Optimal eROSITA bandpasses for the detection of galaxy clusters

P. D. Spencer-Smith[1][1]

[1]*H. H. Wills Physics Laboratory, University of Bristol, Tyndall Ave, Bristol BS8 1TL, UK.*

10 May 2021

ABSTRACT
This study investigates the use of "optimal", signal-to-noise energy bands for the detection of galaxy clusters. The optimal energy bands for the X-ray emission from galaxy clusters were calculated for the eROSITA observatory. Using a typical eROSITA background spectrum, it was found that the conventionally used $0.5 - 2$ keV band is close to optimal for clusters with plasma temperatures of $kT > 2$ keV and redshifts of $z < 1$. For galaxy clusters with a plasma temperature of $kT < 2$ keV it is shown that, relative to an optimal band, the conventionally used band suffers $10\% - 55\%$ signal-to-noise loss. For these lower temperature systems, the conventionally used band suffers a further deterioration in signal-to-noise for clusters with a redshift of $z > 0$. The performance of the optimal energy bands was tested using a simulation of the eROSITA all-sky-survey. For galaxy clusters with plasma temperatures of $(0.5 < kT < 10)$ keV and redshifts of $z < 2$, the number of clusters detected in an optimal band was found to be either highly comparable to or less than the number detected in the conventional band. The number of false positives detected in an optimal band was found to be either highly comparable to or grater than the number detected in the conventional band. A flux limit in the conventional $0.5 - 2$ keV band for the eROSITA observatory of $(2.1 \pm 0.5) \times 10^{-14}$ erg/s/cm^2 was also recorded.

1 INTRODUCTION

The study of galaxy clusters plays an essential role in developing our understanding of the cosmos. The extensive use of clusters as cosmological probes has provided some of the most accurate constraints on cosmological parameters ever recorded (Voit 2005; Allen et al. 2011; Weinberg et al. 2013). By continuing to develop the techniques which allow us to study these objects we can likewise maximise the significance of the cosmological results they provide.

Clusters of galaxies are the largest known gravitationally bound objects in the universe. Their discovery (Abell 1958; Felten et al. 1966; Abell et al. 1989) showed them to consist of populations of galaxies numbering from hundreds to thousands. The formation of a cluster occurs due to the gravitational collapse of cosmic material which traces peaks in the primordial density distribution of the early universe. The amplitude of the perturbations in this density distribution is quantified by the so called σ_8 parameter and arose during an early inflationary period (Bond et al. 1996). During gravitational collapse, cosmic baryons, which make up approximately 15 % of the mass in the universe, conform to the same spacial distribution as the dark matter in the universe. The formation of a superheated plasma, known as the intracluster medium (ICM), then follows from adiabatic compression and virialization of the baryonic matter. Through the virial theorem, the mass of a cluster can be related to the temperature of its ICM giving temperatures of $kT \sim (0.2 - 10)$ keV for cluster masses of $M \sim (10^{14} - 10^{15})$ M$_\odot$ (solar masses). At these temperatures, the ICM becomes fully ionised and thus produces X-ray radiation due to thermal bremsstrahlung. These emissions, first observed during the 1970's (Gursky et al.), had been proposed to originate

due to thermal bremsstrahlung several years earlier by Felten et al. (1966). The collision of elements present within the ICM leads to the conversion of their kinetic energy to X-ray emission lines which overlay the bremsstrahlung continuum. For cluster temperatures of $kT > 2$ keV the bremsstrahlung continuum dominates over these emission lines. For temperatures of $kT < 2$ keV the opposite is true and is due to the incomplete ionisation of the gasses.

1.1 Studying the cosmos with galaxy clusters

The growth of large scale structure in the universe follows directly from the amplitude of the perturbations in the primordial density distribution, quantified by the σ_8 parameter. As such, galaxy clusters provide a map to this growth of structure and the number density of the clusters themselves provide insight into a range of proposed cosmological models (Press & Schechter 1974; Kofman et al. 1993; Bahcall & Cen 1993; White et al. 1993). Typically, a cosmological model will be defined in terms of a set of cosmological parameters such as σ_8 and the cosmological constant and matter density parameters Ω_Λ and Ω_m respectfully. The various models of cosmology give us direct predictions of how the large scale structure of the universe evolves. This evolution is quantified by measuring the cluster mass function $n(M)$, which describes the number of clusters of mass M, per unit co-moving volume element. The mass function which arises from a range of different cosmologies can be simulated and then statistically fit to a physically measured mass function (Bode et al. 2001; Vikhlinin et al. 2009; Voit 2005; Paranjape 2014). In order to measure this mass function, galaxy clusters must essentially be counted and binned at their respective mass and redshift. This gives us three criteria which must be met in order

for a survey of galaxy clusters to provide useful cosmological results; cluster identification over a large range of redshifts, a reliable estimator of cluster mass and a precise knowledge of any observational biases. Hence, by optimising the data analysis protocols that maximise the potential to recover clusters over a wide range of masses and redshifts, we can likewise maximise the significance of the cosmological results obtained.

The use of X-ray data in constructing statistically complete catalogs of clusters has significantly contributed to our picture of their evolution. The refinement and extension of this work continually benefits from the advent of new and improved X-ray observatories. The first large orbiting X-ray telescope to make use of mirrors was the Einstein Observatory which began operating in 1978. This provided the first X-ray cluster catalogs which were constructed in order to answer cosmological questions (Abramopoulos & Ku 1983). Launched in 1990, the ROSAT all-sky survey was the first mission to survey the entire X-ray sky. The 380 clusters detected by ROSAT provided the basis for multiple subsequent, long exposure or 'pointed', observations (Bohringer et al. 2000). Following a line of subsequent missions, the Chandra and XMM-Newton observatories have provided the leading advancements in the field of X-ray astronomy. The significant improvement in angular resolution provided by the Chandra telescope, enabled it to observe X-ray sources 100 times fainter than any of its predecessors (Schwartz 2014). Deployed in the same year of 1999 XMM-Newton, with its large collecting area (Jansen et al. 2001), provided constraints on $\Omega_m = 0.316 \pm 0.060$ and $\sigma_8 = 0.814 \pm 0.054$ (Pacaud et al. 2018). These observatories, both launched in 1999, could however only perform pointed observations; due to their limited field of view, it was unfeasible to perform large area surveys with those telescopes. This led to the proposal of the eROSITA instrument, a hard X-ray imaging telescope with all-sky surveying capabilities (Merloni et al. 2012).

The eROSITA telescope is the primary instrument placed on board the Spectrum-Roentgen-Gamma satellite. Successfully launched on July 13th 2019, eROSITA is placed in a halo orbit around the L2 Lagrange point. The primary experimental aim of eROSITA is to measure the location and signatures of 100 000 galaxy clusters over the entire sky (Predehl et al. 2021). This sample will include all clusters with masses $M > 3 \times 10^{14} \, \mathrm{M_\odot}$ at redshifts out to $z \sim 1.5$. In addition to this, somewhere in the region of $\sim 3,000,000$ active galactic nuclei (AGN) are expected to be detected. Over a four year period, the scanning strategy employed by eROSITA will provide full-sky coverage a total of eight times over in the $0.5 - 10$ keV X-ray band (Predehl 2017). The state of the art technology employed by eROSITA will provide a sensitivity which is ~ 30 times grater than ROSAT, dramatically improving upon previous missions.

The primary science driver of the eROSITA mission is the study of dark energy through the discovery of all massive galaxy clusters in the observable universe (Albrecht et al. 2006). The two competing approaches to understanding dark energy require either an extension to the laws of gravity or the introduction of a new particle (Carroll 2001). Growth of structure measurements are currently the only known method to determine which of these approaches is correct (Merloni et al. 2012). From these measurements the eROSITA mission will look to constrain cosmological parameters to within $1 - 2$ orders of magnitude of the currently available X-ray cluster samples (Pillepich et al. 2018). Forecasts for constraints expected from the eROSITA cluster counts include $\Delta\Omega_m = 0.0026$ and $\Delta\sigma_8 = 0.006$ (Pillepich et al. 2012). The scope of the growth of structure measurements will be facilitated

by the incredible sensitivity of eROSITA which will allow the detection of the most massive clusters out to a redshift of $z \sim 1.6$. The detection of these objects presents a significant challenge and will require robust and effective detection algorithms. Multiple other mission objectives include studying the evolution of the thermal structure and chemical enrichment of clusters, deepening our understanding of inflationary theory and creating catalogs of superclusters of galaxy clusters (Guth 1981; Linde 1982; Bartolo et al. 2004; Chen 2010; Einasto et al. 2008).

All of the goals set by the eROSITA mission require precise knowledge of observational biases. These biases are quantified by the so called galaxy cluster 'selection function' which gives us the probability that source of a given mass and redshift will be included in the final cluster catalogue. In general, we see a decrease in probability as the redshift increases and below a flux limit we see zero objects included. Without precise knowledge of the selection function, the experimentally measured mass function will be skewed and the corresponding cosmological parameters biased. The eROSITA selection function will depend upon many factors which include, the detection pipeline used and the nature of the background noise to which eROSITA is exposed. The detection pipeline simply describes the data analysis protocols used to identify galaxy clusters from the raw X-ray data. By improving the methods used in the detection pipeline, the skew quantified by the selection function will be reduced and biasing effects minimised.

When characterising a selection function it is also important to consider and define the characteristics of purity and completeness (White & Kochanek 2002). Plainly speaking, purity is a measure of the number of false positives present in a sample of candidate sources whilst, relative to some true number, completeness is a general measurement of how many clusters the survey was able to detect. These measures therefore tell researchers not only how complete their sample is, but also how contaminated it is.

1.2 Background interference

The background noise experienced by X-ray observatories, is composed of two primary elements, an X-ray background (XRB) and a particle background. The X-ray background emanates from sources which lie both within and outside of the Milky Way (Kuntz & Snowden 2000). The extragalactic sources originate primarily from AGN, which constitutes 74% of the soft XRB (Lehmann et al. 2001). This extragalactic X-ray background presents a diffuse emission which is evenly distributed on the sky. The X-rays which emanate from these sources are subject to absorption due to the elemental abundance present in the interstellar medium. The number density of neutral hydrogen N_H traces the distribution of these elements and therefore provides an accurate indicator of the galactic foreground absorption (Morrison & McCammon 1983). The X-ray background which emanates from within the Milky Way originates from a range of objects such as supernovae remnants and stars. This galactic foreground emission provides a source of soft X-rays which varies with position on the sky.

The particle background is composed of soft protons originating from solar flares (O'Dell et al. 2000) and hard protons in the form of cosmic rays. When particles with energies comparable to X-rays make contact with a detector, the energy is recorded and incorrectly measured as an X-ray event. In addition to this, a florescent background is produced by particles hitting parts of the telescope. This results in the presence of emission lines in the background spectrum. The exposure of the telescope to soft protons increases in intensity as the satellite transitions through the radiation
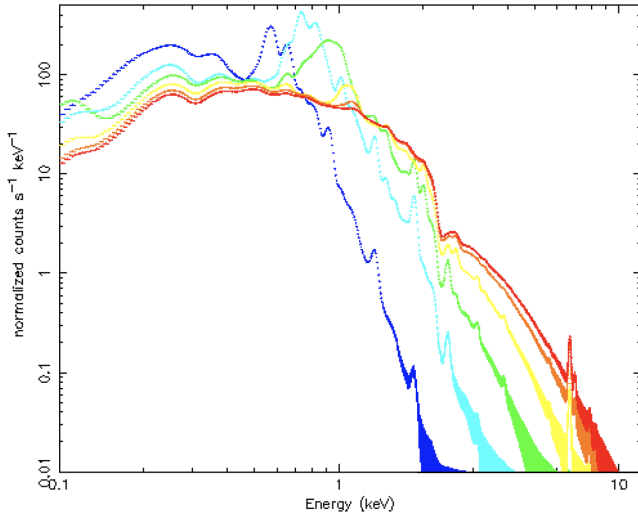
Figure 1. Simulated spectra of galaxy cluster X-ray emission. Plasma temperatures are, 0.2 keV (blue), 0.5 keV (turquoise), 1 keV (green), 2 keV (yellow), 4 keV (orange) and 6 keV (red). All sources have a redshift of $z = 0$ and an arbitrary normalisation set to provide the clearest spectral profiles possible.

belt (O'Dell et al. 2000). As such, soft protons present a highly variable component of the background. The presence of soft protons must be carefully managed as evidenced by their impact on previous missions which, in the case of XMM-Newton, managed to contaminate 40% of its observing time (Gastaldello et al. 2017).

Relative to a low earth orbit, the location of eROSTIA at the L2 point comes at the cost of a substantial increase in cosmic radiation (Perinati et al. 2012). The resulting increase in particle background is primarily due to the lack of shielding provided by the earth's magnetic field. Preliminary eROSITA measurements have however recorded the particle background to be even grater than expected from pre-launch estimates (Merloni et al. 2012; Predehl et al. 2021). This increase is believed to be due to the pre-launch date coinciding with a solar minimum when a date closer to solar maximum was expected. The anti-correlation between solar cycle and instrumental background has been previously established from XMM-Newton and Chandra observations (Gonzalez-Riestra & Rodriguez-Pascual 2019). Line features have also been detected within the eROSITA background spectrum. Most notably a strong Al-K$_\alpha$ line at ~ 1.5 keV originating due to the 4 mm thick aluminium filter wheel located between the optical path and CCD's.

1.3 Motivation for the study

A fundamental component of any detection pipeline will be the choice of X-ray energies, or "energy band", used to create an image of a galaxy cluster. By limiting the image to a specific energy band, it is possible to reduce the level of background interference whilst maximising the level from the source. Galaxy clusters of different temperatures and at different redshifts will produce different emission spectra and different X-ray telescopes will be exposed to different levels of background radiation. As such, the energy band which provides the greatest signal-to-noise ratio (SNR) will be different for each specific telescope and galaxy cluster. Figure 1 displays the emission spectra produced by clusters over a range of temperatures and highlights the degree of the spectral

changes. The detection of clusters from their X-ray emission has commonly made use of the $0.5 - 2$ keV bandpass. This "conventionally used" band typically helps minimize the contribution of background noise whilst maximising the sensitivity to general cluster emission (McHardy et al. 1998; Rosati et al. 1995).

In 2002, Scharf examined the SNR provided by the conventionally used band for the Chandra ACIS-I, ACIS-S, XMM-Newton MOS, and PN instruments. For each instrument a set of 'optimal' bands was recorded. These bands were optimised to provide the maximum SNR for clusters at specific temperatures and redshifts. For cooler plasma temperatures of $kT < 2$ keV and higher redshifts $z > 1$, it was shown that, relative to an optimal band, the conventionally used $0.5 - 2$ keV band incurs a signal-to-noise loss of between $20 - 30\%$. It was concluded that, as future missions seek to probe these lower mass and higher redshift clusters, the de-facto use of the conventionally used band should be reassessed. During their study it was however not tested to see if an increase in the SNR of an image of a galaxy cluster would make that cluster easier to detect. The next step in furthering this work would logically be to test the use of optimal, SNR energy bands for the forthcoming eROSITA instrument. It is this investigation on which the study presented here is founded.

In this paper the optimal, SNR energy bands for the eROSITA instrument are calculated. The optimal, SNR energy bands are then tested on the recently released eROSITA full-sky photon simulation (Comparat et al. 2020). Access to the data for this simulation is made available at MPE/MPG, Germany [2]. The eROSITA full-sky photon simulation was constructed by generating an N-body dark matter only set of full-sky light cones using the MultiDark (MDPL2) simulation (Klypin et al. 2016). The X-ray catalogue of galaxy clusters generated by this procedure accurately mimics the eROSITA scanning strategy. Due to artefacts arising from the alignment of structures in the MultiDark (MDPL2) simulation (Merson et al. 2013), only clusters located below a redshift of $z < 2$ are presented in the final catalog. From scans taken of this cluster catalog, further work estimating the flux limit of eROSITA is also conducted.

2 METHODOLOGY

A range of optimal, SNR energy bands for the X-ray emission from galaxy clusters was calculated for the eROSITA observatory. Galaxy clusters were generated over a range of temperatures, $kT < 10$ keV and over a range of redshifts, $z < 2$. An optimal SNR energy band was then calculated for each cluster. The effectiveness, relative the the conventionally used band, of the optimal energy bands in detecting galaxy clusters was then determined.

2.1 Optimal bandpass determinations

An optimal SNR energy band is defined as the range of X-ray photon energies which provide the maximum SNR for an image of a galaxy cluster of a given redshift and temperature. From here on I shall simply refer to a galaxy cluster as a 'source' and an optimal SNR energy band as an 'optimal' band. The SNR criterion used here is

$$SNR = \frac{st}{\sqrt{bt}} \qquad (1)$$

---

[2] https://www.mpe.mpg.de/~comparat/eROSITA_mock/

where s is the count rate of a source, b the background count rate and t is the observation time. This form of the SNR is chosen in order to make a direct comparison with the results obtained by Scharf (2002).

Explicit details of the protocols used to calculate the optimal bands are provided in this section. Here, I start with a brief summary of those procedures. The optimal energy bands were calculated by initially deriving a typical background spectrum for the eROSITA observatory. Sources over a range of temperatures and redshifts were then simulated. For each source spectra the count rates s and b were determined over a grid of multiple energy bands. The SNR was calculated for each of the energy bands and the band which provided the greatest SNR was recorded as the optimal band for that source. The ratio of the SNR measured in the conventional band to that in an optimal band was also recorded and will be referred to as b_r, the 'band ratio'. This procedure was repeated for each source providing a value of b_r and an optimal band for each one. All procedures were embedded within a Python automation script which facilitated the data collection protocols.

The background spectrum was created by firstly emulating the model presented in Merloni et al. (2012) and then modifying the parameters of the model to match the measured spectrum presented in Predehl et al. (2021). The background spectrum comprises two primary elements, an X-ray background and a particle background. All models described in the following were implemented within NASA's 'XSPEC', X-ray spectral fitting software package.

The particle background is able to circumvent the main body of the telescope and is therefore unaffected by the attenuation of certain energies mapped by the aperture response function (ARF). The ARF describes the sensitivity of the telescope to certain energies and characterises the instrumental response to incoming X-rays. The interaction of the particles with the detector was emulated by folding the particle model through the response matix file (RMF), which gives the probability that a photon will be correctly binned by the CCD. The particle background was modelled using two simple power laws with photon indexes of $\alpha_1 = 3$ and $\alpha_2 = 0.2$ and normalisations of 4×10^{-6} and 6×10^{-5} photons/keV/cm^2/s at 1 keV respectfully. The second power law was added in order to emulate the grater than expected particle background present at energies below about 0.4 keV (Predehl et al. 2021). In order to reproduce the presence of the Al-K$_\alpha$ line present in the measured spectrum a simple Gaussian line profile was added at a frequency of 8.34 Angstrom.

The X-ray background passes through the main body of the telescope prior to interacting with the detector. As such, the models which emulate its composition were folded through the both the RMF and ARF. The extragalactic unresolved emission and galactic foreground absorption were modelled using a Wabs photoelectric absorption, multiplied by a power law (Morrison & McCammon 1983; Lumb et al. 2002). The photon index and normalization were set to $\alpha_3 = 1.42$ and 1.5×10^{-6} photons/keV/cm^2/s at 1 keV respectfully and a galactic foreground absorption of $N_H = 1.7 \times 10^{20}$ cm^2 was assumed. The optically thin background emission from within the Milky Way was modeled using two Mekal models (Mewe et al. 1986; Liedahl et al. 1995) with temperatures of $kT_1 = 8.1 \times 10^{-2}$ keV and $kT_2 = 0.204$ keV and normalisations of 7.59×10^{-12} and 8×10^{-7} photons/keV/cm^2/s at 1 keV respectfully. Figure 2 displays the eROSITA background spectrum generated by the above procedures.

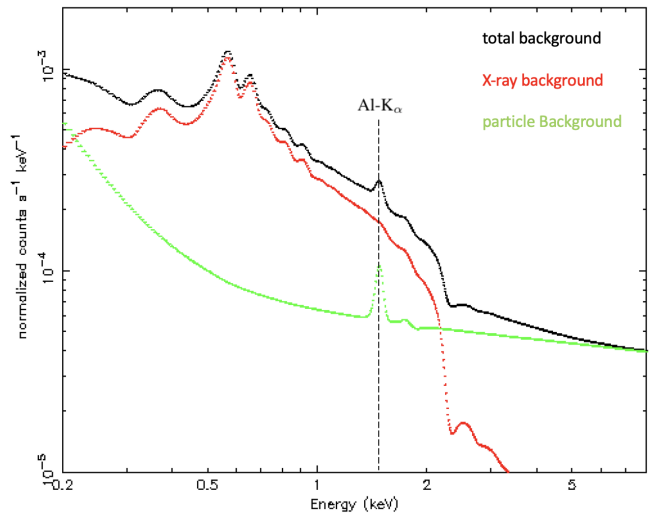Galaxy clusters were generated using XSPEC with an APEC

Figure 2. Simulated eROSITA background spectral components. In red the X-ray background (galactic foreground plus extragalactic), in green the particle background and in black the total background.

plasma emission spectrum model set to a metal abundance of 0.3 solar units. A galactic foreground absorption of $n_H = 1.7 \times 10^{20}$ cm^2 was again assumed. The normalisation for each source was set at 0.1×10^{-2} and was derived by calculating the value required to give a SNR ~ 10 for an observation time of $t = 2548$ seconds. The observation time was selected as the average exposure reached for eROSITA after the completion of the four year all-sky survey (Merloni et al. 2012). Sources were generated at temperatures of 0.2, 0.5, 1, 2, 4 and 6 keV. For each temperature a set of sources was produced at redshifts from 0.1 to 1.5 in steps of 0.1. These ranges were selected in order to make a direct comparison with the work conducted by (Scharf 2002). Each source spectra was then folded through the same on-axis RMF and ARF eROSITA files used in the creation of the background spectrum. Figure 1 displays the source spectra generated for all six of the plasma temperatures at a redshift of $z = 0$.

The grid of energy bands was generated through the cyclic permutation of a lower E_1 and upper E_2 band-limit, which defined the boundaries of an energy band. E_1 and E_2 were varied independently from 0 to 6 keV in steps of 0.1 keV. This procedure resulted in a set of energy bands which covered every possible width and position within an energy range of $(0-6)$ keV with a resolution of 0.1 keV.

2.2 Detection of galaxy clusters

Explicit details of the protocols used to the test the optimal bands are provided in this section. Here, I begin with a brief summary of those procedures. The eROSITA full-sky photon simulation of galaxy clusters was processed into a set of X-ray images. Each image was confined to an optimal band which had been selected to proivide the maximum SNR for soruces at specific temperature and redshift ranges. These specific sources I shall from now on refer to simply as 'band specific sources'. Each image, which I shall now refer to as an 'optimised image' was then scanned by a source detection algorithm. The number of band specific sources detected in the corresponding optimised image was then counted. From these counts a 'purity' and 'completeness' was determined. Finally, the
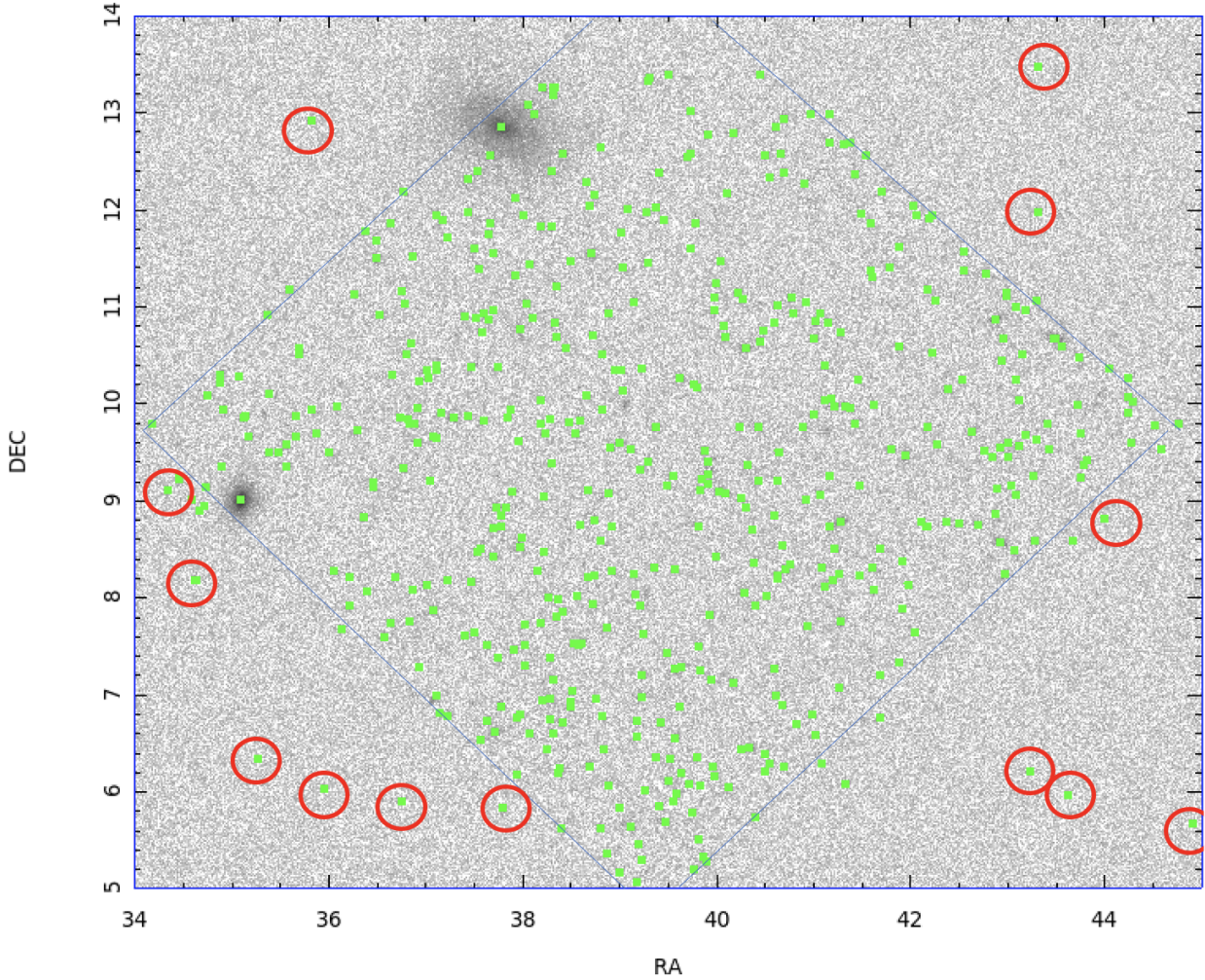
Figure 3. Declination vs. Right Ascension of the photons observed (black dots) in the conventional $0.5 - 2.0$ keV band for a typical $t = 2500$'s eROSITA observation. Green dots represent galaxy cluster detections made by Wavdetec. Background noise fills the entire square image and simulated eROSITA observation have been added to the central diamond tile. Detections made beyond the tile are entirely due to Wavdetect mistaking background noise for cluster emissions (red circles).

flux of every source detection made by the conventional band was recorded and a flux limit was derived.

For the purposes of this investigation I define the purity of an optimised image as

$$P = 1 - \frac{F}{T} \tag{2}$$

where F is the number of false positives and T the total number of sources of all redshifts and temperatures identified by the detection algorithm in that optimised image. The completeness of an optimised image I define as

$$C = \frac{N_{\text{band}}}{N_{\text{true}}} \tag{3}$$

where N_{band} is the number of band specific sources identified by the detection algorithm in that optimised image and N_{true} the true number of band specific sources present.

The simulated eROSITA data was composed of a source catalog and a set of events lists. The source catalog provided the precise location, temperature, redshift and flux of all the sources within the simulation. The events lists contained the position, time and energy of every X-ray event produced by the sources within the simula-

tion. A total of 7 events lists, one for each of eROSITA's ccd's, were grouped into a total of 768 tiles which covered the entire sky. Based on the lengthy computational time required to process each image, a total of ten such tiles were selected for analysis.

Four temperature ranges were selected, targeted around 1, 2, 4 and 6 keV. These values correspond to the temperatures at which the optimal energy bands had been previously calculated. For each temperature range, four redshift ranges were selected from $0 - 2$ in steps of 0.5. This provided a total of 16 discrete subsections within the temperature and redshift parameter space. Each subsection was assigned an energy band which had been optimised to provide the maximum SNR for sources at the median point within that subsection. The left hand column of table 1 displays the precise boundaries of each temperature range. Ranges vary depending on the weighting of sources present within the source catalog. For the upper two temperature ranges where the median points does not lie precisely at 4 and 6 keV respectfully, the corresponding optimal bands were calculated by interpolation of the previously recorded energy band data. The median points for these two bands is 3.75 keV and 7.5 keV respectfully.

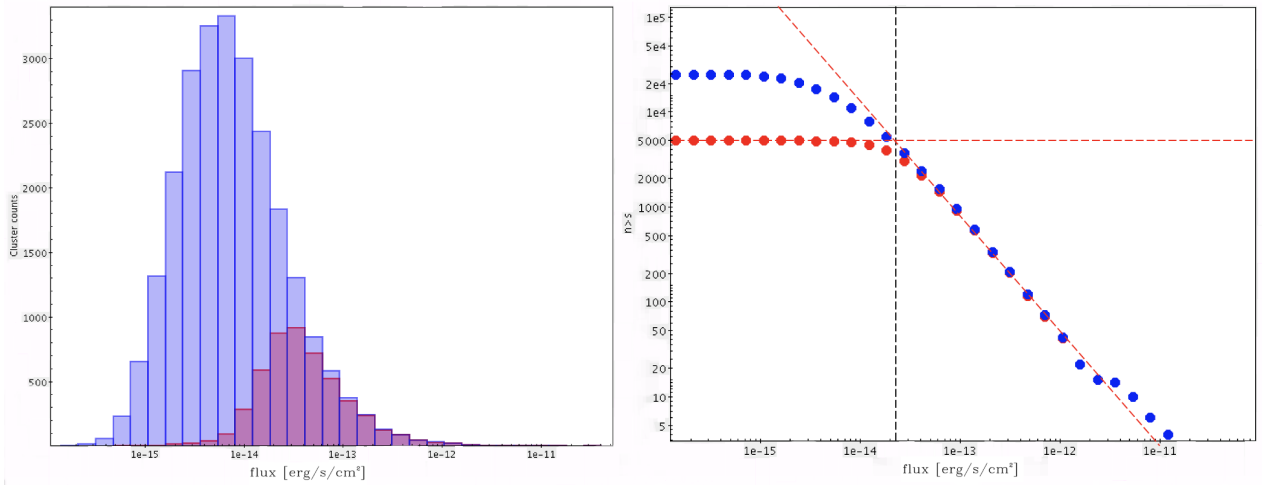For each of the ten tiles, an optimised image was created for

Figure 4. Left panel: Histogram displaying cluster counts recorded as a function of their flux. The cluster counts recorded directly from the source catalog are displayed in blue and the cluster counts detected in the conventional band are displayed in red. Right panel: Displays the number of sources n, with a flux grater than some value s, recorded as a function of that value. Blue data points display $(n > s)$ for cluster counts taken from the source catalog. Red data points display $(n > s)$ for cluster counts detected in the conventional band.

each of the 16 optimal energy bands and one in the conventional band. To construct an optimised image, X-ray data from each of the 7 ccd events lists was constrained to the required energy band and then stacked. A background image was added to the stack by applying the background spectrum count rate, in a given band, to each pixel in the image. Poisson noise was then applied across the entire background image.

Optimised images were processed using the Wavdetect detection algorithm which provided a list of source detections and corresponding location data. The total number of entries in this list provided the value of T. The location data from this list was indexed to the source catalog which provided each detected source with a corresponding value of redshift, temperature and flux. The number of false positives, F recorded in a given optimised image could then be calculated by counting the number of sources which could not be matched to an object in the source catalog. The purity of each optimised image, P_o could then be calculated. This procedure was then repeated for an image which had been created in the conventional band to give a value of P_c. A Purity ratio was then determined and is defined as

$$P_r = \frac{P_c}{P_o} \qquad (4)$$

This procedure was repeated for all of the 16 optimised images providing a value of P_r for each of the corresponding 16 subsections of parameter space.

The number of band specific sources present in the list of source detections, N_{band} was then counted. The true number of band specific sources N_{true}, which included the number missed by Wavdetect, was counted directly from the source catalog. The completeness of each optimised image, C_o could then be calculated. This procedure was then repeated for an image which had been created in the conventional band to give a value of C_c. A completeness ratio was then determined and is defined as

$$C_r = \frac{C_c}{C_o} = \frac{N_c}{N_o} \qquad (5)$$

where N_o is the number of band specific sources identified in the optimal band and N_c is the number of those same band specific

sources identified in the conventional band. This procedure was repeated for all of the 16 optimised images providing a value of C_r for each of the corresponding 16 subsections of parameter space.

The measurements of P_r and C_r were then repeated for each of the ten tiles providing a set of average $\bar{C}_r$ and $\bar{P}_r$ values. Figure 3 displays a typical example of an image generated in order to make these determinations.

Finally, the flux was recorded for every source detected by the conventional band across all 10 of the tiles. The left hand panel of figure 4 displays this data as a histogram. The number of clusters of a given flux recorded by the conventional band is displayed in red and the number recorded directly from the source catalog are displayed in blue. From this data, a function of the number of sources n, with a flux s, grater than some value of s, was derived. The right hand panel of figure 4 displays this function for cluster counts recorded directly from the source catalog (blue) and for counts detected in the conventional band (red). The median point in the "knee" of the red curve represents where the completeness rapidly drops off and the secure identification sources ceases. It is this point where the flux limit itself is defined.

3 RESULTS

Optimal energy bands were recorded for sources with plasma temperatures of 0.2, 0.5, 1, 2, 4 and 6 keV and over redshifts from 0.1 to 1.5 in steps of 0.1. The left hand panel of figure 5 displays the optimal energy band recorded as a function of redshift for each of the six plasma temperatures. The right hand panel displays the band ratio b_r, for each of the six plasma temperatures also recorded as a function of redshift. Lower values of b_r represent an increase in the SNR of an optimal band relative to the conventional band. For sources with a plasma temperature $kT < 2$ keV and at redshift $z = 0$, the improvement in SNR provided by an optimal band ranges from $\sim 55\% - 10\%$. For these lower temperature systems, the conventionally used band suffers a further deterioration in SNR for sources with redshift $z > 0$. For sources with a plasma temperature $kT > 2\,\text{keV}$, he improvement in SNR provided by an optimal

band ranges from between $\sim 10\%$ at $z = 0$ and $\sim 10\% - 30\%$ at $z = 1.5$.

Sixteen completeness ratios $\bar{C}_r$, were recorded, one for each of the 16 subsections of the temperature and redshift parameter space. Figure 6 displays the results of the completeness ratio calculations (left panel) and their associated errors (right panel). In the left hand panel 'hotter' values indicate a decrease in the performance of the optimal band relative to the conventional band. A value of $\bar{C}_r = 1$ indicates that the number of source detections was the same for both the conventional and the corresponding optimal band. The blank region in the $1.5 < z < 2$ and $0.5 < kT < 1.5$ keV subsection of the parameter space is due to a complete lack of source detections in the corresponding 0.2 - 0.5 keV optimal band. The subsection which displays the greatest improvement in the number of sources detected by an optimal band over the conventional band is at $1.5 < z < 2$ and $1.5 < kT < 2.5$ keV with $C_r = 0.8 \pm 0.4$. The error associated with this results is however significant and therefore rules out any improvement by the optimal band at that subsection. At plasma temperatures of $kT < 2$ keV, the optimal bands produced less source detections than the conventional band. The worst performance by an optimal band is in the $0.5 < kT < 1.5$ keV and $0.5 < z < 1$ subsection with $C_r = 3.2 \pm 0.7$. Overall, the completeness ratios display results which in all cases indicate that, relative to the conventional band, an optimal band produces either the same or less source detections. Appendix A provides a full derivation of the $\Delta \bar{C}_r$ values.

As with the completeness ratios, Sixteen purity ratios $\bar{P}_r$, were recorded, one for each of the 16 subsections of the temperature and redshift parameter space. Figure 7 displays the results of the purity ratio calculations (left panel) and their associated errors (right panel). The values of $\bar{P}_r$ indicate that for all subsections of the parameter space, with the exception of the $kT < 2$keV and $z > 0.8$ subsection, where a value of $\bar{P}_r = 1.25 \pm 0.05$ was recorded, the optimal bands produce a purity which is highly comparable to that produced by the conventional band. Appendix A provides a full derivation of the $\Delta \bar{P}_r$ values, which were calculated as the standard error of the $\bar{P}_r$ values.

The results of the band ratio calculations clearly show that, relative to the conventional band, optimal bands provide a significant increase in the SNR of band specific sources. The completeness ratio calculations measure, relative to the conventional band, the number of band specific sources that were detected by Wavdetect. The fundamental assumption on which this study is based is that increasing the SRN of band specific sources will make them easier to detect. By measuring the correlation between $\bar{C}_r$ and b_r this assumption can be tested statistically. Further to this, a possible dependence of the number of sources detected in a given band and the width of that band is also examined.

The left hand panel of figure 8 displays the correlation between the completeness ratio $\bar{C}_r$ and the corresponding band ratio b_r. Data points were fit using a linear regression model which gives a slope for the line of best fit of -0.4 ± 0.4. The Pearson's-r (product moment correlation coefficient) for the model is -0.33 and the reduced χ^2 is 1.5. The Pearson's-r gives a low strength of association between $\bar{C}_r$ and b_r. The reduced χ^2 is > 1 also indicating that the fit has not fully captured the data. The incomplete description of the data by the model, poor correlation between data points and the relatively large error on the slope all show that no significant scaling between $\bar{C}_r$ and b_r can be concluded.

The right hand panel of figure 8 displays the correlation between the completeness ratio $\bar{C}_r$ and the bandwidth of the corresponding optimal band. Data points were fit using a linear re-

gression model which gives a slope for the line of best fit of -0.16 ± 0.07. The Pearson's-r for the model is -0.67 and the reduced χ^2 is 1.6. The Pearson's-r gives a moderate strength of association between $\bar{C}_r$ and the bandwidth, however the reduced χ^2 is > 1. The incomplete description of the data by the model and the relatively large error on the slope show that no significant scaling between $\bar{C}_r$ and bandwidth can be concluded.

A flux limit in the conventional $0.5 - 2$ keV band of the eROSITA observatory of $(2.1 \pm 0.5) \times 10^{-14}$ erg/s/cm^2 was recorded. The right hand plot of figure 4 displays the results of the flux limit calculation which was measured for an observation time of $t = 2.5$ ks.

Table 1 displays the numerical results and errors of the purity ratio $\bar{P}_r$, completeness ratio $\bar{C}_r$ and band ratio b_r, calculations. The first and second columns define the subsection boundaries of the temperature and redshift parameter space respectfully. The third and fourth columns give the band limits and bandwidth respectfully of the optimal bands.

4   DISCUSSION

4.1   eROSITA background spectrum

The effect of galactic foreground absorption on the extra-galactic X-ray component of the background was assumed to be constant. This assumption is incorrect, with measurements of N_H varying from between 10^{22} cm^{-2} near the galactic core to 10^{19} cm^{-2} at a galactic latitude of $\sim 90^\circ$ (HI4PI Collaboration et al. 2016). The value N_H used for the derived eROSITA background spectrum of $N_H = 1.7 \times 10^{20}$ cm^{-2} was taken from the preflight estimates presented by Merloni et al. (2012). This value was calculated from the average measurements performed by Lumb et al. (2002) as part of background measurements taken at high galactic latitude for the XMM-Newton EPIC cameras. The locations on the sky at which they recorded N_H ranged from $40^\circ < l < 170^\circ$ and $55^\circ < b < 80^\circ$. This section of the sky lies at a grater galactic latitude than the tiles examined by my investigation which lie in a range $144^\circ < l < 190^\circ$ and $-25^\circ < b < -48^\circ$. The HI4PI full-sky hydrogen column density survey (HI4PI Collaboration et al. 2016) provides access to data from the recently completed Effelsberg-Bom H1 survey (EBHIS) which maps atomic neutral hydrogen across the entire sky. Their study provides accurate values of $(2.5 \times 10^{20} < N_H < 28 \times 10^{20})$ cm^{-2} for the tile locations used in my study. This gives values of N_H which are between 1.5 and 16 times grater the assumed value of $N_H = 1.7 \times 10^{20}$ cm^{-2}. The significant increase in the absorption column on the eROSITA background spectrum would result in a reduction in the intensity of the extra-galactic X-ray component. This would, in-turn, reduce the sensitivity of the optimal SNR energy bands on the $E < 2$ keV part of the background spectrum where the extra-galactic X-ray component dominates.

The galactic foreground emission was also assumed to constant with sky position. This assumption is also incorrect with both spatial and spectral variations arising with sky position. Beyond the galactic ridge (GR), at galactic latitudes of $b > 5^\circ$, the diffuse X-rays exhibit a surface brightness of less than $\sim 10\%$ of the extra-galactic X-ray background emission (Koyama 2018). As such, any quantitative study into this galactic region has been limited. The spectral measurements taken by Lumb et al. (2002) therefore provide an acceptable characterization of the galactic foreground emission at galactic latitudes of $b > 5^\circ$. The data they collected was
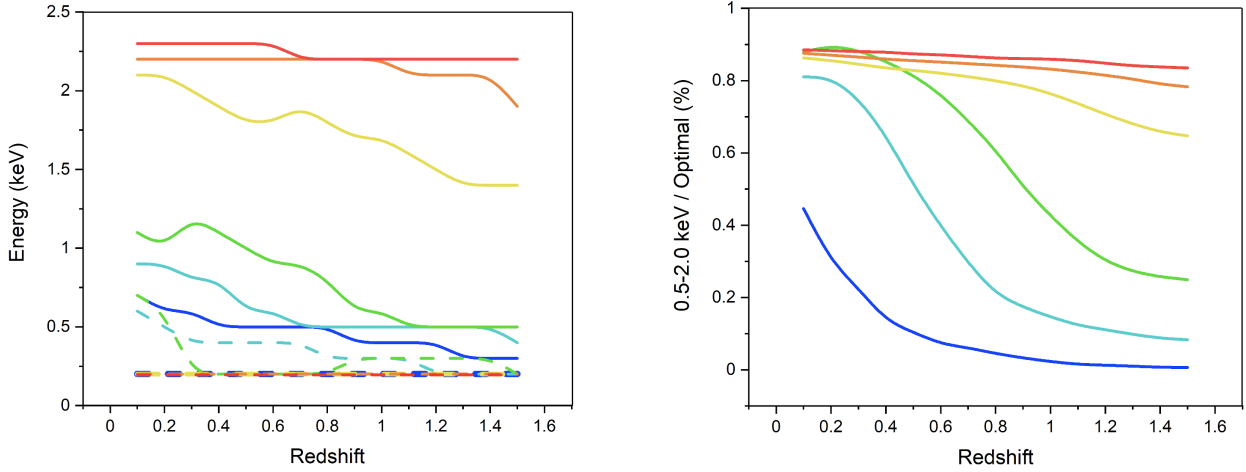
Figure 5. Results of an optimal band search for the eROSITA instrument. Left panel: Upper (solid) and lower (dashed) energy band limits (E_2 and E_1) are plotted as a function of redshift. Plasma temperatures are, 0.2 keV (blue curves), 0.5 keV (turquoise curves), 1 keV (green curves), 2 keV (yellow curves), 4 keV (orange curves) and 6 keV (red curves). Right panel: The band ratio, b_r is plotted (as a percentage) vs. redshift for the six plasma temperatures.
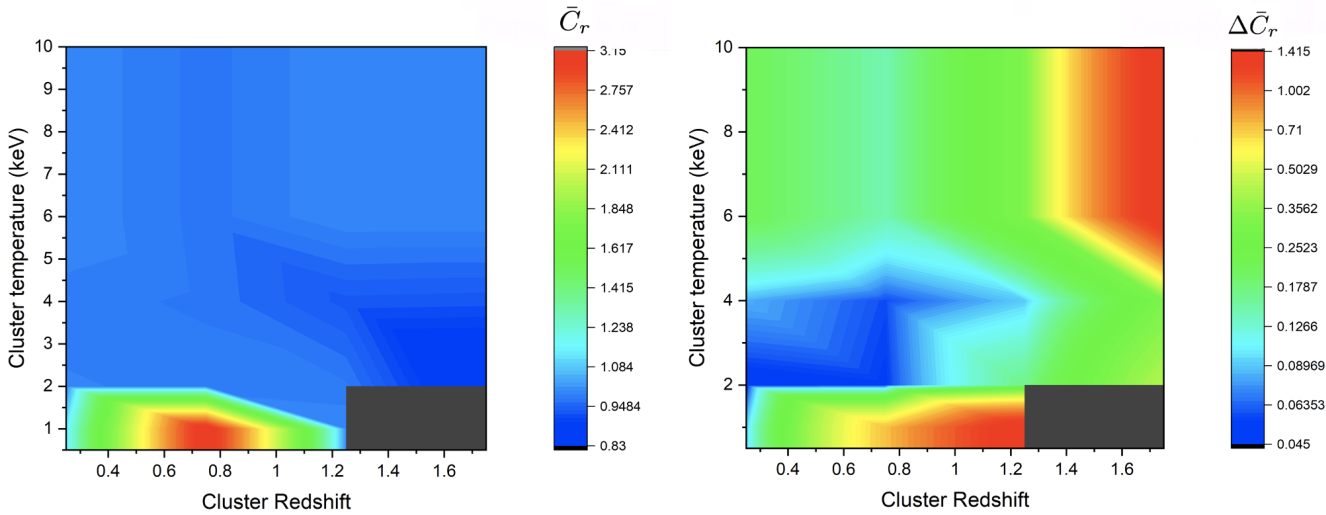

Figure 6. Contour plots displaying the Completeness ratios $\bar{C}_r$ (left panel) and their associated errors, $\Delta\bar{C}_r$ (right panel) recorded as a function of temperature and redshift.

statistically fit to two MEKAL models, providing the parameters used in my study. The modifications made to these values in order to match the spectrum observed by eROSITA (Predehl et al. 2021), reflect a background spectrum characteristic of an all-sky survey and are therefore averaged over the entire sky. As such, the galactic foreground emission derived by my study is not precisely reflective of the sky location I examined. In order to quantify this difference, pointed eROSITA observation at galactic galactic latitudes of $-25° < b < -48°$ will be required.

The pre-flight measurements of the particle background (Predehl et al. 2021) were recorded at a local solar minimum, resulting in a corresponding local maximum in particle background intensity (Gonzalez-Riestra & Rodriguez-Pascual 2019). The simulated eROSITA all-sky survey data was designed to mimic the 4 year duration of the mission (Comparat et al. 2020). As such, a more accurate particle background measurement would have been one which was taken at the median point between the solar maximum and four years into a typical 11 year solar cycle (Brehm et al. 2021).

In the absence of direct measurements, the preflight estimates of the particle background represent the best approximation of the minimum levels, expected when the estimates were made (Merloni et al. 2012). A revised level at the median point between the two could therefore be deduced and would naturally be lower than the levels used in this study. A reduction in particle background intensity would effect the entire $(0-10)$ keV energy range of the background spectrum. Such a reduction would be expected to result in broader optimal energy bands overall, as the SNR calculations would be less contaminated by the background and therefore be able to capture more of the source spectrum.

4.2 Optimal energy bands

The optimal bands display a dependence on both temperature and redshift. In order to understand the temperature dependence, we consider figure 1, which demonstrates that as the cluster temperature increases, the spectrum flattens and the peek is shifted towards
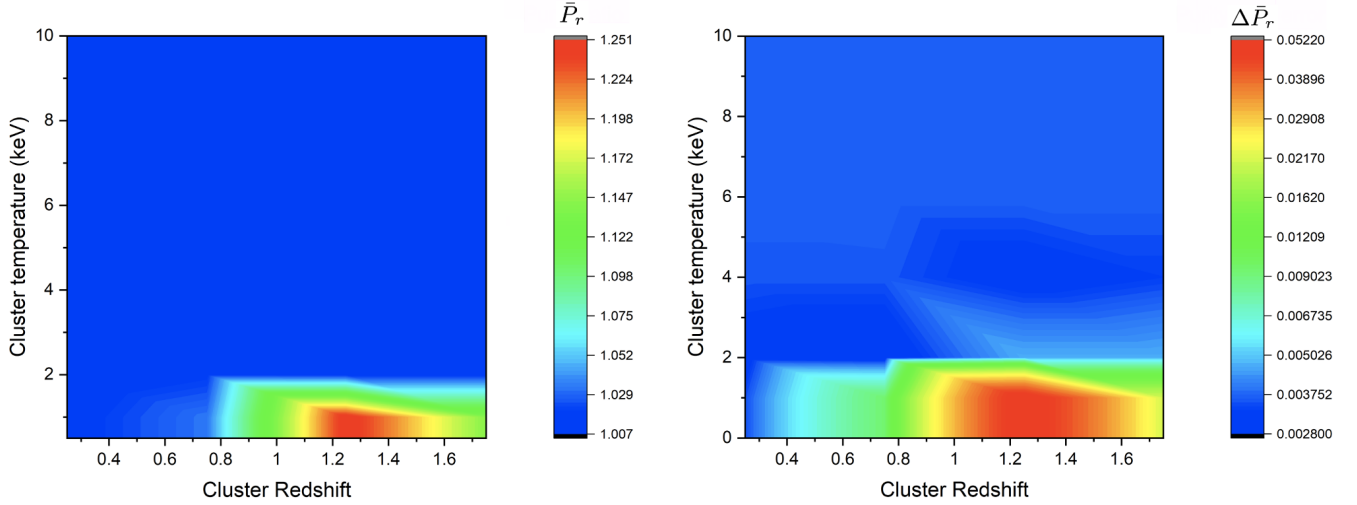
Figure 7. Contour plots displaying the Purity ratios $\bar{P}_r$ (left panel) and their associated errors, $\Delta\bar{P}_r$ (right panel) recorded as a function of temperature and redshift.

Table 1. Numerical results and errors of the purity ratio $\bar{P}_r$, completeness ratio $\bar{C}_r$ and band ratio b_r, calculations. The first and second columns define the subsection boundaries of the temperature and redshift parameter space respectfully. The third and fourth columns give the band limits and bandwidth respectfully of the optimal bands.

Temperature range (keV)	Redshift range	Optimal band (keV)	Bandwidth (keV)	$\bar{P}_r$	$\bar{C}_r$	b_r
0.5 - 1.5	0 - 0.5	0.2 - 1.2	1.0	1.015 ± 0.004	1.14 ± 0.09	0.8839
	0.5 - 1	0.2 - 0.8	0.6	1.033 ± 0.009	3.2 ± 0.7	0.6081
	1 - 1.5	0.3 - 0.5	0.2	1.25 ± 0.05	1 ± 1	0.2719
	1.5 - 2	0.2 - 0.5	0.3	1.14 ± 0.02	- -	0.2125
1.5 - 2.5	0 - 0.5	0.2 - 2.1	1.9	1.010 ± 0.003	0.99 ± 0.05	0.8456
	0.5 - 1	0.2 - 1.8	1.6	1.019 ± 0.003	0.98 ± 0.05	0.7999
	1 - 1.5	0.2 - 1.4	1.2	1.017 ± 0.005	1.0 ± 0.2	0.6801
	1.5 - 2	0.2 - 1.4	1.2	1.017 ± 0.005	0.8 ± 0.4	0.6154
2.5 - 5	0 - 0.5	0.2 - 2.2	2.0	1.015 ± 0.004	0.98 ± 0.08	0.8653
	0.5 - 1	0.2 - 2.2	2.0	1.015 ± 0.004	0.97 ± 0.06	0.8424
	1 - 1.5	0.2 - 2.1	1.9	1.010 ± 0.003	0.93 ± 0.09	0.8049
	1.5 - 2	0.2 - 1.9	1.7	1.008 ± 0.003	0.9 ± 0.3	0.7775
5 - 10	0 - 0.5	0.2 - 2.3	2.1	1.008 ± 0.004	0.9 ± 0.2	0.8851
	0.5 - 1	0.2 - 2.2	2.0	1.014 ± 0.004	1.0 ± 0.1	0.8621
	1 - 1.5	0.2 - 2.2	2.0	1.014 ± 0.004	1.0 ± 0.3	0.8409
	1.5 - 2	0.2 - 2.2	2.0	1.014 ± 0.004	1.0 ± 1.0	0.8231

higher energies. The shift towards higher energies pushes the peak over the region of the background where the particle component dominates. The particle component lies approximately one order of magnitude in intensity below the X-ray component and thus impacts the SNR to a significantly reduced degree. The upper band limit for higher temperature clusters is thus less restricted and able to capture more of the source spectrum, thus resulting in the borader bands we see at temperatures $kT > 2$ keV. In understanding the redshift dependence of the optimal bands, we consider that the peak of a cluster spectrum broadly represents the region where the SNR is greatest. Therefore, as this peak is redshifted towards lower energies, the optimal band will track this shift to maintain the optimal SNR.

The features we see present in the band-limits E_1 and E_2, flatten out as the plasma temperature increases. In order to understand

why this happens, we consider firstly that as the source spectrum is redshifted the background spectrum remains static. When peaks in the source and background spectra cross one another, features form as the band limits jump in order to avoid relative increases in the background spectrum. For lower temperature clusters, this effect becomes more pronounced as the bremsstrahlung continuum becomes less dominant and so the emission lines become more prominent. The results presented by Scharf (2002) provide a strong example of this phenomenon. For the Chandra ACIS-S BI chips, they report a sharp drop in the upper band-limit for the 6 keV temperature clusters as the spectrum redshifts from 0.6 to 0.7. This is entirely due to the location of a fluorescent Si-K$_\alpha$ line in the background at ~ 1.7 keV which the band-limit jumps to avoid. In contrast to this study, we see no such feature at an energy of 1.5 keV where the Al-K$_\alpha$ line is present. The band limits used by Scharf

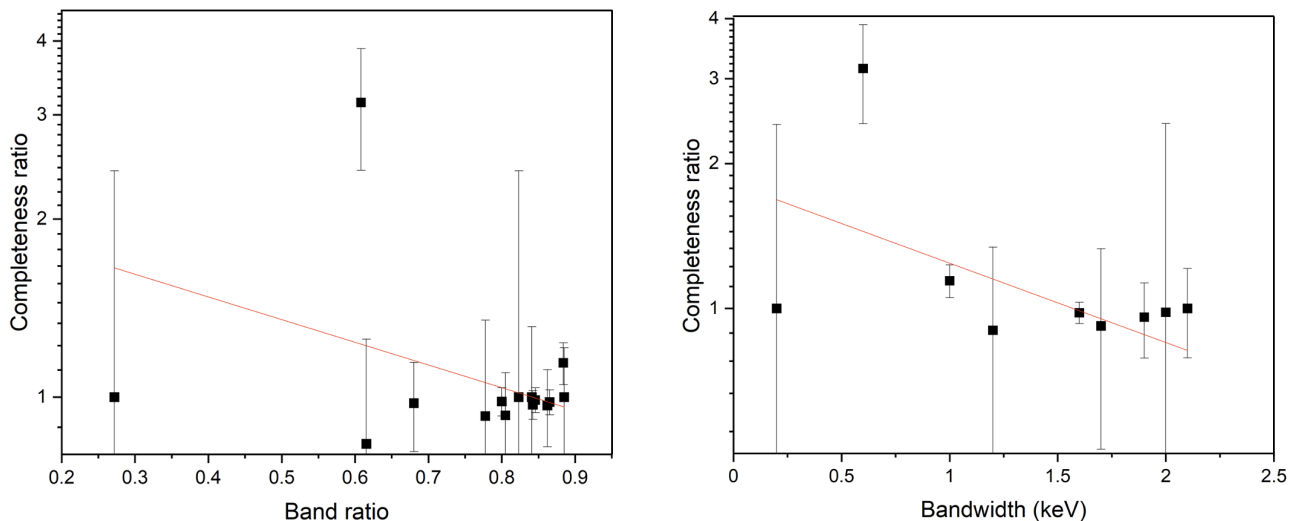**Figure 8.** Left panel: Correlation between $\bar{C}_r$ and b_r. Right panel: Correlation between $\bar{C}_r$ and bandwidth.

(2002) were recorded at steps of 0.01 keV providing an energy resolution which was approximately one order of magnitude better than that used in this study. This relative lack of resolution could conceivably result in the omission of fine detail and obscure the effect of the Al-K$_\alpha$ emission line in the band limits.

As with the galactic foreground absorption on the extra-galactic X-ray component of the background, N_H was assumed constant for all galaxy clusters. Figure 9 displays a comparison between a typical source spectra subject to two levels of galactic foreground absorption. The black curve has been subjected to the level assumed in this study of $N_H = 1.7 \times 10^{20}$ cm^{-2} and the red curve has been subjected to the upper bound on the corrected level of $N_H = 28 \times 10^{20}$ cm^{-2} derived in section 4.1. The impact of this systematic overestimation of the intensity of the source spectra would be a relative decrease in the level of background noise present in the energy band calculations. Specifically, the particle and galactic foreground emission would be less prominent. The extra-galactic X-ray component would remain unchanged as it too would be subject to the increased degree of attenuation. Ultimately, as discussed above, a decrease in relative background intensity would result in broader optimal bands. Given however that the particle background was also overestimated to an unknown degree and the correct levels of galactic foreground emissions are unknown, the net effect on the band limits is difficult to quantify without conducting the further investigations previously suggested.

The band ratios recorded a relative deterioration in the SNR provided by the conventional band for low temperature, $kT < 2$ keV, high redshift $z > 1$, sources. The primary reason for this deterioration is due to excess noise being captured by the conventional band. For higher temperature $kT > 2$ keV, low redshift $z < 1$ sources, the deterioration becomes negligible as there is very little difference between the band limits of the conventional and optimal bands. These results display a high degree of correlation to the work carried out by Scharf (2002) for the Chandra ACIS-I, ACIS-S, XMM-Newton MOS, and PN instruments. For these observatories an average $20\% - 60\%$ SNR loss compared to an optimal band was recorded. The confirmation of this result for the eROSITA observatory provides compelling evidence that the use of optimal bands in detecting these sources should be carefully considered.

4.3 Purity ratio calculations

The primary source of false positives detected in a given optimised image was due to the presence of the background noise added to that image. Figure 3 highlights this phenomenon and displays a typical X-ray image used in this study. Background noise fills the entire square image and a simulated eROSITA observation has been added to the central diamond tile. Outside of the tile only background noise exists and so detections made in this region are entirely due to Wavdetect mistaking background noise for cluster emissions.

In order to understand why the purity ratios were so consistent across the parameter space, we firstly consider that as the bandwidth of an optimal band increases, the level of background noise also increases. This increase is not linear as the background spectrum itself is not constant. The noise added to an image created in the conventional band captured the majority of the most intense region in the background spectrum. An optimal band with a bandwidth grater than the conventional band would therefore present a negligible addition to the noise added to an image. The noise added to an image which was created using an optimal band was in all cases also taken from the most intense region of the background spectrum. An optimal band with a bandwidth which was less than the conventional band would therefore also present a negligible addition to the noise added to an image. In summary, the variation in noise added by different bands was in all cases negligible and therefore did not significantly skew the results obtained by Wavdetect.

A major limiting factor in testing the true level of purity provided by different bands, was the lack of AGN in the simulated eROSITA data (Comparat et al. 2020). Typically, AGN will present a contribution to the number of false positives recorded by detection algorithms. This contribution may be caused by clusters hosting bright X-ray AGN or by point source projection along the line of sight of diffuse cluster emission (Rosati et al. 2002). The effect of the former case has been shown to be negligible, were the number of clusters which host large numbers of bright AGN at their centre has been shown to be $< 5\%$ (Böhringer et al. 2001). The latter contribution has however shown to be significant, where Stern et al. (2002) showed Chandra observations of distant ROSAT se-

lected clusters revealed up to 50% flux contamination. The omission of AGN from the simulated eROSITA data therefore translates to a significantly biased set of results for the absolute purity levels P_o and P_c. As this study looked at the relative levels of purity P_r, this effect would in principle present a negligible impact with any biasing effectively cancelling in the ratio of equation 4. This of course assumes that the presence of AGN would result in the same levels of degradation in purity between an optimal and the conventional band.

4.4 Completeness ratio calculations

The band ratios indicated that an optimal band would provide the largest improvement in SNR for low temperature, $kT < 2$ keV, high redshift, $z > 1$ sources. This region of the parameter space is precisely where the values of $\Delta \bar{C}_r$ were the greatest. This is due to a combination of the rarity of these types of clusters in the source catalog and the low number of photon counts produced by them. In order to obtain a large enough sample to generate a statistically complete and significant result for $\bar{C}_r$, a conservative estimate of ten times the number of tiles examined would be required. This is based on the scaling of $\Delta \bar{C}_r$ with the sample size n and the average size of the error for the current sample size of 10 tiles. The measurements of $\bar{C}_r$ indicate that any performance gains by an optimal band are at best small. As such, the corresponding values of $\Delta \bar{C}_r$ must be equivalently small in order to measure these gains precisely and with confidence.

In order to understand why, relative to the conventional band, no clear improvement in source detection by an optimal band was recorded, we must examine the Wavdetect algorithm used to scan the X-ray images. Wavdetect is a Wavelet-based source detection algorithm which involves the transformation of scaled wavelet functions to two-dimensional image data (Freeman et al. 2002). The primary function of the Wavelet Transform is the isolation of point sources from an image to reveal the smoothed, diffuse source emission of the galaxy cluster (Grebenev et al. 1995). A major advantage to this method is that it picks out structures in an extended object by their scale, dramatically enhancing their identifying features (Starck & Pierre 1998). As such, X-ray images of galaxy clusters become significantly clearer after being processed by the Wavelet Transform. With this in mind, we consider that during the course of this study, the effectiveness of the Wavelet Transform has essentially been tested. From the results of the completeness ratio calculations, it would appear quite clear that the X-ray images presented to Wavdetect during the course of this study presented no significant challenge to its detection capabilities. As such, the lack of any clear improvement in source detection by an optimal band can most likely be attributed to the robustness of the Wavdetect algorithm.

The emission of X-ray photons from galaxy clusters is a random process which follows a Poisson probability distribution (Weinberg et al. 2013). Sources produce X-rays independently from one another and so the mean count rate of a given source will be unique. In running the mock eROSITA simulation multiple times (Comparat et al. 2020), one could obtain samples that varied in a way which was reflective of this underlying Poisson probability distribution. The number of source detections made for each simulation would therefore vary according to this statistical noise. This approach was unfortunately not possible and the method employed by this study took data from multiple spatial locations in the form of the 10 tiles. Using this approach, the most obvious way to calcu-
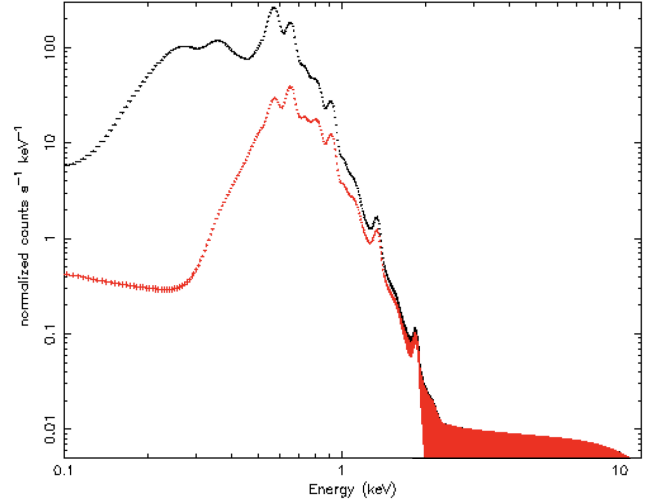

Figure 9. Source spectra for a cluster of temperature $kT = 0.2$ keV and redshift $z = 0$ subject to galactic foreground absorption of $N_H = 1.7 \times 10^{20}$ cm^{-2} (black) and $N_H = 28 \times 10^{20}$ cm^{-2} (red).

late $\bar{C}_r$ would be to calculate $C_r = N_c / N_o$ in each tile, and then use equation A1 which would give the mean of this ratio. Due to the low numbers of detections in multiple subsections of the parameter space, this approach resulted in many values of $\bar{C}_r$ being undefined and multiple omission in the parameter space. In many cases, only one detection was made in both N_c and N_o. In such cases if both detections were not recorded in the same tile, then the value of C_r would be undefined for every tile. By taking $\bar{C}_r$ as the ratio of the means, $\bar{N}_c / \bar{N}_o$, this issue is avoided as it did not require counts in N_o and N_c to come from the same tile. In order to calculate the error of the ratio of the means, the method of relating the error of a function to the differential of that function was used and is derived in appendix A.

Calculating the ratio of the means instead of the mean of a ratio does however come at a cost. For this specific case, Jensen's inequality can be used to prove that the mean of the ratio is always less than or equal to the ratio of the means (Needham 1993). The effect of this is that of a systematic overestimation of the values $\bar{C}_r$. An examination between the two methods of calculating the mean revealed differences of $\sim 2\%$. The question of whether to take the mean of the ratio or the ratio of the means have been thoroughly explored by statisticians and the trade of between the accuracy of the mean and the number of omissions in your resulting data set must be carefully considered (Rao 2002).

4.5 Correlations and Flux limit

The two measures of correlation between $\bar{C}_r$ against b_r and $\bar{C}_r$ against bandwidth both assume a linear model. In the absence of any prior theoretical model which describes these relationships, the linear models assumed serves as a reasonable starting point in understanding any potential dependencies. In considering the methods used to measure the validity of the linear model, it should be noted that the reliability of the reduced chi squared as a measure of the fit is somewhat dubious. Due to random noise in the data, the value of reduced chi-squared itself is subject to noise (Andrae et al. 2010). However, when considering the general agreement be-

tween the statistical measures of Pearson's-r and the error on the slope, any linear dependency between $\bar{C}_r$ against b_r and $\bar{C}_r$ against bandwidth can be ruled out with a high degree of confidence.

The error derived for the flux limit calculation was taken as the width in flux between data points at the median point in the "knee" of the red curve displayed in the right hand panel of figure 4. Qualitatively, the knee is where the number of source detections rapidly drops off to zero. As such, the median point in the knee provides a sensible definition of where the flux limit is located (Mateos et al. 2008). The value of the flux limit predicted as part of the pre-flight estimated derived by Merloni et al. (2012) lies within the error of the results presented in this paper. Using a Poisson probability model to estimate the flux limit as a function of observation time, Merloni et al. (2012) report a flux limit of 3.5×10^{-14} erg/s/cm^2 (Weisskopf et al. 2007). Their result is purely theoretical and was derived for the 2.5 ks observation used by eROSITA. As part of their study it was concluded that a follow up investigation using simulated data would be required to ratify the result. As such, the results provided by the analysis of the simulated eROSITA events lists used in this study are highly relevant to confirming their result and quantifying the flux limit of the eROSITA selection function.

5 CONCLUSION

Optimal, SNR energy bands were determined for the eROSITA observatory . Data taken from the simulated eROSITA all-sky-survey (Comparat et al. 2020) was used to test the performance of the optimal bands for sources with temperatures $(0.5 < kT < 10)$ keV and redshifts $0 < z < 2$. The performance of optimal bands was measured in terms of purity and completeness ratios, $\bar{P}_r$ and $\bar{C}_r$ respectfully. The flux limit of the eROSITA selection function in the conventional $0.5 - 2$ keV band for an observation time of $t = 2.5$ ks was also measured. The main conclusions of these investigations is as follows.

• For sources with a plasma temperature $kT < 2$ keV and at redshift $z = 0$, the improvement in SNR provided by an optimal band ranges from $\sim 55\% - 10\%$. For these lower temperature systems, the conventionally used band suffers a further deterioration in SNR for sources with redshift $z > 0$. For sources with a plasma temperature $kT > 2$ keV, the improvement in SNR provided by an optimal band ranges from between $\sim 10\%$ at $z = 0$ and $\sim 10\% - 30\%$ at $z = 1.5$. These results lie in excellent agreement with the results provided by Scharf (2002) for the the Chandra ACIS-I, ACIS-S, XMM-Newton MOS, and PN instruments.

• Systematic biases in the derived eROSITA background spectrum resulted in a general overestimation in the intensity of the particle and extra-galactic components. The magnitude of the galactic foreground absorption was found to be between 1.5 and 16 times grater the assumed value of $N_H = 1.7 \times 10^{20}$ cm^{-2} used in this study. The intensity of the source spectra were likewise overestimated due to the same effect resulting in a negligible net effect on the resulting optimal bands. Uncertainties remain in the precise levels of particle and galactic foreground emission used and require further investigation in order to fully quantify the true effect on the optimal band limits.

• The completeness ratios vary from between $\bar{C}_r = 3.2 \pm 0.7$ to $\bar{C}_r = 0.8 \pm 0.4$ which indicate that, relative to the conventional band, an optimal band produces either the same or less numbers of band specific source detections. However, due to the limited scope of the study, low numbers of the low temperature, $kT < 2$ keV, high redshift, $z > 1$ sources were detected resulting in statistically

insignificant results for these objects. In order to compensate for this, an alternative method was used to derive the mean values, $\bar{C}_r$. This resulted in a systematic overestimation of the values $\bar{C}_r$ by $\sim 2\%$. A conservative estimate of ten times the number of tiles examined would be required in order to mitigate this statistical bias.

• The values of $\bar{P}_r$ indicate that for all subsections of the parameter space, with the exception of the $kT < 2$keV and $z > 0.8$ subsection, the optimal bands will return a number of false positives in an X-ray image which is highly comparable to that returned by the conventional band. The primary source of false positives detected in a given energy band was attributed to the presence of background noise.

• It was shown that improvements in the SNR provided by an optimal band did not make sources easier to detect. This was quantified by measuring the correlation between $\bar{C}_r$ and b_r where a Pearson's-r of -0.33 and reduced χ^2 of 1.5 was recorded.

• A possible dependence of $\bar{C}_r$ on the width of the optimal energy band was explored. No such dependence was recorded with the correlation between $\bar{C}_r$ and bandwidth yielding a Pearson's-r of -0.67 and reduced χ^2 of 1.6.

• A flux limit in the conventional $0.5 - 2$ keV band for the eROSITA observatory of $(2.1 \pm 0.5) \times 10^{-14}$ erg/s/cm^2 was recorded. This lies is excellent agreement with the value derived by Merloni et al. (2012) of 3.5×10^{-14} erg/s/cm^2. As such, the results provided by the analysis of the simulated eROSITA events lists used in this study are highly relevant to characterising the flux limited eROSITA selection function.

As the forthcoming eROSITA all-sky-survey looks to expand the detection number of lower mass higher redshift galaxy clusters, the techniques applied in the detection pipeline must be refined and enhanced. Although statistical uncertainties still remain, it is the primary conclusion of this study that no such enhancements are likely obtainable through the use of optimal, SNR energy bands.

REFERENCES

Abell G. O., 1958, ApJS, 3, 211
Abell G. O., Corwin Harold G. J., Olowin R. P., 1989, ApJS, 70, 1
Abramopoulos F., Ku W. H. M., 1983, ApJ, 271, 446
Albrecht A., et al., 2006, arXiv e-prints, pp astro–ph/0609591
Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Andrae R., Schulze-Hartung T., Melchior P., 2010, arXiv e-prints, p. arXiv:1012.3754
Bahcall N. A., Cen R., 1993, ApJ, 407, L49
Bartolo N., Komatsu E., Matarrese S., Riotto A., 2004, Phys. Rep., 402, 103
Bode P., Bahcall N. A., Ford E. B., Ostriker J. P., 2001, ApJ, 551, 15
Bohringer H., et al., 2000, VizieR Online Data Catalog, p. J/ApJS/129/435
Böhringer H., et al., 2001, A&A, 369, 826
Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603
Brehm N., et al., 2021, Nature Geoscience, 14, 10
Carroll S. M., 2001, Living Reviews in Relativity, 4, 1
Chen X., 2010, Advances in Astronomy, 2010, 638979
Comparat J., et al., 2020, arXiv e-prints
Einasto M., et al., 2008, ApJ, 685, 83
Felten J. E., Gould R. J., Stein W. A., Woolf N. J., 1966, ApJ, 146, 955
Freeman P. E., Kashyap V., Rosner R., Lamb D. Q., 2002, ApJS, 138, 185
Gastaldello F., et al., 2017, Experimental Astronomy, 44, 321
Gonzalez-Riestra R., Rodriguez-Pascual P., 2019, Tech, 33, 39
Grebenev S. A., Forman W., Jones C., Murray S., 1995, ApJ, 445, 607
Gursky H., Solinger A., Kellogg E. M., Murray S., Tananbaum H., Giacconi R., Cavaliere A., 1972, ApJ, 173, L99
Guth A. H., 1981, Phys. Rev. D, 23, 347
HI4PI Collaboration et al., 2016, A&A, 594, A116

Jansen F., et al., 2001, A&A, 365, L1

Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340

Kofman L. A., Gnedin N. Y., Bahcall N. A., 1993, ApJ, 413, 1

Koyama K., 2018, PASJ, 70, R1

Kuntz K. D., Snowden S. L., 2000, in AAS/High Energy Astrophysics Division #5. p. 32.26

Lehmann I., et al., 2001, in White N. E., Malaguti G., Palumbo G. G. C., eds, American Institute of Physics Conference Series Vol. 599, X-ray Astronomy: Stellar Endpoints, AGN, and the Diffuse X-ray Background. pp 189–198, doi:10.1063/1.1434632

Liedahl D. A., Osterheld A. L., Goldstein W. H., 1995, ApJ, 438, L115

Linde A. D., 1982, Physics Letters B, 108, 389

Lumb D. H., Warwick R. S., Page M., De Luca A., 2002, A&A, 389, 93

Mateos S., et al., 2008, A&A, 492, 51

McHardy I. M., et al., 1998, MNRAS, 295, 641

Merloni A., et al., 2012, arXiv e-prints, p. arXiv:1209.3114

Merson A. I., et al., 2013, MNRAS, 429, 556

Mewe R., Lemen J. R., van den Oord G. H. J., 1986, A&AS, 65, 511

Morrison R., McCammon D., 1983, ApJ, 270, 119

Needham T., 1993, The American Mathematical Monthly, 100, 768

O'Dell S. L., et al., 2000, in Flanagan K. A., Siegmund O. H. W., eds, Vol. 4140, X-Ray and Gamma-Ray Instrumentation for Astronomy XI. SPIE, pp 99 – 110, doi:10.1117/12.409104

Pacaud F., et al., 2018, A&A, 620, A10

Paranjape A., 2014, Phys. Rev. D, 90, 023520

Perinati E., Tenzer C., Santangelo A., Dennerl K., Freyberg M., Predehl P., 2012, Experimental Astronomy, 33, 39

Pillepich A., Porciani C., Reiprich T. H., 2012, MNRAS, 422, 44

Pillepich A., Reiprich T. H., Porciani C., Borm K., Merloni A., 2018, MNRAS, 481, 613

Predehl P., 2017, Astronomische Nachrichten, 338, 159

Predehl P., et al., 2021, A&A, 647, A1

Press W. H., Schechter P., 1974, ApJ, 187, 425

Rao T., 2002, Journal of Statistical Planning and Inference, 102, 129

Rosati P., Della Ceca R., Burg R., Norman C., Giacconi R., 1995, ApJ, 445, L11

Rosati P., Borgani S., Norman C., 2002, ARA&A, 40, 539

Scharf C., 2002, ApJ, 572, 157

Schwartz D. A., 2014, Review of Scientific Instruments, 85, 061101

Starck J. L., Pierre M., 1998, A&AS, 128, 397

Stern D., et al., 2002, The Astronomical Journal, 123, 2223

Vikhlinin A., et al., 2009, ApJ, 692, 1033

Voit G. M., 2005, Reviews of Modern Physics, 77, 207

Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys. Rep., 530, 87

Weisskopf M. C., Wu K., Trimble V., O'Dell S. L., Elsner R. F., Zavlin V. E., Kouveliotou C., 2007, The Astrophysical Journal, 657, 1026

White M., Kochanek C. S., 2002, ApJ, 574, 24

White S. D. M., Navarro J. F., Evrard A. E., Frenk C. S., 1993, Nature, 366, 429

APPENDIX A: ERROR DERIVATIONS

The mean values $\bar{N}_o$, $\bar{N}_c$, $\bar{N}$ and $\bar{P}_r$ were calculated via

$$\mu_x = \frac{\Sigma_{i=1}^n (x_i)}{n} \qquad (A1)$$

with μ_x the respective mean and n the sample size which, in our case, was the the number of tiles examines, which was 10. The error on these values $\Delta\bar{N}_o$, $\Delta\bar{N}_c$, $\Delta\bar{N}$ and $\Delta\bar{P}_r$ was taken as the standard error

$$\Delta = \frac{\sigma}{\sqrt{n}} \qquad (A2)$$

where σ is the sample standard deviation given by

$$\sigma = \sqrt{\frac{\Sigma_{i=1}^n (x_i - \mu_x)^2}{n-1}} \qquad (A3)$$

The mean completeness in an optimal band was taken as

$$\bar{C}_o = \frac{\bar{N}_o}{\bar{N}} \qquad (A4)$$

which is dependent upon $\bar{N}_o$, the average number of band specific detections in an optimal band and $\bar{N}$ the average number of band specific sources present in the source catalog. The mean completeness in the conventional band is given as

$$\bar{C}_c = \frac{\bar{N}_c}{\bar{N}} \qquad (A5)$$

with $\bar{N}_c$, the average number of band specific detections in the conventional band. The mean completeness ratio is then given as

$$\bar{C}_r = \frac{\bar{C}_c}{\bar{C}_o} = \frac{\bar{N}_c}{\bar{N}_o} \qquad (A6)$$

Through the general relation between the error of a function and the error of its dependent variables given in differential form by

$$df = \sqrt{\left(\frac{\partial f}{\partial x}dx\right)^2 + \left(\frac{\partial f}{\partial y}dy\right)^2} \qquad (A7)$$

the error on the mean completeness in an optimal band was determined to be

$$\Delta\bar{C}_o = \sqrt{\left(\frac{\Delta\bar{N}_o}{\bar{N}}\right)^2 + \left(\frac{\bar{N}_o}{\bar{N}^2}\Delta\bar{N}\right)^2} \qquad (A8)$$

the error on the mean completeness in the conventional band by

$$\Delta\bar{C}_c = \sqrt{\left(\frac{\Delta\bar{N}_c}{\bar{N}}\right)^2 + \left(\frac{\bar{N}_c}{\bar{N}^2}\Delta\bar{N}\right)^2} \qquad (A9)$$

and the error on the mean completeness ratio by

$$\Delta\bar{C}_r = \sqrt{\left(\frac{\Delta\bar{N}_c}{\bar{N}_o}\right)^2 + \left(\frac{\bar{N}_c}{\bar{N}_o^2}\Delta\bar{N}_o\right)^2} \qquad (A10)$$