of the full-band (0.5–10 keV) X-ray flux and the optical flux at  $r$ -band. Left – Full-band X-ray flux versus the near-IR flux at  $Ks$ -band. In both plots, the shaded regions mark the “AGN” regime as defined by the  $\log_{10} f_x/f_o > -1$  (right) or the  $\log_{10} f_x/f_{Ks} > -1.2$  (left) thresholds as described in Sec. 4.4 of Xue et al. (2011). For sources with spectroscopic redshift measurements, we also make those with  $L_X > 3 \times 10^{42}$  erg cm $^{-2}$  s $^{-1}$  as the green triangles. Source that are identified as broad-line AGNs are also marked as open red circles.

sources with a secure redshift, we calculate their rest-frame 2–10 keV luminosity assuming a  $\Gamma = 1.7$  power-law and corrected for the Galactic absorption. We compare the luminosity, flux, and redshift distributions of our sample to those from archival X-ray surveys, including *XMM-COSMOS*, COSMOS-legacy, and Stripe 82X surveys. We show the comparison in Fig. 15. The comparisons in the middle and right panels are limited to the availability of spectroscopic redshifts in Stripe-82 and X-SERVS-LSS regions. We note that high-quality photometric redshifts are already available in a  $\sim 1 \text{ deg}^2$  area within the X-SERVS-LSS region (Nyland et al. 2017). The photometric redshift catalog for the full survey region will also be available soon (Pforr et al., in preparation). Nonetheless, the left-panel of Fig. 15 demonstrates that our X-SERVS-LSS survey occupies a unique parameter space in X-ray surveys by more than doubling the source counts of the *XMM-COSMOS* survey, which will enable a wide range of science that was previously limited by either survey sensitivity or cosmic variance. For our sample, the  $L_X - z$  distribution is shown in Fig. 16, along with the  $L_X$  vs. HR, HR vs. Redshift, and full-band flux vs. Redshift distributions.

We also include the basic source classifications in our catalog. For sources with spectroscopic observations, we directly make use of the spectroscopic classifications when available. A total of 831 sources are classified as AGNs based on the presence of the broad lines in their optical spectra from SDSS, VIPERS, or VVDS. For the other sources, we use the criteria described in Luo et al. (2016) to select AGNs, including: 1. A X-ray luminosity threshold where we regard sources with rest-frame  $L_{2-10\text{keV}} > 3 \times 10^{42}$  erg s $^{-1}$  as an AGN. A total of 1504 sources satisfy this criterion. 2. X-ray bright sources with X-ray-to-optical and X-ray-to-near-IR flux ratios larger than  $\log f_x/f_R > -1$  and  $\log f_x/f_{Ks} > -1.2$ , respectively. There are 3409 sources with  $\log f_x/f_R > -1$  and 4071 sources with  $\log f_x/f_{Ks} > -1.2$ . The total number of

sources classified as AGN is 4584, or  $\sim 88\%$  of the total sample. We show the flux ratio distributions in Fig. 17.

## 6 CATALOG DESCRIPTION

Here we describe the columns of the main X-ray source catalog, Table A. Null values are filled with -99.

- Column 1: the unique source ID (XID) assigned to each X-ray source.
- Columns 2–3: RA and DEC in degrees of the X-ray source.
- Column 4: X-ray positional uncertainty ( $\sigma_x$ ) in arcsec. Note that this is not the  $\sigma$  of Gaussian distribution but rather the scaling parameter of the univariate Rayleigh distribution (see Sec. 3.3 and Pineau et al. (2017) for details). The positional uncertainties are based on the those of the full-band. For sources without a full-band detection, the soft or hard-band positional uncertainties are used.
- Columns 5–6: 68% and 99.9% X-ray positional uncertainties in arcsec, see Sec. 3.3 for details.
- Columns 7–14: RA/DEC in degrees of the source at the full-band, soft-band, and hard-band, respectively.
- Columns 15–18: DET\_ML at different bands.
- Columns 23–34: Exposure times in seconds (total, PN, MOS1, MOS2) at different bands.
- Columns 35–50: Value of the background map (counts per pixel) at different bands.
- Columns 51–58: Source counts at different bands (PN+MOS)
- Columns 59–66: Total count rates at different bands.
- Columns 67–78: Source counts at different bands for PN, MOS1, and MOS2.
- Columns 79–90: Count rates at different bands for PN, MOS1, and MOS2. In count s $^{-1}$ .
- Columns 91–98: Flux and flux uncertainty (error-weighted average of different EPIC detectors) for different bands. In erg s $^{-1}$  cm $^{-2}$ .

14. Columns 99–101: Hardness ratio and its upper and lower limits.
15. Columns 102: Rest-frame 2–10 keV X-ray luminosity
16. Columns 103–106: Number of sources within the 99% positional uncertainties.
17. Columns 107–110: Number of sources that satisfies  $LR \geq LR_{\text{th}}$ .
18. Column 111: Whether an X-ray source has at least one reliable counterpart with  $LR > LR_{\text{threshold}}$  from any of the four OIR catalogs.
19. Column 112: Whether an X-ray source has no reliable counterparts but has at least an acceptable counterpart from any of the four OIR catalogs (see Sec. 5 for details).
20. Column 113: Whether an X-ray source has no reliable or acceptable counterparts.
21. Column 114: Whether an X-ray source has no counterparts from any of the four OIR catalogs.
22. Column 115: Catalog from which the primary counterpart is selected.
23. Column 116–118: RA, DEC of the primary counterpart and its separation from the X-ray source.
24. Column 119: The matching likelihood ratio.
25. Columns 120–124: The same as Columns 114–118, but for the secondary counterpart (only if the secondary counterpart is also reliable, see Sec. 5.1 for details).
26. Columns 125–129: The same as Columns 114–118, but for the tertiary counterpart (only if the tertiary counterpart is also reliable).
27. Columns 130–141: RA, DEC, and Object ID from the original OIR catalog for the primary counterpart.
28. Columns 142–153: Same as Columns 126–137 but for the secondary counterpart.
29. Columns 154–165: Same as Columns 126–137 but for the tertiary counterpart.
30. Columns 166–167: SERVS 3'' aperture photometry at  $3.6\mu\text{m}$  and  $4.5\mu\text{m}$ .
31. Columns 167–171: SWIRE 3'' aperture photometry at  $3.6\mu\text{m}$ ,  $4.5\mu\text{m}$ ,  $5.8\mu\text{m}$ , and  $8.0\mu\text{m}$ .
32. Columns 172–181: VIDEO psf photometry  $Y$ ,  $J$ ,  $H$ , and  $Ks$ -band magnitude and uncertainties.
33. Columns 182–191: CFHTLS PSF photometry at  $u$ ,  $g$ ,  $r$ ,  $i$ ,  $z$  bands.
34. Columns 192–201: HSC cmodel photometry at  $g$ ,  $r$ ,  $i$ ,  $z$ ,  $y$  bands.
35. Columns 202–204: RA, DEC, and Object ID from the original redshift catalogs for the primary counterparts.
36. Column 205: Redshift.
37. Column 206: The catalog from which the redshift is culled from.
38. Column 207: Original redshift flag from one of the redshift catalogs. For SDSS, see <http://www.sdss.org/dr14/algorithms/bitmasks/#ZWARNING>. For VVDS, see Sec. 3.4 of Fevre et al. (2013). For VIPERS, see Sec. 4.3 of Garilli et al. (2014). For PRIMUS, see <http://primus.ucsd.edu/version1.html#ztags>. For CSI, see Sec. 4.6 of Kelson et al. (2014). For UDSz, see McLure et al. (2013). For the 3D-HST catalog, we only select redshifts with  $\sigma_z/(1+z) \leq 0.003$  thus no redshift flags are included.
39. Column 208: Source classification. See Sec. 5.5 for details.

We also include a supplementary table for the matching results by NWAY. The columns of Table B are listed below.

The first-half of the table presents the matching results that included magnitude priors for each OIR catalog and a prior of IRAC color for SERVS.

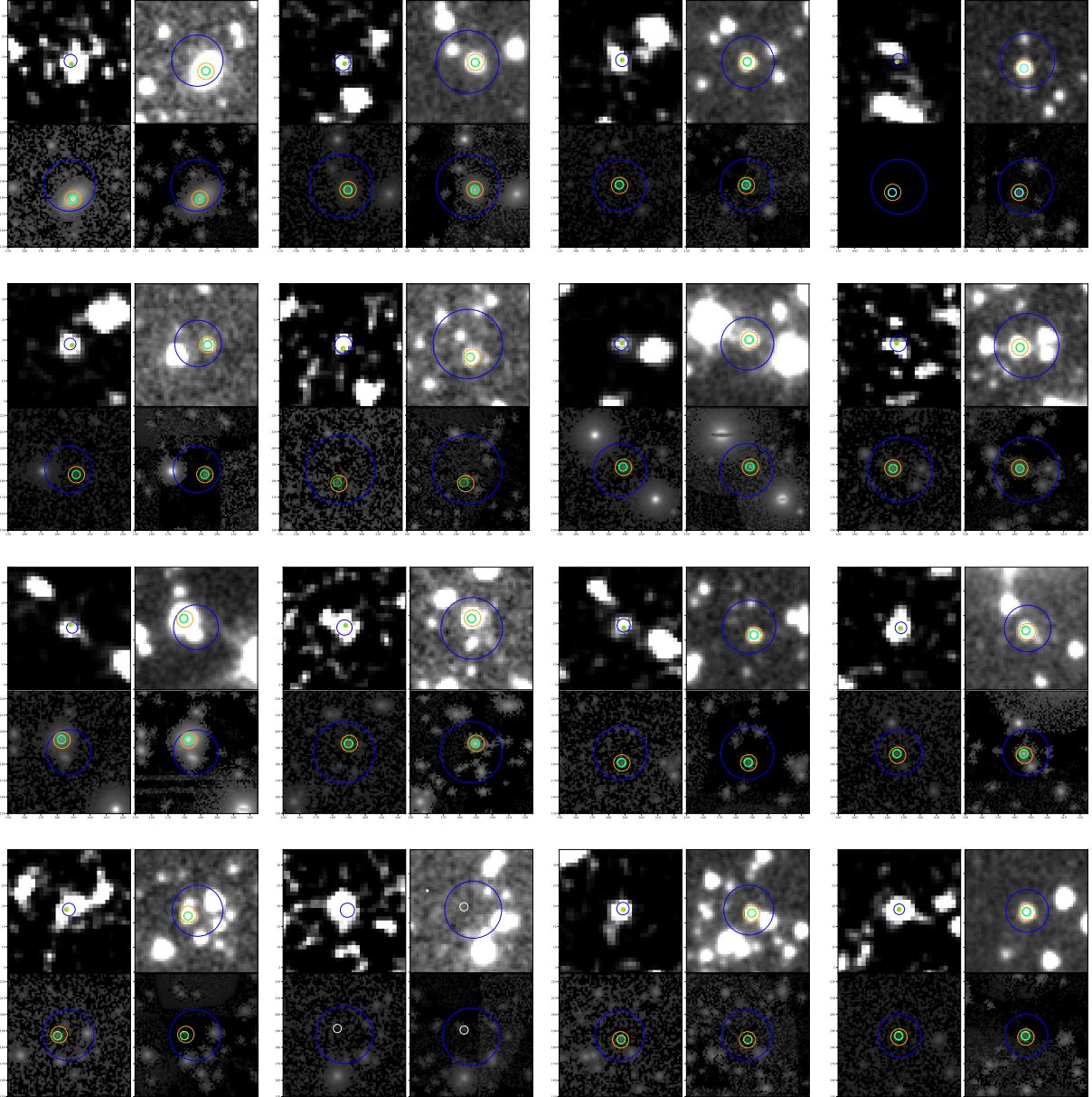
1. Column 1: the unique source ID (XID) assigned to each X-ray source.
2. Columns 2–13: RA, DEC, separation to the X-ray position, and the original Object ID of the matched counterpart from each OIR catalog.
3. Column 14:  $p_i$
4. Column 15:  $p_{\text{any}}$
5. Column 16:  $p_{\text{single}}$
6. Column 17: MATCH\_FLAG. The most probable counterparts have MATCH\_FLAG=1. Other counterparts that are almost as good have MATCH\_FLAG=2.

For each source, we also generate postage-stamp images in the X-ray, mid-IR, near-IR, and optical wavelengths. For illustration, we show a random collection of 16 sources in Figure 18.

## 7 SUMMARY

In this work, we present the X-ray source catalog in the X-SERVS-LSS region constructed with both archival and new *XMM* AO15 data. We summarize the main results as the following:

1. Our X-SERVS-LSS catalog is constructed based on data in a  $\sim 5.3$  deg $^2$  rectangular region centered at RA=35.58°, DEC=−4.965°. A total of 153 pointings from 149 different *XMM* obsIDs are used, with a total of  $\sim 2.7$  Ms background-filtered exposure time (1.1 Ms from AO15).
2. The main X-ray source catalog is generated using the EWAVELET and EMLDETECT. We list all 5218 sources with EMLDETECT DET\_ML > 6 in either the full-band (0.5–10 keV), soft-band (0.5–2 keV), or the hard-band (2–10 keV). Of the 5218 sources, we find 2843 of them to be the same X-ray sources identified in previous X-ray surveys in our survey area (the XXL-N surveys, Liu et al. 2016) and 2375 are newly discovered X-ray sources.
3. The median fluxes in erg s $^{-1}$  cm $^{-2}$  for the three X-ray bands are  $2.7 \times 10^{-15}$  (0.5–2 keV),  $1.5 \times 10^{-14}$  (2–10 keV),  $8.8 \times 10^{-15}$  (0.5–10 keV). There are 2928 sources with more than 100 X-ray counts in full-band (PN + MOS), and 126 sources with more than 1000 X-ray counts.
4. Monte Carlo simulation suggests that the number of spurious sources should be  $\sim 41$  with a DET\_ML=6.0 threshold, which is  $\sim 99.2\%$  reliable. If we raise the detection threshold to DET\_ML=10.8, the expected number of spurious sources would be 12, or 99.8%.
5. The absolute astrometry of the *XMM* catalog is registered to the WCS frame of the Subaru HSC-PDR1 survey. The positional uncertainties for the X-ray sources are determined based on an empirical relation between the X-ray-to-optical positional offsets and the X-ray source counts, which is well-characterized by the Rayleigh distribution. The median positional uncertainties at the full, soft, and hard-bands are: 1.31'', 1.35'', and 1.37'', respectively.
6. We search for OIR counterparts in SERVS, VIDEO, CFHTLS, and HSC-PDR1 surveys and we find that 99.5% (5196/5218) of the X-ray sources have at least one OIR counterparts within the 99% positional uncertainties. We



**Figure 18.** Postage-stamp images for 16 randomly selected X-ray sources. For each source, we show 1. X-ray image (upper-left panel) 2. mid-IR image from SERVS (upper-right panel) 3. near-IR image from VIDEO (lower-left panel) 4. optical image from HSC-PDR1 (lower-right panel). In each image, the X-ray position is marked as the blue circle with the corresponding 99% error radius. The position of the most probable mid-IR counterpart is marked as the orange circle with a 2'' radius. The positions of VIDEO, HSC, and CFHTLS counterparts are marked as green, white, and cyan circles with a 1'' radius. The entire set of postage-stamp images are available separately.

also find that  $\sim 93\%$  (4854/5218) of the X-ray sources have at least one reliable OIR counterpart. We also collect 1762 secure redshifts from SDSS, VIPERS, VVDS, UDSz, PRIMUS, CSI, and 3D-HST.

7. We test the matching results using a subsample of 241 X-ray sources with a reliable *Chandra* counterpart from CSC 2.0. We find that  $\sim 96\%$  of the matching results from *XMM* and *Chandra* are identical, suggesting our multi-wavelength matching results are highly reliable.
8. We classify 4584 X-ray sources as an AGN. The classification is based on their optical spectral from SDSS, VIPERS, or VVDS (831); X-ray luminosity larger than  $3 \times 10^{42} \text{ erg s}^{-1} \text{ cm}^{-2}$  (1504); and large X-ray-to-optical and X-ray-to-NIR flux ratios (4479).

The *XMM*-SERVS-LSS source catalogs presented in this work will provide the first step in understanding the details of the relation between AGN evolution and environment over the largest cosmic structure different from the existing COSMOS datasets. In the near future, the inclusion of COSMOS, deeper *Chandra* survey in Boötes, and additional X-ray surveys of similar depth and high-quality multiwavelength coverages can further increase the AGN sample size and will

allow for a better understanding of the relation between different large-scale structures and AGN evolutions.

## ACKNOWLEDGMENTS

**To be completed** This work is based on observations taken by the 3D-HST Treasury Program (GO 12177 and 12328) with the NASA/ESA HST, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555.

Funding for PRIMUS is provided by NSF (AST-0607701, AST-0908246, AST-0908442, AST-0908354) and NASA (Spitzer-1356708, 08-ADP08-0019, NNX09AC95G).

NAOJ / HSC Collaboration

This paper uses data from the VIMOS Public Extragalactic Redshift Survey (VIPERS). VIPERS has been performed using the ESO Very Large Telescope, under the "Large Programme" 182.A-0886. The participating institutions and funding agencies are listed at <http://vipers.inaf.it>

This research uses data from the VIMOS VLT Deep Survey, obtained from the VVDS database operated by Cesam, Laboratoire d'Astrophysique de Marseille, France.

SDSS :

CSI :

**Table A.** Main X-ray catalog

XID (1)	RA (2)	DEC (3)	$\sigma_x$ (4)	FB_DET_ML (15)	FB_EXP (23)	FB_BKG (35)	FB_SCTS (51)	FB_FLUX (91)	HR (99)	$L_{2-10\text{keV}}$ (102)	LR_FLAG_RE (111)	LRA_CATALOG (115)	LRA_ZBEST (205)	LRA_ZSOURCE (206)	CLASS (208)
XMM00000	34.200218	-4.035255	1.44	19.0	59076.2	1.74	83.04	8.3e-15	-99	-99	False	SERVS	-99	-99	AGN
XMM00001	34.200713	-4.933734	1.45	63.0	61051.8	1.0	82.03	7.93e-15	-99	8.983e+43	True	SERVS	1.82	UDSz	AGN
XMM00002	34.201454	-5.556716	1.96	16.4	29731.6	0.8	29.64	5.31e-15	-99	2.46e+42	True	SERVS	0.459	VIPERS	AGN
XMM00003	34.201466	-4.499315	1.5	23.3	72553.8	1.76	72.37	5.32e-15	-99	1.37e+43	True	SERVS	0.959	PRIMUS	AGN
XMM00004	34.201949	-4.555523	0.93	316.8	87846.9	1.81	351.91	2.666e-14	-0.43	9.57e+42	True	SERVS	0.41	SDSS	AGN
XMM00005	34.202636	-5.690719	1.66	16.5	26430.1	1.01	52.23	1.301e-14	-99	1.6854e+44	True	CFHTLS	1.932	VIPERS	AGN
XMM00006	34.203276	-4.315289	1.55	29.2	107957.7	1.79	65.42	2.94e-15	-99	-99	True	SERVS	-99	-99	AGN
XMM00007	34.203748	-5.433790	1.77	11.3	78270.8	1.54	41.87	4.71e-15	-99	-99	True	VIDEO	-99	-99	AGN
XMM00008	34.203820	-4.595275	1.17	114.8	83485.0	1.49	168.25	1.214e-14	-0.48	2.05e+42	True	VIDEO	0.294	SDSS	AGN
XMM00009	34.204668	-5.378238	1.35	57.1	93769.9	1.4	101.92	6.8e-15	-99	-99	False	SERVS	-99	-99	AGN
XMM00010	34.204772	-4.520794	1.51	32.0	77664.5	1.51	71.54	5.7e-15	-99	-99	True	SERVS	-99	-99	AGN
XMM00011	34.206732	-4.469321	1.25	81.2	63040.8	1.74	132.09	1.579e-14	-99	2.61e+42	True	SERVS	0.291	SDSS	
XMM00012	34.207426	-4.585313	1.54	17.3	94203.8	1.85	67.16	4.22e-15	-99	6.8e+41	True	SERVS	0.289	PRIMUS	
XMM00013	34.208245	-5.295083	1.28	66.6	91655.9	1.22	124.63	7.18e-15	-0.32	-99	True	SERVS	-99	-99	AGN
XMM00014	34.209443	-4.012790	1.34	34.8	63642.5	1.92	106.5	9.3e-15	-99	-99	True	SERVS	-99	-99	AGN
XMM00015	34.209479	-4.028269	0.88	349.3	68619.6	1.94	426.86	4.119e-14	-0.12	1.2566e+44	True	SERVS	1.031	SDSS	AGN
XMM00016	34.209496	-4.421868	1.48	89.5	23041.3	0.22	76.5	2.309e-14	-99	0.0	True	SERVS	0.0	SDSS	
XMM00017	34.209799	-4.328703	1.12	84.6	115477.7	2.71	194.51	1.317e-14	-99	2.17e+42	True	SERVS	0.291	SDSS	AGN
XMM00018	34.209992	-4.563673	1.14	187.9	88839.2	1.41	181.45	1.38e-14	-0.25	-99	True	SERVS	-99	-99	AGN
XMM00019	34.210413	-3.890166	1.53	37.9	44730.1	0.95	68.48	7.55e-15	-99	-99	True	CFHTLS	-99	-99	AGN
XMM00020	34.210863	-5.410119	1.76	11.6	89369.8	1.53	43.01	2.63e-15	-99	-99	True	SERVS	-99	-99	AGN
XMM00021	34.211423	-5.194363	1.09	104.2	85216.5	1.93	207.87	1.356e-14	-0.36	1.58e+42	True	SERVS	0.249	SDSS	
XMM00022	34.211535	-3.892733	1.62	-99	-99	-99	-99	-99	-99	-99	False	HSC	-99	-99	
XMM00023	34.211690	-5.313163	1.44	34.3	96622.2	1.24	83.2	4.54e-15	-99	4.33e+42	True	SERVS	0.627	VIPERS	AGN
XMM00024	34.213193	-4.355747	1.64	11.1	100517.8	2.9	54.4	4.4e-15	-99	-99	True	SERVS	-99	-99	AGN
XMM00025	34.214208	-5.417995	1.75	10.7	88372.6	1.56	43.75	5.35e-15	-99	-99	True	SERVS	-99	-99	AGN
XMM00026	34.215117	-4.937388	1.82	13.5	64067.0	1.07	38.16	5.06e-15	-99	-99	True	SERVS	-99	-99	AGN
XMM00027	34.215205	-4.388650	1.73	16.5	67996.2	2.29	45.43	3.98e-15	-99	4.7e+41	True	SERVS	0.251	PRIMUS	
XMM00028	34.215240	-3.918303	1.62	15.1	33523.0	0.94	56.51	6.71e-15	-99	-99	True	SERVS	-99	-99	AGN
XMM00029	34.217465	-4.087917	1.48	22.2	82858.6	1.96	75.51	6.05e-15	-99	7.65e+42	True	SERVS	0.707	PRIMUS	AGN
XMM00030	34.217481	-5.601158	1.71	17.5	24629.1	1.0	47.25	1.061e-14	-99	7.024e+43	True	VIDEO	1.44	SDSS	AGN

The main X-ray source catalog with a selection of columns. Empty or null values are marked as -99. The numbers listed on the second row of this table is the column number of the full X-ray catalog with 208 columns. See Sec. 6 for a detailed descriptions of each column.

**Table B.** NWAY Bayesian matching results (To be updated)

XID	NWA_PANY	NWA_PI	NWA_SERVS_ID	NWA_SERVS_MAG1	NWA_VIDEO_ID	NWA_VIDEO_KSMAG	NWA_CFHT_ID	NWA_CFHT_IMAG	NWA_HSC_ID	NWA_HSC_IMAG
XMM00000	0.99	0.96	172727	20.5	644245983190	21.08	1114.031255	21.93	37484971320960508	22.04
XMM00001	0.94	1.0	-99	-99	644245946363	17.51	1123.223188	19.29	37484833882002514	19.22
XMM00002	1.0	0.99	174016	17.13	644245975018	17.08	1114.024092	17.85	-99	-99
XMM00003	0.99	0.6	172557	19.57	644245974364	19.35	1114.023481	20.58	37485108759912843	20.5
XMM00004	0.99	0.97	130124	18.41	644245971200	18.41	1114.021145	19.78	37484833882008818	19.85
XMM00005	0.99	0.98	-99	-99	-99	-99	1123.211260	20.41	37484692148084612	20.56
XMM00006	0.99	1.0	-99	-99	-99	-99	1123.212886	20.98	37484692148084849	21.09
XMM00007	1.0	0.99	159785	19.4	644245968189	20.16	1114.018595	21.27	37485108759911057	21.16
XMM00008	0.15	0.37	-99	-99	644245967227	19.97	-99	-99	-99	-99
XMM00009	0.94	1.0	-99	-99	644246423098	19.85	1114.015512	20.02	-99	-99
XMM00010	0.56	0.67	-99	-99	644245971580	21.66	1114.021242	22.87	37484971320960299	21.79

The NWAY matching results with a selection of columns. Empty or null values are marked as -99. See Sec. 6 for a detailed descriptions of each column.

XSERVS-LSS

## REFERENCES

- Aihara H., et al., 2017  
 Bradshaw E. J., et al., 2013, *Mon. Not. R. Astron. Soc.*, 433, 194  
 Brandt W. N., Alexander D. M., 2015, *Astron. Astrophys. Rev.*, 23, 1  
 Broos P. S., Feigelson E. D., Townsley L. K., Getman K. V., Wang J., Garmire G. P., Jiang Z., Tsuboi Y., 2007, *Astrophys. J. Suppl. Ser.*, 169, 353  
 Broos P. S., et al., 2011, *Astrophys. J. Suppl. Vol. 194, Issue 1, Artic. id. 2, 19 pp. (2011)*, 194  
 Brusa M., et al., 2007, *Astrophys. J. Suppl. Ser.*, 172, 353  
 Cappelluti N., et al., 2007, *Astrophys. J. Suppl. Ser.*, 172, 341  
 Cappelluti N., et al., 2009, *Astron. Astrophys.*, 497, 635  
 Chiappetti L., et al., 2005, *Astron. Astrophys.*, 439, 413  
 Coil A. L., et al., 2010, *Astrophys. Journal, Vol. 741, Issue 1, Artic. id. 8, 15 pp. (2011)*, 741  
 Evans I. N., et al., 2010, *Astrophys. J. Suppl. Ser.*, 189, 37  
 Fevre O. L., et al., 2013, *Astron. Astrophys.*, 559, A14  
 Garilli B., et al., 2014, *Astron. Astrophys.*, 562, A23  
 Hudelot P., et al., 2012, *VizieR On-line Data Cat. II/317. Orig. Publ. SPIE Conf. 2012*, 2317  
 Jarvis M. J., et al., 2013, *Mon. Not. R. Astron. Soc.*, 428, 1281  
 Kelson D. D., et al., 2014, *Astrophys. J.*, 783, 110  
 Kim M., Wilkes B. J., Kim D., Green P. J., Barkhouse W. A., Lee M. G., Silverman J. D., Tananbaum H. D., 2007, *Astrophys. J.*, 659, 29  
 LaMassa S. M., et al., 2016, *Astrophys. J.*, 817, 172  
 Liu Z., et al., 2016, *Mon. Not. R. Astron. Soc.*, 459, 1602  
 Lonsdale C. J., et al., 2003, *Publ. Astron. Soc. Pacific, Vol. 115, Issue 810, pp. 897-927.*, 115, 897  
 Luo B., et al., 2010, *Astrophys. J. Suppl. Ser.*, 187, 560  
 Luo B., et al., 2016, *Astrophys. J. Suppl. Ser.*, 228, 2  
 Marchesi S., et al., 2016, *Astrophys. J.*, 817, 34  
 Mauduit J. C., et al., 2012, *Publ. Astron. Soc. Pacific, Vol. 124, Issue 917, pp. 714 (2012)*, 124, 714  
 McLure R. J., et al., 2013, *Mon. Not. R. Astron. Soc.*, 428, 1088  
 Menzel M. L., et al., 2016, *Mon. Not. R. Astron. Soc.*, 457, 110  
 Momcheva I. G., et al., 2016, *Astrophys. J. Suppl. Ser.*, 225, 27  
 Pierre M., et al., 2016, *Astron. Astrophys.*, 592, A1  
 Pineau F.-X., et al., 2017, *Astron. Astrophys.*, 597, A89  
 Ranalli P., et al., 2013, *Astron. Astrophys.*, 555, A42  
 Ranalli P., et al., 2015, *Astron. Astrophys.*, 577, A121  
 Rosen S. R., et al., 2016, *Astron. Astrophys.*, 590, A1  
 Salvato M., et al., 2017  
 Skelton R. E., et al., 2014, *Astrophys. J. Suppl. Ser. Vol. 214, Issue 2, Artic. id. 24, 49 pp. (2014)*, 214  
 Sutherland W., Saunders W., 1992, *Mon. Not. R. Astron. Soc.*, 259, 413  
 Ueda Y., et al., 2008, *Astrophys. J. Suppl. Ser.*, 179, 124  
 Vaccari M., 2016, Proc. "The many Facet. extragalactic radio Surv. Towar. new Sci. challenges" (EXTRA-RADSUR2015). 20-23 Oct. 2015. Bol. Italy. Online <http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=267>, id.27  
 Watson M. G., et al., 2008, *Astron. Astrophys.*, 493, 339  
 Xue Y. Q., et al., 2011, *Astrophys. J. Suppl. Ser.*, 195, 10

This paper has been typeset from a TeX/LaTeX file prepared by the author.