



HAL
open science

Establishing the X-ray source detection strategy for eROSITA with simulations

Teng Liu, Andrea Merloni, Johan Comparat, Kirpal Nandra, Jeremy S. Sanders, Georg Lamer, Johannes Buchner, Tom Dwelly, Michael Freyberg, Adam Malyali, et al.

► To cite this version:

Teng Liu, Andrea Merloni, Johan Comparat, Kirpal Nandra, Jeremy S. Sanders, et al.. Establishing the X-ray source detection strategy for eROSITA with simulations. *Astronomy and Astrophysics - A&A*, 2022, 661, pp.A27. 10.1051/0004-6361/202141178 . hal-03285716

HAL Id: hal-03285716

<https://hal.science/hal-03285716>

Submitted on 22 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Establishing the X-ray source detection strategy for eROSITA with simulations

Teng Liu (刘腾)¹ , Andrea Merloni¹, Johan Comparat¹, Kirpal Nandra¹, Jeremy Sanders¹, Georg Lamer², Johannes Buchner¹, Tom Dwelly¹, Michael Freyberg¹, Adam Malyali¹, Antonis Georgakakis³, Mara Salvato¹, Hermann Brunner¹, Marcella Brusa^{4,5}, Matthias Klein¹, Vittorio Ghirardini¹, Nicolas Clerc⁶, Florian Pacaud⁷, Esra Bulbul¹, Ang Liu (刘昂)¹, Axel Schwoppe², Jan Robrade⁸, Jörn Wilms⁹, Thomas Dausser⁹, Miriam E. Ramos-Ceja¹, Thomas H. Reiprich⁷, Thomas Boller¹, and Julien Wolf¹

¹ Max-Planck-Institut für extraterrestrische Physik, Giessenbachstraße 1, 85748 Garching bei München, Germany
e-mail: liu@mpe.mpg.de

² Leibniz-Institut für Astrophysik, An der Sternwarte 16, 14482 Potsdam, Germany

³ Institute for Astronomy and Astrophysics, National Observatory of Athens, V. Paulou and I. Metaxa 11532, Greece

⁴ Dipartimento di Fisica e Astronomia, Università di Bologna, Via Piero Gobetti 93/2, 40129 Bologna, Italy

⁵ INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy

⁶ IRAP, Université de Toulouse, CNRS, UPS, CNES, 31028 Toulouse, France

⁷ Argelander-Institut für Astronomie, Rheinische Friedrich-Wilhelms-Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany

⁸ Hamburger Sternwarte, University of Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany

⁹ Dr. Karl Remeis-Sternwarte & Erlangen Centre for Astroparticle Physics, Sternwartstr. 7, 96049 Bamberg, Germany

Received 25 April 2021 / Accepted 23 January 2022

ABSTRACT

Context. The eROSITA X-ray telescope on board the Spectrum-Roentgen-Gamma satellite has started to detect new X-ray sources over the full sky at an unprecedented rate. Understanding the performance and selection function of the source detection is important for the subsequent scientific analysis of the eROSITA catalogs.

Aims. Through simulations, we test and optimize the eROSITA source detection procedures, and we characterize the detected catalog quantitatively.

Methods. Taking the eROSITA Final Equatorial-Depth Survey (eFEDS) as an example, we ran extensive photon-event simulations based on our best knowledge of the instrument characteristics, the background spectrum, and the population of astronomical X-ray sources. We introduce a method of analyzing source detection completeness, purity, and efficiency based on the origin of each photon.

Results. According to the source detection efficiency measured in the simulation, we chose a two-pronged strategy to build eROSITA X-ray catalogs, creating a main catalog using only the most sensitive band (0.2–2.3 keV) and an independent hard-band-selected catalog using multiband detection in a range up to 5 keV. Because our mock data are highly representative of the real eFEDS data, we used the mock catalogs to measure the completeness and purity of the eFEDS catalogs as a function of multiple parameters, such as detection likelihood, flux, and luminosity. These measurements provide a basis for choosing the eFEDS catalog selection thresholds. The mock catalogs (available with this paper) can be used to construct the selection function of active galactic nuclei and galaxy clusters. A direct comparison of the output and input mock catalogs also gives rise to a correction curve that converts the raw point-source flux distribution into the intrinsic number counts distribution.

Key words. surveys – catalogs – X-rays: galaxies – X-rays: galaxies: clusters – X-rays: diffuse background – galaxies: active

1. Introduction

Significant developments have been made in X-ray surveys in the past decades. More and more X-ray sources are detected, resolving an increasingly larger fraction of the cosmic X-ray background. With the largest grasp in the soft X-ray band of current X-ray imaging telescopes, the eROSITA telescope is detecting new X-ray sources at an unprecedented rate. It is expected to detect millions of active galactic nuclei (AGN) and more than 10 000 galaxy clusters in the ongoing eROSITA all-sky survey (eRASS; Merloni et al. 2012; Predehl et al. 2021). To maximize the impact of the survey, it is important to choose a source detection strategy that is appropriate and ideally optimized for the main scientific applications of the survey.

The eROSITA is extremely efficient in X-ray imaging surveys not only because of its large grasp, but also because of its scanning observation mode. During the eRASS surveys, it continuously scans the sky and covers the full sky every six months. Ahead of the four-year eRASS survey, we performed the eROSITA Final Equatorial-Depth Survey (eFEDS; Brunner et al. 2022), which was designed as a prototype survey of eRASS. Using raster-scanning mode, the eFEDS survey observes a 140 degree² field centered at RA 136°, Dec 1.5° in ~100 hours. With a scanning speed of ~13"/s, the field has a relatively uniform exposure depth of ~2.2 ks (0.2–2.3 keV vignettted depth ~1.2 ks), which is about 50% higher than the final depth of the four-year eRASS survey at the same position. The average point spread function (PSF) in the scanning mode

has a half-energy width of $\sim 26''$ at 1.5 keV (Predehl et al. 2021). The eFEDS survey is currently the largest continuous X-ray survey and highly representative of the final eRASS data in extragalactic regions. Therefore, we can investigate the eROSITA source detection through this survey.

The detection of astrophysical sources in imaging X-ray surveys is challenging because the ultimate sensitivity of these surveys generally probes the low count-rate regime, leading to high Poisson fluctuations. In addition, the spatial resolution of X-ray telescopes is often relatively high compared to the sky density of potential X-ray emitting sources. The eROSITA Science Analysis Software System (eSASS; version eSASS_users201009; Brunner et al. 2022) is employed in the source detection. It provides at least two methods of defining the source detection likelihood: PSF-fitting likelihood, based on maximum likelihood image fitting with the PSF model, and a Poissonian likelihood, based on aperture count extraction. The detection likelihood is always defined as corresponding to a probability of being false (probability = $\exp(-\text{likelihood})$). However, in practice, the false rate cannot be accurately predicted theoretically. The performance of source detection algorithms depends on instrumental characteristics, observing strategy, and background, combined with the nature of the target sources (e.g., brightness, spectral shape, extent, and number density). In dealing with these effects, every algorithm has adjustable parameters that need to be optimized. Realistically, the source detection problem is sufficiently complex that, to fully characterize an X-ray detection scheme or catalog in terms of completeness and purity, realistic simulations are needed. In this work, we simulate the eFEDS survey with two goals: 1) to investigate and optimize the eROSITA source detection strategy, and 2) to quantify the completeness and purity of the eFEDS catalogs.

Simulation tests have often been used in previous X-ray surveys (e.g., LaMassa et al. 2013; Liu et al. 2013) and were employed for eROSITA pre-launch in order to forecast the instrument capabilities for detecting galaxy clusters and AGN (Clerc et al. 2018). Typically, simulation results are analyzed at the catalog level, that is, by comparing the input and output catalogs on the basis of source positions and fluxes. In the case of deep X-ray surveys with relatively modest spatial resolution, the source detection process is not just a question of distinguishing real sources from background fluctuations. The blending of point sources and the unknown profile of extended sources introduce additional complexity and uncertainty, which are hard to quantify at the catalog level. To address these issues, we here analyze simulation results at the event level. We use the SIXTE software to simulate the X-ray event files. SIXTE is the official eROSITA end-to-end simulator (Dauser et al. 2019; provided by ECAP/Remeis observatory¹). It takes an input catalog to create photons, which are then propagated through the mirrors and incident on the detector. It uses the measured energy-dependent PSF and vignetting to simulate the mirrors as close as possible to the flying detector. The incident photons create a charge cloud and are then read out as single events and are then reconstructed. This approach allows correctly predicting the split patterns and also including detector effects such as pile-up (Dauser et al. 2019). Because of the nature of these simulations, the origin of each X-ray photon in the event list is known, encoded in the ID of the input source or ID of the background component. Based on the origin of each photon, we introduce a strategy of attributing detected sources to input sources, which could reveal

any potential source blending or misclassification (extended or unresolved). This allows us to analyze the properties of the detected sources at a more detailed level.

The content of the paper can be summarized as follows: in Sect. 2 we use the measurements of the eROSITA background based on the real eFEDS data to create a mock eFEDS dataset. In Sect. 3, we describe the method of characterizing the nature of detected sources and measure the purity of the eFEDS catalog in various manners, for instance, the fraction of spurious sources, the fraction of misclassified clusters, and the fraction of misclassified AGN. In Sect. 4, we investigate the selection function of AGN and clusters in terms of detected fraction as a function of source properties and the source detection efficiency in terms of completeness–contamination parametric curve. We optimize the eFEDS source detection according to the source detection efficiency. In Sect. 5, we test the construction of point-source number counts. The results are summarized in Sect. 6.

2. Simulating eFEDS data

2.1. Input

To simulate the eFEDS source detection, we began with the inputs we describe below.

1. Hardware characteristics: The currently available updated eROSITA calibration files were used, including the 2D PSF model version 190219v05, the vignetting model version 4.0, and the normalized single-pattern redistribution matrix file (RMF) version 20170725 (Dennerl et al. 2020; Brunner et al. 2022).
2. Observing strategy: We used the telescope attitude (at every instant) of the eFEDS cleaned event file (version c001) that was used to build the eFEDS catalogs (Brunner et al. 2022). As a result of the raster-scanning mode, the field has a sharp drop in exposure depth at the border, where it has not only a much lower exposure depth, but also a different vignetting and spatial resolution from the main field. We simulated the full eFEDS field, but the tests for catalog completeness and purity were made within the inner 90% area (126.6 degree²) region where the 0.2–2.3 keV vignettted exposure value is above 500s. In this region, the exposure depth of eFEDS is relatively flat. The number count analysis of the real eFEDS catalog is also limited to this region (Brunner et al. 2022).
3. Input sources and their background: As input sources, we have one cluster catalog and two point-source catalogs, one for AGN and one for stars. Based on the Universe N -body simulations for the Investigation of Theoretical models from galaxy surveys (UNIT; Chuang et al. 2019), Comparat et al. (2019, 2020) created a full-sky light cone for AGN and clusters, extending to redshift 6.1. We extracted 18 nonoverlapping regions of the eFEDS-field shape from the full-sky mock catalogs. Particularly for AGN, the adopted spectral model is composed of an absorbed power law with $\Gamma = 1.9$ plus three additional weak components (TBabs(p1cabs + constant*powerlaw + pextrav + zgauss) in Xspec), including a soft scattered power law with the same Γ , but 1% normalization, a cold reflection component with reflection scaling factor 1, and a narrow 6.4 keV Gaussian emission line with an equivalent width of 0.3 keV against an unabsorbed power law. The AGN population follows a realistic luminosity function and obscuring column density distributions as measured by previous X-ray surveys (Comparat et al. 2019). The soft fluxes of the input AGN catalog extend to much lower values $\sim 10^{-17}$ erg cm⁻² s⁻¹ (Comparat et al. 2019) than the source detection limit of eFEDS ($\sim 7 \times 10^{-15}$ erg cm⁻² s⁻¹).

¹ <https://www.sternwarte.uni-erlangen.de/research/sixte/>

The cluster fluxes and surface brightness profiles were constructed based on the dark matter halo properties as described in Comparat et al. (2020). We summarize the method here. Given a set of dark matter halo properties (mass, redshift, ellipticity, and offset parameter), we assigned an X-ray emissivity profile and image to each halo using a profile generator trained on a set of observed clusters. For the complete cluster population, the emissivity profiles reproduce the distribution and scatter as a function of scale as measured in Ghirardini et al. (2019). The nature of the core (cool vs. noncool) devised by the central part of the profile was mapped on the offset parameter following a tentative physical correlation (Seppi et al. 2021). Simulated clusters are isothermal. Compared to Comparat et al. (2020), we lower the mass limit to $M_{500c} = 10^{13} M_{\odot}$. M_{500c} is the halo mass inside the radius R_{500c} , where the mass density is 500 times the universe critical density. This low limit leads to a strong Malmquist bias on the X-ray luminosities of clusters, as discussed in Sect. 4.1.1 of Comparat et al. (2020). We performed an empirical correction to the luminosities in order to align the scaling relation between luminosity and stellar mass with that measured by Anderson et al. (2015). The obtained population reproduces the cluster X-ray luminosity function and number counts (see Sect. 2.3). A typical Galactic N_{H} of $3.26 \times 10^{20} \text{ cm}^{-2}$ was adopted for the X-ray spectra of the AGN and clusters. The stellar catalog follows the logN-logS observed in the eFEDS field.

As described below in Sect. 2.2, we measured the background spectra from the real eFEDS data and decomposed the vignettted and unvignettted components. They were simulated separately using the proper vignetting models. The mock and real data are compared in detail in Sect. 2.3.

4. Software settings: We used the same source detection pipeline as was used to build the real eFEDS catalogs (Brunner et al. 2022).

Based on these four inputs, we used sixte-2.6.2 to create mock event files. Each input source and each background component has a unique ID, and each simulated photon has a flag of the input ID.

2.2. eFEDS background spectra

Background plays a crucial role in the detection of faint sources. We considered the background as two components, a vignettted and an unvignettted component. The vignettted component corresponds to the X-ray photons from the sky, that is, diffuse Galactic X-ray emission and cosmic X-ray background, which are transmitted through the mirror before hitting the CCD. The unvignettted component is mainly due to high-energy particles, which hit the camera and generate secondary X-ray emissions inside the camera. Electronic noise also contributes to the unvignettted component.

In order to make the mock data as representative as possible of the real data, we measured the background from the eFEDS data (version c001) that were used to build the eFEDS catalogs (Brunner et al. 2022), adopting all the valid events (single, double, triple, and quadruple patterns) and all the seven telescopes. Liu et al. (2022b) extracted the spectra of all the eFEDS sources. We used their source-excluding regions and extracted background spectra from eight circular source-free regions with a radius of $20'$ as displayed in Fig. 1. The eight regions were chosen to be evenly distributed in the field and are representative of both the region in pure scanning-mode (regions 1, 3, 5, and 7) and of the region in which the pointing stays still for a long time (regions 2, 4, 6, and 8). These two types of regions have different scanning strategies and thus different vignetting.

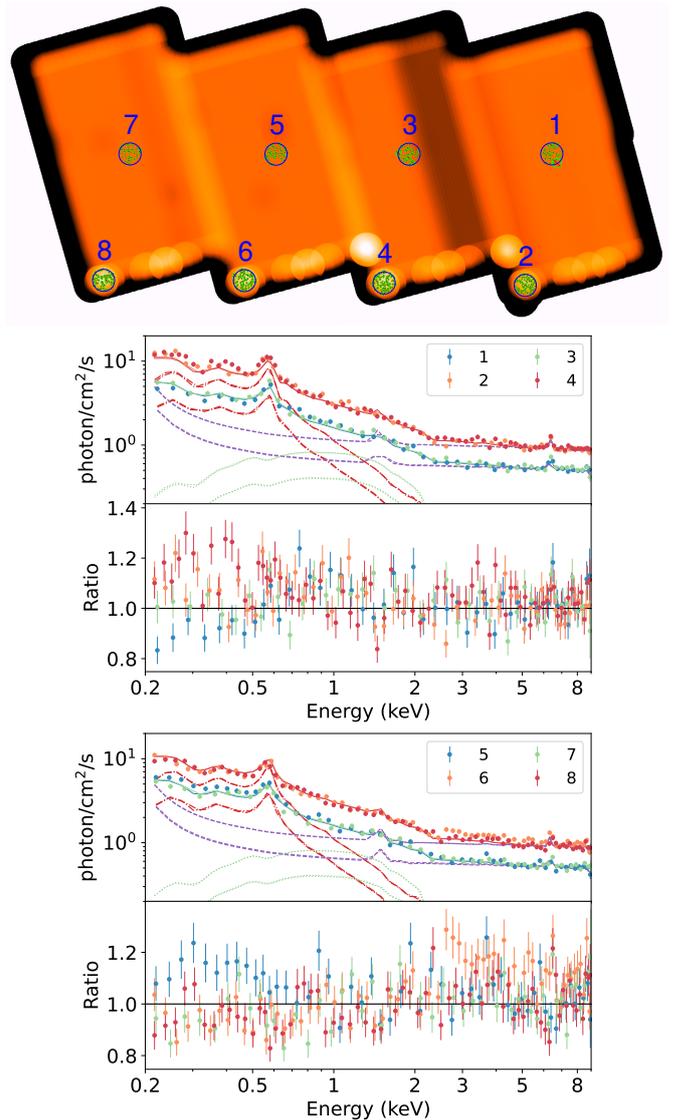


Fig. 1. Eight background extraction regions with radius $20'$ overlaid on the 0.2–2.3 keV exposure map of eFEDS (top). To avoid overlapping, the eight spectra are displayed separately in terms of the data (points), the model (solid lines), and the data-to-model ratio. The purple, green, and red lines indicate the averaged particle background, the mock cosmic X-ray background, and the diffuse Galactic X-ray background, respectively. All the lines correspond to the same model with the same flux for the eight spectra. The count rate in the spectra was calculated based on the total exposure time, which is longer than the effective exposure depth in survey mode because the extraction region is not always covered by the moving FOV. The difference (by a factor of ~ 2) between regions 1, 3, 5, and 7 and regions 2, 4, 6, and 8 reflects the different effective exposure depths. We did correct for the count rate because the uncorrected data look better for the representation.

Because our input AGN catalog has a very low flux limit ($\sim 10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1}$), some faint sources do not contribute any signal at all, while a large number of them contribute a few photons that are too few to make the source detectable. These photons from undetectable AGN compose a mock cosmic X-ray background. This is not necessarily the same as the real cosmic X-ray background (Brandt & Hasinger 2005; Brandt & Alexander 2015), as it depends on the assumption of the AGN

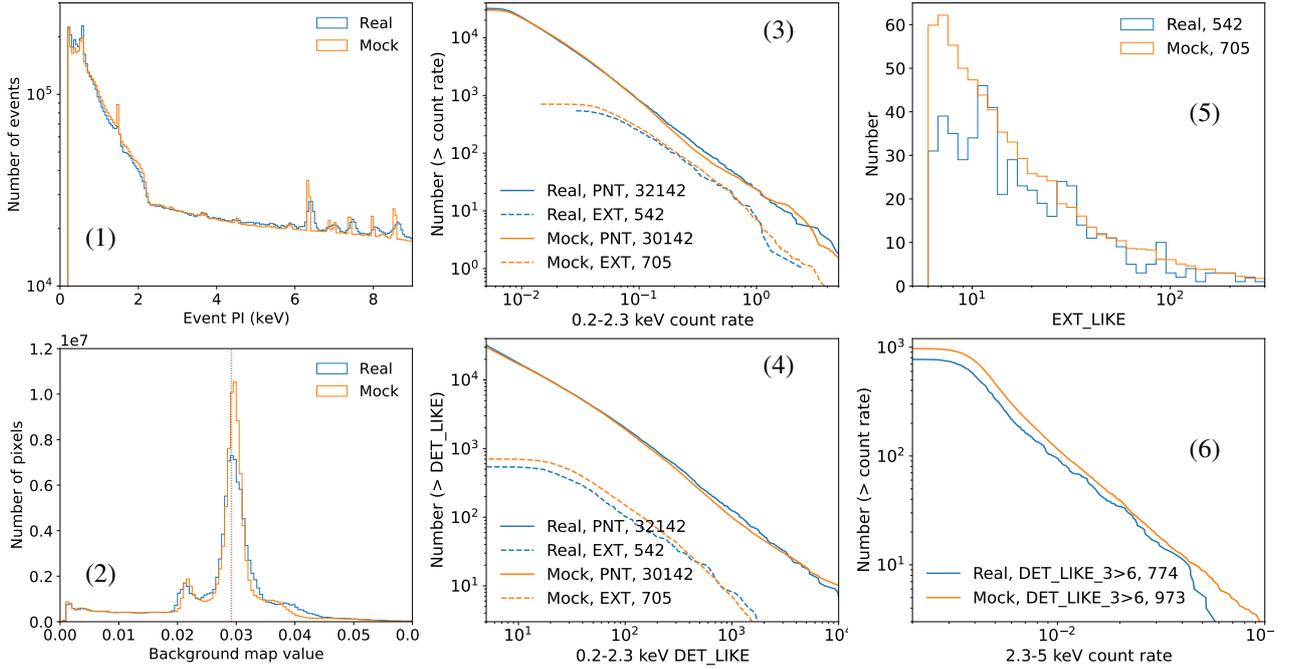


Fig. 2. Comparisons between the real (blue) and the mock (orange) eFEDS data. *Panel 1* compares the event PI (photon energy) distributions. *Panel 2* compares the distributions of the 0.2–2.3 keV background map value (counts/pixel) in each pixel. The dotted vertical line corresponds to the median values of the two data sets, which are identical. *Panels 3–4* display the total number of sources in the 0.2–2.3 keV single-band-detected catalog above a given count rate (panel 3) or above a given detection likelihood (DET_LIKE; panel 4) for point sources (“PNT”; EXT_LIKE=0; solid lines) and extended sources (“EXT”; EXT_LIKE>0; dashed lines), respectively. *Panel 5* displays the extent likelihood (EXT_LIKE) distribution of the single-band-detected extended sources. *Panel 6* displays the 2.3–5 keV (band 3) count rate of sources that are detected in this band (DET_LIKE_3 > 6) from the three-band detection. The number of sources is printed in the label of each entry.

luminosity function adopted when creating the mock AGN catalog (Comparat et al. 2020). However, in the background spectral model described above, the real cosmic X-ray background is already included, thus this mock cosmic X-ray background is a duplication. Therefore, we measured this background component in order to exclude it from the background spectral model. Because we simulated the source and background events separately, we created a special version of mock data with source events alone. We performed a similar background extraction from the pure-source-event data as was done above for the real data, that is, we ran the source detection on it and then extracted a background spectrum from a source-free region. This background component from the undetectable input AGN can be fit with a partially covering absorbed power law with $\Gamma = 1.41$, $N_{\text{H}} = 3.5 \times 10^{21} \text{ cm}^{-2}$, and a covering factor of 52%. This spectral component was only measured in order to be excluded from our background model, so that this component is not duplicated in the simulation.

Then we fit each of the real background spectra with three components, as shown in Fig. 1. The first component was the cosmic X-ray component measured above from the undetectable input AGN. The second component was the particle background. We adopted the phenomenological spectral shape of the eROSITA Filter Wheel Closed (FWC) data and normalized it to the eFEDS background spectrum in the 4.5–9 keV band. In this band, the spectrum is dominated by particle background and can be well fit with the FWC spectral model. Then we fixed the parameters of the above two components and fit the residual signal, which is the foreground diffuse X-ray emission, using a phenomenological model composed of a power law, an APEC (Smith et al. 2001) plasma model, and a Gaussian component.

The eROSITA background has been found to be relatively constant in time and thus flat in the scanned field (Brunner et al. 2022; Predehl et al. 2021). In Fig. 1, we also test the background variability by comparing the eight background spectra with the averaged model. The background is highly constant in the hard band. It is only relatively higher in region 6 by a factor of <20%. In the very soft band below 0.5 keV, the background is relatively more variable, which might be caused by the spatial variability of the Galactic emission and/or the time variability of light leak (Predehl et al. 2021). Overall, the variability has a moderate amplitude and its spatial scales are much larger than the PSF size, thus the impact on source detection is small. We therefore assumed a flat background in the simulation. The particle background components measured from the eight regions were averaged and converted into a SIXTE particle-background file (available with the SIXTE package). The foreground diffuse X-ray background components from the eight regions were also averaged, creating an X-ray background spectral model. We ran SIXTE to simulate the particle background events and X-ray background events separately, and we finally merged them into the source signal event file.

2.3. Mock versus real data

We succeeded in making the mock data as representative as possible to the real eFEDS data in many aspects, that is, X-ray background, particle background, and distributions of source fluxes and spectral shapes. The event energy distribution of the mock data is therefore similar to that of the real data, as displayed in panel 1 of Fig. 2. Compared with the real data,

the background emission lines in the simulation are not sufficiently smoothed because the RMF adopted by SIXTE is slightly different from the real RMF. This minor difference in energy response has no impact on the broadband source detection.

In panel 2 of Fig. 2, we compare the (exposure-uncorrected) background maps measured from the real and mock data using the same method (Brunner et al. 2022). The distribution of the background counts per pixel shows a strong peak with two tails on both sides. The low-value tail corresponds to the field border, where the exposure depth drops sharply. The high-value tail corresponds to a small portion of regions in which the depth is much higher than average (e.g., regions 2, 4, 6, and 8 in Fig. 1). As we treated the X-ray background and particle background separately using proper vignetting models in the simulation, the background maps are highly similar. They have identical median values, but the real data show a slightly larger scatter. This is because of both the background variation in the real data and the averaging of the mock data among the 18 realizations.

We used eSASS to detect sources in the mock eFEDS data. The source detection is described in detail in Brunner et al. (2022). It is done in two steps, first detecting a preliminary catalog using the eSASS task `erbox` and then performing PSF fitting for each source in the preliminary catalog using the task `erlmdet`. The PSF fitting measures a detection likelihood `DET_LIKE` for each source, which corresponds to the probability of the source being spurious (background fluctuation) in terms of $\text{probability} = \exp(-\text{likelihood})$. By applying a threshold on `DET_LIKE`, the final catalog is selected. In addition to `DET_LIKE`, `erlmdet` also measures an extent likelihood `EXT_LIKE`, which corresponds to the probability of the source being point-like in terms of $\text{probability} = \exp(-\text{likelihood})$. Point sources or extended sources can be selected by thresholding `EXT_LIKE`. As we adopted a minimum `EXT_LIKE` parameter of 6 for `erlmdet`, all the sources with `EXT_LIKE` < 6 were set as `EXT_LIKE` = 0. We performed the PSF-fitting source detection adopting two different sets of energy bands, a single-band PSF fitting in 0.2–2.3 keV, and a three-band (1: 0.2–0.6, 2: 0.6–2.3, and 3: 2.3–5 keV) simultaneous PSF fitting. In the latter case, `erlmdet` measured a detection likelihood `DET_LIKEn` for each band n (1, 2, or 3) and a summary detection likelihood `DET_LIKE0`. Performing post hoc aperture photometry for each source in the catalog using the eSASS task `apetool`, we also measured the probability of the source being background fluctuation based on the source and background counts in the aperture. Similarly, this probability can also be converted into an aperture-photometry-based likelihood `APE_LIKE` ($= -\ln \text{probability}$).

As displayed in Fig. 2, the single-band detected point sources have similar distributions in the mock and real data in the soft band (panels 3 and 4). The mock data have ~25% more hard sources than the real data (panel 6), likely because our assumed spectral model is slightly harder than the average spectral shape of the eFEDS sources. The detected extended sources from the real and mock data also show similar distributions, but ~30% more extended sources are detected in the mock data. Particularly, these mock extended sources show a strong peak at `EXT_LIKE` < 11 that does not exist in the real catalog. This is because of the uncertainty in the cosmological and astrophysical model used to create the mock cluster catalog (Comparat et al. 2020). As the clusters are drastically outnumbered by AGN and have a low number density, they do not impact the detection of point sources.

3. Characterizing the catalog

3.1. Characterizing each source

Based on the flag of the input-source ID on each photon, we examined the input-output source association in terms of the following four characteristics:

1. Does a detected source have an input counterpart? The association was made within a circular aperture of a radius of 20'' at the position of the detected source. This radius corresponds to the 60% enclosed-energy fraction (EEF) radius of the PSF, which is the optimized aperture photometry size that leads to the best efficiency in distinguishing between source signal and background fluctuation (see Sect. 5.1). The input source that contributes the largest number of photons in this aperture was considered the input counterpart (with the `ID_Any`; a negative value means that the counterpart was not found) of the target source. As the fluxes of our input catalog extend far lower than the detection limit, none or only a few photons are captured by the camera for a large number of input sources. We only considered input sources for which at least 3 photons were captured at any position. Still, a spuriously detected source might coincidentally be attributed to a very faint input source. To exclude these coincident associations, we extracted all the X-ray and particle background photons into a background image and smoothed it to increase its signal-to-noise ratio (S/N), and then we defined a lower limit of the source aperture counts according to the local aperture background counts as its 2σ Poissonian upper limit (97.725% point of the Poisson distribution). An input source was considered a meaningful counterpart only if its aperture counts exceeded this limit.

2. Does a detected source have a secondary input counterpart? If multiple input sources meeting the above requirements were found within the 20'' aperture of a detected source, the input source that contributed with the second-highest number of photons was considered the secondary counterpart (`ID_Any2`; a negative value means that no secondary counterpart was found).

3. Is an input source contaminated by another source? If a second input source contributes at least 3 photons within 60'' (the PSF-fitting radius used in the eFEDS source detection) of an input source, and if its number of photons is larger than the square root of the target photon counts, we considered this second source a contamination of the target source and saved it as `ID_contam` (a negative value means that it does not exist).

4. Unique input-output association: One input source might be falsely detected as multiple sources, resulting in duplicated `ID_Any`. In these cases, we selected the input source that contributed the largest number of photons in the 20'' aperture of the detected source and considered this input source a unique counterpart (`ID_Uniq`). If the `ID_Any` of a detected source was not a unique counterpart, we took the secondary input counterpart (`ID_Any2`) if it existed and was not assigned to other detected sources. With `ID_Uniq`, a one-to-one association between the input and output sources was built.

3.2. Classifying detected sources

Based on the source-matching characteristics defined above, we divided the detected sources into five classes as follows. We plot their distributions in Fig. 3 in five colors.

1. Primary counterpart of a point source (PNT, in Fig. 3): A detected source is classified as the primary counterpart of an input point source if it has a unique input counterpart (`ID_Uniq`) of an AGN or star.

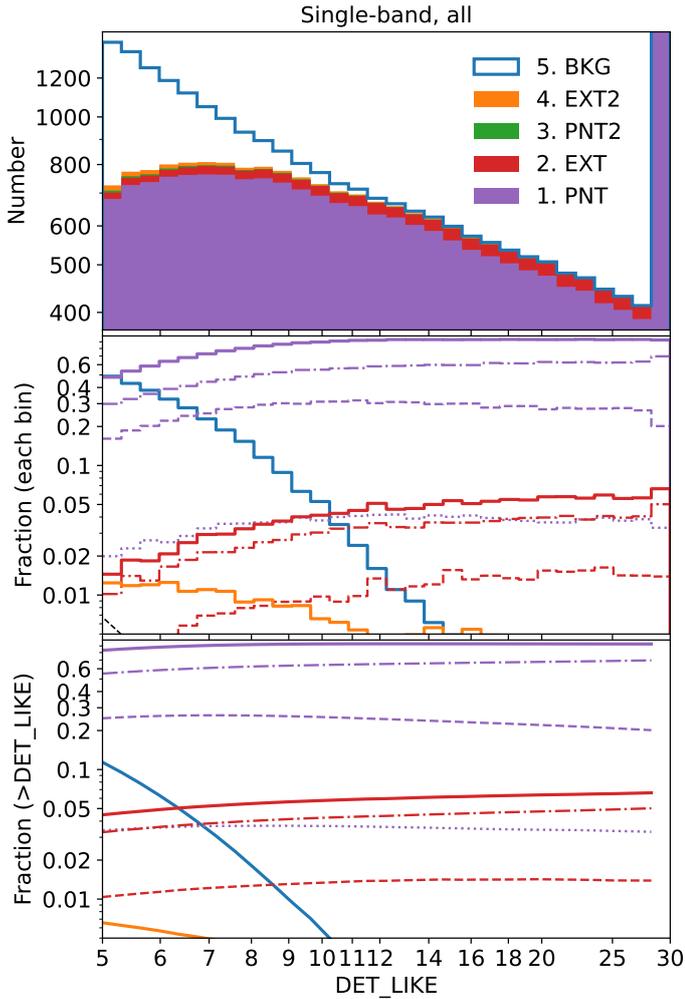


Fig. 3. Distributions of all the single-band-detected sources as a function of detection likelihood. Purple, red, green, orange, and blue correspond to the five classes (1~5) of detected sources. The *top*, *middle*, and *bottom panels* display stacked histograms, differential fractions, and cumulative fractions, respectively. The dot-dashed, dashed, and dotted lines in purple and red indicate the three subclasses of no contamination, point-source contamination, and extended-source contamination, respectively. See Sect. 3.2 for details. The dashed black line indicates $\text{probability}=\exp(-\text{likelihood})$. It is too low and almost drops out of the plotting scope of the *middle panel*.

2. Primary counterpart of an extended source (EXT): A source is classified as the primary counterpart of an input cluster if it has a unique input counterpart (ID_Uniq) of a cluster.

3. Secondary counterpart of a point source (PNT2): A source with an input counterpart (ID_Any) of an AGN or star but without a unique counterpart (ID_Uniq < 0) is classified as the secondary counterpart of an input point source. In these cases, a primary counterpart (class 1 or 2) for the input point source is already detected at the correct (brightest) location; the detected source contains point-source signals that correspond to a fragment of an input point source in the outer wing of the PSF.

4. Secondary counterpart of an extended source (EXT2): A source with an input counterpart (ID_Any) of a cluster but without a unique counterpart (ID_Uniq < 0) is classified as the secondary counterpart of an input cluster. In these cases, a primary counterpart (class 1 or 2) is already detected for the input cluster at the correct (brightest) location; the detected

source contains cluster photons that represent a substructure or fluctuation in the cluster.

5. Background fluctuation (BKG): A source without any input counterpart (ID_Any < 0) is classified as a spurious source due to background fluctuation.

The secondary input counterpart of a detected source ID_Any2 is useless here in characterizing the detected sources. It is useful, on the other hand, in characterizing the input sources. If an input source has no detected primary counterpart, it might be considered as the secondary input counterpart of any detected source. In this case, the signal of this input source is detected but not as an independent source, that is, it is blended with another brighter source. Alternatively, another indicator of source blending is ID_contam, which is defined in a larger region (60'' instead of 20'') around each input source. This indicator is more sensitive in selecting the cases of blended sources in the sense that all the photons involved in the PSF fitting (within 60'') were taken into account. For input sources or primary detected counterparts (class 1 or 2) of input sources, we furthermore divided them into three subclasses: (1) without contamination (ID_contam < 0), (2) contaminated by a point source, and (3) contaminated by a cluster.

As illustrated in Fig. 3, which displays the whole single-band-detected sample as a function of DET_LIKE, we can display the distributions of the five classes of detected sources in three terms: (1) a stacked histogram as a function of a particular source property, (2) a differential fraction (in each bin of the histogram) of each class in the selected sample, and (3) a cumulative fraction of each class in the selected sample, calculated using subsamples above any given value of the concerned source property. The stacked histogram was used to choose the bins and avoid measurements of fractions with only a few sources. The differential fraction distributions are helpful in understanding the source detection performance in various aspects and accordingly in making adjustments, for example, whether an adjustment is helpful in suppressing spurious detections at low detection likelihood (DET_LIKE), distinguishing point and extended sources at low extent likelihood (EXT_LIKE), or resolving blended sources at high fluxes. The cumulative fraction distributions were used to characterize a selected sample quantitatively in terms of various types of false rates (or purity).

Given any sample selection criteria, we can display the distributions of various classes of sources as a function of any source property, for instance, detection likelihood, extent likelihood, or source flux, as illustrated in Figs. 4 and 5, and as discussed Sects. 3.3 and 3.4. Figure 4 displays the single-band-detected catalog, and Fig. 5 displays the three-band-detected catalog. We provide the single-band-detected and the three-band-detected mock catalogs here. They include the classification information (Appendix B). With this information, the fraction of any class in the eFEDS catalog under the sample selection criteria required by a particular scientific goal can be measured.

3.3. Spurious sources

By definition, the source detection likelihood, either measured through PSF fitting (DET_LIKE) or through Poisson tests of the aperture photon counts (APE_LIKE), reflects the probability of a source being background fluctuation, which equals $\exp(-\text{likelihood})$. However, these definitions assume an ideal situation of a single source with pure Poissonian fluctuation, but do not account for any additional uncertainty or bias introduced during background estimation, PSF fitting, or aperture photometry, not to mention potential uncertainties of the facility hardware,

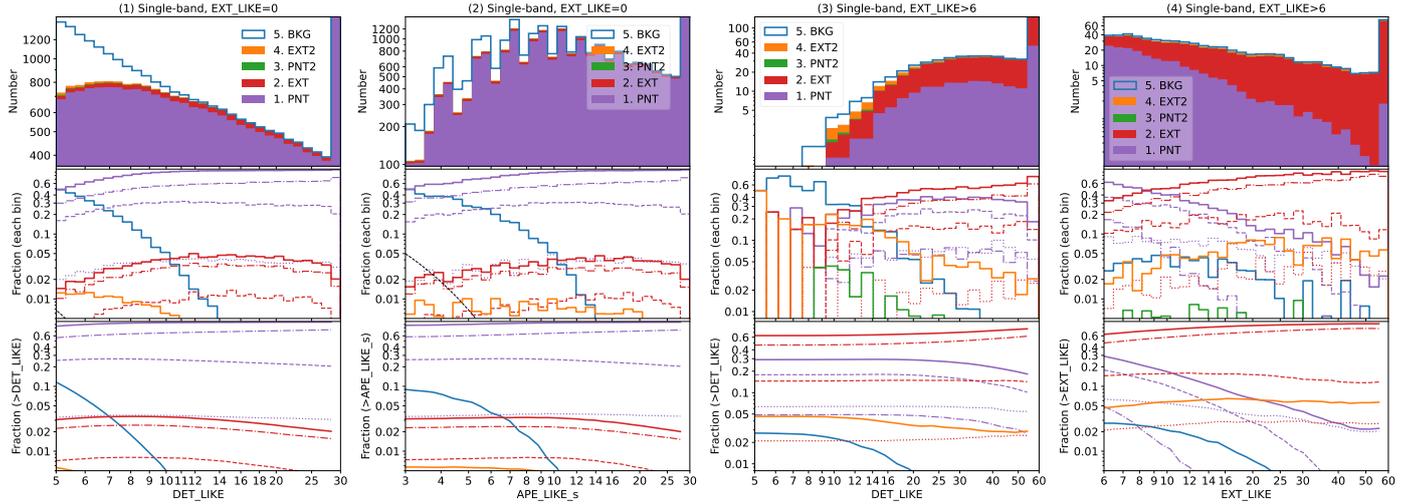


Fig. 4. Same as Fig. 3, but for different sample selections and/or different variables. The panel number and selection rules are printed in the figure titles, i.e., *panels 1–2* for point sources ($\text{EXT_LIKE}=0$) and *panels 3–4* for extended sources ($\text{EXT_LIKE}>6$), both of which are selected from the single-band-detected catalog. The *four panels* display the selected subsamples as a function of (1) single-band detection likelihood (DET_LIKE), (2) 0.5–2 keV aperture-photometry-based likelihood (APE_LIKE_s), (3) DET_LIKE , and (4) extent likelihood (EXT_LIKE). The dashed black lines in the middle subpanel of *panels 1–2* correspond to $\text{probability}=\exp(-\text{likelihood})$.

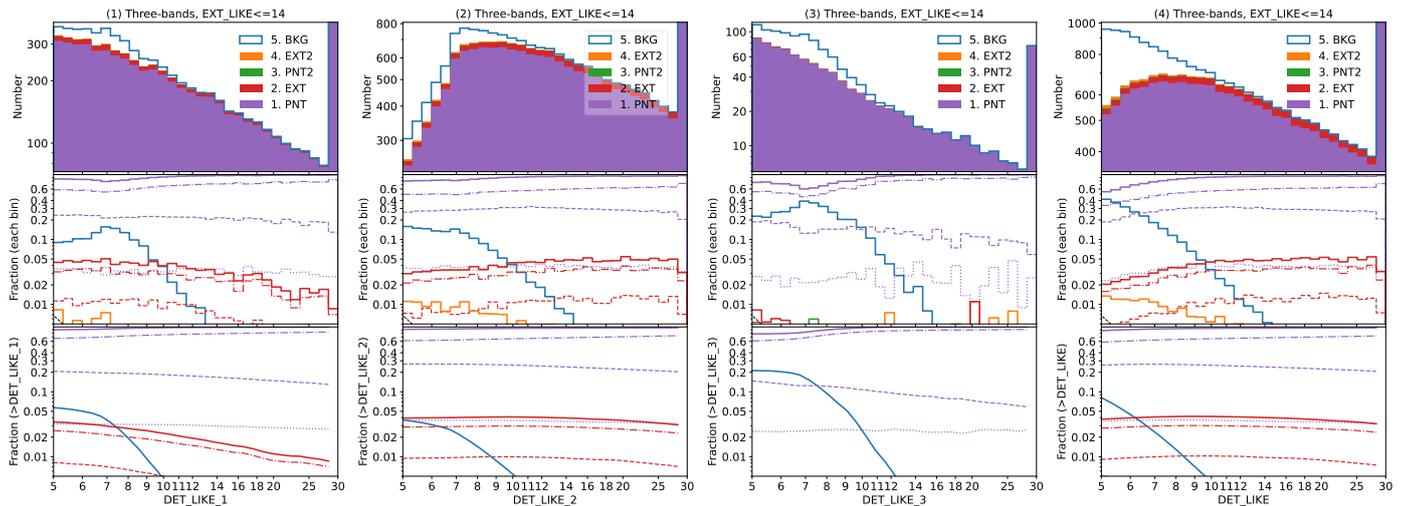


Fig. 5. Same as Fig. 4, but for the three-band-detected sources with $\text{EXT_LIKE}\leq 14$ and plotted as a function of the detection likelihood DET_LIKE_n , where n indicates the energy bands 1, 2, 3, or 0. In *panels 1–3*, the DET_LIKE_n corresponds to the individual band detection likelihood in band 1: 0.2–0.6, band 2: 0.6–2.3, and band 3: 2.3–5 keV. In *panel 4*, the DET_LIKE_0 corresponds to the three-band-summary likelihood.

calibration, and even software numerical issues. As displayed in panels 1 and 2 of Fig. 4, which display the distribution of the single-band-detected point sources, the measured spurious fractions (blue lines) are much higher than expected from the likelihood definition.

As suggested by Liu et al. (2020), a relatively low source detection likelihood should be adopted in detecting the candidates (not necessarily the final catalog) with PSF fitting, so that many faint but potentially interesting sources can be detected, potential cases of blended faint sources can be checked by multiple PSF fitting, and faint sources can be effectively masked out when measuring the properties of nearby sources. We adopted a threshold of $\text{DET_LIKE}>5$ in eFEDS source detections. As shown in Fig. 3, this threshold corresponds to a spurious fraction of 11.5% in the single-band-detected catalog, which is too high to be used for most scientific works. A further selection of subsamples is needed according to particular scientific goals.

To select a sample with relatively balanced completeness and purity, $\text{DET_LIKE}>6$ (for the main eFEDS catalog; Brunner et al. 2022) might be adopted, for instance, which corresponds to a spurious fraction of 6.3%. Finally, to select a cleaner sample, a higher threshold such as $\text{DET_LIKE}>8$ would be needed (1.8% spurious).

Figure 5 displays the likelihood distributions of the three-band-detected catalog. The samples selected from the three-band detection based on the 0.2–0.6 and 0.6–2.3 keV individual band likelihoods seem to have relatively lower spurious fractions (panels 1 and 2 in Fig. 5) than the single-band detection. However, this is only because they result from two-pass selections, that is, first requiring the summary likelihood $\text{DET_LIKE}_0>5$ and then the individual band likelihood >5 , and does not necessarily indicate a higher selection efficiency in the three-band detection. Even with the two-pass selection, the sample selected with the 2.3–5 keV individual band

likelihood still has a much higher spurious fraction because of the relatively small effective area and relatively high background in the hard band of eROSITA. Selecting sources with $\text{DET_LIKE}_3 > 10$ and $\text{EXT_LIKE} < 14$ (for the hard eFEDS sample; Nandra et al., in prep.) leads to a spurious fraction of 2.5%. When $\text{DET_LIKE}_0 > 5$ is adopted, the spurious fraction appears lower in the three-band detection than in the single-band detection. This is because when summing the individual-band likelihoods ($\text{DET_LIKE}_1, \text{DET_LIKE}_2, \text{DET_LIKE}_3$) to calculate DET_LIKE_0 , the additional degrees of freedom used in the summing reduce the likelihood of a source. Therefore, $\text{DET_LIKE}_0 > 5$ is a stricter rule in the three-band detection than in the single-band detection (see also the discussion in Sect. 4.2.)

3.4. Blending and misclassification

We have so far only discussed the spurious fraction (or false rate) introduced in the detection. When classification is involved, “false rate” could also refer to the fraction of misclassified sources. For a deep survey without high spatial resolution, it is hard to distinguish between a compact cluster and a point source, between blended point sources and an individual cluster, or between a point source in a cluster and the substructure of the cluster. The photon-flag-based input-output association allows us to investigate these cases. As shown in panels 1 and 2 in Fig. 4, the single-band-detected point-source catalog ($\text{EXT_LIKE}=0$ and $\text{DET_LIKE}>5$) contains $\sim 3\%$ clusters (class 2). It can be considered as a prediction of the fraction of clusters in the real eFEDS point-source catalog, which is only an approximation because the distributions of flux and brightness profile of clusters and groups are highly uncertain in the faint and compact regime. The extended source catalog selected with $\text{EXT_LIKE}>6$ also contains a significant fraction (29.1%) of AGN and stars, as shown in panels 3 and 4 in Fig. 4.

Another problem beyond source detection is source blending. As shown in panel 1 in Fig. 4, of the 84.6% genuine point sources (class 1) in the single-band-detected point-source catalog, 30% (25.0% of the whole point-source catalog) are contaminated by nearby point sources and 4% (3.3% of the whole catalog) are contaminated by nearby clusters. The genuine clusters also contain a significant fraction (23%) that is contaminated by nearby point sources (Fig. 4 panel 3). As shown in panel 3 in Fig. 4, the 29.1% genuine point sources (class 1) in the single-band-detected extended-source catalog include 61.2% (17.8% of the whole extended-source catalog) that are contaminated by nearby point sources and 22.0% (6.4% of the whole catalog) that are contaminated by nearby clusters. The much higher fractions of blended cases in the misclassified point sources indicate that blending of nearby sources is a main reason for this misclassification. Source blending causes not only misclassifications, it also results in sample incompleteness (fewer sources). Unless all the blended sources are fitted simultaneously with proper models, which is not usually the case, the source flux might be overestimated because of the contamination (see also the discussion in Sect. 5.2).

Panels 3 and 4 in Fig. 4 display the single-band-detected extended sources as a function of detection likelihood DET_LIKE and extent likelihood EXT_LIKE . Point sources are misclassified as extended (class 1), mostly because of source blending, as discussed above. This occurs at any detection likelihood (or brightness). The extent likelihood of these false clusters is significantly lower than that of genuine clusters (class 2), however. These false clusters included only a few separated (nonblended) point sources (dash-dotted purple lines) that are concentrate at

low EXT_LIKE ; there are more cases of blended point sources (dashed purple lines) and of cluster-contaminated point sources (dotted purple lines), and they have largely different EXT_LIKE distributions. The blended point sources mostly result in false clusters at low EXT_LIKE . At high EXT_LIKE (e.g., > 20), most of the misclassified point sources have underlying cluster emission. Because we identified the input counterpart of detected sources using the photons within a small radius of $20''$, in the cases of point sources inside clusters, the point sources with more concentrated signals are more easily considered to be the primary counterpart than clusters whose signals are diffusely distributed. These cases of cluster-contaminated AGN can also be considered correctly identified clusters, but with an AGN inside.

In order to evaluate the spurious fraction in the real eFEDS extended source catalog, we predicted the number of spurious clusters in the single-band-detected catalog. We defined spurious clusters as detected sources that are not attributed to any input clusters (either ID_Any or ID_Any2). In the eFEDS 90% area region, we detect 196.7 spurious clusters with $\text{EXT_LIKE}>6$. However, this value must be higher than the real eFEDS data because the detected extended sources from the mock data show a strong peak at the lowest EXT_LIKE (< 11) that does not exist in the real data, as shown in Fig. 2. At $\text{EXT_LIKE}>12$ and $\text{DET_LIKE}>12$, the simulation predicts 42.9 spurious clusters. At $\text{EXT_LIKE}>20$ and $\text{DET_LIKE}>20$, the simulation predicts 9.4 spurious clusters. These values are more reliable. Another issue to point out about the simulation-predicted spurious clusters is that some genuine cluster detectable in the optical band might be considered spurious because its X-ray emission was not significant enough in the small $20''$ aperture.

4. Source detection efficiency

4.1. Source detection completeness

The source detection completeness is defined as the detected fraction of sources in the input catalog. We can also call it selection function or sensitivity. According to various scientific goals, the completeness may be evaluated in various manners, for example, as the completeness in a particular flux range or above a flux threshold, the completeness according a particular sample selection filter (e.g., hard-band selection, extent selection), and as the completeness for a particular input population (e.g., obscured AGN). A few examples are shown in this section. All of them can easily be drawn using the mock input catalogs that we present with this paper.

In the left panel in Fig. 6, we display a few levels of AGN completeness as a function of source flux. At the first level, an AGN is detected as the primary counterpart of a source in the point-source catalog, that is, it is detected and correctly classified as a point source ($\text{EXT_LIKE}=0$). At the second level, an AGN is detected as the primary counterpart regardless of the extent classification, so that the AGN misclassified as clusters are also included. This means an increase of $\sim 1\%$ in the completeness at fluxes above 10^{-14} erg cm $^{-2}$ s $^{-1}$. At the third level, an AGN is detected either as the primary or as the secondary counterpart, so that the cases of blended AGN are included. This means a further $1\% \sim 2\%$ increase. Similarly, we plot the cluster completeness in the right panel of Fig. 6. At the first level, the fraction of clusters that are detected and correctly classified is not high (only exceeds 80% at flux $> 10^{-13.1}$ erg cm $^{-2}$ s $^{-1}$). At the second level, however, the completeness is much higher when misclassified sources are included, indicating that the main problem is classification and not detection. At the third level, a small fraction ($2\% \sim 3\%$) of blended cases are also included. There are

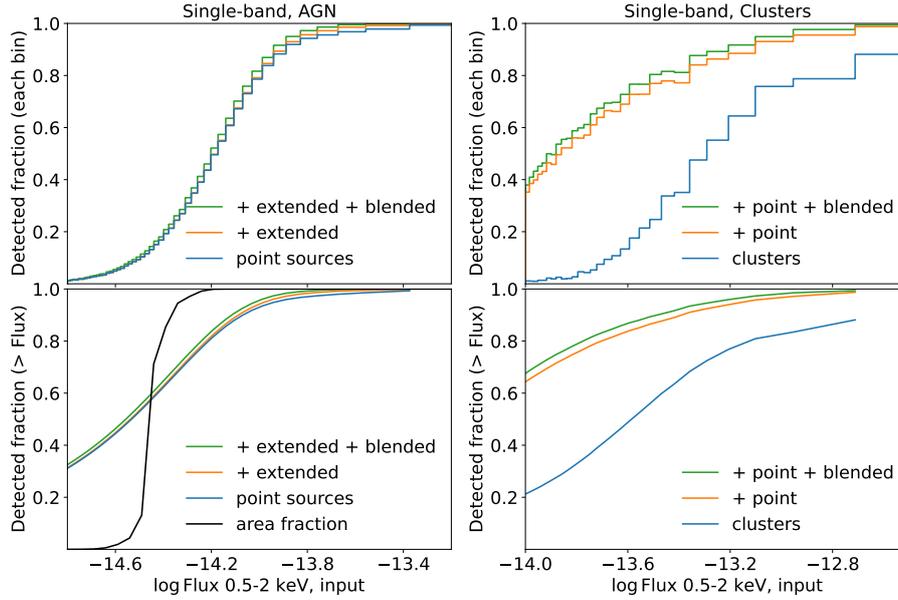


Fig. 6. Completeness of AGN (*left*) and completeness of clusters (*right*) in the single-band-detected catalog as a function of the input 0.5–2 keV flux in differential (*top*) and cumulative manners (*bottom*). The blue lines indicate sources that are detected (as ID_Uniq) and correctly classified as point sources (*in the left panel*) or as extended sources (*in the right panel*). The orange lines indicate sources that are detected (as ID_Uniq) regardless of classification, so that extended sources are also included in the *left panel* and point sources are also included in the *right panel*. The green lines indicate sources that are detected as either primary (ID_Uniq) or secondary (ID_Any2) counterparts, so that the cases of blended sources are included. The black line in the *left panel* shows the sensitivity curve of the single-band-detected catalog calculated using `ersensmap` in terms of fractional area at a given flux limit.

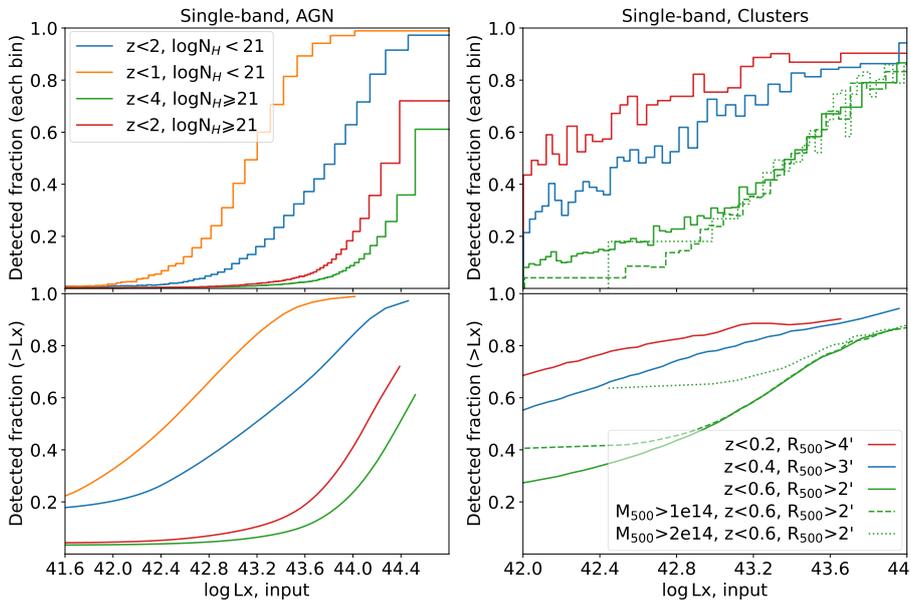


Fig. 7. Completeness of AGN (*left*) and clusters (*right*) in the single-band-detected catalog as a function of the intrinsic 0.5–2 keV luminosity in differential (*top*) and cumulative manners (*bottom*). Only the sources that are detected as primary counterparts (ID_Uniq) and are correctly classified as point sources (for AGN; *left*) or extended sources (for clusters; *right*) are considered. The *left panel* displays four input AGN subsamples selected with different redshift and N_H thresholds. The *right panel* displays input cluster subsamples selected with different thresholds of redshift, apparent scale R_{500} , and halo mass M_{500} (in M_\odot).

46 clusters (in 18 realizations) with fluxes $>10^{-13}$ erg cm $^{-2}$ s $^{-1}$ that are not detected (either ID_Any or ID_Any2) because they have relatively larger scales and thus lower surface brightness than the other clusters.

In Fig. 7, we display the completeness of a few AGN and cluster subsamples as a function of the 0.5–2 keV luminosity. Here we define completeness at the first level, that is,

detected and correctly classified. For AGN, we selected a few subsamples with different redshift and N_H ranges. The huge differences between them clearly show the selection bias against high- z and obscured sources. For clusters, we selected a few subsamples with different ranges of redshift and R_{500} apparent size, which are correlated. The huge differences clearly show the selection bias against high- z and compact sources. We also

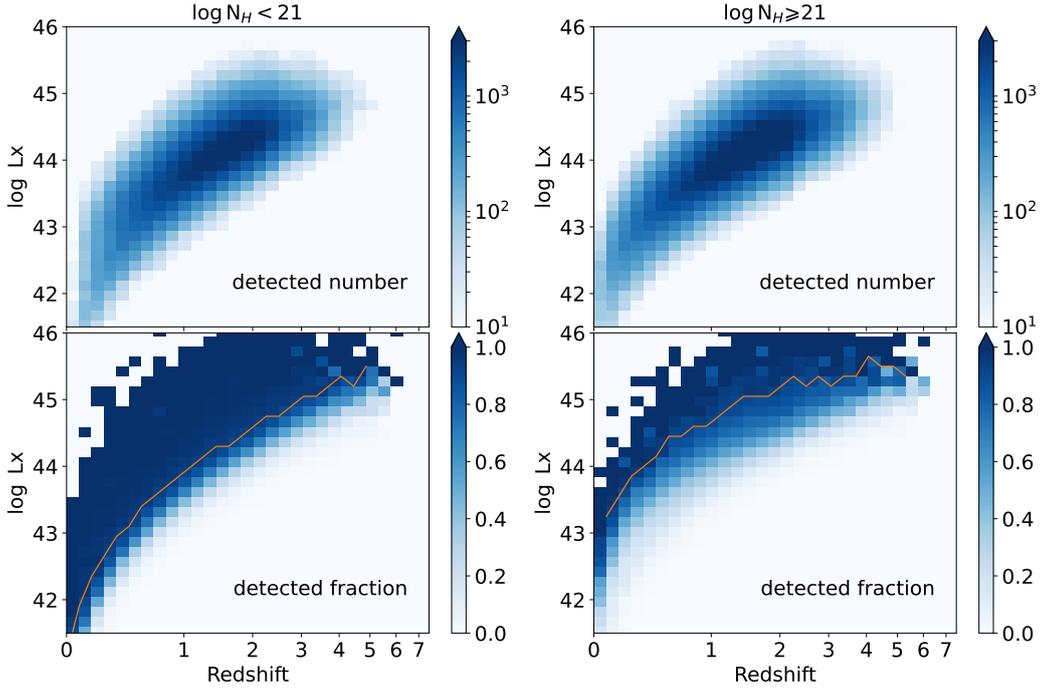


Fig. 8. Source numbers (*top*) and the detection completeness (*bottom*) in a grid of 0.5–2 keV intrinsic luminosity L_X and redshift z for unobscured ($\log N_H < 21$, *left panel*) and obscured ($\log N_H \geq 21$, *right panel*) AGN. The overlaid orange lines indicate 90% completeness.

compared the completeness above a few mass (M_{500}) thresholds among the sources with $z < 0.6$ and $R_{500} > 2'$. As shown in the upper panel, halo mass does not directly impact cluster detection. Higher-mass cluster samples have a higher completeness, as shown in the lower panel, because the cluster mass is strongly correlated with the X-ray luminosity.

Figure 7 only illustrates the AGN and cluster selection functions roughly. To quantify the selection functions accurately, we need to measure the detected fraction in finer parameter bins. Taking AGN, for example, we display the completeness distributions for obscured and unobscured AGN in the space of intrinsic luminosity and redshift in Fig. 8. The large sample size allows us to plot these 2D distributions easily for AGN through a simple binning. To quantify the cluster selection function in narrow parameter bins, more sophisticated sampling methods are needed, for example, through Gaussian process (see Liu et al. 2022a). We provide the mock AGN and cluster catalogs with the information of input-output association (Appendix B). With these catalogs, the selection function may be quantified as needed by specific scientific goals.

Using the eSASS task `ersensmap`, we calculated the flux limit corresponding to the point-source detection likelihood threshold ($\text{DET_LIKE}=5$) and thus the sky coverage area curve as a function of flux limit. When this curve is normalized to a total area of 1 (solid black line in Fig. 6), it also predicts a detectable fraction that only reflects the detectable flux limit across the whole field, however. At high fluxes, the actual detected fraction (sensitivity curve) derived from simulation is lower than that of the `ersensmap` flux limit because a source above the detectable flux limit might still be missed because of the fluctuation and measurement uncertainty in the source and background or because of blending with nearby sources. At low fluxes, the simulation-derived sensitivity curve extends far lower than the flux limit because fluctuation might also make sources below the flux limit detectable in some cases (see also the discussion in Sect. 5.2).

4.2. Source detection efficiency

The essential procedure of source detection is source selection from candidates based on the likelihood of each candidate. A low likelihood threshold leads to high completeness (or sensitivity). However, it also leads to a lower purity. A source detection operation has a high efficiency if it results in high completeness and high purity.

For the eROSITA source detection, we performed a source selection based on the PSF-fitting likelihood (`DET_LIKE`). In the case of multiband PSF fitting, the combined likelihood (`DET_LIKE_0`) was used. All the settings that were adopted in the whole detection procedure, for example, choice of energy bands, estimation of background, and PSF-fitting region size, could affect the final measurement of `DET_LIKE`. Optimizing the source detection is just optimizing the `DET_LIKE` measurement with the aim that a threshold on it leads to the best efficiency. In this section, we compare different settings by means of the completeness–contamination parametric curves as a function of `DET_LIKE`, where completeness can be measured as the detected fraction in a given range of input flux (as displayed in Fig. 6) and contamination can be measured as the fraction of spurious sources in the selected sample (blue cumulative fraction distribution displayed in Fig. 3).

We tested four sets of energy bands for the source detection and describe them below.

1. Single-band detection: We ran source detection in a single 0.2–2.3 keV band, which is the most sensitive band of eROSITA. It spans from the lowest energy to the turning point of the eROSITA effective area curve, above which the effective area is much lower.

2. Three-band detection: To cope with the wide range of spectral shapes of all types of AGN and stars, we ran the PSF fitting simultaneously in three bands: in 0.2–0.6, 0.6–2.3, and 2.3–5 keV. In this way, we did not miss the sources that are significant only in the soft 0.2–0.6 keV band or in the hard 2.3–5 keV band.

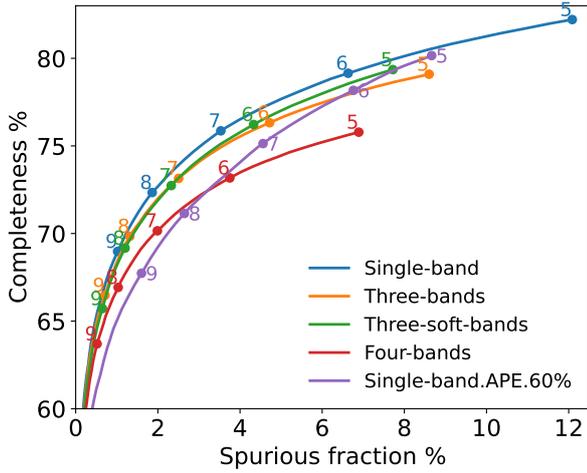


Fig. 9. Completeness–contamination parametric curves as a function of PSF-fitting detection likelihood `DET_LIKE` measured from the four sets of energy bands (blue, orange, green, and red lines). The purple line indicates the selection based on the aperture Poissonian likelihood measured in the 0.2–2.3 keV band within 60% EEf. The points corresponding to the serial of `DET_LIKE` values are marked on the lines.

3. Three soft band detection: In principle, dividing a broad band into multiple narrow bands facilitates detecting sources with a wide range of spectral shapes. However, it might cause additional complexities in combining the detection likelihoods from each individual band. To test this issue, we divided the 0.2–2.3 keV band into three narrow bands: 0.2–0.6, 0.6–1.1, and 1.1–2.3 keV.

4. Four-band detection: eROSITA could detect hard X-ray photons up to 8 keV but with a relatively low sensitivity above 2.3 keV. We have added the 2.3–5 keV band to the three-band detection. In addition to these three bands, we further added the ultra-hard band 5–8 keV to the four-band detection in order to detect ultra-hard sources.

Figure 9 displays the completeness–contamination curves corresponding to these four sets of energy bands. Here the completeness is defined as the detected fraction above an input-flux limit of $10^{-14.2} \text{ erg cm}^{-2} \text{ s}^{-1}$, and the spurious fraction corresponds to the sources due to background fluctuation. A curve with relatively higher completeness and lower contamination indicates a higher source detection efficiency. Although the hard band ($>2.3 \text{ keV}$) provides valuable information, its small effective area and high background result in a negative impact on the source detection efficiency. When the hard band between 2.3 and 5 keV (orange line in the figure) is included, it is not as efficient as using only the single 0.2–2.3 keV band, where eROSITA has the largest effective area. Including the ultra-hard band above 5 keV (red line) has a significant negative impact on the efficiency and thus should be avoided.

Dividing the single band into three narrow bands (green line) could in principle help detect sources with extreme spectral shapes, which make them only visible in the very soft or very hard band. However, such cases are rare in the simulation and in reality, thus it does not make any improvement in the sense of the overall efficiency. Based on the same data (in the same band), these two detections show similar efficiency curves resulting from significantly different `DET_LIKE` measurements. The three soft band likelihood is relatively lower and thus leads to relatively lower completeness at a given likelihood threshold. This is because the likelihood value of a source is reduced when the individual band likelihoods are combined into a summary value.

In addition to the choice of energy bands, we also tested the impact of the essential parameter of PSF fitting, the extraction radius. A radius of 15 pixels (60'') was adopted when the real eFEDS catalog was detected (Brunner et al. 2022). With a given background, a point source has an optimized radius in which the S/N is maximized. Reducing it will cause too much loss of source signal, and enlarging it will include too much background noise. As we discuss in Sect. 5.1, through investigation of aperture photometry, we find an optimized aperture radius of 60% EEf ($\sim 20''$) for the 0.2–2.3 keV band. Liu et al. (2022b) also measured an optimized source spectrum extraction radius that maximized the S/N of each eFEDS source in the 0.2–8 keV band using the eSASS task `srctool`. The median radius is 28''. Therefore, the adopted 60'' PSF-fitting radius is sufficiently large. We performed the source detection procedure by changing the radius to a lower value of 12 pixels and a higher value of 18 pixels. Then we compared the three cases using the completeness–contamination curve as done above. Although the smaller radius leads to slightly higher likelihoods because the S/N is higher, the efficiency is almost identical in the three cases. In other words, the reduced PSF-fitting radius leads to both higher completeness and higher contamination. If completeness is taken as the main figure of merit of the catalog and when spurious sources are allowed to be eliminated through multiband follow-up, a smaller PSF-fitting radius might be adopted in order to detect more faint point sources.

We also tested whether excluding the field of view (FOV) border outside a radius of 180 pixels in the single-band detection improves the detection efficiency. Here the spatial resolution is poor, the background is relatively high, and the calibration is relatively uncertain. Excluding the border region reduces both source signal and background. By comparing the completeness–contamination curve, we did not find any improvement, indicating that at least in the 0.2–2.3 keV band, the border region does not contribute negatively to the whole data set.

In addition, we tested the idea of repeating the PSF fitting using the PSF-fitting selected catalog as input instead of using the dirty (with many spurious sources) preliminary catalog. By comparing the completeness–contamination curve, we found no improvement either.

4.3. Classification efficiency

As discussed in Sect. 4.1, the main problem for galaxy clusters is source classification, that is, distinguishing between point and extended sources, and not detection. In this section, we investigate the classification efficiency through the completeness–contamination parametric curve as a function of extent likelihood `EXT_LIKE` instead of detection likelihood `DET_LIKE`. Here, completeness is defined as the fraction of detected clusters with input 0.5–2 keV fluxes above $5 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$, and the spurious fraction corresponds to the extended sources that are not associated with any input clusters (either `ID_Any` or `ID_Any2`). This completeness–contamination curve reflects the combined efficiency of source detection and classification. Because most of the clusters are significantly detected (panel 3 in Fig. 4), the main factor is classification. As displayed in Fig. 10, the efficiency is highly dependent on the PSF-fitting radius. The larger the region involved in the PSF fitting, the more efficiently clusters and point sources can be distinguished. Comparison of the three cases at the same levels of `EXT_LIKE` thresholds shows that the PSF-fitting radius impacts the completeness more than the contamination. This is because point sources that are misclassified as extended sources (mostly blended, discussed in Sect. 3.4)

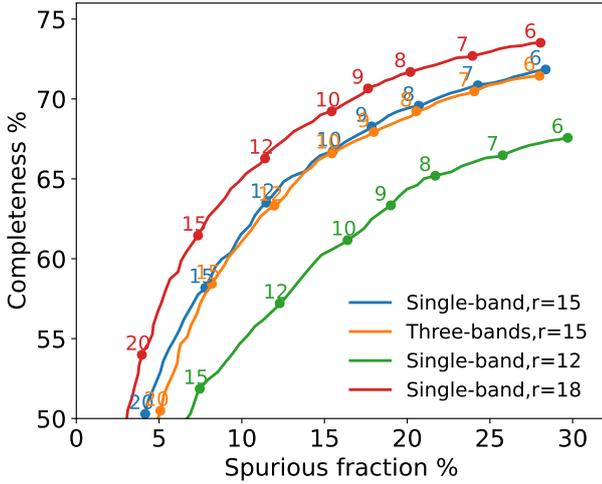


Fig. 10. Completeness–contamination parametric curves as a function of extent likelihood EXT_LIKE measured from single-band PSF fitting adopting three different radii (orange: 12 pixels; blue: 15 pixels; red: 18 pixels) and from three-band PSF fitting with a radius of 15 pixels (orange). The points corresponding to the serial of discrete EXT_LIKE values are marked on the lines.

have small scales not much larger than the PSF HEW. Some compact clusters are hard to be distinguished from point sources unless the PSF-fitting radius is extended to a large scale at which the PSF drops to extremely low values and thus the difference between PSF and the cluster profile can be revealed. We recommend using a larger PSF-fitting radius or performing post hoc analysis within a larger radius in future eROSITA surveys to improve the completeness of the galaxy clusters.

We also compare the single-band detection with the three-band detection in Fig. 10. Adding the 2.3–5 keV band is not helpful for the detection and classification of clusters because the cluster emission is more prominent in the soft band.

5. AGN number counts

5.1. Aperture-photometry-based likelihood

In addition to the PSF-fitting likelihood DET_LIKE , the eSASS task `apetool` also measures a likelihood for each source by comparing the aperture source photon counts with the Poisson distribution of the aperture background counts. This likelihood can also be used for source selection. Figure 9 compares the selection efficiencies based on the 0.2–2.3 keV PSF-fitting likelihood and the 0.2–2.3 keV aperture Poissonian likelihood (APE_LIKE). We only ran aperture photometry for the sources selected with the PSF fitting ($DET_LIKE \geq 5$), so that the APE_LIKE selection is in fact a two-step filtering. When a low APE_LIKE threshold of 5 is adopted, this two-step selection shows a similar completeness and purity as the PSF-fitting likelihood selection because the $DET_LIKE \geq 5$ preselection plays the main role. At a high likelihood (e.g., >8), where the impact of the $DET_LIKE \geq 5$ preselection becomes minor, the aperture likelihood selection shows a relatively lower efficiency. Compared with the aperture likelihood based on source and background aperture counts, making use of additional knowledge about the source image profile and the PSF model therefore improves the efficiency of source detection. We used the PSF-fitting likelihood for the source detection and only used the aperture photometry results to recover the number counts of point sources.

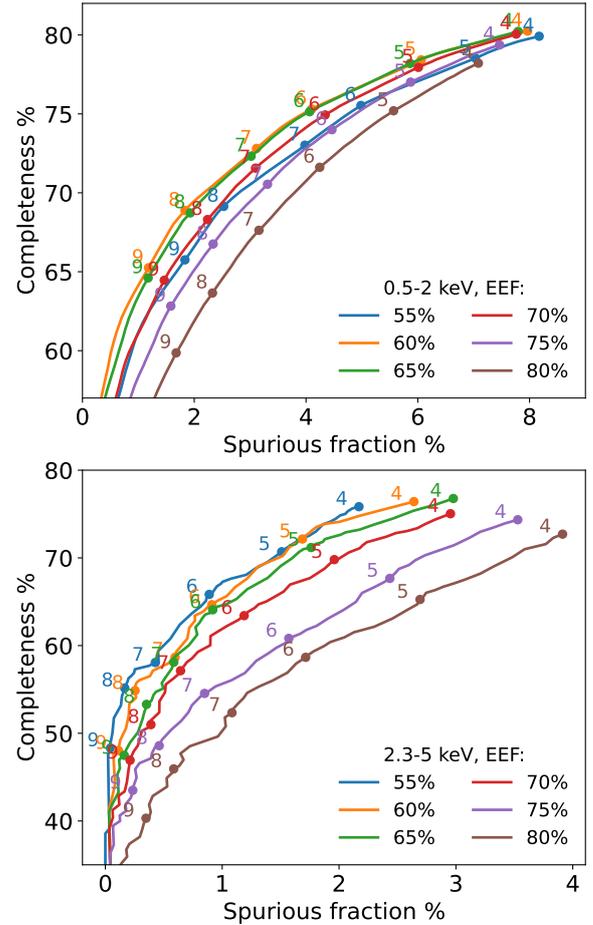


Fig. 11. Completeness–contamination parametric curves as a function of aperture Poissonian likelihood APE_LIKE measured within a serial of aperture radius (in terms of EEF) in the 0.5–2 keV band (*top*) and in the 2.3–5 keV band (*bottom*). The points corresponding to the serial of discrete APE_LIKE values are marked on the lines.

The essential parameter in aperture photometry is the aperture radius. We ran `apetool` with radii of EEFs of 55%, 60%, 65%, 70%, 75%, and 80% in the soft 0.5–2 keV band and the hard 2.3–5 keV band, and compare the completeness–contamination parametric curves as a function of the APE_LIKE in Fig. 11. In the soft band, an EEF of 60% or 65% leads to the best efficiency in distinguishing source signal from background. This aperture can be adopted as the optimized size for aperture photometry. In the hard band, although the PSF has a wider shape, we find that the most efficient aperture (55% or 60%) is smaller than that in the soft band. This is because of the higher noise in the hard band: a smaller region around the PSF core leads to a higher S/N inside it. Based on these tests, we suggest adopting 60% EEF, which corresponds to $\sim 20''$, in the aperture photometry for eFEDS.

5.2. AGN number counts

As an extragalactic field, eFEDS has a majority of AGN in the X-ray sources (Liu et al. 2022b). From the single-band-detected sources located inside the 90%-area region, we selected an AGN catalog as significantly detected ($DET_LIKE \geq 8$) point sources ($EXT_LIKE = 0$) that are not attributed to any input stars (according to ID_Any). We constructed the cumulative distribution of their 0.5–2 keV fluxes measured by forced PSF fitting and divided it by the observed area (126.6 degree²). These raw

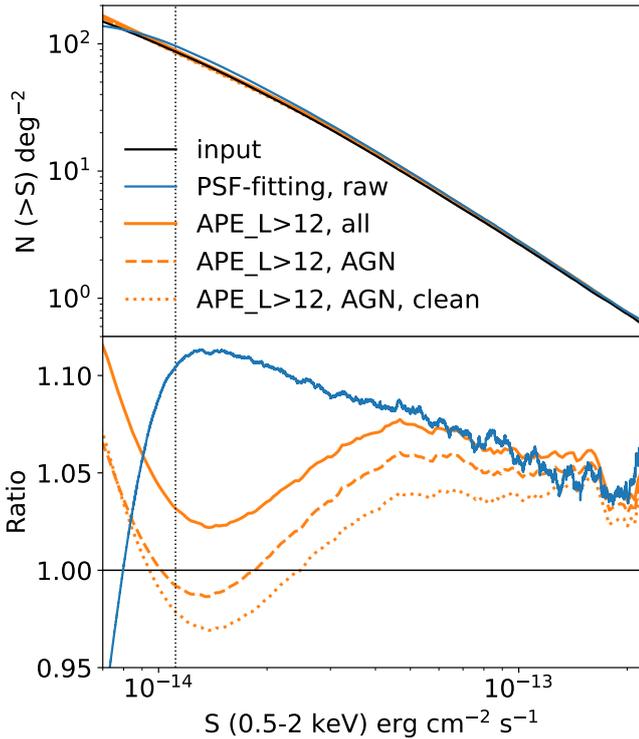


Fig. 12. Input (solid black line) and output number counts of AGN (*upper panel*) and the output-to-input ratios (*lower panel*). The blue lines correspond to the raw cumulative distribution of 0.5–2 keV fluxes measured by forced PSF fitting divided by the total area. The solid orange line shows the number counts we constructed based on the aperture Poissonian likelihood APE_LIKE using all the sources with $APE_LIKE > 12$. When only the genuine AGN instead of all the sources (including misclassified clusters) is used, the result is displayed as the dashed orange line. After a further flux correction when constructing the number counts of genuine AGN, the result is displayed as the dotted orange line. The vertical dotted line ($1.1 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$) corresponds to a sensitivity of 50%.

number counts per degree² are compared with the input AGN number counts in Fig. 12. Their ratios as displayed in the lower panel of Fig. 12 can be used to convert the raw distribution of the real eFEDS catalog into its intrinsic number count distribution. The output catalog is significantly overpopulated. This has three reasons: (1) catalog contamination, especially from galaxy clusters that are misclassified as AGN, (2) source flux overestimation caused by blended nearby sources, and (3) Eddington bias due to Poisson fluctuation. These factors are so strong that they lead to significantly more output sources than the input after compensating for source detection incompleteness (Fig. 6).

Then we tested the method we used to construct the point source number counts for the real eFEDS catalog (Brunner et al. 2022). We measured an APE_LIKE for each source using `apetool`. Selecting sources above an APE_LIKE threshold (e.g., 12), we stacked them to construct the number counts considering each source as a random variable following a probability density distribution of flux rather than a single count (Georgakakis et al. 2008). In this way, the probability of a source having a flux below the detection limit was taken into account and the Eddington bias was corrected for. As displayed in Fig. 12, this method (solid orange line) largely eliminates the bump of the raw distribution (blue line) near the detection limit and leads to a number count distribution that is more consistent with the input, although still higher than it by a few percent.

We listed three reasons for the overpopulated output catalog above. As the third problem (Eddington bias) has already been addressed by the method based on aperture photometry, the main reasons for the number count overestimation are the first two, that is, contamination and source blending. These two issues are impossible to address for real data; they can only be quantified through simulation. We explain these two factors through the following two experiments. First, when constructing the number counts, we selected only the sources that are attributed to genuine AGN (according to ID_Uniq). The spurious AGN due to compact clusters, fragments of AGN or cluster wing region, or background fluctuations were excluded. Above a likelihood threshold of $APE_LIKE > 12$, the main source of spurious AGN is the contamination of clusters (Fig. 4). As displayed in Fig. 12, excluding this contamination improves the derived number counts significantly. Second, we corrected the source flux for the blending of nearby sources. When a detected source has a secondary counterpart (ID_Any2) and this secondary counterpart is not identified as the unique counterpart (ID_Uniq) of any detected source, we calculated a flux correction factor for each of these sources as the ratio of the counts of the primary counterpart and the total counts of the two sources because we know the photon counts of both of them in the $20''$ circular region (approximately the same size for aperture photometry) of the detected source. As displayed in Fig. 12, when this flux correction is applied to the relevant sources in the subsample of genuine AGN, the derived number counts are further reduced, confirming that source flux enlargement caused by blending causes an overestimation of number counts. It is impossible to recover the input number counts precisely because of the additional uncertainty in the flux measurement. The input sources have a variety of spectral shapes, while for the output sources, a fixed spectral shape was assumed when we converted the count rate into flux.

The simulation provides a correction curve (blue line in the lower panel of Fig. 12) that can be used to convert the raw distribution of the real catalog into the intrinsic number counts and to predict that the intrinsic number counts derived in this way are lower than the counts derived using the method based on aperture photometry. This prediction has been confirmed with real eFEDS data (Brunner et al. 2022). It does not indicate any problem with the method based on aperture photometry itself because the information of contamination and blending is invisible for the real data and can only be revealed through simulation. Therefore, we recommend taking these factors into account only when an accurate measurement of AGN luminosity function is needed.

6. Discussions and conclusions

6.1. eFEDS source detection strategy

X-ray source detection can be considered as a two-step procedure: first, a search for source candidates, and second, a down-selection of these candidates to exclude those with a high false-positive rate. The detected sources are often characterized either simultaneously in the down-selection (e.g., Hasinger et al. 1994; Liu et al. 2020) or separately in an additional process (e.g., Liu et al. 2013), including brightness measurement and classification of point or extended source. The goal of the first step is to find possible candidates in as complete a way as possible. This can be accomplished using various algorithms such as the sliding-box detection algorithm (Deponte & Primini 1993;

Calderwood et al. 2001), the Mexican hat wavelet detection algorithm (Freeman et al. 2002), or Voronoi tessellation plus a friends-of-friends detection algorithm (Ebeling & Wiedenmann 1993; Liu et al. 2013). The sample selection in the second step determines the essential figure of merit for the source detection, that is, the efficiency with which sources can be distinguished from background fluctuations. The false-positive rate of a source candidate is usually expressed as the source likelihood, which is used to threshold samples. The ideal source detection strategy should define a source likelihood such that a threshold on the likelihood leads to both an optimized completeness and an optimized purity.

With the eSASS tasks, we can measure two types of source likelihoods. The PSF-fitting likelihood is measured through a maximum-likelihood PSF-fitting, which uses an algorithm that was first developed for ROSAT (Hasinger et al. 1994) and is still in use for *XMM-Newton* (e.g., Liu et al. 2020). The aperture-based likelihood is calculated through a comparison of the source and background photon counts extracted in a small aperture (Georgakakis et al. 2008) to the expectation of Poisson statistics. By comparing the source detection efficiency based on these two types of likelihood, we confirm the good performance of the PSF-fitting algorithm in the sense that involving the known PSF shape information leads to a higher efficiency than using only aperture photometry results (Fig. 9). We also remark that in a broad sense, the likelihood is defined by all the methods and parameters adopted in the procedure of source detection because they all affect it; and the likelihood definition based either on PSF-fitting or on aperture photometry underestimates the false-positive rate (Figs. 3, 4) because they are based on the assumption that there are no additional uncertainties.

Based on the experience and rich legacy of previous X-ray surveys, the eSASS scheme uses a multistage source detection procedure around a core algorithm of PSF fitting, that is, it first detects a preliminary catalog and then performs PSF fitting to make a final selection. Since the preliminary catalog requires high completeness but not high purity, which can be controlled in subsequent steps, a simple sliding-box algorithm is used for it, and in this work, we focused on the PSF-fitting step. A first choice must be made as to whether to use all the data or not. The simple combination of more data does not necessarily result in a higher contrast of source to background, as it might introduce more noise rather than signal. For instance, at the border of the FOV, Wolter-I X-ray telescopes suffer from strong vignetting and degraded spatial resolution, and at high energies, X-ray telescopes usually capture fewer X-ray photons but gain more noise due to instrumental background. To understand how these effects contribute to the data set as a whole, we measured and compared the source detection efficiencies through simulations adopting each of the candidate strategies (Fig. 9). We find that including hard bands in the detection (three band and four band) introduces more noise and thus reduces the overall efficiency. Also considering that a single-band detection leads to a more straightforward estimation of selection function, we chose the single-band detection to create the main eFEDS catalog (Brunner et al. 2022). However, to detect hard sources (especially obscured AGN), we also make use of the three-band (0.2–0.6, 0.6–2.3, and 2.3–5 keV) detection, and select sources based on the 2.3–5 keV individual band detection likelihood (Brunner et al. 2022, Nandra et al., in prep.). The key parameter of the PSF fitting is the fitting aperture radius. We find that the choice of this radius affects the measured detection likelihood, but does not impact the detection efficiency of point sources. Enlarging

the PSF-fitting radius, however, might improve the efficiency of identifying clusters (Fig. 10). A moderate radius of 15 pixels (60'') was adopted when the real eFEDS catalog was created (Brunner et al. 2022). For future eROSITA surveys, PSF fitting with a larger radius or a post hoc analysis with a larger radius is recommended. For aperture photometry, the simulation suggests that an optimized aperture radius corresponds to 60% EEf. Excluding the FOV border or repeating the PSF fitting might improve the detection of faint sources, but no visible impact is found for the whole eFEDS sample in the soft band.

The eFEDS catalog is dominated by soft point sources (Liu et al. 2022b), containing only < 2% galaxy clusters (Liu et al. 2022a). When we chose the source detection strategy for eFEDS, we aimed at the overall quality of the entire X-ray catalog rather than focusing on any particular type of astronomical object. We remark that the method described in this paper allows us to fine-tune the detection accordingly to improve the detection of any particular type of objects such as obscured AGN or galaxy clusters, and different types of objects will require different strategies or parameters. We also remark that the simulation results reported in this paper are subject to the observing strategy of eFEDS. There will be differences in pointing-mode observations and in the eRASS observations of regions with a largely different exposure depth, for instance, the south ecliptic pole (see also the discussion in Clerc et al. 2018). The results of pointing-mode simulation are reported in a following paper about the extragalactic serendipitous X-ray catalog detected from several pointing-mode observations (Liu et al. 2022c).

6.2. Characterizing the eFEDS catalog

In order to simulate eFEDS as representative of the real data as possible, we extracted the background spectrum from the real eFEDS data and decomposed it into an X-ray component and a particle component. The two components were simulated separately with and without vignetting. We compared the mock data with the real data in a few aspects in Fig. 2. The sources and the background are largely similar between them, except that the cluster surface profile distribution has a relatively large uncertainty and possibly differs from the real eFEDS clusters. With this detailed construction of mock data and performing on it the identical source detection pipeline used for the real eFEDS catalog (Brunner et al. 2022), we can quantify the completeness and purity of the real eFEDS catalog with the mock catalog.

We introduced a detailed strategy of analyzing detected sources in a simulation based on the origin of each photon. By verifying the source types (point source or cluster) of the primary and secondary (if exists) input counterparts of each detected source, we identified the nature of the detected sources as either correctly detected point or extended sources or due to blended sources, fragmentation of a large source, or misclassified sources. By plotting the distributions of various classes of detected sources as a function of source properties (e.g., detection likelihood and extent likelihood), the fractions of spurious sources or any particular cases as mentioned above can be measured quantitatively. For the single-band-detected catalog, adopting a DET_LIKE threshold of 5, 6, and 8 resulted in a spurious fraction of 11.5%, 6.3%, and 1.8%, respectively. Selecting hard sources from the three-band-detected catalog with the 2.3–5 keV individual band likelihood $\text{DET_LIKE}_3 > 10$ and the extent likelihood $\text{EXT_LIKE} < 14$ resulted in a spurious fraction of 2.5%. For the single-band detection, the point source catalog ($\text{EXT_LIKE} = 0$) includes 3% misclassified clusters, and the

cluster catalog ($\text{EXT_LIKE} > 0$) includes 29% misclassified point sources. The fraction of spurious clusters is likely overestimated because the mock cluster catalog shows a strong peak near the lower limit of EXT_LIKE , which does not exist in the real eFEDS catalog (Fig. 2). In the single-band-detected catalog, a fraction of 34% of the correctly classified point sources is contaminated by nearby point sources or clusters; but for the misclassified point sources, this fraction increases to 83%, indicating that source blending is the main reason for misclassifying point sources as extended. In the eFEDS 90% area region, the simulation predicts 42.9 spurious clusters ($\text{EXT_LIKE} > 0$) when we adopt $\text{DET_LIKE} > 12$ and $\text{EXT_LIKE} > 12$ and predicts 9.4 spurious clusters when we adopt $\text{DET_LIKE} > 20$ and $\text{EXT_LIKE} > 20$. Using the mock output catalogs (Appendix B) presented with this paper, the fraction of any particular case in the eFEDS catalog under any particular sample selection criteria can be measured.

On the output side, we characterized the catalog purity. On the input side, we measured the detection completeness (selection function) in various manners, including the detected fraction of AGN or clusters as a function of input flux (Fig. 6), the detected fraction as a function of luminosity for AGN in particular redshift and column density ranges and for clusters in particular redshift, scale, and mass ranges (Fig. 7), and the distribution of AGN completeness in 2D luminosity–redshift space. The AGN completeness exceeds 80% at 0.5–2 keV fluxes above $10^{-14.2} \text{ erg cm}^{-2} \text{ s}^{-1}$, and the cluster completeness exceeds 80% at 0.5–2 keV fluxes above $10^{-13.1} \text{ erg cm}^{-2} \text{ s}^{-1}$. The main cause of the cluster incompleteness is misclassification. Including the misclassified sources, the cluster completeness exceeds 80% at fluxes above $10^{-13.7} \text{ erg cm}^{-2} \text{ s}^{-1}$. In this paper, we only display the AGN and cluster selection functions approximately. More sophisticated sampling methods are needed in order to construct the selection function of AGN and clusters more accurately as a function of multidimensional intrinsic source properties from the mock catalogs, which are available with this paper (Appendix B). In a few accompanying papers, detailed analyses of the selection functions are performed in the demography studies of AGN (e.g., Buchner et al., in prep.) and galaxy clusters (e.g., Liu et al. 2022a).

We compared the point-source number counts constructed with the method based on aperture photometry, which is used for the real eFEDS catalog, with the input number counts, and confirmed that this method can broadly recover the input number counts. However, the recovered number counts are slightly (a few percent) higher than the input. Our sophisticated analyzing method based on the origin of each photon enabled us to determine that this is caused by contamination of misclassified clusters and by source blending. These factors can only be revealed by simulation and not in any analytical way.

Through the simulation, we also realized a few minor issues with the current (version c001) eROSITA PSF models. As described in Appendix A, the source positions measured through PSF fitting have a tiny offset and the positional uncertainties are slightly underestimated. Therefore, we recommend post hoc astrometric corrections to the real eFEDS catalog (done

in Brunner et al. 2022). We also noted a $\sim 5\%$ uncertainty in the PSF fitting measured source counts (or flux) caused by the normalization of the PSF models (Appendix A).

Acknowledgement. This work is based on data from eROSITA, the soft X-ray instrument aboard Spectrum-Roentgen-Gamma (SRG), a joint Russian-German science mission supported by the Russian Space Agency (Roskosmos), in the interests of the Russian Academy of Sciences represented by its Space Research Institute (IKI), and the Deutsches Zentrum für Luft- und Raumfahrt (DLR). The SRG spacecraft was built by Lavochkin Association (NPOL) and its subcontractors, and is operated by NPOL with support from the Max Planck Institute for Extraterrestrial Physics (MPE). The development and construction of the eROSITA X-ray instrument was led by MPE, with contributions from the Dr. Karl Remeis Observatory Bamberg and ECAP (FAU Erlangen-Nuernberg), the University of Hamburg Observatory, the Leibniz Institute for Astrophysics Potsdam (AIP), and the Institute for Astronomy and Astrophysics of the University of Tübingen, with the support of DLR and the Max Planck Society. The Argelander Institute for Astronomy of the University of Bonn and the Ludwig Maximilians Universität Munich also participated in the science preparation for eROSITA. The eROSITA data shown here were processed using the eSASS/NRTA software system developed by the German eROSITA consortium. Some of the results in this paper have been derived using the HEALPix (Górski et al. 2005) package.

References

- Anderson, M. E., Gaspari, M., White, S. D. M., Wang, W., & Dai, X. 2015, *MNRAS*, **449**, 3806
- Brandt, W. N., & Alexander, D. M. 2015, *A&ARv*, **23**, 1
- Brandt, W. N., & Hasinger, G. 2005, *A&ARv*, **43**, 827
- Brunner, H., Liu, T., Lamer, G., et al. 2022, *A&A*, **661**, A1 (eROSITA EDR SI)
- Calderwood, T., Dobrzycki, A., Jessop, H., & Harris, D. E. 2001, in *ASP Conf. Ser.*, **238**, *Astronomical Data Analysis Software and Systems X*, eds. J. Harnden, F. R., F. A. Primini, & H. E. Payne, 443
- Chuang, C.-H., Yepes, G., Kitaura, F.-S., et al. 2019, *MNRAS*, **487**, 48
- Clerc, N., Ramos-Ceja, M. E., Ridl, J., et al. 2018, *A&A*, **617**, A92
- Comparat, J., Merloni, A., Salvato, M., et al. 2019, *MNRAS*, **487**, 2005
- Comparat, J., Eckert, D., Finoguenov, A., et al. 2020, *Open J. Astrophys.*, **3**, 13
- Dauser, T., Falkner, S., Lorenz, M., et al. 2019, *A&A*, **630**, A66
- Dennerl, K., Andriutschke, R., Bräuninger, H., et al. 2020, *SPIE Conf. Ser.*, **11444**, 114444Q
- Deponte, J., & Primini, F. A. 1993, in *ASP Conf. Ser.*, **52**, *Astronomical Data Analysis Software and Systems II*, eds. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes, 425
- Ebeling, H., & Wiedenmann, G. 1993, *Phys. Rev. E*, **47**, 704
- Freeman, P. E., Kashyap, V., Rosner, R., & Lamb, D. Q. 2002, *ApJS*, **138**, 185
- Georgakakis, A., Nandra, K., Laird, E. S., Aird, J., & Trichas, M. 2008, *MNRAS*, **388**, 1205
- Ghirardini, V., Eckert, D., Ettori, S., et al. 2019, *A&A*, **621**, A41
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, **622**, 759
- Hasinger, G., Johnston, H. M., & Verbunt, F. 1994, *A&A*, **288**, 466
- LaMassa, S. M., Urry, C. M., Cappelluti, N., et al. 2013, *MNRAS*, **436**, 3581
- Liu, T., Tozzi, P., Tundo, E., et al. 2013, *A&A*, **549**, A143
- Liu, T., Merloni, A., Simm, T., et al. 2020, *ApJS*, **250**, 32
- Liu, A., Bulbul, E., Ghirardini, V., et al. 2022a, *A&A*, **661**, A2 (eROSITA EDR SI)
- Liu, T., Buchner, J., Nandra, K., et al. 2022b, *A&A*, **661**, A5 (eROSITA EDR SI)
- Liu, T., Merloni, A., Wolf, J., et al. 2022c, *A&A*, submitted [arXiv:2202.09430]
- Merloni, A., Predehl, P., Becker, W., et al. 2012, *ArXiv e-prints* [arXiv:1209.3114]
- Predehl, P., Andriutschke, R., Arefiev, V., et al. 2021, *A&A*, **647**, A1
- Seppi, R., Comparat, J., Nandra, K., et al. 2021, *A&A*, **652**, A155
- Smith, R. K., Brickhouse, N. S., Liedahl, D. A., & Raymond, J. C. 2001, *ApJ*, **556**, L91

Appendix A: Accuracy of the source property measurement

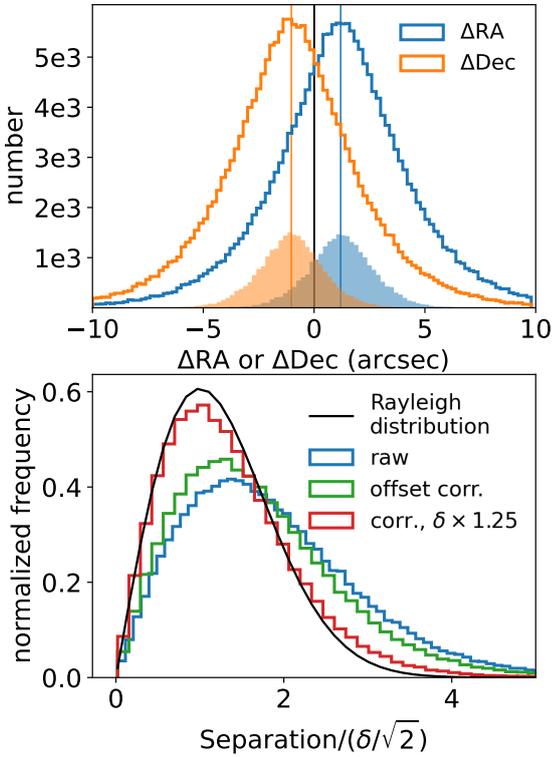


Fig. A.1. Distributions of $RA_{out}-RA_{in}$ (blue, top) and $Dec_{out}-Dec_{in}$ (orange, top). The filled histograms indicate subsamples of very bright sources with detection likelihoods >100 . The full samples and the subsamples have a median $\Delta RA = 1.2''$ (vertical blue line) and a median $\Delta Dec = -1.0''$ (vertical orange line). The lower panel displays the distribution of input-output separation in terms of the ratio to the 1D positional uncertainty in comparison with the Rayleigh distribution. The blue histogram shows the raw measurements. The distribution after correcting the detected positions for the median offsets is plotted in green. Then we further multiply the measured positional uncertainty by a factor of 1.25 and plot the distribution in red.

In addition to the source detection efficiency, we also tested the measurement accuracy of point-source properties. From the single-band-detected mock catalog, we selected the point sources ($EXT_LIKE=0$) that were uniquely matched to input non-contaminated ($ID_contam < 0$) point sources and compared the input and output positions and source counts.

The upper panel of Fig. A.1 compares the input and output positions of the bright sources with $DET_LIKE > 10$. Offsets of $RA_{out}-RA_{in} = 1.2''$, $Dec_{out}-Dec_{in} = -1.0''$ are found, which are independent of the source brightness. Compared to the native eROSITA CCD pixel sizes of $9.6''$, these offsets are very small. One possible reason is the difference between the mock data and the real data, particularly in the process of converting the accurate photon coordinates into pixel coordinates. Subpixel calibration of a photon position is achieved for eROSITA based on event patterns (Dennerl et al. 2020), but these details are not reproduced by SIXTE. Alternatively, it might also be caused by the numerical accuracy of the current PSF model, which has no subpixel resolution. To guarantee maximum positional accuracy, we suggest a post hoc astrometric correction of the X-ray cata-

logs by comparing the X-ray positions and optical-IR positions of distant AGN (e.g., as done in Brunner et al. 2022).

A good measurement of positional uncertainty is essential to identify the optical counterpart of an X-ray source from deep optical surveys, as there are often multiple candidates consistent in position within the uncertainty. The lower panel of Fig. A.1 displays the distribution of separation of all the matched input-output pairs in terms of σ level (ratio to the positional uncertainty). After correcting the offsets found above, this distribution is still broader than expected (the Rayleigh distribution), indicating a slight underestimation of the positional uncertainty. We found that by enlarging the measured positional uncertainty by a factor of 1.25, the distribution becomes comparable with the Rayleigh distribution.

For the simulation, SIXTE adopts the 2D PSF image library. To detect the real eFEDS catalog, we performed photon-mode PSF fitting using `ermldet`, which accounts for the PSF of each photon based on a shapelet PSF library (Brunner et al. 2022). The same photon-mode PSF fitting was performed on the mock data in this work, providing a measurement of source counts, which was corrected for the PSF loss outside the extracting circle. Comparing these measured counts with the input photon counts of each source in the soft (0.2–2.3 keV, in the single-band detection) and hard (2.3–5 keV, in the three-band detection) band, we find that the PSF-corrected source counts are $\sim 5\%$ lower than the input. We repeated the `ermldet` PSF fitting of the mock data in image-mode, that is, adopting the 2D PSF image library, and find that the input and output source counts are perfectly consistent. Therefore, the current (version c001) shapelet PSF library and the 2D PSF library deviate slightly in the normalization. This issue will be addressed in a future version of the eROSITA calibration database.

Appendix B: Catalog description

We provide the input AGN and cluster catalogs and the output single-band-detected and three-band-detected catalogs, whose columns are described in Table. B.1. These mock catalogs are available on the eROSITA Early Data Release website². The input catalogs contain a number of columns of AGN or cluster properties drawn from Comparat et al. (2019, 2020) and a few columns of the output-counterpart ID (-99 means that it does not exist). The very faint input AGN with fluxes below 5×10^{-16} erg $cm^{-2} s^{-1}$ are omitted in the input AGN catalog. The output catalogs contain a few essential columns that are the same as in the real eFEDS catalog (Brunner et al. 2022) and a few additional columns of the input-counterpart ID (-99 means that it does not exist).

² <https://erosita.mpe.mpg.de/edr/eROSITAObservations/Catalogues/>

Table B.1. Columns of the mock input and output eFEDS catalogs

Name	Description
For both input AGN and clusters	
SRC_ID	Source ID; always $< 5 \times 10^6$ for AGN; always $> 10^7$ for clusters
RA	Right ascension (deg)
DEC	Declination (deg)
z	redshift (in redshift space)
dL	luminosity distance (cm) in the cosmology adopted by Comparat et al. (2019, 2020) to create the mock catalogs
ID_Uniq	source ID in single-band catalog whose ID_Uniq is this source
ID_Any	source ID in single-band catalog whose ID_Any is this source
ID_Any2	source ID in single-band catalog whose ID_Any2 is this source
ID_contam	ID of contaminating input source
ID_Uniq_T	source ID in three-band catalog whose ID_Uniq is this source
ID_Any_T	source ID in three-band catalog whose ID_Any is this source
ID_Any2_T	source ID in three-band catalog whose ID_Any2 is this source
Nreal	number of realization (1–18)
For input AGN only, see Comparat et al. (2019) for more details	
logNH	AGN column density (cm^{-2})
FLUX	0.5-2 keV observed flux ($\text{erg cm}^{-2} \text{s}^{-1}$); the input flux in simulation
FLUX_GalCorr	0.5-2 keV flux corrected for Galactic absorption ($\text{erg cm}^{-2} \text{s}^{-1}$)
Lx_hard	2-10 keV intrinsic luminosity (erg/s)
agn_type	AGN type
galaxy_SMHMR_mass	Host galaxy stellar mass (M_{\odot})
For input cluster only, see Comparat et al. (2020) for more details	
kT	cluster temperature (keV)
Lx_soft	rest-frame 0.5-2 keV luminosity within R500c (erg/s)
Lx_soft_2R500	rest-frame 0.5-2 keV luminosity within twice R500c (erg/s)
Fx_soft	observed 0.5-2 keV flux within R500c
Fx_soft_2R500	observed 0.5-2 keV flux within twice R500c; the input flux in simulation
M500c	halo mass within R500c (M_{\odot})
R500c_arcmin	apparent radius where the mass density is 500 times the universe critical density (arcmin)
R500c_kpc	physical radius where the mass density is 500 times the universe critical density (kpc, proper distance)
pid	flag; a value of -1 means the halo is a distinct halo
Mvir	halo mass within the virial radius (M_{\odot})
Rvir	halo virial radius (kpc)
Xoff	halo offset parameter (kpc)
b_to_a_500c	halo ellipticity
Single-band and three-band detected output catalogs	
ID_SRC	source ID
RA	Right ascension (deg)
DEC	Declination (deg)
RADEC_ERR	combined positional error (arcsec)
EXT	source extent (arcsec)
EXT_ERR	extent error (arcsec)
EXT_LIKE	extent likelihood measured by PSF-fitting
ML_RATE_n	source count rate (cts/s)
ML_RATE_ERR_n	count rate error (cts/s)
DET_LIKE_n	detection likelihood measured by PSF-fitting
ML_BKG_n	background at source position (cts/arcmin^2)
ML_EXP_n	vignetted exposure (s) at source position
ID_Uniq	ID of the unique input counterpart
ID_Any	ID of the brightest input counterpart, allowing duplicate
ID_Any2	ID of the secondary input counterpart
ID_contam	ID of the input source contaminating the unique input counterpart
Nreal	number of realization (1–18)

Notes. The input AGN and clusters catalogs have some columns in common. They are listed in the first section. Some exclusive columns are listed in the second and third sections. The output single-band-detected and three-band-detected catalogs share basically the same columns, except for a few columns from PSF fitting with an optional suffix *_n* in the name. The columns with $n = 1, 2,$ or 3 correspond to the three individual energy bands 0.2–0.6, 0.6–2.3, and 2.3–5 keV used in the three-band detection; $n = 0$ indicates the summary value in the three-band detection; if without the *_n* suffix, the column corresponds to the single-band detection.