FULL-SKY PHOTON SIMULATION OF CLUSTERS AND ACTIVE GALACTIC NUCLEI IN THE SOFT X-RAYS FOR EROSITA

Johan Comparat1, Dominique Eckert2, Alexis Finoqueiro3, Robert Schmidt4, Jeremy Sanders5, Daisuke Nagai6, Erwin T. Lau6, Florian Kaefer7, Florian Pacaud8, Nicolas Clerc9, Thomas H. Reiprich10, Esra Bulbul1, Jacob Ider Chitham1, Chia-Hsun Chuang12, Vittorio Ghirardini1, Violeta Gonzalez-Perez13, Ghassem Gozaliasl14, Charles C. Kirkpatrick5, Anatoly Klypin15, Andrea Merloni16, Kirpal Nandra17, Teng Liu18, Francisco Prada19, Miriam E. Ramos-Ceja20, Mara Salvato21, Riccardo Seppi22, Elmo Tempel23, and Gustavo Yepes24

1Max-Planck Institut für extraterrestrische Physik, Postfach 1312, D-85741 Garching bei München, Germany
2The Astronomy Department, University of Geneva, Ch. de Ecogia 16, CH-1290 Versoix, Switzerland
3Department of Physics, University of Helsinki, Gustaf Hällström katu 2a, FI-00014 Helsinki, Finland
4Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstrasse 12-14, 69120 Heidelberg, Germany
5Department of Physics, Yale University, New Haven, CT 06520, U.S.A.
6Department of Physics, University of Miami, Coral Gables, FL 33124, U.S.A.
7Argelander-Institut für Astronomie (AIA), Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
8IRAP, Université de Toulouse, CNRS, UPS, CNES, Toulouse, France
9Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA
10Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK
11Astronomy Department, New Mexico State University, Las Cruces, NM, USA
12Laboratoire d’Astrophysique, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland
13Institute for Computational Cosmology and Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham, DH1 3LE, UK
14Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, E-18080 Granada, Spain
15Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France
16IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France
17Tartu Observatory, University of Tartu, Observatooriumi 1, 51022 Tartu, Estonia and
18Departamento de Física Teórica y CIIAFF, Universidad Autónoma de Madrid, 28049 Madrid, Spain

Abstract

The eROSITA X-ray telescope on board the Spectrum-Roentgen-Gamma (SRG) mission will measure the position and properties of about 100 000 clusters of galaxies and 3 million active galactic nuclei over the full sky. To study the statistical properties of this ongoing survey, it is key to estimate the selection function accurately.

We create a set of full sky light-cones using the MultiDark and UNIT dark matter only N-body simulations. We present a novel method to predict the X-ray emission of galaxy clusters. Given a set of dark matter halo properties (mass, redshift, ellipticity, offset parameter), we construct an X-ray emissivity profile and image for each halo in the light-cone. We follow the eROSITA scanning strategy to produce a list of X-ray photons on the full sky.

We predict scaling relations for the model clusters, which are in good agreement with the literature. The predicted number density of clusters as a function of flux also agrees with previous measurements. Finally, we obtain a scatter of 0.21 (0.07, 0.25) for the X-ray luminosity – mass (temperature – mass, luminosity – temperature) model scaling relations.

We provide catalogues with the model photons emitted by clusters and active galactic nuclei. These catalogues will aid the eROSITA end to end simulation flow analysis and in particular the source detection process and cataloguing methods.

Keywords: cosmology - simulations - clusters of galaxies - active galactic nuclei

1. INTRODUCTION

As the most massive gravitationally bound objects in the Universe, Galaxy clusters have been widely used as probes for cosmology due to the exponential dependence of their abundance on cosmological parameters (see Voit 2005; Allen et al. 2011; Weinberg et al. 2013, for reviews).

The eROSITA X-ray telescope on board the Spectrum-Roentgen-Gamma (SRG) mission will measure the position and properties of about 100 000 clusters of galaxies and 3 million active galactic nuclei over the full sky (Merloni et al. 2012; Predehl et al. 2016). The large eROSITA cluster sample will improve cosmological constraints significantly (Pillepich et al. 2018b). The improvement is contingent on our ability to constrain the X-ray emitting intracluster medium (ICM) profiles out large radii (Walker et al. 2019) that affect the mass measurements of galaxy clusters (Pratt et al. 2019).

Over the last decade, there have been numerous advances in modeling of the ICM properties and their X-ray emissions. In particular, cosmological hydrodynamical simulations have been the keys for understanding the roles of cluster astrophysics (e.g. Springel et al. 2001; Na-
et al. 2015; Ishiyama et al. 2015, 2020 as they have suf-

Table 1. The MultiDark (UNIT) simulations are com-

Table 1. Characteristics and setup for each N-body dark matter only simulation. L: length of the box [Gpc/h]. $M_p$: Mass of the dark matter particle [$M_\odot$/h]. $M_{vir}$: Cut in halo mass [$M_\odot$/h]. All haloes with $M_{vir} > M_{vir}^{cut}$ are included in the light-cone. The mass limit is rounded a little above 100 particles. $z_{max}$: maximum redshift to the observer. The minimum redshift is 0 for all light-cones. N snap: Number of snapshots available within $z_{max}$. Ref: reference articles for the simulations (1): Klypin et al. (2016), (2): Chuang et al. (2019).

| Name          | $\Gamma_{box}$ | $M_p$      | $M_{vir}^{cut}$ | $z_{max}$ | N Snap | Ref   |
|---------------|----------------|------------|-----------------|-----------|--------|-------|
| SMDPPL        | 0.4            | $9.63 \times 10^7$ | $1 \times 10^{10}$ | 0.43       | 41 (1) |       |
| MDPL2         | 1.51           | $10^{10}$  | $2 \times 10^{11}$ | 6.2        | 91 (1) |       |
| HMDPL         | 7.9            | $10^{10}$  | $1 \times 10^{14}$ | 1.8        | 50 (1) |       |
| UNIT1         | 1.2            | $10^9$     | $2 \times 10^{11}$ | 6          | 89 (2) |       |
| UNIT1i        | 1.2            | $10^9$     | $2 \times 10^{11}$ | 6          | 89 (2) |       |

1 https://www.cosmosim.org
2 https://unitsim.fr.uam.es

On the other hand, phenomenological and semi-

analytical models (e.g. Shaw et al. 2010; Balaguera-

Antolínez et al. 2012; Zandanel et al. 2014; Flender

et al. 2017; Zandanel et al. 2018; Clerc et al. 2018) have

emerged as a complementary modeling approach to dir-

direct hydrodynamical simulations, allowing us to explore

ICM physics in a relatively computationally efficient way.

These models also allow us to generate full-sky photon-

level cluster simulations that are essential for eROSITA

(Zandanel et al. 2018).

In this paper, we present a novel method to model

X-ray cluster profiles. We adopt an empirical method

where we sample ICM profiles from high quality X-ray

observations of representative galaxy clusters. In this

way we are able to generate realistic clusters and produce

realistic X-ray full sky light-cones of dark matter only N-

body simulations for eROSITA. This empirical approach

is complementary to phenomenological and semi-

analytical models, as well as hydrodynamical simulations in that

the modeled ICM profiles are non-parametric other than

halo mass, redshift, and dynamical state.

The structure of the paper is as follows: We discuss

the cosmological N-body simulations used to create the

mock X-ray light-cone catalogues and photons in Sec-

tion 2. We describe models to simulate the X-ray prop-

erties of galaxy groups and clusters in Section 3. Finally,

in Section 4 we discuss the results and predictions of the

model.

We assume a flat $\Lambda$CDM cosmology close to that of

the Planck Collaboration et al. (2014b), detailed in the

following Section. The set of eROSITA mocks is made

public. Access is described in Appendix.

2. N-BODY DATA

We use the MultiDark (HMDPL, MDPL2, SM-

DPL Klypin et al. 2016)\(^1\) and the UNIT (UNIT1, UNIT1i Chuang et al. 2019)\(^2\) dark matter only simula-

tions to create light-cones spanning the full sky, see

Table 1. The MultiDark (UNIT) simulations are com-

puted in a flat $\Lambda$CDM cosmology with $H_0 = 67.77$

(67.74) km s\(^{-1}\) Mpc\(^{-1}\), $\Omega_m = 0.307115$ (0.308900),

$\Omega_\Lambda = 0.048206$, $\sigma_8 = 0.8228$ (0.8147). Alternative

simulations which could be used for this project include those of

Angulo et al. 2012; Skillman et al. 2014; Heitmann et al. 2015; Ishiyama et al. 2015, 2020 as they have suf-

ficiently high resolution for the AGN population to be

accurately modeled, which requires the full population of haloes with mass greater than $10^{12} M_\odot$.

The setup of each light-cone and parameters of each

simulation box used (5 boxes in total) are detailed in

Table 1. All simulations are consistently post-processed with

the ROCKSTAR software (Behroozi et al. 2013) to obtain haloes and sub-haloes. The halo mass function

is correct to a few percent down to halo masses of $M_{vir} \sim 10^{10} M_\odot$ for SMDPL, $M_{vir} \sim 10^{11} M_\odot$ for

MDPL2, $M_{vir} \sim 10^{12} M_\odot$ for HMDPL, see details in Comparat et al. (2017); Chuang et al. (2019).

We consider haloes and sub-haloes with slightly more than 100 bound particles per (sub) halo, see Table 1.

Table 2 gives the number of halos and sub-haloes for the complete halo population and for four mass thresholds:

$M_{vir} > 5, 7, 10, 50 \times 10^{13} h^{-1} M_\odot$. There are about 2 million (40 thousand) haloes (sub-haloes) with a halo mass

$M_{vir} > 5 \times 10^{13} h^{-1} M_\odot$ over the full sky. A quarter of them have $M_{vir} > 10^{14} h^{-1} M_\odot$ and only 2% of them

have $M_{vir} > 5 \times 10^{14} h^{-1} M_\odot$. The comparison of the fraction of sub-haloes in Table 2 shows that the resolution

matters to find sub-haloes in dense environments. For comparison, the eROSITA cluster sample is expected to be

complete for masses greater than $10^{14} (7 \times 10^{13}) M_\odot$ for redshifts $z < 1$ ($z < 0.3$) (Merloni et al. 2012; Clerc et al. 2018).

2.1. Light-cones

From the list of snapshot output ($z_{\text{snap}}$) in each simulation, we compute the comoving distance between the

snapshots redshifts and redshift 0, $d_C(z_{\text{snap}}) = d_C(z_{\text{snap}})$, using the exact cosmological parameters of the simulation:

$$d_C(z) = \int_0^z \frac{c \, du}{H_0 \sqrt{(1 + u)^3 \Omega_m + \Omega_D \Omega_{DE}}},$$  

(1)

We compute the middle point distance between each snapshot. These middle points are the boundaries between the shells; i.e. the distances at which we switch between using the outputs in a snapshot to another. The array of boundaries reads:

$$d_C^{\text{max boundary}} = \text{array (Mid points, } d_C^{\text{snap}}[-1])$$

$$d_C^{\text{min boundary}} = \text{array } (d_C^{\text{snap}}[0], \text{Mid points})$$
We compute the maximum number of replicas needed for each snapshot:

\[ N_{\text{replicas}} = \text{int}(C_{\text{max boundary}} / L_{\text{box}}) + 1. \]  

We replicate all snapshots available \((2 N_{\text{replicas}})^3\) that are within the redshift range of interest. The maximum to reach redshift 6 is \(12^3\)=1728 replicas. We use the periodic boundary condition to have consistent large scale structure across boundaries. It also preserves the angular \((2D)\) and monopole \((3D)\) clustering information on the full sky up to redshift 6. We obtain a regular grid with the observer located at \((0,0,0)\). In each replica, the coordinates \((x_{\text{snap}}, y_{\text{snap}}, z_{\text{snap}})\) are shifted by units of box length:

\[ x \leftarrow x_{\text{snap}} + i \times L_{\text{box}} [\text{Mpc}], \]
\[ y \leftarrow y_{\text{snap}} + j \times L_{\text{box}} [\text{Mpc}], \]
\[ z \leftarrow z_{\text{snap}} + k \times L_{\text{box}} [\text{Mpc}], \]  

where \((i,j,k)\) takes values between \(-N_{\text{replicas}}\) and \(N_{\text{replicas}} - 1\). In each replicated snapshot, we select objects with a comoving distance to the observer between \(d_{C_{\text{min boundary}}}^{}\) and \(d_{C_{\text{max boundary}}}^{}\). We concatenate each set to obtain a full-sky shell. The width of the shells varies between 100 and 120 Mpc, except for SMDPL where it is between 20 and 60Mpc. The properties of objects in a shell are at a fixed time (that of the snapshot). The redshifts of each \((\text{sub})\) halo \((\text{computed with or without line-of-sight velocity projection})\) give an accurate position, but not an accurate proper time, defined by the snapshot redshift. Indeed, we do not interpolate the positions and the velocities between snapshots to obtain a more continuous redshift sampling and set of halo properties. It is not needed in this study. The replication of snapshot and extraction of the shell typically requires 20-60 CPU hours.

As pointed out by Blaizot et al. (2005) and Merson et al. (2013), when replicating the simulations artefacts appear due to the alignment of structures. As we simulate the full sky, we cannot apply the technique proposed by Blaizot et al. (2005), best suited for pencil beam surveys. For HMDPL, the simulation is sufficiently large that there is no such artifact. For the 1Gpc/h simulations, the replications of a structure at redshift 0.1 \((\text{age } 12.48 \text{ Gyr})\) do occur at redshifts 0.1 , 0.49, 0.99, 1.71, 2.82, 4.71 \((\text{or ages } 8.71, 5.91, 3.82, 2.31, 1.27 \text{ Gyr})\). These redshifts are more spaced in time than the dynamical times of the haloes. We find that the effects shown in Fig. 1 and 2 of Blaizot et al. (2005) are hardly visible until redshift 2. Above redshift 2, as the number of replica increases drastically, they become visible. Thus the eROSITA cluster sample located below \(z < 1.5\) is unaffected by these artefacts. The eROSITA AGN sample with an extended high redshift tail is affected when considering the full sky area. For the smallest simulation, SMDPL, we limit the simulation in redshift to have a maximum of two line-of-sight replications. Due to its smaller volume and redshift reach, the use of SMDPL will be limited to studying the galaxy population (sub-haloes) in low redshift groups and clusters. The SMDPL simulation will not be used to create full sky X-ray maps where these artefacts matter.

Using SMDPL, we create one low redshift \((z < 0.43)\) light-cone that resolves groups and clusters and the sub-haloes therein. Using HMDPL, we create a cluster only light-cone as its resolution is too coarse to add AGN or galaxies in clusters. Using the MDPL2, UNIT1, UNIT1i simulations, we create light-cones \((z < 6)\) that contain all clusters, their galaxies (with some degree of incompleteness at low redshift) and the large scale structure sampled by AGN. The construction of the light-cone is close to that used by Merson et al. (2013). In the MDPL2, UNIT1, and UNIT1i light-cones, thanks to a sufficiently high resolution and a large volume, we simulate both the AGN and the clusters (Georgakakis et al. 2018; Comparat et al. 2019). Note that due to the limited volume, the high mass end \((> 10^{15} M_{\odot})\) of the halo mass function remains noisy. Below redshift \(z < 0.1 - 0.2\) the galaxies in the clusters and AGN populations suffer from the limited resolution of the simulation.

### 2.2. Coordinates

For each halo, we compute the angular coordinates (equatorial, galactic, ecliptic) and line-of-sight velocity vector to infer the redshift in real and redshift space. The simulated positions \((x, y, z)\) and peculiar velocities \((v_x, v_y, v_z)\) enable computation of observed coordinates as follows:

\[ d_C = \sqrt{x^2 + y^2 + z^2} [\text{Mpc}] \]

From \(d_C\), we infer the real space redshift: ‘redshift \(R\)’ through the comoving distance – redshift relation given in Eq. 1. Angular coordinates \((\text{in degrees})\) are then computed:

\[ \theta = \arccos(z/d_C) 180/\pi [\circ], \]
\[ \phi = \arctan(y/x) 180/\pi [\circ], \]
\[ \text{R.A.} = \phi + 180 [\circ], \]
\[ \text{Dec.} = \theta - 90 [\circ], \]

and converted to galactic and ecliptic coordinates with ASTROPY. For convenience, we split the sky into 768 equal arc pixels using healpix (NSIDE= 8 = 2^3 in the

| name | \(M_{\text{vir}}^\text{cut}^\text{out}\) | \(N\) haloes (h: haloes, s: sub-haloes) with \(M_{\text{vir}}^\text{cut}\) | \(5 \times 10^{13}\) | \(7 \times 10^{13}\) | \(10^{14}\) | \(5 \times 10^{14}\) |
|------|----------|--------------------------|----------|----------|----------|----------|
| SMDPL | 1,574,501,466 | 296,215,425 (18.8) | 319,746 | 10,796 (3.4) | 192,165 | 5,840 (3.0) | 106,581 | 2,939 (2.8) | 3,580 | 56 (1.6) |
| MDPL2 | 5,711,284,943 | 534,492,535 (9.4) | 2,219,880 | 40,903 (1.8) | 1,124,378 | 18,200 (1.6) | 519,691 | 6,953 (1.3) | 5,897 | 56 (1.3) |
| HMDPL | 28,897,265 | 1,150,831 (4.0) | 2,046,611 | 35,339 (1.7) | 1,051,562 | 14,491 (1.4) | 491,468 | 5,335 (1.1) | 11,615 | 56 (1.0) |
| UNIT1 | 5,923,863,122 | 1,024,611 (9.2) | 2,115,664 | 38,174 (1.8) | 1,060,239 | 16,470 (1.5) | 488,147 | 6,203 (1.3) | 7,590 | 76 (1.3) |
| UNIT1i | 5,917,731,143 | 1,054,039 (9.2) | 2,120,236 | 37,946 (1.8) | 1,069,035 | 16,470 (1.5) | 488,147 | 6,203 (1.3) | 7,415 | 45 (1.0) |
Figure 1. \(\log_{10}(M_{500c}/M_\odot)\) vs redshift. 100,000 clusters with the highest flux in the full sky light-cone extracted from MDPL2 simulation (lila transparent crosses). Including fainter clusters would fill the lower-right part of the diagram. The sample used to create the covariance matrix are shown with the symbols specified in the legend (XXL: red circle, SPT: black cross, HI\textsc{flugcs}: green square, X-COP: blue star).

nested scheme) where the colatitude is \(90° - \text{Dec}\) and the longitude is R.A. (Górski et al. 2005). We project the peculiar velocity along the line-of-sight as follows:

\[
v_\parallel = (v_x x + v_y y + v_z z)/d_C \text{ [km s}^{-1}]\text{,}
\]

and obtain the comoving distance with peculiar motion projected as:

\[
d_C^2 = d_C + v_\parallel / H(z_R) \text{ [Mpc].}
\]

From \(d_C^2\) and Eq. 1, we infer the observed redshift: ‘redshift \(S\)’.

2.3. Foreground absorption

Given each (sub) halo’s position on the sky, we retrieve the Milky Way HI column density from HI4PI Collaboration et al. (2016) and the E(B-V) extinction value taken from Planck Collaboration et al. (2014a) using the Dustmap software (Green 2018). These values are later used to attenuate fluxes.

3. X-RAY CLUSTER EMPIRICAL MODEL

The goal here is to create a model that represents the overall population of clusters and groups to be detected by \(\text{eROSITA}\). A description of its selection function is available in Clerc et al. (2018). The \(\text{eROSITA}\) cluster sample is expected to be complete for masses greater than \(10^{14}(7 \times 10^{13})M_\odot\) for redshifts \(z < 1\) (\(z < 0.3\)). Thus, we consider the (sub) halos with a mass \(M_{500c} > 5 \times 10^{13}M_\odot\) (\(\log_{10}(5 \times 10^{13}) = 13.7\)).

There exists a number of approaches to paint galaxy clusters in the X-ray, for example, phenomenological modelling of cluster pressure and temperature profiles with dependence on dynamical state (e.g. Zandanel et al. 2018), or semi-analytical, physically informed relations (e.g. Shaw et al. 2010; Flender et al. 2017), or machine learning algorithms trained on hydro-dynamical simulations (e.g. Cui et al. 2018). The model created here could be regarded as an upgrade of the first method.

3.1. Surface brightness profiles
To create a consistent library of profiles from the initial heterogeneous data set, we scale each profile according to the self-similar mass and redshift evolution (Neumann & Arnaud 1999; Arnaud et al. 2002) and interpolate them onto a common grid of 20 logarithmic-spaced points in the radial range of $R = [0.02 - 4] R_{500c}$.

When the outskirts of the detected clusters are not covered by the X-ray data, we extrapolate their profile with a power-law. The slope of the profile is estimated from the outermost 3 data points, and the emissivity beyond the maximum detection radius is assumed to follow a power law smoothly connected to the outermost data point. In this analysis, luminosity and flux are estimated by integrating the profiles up to $R_{500c}$, while events are simulated up to $2 R_{500c}$. In observations, the cluster surface brightness is dominated by that of the background at $R > R_{500c}$. So, even if the profile extrapolation method adopted is not the most accurate, it will hardly be visible (blended with noise) in the event maps for this simulation setup. Thus, these mock clusters may not be the best for studying cluster outskirts.

Within $R_{500c}$, we investigate possible biases. We create four mean vector and covariance matrices (see next section) with extrapolation slopes biased by $+20$, $-20$, $+40$, and $-40\%$. We create full sky mock catalogues, so the comparison is not limited by statistics. For $M_{500c} > 7 \times 10^{13} M_\odot$, we measure a relative change in the mean of the mass–luminosity ($M_{500c} - L_{500c}$) scaling relation smaller than $\pm 0.5\%$. The effect of the extrapolation is negligible and thus we find the method robust enough for our purpose.

The emission measurement (or emissivity) along the line-of-sight as a function of radius, $EM(r) = \int n_e n_{\perp} dl$, is defined as in Neumann & Arnaud (1999) and Arnaud et al. (2002):

$$EM(r) \ [cm^{-6} \ Mpc] = \frac{4\pi(1+z)^4 S(r/d_A(z))}{\epsilon(T,z)}.$$

The unit [cm$^{-6}$ Mpc] is used for convenience. $EM(r)$ is proportional to the ratio of the X-ray surface brightness profile, $S(\theta) \ [ct \ s^{-1} \ arcmin^{-2}]$, and the emissivity of the instrument, $\epsilon(T,z) \ [ct \ s^{-1} \ cm^2]$. $\epsilon$ is the integral over the energy band (depending on the survey considered) of the detector effective area [cm$^2$], times an inter-stellar absorption term, times the emissivity of a plasma of temperature $T$, heavy element abundance $A$, and redshift $z$. 

3.1.1. Emissivity profiles

To simulate realistic surface brightness profiles, we start from a collection of measured profiles from several existing samples spanning a wide range in mass and redshift, as shown in Fig. 1. The considered samples are XMM-XXL (Pierre et al. 2016; Adami et al. 2018), HIFLUGCS (Reiprich & Böhringer 2002), X-COP (Eckert et al. 2019) and SPT-Chandra (Sanders et al. 2018). XXL is a serendipitous X-ray survey over an area of 50 square degrees. The detected clusters span a redshift range of $z = (0.05 - 1.1)$ and mainly populate the galaxy group and low-mass cluster regime, with a median mass of $\sim 10^{14} M_\odot$. As such, it is quite similar to the eROSITA survey. HIFLUGCS is a complete sample of the brightest clusters detected in the ROSAT all-sky survey and constitutes a good low-redshift anchor. X-COP was selected from the strongest Sunyaev-Zeldovich detections in the Planck survey and thus samples the massive end of the cluster population. The SPT sample was drawn from a sample of the 90 most massive SPT-SZ detections at $z > 0.3$, for which a homogeneous Chandra follow-up program was conducted (McDonald et al. 2013, 2014). Altogether, our sample comprises a set of 322 ($182 + 45 + 12 + 83$, XXL + HIFLUGCS + X-COP + SPT) profiles. For the details of the data analysis and profile extraction methods, we refer to Eckert et al. (2016, XMM-XXL), Käfer et al. (2019, HIFLUGCS sub-sample limited to $kT \geq 3$ keV clusters), Ghirardini et al. (2019a, X-COP) and Sanders et al. (2018, SPT). The mass and redshift ranges covered by these surveys encompass most of the parameter space that will be probed by eROSITA. The dataset will eventually be complemented with measurements from the eROSITA survey itself.

For each cluster, we additionally have measurements of their redshift, spectroscopic temperature, and mass. The XXL masses are based on a mass-temperature relation calibrated through Subaru/HSC weak lensing measurements of 136 XXL clusters (Umetsu et al. 2020). The X-COP masses are drawn from high-quality hydrostatic mass measurements from combined XMM-Newton and Planck data and are found to be consistent with weak-lensing measurements for a subset of systems (Ettori et al. 2019; Eckert et al. 2019). The HIFLUGCS masses are derived from the Planck $Y_{SZ} - M_{500c}$ relation (Planck Collaboration XXVII 2016) and are corrected for a fiducial hydrostatic bias of 20% for consistency with the other datasets. The SPT masses are determined using the best-fit scaling relation (4 parameters) for a flat $\Lambda$CDM cosmology with $\Omega_m = 0.3$, $h = 0.7$ and $\sigma_8 = 0.8$, see Bocquet et al. (2019) for more details. Fig. 1 shows how the calibration data set compares to the simulated lightcone in the mass redshift plane. We find that the set of clusters used for calibration covers the relevant parameter space, except for the low mass and low redshift clusters and groups, see discussion at the end of this Section.
The isothermal surface brightness profile, \( \rho(x) \), is obtained by:

\[
\rho(x) = \frac{3M_{500c}}{4\pi500\rho_c(z)} \left[ \text{cm} \right].
\]

The iso-thermal surface brightness profile, \( \rho(x) \), is obtained by:

\[
\rho(x) = EM(x)\sqrt{\frac{kT}{10\text{keV}}}E^3(z)\Lambda(T) \left[ \text{erg s}^{-1} \text{cm}^{-2} \right],
\]

where \( x = r/r_{500c} \) and \( \Lambda(T) \) is the cooling function \( \text{cm}^3 \text{erg s}^{-1} \) in the band 0.5-2 keV (Sutherland & Dopita 1993). The luminosity at \( x \) is obtained by the cumulative sum (equivalent to integrating over a cylinder):

\[
L_X(r) = \int_{r=0}^{r} \left( \rho(x)r^2_{500c}2\pi x dx \right) \left[ \text{erg s}^{-1} \right].
\]

For each simulated cluster, we record its redshift, temperature, mass, X-ray luminosity in the band 0.5-2 keV and the surface brightness profile in an image format compatible with the SIXTE simulator (Dauser et al. 2019), see Sec. 3.4.

3.1.2. Covariance matrix

We construct the mean vector of the data: redshift, temperature, mass \( M_{500c} \), and the emissivity as a function radius (20 points, normalized by \( R_{500c} \)) (23 elements). We then construct a covariance matrix between these quantities, see Fig. 2 (the resulting matrix has a shape 23 x 23, the Pearson’s cross-correlation coefficient is also shown). The positive correlation between mass and emissivity as a function of radius is related to the deviation from the self similar model caused by the gas fraction. For the same reason, but to a lesser extent, \( kT \) correlates with emissivity as a function of radius. The anti-correlation between redshift and mass is a selection effect, see Fig 1. The covariance between the emissivity values as a function of radius are expected at small radius, see Fig 4. At large radius, the covariance might be over-estimated. The correlation between \( kT \) and mass is illustrated in Fig. 7, that shows the scaling relation.

Through a Gaussian multivariate random process (using the log of the mean vector and the covariance matrix i.e. values are log-normally distributed), we generate a set of simulated clusters characterized by their X-ray emissivity profile, redshift, mass and temperature. Fig. 3 shows an example set of simulated cluster profiles. The simulated profiles are color-coded by the temperature. We see that our procedure captures the known mass dependence of the X-ray emissivity profiles. Group-scale objects show a lower central emissivity and a flatter profile than cluster-scale systems (Eckert et al. 2016). We predict an intrinsic scatter of the emissivity as a function of radius. The anti-correlation between redshift and mass is a selection effect, see Fig 1. The covariance between the emissivity values as a function of radius are expected at small radius, see Fig 4. At large radius, the covariance might be over-estimated. The correlation between \( kT \) and mass is illustrated in Fig. 7, that shows the scaling relation.

3.1.3. Derived quantities

From the redshift and the halo mass, we deduce for each cluster the radius corresponding to 500 times the critical density, \( r_{500c} \), as follows:

\[
r_{500c} = \left( \frac{3M_{500c}}{4\pi500\rho_c(z)} \right)^{1/3} \left[ \text{cm} \right].
\]

The luminosity function and in the logN-logS. In Section 3.2 Linking simulated profiles to dark matter haloes

In the dark matter only N-body simulation, each halo (sub-halo) is characterized by its redshift and mass \( M_{DM_{halo}} \). The dynamical state of a halo is traced by \( X_{off} \), defined as the distance between the halo highest density peak and its centre of mass divided by its virial radius. Klypin et al. (2016) classified halos as relaxed or disturbed using a combination of Spin and \( X_{off} \). Typically, relaxed haloes have \( X_{off} \in [0.01, 0.07] \) and disturbed haloes have \( X_{off} \in [0.07, 0.3] \). Henson et al. (2017) discussed possible criteria to describe the relaxation state of the cluster: e.g. substructure fraction, spin parameter, etc and retain \( X_{off} \) as the cleanest (and most direct) estimator.

To ensure the availability of simulated profiles from each light-cone shell, we generate 100 times more profiles from the covariance matrix than the number of clusters present in a given shell. In this way, we densely sample the parameter space, avoiding using the same profile twice. We assign to each halo (and sub-halo) a simulated cluster using a nearest neighbour procedure in three dimensions. The near-neighbour search for redshift is direct. For the halo mass, we introduce a mass bias parameter, \( b_M \), and look for the nearest neighbour between \( M_{DM_{halo}} \) and \( b_M M_{DM_{halo}} \). This is the only parameter of this model, set to 1 (unless stated otherwise). This parameter induces significant differences when predicting the density of clusters as a function of their brightness i.e. in the luminosity function and in the logN-logS. In Section...
4.2, we show simulation results for $b_{\text{off}} \in [0.6, 1.1]$, a range similar to that used by Zandanel et al. (2018).

Since the shape of the surface brightness profile is related to the relaxation state of the halo, we preferentially assign flatter profiles to the most disturbed systems, according to their $X_{\text{off}}$ value. We relate the negative logarithm of the central emissivity to $X_{\text{off}}$.

$$-\log_{10}(\text{EM}(r = 0)) \propto X_{\text{off}}. \quad (13)$$

A low $X_{\text{off}}$ (a high EM(0)) corresponds to a cool-core (relaxed) cluster. A high $X_{\text{off}}$ (a low EM(0)) corresponds to a non cool-core (un-relaxed) cluster. $X_{\text{off}}$ and EM(0) have quite different face values, so for the nearest-neighbour match to work, we normalize them. We compute the cumulative distribution function for $X_{\text{off}}$ and EM(0). The cumulative distribution function relates a value (of $X_{\text{off}}$ and EM(0)) to its percentile in its parent distribution function. The nearest neighbour is then chosen with the percentiles of both quantities. Note that relating EM(0) to $X_{\text{off}}$ does not skew or change the intrinsic population of simulated clusters i.e. $X_{\text{off}}$ is not used to preferentially select certain types of clusters over others. Instead, among the simulated clusters (a fair sample), it relates EM(0) to $X_{\text{off}}$ after ordering them. In this manner, the cool-core bias is physically related to the halo properties and possibly its surrounding large-scale structure (Eckert et al. 2011). We force a relation to a parameter of the dark matter halo to possibly help reconstructing the selection function in a continuous fashion (e.g. Käfer et al. 2019, 2020; Seppi et al. 2020). Note that the true distribution of EM(0) is not well known. Nevertheless, Rossetti et al. (2017, Fig. 2) shows a distribution of the concentration of Planck clusters, which is similar to a log-normal distribution. Nurgaliev et al. (2017) reached similar conclusions based on the SPT sample. In the future, with more constraints from observations, we will be able to better model this distribution.

As the model sample is finite, the precision of the nearest neighbour match will depend on its total size. The fractional differences in masses are smaller than 2% for $M_{500c} > 5 \times 10^{13} M_\odot$. The fractional differences in redshift are smaller than 2%. At low redshift and/or at low masses, we see the limitations of the sample we draw from. Indeed, the original sample from which simulated clusters are drawn does not contain enough clusters to smoothly cover the full mass-redshift parameter space of interest. Modeling more clusters at low masses and low redshift does not shrink the differences significantly. The low mass and low redshift regime will benefit from eROSITA observations and should improve over the coming years, see Fig 1. We find the dataset used may need more lower mass, low redshift clusters (e.g. eHFLUGS Reiprich 2017).

At the end of this procedure, each DM halo in the light-cone is linked to a surface brightness profile, a temperature and an X-ray luminosity.

3.2.1. Limitations

We verify that the mean of EM(0) in bins of halos mass does not depend on halo mass. A linear fit gives a correlation coefficient of $-0.013$, indicating no significant dependence on the halo mass. We find that the mean of EM(0) in bins of temperature depends slightly on the temperature. A linear fit gives $\text{EM}(0) = -0.7 \log_{10}(kT) + 5.2$ with a correlation coefficient of 0.17.

Although at times symmetric (low $X_{\text{off}}$) clusters have a high central surface-brightness (high EM(0)), it is not always the case that a high central surface brightness indicates symmetry for a cluster. For example Mantz et al. (2016, Fig. 2) showed that at a given over-density, there is intrinsic scatter in the surface-brightness. In this approach, the scatter is present by construction, but its correlation (via Eq. 13) to the $X_{\text{off}}$ parameter is tighter than it should be. With upcoming eROSITA observations, we will hopefully constraint the relation and its scatter (currently assumed to be zero) between relaxation state and central emissivity, to improve on Eq. 13 (Seppi et al. 2020).

3.3. Flux

We deduce observed X-ray luminosities and fluxes in the soft X-ray band 0.5-2 keV (like eROSITA) using the xspec software combined with an APEC spectrum extracted by a TBABS model. The metallicity is assumed to be 0.3 solar (Asplund et al. 2009). K-correction and attenuation are pre-computed on a large grid of redshifts, temperatures and hydrogen column density ($n_H$), then interpolated and applied to the mock clusters. Fig. 5 shows the conversion from luminosity to flux and its dependence on redshift and temperature. Finally, fluxes are extincted using the local $n_H$ value and we obtain an 'observed' mock catalogue.

3.4. Images

Oguri et al. (2010) measured, with the weak lensing technique, the ellipticity of dark matter haloes hosting galaxy clusters. This measurement is in excellent agreement with the standard collisionless cold dark matter model. Umetsu et al. (2018) found that the ellipticity of the X-ray emission by the ICM in clusters is correlated to the dark matter halo ellipticity, suggesting a tight alignment between the intracluster gas and dark matter. So, having elliptical 2D surface brightness is key for upcoming studies characterizing their detection and selection function. For each model cluster, using the ellipticity of the dark matter halo and the model profile, we create an image (SIMPUT format, Dauser et al. 2019).

We use the axis ratio fitted on the 3D dark matter halo spheroid ($b_{\text{to_a_500c}}$) as the axis ratio on the sky. Because $X_{\text{off}}$ and the axis ratio are correlated (e.g. Lau et al. 2020), the profile shape EM(0) becomes correlated to the ellipticity of the X-ray surface brightness contour. Note that the ellipticity in the simulation is determined on the mass density and not on the potential, which is traced by the X-ray emitting gas (e.g. Lau et al. 2011). Nevertheless, the Quartiles (Q1, 2, 3) of the ($b_{\text{to_a_500c}}$) axis ratio distribution: 0.45, 0.55 (median), 0.65, are in good agreement with observations (e.g. Shin et al. 2018). So even if the physical link is not direct, the face values are following a reasonable distribution. We find that using an axis ratio estimated on a larger aperture (e.g. 200c) leads to a set of clusters that would be too elliptical.

We defer to future work to extend this method to create weak lensing observables (Giocoli et al. 2016; Umetsu 2020).

3 http://www.atomdb.org
3.5. Sub-halos

We also model sub-halos to reproduce substructure in the largest halos. Massive sub-halos are rare, only a few percent of most massive halos contain a very massive substructure. Table 2 gives the number of halos and sub-halos for four mass thresholds: \(M_{\text{vir}} > 5 \times 10^{13}, 7 \times 10^{13}, 5 \times 10^{14}, 5 \times 10^{14} M_\odot/h\). The fraction of sub-halo depends on mass resolution i.e. on the ability to resolve a substructure within a very dense environment. The fraction of sub-halo with \(M_{\text{vir}} > 7 \times 10^{13} M_\odot/h\) varies from 1.4 (HMDPL, lowest resolution) to 3 (SMDPL, highest resolution) per cent.

3.6. Other models

We compare our empirical model with a phenomenological and a semi-analytic model in the literature.

3.6.1. Phenomenological model by Zandanel et al. (2018)

Zandanel et al. (2018, hereafter Za18) developed a phenomenological model, starting from the Planck pressure profiles (Planck Collaboration et al. 2013, 2016) and Chandra temperature profile models adapted from Hudson et al. (2010) and including four different dynamical states for clusters, to implement ICM properties onto dark-matter-only halos. They use the MultiDark (Klypin et al. 2016) simulation suite to create a set of eROSITA mock light cones, available here\(^4\) (Klypin et al. 2017). This phenomenological model reproduces well state-of-the-art X-ray and SZ observations of galaxy clusters.

3.6.2. Semi-analytic model by Shaw et al. (2010)

The semi-analytical model of the ICM presented in Shaw et al. (2010, hereafter Sh10), which was based on (Ostriker et al. 2005). It assumes that for each halo, DM density profile follows the NFW profile which is uniquely determined by halo mass and halo concentration. For this paper we use the mass-concentration relation from Diemer & Kravtsov (2014). The gas pressure is in hydrostatic equilibrium with the NFW gravitational potential. The gas density and the gas pressure is related by a polytropic equation of state with polytropic index of \(\Gamma = 1.2\) outside the cluster core \(\geq 0.2 R_{500c}\), as suggested by observations (e.g. Ghirardini et al. 2019b). The gas density and total pressure profiles are determined by solving the equations of energy and momentum conservation. Star formation, feedback from supernovae and active galactic nuclei are included in the model. In Shaw et al. (2010), the model is extended with the inclusion of non-thermal pressure. The model is further extended with the inclusion of cool-core, where we model a different polytropic index in the cluster core Flender et al. (2017), and gas density clumping in Shirasaki et al. (2020). The model parameters used in the current paper was calibrated with density profiles measured with Chandra observations of galaxy clusters detected by the South Pole Telescope (SPT) and with \(M_{\text{gas}} - M\) relations from Chandra and XMM-Newton observations. We refer the reader to Flender et al. (2017) for details on the calibration. Note that clumping is not included in the current calibration. Also note that the current model does not include intrinsic scatter in the ICM profiles, meaning any two halos with the same mass and redshift will have the same ICM profiles. Intrinsic scatter in the ICM profiles due to the halo formation histories and halo shapes will be implemented in the future.

4. RESULTS

We obtain scaling relations and their scatter from the generated mock catalogues, these are discussed in Section 4.1. In Section 4.2 we present the number of model clusters per square degrees as a function of flux. Finally, in Section 4.3 we discuss the distribution of model photons as they will be observed by eROSITA.

4.1. Scaling relations

We measure a set of scaling relations between the halo properties and X-ray properties from the generated model catalogues. We consider an aperture corresponding to the 500c over-density and a soft band in the X-ray: 0.5-2 keV. The scaling relations obtained are close (but not equal) to the self-similar model, used in the construction of the profiles. Overall, the scaling relations obtained are in good agreement with the literature (Lovisari et al. 2015; Manz et al. 2016; Schellenberger & Reiprich 2017; Adami et al. 2018; Bulbul et al. 2019; Lovisari et al. 2020; Sereno et al. 2020; Umetsu et al. 2020). To obtain XMM-like temperatures and carry out these comparisons, we correct temperatures from the literature following Schellenberger et al. (2015). We also convert literature luminosities into the 0.5-2 keV band using the APEC model. Additionally, the intrinsic scatter around each model scaling relation is in good agreement with measurements from the literature (Lovisari et al. 2015; Bulbul et al. 2019; Lovisari et al. 2020). We also compare our results to those from alternative numerical methods to mock clusters (Sh10, Za18 Shaw et al. 2010; Zandanel et al. 2018). The methods are overall in good agreement, discrepancies are detailed below.

We discuss in detail the results for the mass-luminosity (§ 4.1.1), the mass-temperature (§ 4.1.2) and the luminosity-temperature (§ 4.1.3) relations.

4.1.1. Mass-luminosity relation

In the self-similar model, the luminosity is proportional to the mass as \(L_X \propto E^2(z)M_{500c}^{4/3}\). The scaling relation obtained here is close, but not consistent with, the self similar model as it has a slope of 1.253 ± 0.031. This is a consequence of the profile construction methodology. Fig. 6 top panel shows the scaling relation between mass and luminosity of the MDPL2 light-cone. We compare with the distribution of measurements (Lovisari et al. 2015; Manz et al. 2016; Schellenberger & Reiprich 2017; Adami et al. 2018; Bulbul et al. 2019; Lovisari et al. 2020; Sereno et al. 2020) and find a good agreement. It seems the model relation has a somewhat lower slope and normalization that could be suggested by the data. For XXL, the uncertainty on the mass in the data is of order of 15% and on the X-ray luminosity of order of 10% (represented by the black cross in top panel of Fig. 6). All other surveys have similar uncertainty on the mass and brighter clusters have smaller uncertainties on the X-ray luminosity (a few percents for the brightest). Each survey is subject to a different selection function (flux limit, redshift reach, Malmquist bias etc), so the

\(^4\) http://skiesanduniverses.org/
density of data points is not necessarily an accurate indicator of the location of the scaling relation in the figure. The model cluster sample being complete, we expect the model scaling relation to be lower than the data points. Indeed, the Malmquist bias results in an overestimation of average observed luminosity for given mass.

We predict that the intrinsic scatter around the mean log10(LX) follows a normal distribution with σ = 0.21 with no significant mass dependence, as shown in the middle panel of Fig. 6. In the literature the intrinsic scatter has been measured at 0.24, 0.25, 0.2 (Lovisari et al. 2015; Bulbul et al. 2019; Lovisari et al. 2020).

Mean model scaling relations from Sh10, Za18 are compared to MDPL2 for two values of the b_M: 0.6 and 1.0, see Fig. 6, bottom panel. They are in good agreement. The slope we obtain is shallower than that of Za18.

4.1.2. Mass-temperature relation

In the self-similar model, the temperature scales with mass as kT ∝ E^{2/3}(z)M^{2/3}500c. With a slope of 0.644 ± 0.026, we obtain a scaling relation close to the self similar expectation. Fig. 7 shows the scaling relation between mass and temperature. As for the previous scaling relations, we compare with the distribution of measurements in the literature and generally find a good agreement. The difference with the Sh10 model lies in the different non-thermal pressure fraction. The Sh10 model was calibrated on density profiles and the gas mass – mass relation. The non-thermal pressure fraction was taken from simulations (Nelson et al. 2014), which is found to be relatively high compared to the X-COP measurements (e.g. Eckert et al. 2019). A change of the non-thermal pressure fraction in the Sh10 model shifts its scaling relations towards that obtained with the method described in this article.

The predicted intrinsic scatter on the log10(kT) follows a normal distribution with σ = 0.07. In the literature the intrinsic scatter has been measured at 0.035, 0.18 (Bulbul et al. 2019; Lovisari et al. 2020).

4.1.3. Luminosity-temperature relation

In the self-similar model, the luminosity scales with temperature as LX ∝ E^{2/3}(z)(kT)^2. With a slope of 1.659 ± 0.098, we obtain a scaling relation close to (but not consistent with) that. By construction, the scaling relation predicted is close to that of Giles et al. (2016); Lieu et al. (2016), based on the XXL data. Fig. 8 shows the scaling relation between temperature and luminosity. We compare with the distribution of measurements from (Lovisari et al. 2015; Mantz et al. 2016; Adami et al. 2018; Bulbul et al. 2019; Lovisari et al. 2020) and find a good agreement. The turn over of the relation at low temperature is due to the Malmquist bias in the XXL data used, future inclusion of lower mass objects in the covariance matrix will enable to correct this effect and extend the model to lower masses.

The predicted intrinsic scatter on the log10(LX) follows a normal distribution with σ = 0.25. In the literature the intrinsic scatter has been measured at 0.2 (Lovisari et al. 2015) and 0.2-0.28 (Sereno et al. 2019).

4.2. Number density
Figure 7. Same as Fig. 6. **Top.** Model scaling relation between temperature and mass, the agreement with observations is good. **Middle.** Intrinsic scatter as a function of mass. The mean value of 0.07 is in the middle of values obtained in previous studies: 0.035 and 0.18. **Bottom.** Comparison of model scaling relations.

Figure 8. Same as Fig. 6. **Top.** Model scaling relation between X-ray luminosity and temperature, the agreement with observations is good. **Middle.** Predicted intrinsic scatter as a function of temperature. It is higher than measurements. **Bottom.** Comparison of model scaling relations.
The cosmological parameter used in the N-body simulations (Ω_m ~ 0.31) are known to produce ‘too many’ clusters (e.g., Zandanel et al. 2018), we indeed see that the predicted log N – log S lies at the upper boundary of observations. The b_M parameter has quite an impact here. The higher the b_M, the more bright clusters will be present in the mock, thus their density will increase. The higher the \( m > 500 \), the more clusters will be present in the mock. 

### 4.3. Simulated eROSITA photons

We use the SIXTE software package for X-ray telescope observation simulations (Dauzer et al. 2019)\(^5\). It produces simulated event files for mission studies. We create a full sky event list using the observational strategy from eROSITA. With a healpix pixelisation nested scheme, we split the sky into 768 equal area pixels of approximately 53 deg\(^2\) each (Górski et al. 2005). The co-latitude is 90 – Dec. and the longitude is R.A.

Along with clusters, we model a population of AGN. We release the set of AGN and cluster photons. The AGN model is described in (Comparat et al. 2019). A notable difference in the AGN model w.r.t. its previous incarnation is that they are also assigned to dark matter sub-halos. Figs. 10 and 11 show the events related to the clusters and AGNs on a couple 2° × 2° fields. The diversity of the cluster population is well represented. Understanding the co-evolution of AGN and clusters seems to be key to turn these event lists back into unbiased catalogues of the large scale structure.

Ongoing eROSITA studies detailing the procedure of detection and catalogue creation intensively use these simulations to understand the trade-off between purity and completeness. These also enable an in-depth discussion of the line-of-sight confusion between the various sources as a function of exposure time.

To enable the study of systematic errors throughout the data flow, we aggregate photon statistics (number of photons in a set of apertures emitted by the source itself or other nearby sources) in the simulated catalogues.

### 5. SUMMARY AND OUTLOOK

We presented a novel empirical approach to model X-ray profiles of cluster and group-size halos. We used the MultiDark (HMDPL, MDPL2, SMDPL Klypin et al. 2016) and the UNIT (UNIT1, UNIT1 Chuang et al. 2019) dark matter only simulations to create light-cones spanning the full sky. We have created a new method to simulate the X-ray light emitted by the hot gas in galaxy clusters. This model computes the emissivity and image for the clusters in the light-cones.

We find that the mean scatter around the profile as a function of scale is in agreement with the latest measurements from Ghirardini et al. (2019a). The model scaling relations, their scatter and the number density of clusters are in good agreement with observations and other models. We obtain a scatter of 0.21 (0.07, 0.25) for the X-ray luminosity – mass (temperature – mass, temperature – luminosity) model scaling relations. Finally, using the AGN model from Comparat et al. (2019), we predict the full sky distribution of photons to be observed by eROSITA.

This is the first building block of an end to end validation of the eROSITA work flow towards measuring the large scale structure traced by X-rays. This will enable accurate estimations of various systematic uncertainties associated with cluster astrophysics (e.g., cool cores, non-thermal pressure, gas clumping) and those due to observation and analysis methods (e.g., projection effects). This model data presented here will support upcoming studies of the eROSITA source detection process and selection function computation.

We will extend the method to lower mass halos by extracting X-ray profiles of groups from the deep Chandra on CDFS (Finoguenov et al. 2015) and COSMOS (Gozaliasl et al. 2019). We will extend the discussion on scaling relations to core-excised quantities and count rates in future studies.

The combined X-ray catalogues of AGN and galaxy clusters and groups produced from our model will also open up new avenues in the study of co-evolution between AGN and galaxy clusters and groups with eROSITA and other future X-ray surveys by cross-correlating AGN and cluster signals. This will lead to better understanding of the role of AGN feedback in physics of intra-cluster and intra-group medium. The modelling pipeline presented in the current paper can also applied to other wavelengths, such as the microwave in modeling the Compton-y signals from galaxy clusters and groups, and the modeling of gravitational lensing signals, using the same N-body lightcone. This will be useful and essential for generating predictions for multiwavelength observations for constraining outstanding issues in cluster astrophysics, such as non-thermal pressure and hydrostatic mass bias, and clumping. This will help eROSITA

\(^5\) http://www.sternwarte.uni-erlangen.de/research/sixte/
Figure 10. Declination vs. Right Ascension of the photons observed (black dots) in the 0.2-2.3 keV band after 4 years of eROSITA observations on the 2x2 degrees center of the field 0 (out of 768). A fraction of the sources present in the catalogue are identified by circles. It shows the clusters that emitted more than 8 counts (red), the AGN that emitted more than 6 counts (blue), and the stars (green).

to maximize its scientific returns through synergies with other ongoing microwave and optical surveys.

ACKNOWLEDGEMENTS

GY acknowledges financial support by MI- CIU/FEDER under research grant PGC2018-094975-C21. DN acknowledges the Max-Planck-Institut für Astrophysik for hospitality when this work was initiated.

The authors gratefully acknowledge the Gauss Centre for Supercomputing and the Partnership for Advanced Supercomputing in Europe (PRACE) for funding the MultiDark simulation project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre, Germany.

The UNIT simulations were run at the MareNostrum Supercomputer hosted by the Barcelona Supercomputing Center (BSC), Spain, thanks to the PRACE project grant number 2016163937.

The CosmoSim database is a service of the Leibniz-Institute for Astrophysics Potsdam (AIP). The MultiDark database was developed in cooperation with the Spanish MultiDark Consolider Project CSD2009-00064.

This project made use of gawk\(^6\), python\(^7\), astropy\(^8\) (Astropy Collaboration et al. 2013, 2018), topcat/stilts\(^8\) Taylor (2006), slurm\(^9\).

This research has made use of the SIXTE software package (Dauser et al. 2019) provided by ECAP/Remels

---

\(^6\)https://www.gnu.org/software/gawk/
\(^7\)https://www.python.org
\(^8\)http://www.star.bris.ac.uk/~mbt/stilts/
\(^9\)https://slurm.schedmd.com/publications.html
Figure 11. Same as Fig. 10 for field 129.

observatory (https://github.com/thdauser/sixte).

REFERENCES

Adami C., et al., 2018, A&A, 620, A5
Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Arnaud M., Aghanim N., Neumann D. M., 2002, A&A, 389, 1
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Astropy Collaboration et al., 2013, A&A, 558, A33
Astropy Collaboration et al., 2018, AJ, 156, 123
Balaguera-Antolínez A., Sánchez A. G., Böhringer H., Collins C., 2012, MNRAS, 425, 2244
Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017a, MNRAS, 465, 212
Barnes D. J., et al., 2017b, MNRAS, 471, 1088
Barnes D. J., et al., 2018, MNRAS, 481, 1809
Barnes D. J., Vogelsberger M., Pearce F. A., Pop A.-R., Kannan R., Cao K., Kay S. T., Hernquist L., 2020, arXiv e-prints, p. arXiv:2001.11508
Behroozi P., Wechsler R., Wu H.-Y., 2013, ApJ, 762, 109
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Biffi V., Dolag K., Merloni A., 2018, MNRAS, 487, 159
Bocquet S., et al., 2019, ApJ, 878, 55
Böhringer H., Chon G., Retzlaff J., Trümper J., Meisenheimer K., Schartel N., 2017, AJ, 153, 220
Bulbul E., et al., 2019, ApJ, 871, 50
Cao K., Barnes D. J., Vogelsberger M., 2020, arXiv e-prints, p. arXiv:2006.10752
Chuang C.-H., et al., 2019, MNRAS, 487, 48
Clerc N., et al., 2018, A&A, 617, A92
Comparat J., Prada F., Yepes G., Klypin A., 2017, MNRAS, 469, 4157
Comparat J., et al., 2019, MNRAS, 1335
In this work we upgrade the model from Comparat et al. (2019). We use the Behroozi et al. (Universe Machine 2019) empirical galaxy model to compute and attach the galaxy quantities, such as stellar mass, star formation rate, for each dark matter halo. Note that there exists more galaxy models run on the MDPL2 simulation (Stoppacher et al. 2019). Each model has advantages and disadvantages. The main reason for choosing the Universe Machine
is that this model is constrained to reproduce the fraction of quenched galaxies as a function of stellar mass, which is an important feature for future studies of galaxies in clusters.

For AGNs, we follow the works by Georgakakis et al. (2018); Comparat et al. (2019) without a prior on the fraction of sub-halos (i.e. we treat all sub-halos in the simulation as if they were haloes). We then consider sub-halos as possible AGN hosts in the simulation, when computing the duty cycle. As the periodic boundary condition is respected (while replicating the simulation boxes) these mocks are suited for clustering studies on the full sky.

DATA AND SOFTWARE

The light-cones and event files produced are made available in two locations: MPE/MPG, Germany\textsuperscript{10} and skies and universes, Granada, Spain\textsuperscript{11} (Klypin et al. 2017). We make available the catalogues and event lists produced based on the MDPL2 simulation. The largest part of the pipeline developed is public and divided in two repository. The first to create the light cones and the mock catalogues\textsuperscript{12} and the second to draw cluster model profiles\textsuperscript{13}.