The eROSITA Final Equatorial Depth Survey (eFEDS):

X-ray catalogue

H. Brunner[1]⋆, T. Liu[1] (刘腾)⋆⋆, G. Lamer[2], A. Georgakakis[3], A. Merloni[1], M. Brusa[4,5], E. Bulbul[1], K. Dennerl[1], S. Friedrich[1], A. Liu[1] (刘昂), C. Maitra[1], K. Nandra[1], M. E. Ramos-Ceja[1], J. S. Sanders[1], I. M. Stewart[1], T. Boller[1], J. Buchner[1], N. Clerc[6], J. Comparat[1], T. Dwelly[1], D. Eckert[7,1], A. Finoguenov[8], M. Freyberg[1], V. Ghirardini[1], A. Gueguen[1], F. Haberl[1], I. Kreykenbohm[9], M. Krumpe[2], S. Osterhage[1], F. Pacaud[10], P. Predehl[1], T. H. Reiprich[10], J. Robrade[11], M. Salvato[1], A. Santangelo[12], T. Schrabback[10], A. Schwope[2], and J. Wilms[9]

[1] Max-Planck-Institut für extraterrestrische Physik, Gießenbachstraße 1, 85748 Garching, Germany
[2] Leibniz-Institut für Astrophysik Potsdam (AIP), 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 "Augusto Righi", Alma Mater Studiorum Universit à di Bologna, via Gobetti 93/2, 40129 Bologna, Italy
[5] INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Gobetti 93/3, 40129 Bologna, Italy
[6] IRAP, Université de Toulouse, CNRS, UPS, CNES, Toulouse, France
[7] Department of Astronomy, University of Geneva, Ch. d'Ecogia 16, 1290 Versoix, Switzerland
[8] Department of Physics, University of Helsinki, Gustaf Hllströmin katu 2a, FI-00014 Helsinki, Finland
[9] Dr. Karl Remeis-Sternwarte and Erlangen Centre for Astroparticle Physics, Friedrich-Alexander Universität Erlangen-Nürnberg, Sternwartstraße 7, 96049 Bamberg, Germany
[10] Argelander-Institut für Astronomie (AIfA), Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
[11] Universität Hamburg, Hamburger Sternwarte, Gojenbergsweg 112, D-21029 Hamburg, Germany
[12] Institut für Astronomie und Astrophysik, Universitt Tübingen, Sand 1, D 72076 Tübingen, Germany

ABSTRACT

Context. The eROSITA X-ray telescope on board the *Spectrum-Roentgen-Gamma* (*SRG*) observatory combines a large field of view and a large collecting area in the energy range between ∼0.2 and ∼8.0 keV. This gives the telescope the capability to perform uniform scanning observations of large sky areas.

Aims. *SRG*/eROSITA performed scanning observations of the ∼140 square degree *eROSITA Final Equatorial Depth Survey* field (the eFEDS field) as part of its performance verification phase ahead of the planned four years of all-sky scanning operations. The observing time of eFEDS was chosen to slightly exceed the depth expected in an equatorial field after the completion of the all-sky survey. While verifying the capability of eROSITA to perform large-area uniform surveys and serving as a test and training dataset to establish calibration and data analysis procedures, the eFEDS survey also constitutes the largest contiguous soft X-ray survey at this depth to date, supporting a range of early eROSITA survey science investigations. Here we *i)* present a catalogue of detected X-ray sources in the eFEDS field providing information about source positions and extent, as well as fluxes in multiple energy bands, and *ii)* document the suite of tools and procedures developed for eROSITA data processing and analysis, which were validated and optimised by the eFEDS work.

Methods. The data were fed through a standard data processing pipeline, which applies X-ray event calibration and provides a set of standard calibrated data products. A multi-stage source detection procedure, building in part on experience from *XMM-Newton*, was optimised and calibrated by performing realistic simulations of the eROSITA eFEDS observations. Source fluxes were computed in multiple standard energy bands by forced point source fitting and aperture photometry. We cross-matched the eROSITA eFEDS source catalogue with previous XMM-ATLAS observations, which confirmed the excellent agreement of the eROSITA and XMM-ATLAS source fluxes. Astrometric corrections were performed by cross-matching the eROSITA source positions with an optical reference catalogue of quasars.

Results. We present a primary catalogue of 27910 X-ray sources (542 of which are significantly spatially extended) detected in the 0.2–2.3 keV energy range with detection likelihoods ≥ 6, corresponding to a (point source) flux limit of 6.5×10^{-15} erg/cm^2/s in the 0.5–2.0 keV energy band (80% completeness). A supplementary catalogue contains 4774 low-significance source candidates with detection likelihoods between 5 and 6. In addition, a hard-band sample of 246 sources detected in the energy range 2.3–5.0 keV above a detection likelihood of 10 is provided. In an appendix, we finally describe the dedicated data analysis software package, the eROSITA calibration database, and the standard calibrated data products.

Key words. catalogues – surveys – X-ray: general

1. Introduction

eROSITA (extended ROentgen Survey with an Imaging Tele-scope Array; Predehl et al. 2021) is the primary instrument on

---

⋆ email: hbrunner@mpe.mpg.de
⋆⋆ email: liu@mpe.mpg.de

the *Spektrum-Roentgen-Gamma (SRG)* orbital observatory (Sunyaev et al. 2021). It was designed and built, over a period of about 12 years, with the goal of realising a sensitive X-ray telescope with a wide field of view and a significantly larger grasp[1] at 1 keV than that of either *XMM-Newton* (by about a factor of 4) or *Chandra* (by about a factor of 60). These are the two most sensitive focusing X-ray telescopes currently in operation. In addition, a mission plan was devised that included a long (4 years), uninterrupted all-sky survey program (the eROSITA All-Sky Survey: eRASS; Predehl et al. 2021) in order to guarantee a large volume of accessible discovery space. The *SRG* is operated by the Space Research Institute of the Russian Academy of Sciences (IKI).

The scientific motivation for the eROSITA design was the desire to detect and spatially resolve a large number (about 10^5) of clusters of galaxies over a wide redshift range (up to at least $z \sim 1$) in order to constrain cosmological parameters (see e.g. Allen et al. 2011; Pillepich et al. 2018) at the level of a Stage-IV Dark Energy experiment (Albrecht et al. 2006) by characterising the growth of structure.

Following the successful launch in July 2019, and before the start of the all-sky survey, a number of performance verification (PV) observations were carried out during the early phases of the *SRG* mission, aimed at verifying all different aspects of the instrument capabilities. These early post-commissioning phases of science operations have clearly demonstrated that eROSITA is capable of delivering the design performance in terms of sensitivity, image quality, and spectroscopic capabilities (see e.g. Merloni et al. 2020; Predehl et al. 2021). Nevertheless, the ambitious scientific goals of eROSITA, and in particular its cosmological objectives, require accurate control over a number of elements, including the modelling of the instrumental and astrophysical backgrounds, the reconstruction of the selection function for both point-like and extended X-ray sources over very large sky areas, the identification of reliable counterparts of the X-ray sources at longer wavelengths, the measurement of their distances (redshifts), the calibration of the empirical (or physical) relations between X-ray observables (e.g. cluster luminosities, temperatures, and density profiles) and cosmologically meaningful parameters, particularly cluster masses.

The *eROSITA Final Equatorial Depth Survey* (eFEDS) was the largest investment of observing time during the PV phase (about 360 ks, or 100 hours, in total), and was designed to test the key elements of the science workflow from X-ray photon detection to astrophysics and cosmology. The eFEDS field, an area of approximately 140 deg^2 composed of four individual rectangular raster-scan fields of $\sim$35 deg^2 each, was chosen because its multi-wavelength coverage for an extragalactic field of this size is one of the richest, and because it was visible for eROSITA in the limited time period allocated to PV observations. The field coincides with an area enriched by deep optical and near-infrared (NIR) imaging of the HSC[2] Wide area Survey (Aihara et al. 2018), KIDS-VIKING[3] (Kuijken et al. 2019), DESI Legacy Imaging Survey[4] (Dey et al. 2019), and, among others, GAMA[5] (Driver et al. 2009), WiggleZ (Drinkwater et al. 2018), LAM-OST[6], and the SDSS[7] (Blanton et al. 2017) spectroscopic coverage. A full description of the available multiwavelength data, together with the identification of the multiwavelength counterparts, is presented in Salvato et al. (2021). The field also contains a medium-wide *XMM-Newton* survey field (the XMM-ATLAS; Ranalli et al. 2015), which provides a useful dataset for a comparison and validation of the eROSITA X-ray analysis pipeline.

The eFEDS observational strategy was designed so as to provide uniform exposure over the field about 50% deeper than what is expected for eRASS at the ecliptic equator (i.e. over most of the sky) at the end of the 4-year all-sky survey program, while at the same time covering a sufficiently wide area to provide large statistical samples of different source classes. In particular, enough clusters were to be provided for calibrating the mass-observable relation using the weak-lensing maps provided by the exquisite optical imaging of the HSC survey.

In a series of accompanying papers, we will present the identification of the multiwavelength counterparts of point-like (Salvato et al. 2021; Schneider et al. 2021) and extended (Klein et al. 2021) X-ray sources; the resulting clean catalogues of clusters of galaxies (Liu et al. 2021a; Bulbul et al. 2021), active galactic nuclei (AGN) (Liu et al. 2021b, Nandra et al., in prep.) and stars (Schneider et al. 2021); the X-ray variability properties of the detected sources (Boller et al. 2021; Buchner et al. 2021); the X-ray spectral analysis of the point sources (Liu et al. 2021b); the X-ray morphological analysis of the clusters (Ghirardini et al. 2021a); the optical and lensing analysis of the X-ray clusters (Ramos-Ceja et al. 2021; Chiu et al. 2021; Bahar et al. 2021, Ota et al., in prep.); the properties of non-active galaxies detected by eROSITA (Vulic et al. 2021), and the discovery of some extreme AGN (Toba et al. 2021; Brusa et al. 2021). Earlier works based on the eFEDS data have been published and include the discovery of a supercluster (Ghirardini et al. 2021b) and a very high-redshift quasar (Wolf et al. 2021).

In this paper, we describe in detail the X-ray observations, the data analysis and calibration procedures, the algorithms we used to detect and characterise X-ray sources, and their validation through an extensive simulation analysis. We also release here the resulting catalogues of X-ray sources, and present the description of the dedicated software system (the eROSITA Science Analysis Software System; eSASS) in a series of appendices, which is also made public concurrently with this paper. The appendices include a basic description of the operating principles and algorithms of the data analysis pipeline. A full description of the eSASS software is also available online[8].

2. Observations

The *SRG* mission supports three principal observing modes (Predehl et al. 2021; Sunyaev et al. 2021). As its main objective, *SRG* is in the process of performing a 4-year survey of the full sky in continuous scanning mode, in which the spacecraft pointing direction traces great circles in the sky with a rotation speed of 90 degrees per hour (or $\sim$ 90$''$/s). In addition, *SRG* is capable of making pointed observations of individual targets, as well as extended rectangular fields in so-called field-scanning mode. The pattern of the field-scanning mode consists of parallel scans in alternating directions, offset by 6$'$. Given the eROSITA field of view ($\approx$ 1 degree in diameter), each sky position is thus observed in ten consecutive scans. The maximum supported size

---

[1] In X-ray astronomy, grasp is a primary survey metric that is the product of the field-of-view average effective area times the field of view of the telescope.
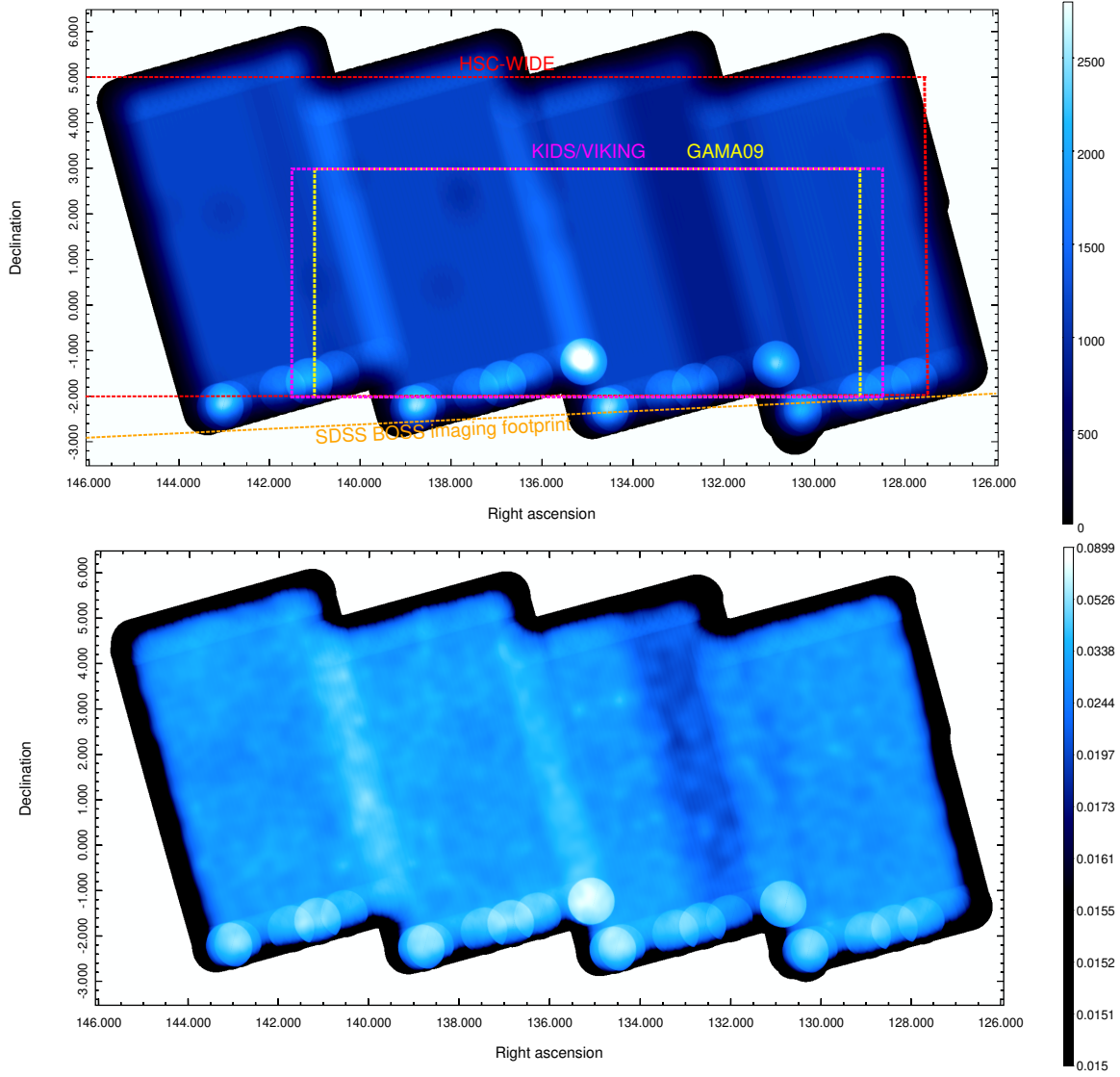
[2] Hyper Suprime-Cam; https://subarutelescope.org/

[3] https://kids.strw.leidenuniv.nl/

[4] https://www.legacysurvey.org/

[5] http://www.gama-survey.org/

[6] http://dr5.lamost.org/doc/release-note-v3

[7] Sloan Digital Sky Survey; https://www.sdss.org/

[8] https://erosita.mpe.mpg.de/edr/DataAnalysis/

Fig. 1. 0.2-2.3 keV exposure map (vignetting corrected, in units of seconds; upper panel) and background map (in units of counts per pixel; lower panel). The bright circles mark high-exposure regions and are a result of the scanning mode used in the observations. They are created due to the waiting time of the spacecraft before inverting its scanning direction. The darker stripe in the second rectangular chunk on the right is due to the malfunction of two of the seven eROSITA cameras during that period. The upper panel also shows the footprints of a few relevant optical and NIR, imaging and spectroscopic surveys.

of the scanned rectangles is 12.5° × 12.5°. In the interest of operational simplicity, the orientation of the rectangles is aligned with the ecliptic coordinate system. For mission planning purposes, individual field-scan observations are scheduled to be not significantly longer than one day, thus fitting between consecutive ground contacts with the SRG spacecraft. These constraints necessitate splitting the observations of the eFEDS field into four sub-fields of size 4.2° × 7.0° (this is the area with the nominal exposure depth; the area is larger by ∼ 0.5° on each side, but the exposure is reduced). A scanning speed of 13″.15/s results in a uniform exposure depth of ∼ 2.2 ks (∼ 1.2 ks after correcting for telescope vignetting) across most of the field. Each source is observed continuously for up to 4.7 min in an individual scan when passing through the central part of the field of view and for shorter time periods in peripheral scans.

The eFEDS observations were carried out at the beginning of November 2019 with all seven eROSITA Telescope Modules (TM1-7) in operation. However, due to an unrecognised mal-

function of the camera electronics, 28% of the TM6 data of eFEDS sub-field I and 48% and 43% of the TM5 and TM6 data, respectively, of eFEDS sub-field II could not be used. This resulted in a reduced exposure depth in the affected areas of up to ∼30% (as visible in Figure 1).

We report an unrecognized calibration error of 1 s in the arrival time correction for approximately half of the TM6 data of eFEDS sub-field I prior to the malfunction. Because of the scanning pattern of the eROSITA field-scan observing mode, this causes photons observed by this camera to be projected onto the sky with an offset of ∼13″ on either side of the true position, resulting in an elongation of the merged all-camera point spread function (PSF) in the scanning direction in the affected area on the eastern side of sub-field I. Some technical details about the cause and correction method of these 1 s time shifts are provided in Appendix B.3.

All eROSITA cameras are protected by light-blocking filters. For five of the seven cameras, a 200 nm Al layer is deposited

directly on the CCDs[9] , while for the remaining two (TM5 and TM7), the suppression of optical light is achieved by a 100 nm Al layer on an external filter in the filter wheel. After launch it was found that a small amount of scattered sunlight can reach the CCDs from the side, bypassing the filter wheel ('light leak'; see Predehl et al. 2021). This causes a spatially inhomogeneous raised background at low energies in TM5 and TM7 that changes with the orientation of the spacecraft with respect to the Sun.

This light-leak noise, which is smoothly distributed at scales >10′, will be included in the background map and thus has a minor impact on the source detection. The optical photons may also cause an energy shift that is negligible for the source detection, however. Therefore, all seven cameras are used in this work.

Table 1 presents the details of the eFEDS field observations. The top panel of Figure 1 shows the final exposure map of the eFEDS survey in the soft band (0.2–2.3 keV), corrected for vignetting[10]. The schematic footprints of a few relevant optical and NIR, imaging and spectroscopic surveys are overlaid on this map.

Figure 2 shows an exposure-corrected RGB image of the entire field. The images were created using the 0.2–0.5 keV (red), 0.5–1 keV (green), and 1–2 keV (blue) energy bands, after smoothing each count rate image with a Gaussian with $\sigma = 40''$.

3. Data analysis

In this section we describe the details of the data processing with the eROSITA eSASS. The eSASS data analysis package and its tasks are described in detail in Appendix A. Further documentation including a description of all command-line parameters is available online[8].

3.1. Data reception and standard pipeline processing

eROSITA science and auxiliary telemetry data received during each daily SRG ground contact are transferred from the IKI to the eROSITA Data Centre at the Max Planck Institute for Extraterrestrial Physics (MPE) in Garching, Germany, where they undergo various processing steps as part of a standard data analysis pipeline. The data are decommutated, packaged, and converted into FITS format (Wells et al. 1981) files in an initial preprocessing step. They are subsequently fed through an event-processing task chain that creates calibrated X-ray event files suitable for scientific analysis. The main software tasks of the event-processing chain are described in Appendix A.1 and A.2.; calibrated event files are described in Appendix C.1. In addition, the standard data processing pipeline provides a range of high-level data products such as exposure, background, and sensitivity maps, as well as X-ray source catalogues and source-specific products such as spectra and light curves. They are described in Appendix C.

This work is based on the pipeline-generated calibrated event files, while higher-level data products are created by an additional pipeline. The data were processed using pipeline version c001, which was released as part the eROSITA Early Data Release[11].

3.2. Astrometric corrections and data preparation

Some artefacts in the data due to a temporary malfunctioning of the camera electronics of TM4 are not yet perfectly removed in this version of the pipeline. In the TM4 data of all the four observations, we therefore removed two bright pixels (pixel coordinate `RAWX`, `RAWY`: 115, 235, and 178, 225). Furthermore, we removed the soft photons with a `PI` (event energy in eV) below 600 and `RAWY` above 120 in a few columns (`RAWX`: 121, 125, 127, 259, 262, 380, and 382). The fraction of removed events is negligible.

We carried out an initial round of astrometric corrections of the X-ray dataset as follows. We created images and detected sources in the 0.2–2.3 keV band separately for each of the four eFEDS ObsIDs (using the methods described below). For each ObsID, we searched for the closest optical and IR AGN in the *Gaia*-unWISE AGN catalogue (Shu et al. 2019) for each detected point source (`EXT=0`) within a maximum separation of 30″using an iterative 3σ clipping algorithm. We found 1600–1900 matches per ObsID and calculated the right ascension (RA; α) and declination (DEC; δ) offsets $\Delta\alpha$, $\Delta\delta$ for each ObsID. This removes the mean linear offsets of the X-ray sources from the *Gaia* DR2 positions. We applied the computed corrections (see Table 1) to the observation attitude and then recalculated the event coordinates using the tasks `evatt` and `radec2xy` (see Appendix A.2).

After applying astrometric corrections to each of the four observations separately, we merged them into one and applied filters of `FLAG=0xc00fff30` (this selects good events from the nominal field of view, excluding bad pixels) and `PATTERN≤ 15` (this includes single, double, triple, and quadruple events). In the merged events, we searched for background flares by running `flaregti` (see Appendix A.5) in the 0.2–5 keV band[12]. Only one short (< 1 ks) significant flare was detected by `flaregti` in the entire dataset (algorithm described in Appendix A.5). We applied the `flaregti` filter with `evtool` (see Appendix A.3) to filter the background flare out.

Using `evtool`, we extracted the events in specific energy ranges and created images with a resolution of 4″0 per pixel. This is a factor 2.4 higher than the size of the physical pixels of the eROSITA cameras (9″6). We created vignetted exposure maps in each band and an unvignetted exposure map using `expmap`. A source detection mask was created with `ermask` on the basis of the 0.2–2.3 keV band vignetted exposure map by applying a minimum cut at 1% of the maximum value. By adopting such a low exposure threshold, we included the field border in the analysis. The depth is significantly shallower than the main field in the border. These shallow border regions can be excluded when necessary (§ 4.2).

3.3. Source detection

We created the main eFEDS catalogue by running a single-band source detection in the 0.2–2.3 keV band using all TMs. This band guarantees the highest sensitivity given the shape of the eROSITA response (Predehl et al. 2021). This was shown by simulation tests (Liu et al. 2021c). In addition, we also ran the same source detection procedure simultaneously in three bands (0.2–0.6, 0.6–2.3, and 2.3–5 keV), in order to select sources with particularly hard (or soft) spectra. As the eSASS tasks are able

---

[9] charge-coupled device sensors of the eROSITA cameras
[10] As the eFEDS field is scanned uniformly, the ratio of the unvignetted exposure time and the 0.2-2.3 keV vignetted exposure time is almost a constant (1.86) across the field, except at the field border.
[11] available for download at
https://erosita.mpe.mpg.de/edr/eROSITAObservations/

---

[12] For `flaregti`, we adopt a source diameter of 120″, a source likelihood threshold of 10, a time bin size of 100s, and 60 grid points per dimension.

Table 1. eROSITA observations of the four sub-fields of eFEDS

	ObsID	Central R.A. [deg]	Central Dec [deg]	t_{exp} [s]	Start time [UTC]	$\Delta\alpha$ [arcsec]	$\Delta\delta$ [arcsec]
I	300007	129.55	+1.5	89642	2019-11-03T02:25:50	5.0	-4.2
II	300008	133.86	+1.5	89642	2019-11-04T04:05:52	5.0	-3.2
III	300009	138.14	+1.5	89642	2019-11-05T05:45:54	5.1	-3.5
IV	300010	142.45	+1.5	89642	2019-11-06T07:25:56	4.6	-4.0

Notes. $\Delta\alpha$, $\Delta\delta$: the corrections that were applied to the raw attitude solutions for each ObsID in order to remove systematic linear offsets of derived X-ray source positions in right ascension (R.A.; α) and declination (DEC; δ), using the *Gaia*-unWISE AGN candidate catalogue of Shu et al. (2019) as an astrometric reference. See section 3.1 for details.
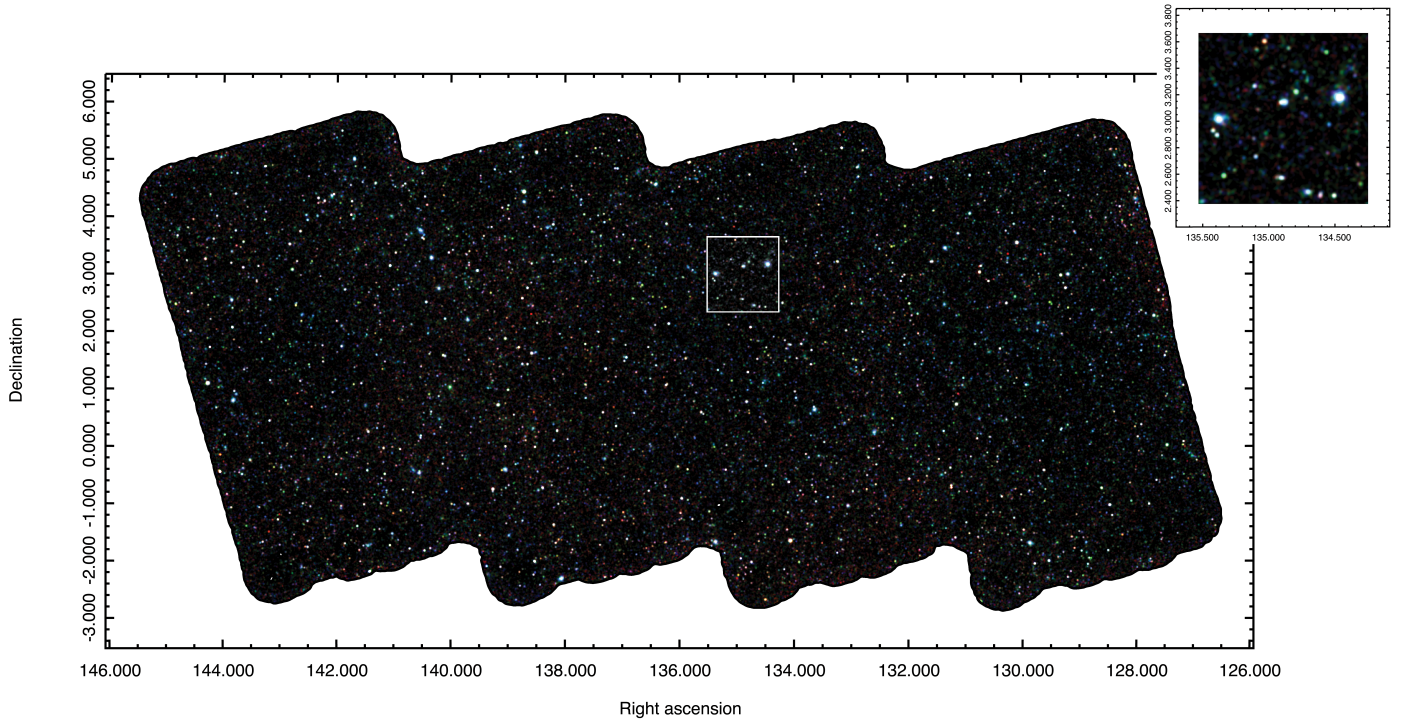
Fig. 2. RGB image of the eFEDS field in X-rays, created using 0.2–0.5 (R), 0.5–1 (G), and 1–2 (B) keV bands. Each count rate image has been smoothed with a Gaussian with $\sigma = 10$ pixels (40″). The inset in the upper right corner shows a zoom-in around a newly discovered supercluster (Liu et al. 2021a).

to handle either one set of files (e.g. event file, image, or background map) in a particular energy band or multiple sets of files in a few bands, the source detection procedure is identical for the single-band detection and the three-band detection. The source detection procedure is illustrated in Figure 3 and is described below. We present the adopted values of the key parameters of the eSASS tasks we used here. They are described in more detail in Appendix A.

Our core source detection algorithm selects source candidates according to the statistics of fitting source images with a PSF-convolved model (beta model or δ function), which is called PSF-fitting hereafter. Before the PSF-fitting, a preliminary catalogue that contains all the potential source candidates must be prepared. To prepare thispreliminary catalogue, a background map must be available. We therefore first ran `erbox`[13] in local mode (without background map) and adopted a detection likelihood threshold of 6 and a box size of 9 pixels (`boxsize=4`) to create an initial catalogue. This initial catalogue, in which the quality is not controlled, was only used to create a background map using `erbackmap` (adopting a likelihood threshold

of 6 and a required signal-to-noise ratio of 40), which masks out the sources in the catalogue and then adaptively smooths the image to create the background map. After preparing the background map prepared, we ran `erbox` in map mode, adopting a detection likelihood of 4 and a box size of 9 pixels, to create an updated catalogue. Considering that this updated catalogue is affected by the settings that were adopted when the initial catalogue was adopted through the background map, we repeated the step of creating a background map and running `erbox` with the same parameters and updated the catalogue again, creating the preliminary catalogue. This preliminary catalogue is determined by the parameter settings used in the background map creation and the map-mode `erbox` detection. Using the preliminary catalogue, we created a final background map using the same parameters as above. The bottom panel of Figure 1 shows the resulting background map in the 0.2-2.3 keV band.

Finally, the preliminary catalogue was used as input to the PSF-fitting in the next step, which selects reliable sources from this catalogue. We ran photon-mode PSF-fitting using `ermldet` and adopted a PSF-fitting radius (`cutrad`) of 15 pixels, a multiple-source searching radius (`multrad`) of 20 pixels, a detection likelihood threshold (`likemin`) of 5, an extent likelihood

---

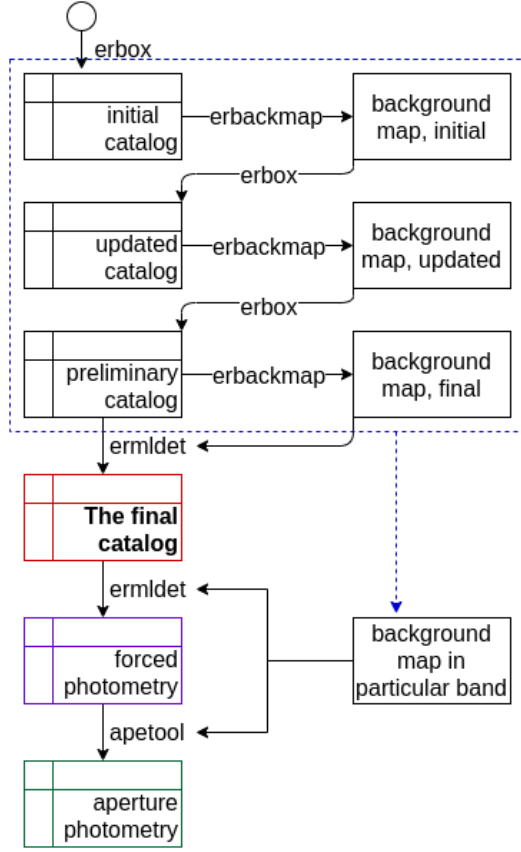[13] For `erbox`, we always adopted two-step image rebinning (`nruns=2`).
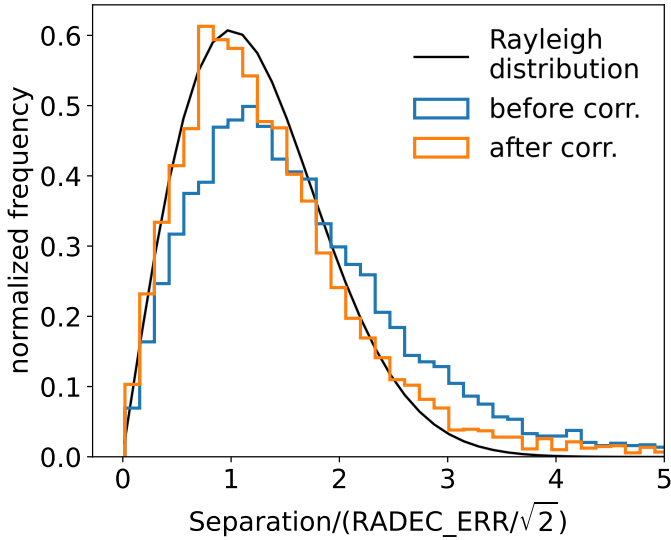
Fig. 3. Source detection procedures.

Fig. 4. Distributions of the separations between the optical positions of the *Gaia*-unWISE AGN catalogue and the X-ray positions before and after the second-pass corrections.

threshold (`extlikemin`) of 6, an extent range between 2 and 15 pixels, a maximum of four sources for simultaneous fitting, and allowed a source to be split into two sources at most. We finally used `catprep` to format the final catalogue. The preliminary catalogue (containing 84565 sources in the single-band detection and 58227 sources in the three-band detection) is much larger than the PSF-fitting output catalogue. As mentioned before, the PSF-fitting itself was performed within a circular region with a
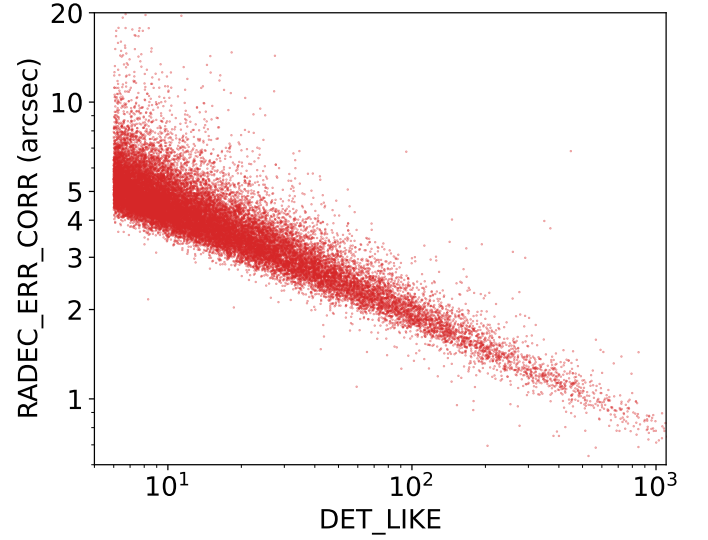

Fig. 5. Positional uncertainty (after astrometric correction, `RADEC_ERR_CORR`) of all point sources in the main catalogue as a function of the detection likelihood in the 0.2-2.3 keV energy band.

radius that was fixed by `cutrad`, which is an essential parameter of `ermldet`. Through simulation tests, we determined that `cutrad=15` is a good choice. Adjusting it does not improve the detection efficiency of point sources significantly.

By comparing the best-fit source model with a zero-flux (pure background) model, `ermldet` calculates a detection likelihood (`DET_LIKE`) L for each source, defined as $L = -\ln P$, where P is the probability of the source being caused by random background fluctuation. By comparing the extended beta model with a δ function, `ermldet` also calculates an extent likelihood `EXT_LIKE` for each source, which is defined corresponding to the probability of a source being unresolved rather than extended. Sources with an extent likelihood above and below `extlikemin` (6) were fitted with the beta model and δ function, respectively. These sources have catalogue values `EXT=0.0` and `EXT_LIKE=0.0`. A low threshold of `extlikemin=6` was chosen in order to achieve a high completeness of the extended source sample, allowing some point sources to be misclassified as extended (Liu et al. 2021c). To increase the completeness of the point source catalogue or the purity of the extended source catalogue, a higher cut on `EXT_LIKE`, such as 8 or 12, can be applied when needed. By minimising the C-statistic, `ermldet` measures the source position, extent, and count rate, together with the 68% confidence intervals. The task determines error margins by varying each fit parameter from the best-fit position in both directions to find the points at which the likelihood function $\Delta C = C - C_{\mathrm{best}}$ reaches the value 1.0. The error values from both directions are averaged and a single error value is written to the output file. For this reason, the error values and the combined errors for faint sources with large uncertainties have to be treated with caution. In the case of multi-band detection, the total-band count rates `ML_RATE_0`, counts `ML_CTS_0`, and fluxes `ML_FLUX_0` have the errors of the single-band rates, counts, and fluxes added in quadrature, treating them as Gaussian errors. The algorithm of `ermldet` is described in more detail in Appendix A.5.

For the single-band detection, the 0.2-2.3 keV event file, image, and vignetted exposure map were used. For the three-band detection, the background maps were created using the same masking catalogue, but separately in the three bands; and the

event files, images, exposure maps, and background maps of the three bands are input simultaneously into `erbox` and `ermldet` for source detection. The vignetted exposure map was used except for the 2.3–5 keV band, where the unvignetted exposure map was used instead because the background in the hard band is dominated by high-energy particles (Predehl et al. 2021).

After source detection, we matched the X-ray sources to the *Gaia*-unWISE AGN catalogue (Shu et al. 2019) , for which we again adopted a maximum separation of 30″. we display the X-ray to optical positional separation in Figure 4. We made a second-pass astrometric correction of the catalogue in right ascension and declination, allowing a linear shear term rather than considering only linear shift (as was done in § 3.2). Simulations show that the direct measurements of the positional uncertainty are slightly underestimated (Liu et al. 2021c). We therefore also computed an empirical correction of the raw positional uncertainty estimates $\Delta\theta$, that brings the distribution of the observed X-ray to reference catalogue offsets closer to the expected Rayleigh distribution. The computed corrections are as follows:

$$\alpha_{\mathrm{CORR}} = \alpha - (-0.2158 \times \delta + 0.4526)/\cos(\delta)/3600$$
$$\delta_{\mathrm{CORR}} = \delta - (0.086 \times \delta + 0.0679)/3600$$
$$\Delta\theta_{\mathrm{CORR}} = 1.15 \times \sqrt{\Delta\theta^2 + 0.7^2}. \quad (1)$$

The second-pass astrometric correction results in a much better consistency between the separation distribution and the Rayleigh distribution, although a tail beyond the Rayleigh distribution is still visible at large separations. This is mainly caused by faint sources, which have relatively higher probabilities to be spurious and relatively larger positional uncertainties, and could more easily have underestimated positional uncertainties or false optical counterparts. The mean and median positional uncertainty after this correction are 4″.73 and 4″.65, respectively. Figure 5 shows the corrected positional uncertainty as a function of the source detection likelihood in the 0.2–2.3 keV band. α_{CORR}, δ_{CORR}, and $\Delta\theta_{\mathrm{CORR}}$ are added to the source catalogues as columns `RA_CORR`, `DEC_CORR`, and `RADEC_ERR_CORR`, respectively.

3.4. Forced photometry

We ran forced PSF-fitting photometry in seven energy bands: 0.2–0.5, 0.5–1, 1–2, 2–4.5, 0.5–2, 2.3–5, and 5–8 keV (Table D.1). The procedure is described as follows.

Before the forced PSF-fitting, a background map has to be created for each band. We used the same method as described above for the source detection band, as illustrated in the dashed blue box in Figure 3. The only difference is that we did not need to create an initial catalogue again, as we were able to use the preliminary catalogue generated in the single-band detection. Running the `erbackmap` background map creating and the `erbox` detection procedure twice and using the same parameters as above, we created a preliminary catalogue for each band. Using this catalogue, we created the final background map for each band. Instead of creating a preliminary catalogue as described above for each band, an even simpler method is using the single-band detected catalogue directly, which will lead to similar results. However, we adopted this more complex method because of the advantage of having the background in each band determined by only the data in the same band. In this way, the background only depends on the adopted task parameters.

The preliminary catalogue of each band was only used to create the background map. When this was done, we input the full single-band detected (or three-band detected) catalogue into

Table 2. Sample selection criteria

Single-band detection [0.2–2.3 keV]			
catalogue	DET_LIKE	EXT_LIKE	Sources
Full	$\geqslant 5$	$\geqslant 0$	32684
Main	$\geqslant 6$	$\geqslant 0$	27910
Supplementary	< 6	$\geqslant 0$	4774
Point Sources	$\geqslant 6$	$= 0$	27369
Extent-Selected	$\geqslant 5$	$\geqslant 6$	542
Three-band detection			
catalogue	DET_LIKE_3	EXT_LIKE	Sources
Hard [2.3-5 keV]	$\geqslant 10$	$= 0$	246

Notes. The single-band detection produces the main X-ray catalogue, together with the supplementary catalogue of faint sources that are less reliable. The point-source catalogue and extended-source catalogues are selected from the main catalogue. The three-band detection results in sources that mostly are already in the single-band detected catalogue, and they are only used to select hard sources focusing on the 2.3-5 keV band.

`ermldet` [14] PSF-fitting, but fixed the source position and extent. This forced PSF-fitting only provides a count rate measurement for each of the chosen bands. The results of the forced fitting are given for all input sources, but we note that many of the single-band measurements are not significant, in particular for the hard bands beyond 2 keV.

For the two harder bands (2.3–5 and 5–8 keV), we used an unvignetted exposure map to generate the background map. In the forced PSF-fitting, however, the vignetted exposure map was used, so that the measured count rate is corrected for vignetting. Fluxes in each forced photometry energy band were computed by assuming a power-law spectrum with a spectral index Γ = 2.0 and a galactic absorption of N_H = 3×10^{20} cm^{-2} (HI4PI Collaboration et al. 2016). Liu et al. (2021b) performed spectral analysis for all the sources in the main catalogue, and found that this choice of (average) spectral index leads to an unbiased flux estimation for the whole sample, although a residual uncertainty is caused by the variety of spectral shapes.

3.5. The catalogues

When a low detection likelihood threshold of 5 is adopted, the single-band detection results in a large sample of 32684 sources, which includes a relatively high fraction of spurious sources (see § 4.1 for details). Despite the high spurious fraction, we adopted this low threshold because many faint but potentially interesting sources can be detected, possible cases of blended faint sources can be verified by multiple PSF-fitting, and faint sources can be effectively masked out when the properties of nearby sources are measured. As discussed in Liu et al. (2020), our strategy is to adopt a low threshold in the source detection and to apply further likelihood filtering on the output as needed. In this paper, we selected the 27910 single-band detected sources with a detection likelihood $\geq$ 6 as constituting the eFEDS Main catalogue. The 4774 sources with detection likelihood < 6 are kept in a supplementary catalogue, which we also publish here. The selection criteria for the sub-catalogues are listed in Table 2. The main catalogue contains 27369 point sources (extent likelihood

---

[14] We used `ermldet`-1.47 here because a bug that affects the forced PSF-fitting in the c001 version of `ermldet` is solved in this updated version.
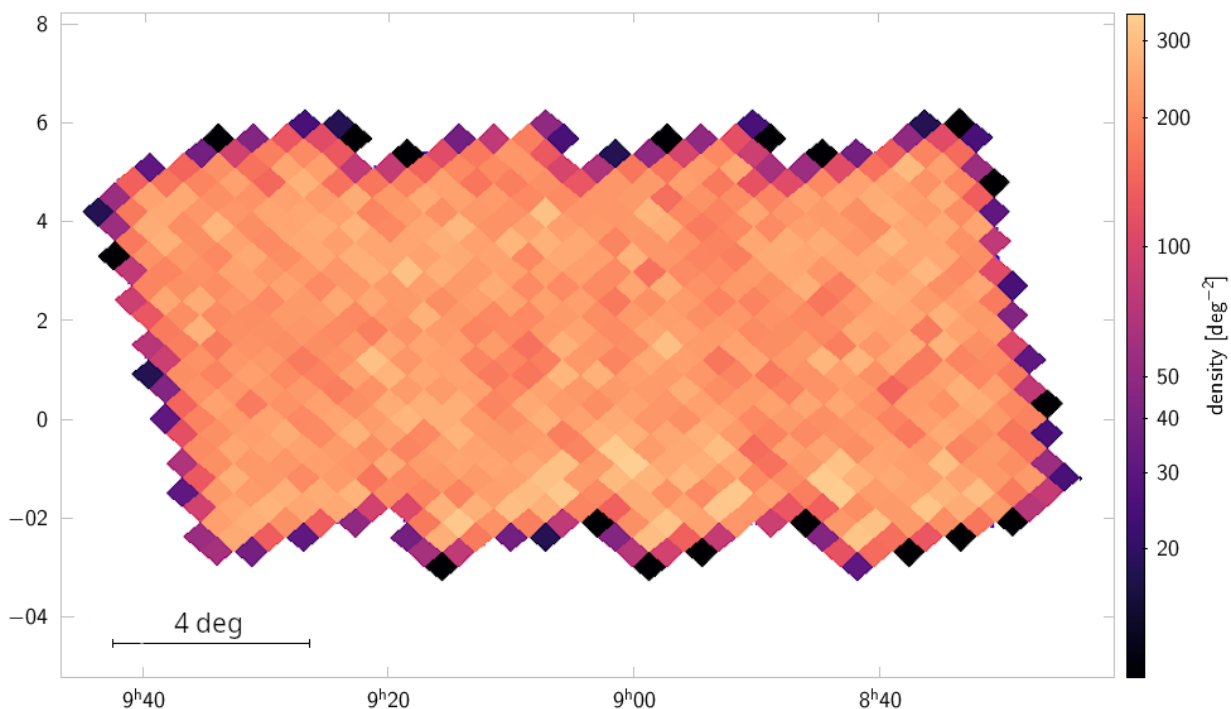
Fig. 6. HEALpix map (`Nside=128`) showing the projected sky density of the eFEDS main catalogue point sources in equatorial coordinates. Each pixel has a size of ≈ 0.2098 deg^2.

=0), and the full single-band detected catalogue contains 542 extended sources.

Figure 6 shows a HEALpix[15] map of the eFEDS field, colour-coded with the projected sky density of the eFEDS main catalogue point sources. In the well-exposed part of the field, the catalogue provides a quite uniform source density across the entire covered area (with an average of about 200 sources per deg^2). As a general term of reference, this average source density is about 70 times higher than that of the ROSAT All-Sky survey (Boller et al. 2016), about 10 times higher than that of eRASS1, and about a factor 2.5 lower than that of the XMM-XXL survey (Chiappetti et al. 2018).

As shown by simulation, the three-band detection is not as efficient as the single-band detection for the main population of X-ray sources, which have soft spectral shapes. In this work, the only aim of the three-band detection is to select hard sources (particularly AGN) that are detected above 2.3 keV. We used three bands rather than only the 2.3-5 keV hard band because most of the hard-band observed sources have much stronger signals in the soft band. Including the soft signals improves the measurement accuracy of source position and extent. The source detection likelihood is meanwhile independently calculated in each of the three bands, allowing us to focus only on the hard band. From the three-band detected sources, we selected a sample of 246 sources with an extent likelihood of 0 and a 2.3–5 keV band detection likelihood $\geqslant 10$ as the *Hard eFEDS* catalogue. The three-band detected sources are not published except for this small *Hard* sample. We adopted a high 2.3–5 keV `DET_LIKE` threshold of 10 in order to guarantee a high purity of the sample. Above this threshold lie eight significantly extended sources with `EXT_LIKE`>47, and all the other 246 sources have `EXT_LIKE`=0. Only 20 of these 246 sources are not included in the single-band detected main catalogue. The number is small because eROSITA has a small effective area and a high particle

---
[15] http://healpix.sourceforge.net

background above 2.3 keV. However, we remark that this sample is valuable for AGN demography studies. Although most of them are already in the main catalogue, these sources are detected independently and thus have independent position and fluxe measurements. Most importantly, they have a well-defined selection function, which is different from that of the main sample and is quantified through simulation.

The content of the catalogues is described in detail in Appendix D. Figure 7 displays the distributions of fluxes converted from the count rates. The energy conversion factor (ECF; listed in Table. D.1) between the count rates and fluxes is based on a power-law model with Γ =2.0 and with Galactic absorption (N_H=3 × 10^{20} cm^{-2}).

Using `ersensmap`, we calculated the sensitivity for point sources in the 0.2–2.3 keV band. As a reference, we compare in Figure 8 the sky coverage of the main sample (detection likelihood $\geqslant 6$) and of the full single-band detected sample (detection likelihood $\geqslant 5$), converted into the 0.5–2 keV band, with that of a few contiguous *Chandra* and *XMM-Newton* surveys, including the XMM-XXL North survey (Liu et al. 2016), the *Chandra* COSMOS Legacy survey (Civano et al. 2016), the XMM-RM survey (Liu et al. 2020), and the CDWFS survey (Masini et al. 2020).

In this catalogue, we detect 542 candidate extended sources with a detection likelihood ≥ 5 and extent likelihood ≥ 6 (see Table 2). This number corresponds to a density of about four extended sources per square degree over the full eFEDS field. The extent and detection likelihoods of this sample are shown in Figure 9. The emission from the detected extended sources was fit using a beta model with a slope fixed to 2/3 and r_{core} as the extent parameter after convolution with the PSF. The fluxes and count rates were then calculated from the normalisation of this fit model and hence correspond to the total integral over the beta model. During this process, a constant temperature of the intra-cluster medium and a constant value of
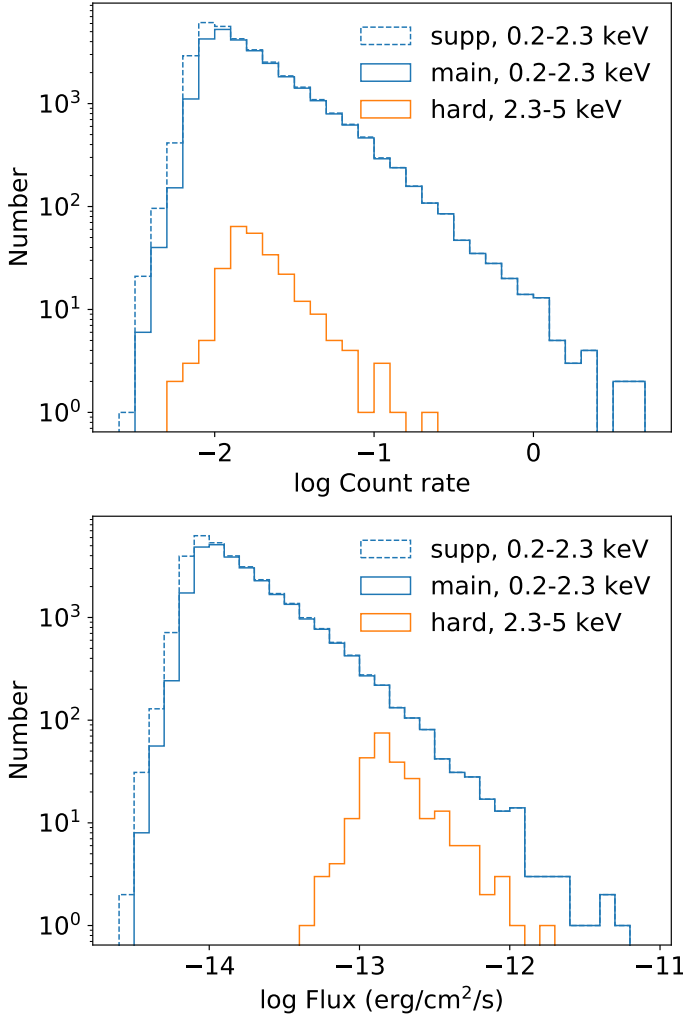
Fig. 7. Count rate (counts/s; measured by PSF-fitting) and flux (derived from count rate) distributions of the main (solid blue) and the supplementary (dashed blue) catalogue in the 0.2–2.3 keV band and of the hard catalogue (orange) in the 2.3–5 keV band.
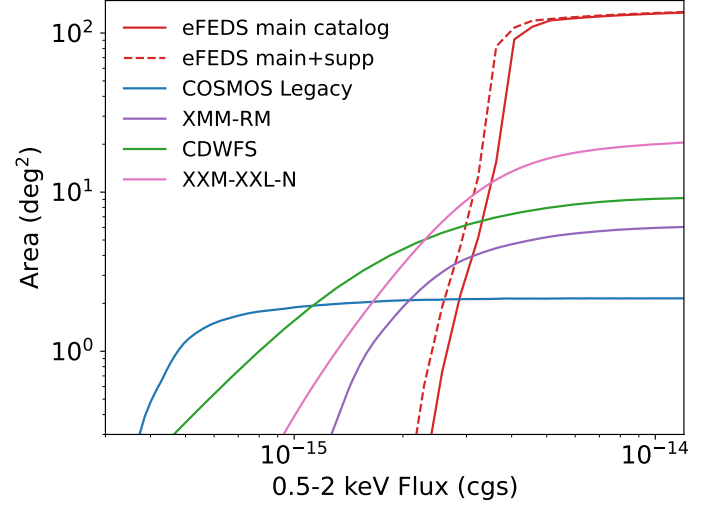
Fig. 8. eFEDS sky coverage area (red) as a function of 0.5–2 keV flux calculated using `ersensmap` at two detection likelihood threshold 5 (for the full single-band sample) and 6 (for the main catalogue). The sky coverage of a few previous contiguous X-ray surveys are plotted for comparison.
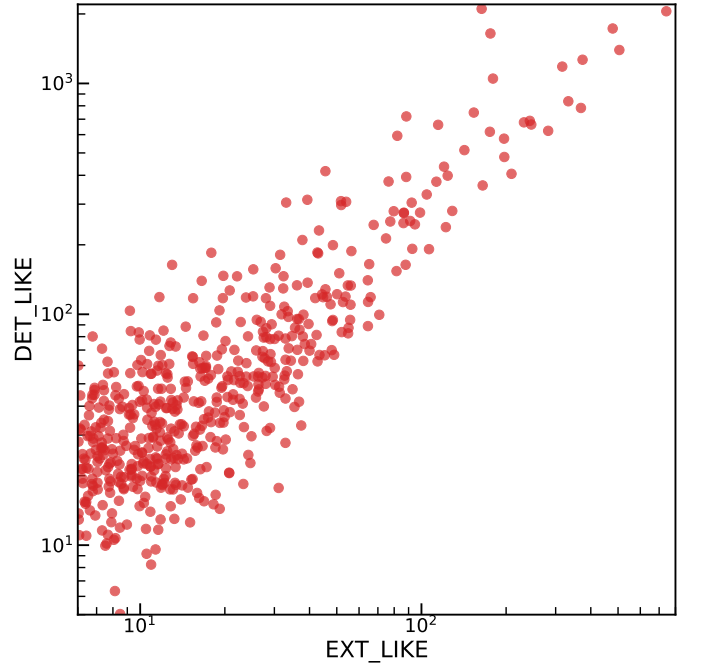

Fig. 9. Detection likelihood as a function of extent likelihood of 542 extended source candidates in the eFEDS catalogue.

the Galactic absorption (see above) were assumed. Our source detection algorithm automatically deconvolves the source flux when a nearby point source lies within a radius marked 'multrad'. If the multrad parameter is smaller than the physical radius within which the flux of the extended source is measured, emission from the point source might contaminate the signal from the extended source. Furthermore, a diligent analysis of the instrumental background and X-ray foreground has to be performed when the extended source flux is calculated because the source detection algorithm might overestimate the background if the source is very extended. Therefore, the fluxes and count rates given in the main eFEDS catalogue must be treated with extreme care. We provide the corrected count rate, flux, and luminosity measurements of the full extent-selected sample at two fixed radial distances (300 kpc and 500 kpc) in Liu et al. (2021a) and at physical overdensity radius of R_{500} in Bahar et al. (2021). The extended sources in the sample are characterised in Ghirardini et al. (2021a), and the optical counterparts are given in Klein et al. (2021).

4. Characterisation of the source detection procedure and catalogue properties

4.1. Completeness and contamination

Detection of celestial sources in the low-count regime (as in this case and in most of X-ray astronomy) is a particular realisation of a stochastic process. Each catalogue generated by such a source detection procedure is statistical in nature, and is always plagued by spurious contaminating sources, as well as by incompleteness or by missed sources. A high level of com-
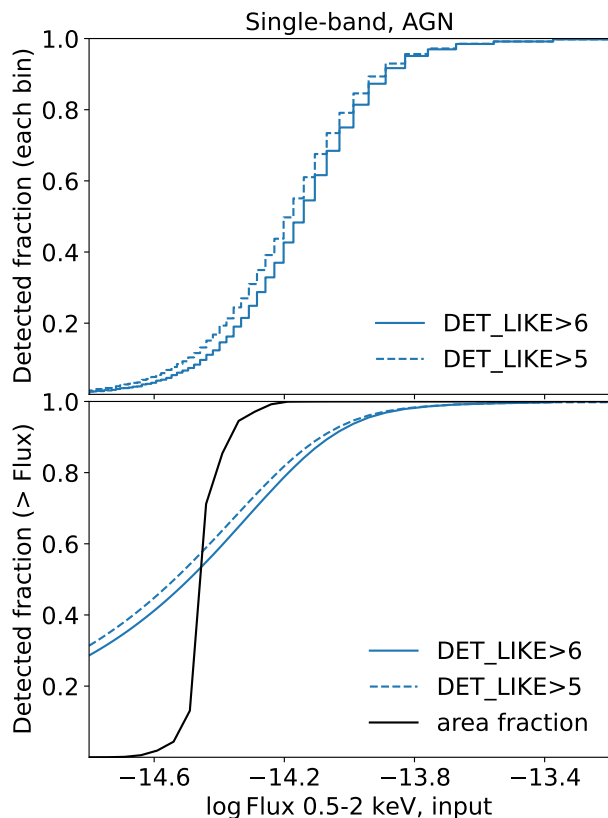
Fig. 10. Simulation-measured completeness of AGN in the single-band detected catalogue as a function of the 0.5–2 keV input flux (erg/cm²/s) in differential (upper panel) and cumulative manners. The solid and dashed blue lines indicate the detected AGN with `DET_LIKE>6` and `DET_LIKE>5`, respectively. For comparison, we also plot the sky coverage area curve measured by `ersensmap` adopting `DET_LIKE⩾5` normalised to a total area of 1 (black line).

pleteness (fraction of detected true sources above a give threshold) achieved at a low contamination level (fraction of spurious sources in the catalogue) is the essential figure of merit of any source detection procedure.

We have measured the completeness and contamination of the eFEDS catalogue through an extensive series of simulations (Liu et al. 2021c, summarised in Appendix E). We briefly describe the main outcome of this analysis here.

Fig. 10 displays the detected fraction of input point sources (AGN and stars) in a differential (top panel) and cumulative (bottom panel) manner. At a threshold of `DET_LIKE> 5` and `DET_LIKE>6`, 94% and 93%, respectively, of the simulated point sources are detected down to a 0.5–2 keV flux limit of 10^{-14} erg cm^{-2} s^{-1}. Down to a flux limit of 4×10^{-15} erg cm^{-2} s^{-1}, the completeness reduces to 63% for `DET_LIKE>5` and 59% for `DET_LIKE>6`. To guarantee a 80% completeness, the flux limit is 6×10^{-15} erg s^{-1} cm^{-2} for `DET_LIKE>5` and 6.5×10^{-15} erg s^{-1} cm^{-2} for `DET_LIKE>6`. Using `ersensmap`, we calculated the flux limit corresponding to the detection likelihood threshold (`DET_LIKE= 5`) and thus the sky coverage area curve as a function of this flux limit. When this curve is normalised to a total area of 1 (solid black line in Fig. 10), this function also predicts the detectable fraction. However, this fraction only reflects the `ersensmap` definition of the detectable flux limit. As displayed in Fig. 10, 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
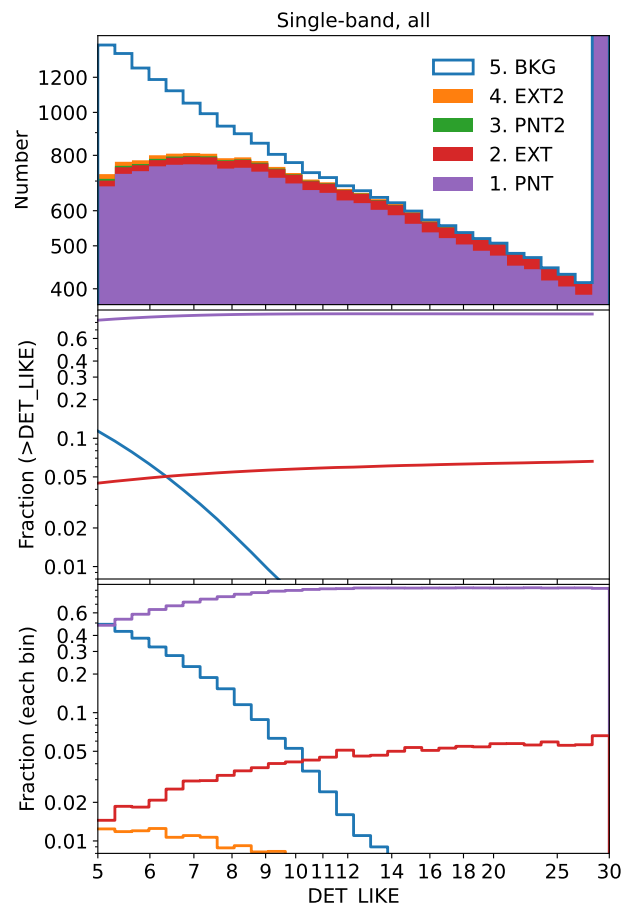
Fig. 11. Distributions of all the single-band detected sources as a function of detection likelihood. The top panel displays the histogram of the detected sources. The bottom panel displays the fraction of each class in each bin (differential distributions). The middle panel displays the fraction of each class above a given detection likelihood (cumulative distributions). The colour code for the various classes considered here is as follows: purple shows primary counterparts of input point sources, red shows primary counterparts of input clusters, green shows secondary counterparts of input point sources, orange shows secondary counterparts of input clusters, and blue shows spurious sources (background fluctuations). The fractions of green and orange sources are so low that they are not always visible.

sources. A source below the limit still has a significant probability of being detected because of fluctuation.

Fig. 11 displays the distribution of all the single-band detected sources from the simulations as a function of their detection likelihood. The detected sources are divided into five classes: the primary counterpart of an input point source (class 1, purple in Fig. 11), the primary counterpart of an input cluster (class 2, red), the secondary counterpart of an input point source (class 3, green), or an input cluster (class 4, orange), and spurious sources due to background fluctuations (class 5, blue). When one input source results in multiple detected sources, the source with the largest number of photons is considered as the primary counterpart (class 1 or 2), and the secondary sources (class 3 or 4) correspond to the signal in the outer wing of an input point sources or to substructures or fluctuations in an input cluster. More details are discussed in Liu et al. (2021c).

According to the definition, the detection likelihood `DET_LIKE` corresponds to a probability of exp(−`DET_LIKE`) of one source being spurious. This probability is too low to be plot-

ted in the bottom panel of Fig. 11; the actual spurious fraction (blue line) is significantly higher. This is because the likelihood is defined in an ideal situation, considering only Poissonian fluctuations. In reality, additional uncertainties such as background measurement or source deblending affect the source detection process in every step. The only way to measure the spurious fraction of a catalogue reliably therefore is through detailed simulation, as we have done here.

The middle panel of Fig. 11 shows that above a detection likelihood of 5, the sample includes 11.5% spurious sources. When a likelihood threshold of 6 or 8 is adopted, the spurious fraction reduces to 6.3% or 1.8%, respectively. The availability of a classification for all detected sources means that the simulation provides the fraction of any type of input in any specifically selected sample. In the suite of accompanying papers, we have made extensive use of this valuable information in order to estimate, for example, the fraction of spurious sources in the hard-band selected eFEDS point source catalogue (Nandra et al., in prep.) and the fraction of true/spurious clusters in the extended source catalogue (Liu et al. 2021a; Klein et al. 2021), and so on (see more detailed discussions in Liu et al. 2021c).

4.2. Aperture photometry and number counts

This section presents the X-ray point-source number count distribution of the eFEDS survey, which requires a knowledge of the selection function, that is, the probability of a source with a given flux to be detected. It has been demonstrated that the selection function can be estimated on the basis of aperture photometry (Georgakakis et al. 2008; Lehmer et al. 2012). Therefore, we ran aperture photometry using the `apetool` task (see Appendix A.5). `apetool` uses the aperture source and background counts to calculate a Poisson false rate for each source, that is, the probability that the source is generated by background fluctuations. This can be expressed in terms of logarithmic likelihood L as $-\ln(probability)$. This likelihood can be used for the sample selection. It can also be converted into a sensitivity map containing the minimum required number of photons to reach a given likelihood. This map can then be used to correct for the incompleteness of the eFEDS point-source catalogue as a function of X-ray flux. During the forced PSF-fitting photometry for each band, we also created a background map and a source map (source extent model convolved with PSF) in addition to calculating the fluxes. Making use of them, we ran `apetool` within a radius of 60% encircled energy fraction (EEF), adopting an aperture likelihood threshold of 12. The aperture size and likelihood threshold were selected according to simulation tests (Liu et al. 2021c). In this analysis, we excluded the border of the field, where the exposure is much lower than the typical depth of eFEDS. We adopted the inner region, in which the 0.2–2.3 keV vignetted exposure value is above 500 s. This region comprises 90% of the total area.

Based on the aperture photometry results, we converted the 0.5–2 keV count rate into flux corrected for Galactic absorption using an ECF of 1.104×10^{12} cm^2/erg, and converted the 2.3–5 keV count rate into 2-10 keV flux corrected for Galactic absorption using an ECF of 5.518×10^{10} cm^2/erg, assuming the same spectrum model as used in § 3.4, that is, a power law with a photon index of 2.0. We selected point sources with an aperture Poissonian likelihood > 12, which resulted in 13457 sources in the 0.5–2 keV band and 151 sources in the 2.3–5 keV band. Based on these two sub-samples and using the method described in Georgakakis et al. (2008), we calculated the number counts in the 0.5–2 and 2–10 keV bands and compared them with those
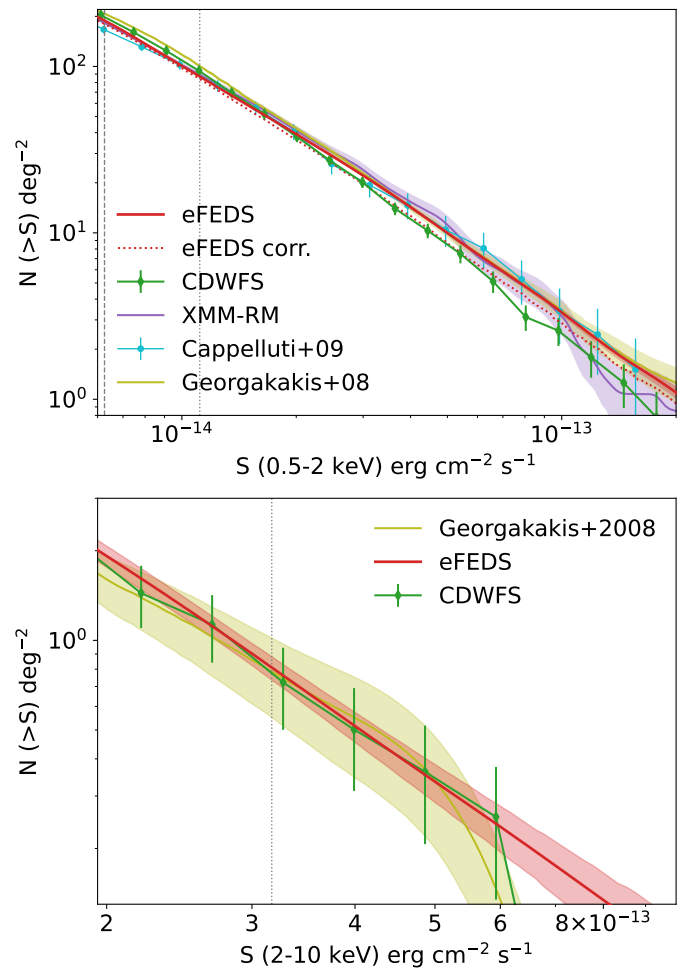
Fig. 12. Point source number counts as a function of fluxes, corrected for Galactic absorption, for the 0.5-2 keV (upper panel) and 2-10 keV (lower panel) bands. The 1σ uncertainties are estimated as the square root of the sources number. In both panels, the eFEDS curves are compared to those of Georgakakis et al. (2008) (compiled from a number of *Chandra* survey fields) and Masini et al. (2020) (from CDWFS). In the soft band, we also plot the results from the XMM-COSMOS survey (Cappelluti et al. 2009) and from the XMM-RM survey (Liu et al. 2020). In both panels, the dashed and dotted vertical lines indicate the flux limits corresponding to 10% and 50% sensitivity, respectively. In the upper panel, the dotted red line displays the number counts derived by applying corrections according to the simulation results to the distribution of the 0.5-2 keV fluxes measured by forced PSF-fitting.

measured in *Chandra* surveys (Georgakakis et al. 2008) and in the XMM-COSMOS survey (Cappelluti et al. 2009) in Fig. 12. They agree well.

Liu et al. (2021c) provided another way of measuring the point source number counts based on the detailed eFEDS simulation. The flux distribution of the output AGN catalogue is different from that of the input AGN catalogue mainly because of (i) sample incompleteness (Fig. 10), (ii) sample contamination (Fig. 11), (iii) flux overestimation caused by source blending, and (iv) Eddington bias. The ratio of the input and output flux distributions of the simulation can be used to convert the flux distribution of the real catalogue into the intrinsic number counts. The soft-band number counts derived in this way from the real point sources with `DET_LIKE>8` are displayed in Fig. 12. This is slightly lower than the number counts measured using the method based on aperture photometry because the simulation-
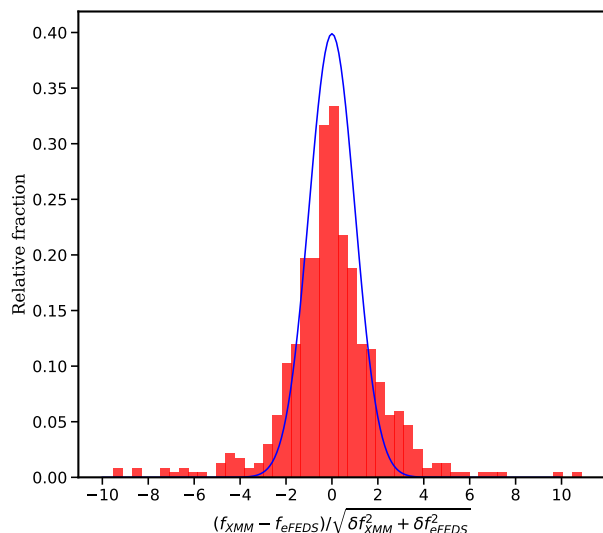
Fig. 13. Distribution of the flux difference of X-ray sources between the XMM-ATLAS and eFEDS observations normalised by the corresponding flux uncertainties added in quadrature. The blue line shows a Gaussian distribution with a mean of zero and a scatter of unity.

based method has corrected for the effects of sample contamination and source blending, which can only be quantified through simulations. These two effects are negligible in deep survey with high spatial resolution such as the CDWFS. The soft-band CD-WFS number counts are slightly lower than those of the other surveys at high fluxes probably for this reason.

The `apetool`-generated catalogues were also used to assess the photometric calibration of the eFEDS field by comparing the fluxes of the detected sources to external catalogues. The choice of aperture photometry for this application enables the statistically robust estimation of the random (shot-noise) and systematic uncertainties (e.g. Eddington bias) affecting source fluxes (e.g. Laird et al. 2009). This means that observational effects can be fully accounted for when fluxes from different experiments are compared to test cross-calibration issues. We also chose to use fluxes in the 0.5–2 keV energy interval because in this standard band, many X-ray catalogues in the literature report fluxes.

The external dataset adopted in this work is the XMM-ATLAS survey (Ranalli et al. 2015). This is one of the wide-area (6 deg^2) and shallow ($F_{0.5-2\,keV} \approx 2 \times 10^{-14}$ erg s^{-1} cm^{-2}) surveys carried out by *XMM-Newton* and it also overlaps with the eFEDS field. A custom reduction of the XMM-ATLAS survey field was used based on the methods described by Georgakakis & Nandra (2011). The advantage of using a custom analysis rather than the publicly available XMM-ATLAS catalogue is control over systematics (e.g. Eddington bias and conversion factor of counts to flux). The identification numbers of the relevant *XMM-Newton* observations are 0725290101, 0725300101, and 0725310101. They were reduced using the *XMM-Newton* Science Analysis System (SAS) version 18. Sources were detected independently in three energy intervals, 0.5–2, 2–8, or 0.5–8 keV to the Poisson false -detection threshold of $< 4 \times 10^{-6}$.

The 987 sources detected in the 0.5–2 keV band to the threshold above are relevant here. They were compared against a total of 985 eFEDS sources that lie within the XMM-ATLAS footprint and have a detection likelihood in the 0.6–2.3 keV band of `DET_LIKE`>10. This threshold was chosen to minimise spurious

detections while keeping number statistics high. The calculation of fluxes for the eFEDS and XMM-ATLAS fields was based on aperture photometry and used the Bayesian method described by Laird et al. (2009) and Georgakakis & Nandra (2011). A description of the basic flux-estimation algorithm is also provided in Boller et al. (2021). The fluxes in both samples are estimated in the 0.5–2 keV spectral band assuming a power-law spectral model with $\Gamma = 1.4$ that is absorbed by a Galactic column density of $\log N_H / \mathrm{cm}^{-2} = 20.3$. We emphasise that this is different from the spectral model adopted for the calculation of fluxes in the main eFEDS catalogue. The reason for this is that the XMM-ATLAS reduction adopts $\Gamma = 1.4$ for the determination of fluxes.

The eFEDS and XMM-ATLAS samples have 616 sources in common. They were identified by matching the two catalogues within a radius of 15″. For the sky density of the *XMM-Newton* and ATLAS sources, this threshold corresponds to $\ll 1$ spurious associations. Figure 13 compares the 0.5–2 keV fluxes of the common sources in the two surveys. It plots the histogram of the flux difference between the eFEDS and XMM-ATLAS normalised to the flux errors (68% confidence interval) added in quadrature. This distribution is compared with a Gaussian with unity variance and zero mean. There is no evidence for strong systematic offsets in Figure 13 , suggesting an overall good photometric agreement between the XMM-ATLAS and eFEDS data analysis. There are more sources with large normalised flux differences than expected for the normal distribution, however. We attribute this excess power at the wings of the histogram in Figure 13 to the intrinsic flux variability of AGN. This effect is discussed further in Boller et al. (2021).
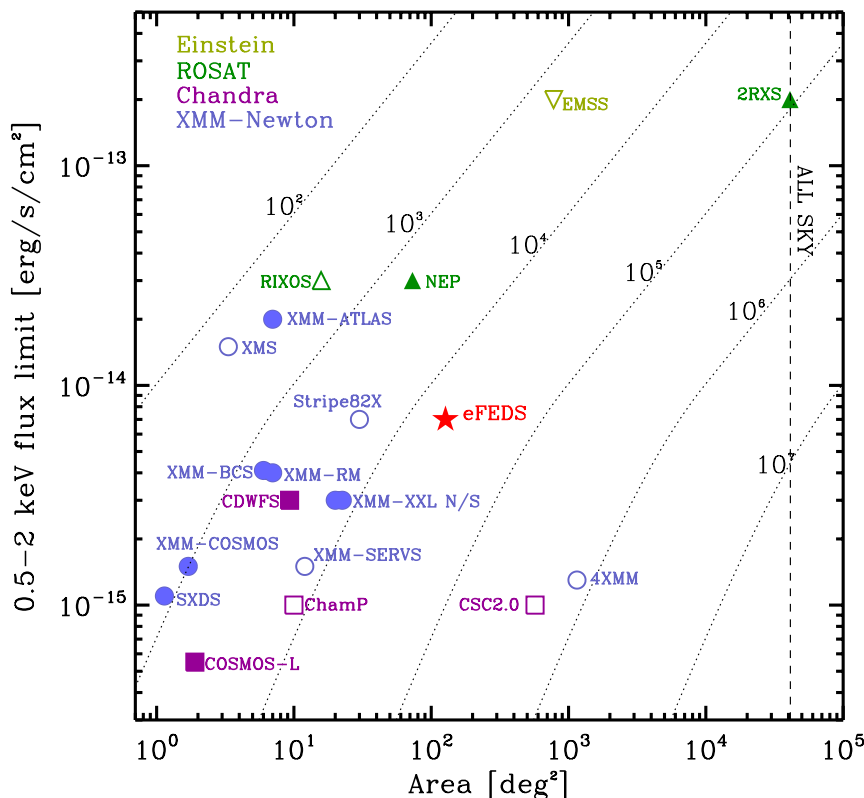
5. Conclusions

By publishing and documenting the full creation process of the catalogue of X-ray sources detected in the eFEDS field, we here complete the verification of the eROSITA design performance. We also demonstrate the ability of the instrument as a powerful survey machine for the X-ray sky.

With the exception of the all-sky surveys, the eFEDS itself represents the largest contiguous X-ray field in the soft X-ray energy range, as illustrated in Figure 14, where we show the point-source flux limit (in the 0.5–2 keV energy range) – area scatter plot of X-ray surveys larger than 1 deg^2. In terms of the sheer number of detected sources excluding all-sky surveys, eFEDS also stands out as the richest contiguous survey field to date.

This paper serves a twofold purpose: it makes the catalogues of X-ray sources detected in the eFEDS field during the *SRG*/eROSITA PV observations of 2019 public for further scientific investigations, and it describes the data processing and analysis of eROSITA field-scanning observations in a comprehensive manner for the first time, including the relevant dedicated software tools. The tools are described in the appendices.

Fig. 14. Point-sources flux limit (in the 0.5–2 keV energy range) vs. area scatter plot of a few selected X-ray surveys larger than 1 deg². Existing surveys from Einstein (light green downward triangle), ROSAT (dark green upward triangles), *XMM-Newton* (blue circles), and *Chandra* (purple squares) are shown for reference. Filled points mark contiguous surveys, and empty points show non-contiguous ones. To allow a fair comparison, we considered the total area that was covered (x-axis) for each survey and a flux limit that corresponds roughly to a completeness level of 66%. The dotted lines mark the loci of constant source numbers based on the number counts in the 0.5–2 keV energy rage of Mateos et al. (2008) (double power-law model). The surveys from left to right are The Subaru/ *XMM-Newton* Deep Survey (SXDS; Ueda et al. 2008); the XMM-COSMOS survey (Cappelluti et al. 2009); the COSMOS-Legacy survey (Civano et al. 2016); the *XMM-Newton* Medium Survey (XMS; Barcons et al. 2003); the *XMM-Newton* Bright Survey (XMM-BCS; Della Ceca et al. 2004); the XMM-RM survey (Liu et al. 2020); the XMM-ATLAS survey (Ranalli et al. 2015); the Chandra Deep Wide Field Survey (CDWFS; Masini et al. 2020); the Chandra Multiwavelength Project (ChamP; Kim et al. 2007); XMM-SERVS (Brandt 2020); the ROSAT International X-ray/Optical Survey (RIXOS; Mason et al. 2000); XMM-XXL N & S (Chiappetti et al. 2018); Stripe82X (LaMassa et al. 2016); NEP (Henry et al. 2006); the second Chandra Serendipitous Sources catalogue (CSC2.0; Civano, priv. comm.); the Einstein Observatory Extended Medium-Sensitivity Survey (EMSS; Gioia et al. 1990); the fourth *XMM-Newton* serendipitous source catalogue (4XMM; Webb et al. 2020), and the second ROSAT all-sky survey catalogue (2RXS Boller et al. 2016).

References

Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, PASJ, 70, S4
Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, arXiv e-prints, astro
Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, AARA, 49, 409
Bahar, Y. E., Bulbul, E., Clerc, N., et al. 2021, arXiv e-prints, arXiv:2110.09534
Barcons, X., Carrera, F. J., Ceballos, M. T., et al. 2003, Astronomische Nachrichten, 324, 44
Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28
Boller, T., Freyberg, M. J., Trümper, J., et al. 2016, A&A, 588, A103
Boller, T., Schmitt, J. H. M. M., Buchner, J., et al. 2021, arXiv e-prints, arXiv:2106.14523
Brandt, W. N. 2020, in American Astronomical Society Meeting Abstracts, Vol. 235, American Astronomical Society Meeting Abstracts #235, 138.01
Brusa, M., Urrutia, T., Toba, Y., et al. 2021, arXiv e-prints, arXiv:2106.14525

Buchner, J., Boller, T., Bogensberger, D., et al. 2021, arXiv e-prints, arXiv:2106.14529
Bulbul, E., Liu, A., Pasini, T., et al. 2021, arXiv e-prints, arXiv:2110.09544
Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, A&A, 497, 635
Cash, W. 1979, ApJ, 228, 939
Chiappetti, L., Fotopoulou, S., Lidman, C., et al. 2018, A&A, 620, A12
Chiu, I.-N., Ghirardini, V., Liu, A., et al. 2021, arXiv e-prints, arXiv:2107.05652
Civano, F., Marchesi, S., Comastri, A., et al. 2016, ApJ, 819, 62
Comparat, J., Eckert, D., Finoguenov, A., et al. 2020, The Open Journal of Astrophysics, 3, 13
Della Ceca, R., Maccacaro, T., Caccianiga, A., et al. 2004, A&A, 428, 383
Dennerl, K., Andritschke, R., Bräuninger, H., et al. 2020, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11444, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 114444Q
Dennerl, K., Burkert, W., Burwitz, V., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8443, Space Telescopes and Instrumentation 2012: Ultraviolet to Gamma Ray, ed. T. Takahashi, S. S. Murray, & J.-W. A. den Herder, 844350
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168
Drinkwater, M. J., Byrne, Z. J., Blake, C., et al. 2018, MNRAS, 474, 4151
Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, Astronomy and Geophysics, 50, 5.12

Ebermayer, S., Andritschke, R., Elbs, J., et al. 2010, in High Energy, Optical, and Infrared Detectors for Astronomy IV, ed. A. D. Holland & D. A. Dorn, Vol. 7742, International Society for Optics and Photonics (SPIE), 254 – 263

Fisher, R. A. 1932, Statistical methods for research workers, 4th ed. (Edinburgh: Oliver and Boyd)

Georgakakis, A. & Nandra, K. 2011, MNRAS, 414, 992

Georgakakis, A., Nandra, K., Laird, E. S., Aird, J., & Trichas, M. 2008, MNRAS, 388, 1205

Ghirardini, V., Bahar, E., Bulbul, E., et al. 2021a, arXiv e-prints, arXiv:2106.15086

Ghirardini, V., Bulbul, E., Hoang, D. N., et al. 2021b, A&A, 647, A4

Gioia, I. M., Maccacaro, T., Schild, R. E., et al. 1990, ApJS, 72, 567

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759

Henry, J. P., Mullis, C. R., Voges, W., et al. 2006, ApJS, 162, 304

HI4PI Collaboration, Ben Bekhti, N., Flöer, L., et al. 2016, A&A, 594, A116

Kim, M., Kim, D.-W., Wilkes, B. J., et al. 2007, ApJS, 169, 401

Klein, M., Oguri, M., Mohr, J. J., et al. 2021, arXiv e-prints, arXiv:2106.14519

Kuijken, K., Heymans, C., Dvornik, A., et al. 2019, A&A, 625, A2

Laird, E. S., Nandra, K., Georgakakis, A., et al. 2009, ApJS, 180, 102

LaMassa, S. M., Urry, C. M., Cappelluti, N., et al. 2016, ApJ, 817, 172

Lehmer, B. D., Xue, Y. Q., Brandt, W. N., et al. 2012, ApJ, 752, 46

Liu, A., Bulbul, E., Ghirardini, V., et al. 2021a, arXiv e-prints, arXiv:2106.14518

Liu, T., Buchner, J., Nandra, K., et al. 2021b, arXiv e-prints, arXiv:2106.14522

Liu, T., Merloni, A., Comparat, J., et al. 2021c, arXiv e-prints, arXiv:2106.14528

Liu, T., Merloni, A., Simm, T., et al. 2020, ApJS, 250, 32

Liu, Z., Merloni, A., Georgakakis, A., et al. 2016, MNRAS, 459, 1602

Masini, A., Hickox, R. C., Carroll, C. M., et al. 2020, ApJS, 251, 2

Mason, K. O., Carrera, F. J., Hasinger, G., et al. 2000, MNRAS, 311, 456

Massey, R., Heymans, C., Bergé, J., et al. 2007, MNRAS, 376, 13

Mateos, S., Warwick, R. S., Carrera, F. J., et al. 2008, A&A, 492, 51

Merloni, A., Nandra, K., & Predehl, P. 2020, Nature Astronomy, 4, 634

Pillepich, A., Reiprich, T. H., Porciani, C., Borm, K., & Merloni, A. 2018, MN-RAS, 481, 613

Predehl, P., Andritschke, R., Arefiev, V., et al. 2021, A&A, 647, A1

Ramos-Ceja, M. E., Oguri, M., Miyazaki, S., et al. 2021, arXiv e-prints, arXiv:2109.07836

Ranalli, P., Georgantopoulos, I., Corral, A., et al. 2015, A&A, 577, A121

Refregier, A. 2003, MNRAS, 338, 35

Refregier, A. & Bacon, D. 2003, MNRAS, 338, 48

Salvato, M., Wolf, J., Dwelly, T., et al. 2021, arXiv e-prints, arXiv:2106.14520

Schneider, P. C., Freund, S., Czesla, S., et al. 2021, arXiv e-prints, arXiv:2106.14521

Shu, Y., Koposov, S. E., Evans, N. W., et al. 2019, MNRAS, 489, 4741

Sunyaev, R., Arefiev, V., Babyshkin, V., et al. 2021, arXiv e-prints, arXiv:2104.13267

Toba, Y., Liu, T., Urrutia, T., et al. 2021, arXiv e-prints, arXiv:2106.14527

Ueda, Y., Watson, M. G., Stewart, I. M., et al. 2008, ApJS, 179, 124

Vulic, N., Hornschemeier, A. E., Haberl, F., et al. 2021, arXiv e-prints, arXiv:2106.14526

Webb, N. A., Coriat, M., Traulsen, I., et al. 2020, A&A, 641, A136

Wells, D. C., Greisen, E. W., & Harten, R. H. 1981, A&AS, 44, 363

Wolf, J., Nandra, K., Salvato, M., et al. 2021, A&A, 647, A5

Appendix A: eSASS data analysis software package

This appendix describes the software tasks comprising the eROSITA Science Analysis Software System (eSASS). eSASS provides a set of command-line tools for pipeline processing the eROSITA data and for performing interactive data analysis tasks. The eSASS tasks interact with the eROSITA calibration database described in Appendix B. Calibrated data products provided by the data analysis pipeline are described in Appendix C. A full description of each eSASS task including usage examples is available online[8].

Appendix A.1: X-ray event processing

evprep. This task generates an event-list file in the format agreed for eROSITA from the raw FITS[16] files created from telemetry by the archiver software. The evprep task performs many other corrections, including the detection of a variety of error and out-of-limit conditions; setting of corresponding bits of the event FLAG masks; the correction of offsets in the Camera Electronics (CE) time stamps as specified in the calibration database; the conversion of times in Good Time Intervals (GTI) and Housekeeping (HK) extensions from the low-cadence but reliable ITC time system to the higher-cadence CE time system; unpacking of the compressed event energy data for observations performed in PMENV2 mode; the exclusion of bad time intervals read from the calibration database from the sequence of GTIs; and the calculation and storage of dead time corrections.

ftfindhotpix. The task ftfindhotpix is a tool for detecting and saving the positions of bad pixels by analysing the data with statistical methods. The task can be run in either 'single' or 'block' mode, where the Poisson mean value of a single pixel (and its immediate neighbours) or a block of pixels is compared to a given threshold that denotes the maximum probability in percent for a false positive of a bad pixel. In addition, a number of additional parameters such as the number of frame bunches to consider, the minimum number of events in a frame bunch, and the minimum fraction of frames for which is a pixel is considered bad, can be optimised to decrease the probability of false positives. In the current version of the pipeline, the detection of bad pixels is turned of,f and ftfindhotpix only sets bits of the FLAG mask for events that lie on or next to a pixel listed in the calibration database as bad (various types are recognised thereof), and writes valid entries from the calibration to a BADPIX extension in the event list file.

pattern. The charge cloud released by the absorption of an X–ray photon may extend over several pixels. The task pattern tries to identify these pixels, so that the total released charge can be reconstructed. This is not always possible in a unique way, however (e.g., when the charge clouds released by two photons overlap). The general driver for the photon reconstruction in pattern is to find the simplest, most likely explanation for the observed charge distribution. If no such explanation can be found, the pattern is marked as invalid. Valid patterns consist of $1-4$ pixels, with 1 and 2 pixel patterns ('singles' and 'doubles') being the most important ones. The task pattern performs an important role in the whole processing chain because it affects key performance parameters of eROSITA, such as the spectral resolution and sensitivity, and also the spatial resolution.

---

[16] Wells et al. (1981)

energy. The task energy tries to reconstruct the energy of each detected photon from the charge distributions found in individual pixels and as a byproduct, also constrains its sub-pixel position. It makes use of the results of the pattern task, which must thus be executed beforehand. The reconstruction of the energy of an incident photon is essentially performed in three steps: First, a charge transfer inefficiency (CTI) correction: the charge loss caused by shifting the charge from its original location to the readout node must be corrected for. Second, a gain correction: the eROSITA CCDs are characterised by parallel readout, where each transfer channel is equipped with an amplifier of its own, and the amplification factors need to be individually determined. Third, a recombination of the reconstructed energies from all the pixels of a pattern. In addition to these essential steps for X–ray spectroscopy, it is also possible to obtain an improved location of the centre of the charge cloud from the individual components, so that the location where the X–ray photon had hit the CCD can be determined with sub-pixel resolution (Dennerl et al. 2012). The task energy is crucial for the whole processing chain because the absolute energy scale and the spectral resolution rely on it.

Appendix A.2: Spacecraft attitude and boresight

attprep. The task attprep converts the attitude time series from the SRG orientation sensors (SED26/1, SED26/2, BOKZ, or Q-Gyro, see also Appendix B.4) into the eSASS intermediate-attitude format. The input attitude is provided as a time series of orientation quaternions for the SED26/1, SED26/2, and Q-Gyro FITS files and as an orientation matrix for the BOKZ FITS file. The nominal cadence is 1 per second, with time tags given in SRG spacecraft clock (SCC). The intermediate attitude describes the orientation of the nominal coordinate system of the central eROSITA camera TM1 as a time series of RA (J2000), DEC (J2000), and roll angle, which is defined as the clockwise angle of the camera X-axis with respect to the north direction. The rotational transformations between the sensor coordinate systems and the TM1 camera coordinate system are stored as rotation quaternions in a calibration file for each sensor. The output FITS file contains the intermediate-attitude dataset from the primary sensor in the the first FITS extension and all sensor specific attitude datasets in additional extensions. The primary sensor is the Q-Gyro, if present, or else the SED26/1, SED26/2, or BOKZ in this order.

telatt. The task telatt calculates attitude time series that are specific for one of the telescope modules TM1 to TM7. The input attitude is read from the intermediate-attitude file written by the task attprep. The camera-specific boresight angles are read from calibration files in the form of Euler angles and are applied to the input attitude. The resulting attitude time series is written as RA (J2000), DEC (J2000), and roll angle either to a separate FITS file or as a FITS extension to the camera-specific events file.

evatt. The task evatt is used to project event positions onto equatorial sky coordinates. The task reads an event list for a single telescope module, the corresponding attitude time series (as written by task telatt), and a GTI table. For all events in the intervals specified by the GTI table, the telescope attitude is interpolated to the arrival time of each event. The event position in the detector coordinate system given in the detector pixel columns

RAWX, RAWY and sub-pixel information in the SUBX, SUBY columns of the event table is projected onto equatorial coordinates by means of a tangential projection with the interpolated attitude and using the plate scale from a calibration file.

radec2xy. The task `radec2xy` computes X and Y sky pixel coordinates (sine projection; pixel size 0″05) corresponding to the right ascension and declination (J2000) event coordinates in the eROSITA event tables. The projection centre must be specified in the command line. Sky pixel coordinates are required for image binning (`evtool`) and may be used for spectrum and light-curve extraction (`srctool`).

Appendix A.3: X-ray event binning

evtool. The task `evtool` provides three capabilities: the merging of several event lists, the filtering of events on a subset of the canonical set of columns, and the creation of images.

Merging is performed on all canonical extensions of the input files. Superfluous members of sets of identical rows found in the event tables are discarded. The task cannot correctly merge input files that have been subject to differing filterings, or that are inconsistent in a number of other tested ways; warnings are issued if such inconsistencies are detected.

The columns and types of event filtering are given as follows. The `FLAG` column can be filtered via the specification of a hexadecimal bit mask, for which the filtering sense can be chosen to be either exclusive or inclusive. The `PAT_TYP` column is filtered via an integer in the range 0-15, which is interpreted as a 4-bit mask, one bit for each of the accepted patterns. The `TM_NR` column is selected via a list of desired numbers. The `PI` column is filtered via lists of lower and upper energy bounds. The `TIME` column is filtered via a GTI specification; either directly, or by giving the base name of a GTI extension, or by specifying an external FITS file. Finally, the `RA` and `DEC` columns can be filtered via a region specification.

When `evtool` is used to create FITS images, the pixel sizes on the sky, the image dimensions in pixels, and the image centre location, are all specifiable. Some auto-sizing operations are also available. Most of the extensions in the input event list(s) are optionally discardable for image output.

srctool. The `srctool` task is responsible for the generation of standard source products, including spectra, response matrices, background spectra, and light curves. It takes as input the calibrated event file, source lists, region files, and the calibration data. The task is designed to create products that take the scanning of the telescope across the sky into account. The mechanism it uses to do this is to take a set of sample points as a function of position and as a function of energy, given a source model and extraction region. The source regions and background regions are both sampled in this way. These samples are propagated through time to take the vignetting and bad pixels as a source scans across the detectors into account, which is included in the computation of the source-specific effective area curves and in the area calculation for light curves. The GTI and exposure for a source are computed from when the sample points for a source enter and exit the field of view. Similarly, the area on the sky of the source, when used for background subtraction, and the geometric area of the extraction region are computed as the average values computed from the sample points during the GTIs. The accuracy to which the output effective areas and exposure

times is calculated depends on the parameters giving the spacing of the sample points on the sky and in time. `srctool` takes the source morphology into account by a model chosen by the user (e.g. a point source, top hat, beta model, or a provided image). If PSF losses are taken into account, the source model is convolved by the PSF to compute how much flux is lost outside the extraction region for each time step. The task supports a number of geometric regions for source extraction, including circles, ellipses, and boxes, in addition to a generic mask image option, all of which can be combined or subtracted. `srctool` also has the ability to compute extraction regions automatically given a source list. In this mode, the aim is to increase the source and background extraction circular regions until the signal-to-noise ratio is maximised, while taking into account excluding neighbouring sources.

Appendix A.4: Map creation

expmap. The `expmap` task generates FITS-format exposure maps to match the location and dimensions of a supplied template image in register. The exposure is expressed in units of the time (seconds) that each sky location was in the field of view. It can optionally be folded with the telescope vignetting function in each energy band. The algorithm samples periods of GTI in the input event list and projects the CCD onto the sky at each sample time according to the record of spacecraft attitude contained in `CORRATT` extensions. Other inputs to the maps are bad pixels as listed in the `BADPIX` extensions, dead time as recorded in the `DEADCOR` extensions, the detector mask, and the vignetting function. While the vignetting function is energy dependent, we followed the approach of *XMM-Newton* exposure maps (XMM-SAS[17]) to not weight the vignetting with an assumed spectral model. This results in systematic errors of the vignetted exposure and derived count rates and fluxes in the energy bands of interest in the few-percent range. Future versions of the `expmap` task will provide an option to fold the vignetting function with a suitable spectral model. Exposure maps for individual telescope modules and also weighted all-eROSITA 'merged' maps are available as outputs.

ermask. The task `ermask` uses the exposure maps to calculate detection masks. The masks are FITS images with the same dimensions as the exposure maps. Pixels in image areas that can be used for source detection are set to 1, all other pixels are set to 0. The area selection is based on two configurable thresholds setting either a lower limit on the exposure as a fraction of the maximum map value or an upper limit for the spatial gradient of the exposure map. In the current pipeline versions only the fractional exposure threshold is set.

erbackmap. The task `erbackmap` calculates a background map based on a photon count image, the corresponding exposure map, the detection mask, and an initial source list. In a first step, the task calculates a mask to blank out circular regions around the sources from the input list. The radii of these circles depend on the source count rate, the PSF, and on source extent parameters from the input list. A second step smooths the remaining images and interpolates the map into the masked out regions around the input sources. For this step either a two dimensional smoothing spline or an adaptive smoothing algorithm (recommended)

---

[17] *Users Guide to the XMM-Newton Science Analysis System*, Issue 16.0, 2021 (ESA: *XMM-Newton* SOC)

can be selected. The adaptive smoothing algorithm convolves the input images with Gaussian smoothing kernels of different scales and chooses the smoothed map for each pixel that reaches the user-defined signal-to-noise ratio.

ersensmap. The task `ersensmap` uses the eSASS exposure maps and background maps to estimate the detection limits in eROSITA observations. For each map pixel, the task calculates the detection limits for aperture methods as employed by the task `erbox` or for the PSF-fitting method as used by task `ermldet`. In the aperture case, the algorithm iteratively determines the source flux necessary to reach the given likelihood threshold according to equation A.1. In the PSF-fitting case, the task calculates the flux necessary to reach the likelihood threshold determined by the C statistic (equation A.2). In both modes, the task can also handle the case where several input images are used for simultaneous source detection as implemented in `erbox` and `ermldet`. For the conversion between fluxes and count rates, energy conversion factors have to be given as task parameters.

Appendix A.5: Source detection and photometry

flaregti. The task `flaregti` creates a set of GTIs that optimise the ability of sources to be detected or created from a fixed count rate threshold. The first step is to remove bright point sources using a simple detection algorithm that identifies bright pixels in a binned image, which are later excluded from the light-curve production. A light curve of the count rate per unit area is then computed from the data in a given energy band, given the input GTIs. A regular grid of points is chosen over the area of the sky that is contained within the event file. For each grid point, light curves are constructed by selecting a subset of time bins from the total point-source masked light curve for those periods when the grid point in question is within the field of view. Each of these lightcurves is analysed to either choose an optimal threshold for that spatial location in standard operation, or to apply a fixed threshold. When an optimal threshold is calculated, the task chooses a threshold that minimizes the flux at which a source of a specified size can be detected against the average sky surface brightness. When it has a set of thresholds for each grid point, the task then identifies for each time bin within the light curve the grid point that is closest to the centre of the field of view. For each time bin, the threshold of this nearest grid point and the measured rate is used to decide whether a time bin is good or bad, and the tool then constructs the GTIs. These flare-based GTIs are merged with the original input GTIs to make new combined GTIs for each TM. The task then repeats the whole processes back to the source detection stage based on these new GTIs for a given number of iterations. `flaregti` writes the output FLAREGTI GTI data back into the original event file as extensions or to a separate GTI file, and optionally writes light curve, point-source mask, and spatial threshold files.

erbox. The task `erbox` is based on a sliding-box algorithm to detect peaks in the input count images. The task can either be run in local mode, which estimates the background from a region surrounding the source box, or in map mode, which uses the background maps generated by `erbackmap` at the position of the source. The algorithm first smooths the input images with a beta-function-type kernel and then searches for peaks in the smoothed images. This peak search can be repeated several times after rebinning the input images by a factor of 2×2, which effectively

doubles the box size. The peaks are then tested for statistical significance by comparing the number of box counts n_i with the expected background counts b_i in each input image i using the (logarithmic) likelihood

$$L_i = -\ln P_\Gamma(n_i, b_i), \tag{A.1}$$

where P_Γ is the regularised incomplete Gamma function

$$P_\Gamma(a, x) = \frac{\int_0^x e^{-t} t^{a-1} dt}{\int_0^\infty e^{-t} t^{a-1} dt}.$$

In the case of multiple input images (e.g. in different energy bands), a combined likelihood is computed using Fisher's method (Fisher 1932). The probability values P_i from n independent tests of the same null hypothesis can be combined as $L' = -2 \sum_{i=1}^n \ln P_i$, which follows a χ^2 distribution with $2n$ degrees of freedom. The combined detection likelihood of a source can therefore be calculated as

$$L_{\mathrm{det}} = -\ln\left(1 - P_\Gamma\left(n, \sum_{i=1}^n L_i\right)\right).$$

The source list is filtered using a threshold on the combined likelihood. For the significant sources, the following parameters are calculated for each input image and for the combined dataset:

– raw box counts,
– PSF-corrected box counts with errors,
– PSF-corrected count rates with errors,
– source fluxes with errors,
– background counts,
– detection likelihood values,
– source exposure values (for each input image only),
– centroid positions in image and equatorial coordinates (combined values only) with errors,
– the image rebinning step in which the source is most significant,
– hardness ratios of the form $HR_{12} = (cr_2 - cr_1)/(cr_1 + cr_2)$ with errors, where cr_1 and cr_2 are the count rates in two of the input energy bands.

ermldet. The task `ermldet` applies a PSF-fitting algorithm to determine source parameters for a list of input positions. It can also apply multi-PSF fits in order to deblend neighbouring X-ray sources. The source PSF can be generated using the following methods:

– from 2D PSF images stored in calibration files for a grid of photon energies and off-axis angles (pointed observations).
– from 2D images of the PSF averaged over the detector area and stored for a grid of photon energies.
– reconstruction of a source-specific PSF by averaging the PSFs of each source photon, the event-specific PSFs are stored in calibration files as sets of shapelet coefficients for a grid of energies and detector positions.
– calculation of event specific PSFs using the shapelet coefficients stored in calibration files ("photon mode").

A maximum likelihood fitting procedure is applied to the sources in the input list, starting with the most significant source and then working on the sources in the order of descending likelihood. Neighbouring sources can be combined for simultaneous fitting, and input positions can also be split into multiple sources.

The source models are generated using PSFs that are optionally folded with an extent model. The extent model is either a Gaussian kernel or a beta model of the form

$$f(x, y) = \left(1 + \frac{(x - x_0)^2 + (y - y_0)^2}{r_c^2}\right)^{-3\beta+1/2}$$

with $\beta = 2/3$. The core radius r_c is a free fit parameter. The models are multiplied with the exposure map to account for instrumental effects, and the background is added. The model parameters are

- x, y positions in image coordinates for each source,
- source extent radius for each source,
- source count rate for each source and input image.

The source parameters are optimised by minimising the C-statistic (Cash 1979),

$$C = 2 \sum_{i=1}^{N} (e_i - n_i \ln e_i), \tag{A.2}$$

where e_i is the expected model value of pixel i, n_i is number of photon events in pixel i, and N is the total number of image pixels used for the fit. When the fit has converged at a set of best-fitting parameters, the significance of each source is tested by calculating

$$\Delta C = C_{\text{null}} - C_{\text{best}},$$

where C_{null} is the C-statistic of the null hypothesis (i.e. model with zero net counts) and C_{best} is the value for the best-fitting model. The fitting algorithm of `ermldet` can optionally obtain an individual PSF for each photon event depending on its energy and detector position; where applicable, the event PSF is convolved with the extent model. The C-statistic is then calculated according to equation A.2 , where e_i is now the event-specific model value in the image pixel in which the event was detected, and C is summed over the N events with index i in the source region using $n_i = 1$. According to Cash (1979), ΔC under certain conditions approximately follows a χ^2 distribution with the number of degrees of freedom ν equal to the number of free model parameters. Hence the probability P that ΔC results from a chance fluctuation of the background can be calculated using the regularised incomplete Gamma function P_Γ,

$$P = 1 - P_\Gamma\left(\frac{\nu}{2}, \frac{\Delta C}{2}\right).$$

Based on the value of P, a logarithmic detection likelihood

$$L = -\ln(P)$$

is obtained for each source. In case of simultaneous multi-source fits, sources falling below the user-defined likelihood threshold are removed and the fitting procedure is iteratively repeated. The conditions for ΔC to follow a χ^2 distribution are not completely

fulfilled in the low count rate regime, and thus P does not correspond to the actual false-alarm probability. We used extensive simulations (Liu et al. 2021c) to determine the false-detection probabilities at various levels of L. An extent likelihood quantifying the probability that an extended model is required to fit the count distribution of a source is calculated in a similar fashion. In this case, the value C_{null} for the null hypothesis is calculated for the best-fitting point source model. Again, the relation between the extent likelihood L_{ext} and the incidence of false classifications as extended sources needs to be calibrated by simulations. If the extended model fit results in an extent likelihood below the user-defined threshold, the fit is repeated with a PSF model and the fit parameters of the point source model are written to the source list. After each source fit, the best-fitting model is added to the internal background map used for the model fits of subsequent (fainter) sources. The final source model + background maps can be written as FITS images.

Assuming ΔC follows a χ^2 distribution, we can calculate the parameter errors corresponding to 68% confidence intervals by varying the parameter of interest from its best value in both directions until $C = C_{\text{best}} + 1.0$ is reached. These boundaries are determined by an iterative procedure, and the errors for both directions are averaged and written to the output table. When for one direction the algorithm fails to converge after the given number of iterations, the errors determined for the other direction are adopted. When no error margin can be determined in either direction, a zero is written to the output table. The positional errors are given in units of image pixels for both X and Y positions and for an error radius in units of arcseconds: $\Delta(\text{RA}, \text{Dec}) = \text{pixelsize} \times \sqrt{\Delta X^2 + \Delta Y^2}$. For derived quantities such as the sum of count rates in multi-band fits or the hardness ratios of energy bands, the count rate errors are propagated by treating them like Gaussian errors.

For each source passing the likelihood thresholds, the following source parameters are written to the output source list:

- best-fit position with errors in image, equatorial, and Galactic coordinates;
- best-fit source extent value with errors;
- best-fit counts, count rates, and fluxes with errors;
- likelihood values for detection and extent;
- exposure and background map values at the source position;
- hardness ratios of the form $HR_{12} = (cr_2 - cr_1)/(cr_1 + cr_2)$ with errors, where cr_1 and cr_2 are the count rates in two of the input energy bands.
- the radius of the sub-region used for fitting, the PSF fraction covered by that region, and the fraction of pixels with valid detection mask values.

The source list contains several rows per source: one per input image, plus summary rows over the energy bands and, for the optional case where the images of different telescope modules are supplied separately, a summary over different instruments.

apetool. The task `apetool` has three main functionalities: *i)* it produces maps of the size (radius in pixels) of the PSF across the eROSITA field of view (PSF maps), *ii)* it performs aperture photometry at a set of user-defined positions, and *iii)* it generates sensitivity maps following the method described by Georgakakis et al. (2008).

The PSF of eROSITA is modelled using the shapelet basis functions (Refregier 2003; Refregier & Bacon 2003). The shapelet coefficients that describe the distribution of photons at a certain energy and at a given position in the eROSITA field of

view are stored in a calibration file. To construct the PSF map, the `apetool` first extracts the GTI from the attitude file. For each time interval, the shapelet coefficients at the positions of a given pixel are retrieved from the calibration file. The average of each shapelet coefficient across all time intervals is used to reconstruct a model of the PSF at the position of interest. This is then used to measure the size of the PSF in pixels at encircled energy fractions (EEF) between 40-95% in steps of 5%. The `apetool` estimates the PSF size in a regular grid of positions in the eROSITA field of view. The density of the grid is a trade-off between speed, size of the final PSF map, and an adequate description of the variations in PSF size across the field of view. The default setup is a grid of 21×21 positions along the X and Y direction.

In the case of aperture photometry, there are two options. Counts are either extracted at source positions defined in a source catalogue generated by the `ermldet` task or at arbitrary user-defined positions. The difference between the two options is the format of the input and output source lists. In the first case, the `ermldet` source catalogue is supplemented with aperture photometry products. These include *i)* the total counts at a given position (source and background) extracted within an aperture of size defined in units of EEF, *ii)* the background counts extracted from the source maps generated by the `ermldet`, *iii)* the mean exposure time, *iv)* the EEF used to define the extraction radius, *v)* the size of the extraction radius in pixels, and *vi)* the Poisson probability that the extracted counts (source + background) are a fluctuation of the background. These quantities can be combined to estimate the source flux. The case of aperture photometry at arbitrary source positions corresponds to science applications related to, for instance, X-ray stacking analysis or searches for faint X-ray emission (i.e. below the formal X-ray detection threshold) associated with counterparts selected at other wavelengths or external (non-eROSITA) X-ray catalogues. The apertures within which counts are extracted at user-defined positions are expressed in terms of EEF. The quantities as listed above are estimated and stored in an output file.

The sensitivity map generated by `apetool` has units of counts. They represent the minimum number of photons (source and background) within the extraction aperture (expressed in EEF units) at a given position on the detector, so that the Poisson false-detection probability is lower than a user-defined value. The Poisson false-detection probability is defined as the probability that the background fluctuats in a random fashion to produce counts within an aperture above a given value. This probability is given by the survival function of the Poisson probability distribution. The resulting sensitivity maps can be combined with the aperture photometry determined by `apetool` for the `ermldet` sources to provide an accurate representation of the point-source selection function (probability of detecting a source) at a given Poisson false-detection threshold (see Georgakakis et al. 2008). The combination of the `apetool` sensitivity maps with the aperture photometry of the `ermldet`-detected sources (also via `apetool`) enables, among others, the accurate estimation of the X-ray point-source number count distribution as a function of X-ray flux, that is, $\log N - \log S$ (see Section 4.2).

The sensitivity map FITS file includes three image extensions: *i)* the sensitivity map itself, that is, the minimum number of counts as described above as a function of detector position, *ii)* the expected background within the extraction radius, and *iii)* the corresponding average exposure time (mean of the exposure map) within the extraction radius. With these three components, it is possible to estimate the detection probability of a source with a given count rate or flux. The resulting area or sensitivity

curve is stored in the fourth (table) extension of the sensitivity map fits files. It includes count rates and the corresponding area in square degrees within which a source of this count rate can be detected. This calculation makes no assumptions on the X-ray spectral shape of the source.

catprep. The task `catprep` re-formats the source lists written by the tasks `ermldet` and `apetool`. The output format contains only one row per source; the energy-band specific values are written as separate columns. The task also assigns a source identification string containing the root name of the input file, the source number from the input list, and the processing version.

Appendix B: eROSITA calibration database

The `eSASS` calibration database (CALDB) contains all information required for calibrating the data collected with the eROSITA telescope modules (TMs). It is based on the HEASARC[18] CALDB framework for managing calibration files. Each of the seven telescope modules of eROSITA has its own set of calibration files. The names of the calibration files follow a naming convention that specifies the telescope module, calibration type, and, if necessary, the type of detector or filter for which this calibration file is intended, as well as the date from which this information is valid and the version number of the file. For example, `tm2_badpix_190712v01.fits` contains the information on the bad pixels of TM2 valid from 12 July 2019 (i.e. before launch) in the first version.

Appendix B.1: Telescope and point spread function

Vignetting. Vignetting is defined here as the flux of a point source contained in a circle of $4'$ radius relative to that flux at the on–axis position. The vignetting database contains parameters for computing this value for any given photon energy and off–axis angle.

These parameters were derived from the same data as were taken to determine the point spread function (see below). Variations in the flux of the X–ray source were considered by performing simultaneous measurements with a monitor counter. This method did not only save measurement time, but also had the advantage that the vignetting centre (i.e. the point of minimum vignetting, which is not necessarily coincident with the point of minimum HEW and which may move with energy) could be determined from the narrow grids of PSF measurements, and that the assumption of azimuthally symmetric vignetting could be verified. Comparisons between the measured values and the modelled vignetting curves are shown in Fig. 7 of Dennerl et al. (2020).

Point spread function: images. In the ground calibration, images of the PSF were obtained at the MPE PANTER facility[19] individually for each mirror assembly (MA) at the energies of the C–K, Cu–L, Al–K, Ag–L, Ti–K, Fe–K, and Cu–K emission lines, covering the energy range from 0.3 to 8.0 keV. The MAs were placed 124 m away from an X–ray point source and mounted together with the X–ray camera TRoPIC on a rigid platform, which could be tilted horizontally and vertically. The TRoPIC X–ray CCD is identical in its pixel size to the eROSITA

---

[18] https://heasarc.gsfc.nasa.gov/
[19] https://www.mpe.mpg.de/heg/panter

CCDs, but consists of only 256 × 256 pixels. In order to sample the full eROSITA focal plane, TRoPIC was mounted on a manipulator, which shifted it to five different positions to cover the central and the four outer parts of the field of view. The four outer TRoPIC positions sampled the PSF on a rectangular 6′ × 6′ grid in 121 images, covering the offset angles from −30′ to +30′. The central TRoPIC position measured the PSF in 36 images on a 6′ × 6′ grid that was shifted by 3′ with respect to the outer grid and covered the off-axis angles between −15′ and +15′. This resulted in 157 images per MA and energy. In order to save measurement time, Cu–L exposures were only done for TM1 and TM2, and the outer part of the field of view was only sampled in the upper right quadrant at Ag–L, Ti–K, Fe–K, and Cu–K, except for TM1, where the full grid was covered in all the seven energies.

In order to obtain the required large number of PSFs within a reasonable time, the measurements were taken at very high photon rates, taking extreme pile–up into account and making use of the fact that, if the dominant incident energy is known, the number of photons per pixel can be reconstructed from the total charged released there. In this mode, however, it is not possible to apply the usual technique to identify and suppress the traces of minimum ionising particles by selecting only valid pixel patterns, because photon pile–up may also create 'invalid' pixel patterns. This problem could be solved by developing specific criteria for spotting suspicious pixel patterns, by removing all the contaminated CCD frames, and by visually checking the remaining frames for each of the 4860 PSF exposures. It turned out that only ∼ 5% of the frames needed to be rejected. In this way, the exposure time per PSF could be reduced from several hours to ∼ 80 seconds, boosting the efficiency by two orders of magnitude.

The further processing of the PSFs required transforming their location on TRoPIC to that on the eROSITA focal plane. This made it necessary to calibrate the geometrical properties of the PANTER setup, which consists of several distances and angles. In total, seven geometrical quantities needed to be determined.

Another challenge was that the PSFs were available only at the native 9″.5 resolution provided by the TRoPIC pixel size (sub-pixel resolution could not be used here because of the high pile–up). In order to increase the spatial resolution, we tried to find plausible flux distributions that agreed with the measured ones after binning them into 9″.5 × 9″.5 pixels. This was done by developing a modified method of bicubic spline interpolation, which was made to be flux conservative (Fig. B.1). These images are the basis for modelling the PSF. More details are presented in Dennerl et al. (2020).

Point spread function: shapelet model. The eROSITA optics and large field of view translate into a complex PSF shape that depends on both off-axis angle and energy. The regular nearly Gaussian PSF on-axis is distorted with increasing off-axis angle, leading to elongated features and asymmetries (see Figure B.2) that are hard to model using analytic functions. This has motivated an approach whereby the PSF at a given position is linearly decomposed into an appropriately chosen two-dimensional set of localised basis functions. The PSF can then be represented by the coefficients of the basis function components. The decomposition uses the shapelet functions proposed by Refregier (2003); Refregier & Bacon (2003), which represent a complete and orthonormal set in the two-dimensional space. The shapelets are weighted Hermite polynomials and correspond to perturbations
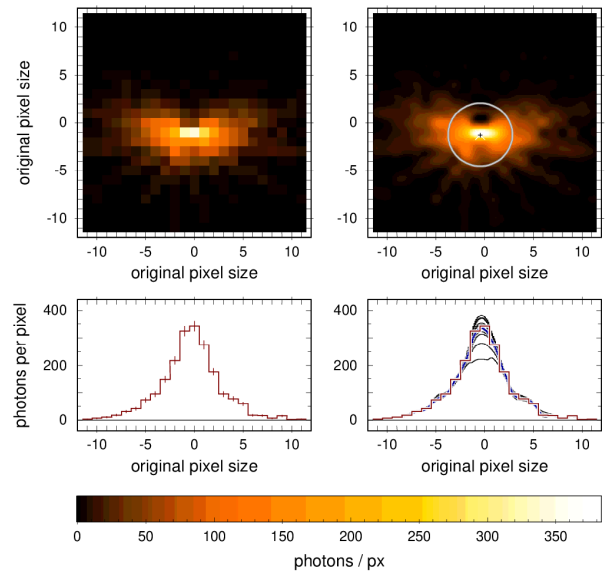
Fig. B.1. Example of the 1.5 keV TM1 PSF at an off–axis angle of 30′ in the original resolution (left) and the reconstructed resolution (right), where each original pixel is sub-divided into 10×10 sub-pixels. The histograms at the bottom show the flux distribution in the brightest pixel row (left) and in the ten corresponding sub-pixel rows (right). The modified bicubic resampling method enhances fine structures. The white circle indicates the minimum HEW.

of a circular Gaussian profile. In one dimension, these are described by the relation

$$\phi_\nu(x;\beta) = \left[2^n \beta \pi^{\frac{1}{2}} n!\right]^{-\frac{1}{2}} H_n(x/\beta)\, e^{-\frac{x^2}{2\beta^2}}, \tag{B.1}$$

where n is a non-negative integer, $H_n(x)$ is the Hermite polynomial of order n, and β is the scale of the shapelet function. In the case of two-dimensional images with coordinates (x, y), the shapelet basis functions are given by the relation

$$\psi_{\nu,\mu}(x, y; \beta) = \phi_\nu(x;\beta) \cdot \phi_\mu(y;\beta). \tag{B.2}$$

In this case, the order of the shapelet function is characterised by the two non-negative integers n, m, and there is a single scale β for both $\phi_\nu(x)$, $\phi_\mu(y)$. The choice of the scale β depends on the size of the features to be modelled. The motivation for using this set of functions is that they have been used to reconstruct the complex shapes of galaxies in optical surveys for weak-lensing applications (Massey et al. 2007).

In the case of the eROSITA PSF, we used three independent sets of shapelet functions, $\psi_{\nu,\mu}(x, y; \beta)$, each of which corresponds to a different scale. They are intended to model the core, the main body, and the wings of the PSF. The scales of each of the three components were fixed to $\beta = 1, 1.5$, and 6 pixels of the ground-calibration images, which have a pixel scale of 9″.5. The number of shapelet orders is set by the requirement $n + m \leq N_{\max}$, where $N_{\max} = 0, 10$, and 8 for each of the three scales $\beta = 1, 1.5$, and 6 pixels, respectively. The choice of $N_{\max} = 0$ for the smallest scale translates into a single Gaussian function to model the core of the PSF. There are 66 and 45 shapelet coefficients for $N_{\max} = 10$ (scale $\beta = 1.5$ pixels) and $N_{\max} = 8$ (scale $\beta = 6$ pixels). The total number of coefficients (free parameters) that describe the PSF shapelet model at a given position on the detector and energy is 112. The choice of the β,
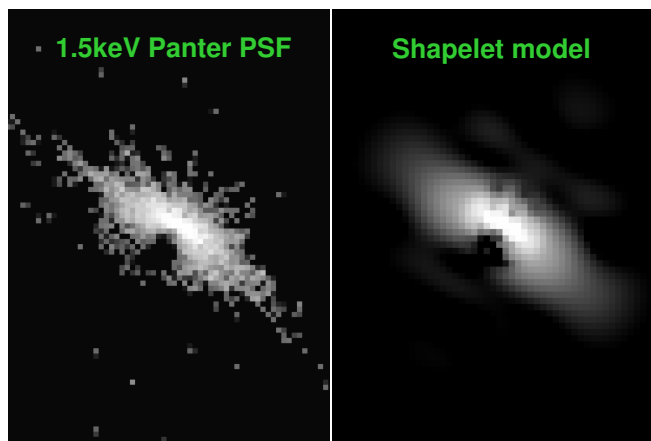
Fig. B.2. Example of the 1.5 keV TM1 PSF at an off-axis angle of about 26″. The image on the left shows the PANTER calibration image. On the right, we show the PSF reconstruction using the shaplet model described in the text.

N_{max} values is the result of experimentation and represents the minimum number of scales and coefficients that provide a reasonable representation of the PSF at all off-axis angles and energies (see Figure B.2). Fixing the scale and N_{max} of the shapelets enables the co-addition (stacking) of PSFs across energies and off-axis angles in the shapelet-coefficient space rather than the image-pixel space and significantly accelerates (factors of 10-100) the calculation of the model PSFs.

The PANTER PSF images for different energies and off-axis angles described earlier in this section with a pixel size of 9″5 were fit with the shapelet model above to determine the maximum likelihood coefficients (112 for each PSF image). The modelled energies were 0.3, 1.5, 2.0, 4.5, 5.4, and 6.4 keV. Figure B.2 shows an example PSF image taken at PANTER for the TM1 module for photons with energy 1.5 keV. For comparison, the shapelet-model reconstruction of the PSF is also presented in this figure. The fraction of photons predicted by the shapelet model within a fixed radius is typically within 10-20% of the PANTER measurement. This difference is related to the maximum adopted shapelet scale of 6 pixels ($\approx$ 1 arcmin). Photons beyond this scale cannot be reproduce by the current model. The shaplet coefficients were packed in calibration files that were accessed by the various `eSASS` tasks. For a given input position on the eROSITA detectors and a given energy, the shapelet coefficients were estimated using linear interpolation in three dimensions (two space dimensions and one energy dimension).

Appendix B.2: Pattern recombination and energy calibration

While no calibration data are needed for the pattern recombination, this is very different for the energy calibration because each CCD column is characterised by an individual CTI and gain, and these values depend on the energy, the CCD temperature, and the time of the observation. The initial database, which did not yet consider the last two dependences, contains already 21616 parameters. In the updated currently used database, which considers the effects of the CCD temperature and the time of the observation, the number of parameters has increased to 29764. Another database, consisting of a total of 1799 sub-pixel maps, is used for a fast reconstruction of the sub-pixel position from the distribution of the charge over the pattern (Dennerl et al. 2012).

Appendix B.3: Detector and camera characterisation

Bad pixels. The bad-pixel calibration files characterise individual bad pixels and groups of bad pixels (specified as rectangles) of each camera by type, affected amplitude range, and time at which they were active. The list of bad pixel is updated continuously as the behaviour of the detectors evolves.

Bad times. These calibration files contain the start and end times of periods when data are either unavailable or cannot be used scientifically because of a malfunction. In addition, time periods are listed in which a TM is offline due to a camera reset (Predehl et al. 2021) or is switched of, for instance, because of an orbit correction. In the calibration and performance verification (Cal-PV) phase, these time periods were set for TM3 and TM6.

Time shifts. Both the Interface and Thermal Controller (ITC) and the Camera Electronics (CE) store a local copy of the onboard time source (OTS) counter (Predehl et al. 2021). These local copies are incremented by a 1 Hz pulse signal from the spacecraft. ITC and CEs can only be updated with the OTS time word from the spacecraft when they are in a specific mode. For the time being, this is only the case when a CE is reset or switched on. On these occasions, it can happen that the time counter of the CE jumps by 1 s.

Time shifts can normally only be detected in survey or fieldscan mode (if the scan velocity is sufficiently high) because data from CEs with incorrect time synchronisation are then projected at incorrect sky locations. With pointed observations, such as those made predominantly in the Cal-PV phase, each event is projected onto the same position in the sky, and time shifts can only be detected by observing a stable clock such as a pulsar. It is known from pulsar observations during the Cal-PV phase that a few time shifts occurred. The only time shift during the Cal-PV phase that could be clearly assigned is from November 24, 2019, for TM2. The `timeoff` calibration files contain the date and the direction of each time shift (0 → 1 or 1 → 0) for each camera from this date on. Time shifts applicable for each camera before this date were determined directly from the eFEDS data and from two all-sky survey test scans performed on September 23 and November 15, 2019. The TM6 time shift in eFEDS subfield I reported in § 2 of the main paper is not corrected.

Detector maps. The detector maps are intended to hold fractional sensitivity values of each detector pixel. While the detector maps are evaluated by eSASS tasks `expmap` and `srctool`, it is currently not foreseen to make use of this calibration mechanism, and all pixel values are set to 1.0.

FOV maps. The field of view maps define the field of view of each camera. Pixels outside the mask will receive the `OUT_OF_FOV` flag (set by task `evprep`), they will not be projected onto the sky, and the corresponding areas will be excluded from the exposure maps.

Appendix B.4: Boresight

Star trackers and gyroscopes. For each of the four attitude sources of the SRG satellite (SED26/1, SED26/2, BOKZ, and Q-Gyro) a calibration file is maintained. Each file holds an orientation quaternion defining the transformation between the co-

ordinate system of the respective device and the nominal camera orientation of telescope module TM1.

X-ray telescopes. The geometrical parameters for each telescope module are stored in the respective instrument calibration files. Each file contains the plate scale and the camera boresight with respect to the nominal orientation of TM1, the boresight is stored in the form of Euler angles (pointing offsets in X, Y, and roll angle).

Appendix C: eROSITA standard calibrated data products

The eROSITA data analysis pipeline provides a set of calibrated data products described below. All products are FITS files complying (where feasible) with established standards such that in addition to eSASS, a range of general-purpose astronomical data analysis tools may be used. A full description of the main product files is available online[8].

Appendix C.1: Calibrated event files

eROSITA calibrated event files are containers that provide a full set of X-ray event and auxiliary data required for most data analysis tasks in multiple FITS extensions. They consist of one `EVENTS` FITS table extension holding data from one or several eROSITA cameras as well as various camera specific extensions.

X-ray events. The `EVENTS` FITS table extension holds information on each observed event and for recombined, calibrated X-ray photons. Event coordinates are provided in units of detector and sky pixel coordinates and as right ascension and declination. Further information includes event arrival times, raw and calibrated event amplitudes, a variety of event pattern type information (see Appendix A.1, `pattern`), and a set of event flags characterising each event.

Good time intervals. Good-time intervals (GTI) are provided for each camera in the `GTIn` FITS table extensions, where n is the TM number. The GTIs indicate the time periods when each camera was operational and collecting event data. Optionally, a set of `FLAREGTIn` FITS table extensions is included. They specify low-background time periods free of background flares (see Appendix A.5, `flaregti`).

Live time. The `DEADCORn` FITS table extensions provide the fractional live time for each 50 ms time interval of each camera. The `DEADORn` extension tracks the fraction of the detector area that is not available for the detection X-ray events due to minimum ionising particles (MIPs) hitting the detector. Where applicable, the `DEADCORRn` extension also records exposure losses when a fraction of the exposure time is intentionally discarded, for instance to prevent count rates exceeding technical limitations.

Telescope specific attitude. Right ascension, declination, and roll angles with respect to the camera (`RAWX`, `RAWY`) coordinate system are provided for each camera in the `CORRATTn` FITS table extensions with a time resolution of one second.

Bad pixel information. The `BADPIXn` FITS table extensions list and characterise bad pixels of each camera. The information provided includes the affected energies and time intervals as well as the bad-pixel type (bright, masked on-board, etc.). The format of the bad-pixel event file extensions is closely related to the bad-pixel calibration files described in Appendix B.3.

Housekeeping information. For each camera, four FITS table extensions are provided that hold a subset of the eROSITA housekeeping data required for event calibration and exposure calculation, such as the observing mode and filter wheel position of each camera.

Appendix C.2: Count images and maps

Images. The `eSASS` science images are created by the task `evtool` and are stored in the primary extension of a FITS file. The FITS header contains the standard `eSASS` keywords copied from the original event file, keywords describing the event selection, and a World Coordinate System (WCS) header. By default, the FITS file also contains the EVENTS and auxiliary extensions from the original events file with the same selections applied as were used to create the image.

Exposure maps. The exposure maps contain the results of the exposure calculation performed by the task `expmap`. The exposure maps can either contain the raw exposure times for each map pixel (unvignetted exposure) or the exposure multiplied with the energy and off-axis angle-dependent telescope vignetting function (vignetted exposure). In both cases, the unit of the pixel values is seconds. The WCS header is identical with that of the input template image.

Background maps. The background maps contain the source free sky + detector background as determined by the task `erbackmap`. The values are stored in units of counts/pixel, and the image size and WCS coordinate system of a background map is identical with that of the respective science image.

Sensitivity maps. The sensitivity maps contain limiting fluxes for sources that can be detected in an eROSITA observation. The nominal unit of the pixel values is erg/(s cm^2), but it depends on the energy conversion factor provided for the task.

Appendix C.3: Source level products

Source catalogues. Source catalogues are written as FITS tables by the tasks `ermldet` and `apetool` and are then reformatted by the task `catprep`. The final format contains one row per source, with a column for each of the source parameters described in section A.5.

Spectra. Spectra are produced by the task `srctool` and are standard OGIP-compliant[20] data products. The task writes output spectra for each TM and a combined spectrum. Each spectrum consists of the number of counts in an extraction region within each PI channel after filtering has been applied. There are some subtleties in how `srctool` defines various output quantities due to the scanning nature of the eROSITA telescope. The

---

[20] OGIP FITS Working Group, https://heasarc.gsfc.nasa.gov/

ONTIME is the total amount of time for which the source is within the field of view during the input GTIs. The time intervals within the ONTIME (source specific GTIs) are normally written to the srctool output products as GTI extensions. The EXPOSURE is the ONTIME reduced to account for dead time fraction given in the input event file. The BACKSCAL is the average area on the sky of the source or background extraction region during the source-specific GTIs. For background subtraction, it is often useful to know the geometric area on the sky of the extraction region, which is written as a srctool-specific keyword called REGAREA. The geometric area of where the source model is non-zero is written as RGDMAREA. In addition to the standard COUNTS column in the output spectra, srctool also writes columns containing the number of single, double, triple or quad patterns in each channel. The total number of counts of each type is written into keywords as CNTS_S, CNTS_D, CNTS_T, and CNTS_Q, and the total as CTS. When combining spectra, srctool adds the counts and computes average ONTIME and EXPOSURE values, while BACKSCAL, REGAREA, and RGDMAREA are exposure-weighted averages.

Light curves. The output light-curve files contain the binned X-ray light curve in several bands, following OGIP recommendations for time-analysis data files. The columns include the midpoint of the time bin (TIME), the width of the bin (TIMEDEL; δ_T), the number of counts in each band (COUNTS; c), the number of counts in the background region in each energy band (BACK_COUNTS; c_B), the fraction of the nominal on-axis effective collecting area computed in each band (FRACAREA; f_A), the fraction of the time bin that overlaps with input GTIs and when the source was visible in the field of view (FRACTIME; f_T), the product of the fractional collecting area and fractional temporal coverage (FRACEXP; $f_E = f_A f_T$), the mean off-axis angle during the bin, and the ratio by which the background counts need to be scaled to give the background contribution in the source aperture (BACKRATIO, r). The RATE column is calculated from the other columns in each band as $(c - r c_B)/(f_E \delta_T)$. The 1σ uncertainty on the rate is given for each energy band in RATE_ERR, providing there are more than 25 counts in the respective counts column, using $\sqrt{c + r c_B}/(f_E \delta_T)$.

Appendix C.4: Detector response and effective areas

Response matrices. Response matrices (RMFs) generated by srctool are OGIP-format response files, containing standard MATRIX and EBOUNDS extensions. The output depends on the pattern selection chosen by the user. Output response matrices are the sum of the response matrices for the different patterns. Similarly to spectra, srctool writes response matrices for the individual TMs and for the combination. The combined response matrix is the exposure-weighted average of the individual response matrices.

The RMFs were derived from measurements that had been performed with a prototype version of the eROSITA CCDs at the synchrotron radiation facility BESSY II, where the energy range $0.08 – 11.0$ keV was sampled with monochromatic measurements at 29 energies (Ebermayer et al. 2010). The measurements were cleaned from artefacts (e.g. higher orders), filtered to contain only single-pixel events, and then used to derive an empirical mathematical function that reproduced the observed spectra. After this function was available, it was straightforward to evaluate it at any energy and to compose an empirical RMF. This method was also applied to the other three valid pattern types (doubles, triples, and quadruples) by using the same generic function and modifying only some parameters. This resulted in four initial RMFs, one for each pattern size (Dennerl et al. 2020). These are the RMFs that are currently used for all TMs. It is planned to adapt these generic RMFs to the properties of the individual TMs by using dedicated calibration observations of suitable astrophysical targets. In this way, it should also be possible to take temporal variations into account.

Auxiliary response files. The auxiliary response files (ARFs) made by srctool are OGIP-compliant. These files include several different effects that are applied to the on-axis effective area, including vignetting, PSF correction, and corrections for times at which the source is only partially visible. The ARF files contain a SPECRESP extension that in turn contains the standard columns of ENERGY_LO, ENERGY_HI , and SPECRESP , which give the effective area curve. In addition, srctool writes columns for diagnostic purposes that give the corrections applied to the on-axis effective area due to PSF (as CORRPSF), vignetting (as CORRVIGN), and the two combined (CORRCOMB).

Appendix D: Description of catalogue entries

Table D.1. Energy bands

Band	energy range keV	ECF cm^2/erg
single-band detection		
	0.2–2.3	1.074×10^{12}
three-band detection		
1	0.2–0.6	1.028×10^{12}
2	0.6–2.3	1.087×10^{12}
3	2.3–5	1.147×10^{11}
post-hoc photometry		
s	0.5–2	1.185×10^{12}
h	2.3–5	1.147×10^{11}
u	5–8	2.776×10^{10}
b1	0.2–0.5	9.217×10^{11}
b2	0.5–1	1.359×10^{12}
b3	1–2	1.014×10^{12}
b4	2–4.5	1.742×10^{11}

Notes. ECF: Energy conversion factors for the four source-detection bands and the seven forced-photometry bands.

The catalogues are available with this paper and on the eROSITA Early Data Release website[21]. Table D.1 lists the energy bands and corresponding ECFs used in this work. The ECFs are calculated assuming an absorbed power law with a slope of 2.0 and with a Galactic absorbing column density of 3×10^{20} cm^{-2}.

The main and the supplementary catalogues from the 0.2–2.3 keV single-band detection share the same columns as described in Table D.2. The seven-band forced photometry is done independently for the single-band and three-band detected sources. In the case of the three-band detected hard catalogue, four sets of output columns from the source detection PSF-fitting (not the forced PSF-fitting) have the same names as listed in section 2 of Table D.2 , but with an energy-band suffix in the column name indicating the three source-detection bands (1, 2, and 3) or the summary of the three bands (0). The only exception is the

---

[21] https://erosita.mpe.mpg.de/edr/eROSITAObservations/Catalogues/

ML_EXP_*Band*, which is only for the three individual bands (1, 2, and 3), but not for the summary case (0). The hard catalogue has an additional column ID_main, which stores the source ID in the main catalogue for the overlapping sources.

The current versions of the catalogues contain a small number of NULL values in the PSF-fitting error measurement columns, for example, ML_RATE_ERR, RADEC_ERR_ERR, and EXT_ERR, because in some cases the `ermldet` algorithm that determines the parameter errors failed to converge. When the determination of the positional error failed (490 sources in the main catalogue; 1.8% of the total), a mean RADEC_ERR_CORR of 4″.0 was assumed, which is the median value of RADEC_ERR.

Appendix E: eFEDS simulations

In order to inspect and optimise the source detection for eROSITA, we ran extensive simulations as described in Liu et al. (2021c). The simulation of eFEDS is briefly summarised here.

We used `sixte-2.6.2`[22] to create mock eFEDS event files and implemented 18 realisations. To make them as representative as possible to the real data, we input the real eROSITA calibration files and eFEDS attitude file, and AGN, star, and cluster catalogues generated from cosmological simulations (Comparat et al. 2020). We measured the mean X-ray and particle background spectra from the eFEDS data and simulated the source signal, X-ray background, and particle background separately to account for their different vignetting. We ran the eFEDS source detection pipeline on the mock data and then associated the detected sources with the input on the basis of the source-ID flag on each photon. By checking the number of photons contributed by each input source in the core region of a detected source, we identified its primary and secondary (if existed) input counterparts. The point sources (AGN and stars) and extended sources (clusters) were considered separately. In this way, we were able to identify not only the cases where one input source was uniquely matched to one detected source that was correctly classified (as point- or extended source) or not, but also the cases where one input source was detected as multiple sources and the cases where one detected source was due to a blending of multiple input sources.

Based on the input-output association, we were able to measure the source detection efficiency, that is, the completeness (detected fraction) and contamination level (fraction of spurious sources). The final source detection strategy was chosen according to the impact on source detection efficiency caused by any adjustment of the source detection procedure and parameters.

---

[22] https://www.sternwarte.uni-erlangen.de/research/sixte/

Table D.2. Catalogue column description

Column	Format	Units	Description
1. Source properties from PSF-fitting detection			
Name	22A	..	
ID_SRC	J	..	Source ID
ID_main (ID_hard)	J	..	Counterpart source ID in the main (hard) catalogue for sources in the hard (main) catalogue
RA	D	deg	Right ascension (ICRS), uncorrected
DEC	D	deg	Declination (ICRS), uncorrected
RADEC_ERR	E	arcsec	Combined positional error, uncorrected
RA_CORR	D	deg	Right ascension (ICRS), corrected
DEC_CORR	D	deg	Declination (ICRS), corrected
RADEC_ERR_CORR	D	arcsec	Combined positional error, corrected
EXT	E	arcsec	Source extent parameter
EXT_ERR	E	arcsec	Extent error
EXT_LIKE	E	..	Extent likelihood
DET_LIKE$_*$	E	..	Detection likelihood
ML_RATE$_*$	E	cts/s	Source count rate measured by PSF-fitting
ML_RATE_ERR$_*$	E	cts/s	1-σ count rate error
ML_CTS$_*$	E	cts	Source net counts measured from count rate
ML_CTS_ERR$_*$	E	cts	1-σ source counts error
ML_FLUX$_*$	E	erg/cm^2/s	Source flux in the detection band
ML_FLUX_ERR$_*$	E	erg/cm^2/s	1-σ source flux error
ML_EXP$_*$	E	s	Vignetted exposure time at the source position
ML_BKG$_*$	E	cts/arcmin^2	Background at the source position
inArea90	L		True if inside the inner region with 0.2-2.3 keV vignetted exposure above 500s, which comprises 90% of the total area
2. Forced PSF-fitting results for seven energy *Band*s (Table. D.1.1); 7×15 columns			
DET_LIKE_*Band*	D	..	Detection likelihood
ML_RATE_*Band*	D	cts/s	Source count rate
ML_RATE_ERR_*Band*	D	cts/s	1-σ combined count rate error
ML_RATE_LOWERR_*Band*	D	cts/s	1-σ lower count rate error
ML_RATE_UPERR_*Band*	D	cts/s	1-σ upper count rate error
ML_CTS_*Band*	D	cts	Source net counts
ML_CTS_ERR_*Band*	D	cts	1-σ combined counts error
ML_CTS_LOWERR_*Band*	D	cts	1-σ lower counts error
ML_CTS_UPERR_*Band*	D	cts	1-σ upper counts error
ML_FLUX_*Band*	D	erg/cm^2/s	Source flux
ML_FLUX_ERR_*Band*	D	erg/cm^2/s	1-σ combined flux error
ML_FLUX_LOWERR_*Band*	D	erg/cm^2/s	1-σ lower flux error
ML_FLUX_UPERR_*Band*	D	erg/cm^2/s	1-σ upper flux error
ML_EXP_*Band*	D	s	Vignetted exposure time at the source position
ML_BKG_*Band*	D	cts/arcmin^2	Background at the source position
3. Aperture photometry results for seven energy *Band*s (Table. D.1.1); 7×5 columns			
APE_CTS_*Band*	J	cts	Total counts extracted within the aperture
APE_EXP_*Band*	D	s	Exposure map value at the given position
APE_BKG_*Band*	D	cts	Background counts extracted within the aperture, excluding nearby sources using the source map
APE_RADIUS_*Band*	D	pixels	Extraction radius
APE_POIS_*Band*	D	..	Poisson probability that the extracted counts (APE_CTS) are a background fluctuation

Notes. The suffix in the column name of forced PSF fitting and aperture photometry results (section 2 and 3) indicates the energy band (Table. D.1). The columns of the three-band detected hard catalogue are almost identical to those of the single-band detected main and supplementary catalogues, except that the PSF-fitting output columns (those marked with *a* subscript asterisk) have four sets of data with an energy-band suffix (0, 1, 2, and3) in the column name rather than only one set of data without this energy-band suffix (as listed in section 1). Moreover, the ID_main and ID_hard columns are only for the hard and main catalogues, respectively.