

Cross-correlation between soft X-rays and galaxies

A new benchmark for galaxy evolution models

Johan Comparat^{1,*}, Andrea Merloni¹, Gabriele Ponti^{2,3,1}, Soumya Shreeram¹, Yi Zhang¹,
Thomas H. Reiprich⁴, Ang Liu^{1,5}, Riccardo Seppi⁶, Xiaoyuan Zhang¹, Nicolas Clerc⁷, Andrina Nicola⁴,
Kirpal Nandra¹, Mara Salvato¹, and Nicola Malavasi¹

¹ Max-Planck-Institut für extraterrestrische Physik (MPE), Gießenbachstraße 1, D-85748 Garching bei München, Germany

² INAF-Osservatorio Astronomico di Brera, Via E. Bianchi 46, I-23807 Merate (LC), Italy

³ Como Lake Center for Astrophysics (CLAP), DiSAT, Università degli Studi dell'Insubria, Via Valleggio 11, I-22100 Como, Italy

⁴ Argelander-Institut für Astronomie (AIfA), Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany

⁵ Institute for Frontiers in Astronomy and Astrophysics, Beijing Normal University, Beijing 102206, China

⁶ Department of Astronomy, University of Geneva, Ch. d'Ecogia 16, CH-1290 Versoix, Switzerland

⁷ IRAP, Université de Toulouse, CNRS, UPS, CNES, F-31028 Toulouse, France

Received 21 February 2025 / Accepted 25 March 2025

ABSTRACT

This paper presents the construction and validation of complete stellar mass-selected, volume-limited galaxy samples using the Legacy Survey (data release 10) galaxy catalogs, covering $\sim 16\,800\text{ deg}^2$ of extra-galactic sky and extending to redshifts of $z < 0.35$. We used companion mock catalogs to ensure a controlled galaxy selection. We measured the two-point correlation function of these galaxies with tiny statistical uncertainties at the percent level and systematic uncertainties up to 5%. We fitted a four-parameter halo occupation distribution (HOD) model to retrieve the population of host halos, yielding results on the stellar to halo mass relation that are consistent with the current models of galaxy formation and evolution. Using these complete galaxy samples, we measured and analyzed the cross-correlation between galaxies and all soft X-ray photons observed by SRG/eROSITA in the 0.5–2 keV band over $\sim 13\,000\text{ deg}^2$. The cross-correlation measurements have an unprecedented sub-percent statistical uncertainty and 5–10% systematic uncertainty. We introduced a novel extension to the halo model to interpret the cross-correlation, decomposing contributions from X-ray point sources, hot gas, satellites, and the two-halo term. The model offers a new comprehensive view of the relation between the complete 0.5–2 keV X-ray photon field and complete sets of galaxies at low redshift and their host halos. For low stellar mass thresholds ($\log M^*/M_\odot > 10$, 10.25, 10.5), we find that the point source emission dominates the cross-correlation at small separation ($r < 80\text{ kpc}$). Then, in the range of $80 < r < 2\text{ Mpc}$, the emission from large halos hosting satellite galaxies dominates. Finally, on scales beyond those considered here ($r > 2\text{ Mpc}$), the two-halo term becomes dominant. Interestingly, there is no scale at which the hot gas dominates. In the range ($20 < r < 200\text{ kpc}$), the hot gas contributes to more than 10% of the signal. Progressively, with the minimum stellar mass increasing, the hot gas emission increases. For the $\log M^*/M_\odot > 10.75$ sample, in the range 50–60 kpc, the three components contribute each the same surface brightness. For the $\log M^*/M_\odot > 11$ sample, the hot gas is the dominating emission source over the range of 30–200 kpc. Finally, for the $\log M^*/M_\odot > 11.25$ and (11.5) samples, the hot gas emission dominates over other components until 400 (700) kpc. We constrained the slope of the scaling relation between halo mass and X-ray luminosity (over three orders of magnitude in mass) at the 5% level, using the samples with the lowest mass threshold. We find a slope of $1.629^{+0.091}_{-0.089}$. Additional analyses explore the energy dependence of the cross-correlation and differences between red sequence and blue cloud galaxies, revealing sensitivity to galaxy quiescent fractions and opening avenues for a more complex, unified modeling of galaxies, active galactic nuclei (AGNs), and hot gas in the optical and X-rays.

Key words. galaxies: general – large-scale structure of Universe – X-rays: general

1. Introduction

The large-scale structure of the Universe is observable at different wavelengths, each related to different sources, objects, and physical processes. In this work, we measure and interpret the cross-correlation between galaxies selected in the optical in the legacy surveys' tenth data release (LS10 hereafter, Dey et al. 2019) and the set of X-ray photons gathered by SRG/eROSITA (Predehl et al. 2021; Sunyaev et al. 2021) over half of the sky. These two wavelength ranges provide complementary information about the cohabitation among galaxies, their surround-

ing circumgalactic medium, and nearby active galactic nuclei (AGNs).

We use the combination of four eROSITA full-sky scans (eRASS:4) accessible to the German eROSITA consortium (Western Galactic hemisphere, eROSITA-DE hereafter). These constitute our Universe's deepest and most homogeneous soft X-ray map. Thanks to its scanning strategy, the X-ray background is well covered over large areas (see e.g., Merloni et al. 2024; Locatelli et al. 2024; Zheng et al. 2024a; Yeung et al. 2024). As demonstrated by the spectral analysis of Ponti et al. (2023a,b), this set of events mainly traces (ordered by the total number of detected photons): the cosmic X-ray background (Hopkins et al. 2006), our Milky-Way's hot circumgalactic

* Corresponding author; comparat@mpe.mpg.de

medium (Locatelli et al. 2024; Zheng et al. 2024a,b), the local hot bubble (Yeung et al. 2024), solar wind charge exchange (Merloni et al. 2024), detected point sources made of AGNs and stars (Schwope et al. 2024a,b; Schneider et al. 2022; Freund et al. 2024; Waddell et al. 2024b,a), detected extended source made of hot gaseous halos (Liu et al. 2024; Bahar et al. 2024; Seppi et al. 2024; Bulbul et al. 2024; Kluge et al. 2024), and resolved star formation in near-by galaxies (Kyritsis et al. 2025).

In the past, the galaxies-X-ray cross-correlation was used to characterize the composition of the X-ray background (e.g., Soltan et al. 1997; Refregier et al. 1997; Newsam et al. 1999; Kaminsky et al. 2025). More recently, analyses stacking X-rays (in most cases detected by SRG/eROSITA) at the position of optically selected galaxies managed to grasp the relation between galaxies, their host halos, and their hot gas content (Anderson et al. 2015; Comparat et al. 2022; Zhang et al. 2024a,b; Popesso et al. 2024a). They leveraged enhanced galaxy catalogs to select central galaxies in their halos (Yang et al. 2005; Robotham et al. 2011; Tempel et al. 2016, 2018; Tinker 2021, 2022). The inference for an individual galaxy of its host halo mass and its position in the halo (central or satellite) is complex. In the Milky-Way mass range (halos of $\sim 3 \times 10^{12} M_{\odot}$), recent work on simulations (Popesso et al. 2024b) indicate that the scatter on the determination of the halo mass may be as large as one dex. Thus, when selecting a set of galaxies in a bin of inferred host halo mass, the obtained sample suffers from a high fraction of contaminants. Specifically, Popesso et al. (2024b) found that contamination fractions may be as large as 40%. To interpret these stacking experiments, we require solid knowledge of the selection function made on the inferred host halo mass, which may limit, or even systematically bias, our understanding of the observations. Furthermore, Comparat et al. (2022) and Zhang et al. (2024a) used observations and simulations to model the surface brightness profiles of contaminants: point source emission from AGNs and X-ray binaries (XRB) and satellite contamination (or misclassified centrals). They then subtracted these values to derive a measurement focused on the X-ray surface brightness profile of the hot gas and the scaling relation between, for example, X-ray luminosity and halo mass. This fruitful (albeit complex) approach is limited to areas of the sky where nearly complete spectroscopic surveys are available, which currently impedes using the complete eROSITA-DE extra-galactic sky. The study presented here generalizes and extends these earlier works, thanks to larger galaxy catalogs and X-ray data covering most of the eROSITA-DE extra-galactic sky ($13\,000 \text{ deg}^2$). It leverages on the proportionality between the signal-to-noise ratio (S/N) of the cross-correlation (or stacks) and sky coverage.

Comparat et al. (2022) and Zhang et al. (2024a) showed that both GAMA and SDSS, nearly complete spectroscopic surveys (Driver et al. 2022; Strauss et al. 2002), are well suited to the eROSITA depth for stacking analysis. The overlapping area between eROSITA-DE and the GAMA (SDSS) spectroscopic surveys is limited to ~ 200 (~ 4000) square degrees where eRASS exposure is shallowest, thereby limiting the total S/N in the stacks. Complete spectroscopic galaxy surveys overlapping with eROSITA-DE over a wider area will only be available in the 2030s after 4MOST completes its first tier of public surveys (de Jong 2011; de Jong et al. 2019; Finoguenov et al. 2019). In a slightly shorter timescale, DESI (DESI Collaboration 2016) will provide a complete galaxy spectroscopic survey with its Bright Galaxy Sample (BGS Hahn et al. 2023) but with a limited overlap with the eROSITA coverage (about 4000 deg^2 at positive

Declination). Therefore, since complete spectroscopic surveys do not yet cover the full eROSITA-DE area, here we follow a different and complementary methodology to understand how the low-redshift large-scale structure appears in X-rays.

Luckily, the quality and depth of the LS10 in the southern hemisphere (Dey et al. 2019) enables the inference of precise and unbiased photometric redshifts for galaxies brighter than $r < 19.5$ (Zhou et al. 2021). These redshift estimates allow for accurate galaxy clustering measurements, provided that there is access to projected summary statistics integrated up to $100 h^{-1} \text{ Mpc}$. To reach our goal of maximal S/N for a galaxy-X-rays cross-correlation experiment, we constructed galaxy samples following the DESI BGS prescription ($r < 19.5$; Hahn et al. 2023), resulting in a sample similar (but slightly brighter) to the GAMA one (Driver et al. 2022) and significantly deeper than the SDSS MAIN sample ($r < 17.77$; Strauss et al. 2002). We have covered the entire extra-galactic sky from eROSITA-DE with the LS10' photometric sample. We designed nine stellar mass-selected galaxy samples spanning redshifts between 0.05 and 0.35 and stellar masses between 10^9 and $10^{12} M_{\odot}$. We selected complete galaxy samples, close to volume-limited samples, regardless of the position of galaxies in their halo (central or satellite). This choice has guaranteed a simple galaxy selection function.

Here, we present and interpret the measurements of three summary statistics: (i) auto-correlation of stellar-mass selected galaxy samples over $\sim 16\,000 \text{ deg}^2$ ($w_p(r_p)$); (ii) cross-correlation between these galaxies and soft X-ray events over $\sim 13\,000 \text{ deg}^2$; and (iii) stacked X-ray surface brightness profiles around galaxies on the same area. We consider the complete photon field in the 0.5–2 keV band without masking detected sources, making the cross-correlation measurement reproducible. To interpret the auto-correlation, we use a standard halo occupation distribution model (HOD, e.g., Asgari et al. 2023, for a recent review). With the $w_p(r_p)$, we are able to unambiguously link galaxies to their host halos. To enhance our control and understanding of uncertainties related to the galaxy samples, we constructed mock catalogs using the Uchuu simulation augmented by the UniverseMachine empirical model of galaxy evolution (Ishiyama et al. 2021; Behroozi et al. 2013, 2019; Aung et al. 2023).

By cross-correlating low-redshift galaxies with soft X-ray event maps, we filtered out the signal produced by the Solar System and Milky Way foregrounds and traced these galaxies' nuclear activity (AGNs), X-ray binaries (XRB), and hot gaseous halos instead. The complexity of this approach lies in modeling the observed cross-correlation coming from multiple X-ray sources, together (and consistently) with the underlying host dark matter halos. To do so, we can augment the HOD model to predict X-ray events (from the hot gas and point sources) to extract the information from the cross-correlation and the stacks. This model directly interprets the correlation between the complete galaxy and X-ray fields. The strength of this approach is that the uncertainties in the measurement process due to the galaxy selection, the complexities arising from masks on the X-ray map, and the dependency on the stellar-to-halo mass relation will not impede the interpretation in any way (e.g., see discussion in Comparat et al. 2022; Zhang et al. 2024b; Popesso et al. 2024a; Marini et al. 2024).

We structure the paper as follows. We describe the observations we used in Sect. 2 (and the creation of galaxy mock catalogues in Appendix A). We detail the measurement of the galaxy auto-correlation, the galaxy X-ray cross-correlation, and the stacks in Sect. 3. We describe the HOD model in Sect. 4 and the results in Sect. 5. We assume a standard cosmology model

(flat LCDM) from [Planck Collaboration VI \(2020\)](#). Magnitudes are given in the AB system.

2. Observations

We note that this analysis requires X-ray data, which are described in Sect. 2.1. It also draws from the galaxy catalogs described in Sect. 2.2.

2.1. eROSITA soft X-ray events

Overall, eROSITA (extended ROentgen Survey with an Imaging Telescope Array) is a wide-field X-ray telescope on board the Russian-German “Spectrum-Roentgen-Gamma” (SRG) observatory ([Merloni et al. 2012](#); [Predehl et al. 2021](#); [Sunyaev et al. 2021](#)). With its seven identical Wolter-I mirror modules, eROSITA serves as a sensitive wide-field X-ray telescope capable of delivering deep, sharp images over vast sky areas in the energy range of 0.2–8 keV, with a maximum sensitivity in the 0.3–2.3 keV range. We use the first four all-sky scans of eROSITA (eRASS:4) reduced with the pipeline version c030 ([Brunner et al. 2022](#); [Merloni et al. 2024](#)); the data are organized and processed in “fields” of approximately 3.6×3.6 square degrees. In this version of the observations, the times at which solar flares occurred are masked. Then, we need to select events in the soft X-ray 0.5–2 keV range (no flag selection). We did not apply any further masking throughout the analysis and we considered the complete soft X-ray event field.

2.2. LS10 galaxies

We used the LS10 to select a flux-limited galaxy sample ([Dey et al. 2019](#)). LS10 provides source catalogs for Dec $< 32^\circ$, from which we selected low-redshift galaxies using the DESI BGS algorithm ([Hahn et al. 2023](#)) with a magnitude cut of $r_{AB} < 19.5$. Following their criterion, we separated stars from galaxies and selected galaxies with $13 < r_{AB} < 19.5$ to obtain a set of BGS-like galaxies.

We required the galactic reddening to be small ($E(B-V) < 0.1$) and to have at least one observation in each of the g, r, z bands. With the mask bits, we removed any secondary detections (MASKBIT 0), as well as sources that touch Tycho sources with MAG_VT < 13 and Gaia stars with G < 13 (MASKBIT 1), and sources touching a pixel in a globular cluster (MASKBIT 13). These cuts removed 1.7% of the LS10 area. To further refine the selection, we compared the obtained source list with that of GAMA (fourth data release [Driver et al. 2022](#)) as a benchmark. Indeed, GAMA is almost complete to r magnitude of 19.8, deeper than this selection extending to 19.5. By looking at the fraction of sources without a GAMA counterpart and with any of the FITBITS on, we found that sources with the FITBITS 1, 6, 10, 11, 12, and 13 are primarily spurious, so we removed them. Finally, we removed Gaia duplicates (TYPE="DUP"). We followed the BGS selection to discard artifacts, but with a stricter criterion in color-color space, namely: we kept sources with $-0.2 < g - r < 2$ and $-0.2 < r - z < 1.6$.

We compared the sources selected to GAMA DR4 in the 9 h field and found that 92% (93%) of these have a GAMA spectroscopic counterpart within 1 (2) arc seconds. For the sources with a match, the mean difference in r -band magnitude is ~ 0.1 mag, while the scatter is 0.15 mag, which we might expect given the different filters, images, and pipelines used. The remaining 7–8% of target sources closely follow the distributions of the

ones matched to GAMA, simply reflecting the incompleteness of the GAMA survey if it were targeted with LS10. We carried out this selection on the LS10 data and random data to obtain the samples for auto and cross-correlations. Ultimately, we selected 13 881 761 sources over a footprint of 16 796 square degrees.

The observed number densities of the selected sources (counting sources per square degree), after a correction of the mean r -mag offset, are within a few percent from the GAMA densities for $16.5 < r_{AB} < 19.5$ (see Table B.1 and Fig. B.1). This level of disparity is compatible with the cosmic variance of about 2% for 60 square degrees of the volume subtended by a $r_{AB} = 19.8$ band flux-limited survey ([Driver & Robotham 2010](#)).

For this source sample, we measured the density variations across the footprint as a function of EBV, stellar density, depth, and PSF. We found variations within $\sim 10\%$ of the average density. This is sufficient to obtain an unbiased clustering measurement to feed to HOD models (e.g., [Delubac et al. 2017](#); [Kong et al. 2020](#)).

2.2.1. Galaxy properties

We used the photometric redshifts from [Zhou et al. \(2021\)](#). The photometric redshifts of the aforementioned selection ($r_{AB} < 19.5$) are excellent, with 0.15% bias and 3% scatter. We computed the stellar mass on the $g, r, i, z, W1, W2$ set of bands using LePhare ([Ilbert et al. 2006](#)).

Compared to previous estimates using deeper observations and more bands (COSMOS, GAMA [Ilbert et al. 2013](#); [Driver et al. 2022](#)), our estimated stellar masses are found to be less accurate, with a scatter of 0.18 dex. This accuracy is similar to that obtained by [Zou et al. \(2019\)](#) using the ninth data release of legacy surveys. However, the stellar mass values are, on average, almost unbiased (-0.085 dex). The derived stellar mass functions agree with the literature (see following sections and Fig. B.2). This indicates that it is a fair quantity to use when selecting galaxy samples.

2.2.2. Stellar-mass selected samples

We followed [Zu & Mandelbaum \(2015\)](#), [Farrow et al. \(2015\)](#) to design the stellar mass selected volume-limited samples. The most notable difference is the minimum redshift cut. We imposed $z > 0.05$ instead of 0.02 in GAMA. Indeed, the cataloging methods used in LS10 come with some uncertainties when handling sources with a large extent on the sky. This is, in part, handled in the Siena Galaxy Atlas ([Moustakas et al. 2023](#)), but unfortunately, it does not cover the complete footprint of LS10 (only the footprint of the ninth data release). This entails incompleteness in the sample at low redshift, which becomes small above $z = 0.05$, then galaxies appear smaller in the sky. The volume loss 0.042 Gpc^3 is small and negligible (at least ten times smaller) for the samples selected with a stellar mass larger than $10^{10} M_\odot$. For the low-mass selection, adding it would increase the volume and the significance of the correlation function, but at the cost of completeness and interpretation. The upcoming Rubin/LSST observations ([Ivezić et al. 2019](#)) will enable us to push such investigations to lower redshifts and enhance the future results for lower mass galaxies.

The sample selections are illustrated in Fig. 1, where the galaxy sample is shown in the redshift-stellar mass plane, where the boxes identify the selections. Table 1 summarizes the exact boundaries used and gives the general properties of the selections (redshift and stellar mass ranges, number densities, and completeness). Since the total number of galaxies and the area

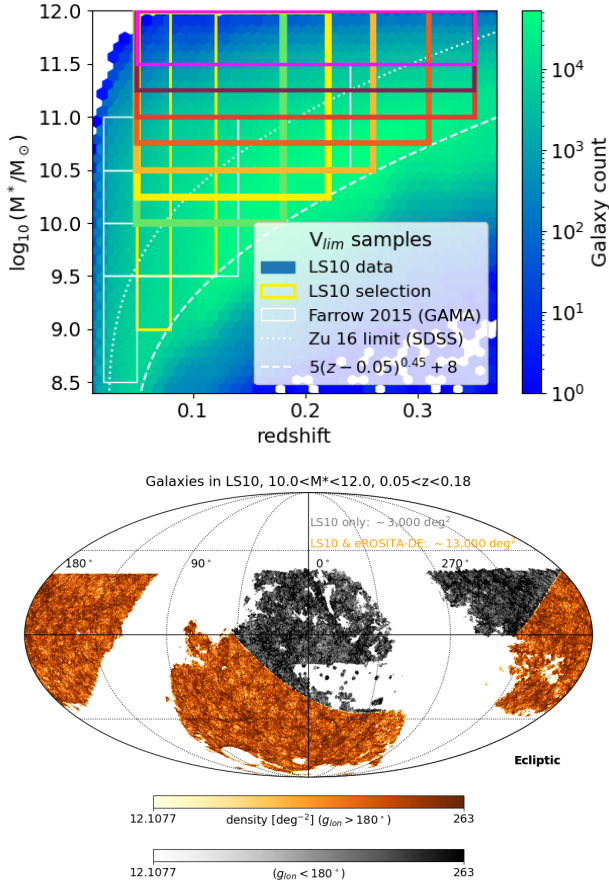


Fig. 1. *Top:* Sample definitions in the redshift-stellar mass plane. In the analysis, we considered the samples encompassed by the colored boxes. The white boxes depict previously defined SDSS and GAMA-based galaxy samples (Farrow et al. 2015; Zu & Mandelbaum 2016). The line $\log_{10}(M^*) = 5(z - 0.05)^{0.45} + 8$ corresponds to the volume limit used in this work. The colorbar depicts the number of galaxies in hexagonal (for better visualization) bins. *Bottom:* Area covered by the sample with $\log_{10}(M^*/M_\odot) > 10$ and $0.05 < z < 0.18$. The filamentary nature of the large-scale structure is visible. The orange (grey) area is (not) included in the eROSITA_DE footprint.

covered are large, we simplified the definitions compared to the GAMA ones. We chose a set of minimum stellar mass threshold (9, 9.5, 10, 10.25, 10.5, 10.75, 11, 11.25, and 11.5) that we mapped; with $\log_{10}(M^*/M_\odot) = 5 \times (z - 0.05)^{0.45} + 8$ (equation form taken from Zu & Mandelbaum 2015 and adapted to our specific case); to a maximum redshift to obtain something close to volume-limited samples. The line used is close to the one used in the GAMA survey (see the GAMA selection in white boxes). This line splits the sample where its density starts to decrease. We obtained high levels of completeness between 96 and 100% when using GAMA as a benchmark (which is deeper and extends to 19.8; see values for each sample in Table 1). These completeness levels are trustworthy at the high mass end, above the turnover of the stellar mass distribution for a given sample, but certainly not at low mass and moderate redshift (below 9 and above $z > 0.05$); in that case, the magnitude limit of GAMA (19.8) will, by construction, deliver an incomplete sample. The stellar mass functions (without weights applied) are in good agreement with similarly estimated (without completeness correction, only volume weighted number count) stellar mass functions from SDSS and GAMA (and from the mock), as shown in Fig. B.2.

These selections are beyond the capabilities of the SDSS spectroscopic survey illustrated by the dotted line in Fig. 1 (see also Figs. B.1, B.2). Compared to Zhang et al. (2024a, 2021), the selections described here cover (at a given redshift) galaxies with significantly lower stellar masses than that from the SDSS group catalog from Tinker (2021, 2022). We obtained a set of GAMA-like galaxy selections over 16 000 deg^2 . The footprint of the $10 < \log_{10}(M^*/M_\odot) < 12$ samples is shown in Fig. 1. Within a narrow redshift range, we can see the largest cosmic structures by eye. The other samples occupy the same sky area.

3. Measurements

This section describes the method to measure the galaxy and event auto-correlation, the galaxy-event cross-correlation, and the stacks. We measured three clustering summary statistics: (i) the projected galaxy correlation function ($w_p(r_p)$), (ii) the galaxy-event angular cross-correlation ($w_{LS93}^{X-\text{corr}}$), and (iii) the stacked surface brightness profiles (w_{DP83}^{stack}).

3.1. Galaxies

We used the LS10's random catalog¹, created following the procedure from Myers et al. (2023). We applied the same cuts as in the data to create a random catalog with five times the total number in the galaxy samples. We assigned redshifts to the randoms by shuffling the measured photometric redshifts (Ross et al. 2020).

We used the Landy & Szalay (1993) estimator and the CorFunc software (Sinha & Garrison 2020) to estimate the clustering. We measured the projected clustering (with photometric redshifts and $\pi_{\text{max}} = 100 h^{-1} \text{Mpc}$):

$$w_p = \frac{G \times G - 2G \times R + R \times R}{R \times R}, \quad (1)$$

where G represents the galaxies and R the randoms. The measurements are then corrected from systematic effects using the mock catalogs. The $w_p(r_p)$ measurements from the observations and the mocks are in good agreement (observations on Fig. 2, mocks on Fig. C.1).

We estimated the clustering signal in equal area pixels (of different sizes from tens to hundreds of square degrees). We constructed a distribution of the clustering amplitude around a projected separation of 5 Mpc/h (well in the two-halo term). We excluded all 3σ outliers and repeat the procedure another time. We excluded a few (on the order of 10) pixels (losing a few hundred square degrees) from the analysis. They reside at the boundary of the survey area either towards the Galactic center or the northern boundary (around a declination of 32°).

The samples defined here do not match those defined in previous analyses (Zehavi et al. 2011; Zu & Mandelbaum 2015; Farrow et al. 2015), so a face-to-face comparison of the measurements is tedious. We find that the measurements are in fair agreement for the same mean stellar mass. When we apply the same measurement procedure to the GAMA catalogs, we can obtain clustering measurements in agreement, albeit with large uncertainties (for the GAMA samples).

The clustering measurements for the lowest stellar mass thresholds are poorly behaved. We would require more accurate photometric redshifts to obtain a robust correlation function in narrow, low-redshift bins. Indeed, in these bins, the photometric redshift uncertainty is comparable to the bin width,

¹ <https://www.legacysurvey.org/dr10/files>

Table 1. Description of the stellar-mass selected samples.

Redshift			Stellar mass				N	Volume Gpc ³	Density		Completeness	
Mean	Median	Max	Min	Mean	Median	Max			deg ⁻²	10 ⁻⁵ Mpc ⁻³	%	σ [%]
0.067	0.068	0.08	9.0	9.89	9.81	12.0	530 387	0.05	31.6	988.4	99.3	2.2
0.094	0.097	0.12	9.5	10.19	10.14	12.0	1 383 581	0.22	82.4	639.1	99.0	1.6
0.136	0.142	0.18	10.0	10.49	10.43	12.0	2 811 951	0.74	167.4	381.0	97.9	1.1
0.162	0.169	0.22	10.25	10.66	10.6	12.0	3 280 777	1.32	195.3	248.4	97.2	0.9
0.191	0.202	0.26	10.5	10.83	10.77	12.0	3 287 997	2.12	195.8	154.8	97.5	0.9
0.226	0.238	0.31	10.75	11.02	10.97	12.0	2 768 066	3.47	164.8	79.7	96.4	0.9
0.252	0.265	0.35	11.0	11.21	11.17	12.0	1 611 928	4.85	96.0	33.3	97.6	1.3
0.255	0.268	0.35	11.25	11.41	11.37	12.0	541 919	4.85	32.3	11.2	99.9	2.7
0.261	0.274	0.35	11.5	11.62	11.59	12.0	1 21 044	4.85	7.2	2.5	100.0	8.4

Notes. The minimum redshift is 0.05 for all samples. The values in the stellar mass columns are in $\log_{10}(M_{\odot})$. N gives the total number of galaxies present in each sample. The “volume” column gives the comoving cosmological volume covered by each sample. We compute the completeness value (and its uncertainty) with GAMA as a reference that extends 0.3 mag fainter in magnitude than the selection done here.

and its effects dominate. In the following, we consider only the samples with a stellar mass threshold larger than $10^{10} M_{\odot}$. We defer the investigation of lower mass samples to spectroscopic surveys.

Uncertainties

To estimate the uncertainties (covariance matrix), we used 1000 re-sampling of the summary statistic, each time leaving out a randomly chosen a 10% fraction of the sample. Figure C.4 shows the correlation coefficient matrix and the diagonal uncertainty. Depending on the sample considered and the scale, the statistical uncertainty, σ_{stat} , varies between 5% (at small separation) and 0.1% (at large separations).

To ascertain that this measurement is only limited by the sub-percent statistical uncertainties, we would need to control (e.g., with weights) density variations across the footprint to the percent level. Reaching this level of precision is difficult and requires complex analysis (e.g., Kong et al. 2020; Berlfein et al. 2024). Given the variation of densities measured across the footprint (see previous paragraphs), we assumed a conservative $\sigma_{\text{syst}} = 5\%$ systematic uncertainty on all $w_p(r_p)$ measurements. We added it in quadrature to the statistical uncertainties derived by resampling to obtain the total uncertainty, $\sigma_{w_p(r_p)}^2 = \sigma_{\text{syst}}^2 + \sigma_{\text{stat}}^2$. An exhaustive study of systematic effects impacting this measurement is beyond the scope of this study. Indeed, massive spectroscopic follow-up is required, which is currently ongoing in DESI.

3.2. X-rays

We measured the cross-correlation between galaxies and events for each eROSITA field, where galaxies are available. A few eROSITA fields are removed due to contamination from bright galactic (extended) foregrounds (see details in Appendix C.1).

The exact area and number of fields removed varies between galaxy samples in this procedure. After removing 100–200 fields, with the remaining 1400–1500 fields, corresponding to approximately 12 000–13 000 deg², we measured the angular cross-correlation between galaxies and soft X-ray events. We show in Fig. D.1 the footprint of the set of eROSITA fields used for the cross-correlation between X-rays and the sample with $\log_{10}(M^*/M_{\odot}) > 10$ and $0.05 < z < 0.18$.

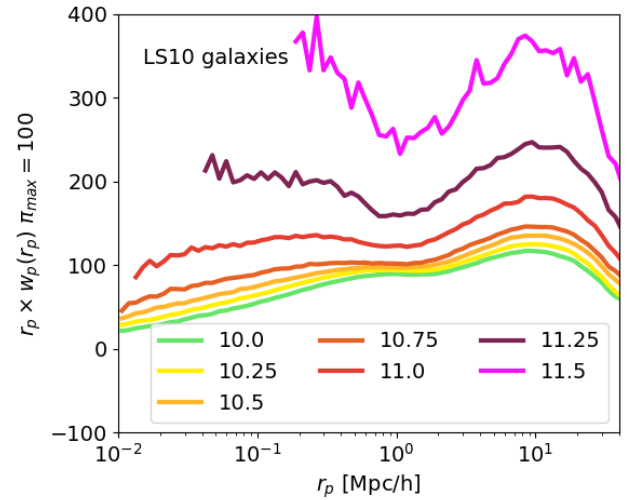


Fig. 2. Projected galaxy auto-correlation function ($w_p(r_p)$) obtained with the LS10 galaxy samples on $\sim 16\,000$ deg². We fit HOD models on these measurements. Differently colored lines represent different samples, identified here by the logarithm of their minimum stellar mass, in solar masses. The color scheme follows that of Fig. 1.

The cross-correlation is obtained with the estimator from Landy & Szalay (1993) with the same set of random points as that used in the galaxy auto-correlation

$$w_{LS93}^{X-\text{corr}}(\theta) = \frac{G \times E - G \times R - E \times R + R \times R}{R \times R}, \quad (2)$$

Here, G represents the galaxies, E the events, R the randoms, and $G \times E$ the normalized pair counts; all are functions of the angular separation. Figure 3 shows the obtained estimated cross-correlations for each galaxy sample. Given the different mean redshifts of each sample, a direct, quantitative comparison is not straightforward. Using the stacking method (as explained below), we obtained the cross-correlation through a different estimator. We found an excellent agreement between the two methods (see the next section and Fig. 4). The cross-correlation is converted to physical units and shown in Fig. 5 (dashes). Therefore, the ordering of the functions is clear: the higher the stellar mass threshold, the higher the correlation.

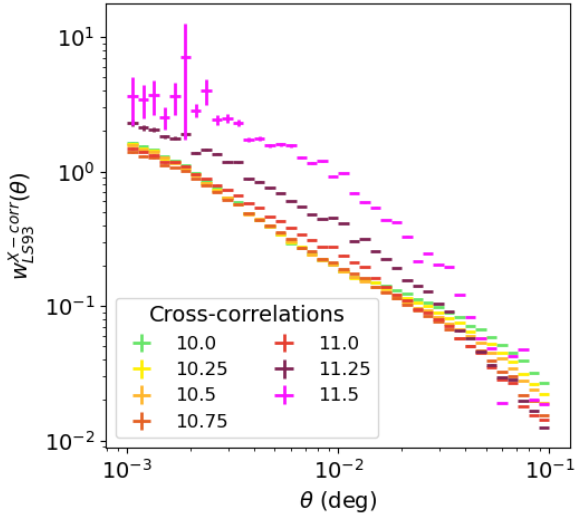


Fig. 3. Measured angular cross-correlation between galaxies and X-ray event angular positions. It is estimated in angular space with Eq. (2). We fit the extension of the HOD model to these measurements. Differently colored lines represent different samples, identified here by the logarithm of their minimum stellar mass, in solar masses. The color scheme follows that of Fig. 1.

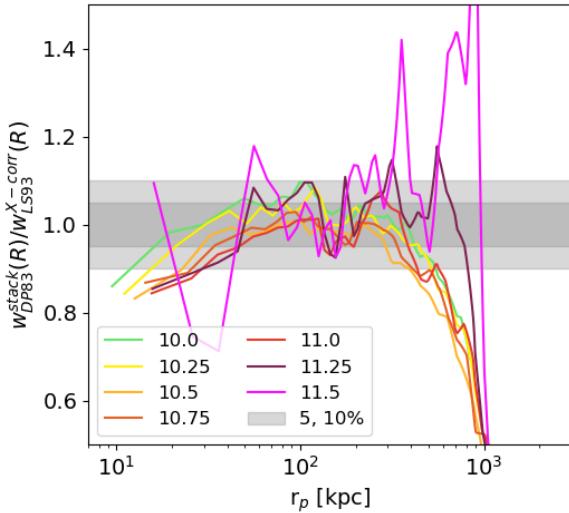


Fig. 4. Ratio of galaxy-event cross-correlation function estimators: $w_{DP83}^{stack} / w_{LS93}^{X-corr}$ (see Eqs. (2), (3)) as a function of separation in proper kpc. The discrepancy between estimators is smaller than 5% in the range ~ 30 – ~ 300 kpc (exact boundaries depending on the samples). For each sample, we fit models on the measurements where the agreement between estimators is better than 5%. For the two highest mass selections, we use the range where the agreement is better than 10%. Differently colored lines represent different samples, identified here by the logarithm of their minimum stellar mass, in solar masses. The color scheme follows that of Fig. 1.

3.2.1. Stacked X-ray surface brightness profiles

As an alternative estimator of the cross-correlation, to verify and strengthen the results, we stacked the X-rays around the same galaxy sample following the method described in Comparat et al. (2022). By counting pairs of galaxies and events as a function of separation, we obtained a surface brightness profile S_X^G . It is similar to $G \times E$ in Eq. (2), however, here we used the redshifts to obtain a result in surface brightness units and proper

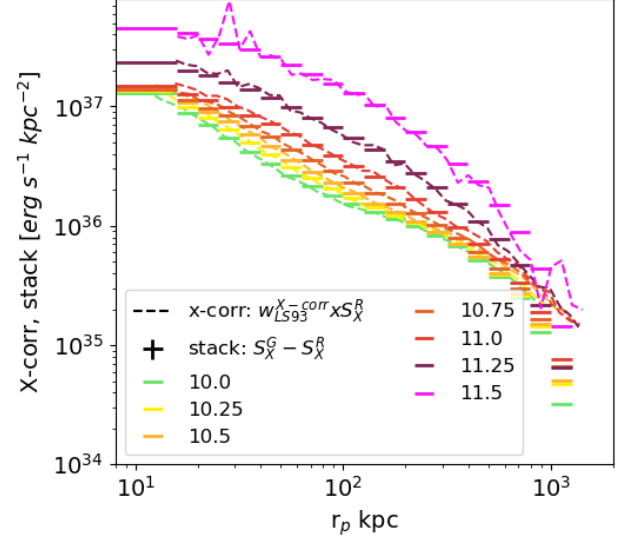


Fig. 5. Comparison of the two cross-correlation estimators in physical units for the stellar mass selected galaxy samples. Differently colored lines represent different samples, identified here by the logarithm of their minimum stellar mass, in solar masses. The color scheme follows that of Fig. 1. The background subtracted surface brightness profiles in the 0.5–2 keV soft band as a function of proper separation (r_p , kpc) is shown in steps ($w_{DP83}^{stack} \times S_X^R$, Eq. (3)). The cross-correlation between galaxies and X-ray events is shown with dashed lines ($w_{LS93}^{X-corr} \times S_X^R$, Eq. (2)). The brightness of the emission is correlated with the stellar mass threshold. The shape of the profiles changes with the stellar mass threshold. The individual best-fit models for these surface brightness profile are presented in Figs. 8, 9, and 10.

Table 2. Background values, S_X^R , for each stack, their uncertainty, and in parenthesis the relative uncertainty in %.

M_0^*	$S_X^R \times (1+z)^{-2}$ [erg kpc $^{-2}$ s $^{-1}$] (%)
10.0	$6.859 \pm 0.032 \times 10^{36}$ (0.47)
10.25	$6.799 \pm 0.028 \times 10^{36}$ (0.41)
10.5	$6.772 \pm 0.025 \times 10^{36}$ (0.37)
10.75	$6.712 \pm 0.027 \times 10^{36}$ (0.39)
11.0	$6.693 \pm 0.051 \times 10^{36}$ (0.76)
11.25	$6.706 \pm 0.078 \times 10^{36}$ (1.16)
11.5	$6.727 \pm 0.109 \times 10^{36}$ (1.61)

separation. In addition, we extracted a surface brightness profile around random points by counting pairs of random points and events as a function of separation S_X^R (similar to $E \times R$ in Eq. (2)). We took the random points from the same set for the galaxy clustering. With that, we were able to estimate the “background” of the stacking experiment. We report the values of S_X^R in Table 2. These values are similar to that derived by Zhang et al. (2024a) $\in 0.8 - 1.2 \times 10^{37}$ erg kpc $^{-2}$ s $^{-1}$, when stacking around different SDSS galaxy samples covering a different redshift range and on a different area of the sky than that used here.

Then, we form the Davis & Peebles (1983) estimator of the cross-correlation

$$w_{DP83}^{stack}(R) = \frac{S_X^G - S_X^R}{S_X^R}. \quad (3)$$

From $w_{LS93}^{X-corr}(\theta)$, we can convert the angular separations into proper distances at the mean redshift of each sample to obtain

$w_{LS93}^{X-corr}(r_p)$ and compare it to $w_{DP83}^{stack}(r_p)$ (Fig. 4). We find the two estimators in 10% (5%) agreement for separations between 20–500 (40–300) kpc. Outside this range, the estimators slowly diverge from one another, with the DP83 lower than the LS93. The large-scale difference may come from excluding the $G \times R$ term as inaccuracies in the photometric redshifts. The difference on small scales may stem from the use of photometric redshifts (in the DP83 estimator) that have a lower accuracy for AGN. To investigate this in detail, one would need event-level simulation with reliable correlation function predictions over Gpc volumes for hot gas, AGN, and XRB events, which is not yet available, but hopefully will be in the future.

When comparing the two estimators in physical units, the cross-correlation (dashes) and stacked (solid steps) measurement in a log-log representation are in excellent agreement in the ranges mentioned above; see Fig. 5.

3.2.2. Uncertainties

Thanks to the large area and number of galaxies used, numbers are sufficient to estimate with the Jackknife method the diagonal uncertainties (Fig. D.2 top panel) and the correlation coefficient of the stacked profiles (Fig. D.2 bottom panel). They are below 2% for the samples with a stellar mass minimum below 11. Only for the two highest stellar mass selections are the uncertainties larger, mainly driven by the smaller sample sizes. The correlation coefficient matrix is diagonal with some noise, so we will only use the diagonal uncertainties. They are smaller than the uncertainty on the background values (Table 2; i.e., on the order of a few percent) needed to convert the cross-correlation to a physical value.

Given the difference obtained with the two estimators of the cross-correlation (Fig. 5), until more simulations are available to quantify the accuracy at which each estimator enables us to retrieve the cross-correlation, we deem necessary to add a $\sigma_{\text{syst}} = 5\%$ (10%) systematic uncertainty to encompass the possible uncertainty on the convergence of the cross-correlation estimator for the samples with threshold in $\log_{10}(M^*/M_{\odot})$ of 10, 10.25, 10.5, 11 (11.25, 11.5). A complete assessment of systematics uncertainties impacting this measurement (and its estimators) requires extensive work on the models linking the large-scale structure to X-ray emission. These new measurements are key to creating new and accurate such models to enhance our understanding of possible systematic uncertainties impacting it.

In what follows, we explain how we used the measurements obtained with the LS93 estimator of the cross-correlation; here, the total diagonal uncertainties considered are the statistical ones, with the systematic one added in quadrature. It is conservative, but it allows us to mine the information contained in the measurements.

When we are working in physical space, the multiplication by the background value incurs an additional statistical uncertainty reported in Table 2 (the last value in percent). There is also a systematic uncertainty in the estimation of the background. Indeed, the estimates on the different samples should all be consistent with the actual value of the background. The seven estimates point to values between 6.7 and 6.85 with a $\sim 2\%$ variation, larger than the statistical uncertainties reported in the table. To remain conservative, we added a 2% systematic uncertainty on the background value (when used).

4. The halo model

Halo occupation distribution (HOD) models describe how galaxies populate dark matter halos, predicting the number and types of galaxies within a given halo based on its mass. These models are essential for linking observations of galaxy clustering to the underlying dark matter distribution. By incorporating statistical relationships between halo properties and galaxy formation, HOD models help refine cosmological simulations and improve our understanding of large-scale structure formation (Cooray & Sheth 2002; Asgari et al. 2023).

We followed Murray et al. (2021), Nishimichi et al. (2019), Asgari et al. (2023), Mead & Asgari (2023) to implement the HOD model and link the galaxy samples to their host dark matter halo population (Sect. 4.1). Using ultranest (Buchner 2016, 2019, 2021), we fit a classic HOD model on the $w_p(r_p)$ measurement, shown on Fig. 2. We added an X-ray dimension to the HOD model (Sect. 4.2) to fit the cross-correlation measurement, shown in Fig. 3.

4.1. Model for the galaxy auto-correlation

To model the projected clustering $w_p(r_p)$ of a galaxy sample, we assumed the measurement to be at the mean redshift of the sample. We relied on the theoretical framework from van den Bosch et al. (2013), More et al. (2015).

4.1.1. Cosmological setup

The cosmological parameters are fixed to $\Omega_c = 0.26069$, $\Omega_b = 0.04897$, $\Omega_k = 0.0$, $h = 0.6766$, $As = 2.09 \times 10^{-9}$, $n_s = 0.965$, $w = -1.0$, and $w_a = 0.0$, $m_\nu = 0.0$ (Planck Collaboration VI 2020). The deduced parameters are $\sigma_8 \sim 0.811$, $\omega_b = \Omega_b h^2$, $\omega_c = \Omega_c h^2$, and $\Omega_m = \Omega_c + \Omega_b = 0.30966$. The linear dark matter power spectrum is then predicted at the redshift of interest using CAMB (Lewis et al. 2000).

For the halo model, we used the Tinker et al. (2010) halo mass function with masses expressed in units of 200 times the mean density of the Universe (M_{200} in M_{\odot}/h). Halo quantities are computed for halos with a mass starting at $10^9 M_{\odot}/h$. Given the halo mass, we computed the virial radius, r_v , based on the halo mass and overdensity condition and the average concentration of the halo $c_s(M)$ using the (Duffy et al. 2008) relation. We performed the Fourier transform of all corresponding halo profiles (assuming Navarro et al. 1997, NFW hereafter, for the set of virial radii) on the wavenumber grid used for the power spectrum to obtain the halo window function $U_k(k, r_v, c_s)$. With these, we predict the dark matter halo profile.

4.1.2. Parameter grid and occupation function

Conceptually, we followed the HOD parametrization of Zheng et al. (2007, Eqs. (2), (5)). Regarding the complete galaxy samples considered here, the satellite galaxies start populating halos at higher halo masses than central ones. We can replace the $\langle N_c \rangle(M)$ factor in front of the satellite occupation with a Heaviside step function for computational speed. Indeed, in that case, the two formulations are numerically very close. We parametrize the occupation number of central (satellite) galaxies in distinct halos, as follows,

$$\langle N_c \rangle(M) = 0.5 \left[1 + \text{erf} \left(\frac{\log_{10}(M) - \log_{10}(M_{\min})}{\sigma_{\log(M)}} \right) \right], \quad (4)$$

$$\langle N_s \rangle(M) = \mathcal{H}(M - M_0, 1.) \left(\frac{M - M_0}{M'_1} \right)_{\text{sat}}^{\alpha} \quad (5)$$

erf is the error function and \mathcal{H} is the Heaviside function.

We tie the two parameters M'_1 and M_0 with: $\log_{10}(M'_1) = \log_{10}(M_0) + 1$. Indeed, theoretical (Zheng et al. 2005) and observational (Zheng et al. 2007; Zehavi et al. 2011) application of this model showed that M_1 and M_0 typically vary by a factor of 10. To avoid sampling un-physical values of M_0 , the parameter passed to the fitting routine is $\Delta_{M_0} = \log_{10}(M_0) - \log_{10}(M_{\min})$. This parametrization can capture the clustering of the (almost complete) samples designed here. We adopted a uniform prior with ranges: $\log_{10}(M_{\min}) \in [10, 15]$, $\alpha_{\text{sat}} \in [0.01, 1.5]$, $\sigma_{\log(M)} \in [0.01, 1.5]$, $\Delta_{M_0} \in [-1, 2]$. We limited M_{\min} to a smaller interval for the higher mass samples.

4.1.3. Fitting function

Given a set of parameters and at each fit iteration, the following is evaluated and compared to the measured signal. We computed the expected (mean) number of central and satellite galaxies as a function of halo mass: $N_{\text{cen}}, N_{\text{sat}}$. We computed the expected variance (and covariance) in the numbers of central and satellite galaxies assuming Bernoulli statistics for centrals and Poisson statistics for satellites: $V_{\text{cen}}, V_{\text{sat}}$. Then we integrated the occupation function with the halo mass function to retrieve the mean density of galaxies: $\rho_g[(\text{Mpc}/h)^{-3}] = \int \langle N \rangle(M) n(M) dM$.

For central galaxies, we used a delta function profile. Satellite galaxies follow the NFW profile. The central profile depends on $N_{\text{cen}}, \rho_g, V_{\text{cen}}$, and the satellite profile on the corresponding quantities. Finally, with all the ingredients mentioned above, we predict the galaxy power spectrum contributed by central satellites and their cross terms: $P_{gg}(k) = P_{cc}(k) + 2P_{cs}(k) + P_{ss}(k)$. Each comprises two components: a one-halo term and a two-halo term. The implementation used to compute the power spectrum is that of Asgari et al. (2023, Eqs. (20)–(29)). With the Hankel transform, we obtained the configuration space real space correlation function ξ_{gg} , which we project into a projected correlation function $w_p(r_p)$ by doing the Abel transform up to $\pi_{\max} = 100 h^{-1} \text{Mpc}$. We interpolated the obtained function on the same grid as the measurement and compute a χ^2 summary statistic.

4.1.4. Angular clustering prediction

The angular clustering was predicted using the Limber (1953) approximation. We used the implementation of Murray et al. (2021).

4.1.5. Validation with the mock catalogs

We tested the HOD implementation with the mock catalogs (see Appendix A). We retrieved the input HOD within a few per cent after fitting the predicted correlation function. The range of scales where the algorithm works seamlessly is 0.01–30 Mpc/h and at the few percent precision level. To reach beyond that precision more detailed HOD models are required (e.g., Zu & Mandelbaum 2015, 2016; Contreras et al. 2021; Yuan et al. 2022). In our case, an accuracy of a few percent is sufficient.

4.2. A model for the galaxy-event cross-correlation

This section aims to obtain a proof-of-concept model that accounts for the cross-correlation measurement to grasp the

information content to the first order. To predict the cross-correlation between soft X-ray photons and galaxies, we fix the parameters of the HOD models to the ones obtained when fitting the $w_p(r_p)$ (see Sect. 5.1).

4.2.1. Hot gas component

In this model, the hot gas emission comprises the circumgalactic, intragroup, and intracluster medium. It corresponds to any X-ray emission coming from a hot gas component. In the following, for simplicity, we designate the hot gas filling the dark matter halo volume, regardless of halo mass (or temperature), by “hot gas” or the circumgalactic medium (CGM). Recent advances in the studies of cosmic rays indicate that a fraction of the X-ray flux may find its origin there (e.g., Hopkins et al. 2025). We opt here for a classic hot gas model with only Bremsstrahlung. We leave the implementation of cosmic rays in the HOD model to future works.

We converted (at mean redshift) the halo mass distribution from the HOD from a 200m definition to a 500c definition using the COLOSSUS software (Diemer 2018). Instead of using the Duffy et al. (2008) relation (as in the halo model), we use that of Ishiyama et al. (2021) that has a native mass definition on both M500c and M200c and thus provides a slightly more precise conversion. Then, we predict the average X-ray luminosity for each halo with the following scaling relation with

$$\log_{10}(L_X^{0.5-2 \text{ keV}}) = L_0 + \alpha_{SR}(\log_{10}\left(\frac{M_{500c}}{10^{15}}\right) + 2 \log_{10}(E(\bar{z})), \quad (6)$$

where α_{SR} and L_0 are the slope and amplitude of the scaling relation and $E(\bar{z}) = H(\bar{z})/H(0)$ is the normalized Hubble parameter.

We convert the mass-luminosity relation into a mass-count relation to predict the average number of counts received as a function of halo mass.

$$\log_{10}(N_{ct}) = \log_{10}\left(\frac{L_X^{0.5-2 \text{ keV}} t_{\text{exp}} \text{ARF}}{4\pi C}\right) - 2 \log_{10}(d_L). \quad (7)$$

There, we denote d_L as the luminosity distance to the mean redshift of the sample. We convert the individual photon energy in keV to erg with a constant factor $C = 1.602177 \times 10^{-9} [\text{erg/keV}]$ without assuming a spectral model and virtually giving all the counts an energy of one keV. Using the mock catalogs from Comparat et al. (2020), Seppi et al. (2022), where detailed spectra are used to convert the rest-frame 0.5–2 keV luminosity into an observed frame flux, we find that a more accurate C has minimal effect on the predicted average count. The scatter increases by $\sim 10\%$, while the mean and the slope remain unchanged. We take the vignetting corrected Field-of-view-averaged Auxiliary Response File (ARF) for all seven telescope modules combined at 1 keV of 1000 cm² to obtain at a given mass the average number of counts expected in the band. We assume the median unvignetted exposure time of eRASS:4 of $t_{\text{exp}} = 550$ s. The cross-correlation is sensitive to the slope of the scaling relation, not to its amplitude, so the exact values of $t_{\text{exp}}, \text{ARF}$ matter little in the prediction. If we were also to model the luminosity functions, one would need to fold in the complete distribution of exposure times. As the samples studied are at low redshift, we ignore the impact of K-correction (and temperature) as a function of mass. As stated above, this assumption has a minimal impact. The sky area considered has low foreground absorption; we ignore its effect. In Sect. 6.2 we show the energy dependence of the cross-correlation. In the future, we will use models with

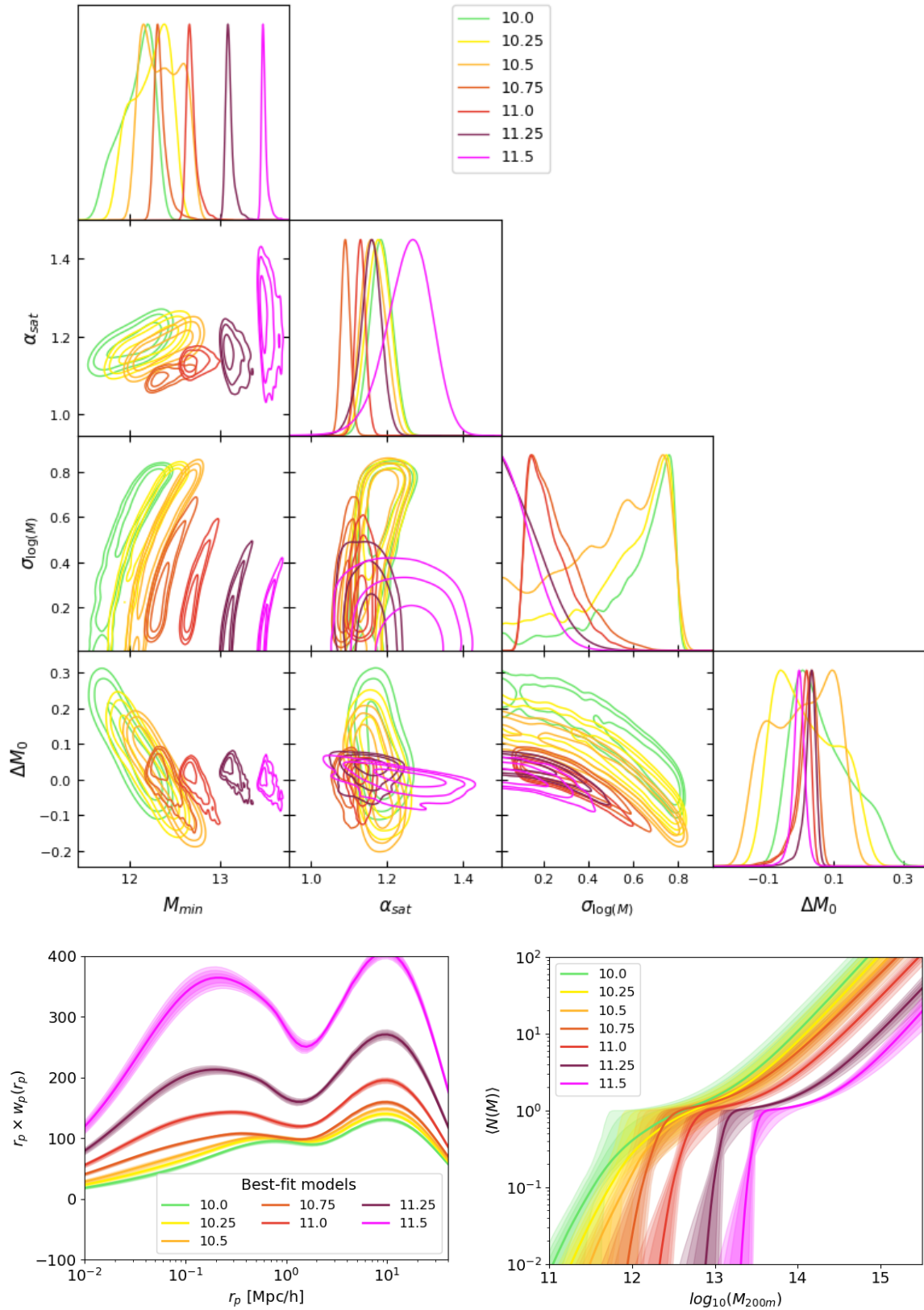


Fig. 6. HOD best-fit model parameters obtained by adjusting the model from Sect. 4 on the $w_p(r_p)$ measurements obtained in Sect. 3 and shown in Fig. 2. Differently colored lines represent different samples, identified here by the logarithm of their minimum stellar mass, in solar masses. The color scheme follows that of Fig. 1. We show the corner plot of the parameters and their 1,2,3 σ contours in the top panel. The trends of the parameters obtained and their correlation with the stellar mass threshold of the different galaxy samples is sensible: M_{\min} is correlated with the minimum stellar mass, while $\sigma_{\log(M)}$ is anti-correlated. Model $w_p(r_p)$ obtained and their 1,2,3 σ contours (bottom left). Posterior HOD with 1,2,3 σ contours (bottom right). The ordering of the HOD curves follows the stellar mass selection and follows expectations. An individual comparison of the models and the measurements are shown in Fig. C.5.

Table 3. Best-fit parameters obtained by fitting the HOD model on the galaxy auto-correlation functions.

Sample M_0^*	Halo occupation distribution model					Deduced quantities					
	Best-fit parameters				$\frac{\chi^2}{\text{nd.o.f.}}$	Densities		f_{sat}	b	$\log_{10}(\bar{M}_{200m})$	
	M_{min}	α_{sat}	$\sigma_{\log(M)}$	ΔM_0		deg^{-2}	10^{-5} Mpc^{-3}	%		\bar{M}_{cen}	\bar{M}_{sat}
10	$12.113^{+0.409}_{-0.257}$	$1.184^{+0.059}_{-0.029}$	$0.666^{+0.335}_{-0.236}$	$0.052^{+0.182}_{-0.067}$	0.49	103.9	235.8	23.0	1.22	12.66	14.14
10.25	$12.26^{+0.468}_{-0.276}$	$1.178^{+0.062}_{-0.031}$	$0.619^{+0.437}_{-0.303}$	$0.014^{+0.204}_{-0.083}$	0.43	128.4	162.9	23.0	1.29	12.79	14.13
10.5	$12.362^{+0.483}_{-0.23}$	$1.163^{+0.063}_{-0.03}$	$0.538^{+0.506}_{-0.312}$	$0.016^{+0.208}_{-0.113}$	0.57	148.5	117.1	22.1	1.34	12.9	14.12
10.75	$12.327^{+0.132}_{-0.047}$	$1.091^{+0.023}_{-0.011}$	$0.228^{+0.231}_{-0.09}$	$0.031^{+0.056}_{-0.035}$	0.84	208.6	100.7	25.3	1.43	12.98	14.04
11	$12.674^{+0.105}_{-0.039}$	$1.131^{+0.026}_{-0.013}$	$0.202^{+0.194}_{-0.072}$	$0.018^{+0.046}_{-0.03}$	0.8	123.0	42.5	20.5	1.6	13.24	14.12
11.25	$13.096^{+0.077}_{-0.03}$	$1.159^{+0.053}_{-0.028}$	$0.123^{+0.199}_{-0.082}$	$0.036^{+0.031}_{-0.019}$	1.02	39.2	13.6	13.9	1.89	13.57	14.26
11.5	$13.483^{+0.057}_{-0.021}$	$1.261^{+0.119}_{-0.064}$	$0.1^{+0.167}_{-0.062}$	$0.002^{+0.031}_{-0.016}$	1.93	12.7	4.4	9.9	2.32	13.87	14.42

Notes. Best-fit parameters (with 1σ uncertainties) and deduced parameters obtained by fitting the HOD model on the galaxy auto-correlation functions (Fig. 2). The parameters are described in Sect. 4. The deduced quantities model densities of tracers, satellite fraction, large-scale halo bias, and mean halo mass hosting central and satellite galaxies are computed for the best-fit model.

energy dependence (and spectral models) upon the complete set of cross-correlations as a function of energy.

Using the best-fit parameters of the auto-correlation function, the halo mass function, and the volume covered by the sample, we deduce the number of distinct halos occupied in each halo mass bin. For each mass bin and each distinct halo we draw an observed luminosity from a Gaussian distribution centered on $\log_{10}(L_X^{0.5-2\text{keV}}, M_{500c})$ with a fixed intrinsic scatter $\sigma_{LX} = 0.3$ compatible with the value inferred by Lovisari et al. (2015), Bulbul et al. (2019), Comparat et al. (2020), Seppi et al. (2024). For all masses, we deduce the average number of counts emitted and its standard deviation, which we use as the amplitude and variance of the X-ray profile.

We then need a functional form for the 3D profile of the events. For an isothermal optically thin plasma, the soft X-ray flux (event rate) is proportional to the square of the electron density. The generic profile formulated by Vikhlinin et al. (2006) (or the simple beta profile) informs on the general shape of the electron density. Here, we need a 3D profile for the events, not the electron density, so previous profile parameters are not directly applicable. The most direct way to obtain a simple analytic formula for such a profile is to use the eROSITA digital twin and events emitted by clusters and groups (Seppi et al. 2022) to find a suitable functional form in 3D that will, by construction, in 2D, follow the surface brightness profile expected. We find that the event profiles of distinct halos in the simulation are well-described by

$$p_e(r, r_{200m}, c_{200m}) \propto \frac{1}{\left(\frac{r}{r_{200m}}\right)^{\alpha_{\text{prof}}}} \times \frac{1}{\left(1 + \left(\frac{r}{r_{200m}}\right)^2\right)^{p_2}}, \quad (8)$$

where c_{200m} is the halo concentration and r_{200m} its radius. Their ratio constitutes the scale radius ($\frac{r_{200m}}{c_{200m}}$) of the halo. It is a characteristic radius that defines the transition between the inner, dense core and the outer, more diffuse region of the halo. We use $(\alpha_{\text{prof}}, p_2) = (0.9, 1.6)$, close to the analytical formulae from Vikhlinin et al. (2006). Converting to a surface brightness, this corresponds on large scales to a beta profile with $\beta = 0.53 (= 1.6/3)$, consistent with measurements in clusters (e.g., Neumann & Arnaud 1999; Mantz et al. 2016; Sanders et al. 2018; Eckert et al. 2019; Etori et al. 2019; Liu et al. 2022; Bulbul et al. 2024, albeit with different profile

functional forms, the slope on large scales is broadly consistent), but steeper than values measured for lower mass halos by Zhang et al. (2024a, that find 0.4). Current observations indicate the need a slope varying with mass to guarantee consistency. Here, we use a value consistent with clusters where most of the events come from, but inconsistent with the measurement from Zhang et al. (2024a). We leave to future models the exploration on how to free the profile slope as a function of mass.

In the observation, we cross-correlate all the events with a complete set of galaxies. This is reflected in the profile's normalization, which gives the cross-correlation amplitude. From simulations (Seppi et al. 2022), we estimate the fraction of hot gas events correlated to galaxies to be $w_0 = 1-2\%$ (depending on the redshift range). Knowing the exact value of w_0 is complex and depends on models of the X-ray foregrounds and backgrounds.

The power spectrum is the sum of the cross components: events with central and satellite galaxies. It contains contributions from the one-halo and the two-halo approach in each term. The one-halo term accounts for galaxy correlations within the same halo, which is important on small scales. The two-halo contains galaxy correlations between separate halos, it dominates on large scales. We compute the Hankel transform to obtain the real-space correlation function and the Limber equation to obtain an angular correlation function, denoted $w_{\text{Gas} \times \text{Gal}}$. Finally, we convolve the cross-correlation function with the eROSITA PSF (Merloni et al. 2024). This prediction works when the set of events solely comes from hot gas around galaxies following that profile and scaling relation.

In future model incarnations, the scaling relation and profile parametrizations (and their scatter) should vary with halo mass (Schaye et al. 2023). Using the scaling relation parametrization from Chiu et al. (2022, Eq. (67)), the profile parametrization from Vikhlinin et al. (2006, Eq. (3)), we can have up to $5 + 10 = 15$ parameters (with degeneracies). The X-ray spectral aspect should be taken into account, possibly by parametrizing the mass-temperature scaling relation to deduce the mass-luminosity relation in different energy bands (Lovisari et al. 2021).

4.2.2. Point sources

In observations, there are X-ray point sources (XRB and AGN) in the galaxies that contribute to the signal. Using Aird et al. (2017) and Comparat et al. (2019), we can estimate that the XRB

Table 4. Best-fit parameters obtained by fitting the model on the galaxy-event cross-correlation.

Sample M_0^*	Surface brightness profile model				Deduced L_X [$10^{40} \text{ erg s}^{-1}$]	
	α_{SR}	$\Delta_{M_{\min}}$	w_0	$\frac{\chi^2}{nd.o.f.}$	Point source	hot gas
10.0	$1.629^{+0.091}_{-0.089}$	$0.911^{+0.099}_{-0.101}$	0.011	$0.918^{+0.5114}_{-0.5501}$	$3.59^{+0.16}_{-0.15}$	$7.81^{+0.16}_{-0.16}$
10.25	$1.573^{+0.077}_{-0.07}$	$1.007^{+0.093}_{-0.087}$	0.014	$0.83^{+0.5237}_{-0.5423}$	$4.54^{+0.19}_{-0.18}$	$11.6^{+0.18}_{-0.18}$
10.5	$1.612^{+0.068}_{-0.072}$	$1.131^{+0.109}_{-0.101}$	0.016	$0.752^{+0.5395}_{-0.534}$	$6.2^{+0.22}_{-0.28}$	$15.64^{+0.18}_{-0.18}$
10.75	$1.654^{+0.166}_{-0.104}$	$0.851^{+0.358}_{-0.331}$	0.019	$1.083^{+0.448}_{-0.6232}$	$8.63^{+0.66}_{-0.54}$	$18.27^{+0.56}_{-0.55}$
11.0	$1.634^{+0.066}_{-0.094}$	$1.221^{+0.279}_{-0.421}$	0.021	$1.103^{+0.7386}_{-0.3429}$	$7.6^{+0.75}_{-0.28}$	$42.11^{+0.93}_{-1.08}$
11.25	$1.713^{+0.087}_{-0.143}$	$2.181^{+1.779}_{-2.041}$	0.021	$0.865^{+0.7073}_{-0.3646}$	$7.43^{+0.21}_{-0.05}$	$137.91^{+3.07}_{-3.09}$
11.5	$1.544^{+0.256}_{-0.244}$	$2.054^{+1.906}_{-1.924}$	0.021	$0.788^{+0.2152}_{-0.0819}$	$5.53^{+0.07}_{-0.01}$	$401.2^{+12.32}_{-12.99}$

Notes. Best-fit parameters (α_{SR} , $\Delta_{M_{\min}}$ and their 1σ uncertainties) obtained by fitting the model on the galaxy-event cross-correlation (Fig. 3). The w_0 is parameter fixed. The parameters are described in Sect. 4. Deduced quantities are obtained by integrating the model profiles.

contribution amounts to $\lesssim 10\%$ of the total point source luminosity. We refer to Fig. 11, where the blue circles corresponding to XRB are ~ 10 times lower than the total model emission in red stars. We opted for a simple formulation of the X-ray point sources occupation in galaxies driven by the AGN HOD occupation. In Comparat et al. (2023), we measured the halo occupation distribution (using the formalism of More et al. 2015) of detected AGN. Here, the galaxy sample also contains undetected AGNs and XRBs. We used the HOD parameters given in the Table 2 (top set of values) from Comparat et al. (2023) except for M_{\min} that is lowered by $\Delta_{M_{\min}}$ to include undetected AGN and XRB. For simplicity, we assumed only AGNs in central galaxies, since the way in which AGNs populate satellite galaxies remains an open question (Leauthaud et al. 2015; Comparat et al. 2023; Powell et al. 2024). Possible signals from the cross-correlation with satellite galaxies hosting AGN are degenerate (and thus absorbed) with that of hot gas events correlated with satellite galaxies. We obtained a power spectrum for the point sources that we convert to the angular correlation function. After the fitting procedure, we extracted the average point source luminosity in central galaxies in the sample that contains contributions from AGNs and XRBs. Similarly to the hot gas, a possible X-ray point source model may be (in the future) formulated with physical links to the halos, galaxies, and AGNs as in Zhang et al. (2023, with more than 20 parameters), and XRB as in Aird et al. (2017, with three parameters).

4.2.3. Complete model

The model for the measured cross-correlation ($w_{LS93}^{X-\text{corr}}(\theta)$) is expressed as

$$w^{\text{model}}(\theta) = w_{\text{PSF}}(\theta, \Delta_{M_{\min}}, w_0) + w_{\text{Gas}} \times \text{Gal}(\theta, \alpha_{SR}, w_0), \quad (9)$$

with three free parameters $\Delta_{M_{\min}}$ (X-ray point sources), α_{SR} (hot gas), and w_0 (the fraction of correlated events in the complete soft X-ray field). With this model, we can decompose the signal as a function of scale into its main components. The normalization parameter (w_0) would not be necessary if we could cross-correlate exactly the photons coming from the large-scale structure in which the galaxies reside.

Further efforts in adding physics in each component and freeing parameters in the X-ray model interfaced with the HOD model is a complex task left for future investigation. Indeed, we do need to account for the physics of the hot gas (AGN feedback, metallicity, density profiles) across five orders of magnitude

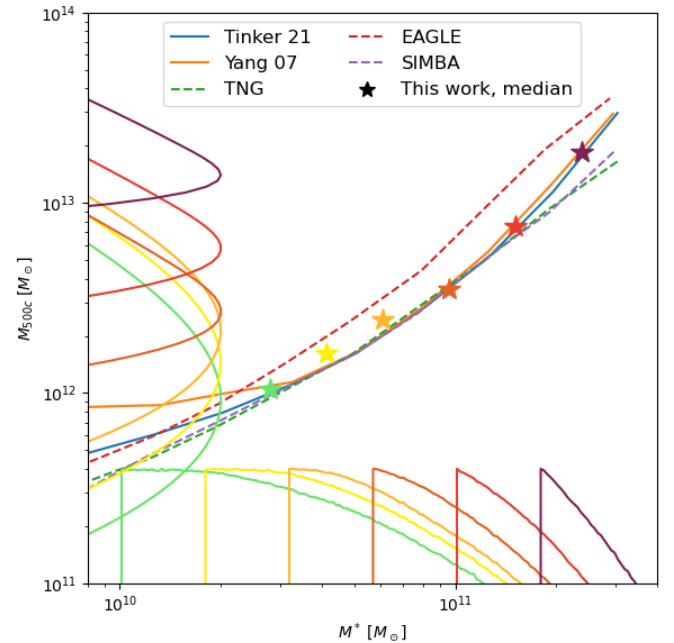


Fig. 7. Inferred stellar-to-halo-mass relation for each sample observed. The stars are located at the median values. The distributions of the halo and stellar masses of the sample studied are shown with colored solid lines. There is a fair agreement with previous models Yang et al. (2005), Tinker (2021) and with the prediction from the TNG and SIMBA hydrodynamical simulations (Pillepich et al. 2018; Davé et al. 2019; Schaye et al. 2015).

in halo mass (e.g., Oppenheimer & Schaye 2013; Flender et al. 2017; Biffi et al. 2018; Comparat et al. 2020; Pop et al. 2022; Oren et al. 2024; Singh et al. 2024; Lau et al. 2024, 2025, expCGM²), the incidence and brightness of AGNs as a function of stellar mass and star formation rate (e.g., Aird et al. 2015; Georgakakis et al. 2017; Biffi et al. 2018; Comparat et al. 2019, 2023; Georgakakis et al. 2019; Allevato et al. 2021), and of the emission of X-ray binaries as a function of stellar mass and star formation rate (e.g., Aird et al. 2017; Lehmer et al. 2016; Vladutescu-Zopp et al. 2023) to construct a comprehensive model.

² <https://gmvoit.github.io/ExpCGM>

This model serves as a proof-of-concept to show the complexity of the data vector considered. Unifying the halo model to predict summary statistics related to galaxies, AGNs, and hot gas both in the optical and X-ray will take significant effort and time. This should serve as a springboard for the community to engage in this complex modeling task.

5. Results

We present here the best-fit models on the galaxy auto-correlation (Sect. 5.1) and on the galaxy-event cross-correlation (Sect. 5.2).

5.1. Best-fit HOD models of the galaxy projected correlation function ($w_p(r_p)$)

The best-fit models account for the observed correlation function within a few percent, with acceptable reduced χ^2 . We show the best-fit model, together with the measurements (see Fig. C.5). A few χ^2 values are quite significantly below 1, hinting that the systematic uncertainty budget of 5% may be too generous in these cases (the ones with stellar mass threshold of 10, 10.25, and 10.5). The best-fit parameters and their 1σ uncertainties are listed in Table 3. The value of M_{\min} ($\sigma_{\log M}$) is correlated (anti-correlated) with the stellar mass threshold. The satellite slope (ΔM_0) is consistent with 1.2 (0) for all samples. The corner plot with the $1, 2, 3\sigma$ contours, shown in Fig. 6 (top panel), illustrates the trends. The bottom panels show the best-fit models for $w_p(r_p)$ and the halo occupation $\langle N(M) \rangle$. The evolution of the parameters with the different stellar mass thresholds is sensible.

The highest mass bin's summary statistic suggests the model does not perfectly account for the observations (see Table 3). The reduced χ^2 is too large at 1.93. In Fig. C.5, bottom right panel, we see deviation between the data and the model at the transition between the one-halo and the two-halo term (1–3 Mpc/h range). The small stellar mass range and the extended redshift range likely play a role in this disagreement. This sample requires further model refinements to fully capture the information present in the measurement.

From the best-fit models, we deduced the fraction of satellites and the large-scale halo bias (Table 3). The deduced parameters' trends are sensible: f_{sat} decreases from 23% to 10% throughout the samples. The large-scale halo bias (mean halo mass hosting central galaxies $\log_{10}(M_{200m}/M_\odot)$) increases from 1.22 to 2.32 (12.66 to 13.87). The samples very well probe the MW-mass galaxies and M31-mass galaxies up to galaxy groups. Throughout the samples, satellite galaxies are hosted by clusters with a mean mass above $\log_{10}(M_{200m}/M_\odot) > 14$. Qualitatively, these parameters and trends are consistent with previous studies (e.g., Zehavi et al. 2005, 2011; Farrow et al. 2015; Zu & Mandelbaum 2015, 2016). Quantitative comparisons with these studies are technically challenging due to the differences in the models used.

The deduced densities (per square degree and cubic Mpc) are in broad agreement with observations. The differences between the two may stem from, in order of importance: (i) redshift evolution of the sample and (ii) uncertainties on photometric redshifts and stellar masses. It indicates a direction for future improvement of the HOD model for these samples.

Each sample, with its model, finds its location on the stellar to halo mass relation (Fig. 7). We get consistent values with previous models (Yang et al. 2005; Tinker 2021) or predictions (Pillepich et al. 2018; Davé et al. 2019; Schaye et al. 2015; McAlpine et al. 2016). Overall, the HOD parameters represent

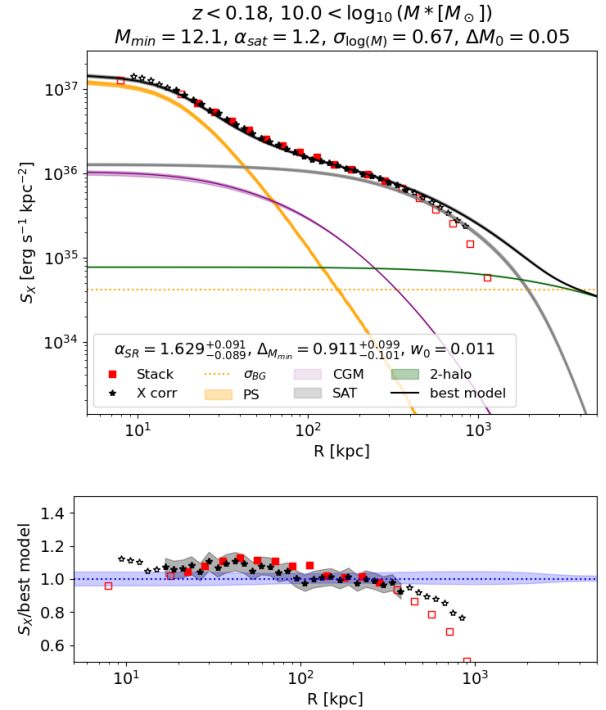


Fig. 8. Best-fit model for the galaxy X-ray cross-correlation obtained with the stellar mass-selected sample with mass $M^* > 10^{10} M_\odot$ (10.0). The black star (red square) symbols indicate the data points obtained with the estimator from the LS93 estimator, see Eq. (2) (DP83 estimator, Eq. (3)). The symbols are filled in the range where good agreement is obtained between estimators. These values are used for fitting the model. Empty symbols are measurements outside of that range. They are not used for fitting. We compare the observed cross-correlation (black stars) and stacks (red squares) to the best-fit model (black line) and its components: circumgalactic medium (purple), point sources (yellow), satellites (grey), and two-halo term (green). The uncertainties are 1σ . The HOD parameters used are fixed at the median of the best-fit model. We indicate the values in the title. The median best-fit cross-correlation parameters are indicated in the legend title (and in Table 4). The reduced χ^2 is 0.918, indicating that the model accounts for the data. The horizontal dotted line shows the uncertainty on the background determination for this sample. The observed cross-correlation is dominated at small separation by the point source emission, then at larger separation by the emission of the hot gas hosted by large halos ($M_{200m} > 10^{14} M_\odot$) seen from the positions of satellite galaxies. Beyond separations of 3 Mpc, the two-halo term dominates. The contribution to the total emission of average emission from the hot gas in distinct halos hosting central galaxies in the inner 200 kpc varies between 5 and 15%. We show the ratio of the cross-correlation (stack) and the best-fit model in the bottom panel. In the range of 20–400 kpc, the residuals are consistent with the model within 1σ uncertainties. Outside of the fitting range, on small scales, the cross-correlation is up to 20% larger than the best-fit model. On large scales, at 1 Mpc, the measurements are smaller than the best-fit model by 10–20%. The result for the other samples are shown in Figs. 9 and 10.

the observed auto-correlation functions well. We fixed the HOD model parameters to their best-fit values to model the cross-correlation.

5.2. Models of the cross-correlation and stacks

We considered three parameters (α_{SR} , ΔM_{\min} , w_0) in this step. We manually adjusted w_0 to match the cross-correlation on large scales (beyond the point sources' influence), giving values

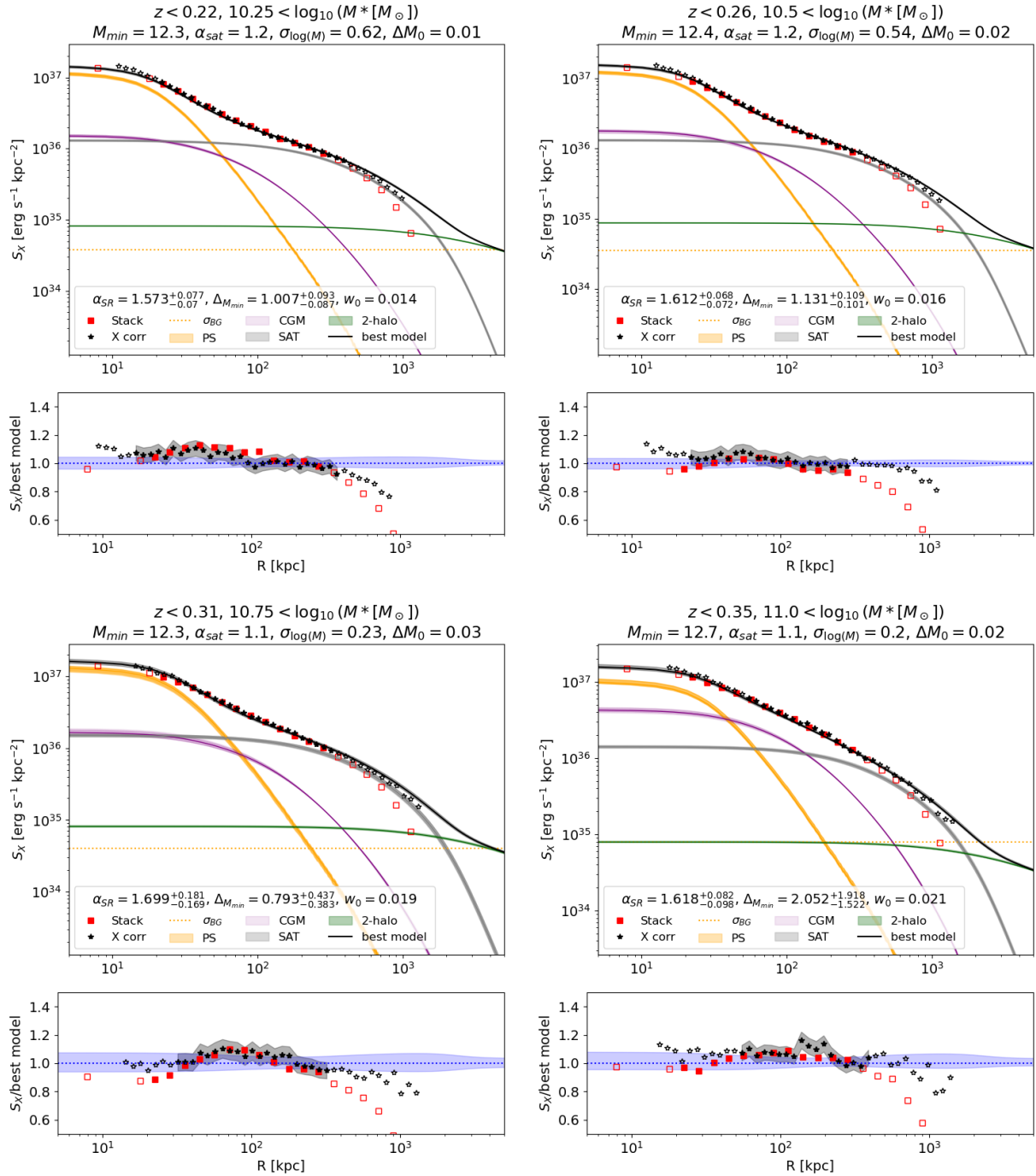


Fig. 9. Fig. 8 continued. Models of the soft X-ray surface brightness profiles for the stellar mass-selected sample with mass $M^* > 1.7 \times 10^{10} M_\odot$ (top left, 10.25), $M^* > 3.1 \times 10^{10} M_\odot$ (top right, 10.5), $M^* > 5.6 \times 10^{10} M_\odot$ (bottom left, 10.75), $M^* > 10^{11} M_\odot$ (bottom right, 11.0). The model accounts well for the measurements. The point source, satellite and two-halo terms show little evolution. The hot gas model brightness increases significantly with the stellar mass threshold.

between 1 and 2 percent as simulations indicate. We report the values found by hand in Table 4. We computed a set of models on a grid of the other two parameters: $\alpha_{SR} \in [1.4, 1.8]$, and $\Delta M_{min} \in [0, 4]$. We fixed the scaling relation amplitude parameter to $\log_{10}(L_0) = 44.7$ to force the compatibility with scaling relations obtained at high mass for galaxy clusters (Seppi et al. 2024; Chiu et al. 2022; Mantz et al. 2016; Lovisari et al. 2015; Bulbul et al. 2019; Pratt et al. 2009).

We obtained a reasonable reduced χ^2 (see Table 4). We show the best-fit models and the cross-correlation measurements in

Figs. 8, 9, and 10. For the lower stellar mass thresholds (10, 10.25, 10.5), we find that the point source (PS, yellow) emission dominates the cross-correlation at small separation ($r < 80$ kpc). In the range ($80 < r < 2$ Mpc), the emission from large halos hosting satellite galaxies dominates (SAT, grey). Finally, on scales beyond that measured here ($r > 2$ Mpc), the two-halo term (green) becomes dominant. Interestingly, there is no scale at which the circumgalactic medium (CGM, purple) dominates. In the range ($20 < r < 200$ kpc), the CGM contributes to more than 10% of the signal. Progressively, with the minimum stellar mass

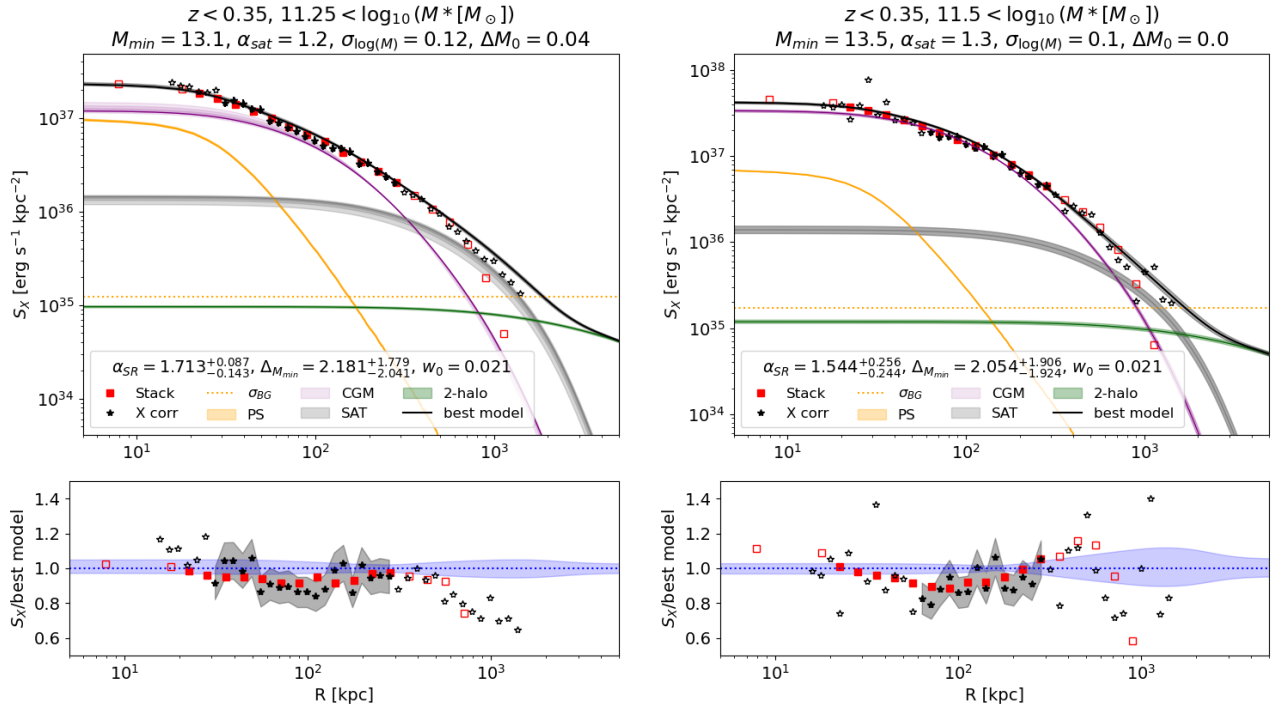


Fig. 10. Fig. 8 continued. Models of the soft X-ray surface brightness profiles for the stellar mass-selected sample with mass $M^* > 1.7 \times 10^{11} M_\odot$ (left, 11.25) and $M^* > 3.1 \times 10^{11} M_\odot$ (11.5, right). The deviations between the model and the data indicate either that the model needs more complexity to account for the data or that the estimator suffers from additional systematic uncertainties. These data sets will constitute an important benchmark for future models. The models indicate that the hot gas emission becomes dominant for these samples.

increasing, the CGM emission increases. For the 10.75 sample, the trend is essentially the same, except in the range 50–60 kpc, where the three components (CGM, PS, SAT) contribute each the same surface brightness. For the mass threshold $M^* > 11$, the CGM is the dominating emission source over the range of 30–200 kpc. Finally, for the mass thresholds 11.25 and (11.5), Fig. 10, the CGM emission dominates over other components until 400 (700) kpc.

5.2.1. Point source emission

The point source emission has a total average luminosity in the soft band (0.5–2 keV) between 3 and 9×10^{40} erg s⁻¹. The luminosity increases with mean stellar mass until $\log_{10}(M^*/M_\odot) = 11$, then stalls and declines (Fig. 11). This result is in disagreement with AGN models where the luminosity correlates with the stellar mass (e.g., Comparat et al. 2019). Indeed, they predict that such a relationship will increase monotonically. This result will constitute a strong benchmark for creating future X-ray AGN models realistically embedded in the galaxy population. It suggests that models sampling the specific accretion rate distribution may perform better in this dimension (e.g., Georgakakis et al. 2019; Powell et al. 2022) and that X-ray AGN activity needs modulation with the environment (Martini et al. 2013; Sabater et al. 2013, 2019; Poggianti et al. 2017; Koulouridis et al. 2024). It opens a new route of exploration for models to jointly predict the entire galaxy population and the average emissions of AGNs in the soft X-rays (e.g., the combination of the UniverseMachine and the Trinity models Behroozi et al. 2019; Zhang et al. 2023). We further discuss this finding in Sect. 6.

5.2.2. CGM emission

The brightness of the model CGM emission correlates with the stellar mass threshold; see Fig. 12. It follows the expectation based on previous measurements indicating that the scaling relation between soft X-ray luminosity and stellar (or halo) mass continues below the group regime (Zhang et al. 2024b, 2025; Popesso et al. 2024b). Since the scaling relation between halo mass and X-ray luminosity is steep, one must interpret these values cautiously. In Fig. 13 (top row of panels), we show the fraction of emission in the profile as a function of its originating halo mass. In the bottom row of panels, we show the posterior scaling relation obtained from the fit (blue shaded area) and, for each halo mass, its contribution to the luminosity obtained when integrating the average model profile. For the $\log_{10}(M^*/M_\odot) > 10$ sample (left column of panels), halos from $10^{11} M_\odot$ to $10^{14} M_\odot$ contribute to the average profile. The most common halo hosting these galaxies ($10^{12} M_\odot$) contributes very little to the average hot gas emission. For the $\log_{10}(M^*/M_\odot) > 11$ sample (right column of panels), halos from $10^{12} M_\odot$ to $10^{14} M_\odot$ contribute to the average profile. The most common halo hosting these galaxies ($10^{13} M_\odot$) contributes more to the average hot gas emission. The model identifies the individual contribution of each dark matter halo to the total luminosity of a cross-correlation measurement. This stresses the importance of considering the complete distribution of halos hosting galaxies when measuring the average luminosity in the presence of a steep slope. The cross-correlation measurements are compatible with a steep scaling relation (slope ~ 1.63). They are compatible with the results of Zhang et al. (2024b, 2025), Popesso et al. (2024b) as well as literature measurements of galaxy clusters as collected in Comparat et al. (2020).

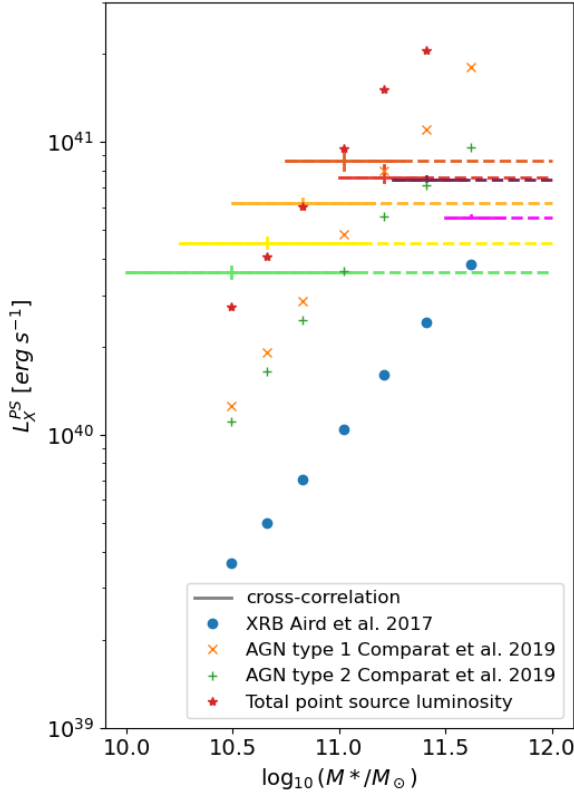


Fig. 11. Comparison of the average point source emission components obtained for each sample as a function of the stellar mass. The horizontal solid line represents the mean and standard deviation of the stellar masses in the samples. The horizontal dashes show the extent of the selection. Different colored crosses represent different samples, identified here by the logarithm of their minimum stellar mass, in solar masses following the color scheme of Fig. 1. The inferred average point source luminosity increases with mean stellar mass from $\log_{10}(M^*/M_{\odot}) = 10.6$ until 11 and stalls, then decreases slightly. The point source luminosities in each sample are consistently more luminous than the predicted XRB luminosity for each sample using Aird et al. (2017). The predicted AGN luminosities from the Comparat et al. (2019) model for type 1 (orange crosses) and type 2 (green pluses) are lower than the observations for the thresholds 10, 10.25, 10.5, 10.75, but they then are higher for the thresholds 11, 11.25 and 11.5. The total predicted luminosity in point sources (red stars) with these models agrees with the observations only for three samples (10.25, 10.5, 10.75 selections) out of seven. These inferred quantities will be used as a benchmark to test upcoming AGN models in the large-scale structure, for example to test mechanisms of X-ray AGN suppression in high halo mass (e.g., Muñoz Rodríguez et al. 2024).

The statements above depend on the formulation of the hot gas profile, the structure of the scaling relation, and their possible variations or evolution with mass and redshift. It shows the importance of embedding detailed models of the hot GCM in the large-scale structure (see e.g., Flender et al. 2017; Comparat et al. 2020; Lau et al. 2025; La Posta et al. 2024; Kéruzoré et al. 2024).

5.2.3. Satellites and two-halo term emission

The emission from the two-halo term dominates on large scales, $R > 3$ Mpc. For the sample $\log_{10}(M^*/M_{\odot}) > 10.5$, the model's two-halo term brightness at 5 Mpc is 4×10^{34} erg $\text{kpc}^{-2} \text{s}^{-1}$ and constitutes $\sim 100\%$ of the model emission. At 1 Mpc, the model two-halo term brightness is less dominant with

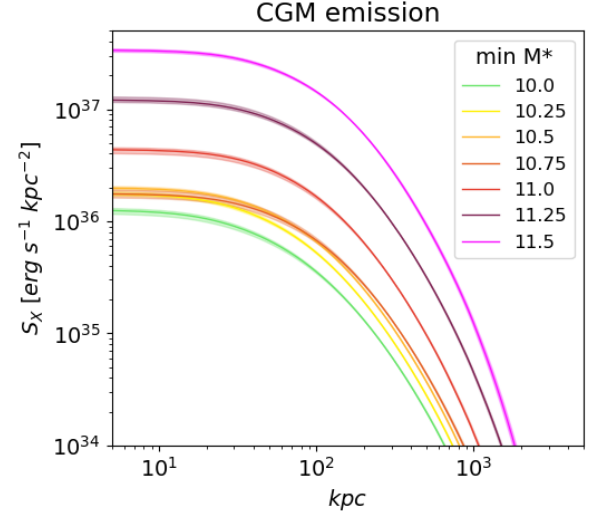


Fig. 12. Comparison of average hot gas surface brightness profiles (cross-correlation between events and central galaxies) inferred with the model for each sample. The differently colored, shaded areas represent different samples, identified here by the logarithm of their minimum stellar mass, in solar masses following the color scheme of Fig. 1. This emission average is made over a large range of halo masses. The average emission increases with the stellar mass threshold. We decompose the average emission as a function of halo mass in Fig. 13 (see top row).

8×10^{34} erg $\text{kpc}^{-2} \text{s}^{-1}$, compared that of the one-halo satellite term at 2×10^{35} erg $\text{kpc}^{-2} \text{s}^{-1}$ (the one-halo central term is much smaller: 5×10^{33} erg $\text{kpc}^{-2} \text{s}^{-1}$). The two-halo term emission is directly related to the large-scale halo bias, defined as the clustering strength of halos compared to that of the underlying dark matter distribution, see Table 3. Naturally, as stellar mass traces halo mass, it also correlates with the stellar mass threshold, see Fig. 14 top panel.

The emission from the cross-correlation between satellite galaxies and hot gas events is nearly identical in every sample (Fig. 14 bottom panel), reflecting the fact that only the satellite galaxies living in clusters contribute to this correlation. Lower mass halos' satellite contributes much less due to the steepness of the mass-luminosity scaling relation. In that regard, masking clusters is a valid approach to discard most of this component.

Importantly, this illustrates how subtracting a constant background value from a surface brightness profile depends on the large-scale structure in which galaxies are embedded (see discussions in Zhang et al. 2024b; Shreeram et al. 2025).

6. Discussion and outlook

The measured galaxy auto-correlation and the galaxy-X-rays cross-correlation are rich in information. With this current combination of surveys, eROSITA + LS10, the statistical uncertainties on these summary statistics are at the percent level (or smaller). The measurement may still suffer from systematic uncertainties of the order of 5% (or even 10% for the higher mass samples). To fully exploit this rich information source, understand its uncertainties and covariance, and gain insights on the connection between the large-scale structure and its dark matter halos, galaxies, hot gas, AGN, and XRB, future efforts must be put into the modeling.

We discuss here two additional sources of information present in this data set, namely: its dependence on the galaxy classification (Sect. 6.1) and on the energy (Sect. 6.2).

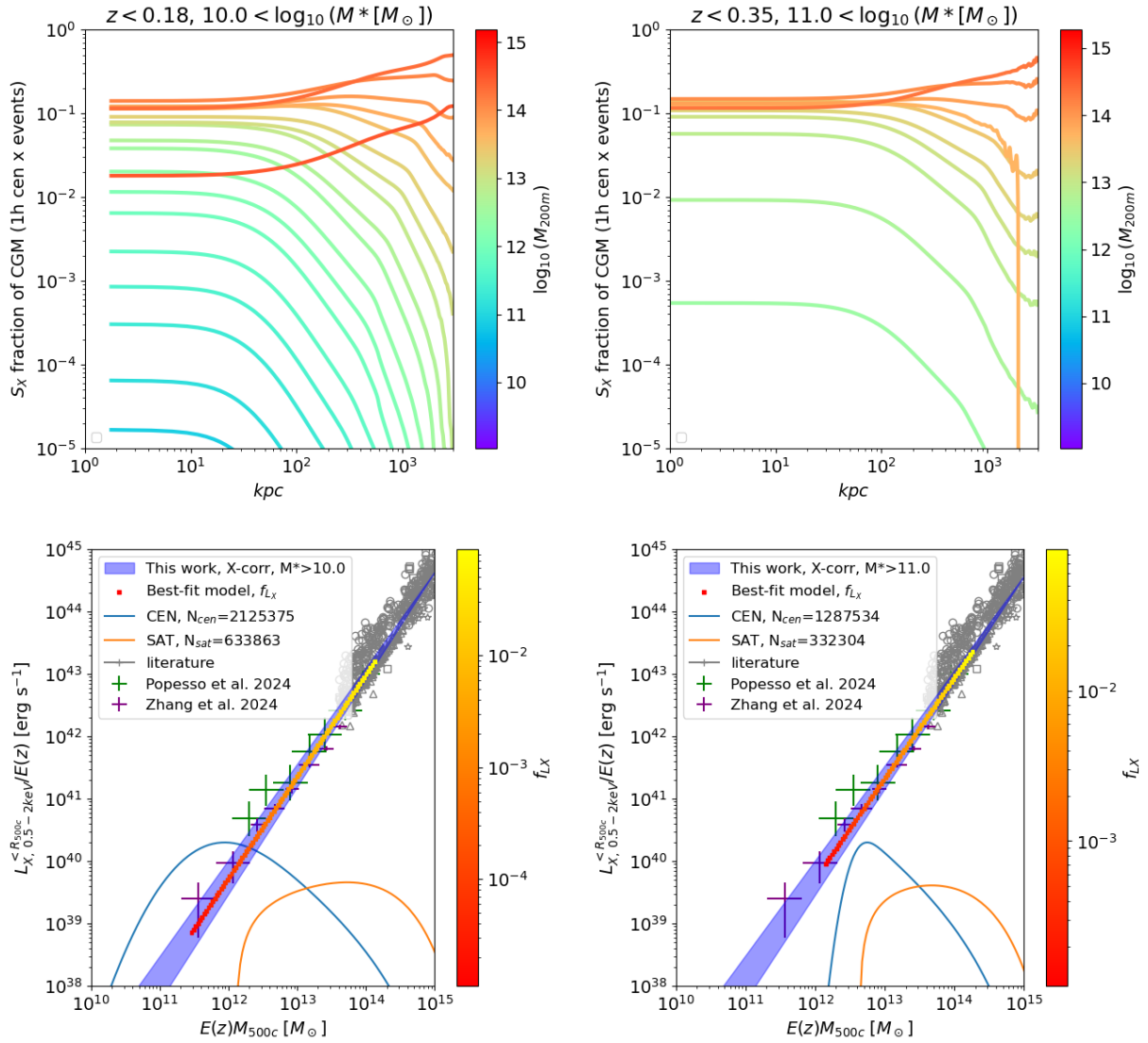


Fig. 13. *Top:* Decomposition in fractions of the average hot gas model surface brightness profile (see Fig. 12) as a function of radius (separation in proper kpc) and halo mass (colorbar) for the samples with $\log_{10}(M^*/M_\odot) > 10$ (left column) and with $\log_{10}(M^*/M_\odot) > 11$ (right column). It shows the different origin of the average emission as a function of separation. The emission at large separation is clearly dominated by higher mass distinct halos. At small separations, smaller distinct halos also contribute to the average signal. *Bottom:* Scaling relation between mass and X-ray luminosity. The posterior on the scaling relation obtained from fitting the cross-correlation (mainly sensitive to its slope) is depicted by the blue shaded area. On the left (right) panel, the width of the contour corresponds to the uncertainty on the scaling relation slope: $\alpha_{SR} = 1.629^{+0.091}_{-0.089}$ ($1.634^{+0.066}_{-0.094}$). The shaded area overlaps with literature measurements shown with crosses and gray symbols (Lovisari et al. 2015, 2020; Mantz et al. 2016; Adami et al. 2018; Schellenberger & Reiprich 2017; Bulbul et al. 2019; Liu et al. 2022; Zhang et al. 2024b; Popesso et al. 2024b). Each galaxy sample is hosted by different populations of dark matter halos. The posterior normalized distribution (multiplied by 2×10^{40} so that it appears in the panel) of halo masses, obtained with the HOD model, hosting central (satellite) galaxies is shown at the bottom of the panel with a blue (orange) solid line. It shows that the constraint on the slope comes mostly from the central (satellite) galaxies at low (high) halo mass. Fitting for a complete sample (including satellite) gives stronger constraints than using central galaxies alone, that are hampered by small statistics at high mass. The integrated X-ray luminosity (one-halo central, see decomposition on the top row of panels) at a given halo mass contributes a fraction (f_{L_X}) to the average luminosity inferred (Table 4, last column) in the cross-correlation fit. This fraction is shown using the red-yellow color bar. On the bottom left panel, the color is limited to 10^{-5} . The galaxies with halo masses smaller than $10^{11} M_\odot$ have almost no contribution to the average luminosity.

Both of these aspects may help in devising relevant future models.

6.1. Dependence on galaxy classification: red-sequence, blue cloud and green valley galaxies

As pointed out by Truong et al. (2021) the CGM and AGN properties may interestingly be correlated to the galaxy's specific star

formation rate. We used a red sequence (RS) model calibrated on clusters of galaxies using redmapper (Rykoff et al. 2014; Kluge et al. 2024) and obtained its mean $g-z$ color (DECam) as a function of redshift, given in Table D.1. Then we selected galaxies as being in the red sequence when they have a $g-z$ color above the mean of the RS minus $c_{RS} = 0.15$ mag with $g_{AB} - z_{AB} > RS(z) - c_{RS}$. By making a comparison to GAMA (Driver et al. 2022), we found that the selection has an 80%

completeness and 66% purity. A completeness of 90% could be reached with $c_{RS} = 0.23$, but would correspond to a lower purity of 62%. Maximizing purity is at the cost of completeness. With this method, the maximum attainable purity is 69% for $c_{RS} = 0.06$, but this corresponds to a relatively low completeness 55%. We selected blue cloud (BC) galaxies, as they are far away from the red sequence, with a $g-z$ color below the mean of the RS minus $c_{BC} = 0.22$ mag with $g_{AB} - z_{AB} < RS(z) - c_{BC}$. This cut entails a completeness (purity) of 80% (94%). We classify the remaining galaxies (in between) as green valley galaxies. Appendix D details the trade-off between completeness and purity.

Figure 15 shows the fraction of emission in the stacks that originate from red-sequence, blue cloud, and green valley galaxies. We see how the profiles differ when selecting red-sequence and blue-cloud galaxies. For the high-mass samples, the emission is entirely dominated by that around red sequence galaxies that reside in groups and clusters. Indeed the quenched fraction steadily increases with stellar mass (e.g., [Ilbert et al. 2013](#); [Muzzin et al. 2013](#)). For the lower mass sample, the decomposition shows interesting trends. The central surface brightness is equally split between the red sequence and the blue cloud. That vastly dominates the outer surface brightness (dominated by satellite emission) around red sequence galaxies. The red sequence galaxies that are not the brightest cluster galaxies (BCG), namely, the low-mass red galaxies, are not in the center of groups and clusters (e.g., [Benson et al. 2007](#); [Moustakas et al. 2013](#)). Therefore, the X-ray emission at their position is lower because the X-ray emission of clusters decreases with radius. Also, the X-ray emission further away from their emission is strong because it (also) includes the region close to the cluster center where the intracluster medium is the brightest. This decomposition can inform a HOD model where the fraction of central and satellite quiescent galaxies is parametrized as a function of stellar and halo mass (e.g., [Zu & Mandelbaum 2016](#); [Behroozi et al. 2019](#); [Tinker 2022](#)).

6.2. Energy dependence of the cross-correlation

The eROSITA telescope measures photons and their energy to about 80 eV precision. For the sample with threshold $\log_{10}(M^*/M_{\odot}) > 11$, we measured the cross-correlation in bins on 100 eV in the range 0.5–2 keV and find genuine differences as a function of scale (Fig. 16). We found that the shape of the relative change of the cross-correlation varies with angular separation. This neatly connects with the physical components inferred by the model (point sources on small scales, hot gas on intermediate scales, and a two-halo term on large scales). The relative change does not follow the ARF, indicating its physical origin. This should be a composite effect. For example, low-energy photons may be dominated by CGM and low-mass halos, high-energy photons from AGNs, and massive halos. The redshift distribution of the sample and k-correction may also play a role here. The cross-correlation at 1 keV is $\sim 20\%$ larger than the broad band, while the one at 0.6 keV is 50% lower. This illustrates the extent of information available in the energy-dependent cross-correlation for interpretation with future models.

6.3. On future observations

In the future and as suggested by [La Posta et al. \(2024\)](#), we anticipate that interesting synergies with lensing to constrain further the link between X-rays, galaxies, and their host dark matter halos (e.g., with the Euclid and LSST experiments

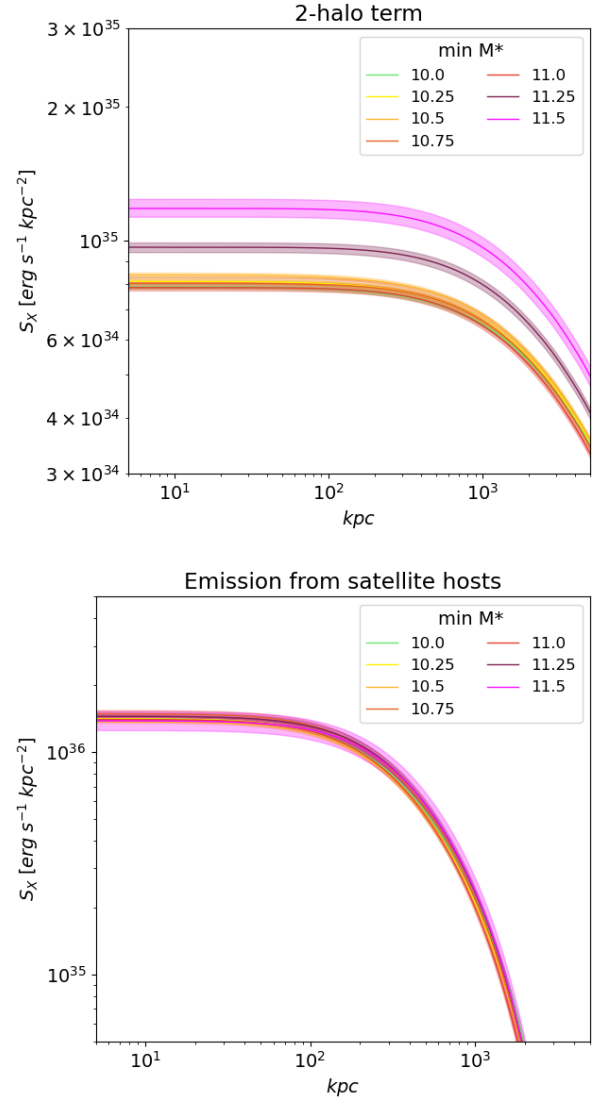


Fig. 14. Comparison of components obtained for each sample: the two-halo term (*top*) and satellite hosts (*bottom*). Differently colored shaded areas represent different samples, identified here by the logarithm of their minimum stellar mass, in solar masses following the color scheme of Fig. 1. The evolution of the two-halo term (satellite) is linked to the evolution of the large-scale halo bias (mean halo mass hosting satellite galaxies), as shown in Table 3, column “b” (\bar{M}_{sat}).

[Euclid Collaboration: Mellier et al. 2025](#); [Ivezić et al. 2019](#)). In addition, synergies with the cross-correlation between the same galaxies and thermal SZ maps will further inform the hot gas profiles (e.g., from ACT, CMB-4, SO, [Das et al. 2023](#); [Abazajian et al. 2016](#); [Ade et al. 2019](#)). Over time, with NewAthena ([Nandra et al. 2013](#); [Cruise et al. 2025](#)) AXIS ([Mushotzky et al. 2019](#)), and HUBS ([Cui et al. 2020](#)), these measurements will be extended to a redshift of $z \sim 1$.

7. Summary

In this paper, we present the construction and validation of stellar mass selected (volume-limited) samples using the Legacy Survey (data release 10) catalogs limited to $r < 19.5$ covering $\sim 16\,800 \text{ deg}^2$ of the extra-galactic sky (Sect. 2, Fig. 1) and companion mock catalogs (Appendix A). The selection of

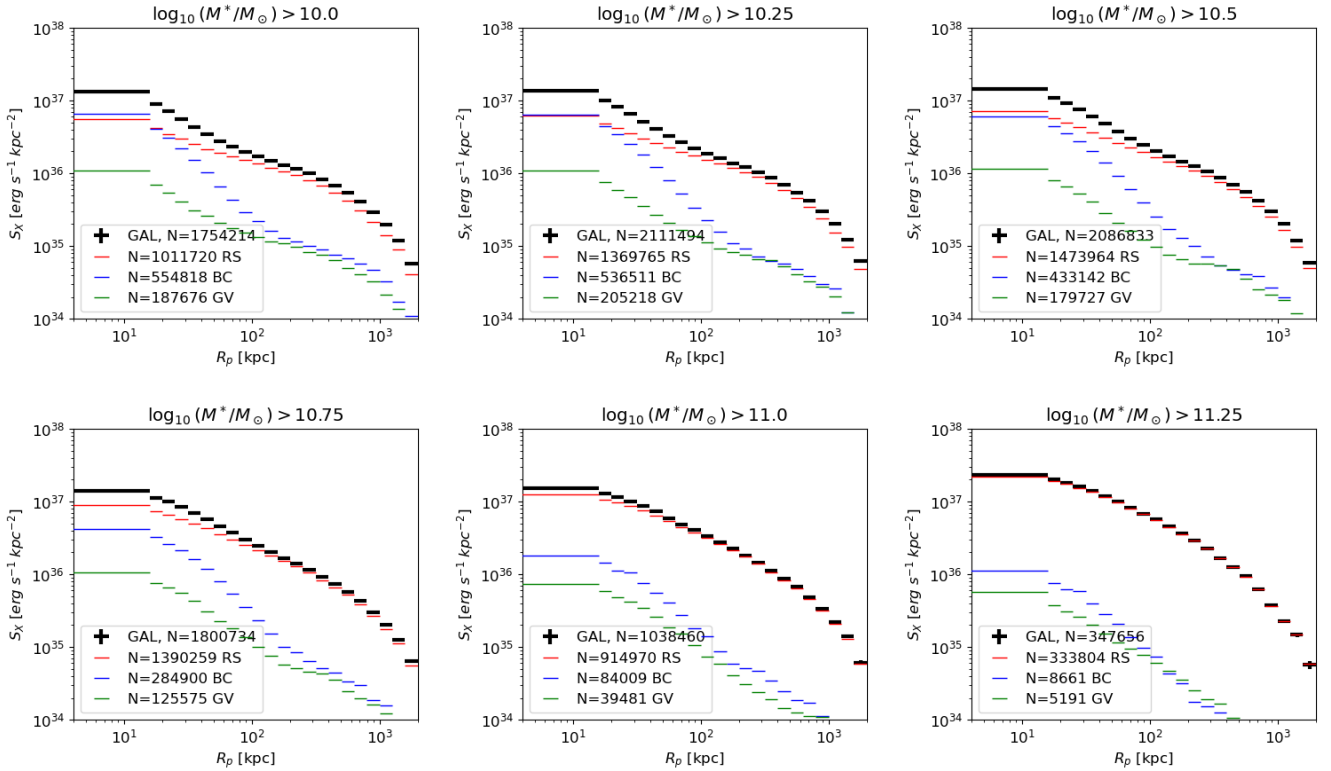


Fig. 15. Stacked surface brightness profiles as a function of radius (proper kpc). The total emission (black lines, identical to that reported in colors in Fig. 5) is decomposed into the fractional contribution of the blue cloud (BC, blue), green valley (GV, green), and red sequence (RS, red) galaxies. The classification into these categories is detailed in Appendix D. This decomposition carries information about quenched fraction as a function of host halo center and halo mass. It may constrain future models of the galaxy-gas-AGN-halo connection as a function of specific star formation rate and in dense environment (e.g., Nishizawa et al. 2018; Aung et al. 2023; Shreeram et al. 2025).

these galaxies was amply controlled (see Appendix B). We measured their two-point correlation function (Sect. 3.1, Fig. 2). The agreement between the simulations and the observations is excellent. The statistical uncertainties are at the percent level, but systematic uncertainties may be as high as 5%. Quantifying the exact systematic uncertainty budget is left for future studies.

We fit a HOD model (Sect. 4) to the galaxy auto-correlation function to retrieve the population of halos hosting them (Sect. 5). We find that the four-parameter HOD model is sufficient to capture the information of interest and reproduce observations (Fig. 6). The deduced stellar-mass-to-halo-mass relation retrieved agrees with literature estimates (Fig. 7). The best-fit parameters obtained (Table 3) are sensible and in agreement with previous studies.

Using this controlled galaxy sample and its HOD model, we then measured and interpreted their cross-correlation, as well as their stacked surface brightness profiles, using the complete field of 0.5–2 keV X-ray photons observed by eROSITA over 2 years of scanning (Sects. 2 and 3). The cross-correlation is measured to unprecedented accuracy thanks to the large area used. Its covariance matrix is mostly diagonal (Fig. 3). With the stacked surface brightness profile, we constructed an alternative estimator of the cross-correlation function and obtained an excellent (better than 5%) agreement for separations between 20–300 kpc (Fig. 4). In this way, we extended the 9000 deg² “FULL photometric” measurement from Zhang et al. (2024a) on the complete eROSITA extra-galactic sky of 13 000 deg². We increased the area by more than 40% to include areas where the eROSITA’s exposure times are longest. We cover the theoretical link between stacked profile and the cross-correlation via a background measured at the

2% percent level precision in detail (see Table 2). The statistical uncertainty on the cross-correlation is unprecedented at the sub-percent level. The systematic uncertainties are estimated to be at the 5% level, except for the higher mass samples, where it reaches the 10% level. We used the complete set of events and a complete set of galaxies, making this measurement extremely stable and robust. These measurements constitute a key benchmark for galaxy evolution models linking galaxies, hot gas, and AGNs to their host dark matter halos.

We wrote a novel extension to the halo model to extract information from the cross-correlation between galaxies and X-rays (Sect. 4). The best-fit models (parameters in Table 4) are presented in Sect. 5 and Figs. 8, 9, and 10. The models account for the observations and permit the decomposition of the observed signal. We obtained interesting posterior on the model components for the X-ray point sources (Fig. 11) and for the hot gas (Figs. 12, 13), the satellite, and two-halo term (Fig. 14). The model posteriors are compatible with previous measurements from Zhang et al. (2024a,b), Popesso et al. (2024b). We obtained a tight 5% constraint on the slope of the mass-luminosity scaling relation over three orders of magnitude. The essential novelty is that we have grasped the complete picture of the hot gas emission within the complete large-scale structure, without any selection biases. This study opens a new venue for writing models that unify galaxies, AGNs, and their hot gas (in optical and X-rays), which can then be constrained by the observations presented here. Finally, we elaborate on the richness of the data presented by stacking red sequence and blue cloud galaxies (Fig. 15), showing the energy dependence of the cross-correlation (Fig. 16). We demonstrate that these additional data are sensitive to the galaxy quiescent fraction.

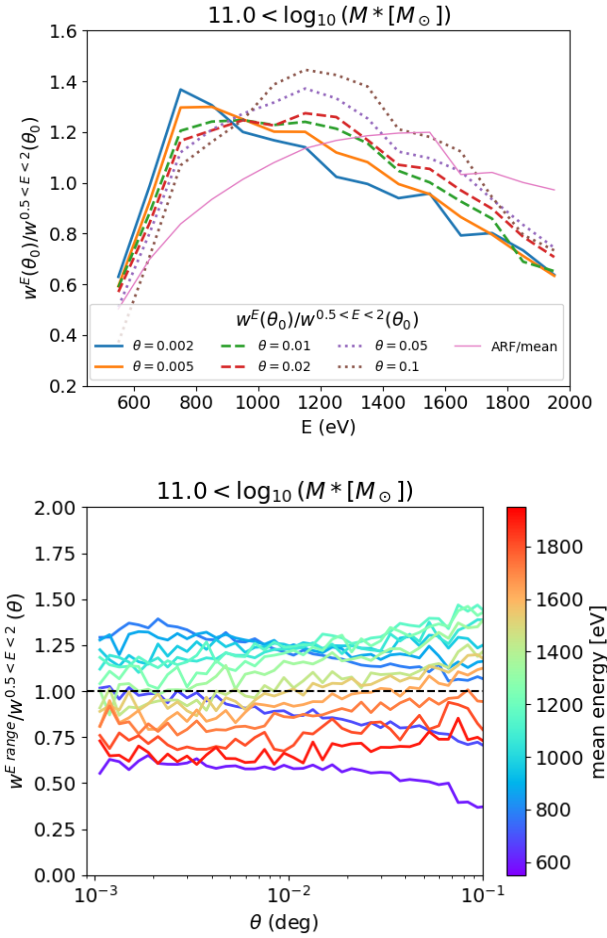


Fig. 16. Illustration of the energy dependence of the galaxy-X-ray cross-correlation with the galaxy sample selected with stellar mass larger than $10^{11} M_\odot$ (see best-fit model of the broad band cross-correlation in Fig. 8 bottom right panel). *Top:* Ratio of the cross-correlations at a fixed set of separations (meaning different sources of emission) as a function of energy. We find that the spectral shape obtained depends on the separation. At small angular separations of 0.002 and 0.005 degree, 30–75 kpc, (blue and orange solid lines), where the point source emission dominates, the ratio to the broad band increases between 600 and 800 eV and decreases thereafter. At large separations (0.05 and 0.1 degree, 700–1400 kpc), where emission coming from satellite and the two-halo dominate (hot gas for large halos), the ratio increases until 1200 eV and decreases thereafter. At intermediate separation (0.01 and 0.02 degree, 150–300 kpc) where the emission comes from half half from the CGM and the satellites, the evolution with energy is in-between the two. Leveraging the energy dependence is key in future models to further ascertain the physical origin of the emitting components. *Bottom:* Same ratio between the cross-correlations in different energy bins as a function of angular separation.

Data availability

The measurements presented are available on [Zenodo](https://zenodo.org). We hand out the relevant data arrays to support the development of future models: (i) galaxy samples (ii) auto and cross correlation functions, and stacked surface brightness profiles.

Acknowledgements. J.C. thanks J. Schaye, B. Oppenheimer, D. Nagai for insightful discussions on this analysis. G.P. acknowledges financial support from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program “Hot Milk” (grant agreement No. 865637) and support from the Framework per l’Attrazione e il Rafforzamento delle Eccel-

lenze (FARE) per la ricerca in Italia (R20L5S39T9). This work is based on data from eROSITA, the soft X-ray instrument aboard SRG, a joint Russian-German science mission supported by the Russian Space Agency (Roskosmos), in the interests of the Russian Academy of Sciences represented by its Space Research Institute (IKI), and the Deutsches Zentrum für Luft- und Raumfahrt (DLR). The SRG spacecraft was built by Lavochkin Association (NPOL) and its subcontractors, and is operated by NPOL with support from the Max Planck Institute for Extraterrestrial Physics (MPE). The development and construction of the eROSITA X-ray instrument was led by MPE, with contributions from the Dr. Karl Remeis Observatory Bamberg & ECAP (FAU Erlangen-Nürnberg), the University of Hamburg Observatory, the Leibniz Institute for Astrophysics Potsdam (AIP), and the Institute for Astronomy and Astrophysics of the University of Tübingen, with the support of DLR and the Max Planck Society. The Argelander Institute for Astronomy of the University of Bonn and the Ludwig Maximilians Universität Munich also participated in the science preparation for eROSITA. The eROSITA data shown here were processed using the eSASS software system developed by the German eROSITA consortium. The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; Proposal ID #2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing-Arizona Sky Survey (BASS; NOAO Prop. ID #2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z -band Legacy Survey (MzLS; Prop. ID #2016A-0453; PI: Arjun Dey). DECaLS, BASS and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF’s NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. Pipeline processing and analyses of the data were supported by NOIRLab and the Lawrence Berkeley National Laboratory (LBNL). The Legacy Surveys project is honored to be permitted to conduct astronomical research on Iolkam Du’ag (Kitt Peak), a mountain with particular significance to the Tohono O’odham Nation. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. LBNL is managed by the Regents of the University of California under contract to the U.S. Department of Energy. This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zurich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l’Espai (IEEC/CSIC), the Institut de Física d’Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF’s NOIRLab, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University. BASS is a key project of the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories of China, the Chinese Academy of Sciences (the Strategic Priority Research Program “The Emergence of Cosmological Structures” Grant # XDB09000000), and the Special Fund for Astronomy from the Ministry of Finance. The BASS is also supported by the External Cooperation Program of Chinese Academy of Sciences (Grant # 114A11KYSB20160057), and Chinese National Natural Science Foundation (Grant # 12120101003, # 11433005). The Legacy Survey team makes use of data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), which is a project of the Jet Propulsion Laboratory/California Institute of Technology. NEOWISE is funded by the National Aeronautics and Space Administration. The Legacy Surveys imaging of the DESI footprint is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH1123, by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; and by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract

No. AST-0950945 to NOAO. The Photometric Redshifts for the Legacy Surveys (PRLS) catalog used in this paper was produced thanks to funding from the U.S. Department of Energy Office of Science, Office of High Energy Physics via grant DE-SC0007914.

References

- Abazajian, K. N., Adshead, P., Ahmed, Z., et al. 2016, ArXiv e-prints [arXiv:1610.02743]
- Adami, C., Giles, P., Koulouridis, E., et al. 2018, *A&A*, **620**, A5
- Ade, P., Aguirre, J., Ahmed, Z., et al. 2019, *JCAP*, **2019**, 056
- Aird, J., Coil, A. L., Georgakakis, A., et al. 2015, *MNRAS*, **451**, 1892
- Aird, J., Coil, A. L., & Georgakakis, A. 2017, *MNRAS*, **465**, 3390
- Allevato, V., Shankar, F., Marsden, C., et al. 2021, *ApJ*, **916**, 34
- Anderson, M. E., Gaspari, M., White, S. D. M., Wang, W., & Dai, X. 2015, *MNRAS*, **449**, 3806
- Asgari, M., Mead, A. J., & Heymans, C. 2023, *Open J. Astrophys.*, **6**, 39
- Aung, H., Nagai, D., Klypin, A., et al. 2023, *MNRAS*, **519**, 1648
- Bahar, Y. E., Bulbul, E., Ghirardini, V., et al. 2024, *A&A*, **691**, A188
- Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013, *ApJ*, **762**, 109
- Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, *MNRAS*, **488**, 3143
- Benson, A. J., Džanović, D., Frenk, C. S., & Sharples, R. 2007, *MNRAS*, **379**, 841
- Berlfein, F., Mandelbaum, R., Dodelson, S., & Schafer, C. 2024, *MNRAS*, **531**, 4954
- Biffi, V., Dolag, K., & Merloni, A. 2018, *MNRAS*, **481**, 2213
- Brunner, H., Liu, T., Lamer, G., et al. 2022, *A&A*, **661**, A1
- Buchner, J. 2016, *Stat. Comput.*, **26**, 383
- Buchner, J. 2019, *PASP*, **131**, 108005
- Buchner, J. 2021, *J. Open Source Software*, **6**, 3001
- Bulbul, E., Chiu, I. N., Mohr, J. J., et al. 2019, *ApJ*, **871**, 50
- Bulbul, E., Liu, A., Kluge, M., et al. 2024, *A&A*, **685**, A106
- Chiu, I. N., Ghirardini, V., Liu, A., et al. 2022, *A&A*, **661**, A11
- Comparat, J., Merloni, A., Salvato, M., et al. 2019, *MNRAS*, **487**, 2005
- Comparat, J., Eckert, D., Finoguenov, A., et al. 2020, *Open J. Astrophys.*, **3**, 13
- Comparat, J., Truong, N., Merloni, A., et al. 2022, *A&A*, **666**, A156
- Comparat, J., Luo, W., Merloni, A., et al. 2023, *A&A*, **673**, A122
- Contreras, S., Angulo, R. E., & Zennaro, M. 2021, *MNRAS*, **504**, 5205
- Cooray, A., & Sheth, R. 2002, *Phys. Rep.*, **372**, 1
- Cruise, M., Guainazzi, M., Aird, J., et al. 2025, *Nat. Astron.*, **9**, 36
- Cui, W., Chen, L.-B., Gao, B., et al. 2020, *J. Low Temp. Phys.*, **199**, 502
- Das, S., Chiang, Y.-K., & Mathur, S. 2023, *ApJ*, **951**, 125
- Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, *MNRAS*, **486**, 2827
- Davis, M., & Peebles, P. J. E. 1983, *ApJ*, **267**, 465
- de Jong, R. 2011, *The Messenger*, **145**, 14
- de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, *The Messenger*, **175**, 3
- Delubac, T., Raichoor, A., Comparat, J., et al. 2017, *MNRAS*, **465**, 1831
- DESI Collaboration (Aghamousa, A., et al.) 2016, ArXiv e-prints [arXiv:1611.00036]
- Dey, A., Schlegel, D. J., Lang, D., et al. 2019, *AJ*, **157**, 168
- Diemer, B. 2018, *ApJS*, **239**, 35
- Driver, S. P., & Robotham, A. S. G. 2010, *MNRAS*, **407**, 2131
- Driver, S. P., Bellstedt, S., Robotham, A. S. G., et al. 2022, *MNRAS*, **513**, 439
- Duffy, A. R., Schaye, J., Kay, S. T., & Dalla Vecchia, C. 2008, *MNRAS*, **390**, L64
- Eckert, D., Ghirardini, V., Ettori, S., et al. 2019, *A&A*, **621**, A40
- Ettori, S., Ghirardini, V., Eckert, D., et al. 2019, *A&A*, **621**, A39
- Euclid Collaboration (Mellier, Y., et al.) 2025, *A&A*, **697**, A1
- Farrow, D. J., Cole, S., Norberg, P., et al. 2015, *MNRAS*, **454**, 2120
- Finoguenov, A., Merloni, A., Comparat, J., et al. 2019, *The Messenger*, **175**, 39
- Flender, S., Nagai, D., & McDonald, M. 2017, *ApJ*, **837**, 124
- Freund, S., Czesla, S., Predehl, P., et al. 2024, *A&A*, **684**, A121
- Georgakakis, A., Aird, J., Schulze, A., et al. 2017, *MNRAS*, **471**, 1976
- Georgakakis, A., Comparat, J., Merloni, A., et al. 2019, *MNRAS*, **487**, 275
- Hahn, C., Wilson, M. J., Ruiz-Macias, O., et al. 2023, *AJ*, **165**, 253
- Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, *ApJS*, **163**, 1
- Hopkins, P. F., Quataert, E., Ponnada, S. B., & Silich, E. 2025, ArXiv e-prints [arXiv:2501.18696]
- Ider Chitham, J., Comparat, J., Finoguenov, A., et al. 2020, *MNRAS*, **499**, 4768
- Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, *A&A*, **457**, 841
- Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, *A&A*, **556**, A55
- Ishiyama, T., Prada, F., Klypin, A. A., et al. 2021, *MNRAS*, **506**, 4210
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, **873**, 111
- Kaminsky, A., Cappelluti, N., Hasinger, G., et al. 2025, ArXiv e-prints [arXiv:2502.09705]
- Kéruszoré, F., Bleem, L. E., Frontiere, N., et al. 2024, *Open J. Astrophys.*, **7**, 116
- Kluge, M., Comparat, J., Liu, A., et al. 2024, *A&A*, **688**, A210
- Kong, H., Burleigh, K. J., Ross, A., et al. 2020, *MNRAS*, **499**, 3943
- Koulouridis, E., Gkini, A., & Drigga, E. 2024, *A&A*, **684**, A111
- Kyritsis, E., Zezas, A., Haberl, F., et al. 2025, *A&A*, **694**, A128
- La Posta, A., Alonso, D., Chisari, N. E., Ferreira, T., & García-García, C. 2024, ArXiv e-prints [arXiv:2412.12081]
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, **412**, 64
- Lau, E. T., Nagai, D., Bogdán, , et al. 2024, ArXiv e-prints [arXiv:2412.04559]
- Lau, E. T., Nagai, D., Farahi, A., et al. 2025, *ApJ*, **980**, 122
- Leauthaud, Jr., A., Benson, A., Civano, F., et al. 2015, *MNRAS*, **446**, 1874
- Lehmer, B. D., Basu-Zych, A. R., Mineo, S., et al. 2016, *ApJ*, **825**, 7
- Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, **538**, 473
- Limber, D. N. 1953, *ApJ*, **117**, 134
- Liu, A., Bulbul, E., Ghirardini, V., et al. 2022, *A&A*, **661**, A2
- Liu, A., Bulbul, E., Kluge, M., et al. 2024, *A&A*, **683**, A130
- Locatelli, N., Ponti, G., Zheng, X., et al. 2024, *A&A*, **681**, A78
- Lovisari, L., Reiprich, T. H., & Schellenberger, G. 2015, *A&A*, **573**, A118
- Lovisari, L., Schellenberger, G., Sereno, M., et al. 2020, *ApJ*, **892**, 102
- Lovisari, L., Ettori, S., Gaspari, M., & Giles, P. A. 2021, *Universe*, **7**, 139
- Mantz, A. B., Allen, S. W., Morris, R. G., & Schmidt, R. W. 2016, *MNRAS*, **456**, 4020
- Marini, I., Popesso, P., Lamer, G., et al. 2024, *A&A*, **689**, A7
- Martini, P., Miller, E. D., Brodwin, M., et al. 2013, *ApJ*, **768**, 1
- McAlpine, S., Helly, J. C., Schaller, M., et al. 2016, *Astron. Comput.*, **15**, 72
- Mead, A. J., & Asgari, M. 2023, Astrophysics Source Code Library [record ascl:2307.025]
- Merloni, A., Predehl, P., Becker, W., et al. 2012, ArXiv e-prints [arXiv:1209.3114]
- Merloni, A., Lamer, G., Liu, T., et al. 2024, *A&A*, **682**, A34
- More, S., Miyatake, H., Mandelbaum, R., et al. 2015, *ApJ*, **806**, 2
- Moustakas, J., Coil, A. L., Aird, J., et al. 2013, *ApJ*, **767**, 50
- Moustakas, J., Lang, D., Dey, A., et al. 2023, *ApJS*, **269**, 3
- Muñoz Rodríguez, I., Georgakakis, A., Shankar, F., et al. 2024, *MNRAS*, **532**, 336
- Murray, S. G., Diemer, B., Chen, Z., et al. 2021, *Astron. Comput.*, **36**, 100487
- Mushotzky, R. F., Aird, J., Barger, A. J., et al. 2019, ArXiv e-prints [arXiv:1903.04083]
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJ*, **777**, 18
- Myers, A. D., Moustakas, J., Bailey, S., et al. 2023, *AJ*, **165**, 50
- Nandra, K., Barret, D., Barcons, X., et al. 2013, ArXiv e-prints [arXiv:1306.2307]
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, **490**, 493
- Neumann, D. M., & Arnaud, M. 1999, *A&A*, **348**, 711
- Newsam, A. M., McHardy, I. M., Jones, L. R., & Mason, K. O. 1999, *MNRAS*, **310**, 255
- Nishimichi, T., Takada, M., Takahashi, R., et al. 2019, *ApJ*, **884**, 29
- Nishizawa, A. J., Oguri, M., Oogi, T., et al. 2018, *PASJ*, **70**, S24
- Oppenheimer, B. D., & Schaye, J. 2013, *MNRAS*, **434**, 1043
- Oren, Y., Sternberg, A., McKee, C. F., Faerman, Y., & Genel, S. 2024, *ApJ*, **974**, 291
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, *MNRAS*, **475**, 648
- Planck Collaboration VI. 2020, *A&A*, **641**, A6
- Poggianti, B. M., Jaffé, Y. L., Moretti, A., et al. 2017, *Nature*, **548**, 304
- Ponti, G., Zheng, X., Locatelli, N., et al. 2023a, *A&A*, **674**, A195
- Ponti, G., Sanders, J. S., Locatelli, N., et al. 2023b, *A&A*, **670**, A99
- Pop, A. R., Hernquist, L., Nagai, D., et al. 2022, ArXiv e-prints [arXiv:2205.11528]
- Popesso, P., Biviano, A., Bulbul, E., et al. 2024a, *MNRAS*, **527**, 895
- Popesso, P., Marini, I., Dolag, K., et al. 2024b, *A&A*, submitted [arXiv:2411.17120]
- Powell, M. C., Allen, S. W., Caglar, T., et al. 2022, *ApJ*, **938**, 77
- Powell, M. C., Krumpe, M., Coil, A., & Miyaji, T. 2024, *A&A*, **686**, A57
- Pratt, G. W., Croston, J. H., Arnaud, M., & Böhringer, H. 2009, *A&A*, **498**, 361
- Predehl, P., Sunyaev, R. A., Becker, W., et al. 2020, *Nature*, **588**, 227
- Predehl, P., Andritschke, R., Arefiev, V., et al. 2021, *A&A*, **647**, A1
- Refregier, A., Helfand, D. J., & McMahon, R. G. 1997, *ApJ*, **477**, 58
- Robotham, A. S. G., Norberg, P., Driver, S. P., et al. 2011, *MNRAS*, **416**, 2640
- Ross, A. J., Bautista, J., Tojeiro, R., et al. 2020, *MNRAS*, **498**, 2354
- Rykoff, E. S., Rozo, E., Busha, M. T., et al. 2014, *ApJ*, **785**, 104
- Sabater, J., Best, P. N., & Argudo-Fernández, M. 2013, *MNRAS*, **430**, 638
- Sabater, J., Best, P. N., Hardcastle, M. J., et al. 2019, *A&A*, **622**, A17
- Sanders, J. S., Fabian, A. C., Russell, H. R., & Walker, S. A. 2018, *MNRAS*, **474**, 1065
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, **446**, 521
- Schaye, J., Kugel, R., Schaller, M., et al. 2023, *MNRAS*, **526**, 4978
- Schellenberger, G., & Reiprich, T. H. 2017, *MNRAS*, **469**, 3738

- Schneider, P. C., Freund, S., Czesla, S., et al. 2022, [A&A](#), **661**, A6
- Schwöpe, A. D., Knauff, K., Kurpas, J., et al. 2024a, [A&A](#), **690**, A243
- Schwöpe, A., Kurpas, J., Baecke, P., et al. 2024b, [A&A](#), **686**, A110
- Seppi, R., Comparat, J., Bulbul, E., et al. 2022, [A&A](#), **665**, A78
- Seppi, R., Comparat, J., Ghirardini, V., et al. 2024, [A&A](#), **686**, A196
- Shreeram, S., Comparat, J., Merloni, A., et al. 2025, [A&A](#), **697**, A22
- Singh, P., Lau, E. T., Faerman, Y., Stern, J., & Nagai, D. 2024, [MNRAS](#), **532**, 3222
- Sinha, M., & Garrison, L. H. 2020, [MNRAS](#), **491**, 3022
- Soltan, A. M., Hasinger, G., Egger, R., Snowden, S., & Truemper, J. 1997, [A&A](#), **320**, 705
- Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, [AJ](#), **124**, 1810
- Sunyaev, R., Arefiev, V., Babyshkin, V., et al. 2021, [A&A](#), **656**, A132
- Tempel, E., Kipper, R., Tamm, A., et al. 2016, [A&A](#), **588**, A14
- Tempel, E., Kruuse, M., Kipper, R., et al. 2018, [A&A](#), **618**, A81
- Tinker, J. L. 2021, [ApJ](#), **923**, 154
- Tinker, J. L. 2022, [AJ](#), **163**, 126
- Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, [ApJ](#), **724**, 878
- Truong, N., Pillepich, A., Nelson, D., Werner, N., & Hernquist, L. 2021, [MNRAS](#), **508**, 1563
- van den Bosch, F. C., More, S., Cacciato, M., Mo, H., & Yang, X. 2013, [MNRAS](#), **430**, 725
- Vikhlinin, A., Kravtsov, A., Forman, W., et al. 2006, [ApJ](#), **640**, 691
- Vladutescu-Zopp, S., Biffi, V., & Dolag, K. 2023, [A&A](#), **669**, A34
- Waddell, S. G. H., Buchner, J., Nandra, K., et al. 2024a, [A&A](#), submitted [arXiv:2401.17306]
- Waddell, S. G. H., Nandra, K., Buchner, J., et al. 2024b, [A&A](#), **690**, A132
- Yang, X., Mo, H. J., van den Bosch, F. C., & Jing, Y. P. 2005, [MNRAS](#), **356**, 1293
- Yeung, M. C. H., Ponti, G., Freyberg, M. J., et al. 2024, [A&A](#), **690**, A399
- Yuan, S., Garrison, L. H., Hadzhiyska, B., Bose, S., & Eisenstein, D. J. 2022, [MNRAS](#), **510**, 3301
- Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2005, [ApJ](#), **630**, 1
- Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, [ApJ](#), **736**, 59
- Zhang, Z., Wang, H., Luo, W., et al. 2021, [A&A](#), **650**, A155
- Zhang, H., Behroozi, P., Volonteri, M., et al. 2023, [MNRAS](#), **518**, 2123
- Zhang, Y., Comparat, J., Ponti, G., et al. 2024a, [A&A](#), **690**, A267
- Zhang, Y., Comparat, J., Ponti, G., et al. 2024b, [A&A](#), **690**, A268
- Zhang, H.-S., Ponti, G., Carretti, E., et al. 2024c, [Nat. Astron.](#), **8**, 1416
- Zhang, Y., Comparat, J., Ponti, G., et al. 2025, [A&A](#), **693**, A197
- Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, [ApJ](#), **633**, 791
- Zheng, Z., Coil, A. L., & Zehavi, I. 2007, [ApJ](#), **667**, 760
- Zheng, X., Ponti, G., Freyberg, M., et al. 2024a, [A&A](#), **681**, A77
- Zheng, X., Ponti, G., Locatelli, N., et al. 2024b, [A&A](#), **689**, A328
- Zhou, R., Newman, J. A., Mao, Y.-Y., et al. 2021, [MNRAS](#), **501**, 3309
- Zou, H., Gao, J., Zhou, X., & Kong, X. 2019, [ApJS](#), **242**, 8
- Zu, Y., & Mandelbaum, R. 2015, [MNRAS](#), **454**, 1161
- Zu, Y., & Mandelbaum, R. 2016, [MNRAS](#), **457**, 4360

Appendix A: Galaxy mock catalog

With mock catalogs, we estimate the impact of possible systematic effects on the clustering measurement. We create mock catalogs using the Uchuu simulation (Ishiyama et al. 2021) and the UniverseMachine model (Behroozi et al. 2019; Aung et al. 2023). We create eight light cones following Comparat et al. (2020) that each spans 1/8th of the sky (by setting the observer at each corner of the simulated box).

A.1. Reference mock catalog

We generate a reference mock catalog for each galaxy sample, applying the same selection as in the data but on the predicted galaxy population (i.e., using redshift and stellar mass). Since the UniverseMachine model is fitted upon different stellar mass functions, the predicted galaxy density as a function of stellar mass function is in excellent agreement with observations (Fig. B.1, B.2). Notably, the correlation functions predicted for each galaxy sample are close to the observed ones, see Fig. C.1. The fact that the measurements are close to the prediction of the UniverseMachine model (Fig. C.1) means that no large systematics have been missed in the measurement procedure.

A.2. Impact of the photometric redshift on the correlation function

We create mock catalogs where we degrade the redshift precision according to the performance of the photometric redshifts. We find that the quality of the redshift considered here does not cause significant distortions in the shape of the correlation function ($w_p^{\pi_{\max}=100}(r_p)$) in the range $0.01 < r_p [\text{Mpc}/h] < 30$. The redshift uncertainty causes a constant loss of amplitude in the correlation function. For each sample (selections in redshift and stellar mass from the complete selection), we estimate the scatter of the photometric redshift. We find that the performances for each sub-sample are better than for the complete sample with a scatter between 1 and 2 percent. Indeed, after discarding low mass, faint sources, and higher redshift sources, it is expected to perform better than the full sample (3 percent scatter). We find an expected loss in the correlation function amplitude between 2 and 7 percent. We use the tabulated value and apply a correction to the measurement for each sample (see $f_{\text{obs}} \in [0.93, 0.98]$ in Fig. C.2 bottom rows of panels).

A.3. Impact of redshift space distortion

With the mock catalog, we fit a model on the $w_p(r_p)$ ratio estimated in real space and redshift space. With that, we correct the observation from this shape bending of the signal. Since we are integrating up to $\pi_{\max} = 100 h^{-1} \text{Mpc}$ on the line of sight, this is a minor correction for the maximum distance we fit to ($\log_{10}(20-30 \text{ Mpc}/h) = 1.6$). The correction would be more prominent for smaller π_{\max} values. In van den Bosch et al. (2013), they typically amount to 2% (4, 10, 14) at $R=4$ (10, 25, 31) Mpc/h . We find similar values here (Fig. C.3). Overall, with the mock catalogs, we derive a theoretical budget of a few percent corrections on each correlation function.

Appendix B: Density of galaxies

This appendix provides additional (large) figures and Tables related to the galaxy samples.

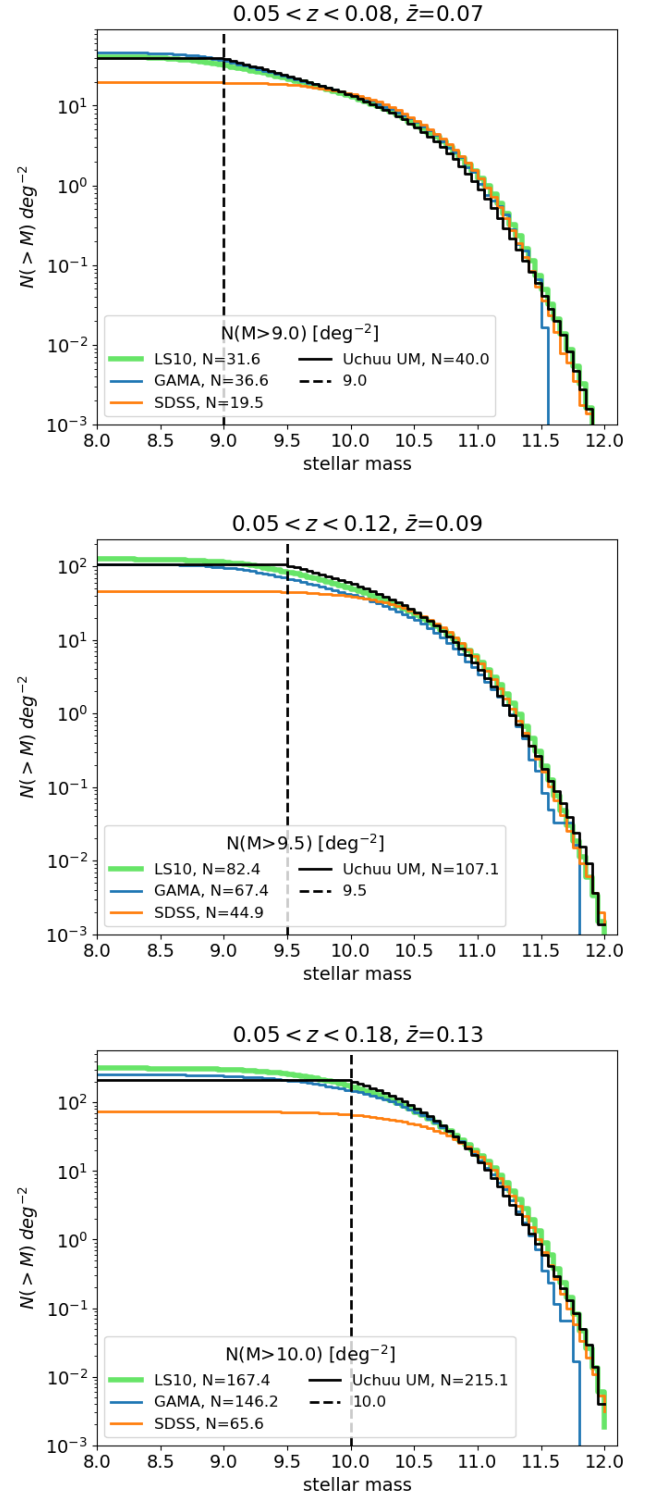


Fig. B.1. Cumulative number density of galaxies per square degree as a function of stellar mass.

Appendix C: Correlation functions

This appendix provides additional information, figures, and tables related to the correlation functions.

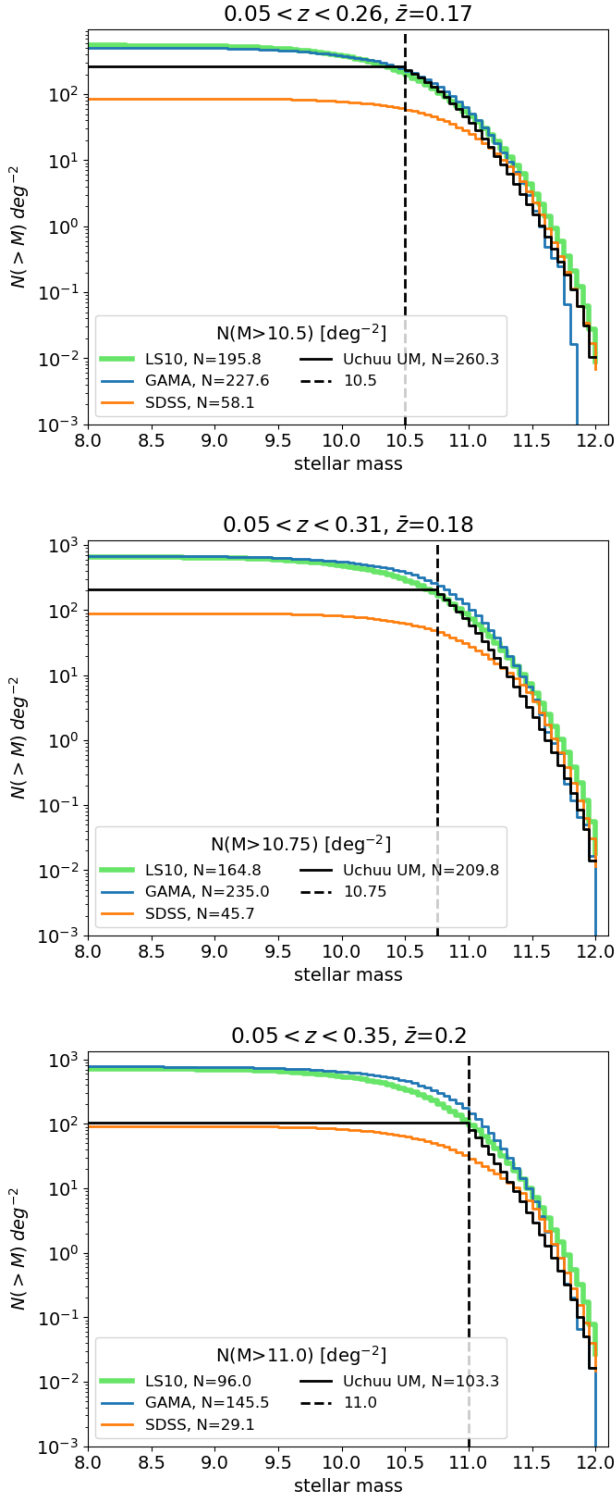


Fig. B.1. Continued.

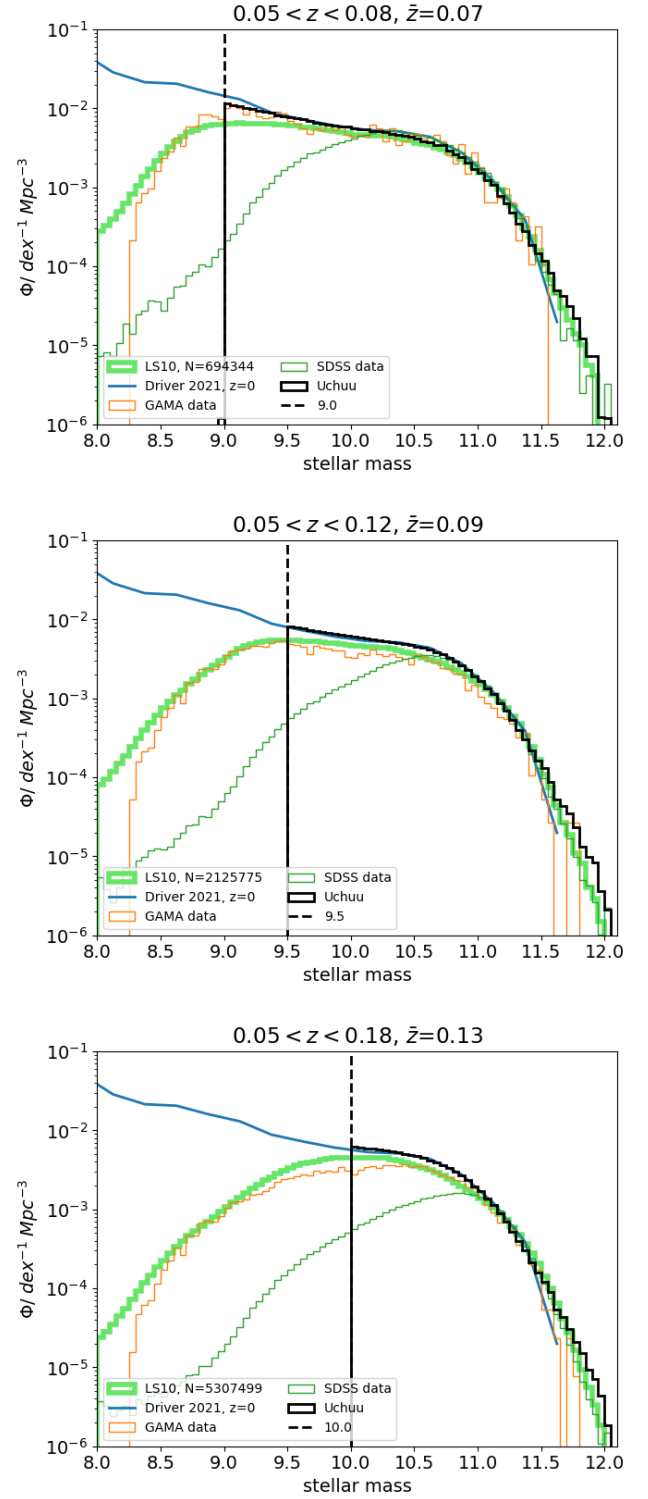


Fig. B.2. Observed number density of galaxies per cubic Mpc as a function of stellar mass.

C.1. Auto-correlation, field selection

We use the X-ray event auto-correlation to determine where we measure the cross-correlation. We measure the auto-correlation of events on each eROSITA field (tiles of $3.6 \times 3.6 \text{ deg}^2$). The obtained distribution of correlation functions is generally well-behaved, albeit with several fields having outlying correlation functions. Indeed, due to bright Galactic foregrounds (resolved or unresolved, e.g., the eROSITA Bubbles (Predehl et al. 2020;

Zhang et al. 2024c), the angular correlation function may have a high power on degree scales. Most of these (~ 50 out of ~ 1600) fields are at low galactic latitudes where our galaxy catalog does not extend. In these cases, the fraction of the correlation function due to point sources or diffuse emission is genuinely interesting but beyond the scope of this paper. In addition, a few outlier fields follow a flaring pattern that was not completely removed in a scan, so we remove these fields from the analysis.

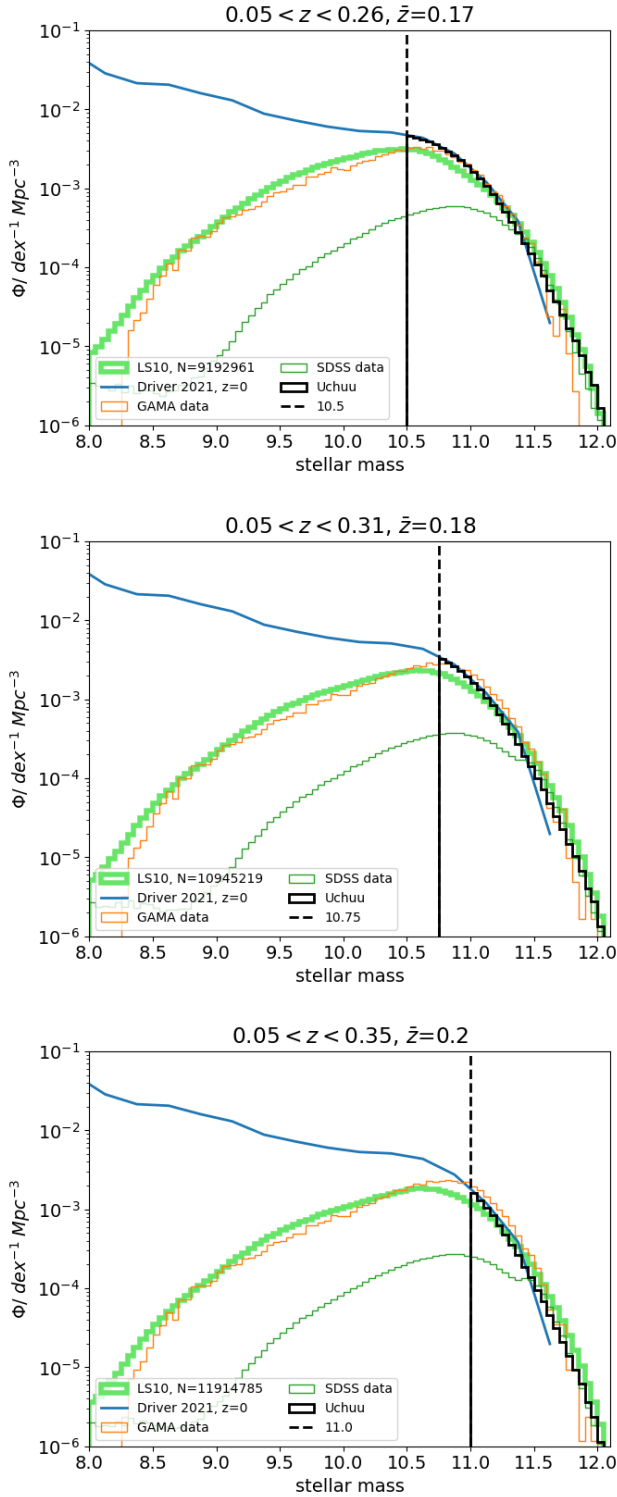


Fig. B.2. Continued.

Cross-correlation, field selection

First, we remove fields that have a correlation function amplitude on large scales (in the two-halo term) at $\theta \sim 0.6^\circ$ (corresponding to 6 (10) comoving Mpc at $z=0.15$ (0.25)) below (or above) the 1st (99th) percentile of the distribution of $w(\theta \sim 0.6^\circ)$, corresponding heuristically to a 2.4σ clipping. For a few more fields, the degree scale cross-correlation function is still a strong positive outlier, which we remove. For example, measuring $w(\theta \sim$

Table B.1. Number density of galaxies selected in LS10 compared to GAMA as a function of r -band magnitude.

r mag		density [deg^{-2}]		ratio
min	max	GAMA	LS10	
16.0	16.5	13.8	14.8	0.929
16.5	17.0	25.2	25.6	0.987
17.0	17.5	51.4	52.6	0.978
17.5	18.0	98.8	99.8	0.99
18.0	18.5	192.9	188.0	1.026
18.5	19.0	359.5	355.0	1.013
19.0	19.5	627.3	629.3	0.997

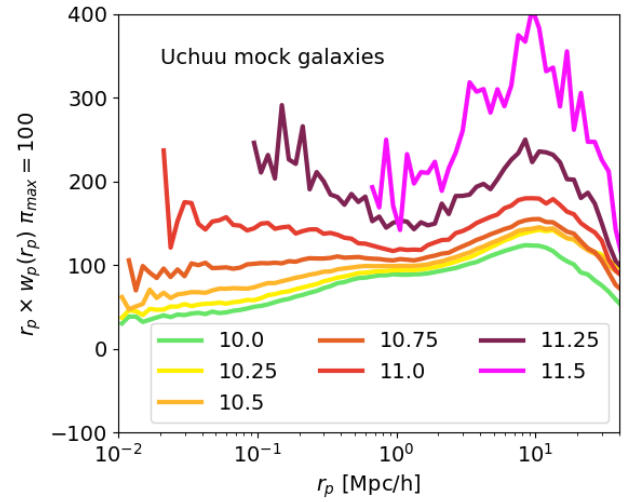


Fig. C.1. Projected galaxy auto-correlation function ($w_p(r_p)$) predicted from the UCHUU mock catalogs on $\sim 5,000 \text{ deg}^2$.

$0.6^\circ) > 0.15$ when the expectation is $w(\theta \sim 0.6^\circ) \sim 2 \times 10^{-3}$ indicates a contamination by stars. These fields are at the edge of the survey, where the stellar density or extinction is high. We remove these fields from the analysis. A cut of the survey footprint using extinction and stellar density is equivalent.

Appendix D: Selection of SF and QU

In the Lephare runs made for this analysis, the SFR values inferred are (very) noisy, with only three or four broad bands. They are not reliable enough to measure sSFR and split the sample into a set of quiescent and star-forming galaxies. Instead, we use a red-sequence model and corresponding colors (IDER Chitham et al. 2020; Kluge et al. 2024). The typical scatter in the red sequence in the LS10 is 0.1 mag.

We define two boundaries in color c_{RS} (c_{BC}) to select red sequence (blue cloud) galaxies as a function of redshift: $g_{AB} - z_{AB} > RS(z) - c_{RS}$ ($g_{AB} - z_{AB} < RS(z) - c_{BC}$). In between, in the green valley, we have a mix of the two populations.

We use the GAMA data, split into SF and QU with a \log_{10} sSFR cut at -11 to optimize the boundary (best values for c_{RS} , c_{BC}) and to estimate completeness, purity and contamination of the obtained samples. In the GAMA-DR10 matched file, there are 46815 QU and 103160 SF galaxies with stellar mass larger than 9 . Table D.2 shows how the completeness and purity evolve as a function of the threshold value.

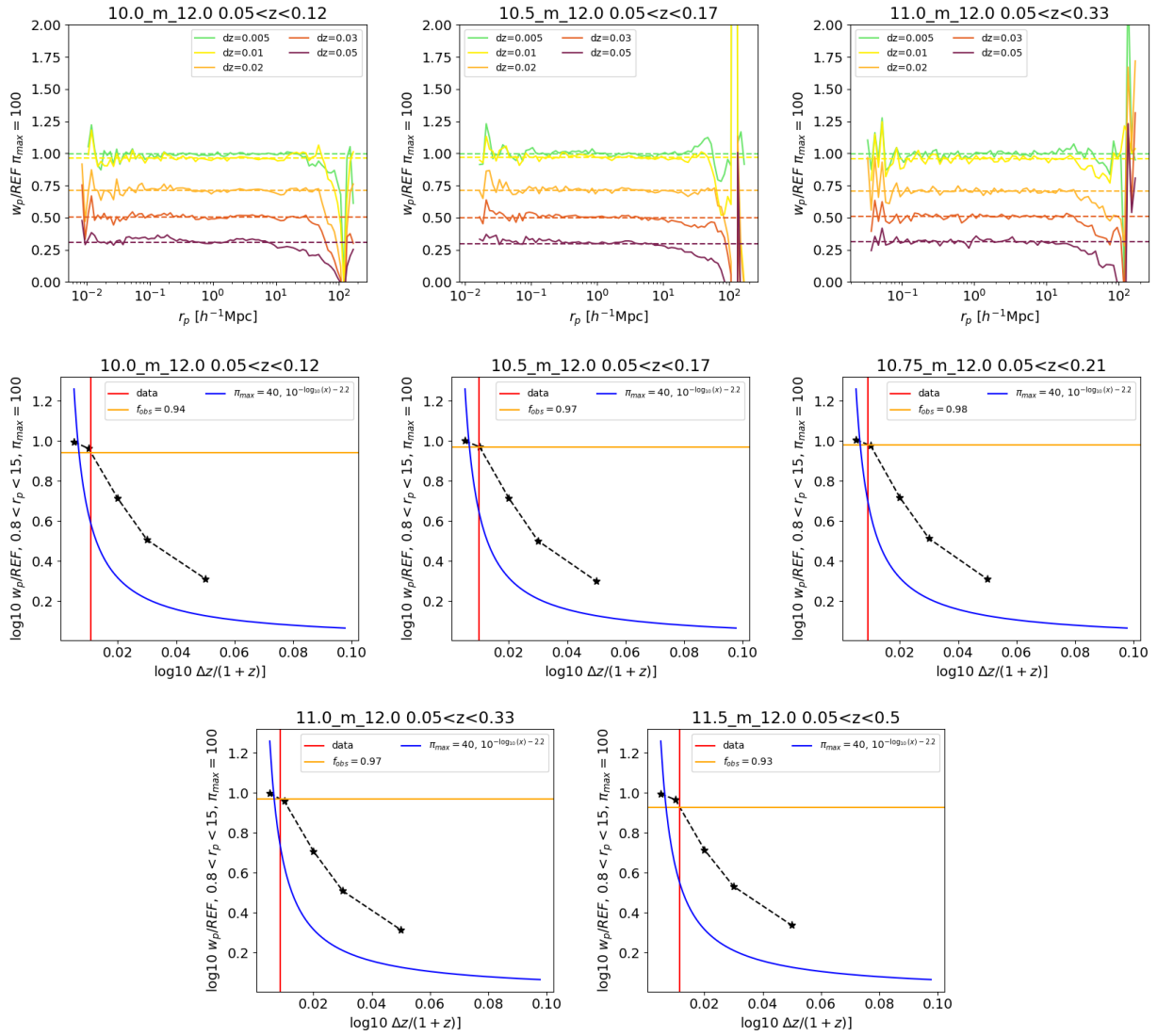


Fig. C.2. Bias in amplitude when computing w_{pr} in the presence of photometric errors for each sample considered here.

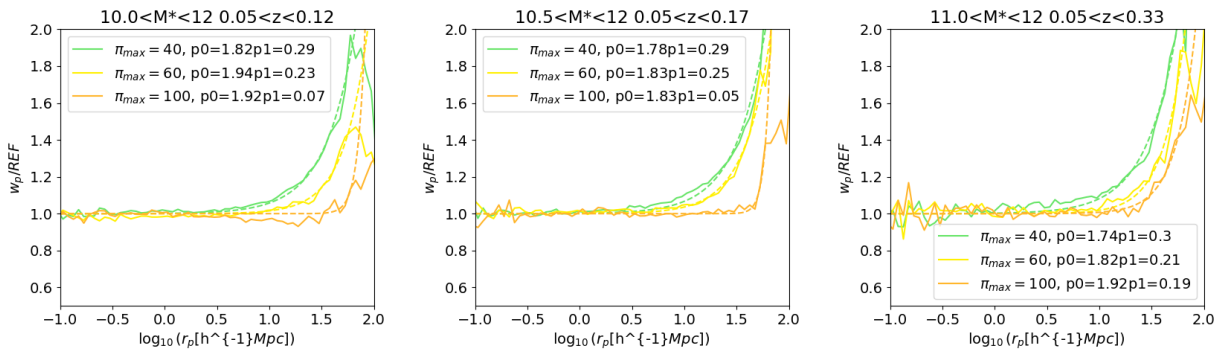


Fig. C.3. Ratio of the real space over redshift space correlation function for all selections. data from Uchuu in solid lines and fitted model in dashed lines. The model is used to correct observations that we fit until 1.6 (40 Mpc/h)

The purity of the obtained RS (BC) sample varies between 59% (68) and 69% (97) and the completeness between 20% (67) and 90% (97). The trade offs to choose the best parameter to cut at is not obvious. For the RS selection the purity is never higher than 70%. A choice could be to cut at 0.1 to obtain a completeness of 68% and a similar purity, leaving less of the

quarter of the selected objects to be SF contaminants. This is already a substantial number that will need to be modeled.

For the BC selection with a $c_{BC} = 0.34$ (0.22, 0.09) parameter, setting a similar completeness at 70.7% (80.3, 90.2) gives a purity of 96.7% (94.1, 84.5). A 5% purity is perfectly acceptable, so we could strive for higher completeness and choose 0.22.

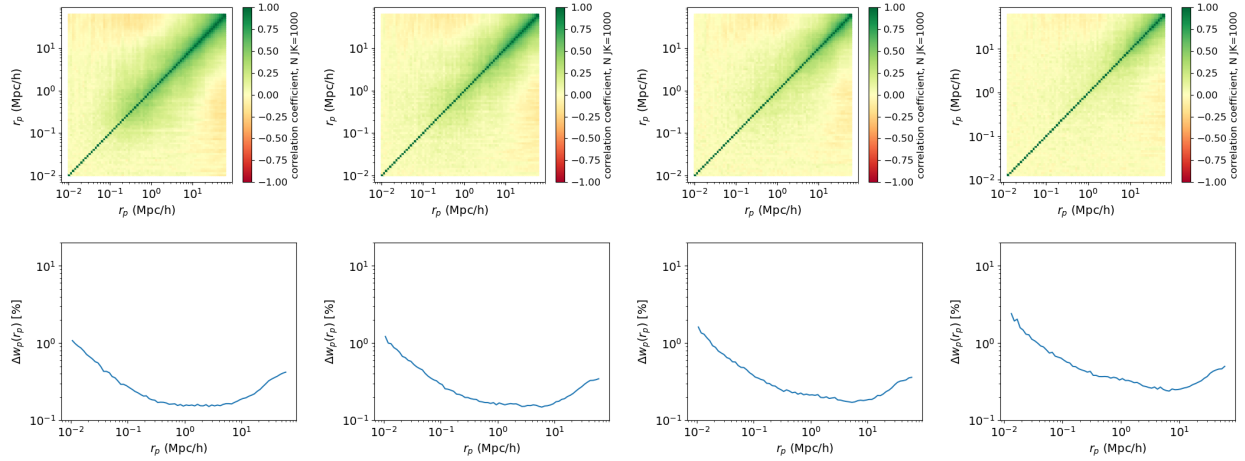


Fig. C.4. correlation coefficient (top), square root of the diagonal of the covariance matrix (bottom) for M 10.0 10.5 (2nd column), 10.75 (3rd column), 11.0 (4th column).

Table D.1. Red sequence g-z color as a function of redshift.

redshift	g-z color DECam
0.05	1.42
0.1	1.59
0.15	1.76
0.2	1.95
0.25	2.15
0.3	2.34
0.35	2.54
0.4	2.68
0.23	2.83
0.5	2.94
0.55	3.06
0.6	3.2

We chose $c_{RS} = 0.1$ and $c_{BC} = 0.22$ to split the sample into RS, BC and GV galaxies. This choice will depend on the application. There is a trend seen for the completeness (and to a lesser extent on the contamination) as a function of stellar mass that depends on the threshold chosen.

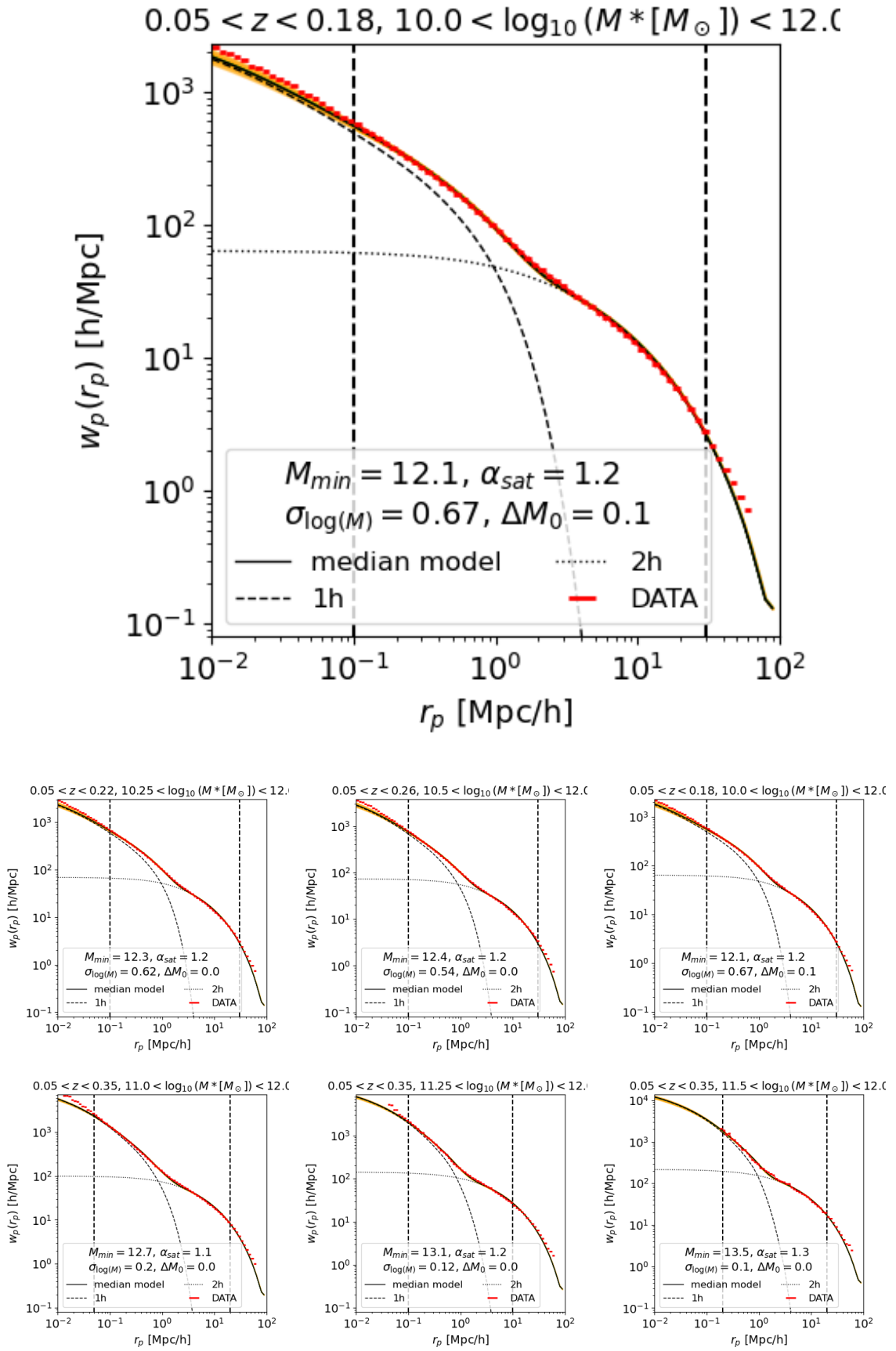


Fig. C.5. Correlation function and best-fit HOD model for the samples considered. The models represent well the observations in the range 0.1–20 Mpc/h .

Table D.2. Red sequence selection parameter, estimated completeness, and purity.

c_{RS} c_{BC}	RS selection $g_{AB} - z_{AB} > RS(z) - c_{RS}$			BC selection $g_{AB} - z_{AB} < RS(z) - c_{BC}$		
	N_{RS} (QU / SF) selected	completeness QU/46815 (Δ)	purity QU/ N_{RS} (Δ)	N_{BC} (QU / SF) selected	completeness SF/103160 (Δ)	purity SF/ N_{BC} (Δ)
-0.09	7655 (4310 / 2671)	9.2 (1.2)	56.3 (1.4)	147183 (42491 / 100448)	97.4 (-0.1)	68.2 (0.3)
-0.08	8610 (4976 / 2846)	10.6 (1.6)	57.8 (1.5)	146228 (41825 / 100273)	97.2 (-0.2)	68.6 (0.4)
-0.07	9687 (5740 / 3045)	12.3 (1.8)	59.3 (1.4)	145151 (41061 / 100074)	97.0 (-0.2)	68.9 (0.4)
-0.06	10945 (6623 / 3266)	14.1 (2.0)	60.5 (1.1)	143893 (40178 / 99853)	96.8 (-0.2)	69.4 (0.5)
-0.05	12376 (7609 / 3540)	16.3 (2.4)	61.5 (1.2)	142462 (39192 / 99579)	96.5 (-0.2)	69.9 (0.5)
-0.04	14018 (8813 / 3797)	18.8 (2.7)	62.9 (1.4)	140820 (37988 / 99322)	96.3 (-0.2)	70.5 (0.7)
-0.03	15849 (10172 / 4081)	21.7 (3.0)	64.2 (1.2)	138989 (36629 / 99038)	96.0 (-0.3)	71.3 (0.8)
-0.02	17793 (11599 / 4382)	24.8 (3.2)	65.2 (0.9)	137045 (35202 / 98737)	95.7 (-0.3)	72.0 (0.8)
-0.01	19973 (13199 / 4746)	28.2 (3.4)	66.1 (0.9)	134865 (33602 / 98373)	95.4 (-0.4)	72.9 (1.0)
-0.0	22181 (14837 / 5102)	31.7 (3.6)	66.9 (0.7)	132657 (31964 / 98017)	95.0 (-0.4)	73.9 (1.0)
0.01	24596 (16604 / 5515)	35.5 (3.9)	67.5 (0.6)	130242 (30197 / 97604)	94.6 (-0.4)	74.9 (1.1)
0.02	27116 (18487 / 5961)	39.5 (3.9)	68.2 (0.5)	127722 (28314 / 97158)	94.2 (-0.4)	76.1 (1.1)
0.03	29642 (20284 / 6460)	43.3 (3.9)	68.4 (0.2)	125196 (26517 / 96659)	93.7 (-0.5)	77.2 (1.2)
0.04	32301 (22198 / 7004)	47.4 (4.1)	68.7 (0.2)	122537 (24603 / 96115)	93.2 (-0.6)	78.4 (1.2)
0.05	34908 (24068 / 7576)	51.4 (3.9)	68.9 (0.1)	119930 (22733 / 95543)	92.6 (-0.6)	79.7 (1.2)
0.06	37482 (25879 / 8172)	55.3 (3.7)	69.0 (0.0)	117356 (20922 / 94947)	92.0 (-0.6)	80.9 (1.2)
0.07	39908 (27505 / 8802)	58.8 (3.5)	68.9 (-0.1)	114930 (19296 / 94317)	91.4 (-0.6)	82.1 (1.2)
0.08	42326 (29122 / 9445)	62.2 (3.5)	68.8 (-0.1)	112512 (17679 / 93674)	90.8 (-0.6)	83.3 (1.2)
0.09	44766 (30739 / 10119)	65.7 (3.2)	68.7 (-0.2)	110072 (16062 / 93000)	90.2 (-0.6)	84.5 (1.1)
0.1	47001 (32135 / 10842)	68.6 (2.9)	68.4 (-0.3)	107837 (14666 / 92277)	89.5 (-0.8)	85.6 (1.0)
0.11	49178 (33470 / 11591)	71.5 (2.8)	68.1 (-0.4)	105660 (13331 / 91528)	88.7 (-0.8)	86.6 (1.0)
0.12	51226 (34693 / 12312)	74.1 (2.5)	67.7 (-0.4)	103612 (12108 / 90807)	88.0 (-0.7)	87.6 (1.0)
0.13	53194 (35806 / 13082)	76.5 (2.2)	67.3 (-0.4)	101644 (10995 / 90037)	87.3 (-0.8)	88.6 (0.9)
0.14	54999 (36781 / 13836)	78.6 (2.0)	66.9 (-0.4)	99839 (10020 / 89283)	86.5 (-0.8)	89.4 (0.9)
0.15	56843 (37737 / 14654)	80.6 (1.9)	66.4 (-0.5)	97995 (9064 / 88465)	85.8 (-0.8)	90.3 (0.8)
0.16	58472 (38557 / 15423)	82.4 (1.7)	65.9 (-0.5)	96366 (8244 / 87696)	85.0 (-0.8)	91.0 (0.6)
0.17	59984 (39263 / 16186)	83.9 (1.4)	65.5 (-0.5)	94854 (7538 / 86933)	84.3 (-0.8)	91.6 (0.6)
0.18	61456 (39902 / 16984)	85.2 (1.2)	64.9 (-0.5)	93382 (6899 / 86135)	83.5 (-0.8)	92.2 (0.6)
0.19	62827 (40444 / 17767)	86.4 (1.1)	64.4 (-0.6)	92011 (6357 / 85352)	82.7 (-0.8)	92.8 (0.5)
0.2	64218 (40971 / 18594)	87.5 (1.1)	63.8 (-0.6)	90620 (5830 / 84525)	81.9 (-0.8)	93.3 (0.5)
0.21	65531 (41455 / 19400)	88.6 (1.0)	63.3 (-0.5)	89307 (5346 / 83719)	81.2 (-0.8)	93.7 (0.4)
0.22	66789 (41852 / 20236)	89.4 (0.8)	62.7 (-0.6)	88049 (4949 / 82883)	80.3 (-0.9)	94.1 (0.4)
0.23	68002 (42202 / 21079)	90.1 (0.8)	62.1 (-0.6)	86836 (4599 / 82040)	79.5 (-0.8)	94.5 (0.4)
0.24	69183 (42533 / 21918)	90.9 (0.7)	61.5 (-0.6)	85655 (4268 / 81201)	78.7 (-0.8)	94.8 (0.3)
0.25	70284 (42827 / 22718)	91.5 (0.6)	60.9 (-0.6)	84554 (3974 / 80401)	77.9 (-0.8)	95.1 (0.3)
0.26	71379 (43120 / 23501)	92.1 (0.5)	60.4 (-0.6)	83459 (3681 / 79618)	77.2 (-0.8)	95.4 (0.2)
0.27	72449 (43355 / 24325)	92.6 (0.5)	59.8 (-0.6)	82389 (3446 / 78794)	76.4 (-0.8)	95.6 (0.2)
0.28	73513 (43576 / 25160)	93.1 (0.4)	59.3 (-0.5)	81325 (3225 / 77959)	75.6 (-0.8)	95.9 (0.2)
0.29	74513 (43743 / 25987)	93.4 (0.4)	58.7 (-0.5)	80325 (3058 / 77132)	74.8 (-0.8)	96.0 (0.1)
0.3	75491 (43908 / 26795)	93.8 (0.4)	58.2 (-0.6)	79347 (2893 / 76324)	74.0 (-0.8)	96.2 (0.2)
0.31	76476 (44078 / 27606)	94.2 (0.4)	57.6 (-0.6)	78362 (2723 / 75513)	73.2 (-0.8)	96.4 (0.1)
0.32	77476 (44234 / 28448)	94.5 (0.3)	57.1 (-0.6)	77362 (2567 / 74671)	72.4 (-0.8)	96.5 (0.1)
0.33	78446 (44359 / 29287)	94.8 (0.2)	56.5 (-0.6)	76392 (2442 / 73832)	71.6 (-0.9)	96.6 (0.1)
0.34	79438 (44464 / 30170)	95.0 (0.2)	56.0 (-0.6)	75400 (2337 / 72949)	70.7 (-0.8)	96.7 (0.1)
0.35	80390 (44560 / 31024)	95.2 (0.2)	55.4 (-0.6)	74448 (2241 / 72095)	69.9 (-0.8)	96.8 (0.1)
0.36	81348 (44675 / 31866)	95.4 (0.2)	54.9 (-0.5)	73490 (2126 / 71253)	69.1 (-0.8)	97.0 (0.1)
0.37	82255 (44764 / 32683)	95.6 (0.2)	54.4 (-0.5)	72583 (2037 / 70436)	68.3 (-0.8)	97.0 (0.0)
0.38	83224 (44847 / 33567)	95.8 (0.2)	53.9 (-0.5)	71614 (1954 / 69552)	67.4 (-0.9)	97.1 (0.1)

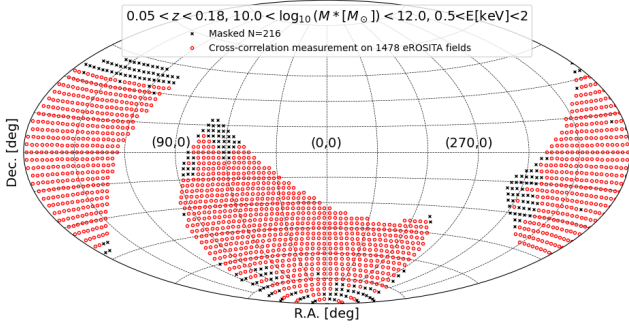


Fig. D.1. The footprint of the set of eROSITA fields used for the cross-correlation between X-rays and the sample with $\log_{10}(M^*/M_{\odot}) > 10$ and $0.05 < z < 0.18$.

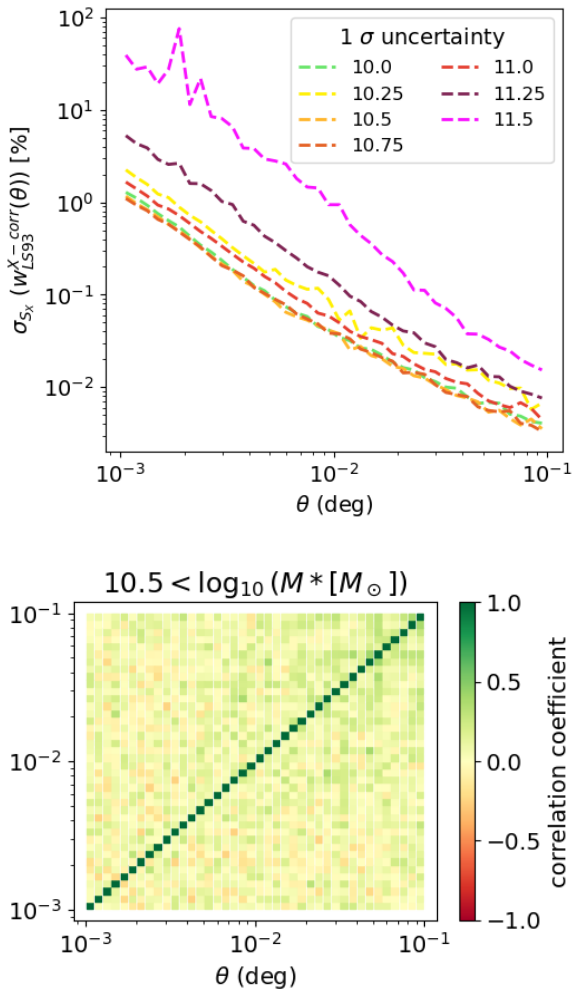


Fig. D.2. Top: Diagonal relative uncertainties on the cross-correlation. **Bottom:** Correlation coefficient for the sample with $\log_{10}(M^*/M_{\odot}) > 10.5$.