A&A 642, A184 (2020)
https://doi.org/10.1051/0004-6361/202038622

Astronomy & Astrophysics

Mock catalogs for the extragalactic X-ray sky: Simulating AGN surveys with ATHENA and with the AXIS probe

S. Marchesi[1,2], R. Gilli[1], G. Lanzuisi[1], T. Dauser[3], S. Ettori[1,4], F. Vito[5], N. Cappelluti[6], A. Comastri[1], R. Mushotzky[7], A. Ptak[8,9], and C. Norman[10,9]

[1] INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti, 93/3, 40129 Bologna, Italy
e-mail: stefano.marchesi@inaf.it
[2] Department of Physics and Astronomy, Clemson University, Kinard Lab of Physics, Clemson, SC 29634, USA
[3] Dr Karl Remeis-Observatory and Erlangen Centre for Astroparticle Physics, Sternwartstr. 7, 96049 Bamberg, Germany
[4] INFN, Sezione di Bologna, viale Berti Pichat 6/2, 40127 Bologna, Italy
[5] Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
[6] Physics Department, University of Miami, Coral Gables, FL 33124, USA
[7] Department of Astronomy, University of Maryland, College Park, MD 20742, USA
[8] NASA Goddard Space Flight Center, Code 662, Greenbelt, MD 20771, USA
[9] Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA
[10] Space Telescope Science Institute, 3700 San Martin Dr., Baltimore, MD 21210, USA

Received 9 June 2020 / Accepted 19 August 2020

ABSTRACT

We present a series of new, publicly available mock catalogs of X-ray selected active galactic nuclei (AGNs), nonactive galaxies, and clusters of galaxies. These mocks are based on up-to-date observational results on the demographic of extragalactic X-ray sources and their extrapolations. They reach fluxes below 10^{-20} erg cm^{-2} s^{-1} in the 0.5–2 keV band, that is, more than an order of magnitude below the predicted limits of future deep fields, and they therefore represent an important tool for simulating extragalactic X-ray surveys with both current and future telescopes. We used our mocks to perform a set of end-to-end simulations of X-ray surveys with the forthcoming ATHENA mission and with the AXIS probe, a subarcsecond resolution X-ray mission concept proposed to the Astro 2020 Decadal Survey. We find that these proposed, next generation surveys may transform our knowledge of the deep X-ray Universe. As an example, in a total observing time of 15 Ms, AXIS would detect ∼225 000 AGNs and ∼50 000 nonactive galaxies, reaching a flux limit of $f_{0.5-2} \sim 5 \times 10^{-19}$ erg cm^{-2} s^{-1} in the 0.5–2 keV band, with an improvement of over an order of magnitude with respect to surveys with current X-ray facilities. Consequently, 90% of these sources would be detected for the first time in the X-rays. Furthermore, we show that deep and wide X-ray surveys with instruments such as AXIS and ATHENA are expected to detect ∼20 000 $z > 3$ AGNs and ∼250 sources at redshift $z > 6$, thus opening a new window of knowledge on the evolution of AGNs over cosmic time and putting strong constraints on the predictions of theoretical models of black hole seed accretion in the early universe.

Key words. X-rays: galaxies – surveys – telescopes – galaxies: active

1. Introduction

Observations of the sky in the X-rays (i.e., in the 0.3–10 keV energy range) are known to play a key role in the detection and understanding of different types of extragalactic sources. For example, X-ray data are an efficient tool to select accreting supermassive black holes (SMBHs), the so-called active galactic nuclei (AGNs). In fact, X-ray surveys are significantly less affected by contamination from non-AGN emission than optical and infrared surveys; this is particularly the case for the one associated to star-formation processes (see, e.g., Donley et al. 2008, 2012; Stern et al. 2012). As a consequence, only the deepest existing X-ray survey, that is, the *Chandra* Deep Field-South (CDF-S) 7 Ms survey (Luo et al. 2017), reaches 0.5–2 keV fluxes where the number of nonactive galaxies becomes non-negligible (i.e., $f_{0.5-2} \lesssim 10^{-17}$ erg s^{-1} cm^{-2}): in all the other X-ray surveys, AGNs make up almost the totality of the detected sources. Characterizing the properties of the diffuse medium in clusters of galaxies, which are the largest virialized structures in the Universe, also requires deep X-ray surveys over large areas of the sky (see, e.g., Finoguenov et al. 2015; Käfer et al. 2020).

X-ray surveys are also effective in detecting obscured sources, that is, objects where the circumnuclear dust and gas absorb the AGN optical emission. A significant fraction of these objects can even be heavily obscured, Compton thick AGNs (i.e., objects with an optical depth to electron scattering of $\sigma_\tau \geq 1$ or column density of $N_{\rm H} \geq 1.5 \times 10^{24}$ cm^{-2}), which are detected in X-ray surveys up to redshifts of $z \sim 2$–3 (see, e.g., Comastri et al. 2011; Georgantopoulos et al. 2013; Buchner et al. 2015; Lanzuisi et al. 2015, 2018), with only one known X-ray selected CT-AGN found at $z > 4$ (XID403, a $z = 4.76$ detected in the *Chandra* Deep field, Gilli et al. 2011, 2014). These sources are especially interesting because they are expected to significantly contribute to the overall AGN population (up to 50%; see, e.g., Gilli et al. 2007; Ananna et al. 2019), and X-rays represent one of the most efficient ways to find and

characterize them. However, most of the CT population is too faint to be detected by current X-ray facilities.

A complete census of AGNs over a wide range of redshifts, column densities, and luminosities requires a multi-tiered approach. For example, intrinsically faint and/or heavily obscured sources need extremely deep surveys, which, as of today, could have only been performed on limited ($<0.5\,\mathrm{deg}^2$) regions of the sky; this is the case, for example, for the CDF-S 7 Ms survey (Luo et al. 2017), AEGIS XD (Nandra et al. 2015), SSA22 (Lehmer et al. 2009), or the J1030 field (Nanni et al. 2020). Rare, luminous quasars can instead be found in significant numbers by only using shallow large-area surveys that cover tens or even hundreds of square degrees (see, e.g., the Stripe 82X survey, LaMassa et al. 2013b,a, or XMM-XXL, Pierre et al. 2016). Finally, intermediate-area surveys (e.g., COSMOS, Civano et al. 2016; Marchesi et al. 2016a) sample relatively deep fluxes over areas of a few deg^2, thus allowing one to detect statistically significant samples of high-z sources and study large-scale structures.

Twenty years of observations with *Chandra* and *XMM-Newton* have provided us with large samples of X-ray selected AGNs, which allowed us to put constraints on the coevolution of the accreting supermassive black holes and their host galaxies up to the peak of the AGN activity. However, our knowledge on the behavior of the first AGNs which formed in the 2 Gyr after the Big Bang (i.e., at redshift $z > 3$) is still limited (see, e.g., Ueda et al. 2014; Aird et al. 2015; Buchner et al. 2015; Miyaji et al. 2015; Ananna et al. 2019). As of today, only ∼300 X-ray selected AGNs have been detected at $z > 3$ (see, e.g., Vito et al. 2014, 2018; Marchesi et al. 2016b), and the farthest spectroscopically confirmed X-ray selected AGN has been found at $z = 5.31$ (Capak et al. 2011). Furthermore, the vast majority of high-z X-ray sources are detected with only a few source counts, thus limiting the scientific outcome of their analysis. In particular, understanding the evolution of the fraction of obscured AGNs, which has been shown to increase with redshift in several works (see, e.g., Treister et al. 2013; Vito et al. 2018), becomes particularly complex and significant incompleteness corrections need to be applied (see, e.g., Lanzuisi et al. 2018). Notably, while obscured AGN are expected to make the majority of the high-z AGN population, at $z > 6$, optical and near-infrared surveys have so far detected only unobscured, luminous quasars (Bañados et al. 2016; Wang et al. 2017). In the X-rays, a proper spectral characterization of heavily obscured AGNs requires the detection of hundreds of source counts (see, e.g., Marchesi et al. 2018, 2019); a task that cannot be achieved by current X-ray telescopes.

For these reasons, new X-ray facilities are required to improve our knowledge on AGNs and complement the information that will become available in the next decade in the other bands. Accurate survey simulations are key to properly assess the required technical layout of future facilities. Such simulations are based on accurate mock catalogs of extragalactic sources, that should be flexible enough to be mission-independent.

One of the major issues when building an AGN mock catalog is which approach to choose to simulate the AGN population which is not detected by current surveys, particularly those sources at redshift $z > 6$. For example, one can extrapolate the available observational evidence at redshift $z \sim 6$ up to redshift $z \geq 10$ (this is for example the approach followed in the development of the ATHENA mission, see, e.g., Nandra et al. 2013). Alternatively, one can use theoretical predictions from black hole seed early accretion models (an approach followed by the *Lynx* mission, The Lynx Team 2018). Notably, since different missions often adopt different approaches to build their mock catalogs, high-z AGN detection predictions cannot be directly compared, thus significantly limiting a cross-mission analysis of future X-ray surveys.

To address this issue, in this work we present a new set of mock catalogs, based on the most recent observational results on the demographic of extragalactic X-ray sources, and we then used them to simulate surveys with future X-ray facilities such as the AXIS probe (Mushotzky et al. 2019) and ATHENA. The mocks are mission-independent and are made available to the public.

The paper is organized as follows: in Sect. 2 we describe `SIXTE`, the software we used to simulate the X-ray surveys, and we discuss how we built our AGN, nonactive galaxies and galaxy clusters mock catalogs. In Sect. 3 we then present the AXIS probe, its technical capabilities and its proposed scientific goals, and in Sect. 4 we present the results of the simulations of three different AXIS surveys. In Sect. 5 we focus on the improvement that ATHENA and AXIS would bring to our knowledge of high-z and of heavily obscured AGNs. We then summarize our results in Sect. 6. In this work, we assume a flat ΛCDM cosmology with $\mathrm{H}_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$.

2. New mock catalogs for X-ray simulations

2.1. The `SIXTE` software

The Monte Carlo code Simulation of X-ray Telescopes (hereafter `SIXTE`, Dauser et al. 2019) allows one to simulate an observation with an X-ray telescope, following a three-step approach. At first, the tool creates a photon list containing the arrival time, energy and position of each photon. To generate this information, `SIXTE` needs to be provided with the instrument effective area, field of view and pointing. The photon list created in the first step is then convoluted with the instrument point spread function (PSF) and vignetting to provide an impact list that contains the energy and arrival time of each photon, as well as its position on the detector. Finally, the charge cloud information reported in the impact list is read-out and re-combined into events, taking into account the properties of the simulated detector (e.g., read-out properties, redistribution matrix file...). The output of this last step is an event file that can be analyzed.

The `SIXTE` webpage[1] contains the configuration files (i.e., telescope setup, response matrices, vignetting, point spread function) for several current facilities, such as *XMM-Newton*, *NuSTAR* and eROSITA. Furthermore, configuration files for new missions are also provided: for example, both ATHENA (Nandra et al. 2013) instruments, that is, the Wide Field Imager (WFI, Meidinger et al. 2017) and the X-ray Integral Field Unit (X-IFU Barret et al. 2016) are available. Finally, the `SIXTE` manual contains all the information to build from scratch the configuration files for facilities that are not currently available on the website.

In order to perform the steps we mentioned above, `SIXTE` needs a source list as input: such a list is generated in the `SIMPUT` data format. `SIMPUT` source lists are detector-independent, that is, the catalogs we originated for this work can be used to simulate any type of X-ray observation.

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

Our catalogs are available online[2] in FITS format and ready to be used, among other software, within SIXTE. In the following sections, we describe how we built them.

2.2. AGN mock catalog

Three AGN mock catalogs are available on our webpage, for survey areas of 1, 10 and 100 deg^2, respectively. Random right ascension and declination have been associated to the SIXTE catalogs.

Our mocks are made of sources with different intrinsic 0.5-2 keV luminosities, redshifts and column densities, that have been extracted by resampling the X-ray luminosity function of unabsorbed AGN given by Hasinger et al. (2005), scaled up by a luminosity–dependent factor to account for the whole AGN population (see Gilli et al. 2007). At $z > 2.7$, a decline in the AGN space density as parameterized in Schmidt et al. (1995) has been assumed, while we did not make any assumption on AGN or host clustering.

The resampling has been checked to reproduce the correct AGN densities as a function of luminosity, redshift and column density. To get the number of sources per unit solid angle, the source populations have been weighted for the volume element dV/dz/dΩ. The obtained catalogs have then been checked to simultaneously reproduce the 0.5–2 keV and 2–10 keV input AGN number counts as a function of column density.

We point out that in Gilli et al. (2007) a distribution of photon indices was assumed for the primary power-law (with dispersion $\sigma_\Gamma = 0.2$), while disk reflection was assumed only for lower luminosity AGN. For simplicity (i.e., to limit the number of spectral templates used in the analysis), in the generation of the mock catalogs we did not include any photon index dispersion and assumed disk reflection everywhere. Therefore, the 2–10 keV k-correction was adjusted to an average value to correctly reproduce the total 2–10 keV number counts.

AGNs have been simulated down to a 0.5–2 keV luminosity $L_{0.5-2} = 10^{40}$ erg s^{-1} and up to $z = 10$: the source density of the mocks is ∼53 500 sources deg^{-2}. It is worth noting that several works have reported detections of AGNs with luminosities below the 10^{40} erg s^{-1} threshold we decided to adopt, both in surveys (see, e.g., Aird et al. 2015) and in the Local Universe (e.g., Bi et al. 2020; Hodges-Kluck et al. 2020). Therefore, our mock is likely missing a fraction of low-luminosity AGN. We further discuss this topic in Sect. 2.5.

We report in Fig. 1 the redshift, 0.5–2 keV luminosity and column density ($N_{\rm H}$) distribution of the sources in the AGN mock catalog. As it can be seen, $N_{\rm H}$ is quantized, following the same approach and column density distribution adopted in the developing of the Gilli et al. (2007) cosmic X-ray background model. More in detail, we assign to each source a column density value taken from the array $\log(N_{\rm H}) = [20.5, 21.5, 22.5, 23.5, 24.5, 25.5]$, each representative of 1-dex wide column density bin. The average Compton thick fraction, that is, the fraction of sources with column density $\log(N_{\rm H}) > 24$, is $f_{\rm CT} = 39\%$. As a reference, we report the average CT fraction values adopted in other AGN population synthesis models, such as, for example, those by Ueda et al. (2014, $f_{\rm CT} = 33\%$), Buchner et al. (2015, $f_{\rm CT} = 38\%$), and Ananna et al. (2019, $f_{\rm CT} = 50\%$). We remark that these are all average values, since all models adopt a luminosity-dependent CT fraction, based on the observational evidence (see, e.g., Ricci et al. 2015; Marchesi et al. 2016a).

Finally, in Fig. 2, left panel, we plot the cumulative AGN number counts as a function of the observed 0.5–2 keV flux for our mock catalog, and we compare them with those obtained in the deep, pencil-beam CDF-S 7 Ms survey (Luo et al. 2017), in the intermediate *Chandra* COSMOS Legacy survey (Civano et al. 2016) and in the wide-area, shallow flux Stripe 82X survey (LaMassa et al. 2013b). Our mock nicely matches with the available observational results, over a wide range of fluxes.

In the same figure, we also report the AGN number counts derived using the Ueda et al. (2014) X-ray luminosity function (XLF) model. As it can be seen, our predictions are consistent with the Ueda et al. (2014) model. The agreement is even stronger considering that the Ueda et al. (2014) model accounts only for AGNs having redshift $z = [0–5]$ and 0.5–2 keV luminosity $L_{0.5-2} = [10^{41}–10^{46}]$ erg s^{-1} cm^{-2}, while our mock contains objects up to redshift $z = 10$ and, more importantly, down to $L_{0.5-2} = 10^{40}$ erg s^{-1}. The inclusion of $L_{0.5-2} < 10^{41}$ erg s^{-1} AGNs is particularly important, considering that our mock contains ∼30 000 of such sources over 1 deg^2 at 0.5–2 keV fluxes $< 10^{-17}$ erg s^{-1} cm^{-2}. Adding these 30 000 sources to the Ueda et al. (2014) AGN number counts, the discrepancy between the two models is <30% at fluxes $f_{0.5-2} \sim 10^{-17}$ erg s^{-1} cm^{-2}, and <10% at $f_{0.5-2} \sim 10^{-19}$ erg s^{-1} cm^{-2}. Therefore, the predictions we make in Sect. 4 on the overall number of AGNs detected by AXIS and ATHENA would not change significantly if we used the Ueda et al. (2014) XLF instead of the Gilli et al. (2007) AGN population synthesis model to generate our mock.

2.2.1. AGN spectral models

SIXTE allows one to associate a spectral model to each source in the mock catalog. This model is then used to convolve the 0.5–10 keV flux which is given to each source when the mock is created.

We created our spectral models using the XSPEC software (Arnaud et al. 1996): in XSPEC, nomenclature, the model is built as follows:

$$phabs * (zcutoffpl + pexmon + zphabs * zcutoffpl), \quad (1)$$

where *phabs* is the Galactic absorption (fixed to $N_{\rm H,Gal} = 1.8 \times 10^{20}$), *zcutoffpl* is a power law with photon index $\Gamma = 1.9$ (a typical value for AGNs, see, e.g., Marchesi et al. 2016c, and references therein) and high-energy cutoff $E_{\rm cut} = 200$ keV, *pexmon* (Nandra et al. 2007) models the emission reprocessed by cold material surrounding the accreting supermassive black hole, including self-consistently generated fluorescence lines such as the Fe Kα, the Fe Kβ, the Ni Kα and the Fe Kα Compton shoulder. In *pexmon*, the photon index and cut-off energy are tied to the *zcutoffpl* one, while the normalization is assumed equal to 37% of the main power-law one, which corresponds to ∼2% of the overall emission in the 2–10 keV band (see, e.g., Gilli et al. 2007). While several more recent models are available to treat the reprocessed component (e.g., MYTorus, Murphy & Yaqoob 2009, or borus02, Baloković et al. 2018), we used *pexmon* because of its consistency with the model originally used in Gilli et al. (2007). Finally, *zphabs* models the photoelectric absorption of a material with column density $N_{\rm H}$.

The first power law component is phenomenological and based on observational evidence, and it models the fraction of emission which is scattered, rather than absorbed by the obscuring material, as well as the unresolved emission lines from

[2] http://cxb.oas.inaf.it/mock.html

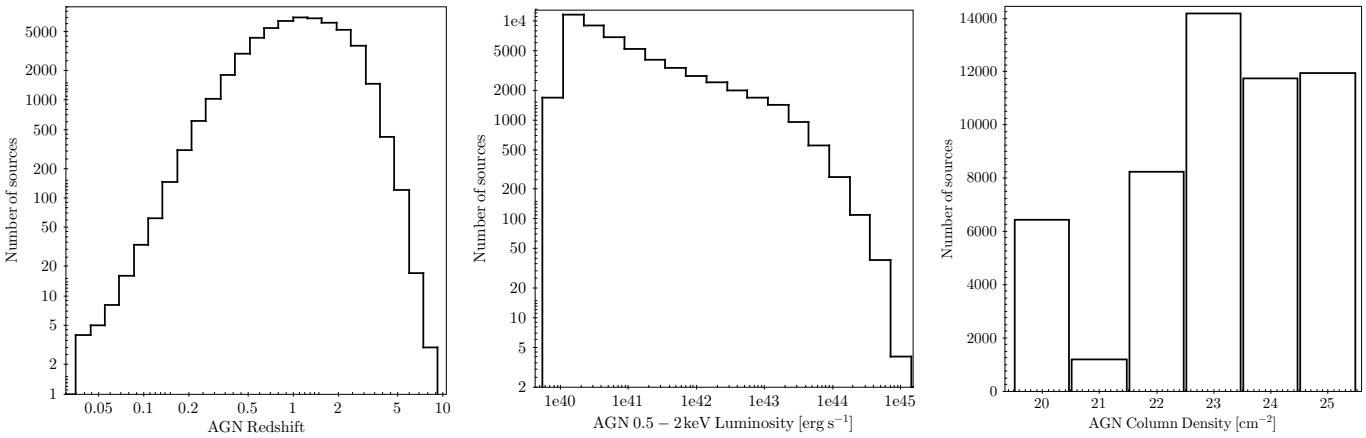

Fig. 1. Distribution of redshift (*left*), 0.5–2 keV luminosity and column density (*right*) for the 53 579 AGNs contained in one of the mock catalogs used in our analysis. The sources cover a 1 deg^2 field and are drawn from the Gilli et al. (2007) model.

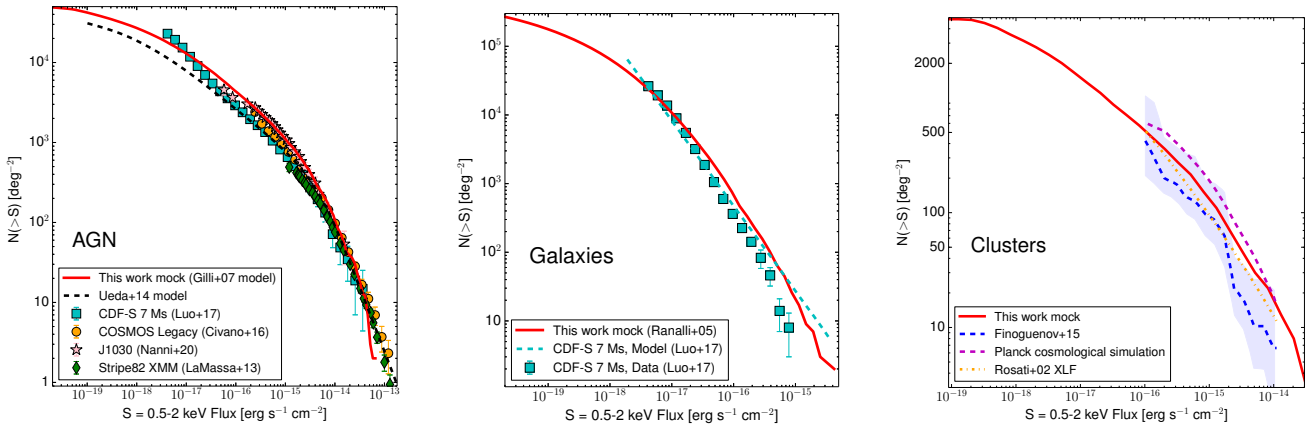

Fig. 2. *Left*: cumulative AGN number counts as a function of the 0.5–2 keV flux. The number counts from our mock catalog, derived using the Gilli et al. (2007) AGN population synthesis model, are plotted as a solid red line. As a comparison, we also plot the number counts derived in the CDF-S 7 Ms (cyan squares; Luo et al. 2017), *Chandra* COSMOS Legacy (orange circles; Civano et al. 2016), J1030 (pink stars; Nanni et al. 2020), and Stripe 82X (green diamonds; LaMassa et al. 2013b) surveys. The number counts derived from the Ueda et al. (2014) X-ray luminosity function are also shown with a dashed black line, for comparison. We note that the difference at faint fluxes between these number counts and those of our mock is almost completely due to the fact that Ueda et al. (2014) did not include AGNs with luminosities $<10^{41}$ erg s^{-1} in their analysis. *Center*: cumulative nonactive galaxies number counts as a function of the 0.5–2 keV flux. The number counts from our mock catalog, derived using the Ranalli et al. (2005) luminosity function, are plotted as a solid red line. The observed (cyan squares) and model-inferred number counts (dashed cyan line) derived in the CDF-S 7 Ms (Luo et al. 2017) are also shown for comparison. *Right*: cumulative galaxy clusters number counts as a function of the 0.5–2 keV flux. The number counts from our mock catalog are plotted as a solid red line. As a comparison, we also plot the number counts derived in the Extended *Chandra* Deep Field–South field (dashed blue line; the 1σ uncertainty area is shown as a blue shaded area; Finoguenov et al. 2015), and those derived combining data from ROSAT, ASCA and *Beppo*-SAX surveys (dash-dotted orange line; Rosati et al. 2002). Finally, we plot as magenta dashed line the predictions of the *Planck* cosmological simulation (Planck Collaboration XVI 2014).

photoionized gas in the narrow line region, and even soft X-ray emission from star forming processes in the host galaxy. The two power laws have same Γ and $E_{\rm cut}$, and the normalization of the unabsorbed power law is set to be equal to 3% of the main one (following what is typically observed in X-ray surveys, see, e.g., Marchesi et al. 2016c). For the most heavily obscured sources, that is, those with $\log(N_{\rm H})$=25.5, we set the normalization of the power law component to 0, assuming that all emission comes from the *pexmon* component: we do so because the *zphabs* component does not accurately describe the absorption caused by material having column density $\log(N_{\rm H}) > 25$.

We generated a set of spectra covering the redshift range $z = [0–10]$, with a redshift bin $\Delta z = 0.1$, adopting the same column density array used in the Gilli et al. (2007) AGN population synthesis model, that is, $\log(N_{\rm H}) = [20.5, 21.5, 22.5, 23.5, 24.5, 25.5]$. We then associated to each source the spectrum with the same $N_{\rm H}$ and the closest z. We show an example of spectrum for each Log($N_{\rm H}$) value in Fig. 3.

While the `pexmon` model generally offers a reliable characterization of the reprocessed emission in heavily obscured AGN, in the past 10 years several models have been developed to treat this complex spectral component in a more self-consistent way (e.g., Murphy & Yaqoob 2009, `borus02` Baloković et al. 2018). These models have as free parameters physical quantities such as the torus average column density, $N_{\rm H}$, which is different from the line-of-sight column density if the torus is inhomogeneous, and the torus covering factor. Thus, in Fig. 3 we also report the spectrum of a CT-AGN with $\log(N_{\rm H})$=24.5 as modeled by the self-consistent `borus02` model (Baloković et al. 2018), assuming a torus covering factor $f_c = 0.5$ and and an almost edge-on viewing angle, $\theta_{\rm obs} = 87°$.

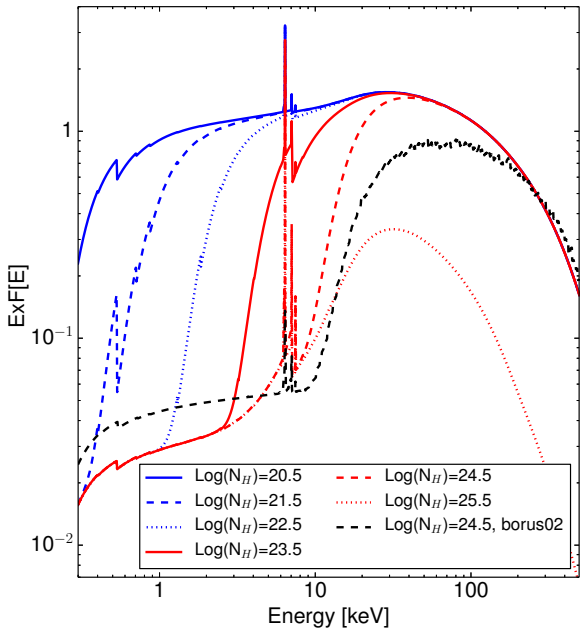

Fig. 3. AGN spectra used in our simulations, at $z = 0$ and for different $\log(N_{\rm H})$ values. We also plot the spectrum of a CT-AGN with $\log(N_{\rm H}) = 24.5$ as modeled using the `borus02` model (Baloković et al. 2018, dashed black line), as a reference.

As shown in Fig. 3, in heavily obscured sources the >10 keV flux predicted by the physically motivated `borus02` model is 30–60% fainter than the one predicted by the `pexmon` one. This discrepancy, while not extreme, can affect the number of heavily obscured sources in simulations with our mocks. We further discuss this effect in Sect. 5.3.

2.2.2. Mock catalog of high-redshift AGNs

One of the fundamental scientific topics that can be addressed by X-ray surveys, particularly with next-generation facilities, is the study and characterization of the high-redshift AGN population. As of today, our knowledge on the $z > 3$ AGN population is extremely limited, since only a few hundreds of these sources have been detected in X-ray surveys (see, e.g., Vito et al. 2014, 2018; Marchesi et al. 2016b; Nanni et al. 2020), and there are no X-ray selected AGN at $z > 6$.

While the Gilli et al. (2007) model is in overall excellent agreement with the observational results obtained by the most up-to-date X-ray surveys (as shown in Fig. 2, left panel), in the high-redshift regime (i.e., at $z > 3$, where the AGN space density starts its decline) there is a more significant discrepancy between the predictions of the AGN population synthesis model and the observational evidence.

In particular, as shown in Fig. 4, at fluxes $<10^{-16}$ erg s^{-1} cm^{-2} the predictions of the Gilli et al. (2007) model lie below the number counts observed in the deepest X-ray surveys currently available (the CDF-S 7 Ms and the CDF-N 2 Ms, see Vito et al. 2018). The Vito et al. (2018) data is instead in close agreement with the predictions of the X-ray luminosity function (XLF) by Vito et al. (2014), which was computed using a sample of 141 X-ray selected high-z AGNs.

For this reason, we generated a second AGN mock catalog, containing only $z > 3$ sources, based on the Vito et al. (2014) XLF. More in detail, the new $z > 3$ mock catalog is obtained by simulating AGN in the redshift range $z = [3–20]$ and down to $\log(L_{0.5-2\rm keV}) = 40$ extrapolating the XLFs of Vito et al. (2014), which were originally derived in the redshift range $z = [3–5]$ and for $\log(L_{0.5-2\,\rm keV}) > 43$. With these redshift and luminosity limits, AGNs in the mock catalog reach fluxes below 2×10^{-20} erg cm^{-2} s^{-1}.

To generate the Vito et al. (2014) mock catalog, AGNs with different intrinsic 2–10 keV luminosities, redshifts and column densities have been extracted by resampling the pure density evolution (PDE) model XLF corrected for redshift incompleteness (see Fig. 7 and Table 5 of Vito et al. 2014) We assumed: (i) a constant obscured AGN fraction with luminosity; (ii) number ratios between unobscured, obscured Compton-thin, and obscured Compton-thick AGNs of 1:4:4; (iii) the same column density distribution of Gilli et al. (2007, see also Fig. 1, right panel). All these assumptions are in agreement with the observational results reported in Vito et al. (2018). Finally, to compute the number of sources per unit of solid angle, the source populations have been weighted for the volume element $\mathrm{d}V/\mathrm{d}z/\mathrm{d}\Omega$. We use the same spectral templates used for the Gilli et al. (2007) mock and presented in Sect. 2.2, do no to include any dispersion in photon indices, and assume disk reflection in each template.

In Fig. 4, left panel, we also report the $z > 3$ number counts derived using the Ueda et al. (2014) XLF, which are in excellent agreement with the predictions of our Vito et al. (2014) mock catalog. The discrepancy at 0.5–2 keV fluxes $<10^{-17}$ erg s^{-1} cm^{-2} is mostly due to the Ueda et al. (2014) model not including AGNs at $z > 5$. As a reference, the Vito et al. (2014) mock contains ~3300 $z > 5$ AGNs at fluxes $f_{0.5-2} < 10^{-17}$ erg s^{-1} cm^{-2}: adding this number of sources to the Ueda et al. (2014) number counts, the discrepancy between the two is <30%, and is dominated by the uncertainties of the XLF models reported by Ueda et al. (2014) and Vito et al. (2014).

Consequently, we expect that simulations with a mock generated using the Ueda et al. (2014) XLF would produce high-z predictions in close agreement with those which we present in Sect. 5 and we obtained using the Vito et al. (2014) mock.

2.3. Galaxy mock catalog

Galaxies (here and in the rest of the paper this is how we define point-like sources for which the X-ray emission is not caused by an accreting supermassive black hole) have been extracted by interpolating between the peakM and peakG model logN-logS by Ranalli et al. (2005). Galaxy spectra have been assumed to be simple power-laws with photon index $\Gamma = 2$, that is, their band ratios do not depend on redshift (see, e.g., Sazonov & Khabibullin 2017; Barger et al. 2019). The Galactic absorption is the same used for AGNs, that is, $N_{\rm H} = 1.8 \times 10^{20}$ cm^{-2}.

Galaxies have been simulated down to a 0.5–2 keV flux limit $f_{0.5-2} = 10^{-20}$ erg s^{-1} cm^{-2}. The mock catalog we use in our analysis contains 294 000 sources per unit of deg^2.

We show in Fig. 2, central panel, the 0.5–2 keV flux number counts of the simulated nonactive galaxies, which are in good agreement with the most recent observational results in the CDF-S 7 Ms (Luo et al. 2017). Since the main focus of this work are active galactic nuclei, we do not associate a redshift (and therefore a luminosity) to the galaxies in our mock catalog.

2.4. Galaxy clusters mock catalog

To build a catalog of X-ray sources associated to galaxy clusters, we start with the predictions of dark matter halos extracted from a numerical mass function for a given set of cosmological

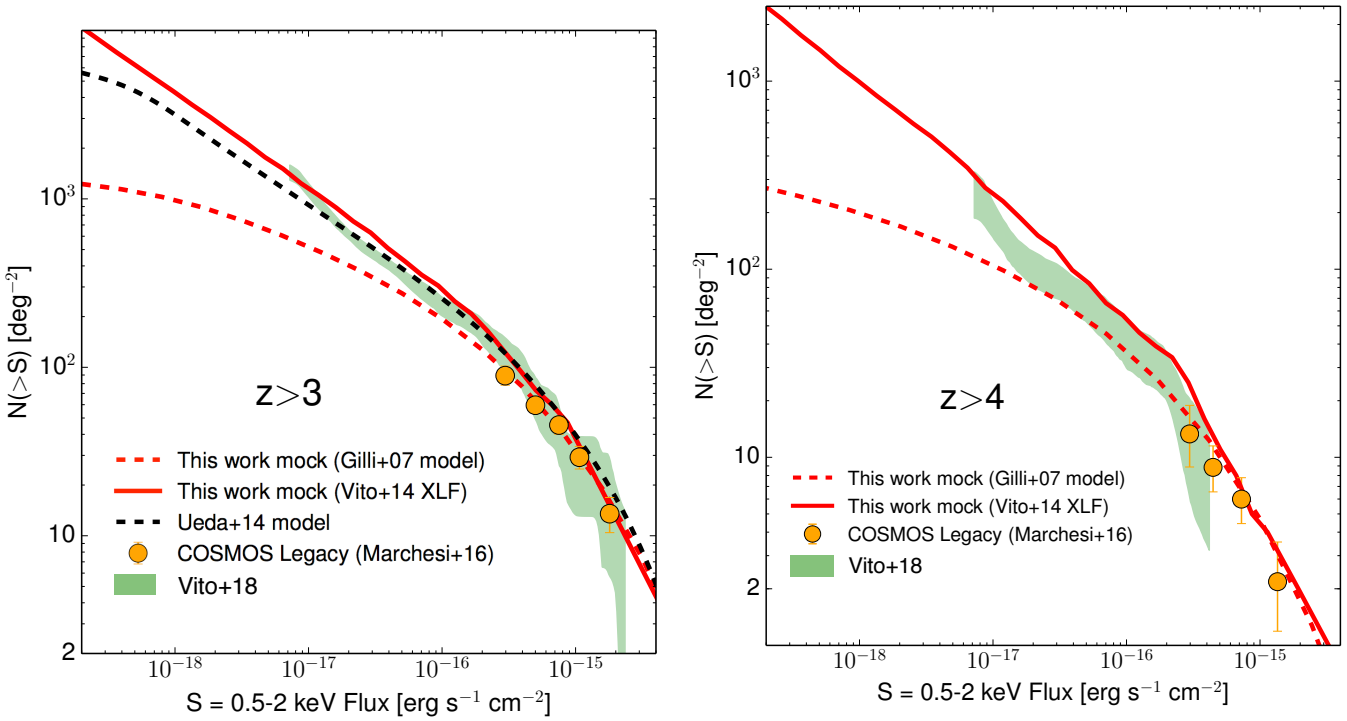

Fig. 4. Cumulative 0.5–2 keV number counts at redshift $z > 3$ (*left*) and $z > 4$ (*right*). The number counts of the mock catalog we derived from the Vito et al. (2014) $z > 3$ XLF are plotted as a solid red line, while those of the mock derived using the Gilli et al. (2007) CXB model are shown using a dashed red line. The $z = [3–5]$ number counts derived with the Ueda et al. (2014, dashed black line) model are also shown for comparison. Finally, the observational results from *Chandra* COSMOS Legacy (orange circles, Marchesi et al. 2016b) and from the CDF-S 7 Ms and CDF-N 2 Ms (green shaded area highligting the 68% confidence region; Vito et al. 2018) are also plotted for comparision.

parameters: we use the mass functions of both Tinker et al. (2008) and Despali et al. (2016). Specifically, we adopt $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_m = 1 - \Omega_\Lambda = 0.3$ and $\sigma_8 = 0.79$ (see present cosmological constraints from galaxy clusters in Pratt et al. 2019). The number density is estimated in the mass range $10^{12} - 4 \times 10^{15} M_\odot$ in 200 redshift bins between $z = 0.03$ and $z = 6$ using the python code `Colossus`[3] (Diemer 2018). Then, it is integrated over the cosmological volume to recover the number of haloes expected at given mass and redshift per square degree. We also associate to each considered mass value, M_{500}, its corresponding radius $R_{500} = M_{500}^{1/3}/(4/3\,\pi\,500\,\rho_{cz})^{1/3}$ in arcminutes[4], and an X-ray luminosity and temperature kT as estimated from available scaling relations (e.g., Reichert et al. 2011). The clusters radial profile is assumed to be a simple β-profile.

We then use `XSPEC` to convert the X-ray luminosities into fluxes in the observed 0.5–2 keV band and in the corresponding surface brightness, including the effect of Galactic absorption, in units of erg s^{-1} cm^{-2} arcmin^{-2}. The cluster spectral model we use is (in `XSPEC` nomenclature) *phabs*∗*apec*, where *phabs* is the same Galactic absorption used for AGNs and nonactive galaxies, while *apec* models emission from a collisionally-ionized diffuse gas with temperature kT and metallicity $Z = 0.3\,Z_\odot$, that is, the average metallicity value for clusters as reported in the literature (see, e.g., Balestra et al. 2007; Maughan et al. 2008; Baldi et al. 2012). Finally, the cluster number counts are obtained by summing the estimated counts in each mass and redshift bin with an associated X-ray surface brightness above a given threshold and multiplying it by the explored area. We simulated clusters down to a 0.5–2 keV surface brightness limit $\mathrm{SB} = 10^{-18}\,\mathrm{erg\,s^{-1}\,cm^{-2}\,arcmin^{-2}}$.

As a consistency test, in Fig. 2, right panel, we report the cumulative number counts derived using our clusters mock catalog, and we compare them with results obtained from both real data and simulations. As it can be seen, the number counts derived using our mocks are consistent with the measurements from the ECDF-S (Finoguenov et al. 2015) and with those derived combining ROSAT, ASCA and *Beppo*-SAX data (Rosati et al. 2002). We also plot the simulated number counts obtained using the *Planck* cosmology (Planck Collaboration XVI 2014) and the Leauthaud et al. (2010) scaling relations: our data always lie slightly below the ones from the *Planck* simulation, thus ensuring that we are not overpredicting the number of clusters in our mocks, at any flux.

Since the focus of this work is the detection of AGNs, in the rest of the paper we do not discuss the detection of galaxy clusters, which for the purposes of our analysis are therefore treated as foreground or background emitters.

2.5. Consistency with the cosmic X-ray background

In Fig. 5 we compare the cumulative 0.5–2 keV and 2–10 keV fluxes per unit of square degree of our mocks, divided by class of sources (i.e., AGNs, nonactive galaxies and galaxy clusters), with the overall surface brightness of the extragalactic cosmic X-ray background (CXB) in the same bands (Cappelluti et al. 2017). Our mocks do not over-estimate the CXB in both the soft and the hard band, an evidence that further validates the reliability of the catalogs we generated.

[3] `http://www.benediktdiemer.com/code/colossus/`

[4] $\rho_{cz} = 3H_0^2 E_z^2/(8\pi G)$ is the critical density of the Universe at given redshift, being $E_z = \left[\Omega_\Lambda - \Omega_m(1+z)^3\right]^{0.5}$ and G the universal gravitational constant.

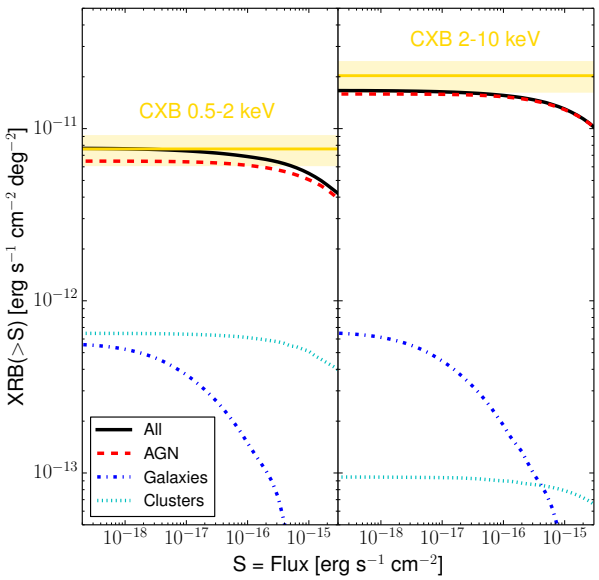

Fig. 5. Cumulative 0.5–2 keV (*left*) and 2–10 keV (*right*) fluxes per unit of square degree of the AGNs (dashed red line), nonactive galaxies (dash-dotted blue line) and galaxy clusters (dotted cyan line) simulated for the mocks used in this work. The overall distribution is plotted as a solid black line. The flux of the cosmic X-ray background measured by Cappelluti et al. (2017) is plotted in yellow, with 20% errors to account for the uncertainties in the CXB absolute flux level measured by different instruments.

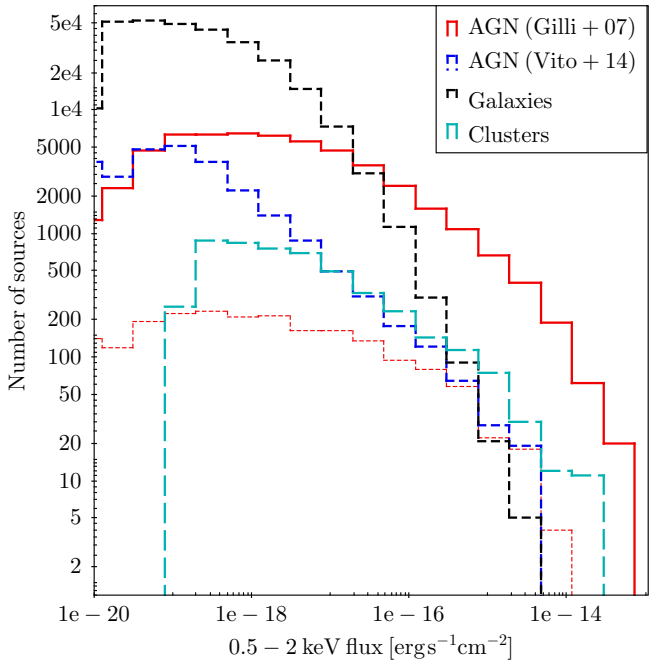

Fig. 6. Distribution of the 0.5–2 keV flux for the four mocks presented in this work: AGN from the Gilli et al. (2007) AGN population synthesis model (red solid line for the whole population; red dotted line for the $z > 3$ subsample); $z > 3$ AGN from the Vito et al. (2014) X-ray luminosity function (blue long-dashed line); nonactive galaxies (black short-dashed line); and clusters of galaxies (cyan long dashed line). All histograms refer to a 1 deg^2 field.

To summarize the properties of the catalogs presented in this section, in Fig. 6 we plot the distribution of the 0.5–2 keV observed flux for our four mocks. As can be seen, the faint end of the flux distribution (i.e., $f_{0.5-2} < 10^{-17}$ erg s^{-1} cm^{-2}, below the flux limit of currently available X-ray surveys) is dominated by nonactive galaxies. However, we expect that a fraction of sources that would be classified as nonactive galaxies are actually low-luminosity AGNs, that is, sources where the luminosity of the accreting supermassive black hole is $\log(L_{0.5-2}) < 42$, and possibly even fainter than the $\log(L_{0.5-2}) = 40$ threshold we adopted in this work. In such sources, the AGN contribution to the overall X-ray luminosity might not be the dominant one, since other processes (e.g., star-formation, diffuse gas emission, emission from ultra-luminous X-ray sources...) can produce larger X-ray luminosities, up to $\log(L_{0.5-2}) \sim 42$ (see, e.g., Ranalli et al. 2005; Lehmer et al. 2016). We therefore expect a significant fraction of $\log(L_{0.5-2}) < 40$ AGNs to be included in our mock of nonactive galaxies, as their X-ray emission would be dominated by the contribution of the galactic sources mentioned above. The number of AGN detections reported in the next sections, as well as those of future simulations with our mocks, can then be treated as a lower limit.

It is also worth noting that, at low fluxes, the AGN population is dominated by low-luminosity AGNs (i.e., objects having 0.5–2 keV luminosity below the 10^{42} erg s^{-1} threshold), which account for ∼92% of the $f_{0.5-2} < 10^{-17}$ erg s^{-1} cm^{-2} AGN population. A large part of low-flux AGNs are also heavily obscured, CT-AGNs: more in detail, ∼52% of the $f_{0.5-2} < 10^{-17}$ erg s^{-1} cm^{-2} AGN sample is made of CT sources, and 45% of the same sample is made of low-luminosity CT-AGNs. These numbers underline how deep surveys with new X-ray facilities will allow us to detect a whole new AGN population of low-luminosity, heavily obscured sources. We further discuss this topic in the next sections.

Notably, moving towards faint fluxes the high-z AGNs in the Vito et al. (2014) mock significantly outnumber those in the Gilli et al. (2007) one, as shown also in Fig. 4. Finally, clusters of galaxies contribute more significantly to the brightest end of the distribution ($f_{0.5-2} > 10^{-15}$ erg cm^{-2} s^{-1}), where they even outnumber nonactive galaxies.

3. The AXIS probe

3.1. A new, sub-arcsecond resolution X-ray facility

As an example of a practical application of our mock catalogs, we simulate a series of potential surveys with the Advanced X-ray Imaging Satellite (AXIS; Mushotzky et al. 2019)[5] mission concept. In this section, we briefly describe AXIS technical layout.

AXIS is a probe-class mission proposed to the Decadal Survey on Astronomy and Astrophysics 2020[6] with 0.3–10 keV observing band pass. The probe is designed to have a subarcsecond resolution over an area of over 500 square arcminutes, that is, over a 24′ × 24′ field of view[7]: this would be an improvement of a factor ∼100 with respect to *Chandra* ACIS-I, which only has subarcsecond resolution at off-axis angles OAA< 2′. AXIS is planned to have a remarkably low detector background, thus increasing sensitivity to extended sources, and high observing efficiency.

In Fig. 7, left panel, we plot the AXIS effective area as a function of energy. We also compare the AXIS effective area with those of *Chandra* and *XMM-Newton* pn, as well as with the predicted one for ATHENA. AXIS is expected represent a major

[5] http://axis.astro.umd.edu
[6] https://www.nationalacademies.org/our-work/decadal-survey-on-astronomy-and-astrophysics-2020-astro2020
[7] The final technical configuration of the AXIS CCDs has not yet been finalized. In this work, we simulate a single square CCD with a 24′ side.

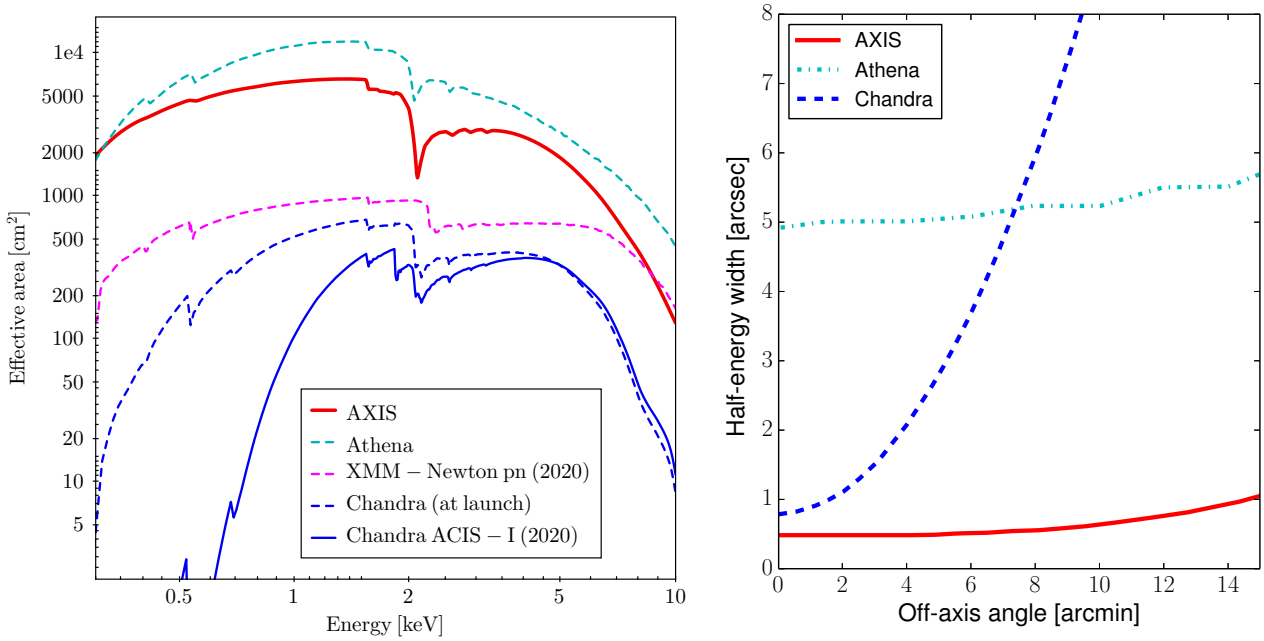

Fig. 7. *Left*: AXIS planned effective area as a function of energy (red solid red line), compared with those of the *XMM-Newton* pn camera (dashed magenta line) and of the *Chandra* ACIS-I camera (dashed blue line) as of early 2020, and with the ACIS-I one at the time *Chandra* was launched (solid blue line). The ATHENA effective area for a 5′ field of view (dashed cyan line) is also plotted for comparison. *Right*: AXIS (solid red line) angular resolution as a function of the off-axis angle. The same quantities are also plotted for ATHENA–WFI (dash-dotted cyan line; Nandra et al. 2013) and *Chandra* ACIS-I (dashed blue line).

improvement with respect to current facilities in terms of collecting area. For example, at 1 keV the planned AXIS effective is a factor ∼25 larger than the one of *Chandra* as of 2020, and a factor ∼10 larger than the one of *Chandra* at launch. AXIS is expected to produce a significant improvement, in terms of grasp, that is, of effective area multiplied by field of view, even with respect to *XMM-Newton* which is currently the best X-ray imaging telescope in that regard. At 1 keV, AXIS would collect ∼5 times more counts than *XMM-Newton* in the same amount of time.

A significant improvement in effective area would also take place at harder energies. At 6 keV (i.e., at the energy of the Iron Kα line), the AXIS effective area is expected to be a factor of ∼5 larger than the *Chandra* one, and a factor of ∼2 larger than the *XMM-Newton* one.

As previously stated, the most remarkable feature of AXIS with respect to *Chandra* is its stable point spread function (PSF) as a function of the off-axis angle: as shown in Fig. 7, right panel, the AXIS PSF (here plotted as the half-energy width) is designed to be almost constant over the whole CCD, being $\lesssim$1″ even at the edge of the field of view. This would represent a large improvement with respect to *Chandra*, whose HEW is <1″ on-axis, but becomes >5″ for off-axis angles (OAAs) >5′. Such an angular resolution would allow AXIS surveys to, for example, locate AGNs within host galaxies even at high redshifts, where mergers are expected to be common, as well as to detect binary AGNs and runaway black holes. This type of science can be performed by *Chandra* only with pointed, more time-consuming observations.

3.2. The complementary strengths of AXIS and ATHENA

If approved for funding, AXIS should be launched in the 2030s. This would allow it to work in synergy with the Advanced Telescope for High ENergy Astrophysics (ATHENA) mission.

ATHENA is the next ESA X-ray observatory mission, selected in the Cosmic Vision program to address the Hot and Energetic Universe scientific theme. It is the second L(arge)-class mission within that program and is due for launch in the early 2030s.

ATHENA will mount two instruments. One is the X-ray Integral Field Unit, a cryogenic X-ray spectrometer with planned 2.5 eV spectral resolution, 5″ pixels and a field of 5′ × 5′ (Barret et al. 2016). The other instrument is the Wide Field Imager (WFI), which will represent the ideal successor of *XMM-Newton* and can be directly compared with AXIS. We therefore plot in both panels of Fig. 7 the planned effective area as a function of energy and HEW as a function of OAA for ATHENA-WFI. The ATHENA WFI (Meidinger et al. 2017), with its 40′–diameter DEPFET chips, will be the instrument with the largest effective area in the 0.3–10 keV band: with respect to AXIS, the ATHENA collecting area will be a factor ∼2–3 times larger, thus making it the ideal instrument for X-ray spectroscopy[8].

ATHENA, similarly to AXIS, is expected to have a remarkably stable PSF over the whole field of view, with a half-energy width HEW∼5–6″. This is a factor of ∼10 larger than the AXIS one, which implies that AXIS, being significantly less affected by confusion issues than ATHENA, would be able to perform deeper surveys, as we show in Sect. 4. The largest field of view and effective area of ATHENA will instead make it ideal for large area (>10 deg^2) surveys: we further discuss the complementarity between the two missions in Sect. 5.

4. Results of the simulations of AGN surveys with AXIS

Using as an input the mock catalogs presented in Sect. 2, we simulate three different types of AXIS surveys, following the approach presented in the AXIS white paper (Mushotzky et al. 2019).

[8] Although, thanks to its low background, AXIS would also be effective in performing X-ray spectroscopy, even for high-redshift, heavily obscured sources. We discuss this in detail in Sect. 5.3.

Table 1. Properties of three reference AXIS surveys simulated in this work.

Survey	Area	Tile exposure	Total exposure	Flux limit (0.5–2 keV)	Number of detections	
	deg^2	ks	Ms	erg s^{-1} cm^{-2}	AGN	Galaxies
Deep	0.16	5000	5	5×10^{-19}	3496	5387
Intermediate	2.5	300	5	3×10^{-18}	32655	22071
Wide	50	15	5	4×10^{-17}	190149	21840

Notes. The flux limit corresponds to 1% of the covered field (see Fig. 9), and is the flux at which 1% of the input sources is detected. The number of detections is computed in the 0.5–7 keV band.

Overall, each survey requires 5 Ms of AXIS observing time: notably, these combined 15 Ms of AXIS time would amount to only 15% of the planned 5 years of observations with AXIS.

The proposed surveys are the following.

1. A deep, pencil-beam survey, that is, a 5 Ms observation of a single AXIS pointing (~0.16 deg^2, that is, ~24 × 24 square arcmin).
2. A moderate–area, moderate–depth survey, that is, 2.5 deg^2 uniformly covered with a tiling of 300 ks AXIS observations.
3. A large-area, shallow-depth survey, that is, 50 deg^2 covered with a tiling of 15 ks AXIS observations.

In Table 1 we report a summary of the properties of the three surveys, as well as the expected number of detections in each of it, based on the results reported in this work and extensively discussed in the following sections.

4.1. Simulations and source detection

Since performing `SIXTE` simulations is a time- and machine-consuming task, we chose to simulate a 0.95 deg^2 field of view[9] with the same exposure of the proposed intermediate survey (i.e., 300 ks) and a 9.5 deg^2 field of view with the same exposure of the proposed wide survey (i.e., 15 ks). We then extrapolate the results obtained on this smaller field to the proposed areas. The deep-area survey is instead simulated in its entirety.

The AGN, galaxy and cluster mock catalogs we presented in the previous sections all have flux limits well below the threshold of the deepest surveys planned with current or future instrumentation, and therefore cause the majority of the unresolved cosmic background in our analysis. The only additional component we include in our simulations is the Galactic diffuse foreground, which we model with a thermal component having surface brightness 1.8×10^{-15} erg s^{-1} cm^{-2} $arcmin^{-2}$, following the same approach described in Cucchetti et al. (2018). The particle background at 1 keV is assumed to be 7×10^{-5} cts keV^{-1} s^{-1} $arcmin^{-2}$ for AXIS (Mushotzky et al. 2019), and 6×10^{-4} cts keV^{-1} s^{-1} $arcmin^{-2}$ for ATHENA (Nandra et al. 2013).

The source detection was performed using the CIAO tool `wavdetect` on the 0.5–7 keV (Full), 0.5–2 keV (Soft) and 2–7 keV (Hard) images. These energy ranges are fairly conservative, and allow us to make a direct comparison with the results obtained by the *Chandra* surveys. It is worth noting, however, that AXIS would collect a significant amount of photons in the 0.3–0.5 keV and in the 7–10 keV energy ranges. The false-probability detection rate `SIGTHRESH` was set to a fairly conservative value, $sig = 5 \times 10^{-8}$, that is, $\sim 1/n_{pix} = 1/4096^2$, where n_{pix} is the number of pixels in our CCD. We then used a $\sqrt{2}$ wavelet sequence (i.e., 1, 1.4, 2, 2.83, 4).

We note that `wavdetect` is commonly used to perform a preliminary source detection and generate a first catalog of candidate objects which is then given as an input to other, more refined detection tools such as ACIS Extract (Broos et al. 2012, for practical applications of this tool see, e.g., Luo et al. 2017; Nanni et al. 2020). In this work we do not perform a second level source detection, but we keep only those sources with significance $\sigma > 3$[10]. Thanks to the low background and the excellent PSF over the whole field of view expected for AXIS, the choice of a $\sigma = 3$ threshold allows us to keep the number of spurious sources well below 0.5% of the total number of detections.

In the following sections, we report the results of our detection in the three simulated fields. The matches between the output catalogs obtained using `wavdetect` and the AGNs and galaxies input ones were performed assuming a maximum positional offset $d = 1''$. We first matched the output catalog with the AGN one, and then matched the sources with no AGN counterpart with the nonactive galaxies catalog. The sources with no match in either of the two input catalogs are classified as spurious and are caused by random background fluctuations.

Thanks to the PSF quality, even at larger off-axis-angles (see Fig. 7, right panel), the chosen maximum distance, while small, is actually a fairly conservative one. For example, in our simulated wide area survey the average distance between the output position and the input one is $d_{AGN} = 0.12''$ (with standard deviation $\sigma_{d,AGN} = 0.11''$) for AGNs and $d_{gal} = 0.15''$ (with standard deviation $\sigma_{d,gal} = 0.10''$) for nonactive galaxies.

4.2. AXIS Deep Field results

We report in Table 2 the results of the simulation of a 5 Ms AXIS observation over a single pointing, that is, an area of ~0.16 deg^2, while in Fig. 8 we show the three-color image of the simulated AXIS Deep field. In the same figure we also show a zoom-in of two different regions of the field: one on-axis (left panel), and one 8′ off-axis: a telescope designed to have subarcsec PSF at large off-axis angles would be an ideal survey instrument, since it would ensure high sensitivity over the whole field of view, and would make the counterpart identification process much easier. As a comparison, in the bottom right panel of the same figure we also show a 8′ off-axis image of the 4 Ms CDF-S (Xue et al. 2011), whose exposure is comparable to the simulated 5 Ms AXIS one. At these off-axis radii, the *Chandra* PSF is >5″ (see also Fig. 7, right panel), thus making the optical counterpart identification process more complicated.

[9] Due to the technical setup we adopted in `SIXTE`, the simulated area is ~5% smaller than the mock one.

[10] σ is the `SRC_SIGNIFICANCE` parameter computed by `wavdetect` and can be treated as the source signal-to-noise ratio.

Table 2. Number of sources, overall and divided by class, detected in the 0.5–7 keV, 0.5–2 keV and 2–7 keV bands in each of the three reference AXIS surveys.

Survey	Band	Total	AGN	Galaxies	Flux 20%	Flux 80%
					erg s^{-1} cm^{-2}	
Deep	0.5–7 keV	9016	3496 (38.8%)	5387 (59.7%)	8.5×10^{-18}	2.9×10^{-17}
	0.5–2 keV	8205	3051 (37.2%)	5065 (61.7%)	1.4×10^{-18}	4.4×10^{-18}
	2–7 keV	5706	3521 (60.7%)	2172 (37.5%)	9.1×10^{-18}	2.4×10^{-17}
Intermediate	0.5–7 keV	54774	32655 (59.6%)	22071 (40.3%)	3.9×10^{-17}	1.5×10^{-16}
	0.5-2 keV	47689	27555 (57.8%)	20124 (42.2%)	8.8×10^{-18}	2.8×10^{-17}
	2-7 keV	26192	23553 (90.0%)	2605 (9.9%)	7.1×10^{-17}	2.0×10^{-16}
Wide	0.5–7 keV	212058	190149 (89.7%)	21840 (10.3%)	3.6×10^{-16}	1.5×10^{-15}
	0.5–2 keV	172690	153669 (89.0%)	18974 (11.0%)	1.0×10^{-16}	2.7×10^{-16}
	2–7 keV	81443	80831 (99.2%)	581 (0.7%)	1.2×10^{-15}	3.1×10^{-15}

Notes. In parentheses, we report the fraction of sources belonging to each class with respect to the total number of sources detected in a given energy range. Spurious detections, not reported in this table, are always <0.5% of the overall detections. "Flux 20%" and "Flux 80%" are the flux limits reached over 20% and 80% of the proposed survey area, respectively: the Full and Hard band fluxes are computed in the 0.5–10 keV and 2–10 keV bands for an easier comparison with previous works, but the detections are obtained in the 0.5–7 keV and 2–7 keV bands.

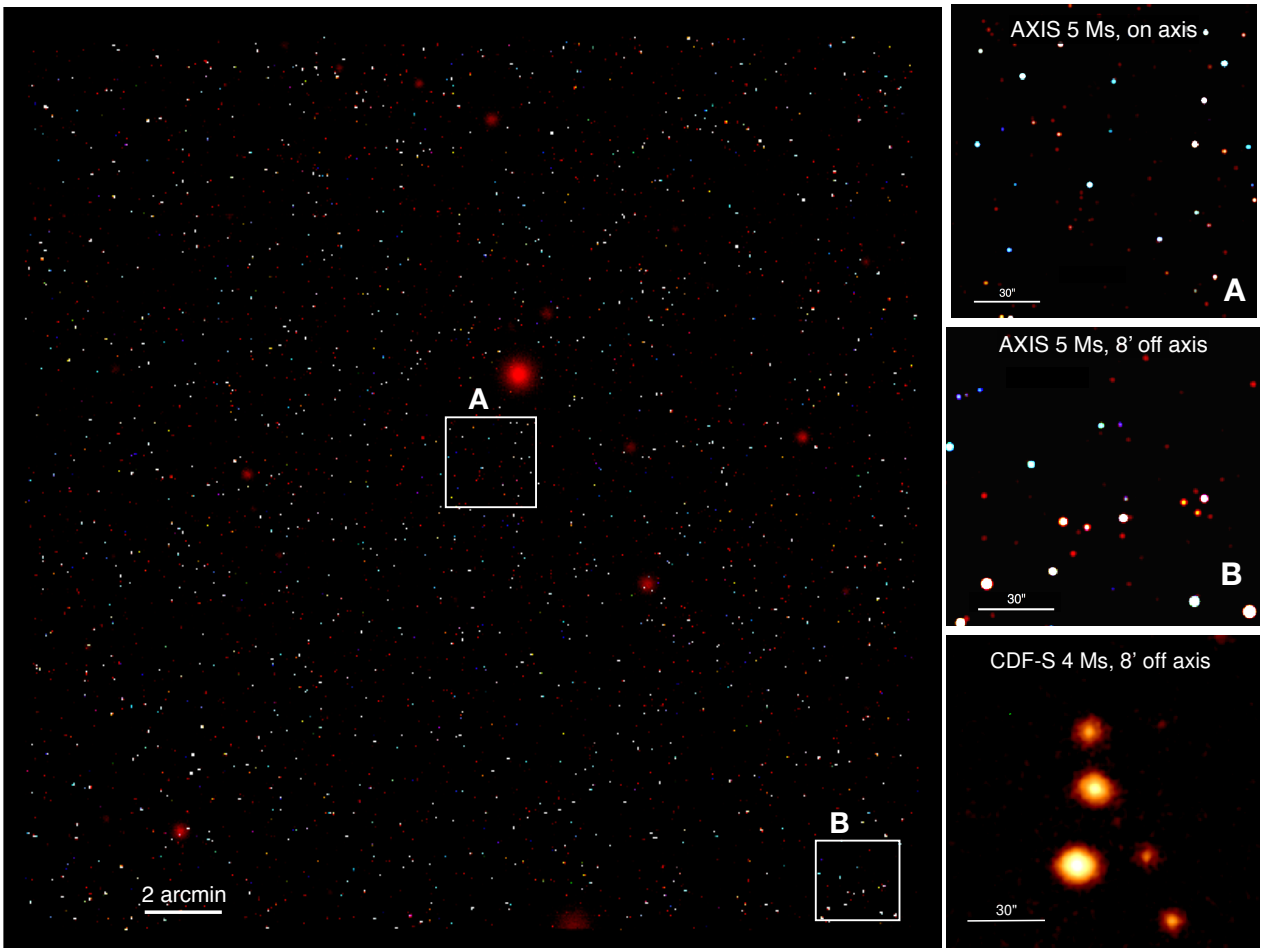

Fig. 8. *Left*: smoothed three-color image (red: 0.5–2 keV; green: 2–4.5 keV blue: 7 keV) of a simulated 5-Ms deep field (24′ × 24′) with AXIS. The extended red structures are galaxy clusters. The white boxes highlight two regions, one on-axis and the one at the boundary of the field of view: we show a zoom-in of these regions in the *top and central right panels*. The AXIS PSF is expected to have remarkably small degradation as a function of the off-axis angle. In the bottom right panel, we show as a reference a ∼2′ × 2′, 8′ off-axis 0.5–7 keV image of the 4 Ms CDF-S.

Based on our simulations, a deep survey with AXIS would detect ∼9000 sources and reach a flux limit $f_{\rm lim} \sim 5\times10^{-19}$ erg s^{-1} cm^{-2} in the 0.5–2 keV band. Here and in the following sections, the "flux limit" is the flux at which 1% of the survey area is covered (see Fig. 9). We also remind that, as mentioned in Sect. 4.1, in our analysis we chose a fairly conservative source detection threshold. It is therefore possible that the actual flux limits would be slightly fainter than the one we obtained. As a comparison, the CDF-S survey (Luo et al. 2017), which covered a field of 0.135 deg^2 with a combined 7 Ms

Chandra observation, contains 1008 sources and reaches a flux limit in the 0.5–2 keV band $f_{\rm lim} \sim 6.5 \times 10^{-18}$ erg s^{-1} cm^{-2}.

Out of these ∼9000 sources, ∼3500 (i.e., ∼39% of the sample) are AGNs. In Fig. 9, left panel, we report the completeness of the survey as a function of the 0.5–10 keV, 0.5–2 keV and 2–10 keV flux, that is, the ratio between the number of detections and the number of input sources. These curves are fully equivalent to the so-called survey sensitivity curves, and can also be used to estimate the fraction of survey area covered at a given flux. In the 0.5–2 keV band 50% of the field is covered down to a flux $\sim 3 \times 10^{-18}$ erg s^{-1} cm^{-2}, while in the 0.5–10 and 2–10 keV bands 50% of the field is covered down to a flux limit $\sim 2 \times 10^{-17}$ erg s^{-1} cm^{-2}.

As already mentioned in Sect. 2.2, our mock does not include AGNs with luminosities $L_{0.5-2} < 10^{40}$ erg s^{-1}, corresponding to the flux limit of the AXIS deep survey at redshift $z \sim 1.1$. Consequently, it is reasonable to assume that the number of AGNs detected in the AXIS deep field would be even larger than the one we obtained from our simulations at $z < 1.1$, while we do not expect the adopted luminosity threshold to impact the number of AGNs detected at higher redshift. More in detail, working under the simple assumption that the number and redshift distribution of AGNs in a dex of luminosity in the range $\log(L_{0.5-2}) = [39–40]$ are equal to those in the $\log(L_{0.5-2}) = [40–41]$ bin, we would expect to detect in the AXIS deep field ∼300–400 additional AGNs at $L_{0.5-2} < 10^{40}$ erg s^{-1}, although their identification as AGN would possibly be challenging, since non-AGN processes can produce similar, if not higher, luminosities. As a reminder, based on our simulations the AXIS Deep survey would contain ∼3500 AGNs with luminosities down to $\log(L_{0.5-2}) = 40$.

As shown in deep *Chandra* surveys such as the CDF-S 7 Ms one (Luo et al. 2017), the deeper an X-ray survey is, the larger is the fraction of nonactive galaxies it detects. As a consequence, in the simulated AXIS deep field the fraction of non-AGN sources detected in the 0.5–7 keV and 0.5–2 keV is ∼60%, that is, as opposed to shallower X-ray surveys, the majority of detected objects are not be AGNs, allowing a direct measurement of the star formation rate over a wide redshift range. In the hard X-ray band, instead, the majority of the detected sources are expected to be AGNs, but it would still be possible to detect a significant number of galaxies (>2000).

4.3. AXIS Intermediate Field results

We report in Table 2 the results of the simulation of a 300 ks AXIS observation. The simulations were performed over a 0.95 deg^2 mock field of view, and the numbers in the table have been obtained by rescaling those obtained from the simulations by 2.5/0.95, to estimate the actual number of objects which would be found by the proposed 2.5 deg^2 survey.

Overall, we expect to detect ∼55 000 (∼48 000) sources in the 0.5–7 keV (0.5–2 keV), reaching a flux limit in that same band $f_{\rm lim} \sim 10^{-17}$ erg s^{-1} cm^{-2} ($\sim 3 \times 10^{-18}$ erg s^{-1} cm^{-2}) (see Fig. 9, central panel). As a comparison, the *Chandra* COSMOS Legacy survey, which covered 2.2 deg^2 with 4.6 Ms of *Chandra* time, contained 4016 sources and reached a flux limit in the 0.5–2 keV band $f_{\rm lim} \sim 2 \times 10^{-16}$ erg s^{-1} cm^{-2}.

In this intermediate survey, the majority of detected sources would be AGNs: more in detail, in the Full band we should detect more than 30 000 AGNs, which account for ∼60% of the overall detections in the same band. In Fig. 9, central panel, we report the completeness of the survey as a function of the 0.5–10 keV, 0.5–2 keV and 2–10 keV flux: in the 0.5–2 keV band 50% of the field is covered down to a flux limit $\sim 2 \times 10^{-17}$ erg s^{-1} cm^{-2}, while in the 0.5–10 and 2–10 keV bands 50% of the field is covered down to a flux limit $\sim 8 \times 10^{-17}$ erg s^{-1} cm^{-2} and $\sim 10^{-16}$ erg s^{-1} cm^{-2}, respectively.

4.4. AXIS Wide Field results

We report in Table 2 the results of the simulation of a 15 ks AXIS pointing: the simulations were performed over a 9.5 deg^2 mock field of view. The numbers in the table have then been obtained by rescaling the simulations one by 50/9.5, to estimate the actual number of objects which would be found by the proposed 50 deg^2 survey.

Overall, we expect to detect more than 210 000 (170 000) sources in the 0.5–7 keV (0.5–2 keV) band, reaching a flux limit in the same band $f_{\rm lim} \sim 10^{-16}$ erg s^{-1} cm^{-2} ($\sim 4 \times 10^{-17}$ erg s^{-1} cm^{-2}). As a comparison, the XMM-XXL Survey (Pierre et al. 2016; Chiappetti et al. 2018) covered 50 deg^2 with a total of over 6 Ms of *XMM-Newton* time to detect 26 056 sources down to a flux limit $f_{\rm lim} \sim 10^{-15}$ erg s^{-1} cm^{-2} in the 0.5–2 keV band. The all-sky X-ray instrument eROSITA (Merloni et al. 2012), which was launched in July 2019 and covers the 0.5–10 keV energy range, is instead expected to detect 2.7 millions of AGNs band over four years of observations: the planned flux limit at 1% of the area covered (i.e., ∼40 deg^2) is $\sim 2 \times 10^{-15}$ erg s^{-1} cm^{-2} in the 0.5–2 keV band (Comparat et al. 2019).

While the AXIS survey plan does not include, as of today, multiple pointings of the same field, it is worth noting that the AXIS wide field would be an ideal region for follow-up observations aimed at finding AGN variability. For example, 1 Ms of AXIS time could be spent covering several square degrees of the AXIS wide field with the same depth of the original survey, one or more times. This would allow us to study AGN variability in samples of thousands of AGNs, at fluxes >1 dex fainter than those sampled by the all-sky telescope eROSITA.

5. The high-redshift Universe as seen by ATHENA and AXIS

5.1. Extragalactic surveys with ATHENA–WFI

As already mentioned in Sect. 3, the technical specifics of ATHENA-WFI, in particular its large field of view and effective area, and its stable PSF, make it an ideal survey instrument. A total of 23.6Ms (nearly 25% of the expected four years of mission lifetime) has been devoted in the Mock Observing Plan for two extragalactic surveys.

These surveys have been designed since the mission proposal (Aird et al. 2013; Nandra et al. 2013) to address several open topics on early SMBH-galaxy evolution. For example, ATHENA will (i) study the mass accretion of the earliest growing SMBHs at $z > 6$; (ii) find distant evolved groups of galaxies with formed hot gaseous atmosphere at $z > 6$; (iii) determine the accretion energy density in the Universe, by measuring the X-ray luminosity function and obscuration properties of the AGN population up to $z = 4$; (iv) determine the incidence of strong and ionized outflowing absorbers (including ultra-fast outflows) among the luminous AGN population from $z = 1$ to $z = 4$.

For each of these scientific goals a specific requirement has been derived: for example, to address scientific goal (*i*) one needs to detect at least 10 AGNs with $43 < \log L_{0.5-2\,\rm keV} < 43.5$ at $z = 6–7$ and at least 10 AGNs with $44 < \log L_{0.5-2\,\rm keV} < 44.5$ at $z = 7–8$. Such a requirement can then be translated into a survey sensitivity goal (in terms of point source sensitivity vs area) for the given mission specifications. All together

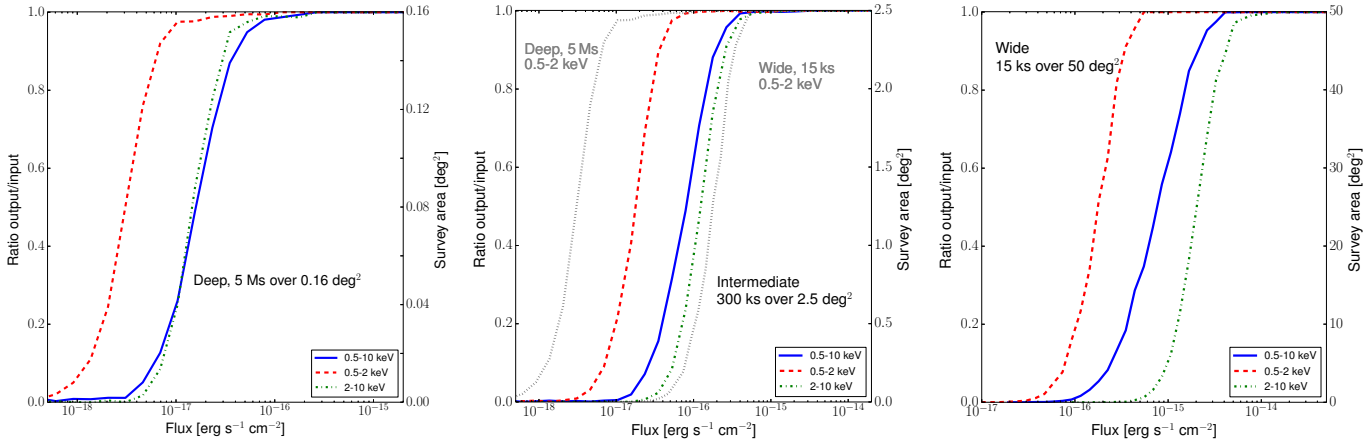

Fig. 9. Survey completeness, that is, ratio between number of detected sources and number of simulated sources, and corresponding survey area for AGNs detected in the 0.5–7 keV (solid blue line), 0.5–2 keV (dashed red line) and 2–7 keV (dash-dotted green line). In the *left panel* we report the curves obtained in the AXIS deep survey, in the *central panel* those obtained in the intermediate-area survey, and in the *right panel* those obtained in the wide-area survey. In the *central panel*, the 0.5–2 keV completeness curves for the deep and wide surveys (dotted gray lines) are also plotted for comparison.

these requirements set the exposure request for the two planned surveys: that is, 4×1.5 Ms $+ 3 \times 1.05$ Ms $+ 5 \times 950$ ks pointings over the 5.28 deg^2 of the deep survey (mainly driven by the high-z AGN search and Compton-thick AGN characterization), and 108×90 ks pointings over the 47.52 deg^2 of the wide one (mainly driven by high-z AGN and first groups search).

To make a comparison with the AXIS results, we have performed `SIXTE` simulations with the same mock catalogs described in Sect. 2, coupled with the most up-to-date calibration files and matrices available for ATHENA. The cosmic background and galactic foreground are the same used for AXIS, while the particle background is modeled using a flat power law with normalization 6×10^{-4} cts keV^{-1} s^{-1} arcmin^{-2}, as defined in the instrument scientific goals (Nandra et al. 2013).

We performed a full `SIXTE` simulation of the deep tier of the survey (12 pointings in total), while for the wide tier we simulated only 10 fields, and rescaled the result to the planned 108 pointings. The detection process is the same described in Sect. 4, except for the false-probability detection rate `SIGTHRESH` in `wavdetect`, which was set to 10^{-6} , that is, $\sim 1/n_{\rm pix} = 1/1024^2$. The detection has been performed in the 0.5–7, 0.5–2, and 2–7 keV bands.

In Fig. 10 we report the area versus 0.5–2 keV (left panel) and 2–10 keV (right panel) flux curve for the simulated AXIS (black lines) and ATHENA (red lines) surveys, as well as for several existing *Chandra* and *XMM-Newton* surveys, namely CDF-S 7 Ms (Luo et al. 2017), CDF-N 2 Ms (Xue et al. 2016), AEGIS XD (Nandra et al. 2015), SSA22 (Lehmer et al. 2009), J1030 (Nanni et al. 2020), XDEEP-2 F1 (Goulding et al. 2012), *Chandra* COSMOS Legacy (Civano et al. 2016), X-Bootes (Murray et al. 2005), Stripe 82X (LaMassa et al. 2013b,a, 2016) and XMM-XXL (Pierre et al. 2016). In the left panel, we also report the 0.5–2 keV predictions for four years of eROSITA observations (i.e., eRASS:8; Merloni et al. 2012; Comparat et al. 2019, dashed light green line).

As it can be seen, a survey program such as those planned for ATHENA and AXIS would represent a major improvement with respect to currently available surveys: particularly, AXIS surveys would be able to sample fluxes ∼25 and ∼50 times deeper than those reached by currently available surveys at 1 and 50 deg^2, respectively. AXIS and ATHENA would complement each other: the larger collecting and effective area of ATHENA would allow one to reach unprecedentedly deep fluxes over 10s of square degrees, while the AXIS PSF quality (<1″ over the whole field of view) would represent a major step forward for deep– and intermediate–area surveys.

5.2. A new X-ray view of the high-redshift Universe

5.2.1. An unprecedented AGN statistic at high redshift

In Table 3 we report the expected number of high-z sources expected in each AXIS survey in the two sets of simulations we performed (i.e., one using the Gilli et al. 2007 mock, the other using the Vito et al. 2014 mock). Regardless of which of the two mock catalogs will turn out to be the most reliable one, all surveys would provide a significant improvement to the $z > 3$ statistic of X-ray selected AGNs: we in fact may expect to detect ∼130–350 $z > 3$ AGNs in the deep survey, ∼1000–2000 $z > 3$ AGNs in the intermediate survey and ∼6000–6700 $z > 3$ AGNs in the wide-area survey. This would be a major leap forward with respect to currently available X-ray datasets. For example, the deepest *Chandra* survey, the CDF-S 7 Ms one, contains ∼70 $z > 3$ AGNs (Vito et al. 2018) over an area of ∼330 arcmin^2; the 2.2 deg^2 *Chandra* COSMOS Legacy survey, which required an overall 4.6 Ms *Chandra* exposure, contains 174 $z > 3$ AGNs (Marchesi et al. 2016b); and the 31 deg^2 Stripe 82X survey, which combined ∼500 ks of *Chandra* time and ∼1 Ms of *XMM-Newton* time for an overall 1.5 Ms X-ray exposure, contains 45 $z > 3$ Ananna et al. (2017).

Even more importantly, the AXIS surveys would allow us to detect for the first time a population of X-ray-selected $z > 6$ AGNs: we expect to detect a total of ∼30–100 of these primordial accreting supermassive black holes from the three surveys. If the predictions of the Vito et al. (2014) XLF are confirmed, we also expect to detect sources up to $z \sim 8$, possibly enabling the direct detection of late-stage accreting SMBHs seeds.

In Table 4 we report the total number of high-z sources detected in the ATHENA simulations. Overall, ATHENA will detect 9000–11 500 $z > 3$ AGNs, that is, about 20% more than what is expected from AXIS surveys (whose reference survey plan is currently 50% shorter).

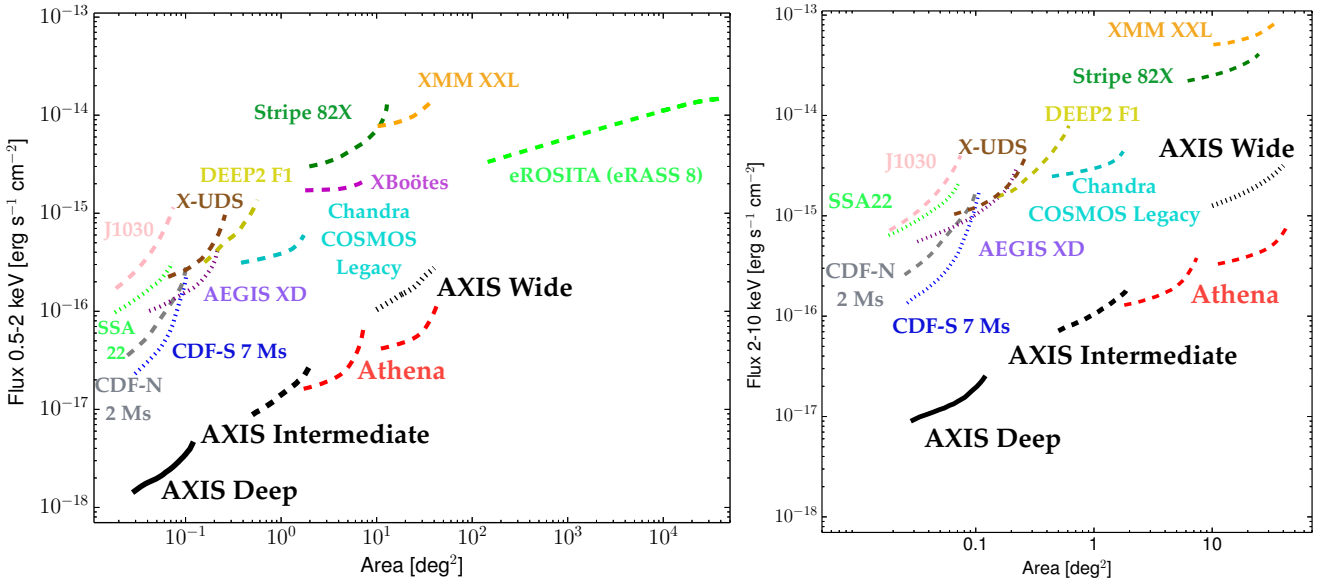

Fig. 10. 0.5–2 keV (*left*) and 2–10 keV (*right*) area-flux curves for the AXIS deep, 5 Ms (solid black line), intermediate, 300 ks (dashed black line) and wide, 15 ks (dotted black line) reference surveys. For comparison, we show the area-flux curves of several existing X-ray surveys: CDF-S 7 Ms (Luo et al. 2017, dotted blue line); CDF-N 2 Ms (Xue et al. 2016, dashed gray line); AEGIS XD (Nandra et al. 2015, dotted purple line); SSA22 (Lehmer et al. 2009, dotted light green line); XDEEP-2 F1 (Goulding et al. 2012, dashed yellow line); J1030 (dashed pink line; Nanni et al. 2020); X-UDS (Kocevski et al. 2018, dashed brown line); *Chandra* COSMOS Legacy (Civano et al. 2016, dashed cyan line); X-Bootes (Murray et al. 2005, dashed magenta line); Stripe 82X (LaMassa et al. 2013b,a, 2016, dashed green line); and XMM-XXL (Pierre et al. 2016, dashed orange line). We also show the predictions made for the ATHENA deep– and wide–area survey (dashed red lines; see the text for more details), and those for four years of eROSITA observations (i.e., eRASS:8; Merloni et al. 2012; Comparat et al. 2019, dashed light green line). The plotted lines have been derived from the 0.5–2 keV survey sensitivity curves in an area range that starts at 20% and stops at 80% of the area covered by the survey.

Table 3. Number of high-redshift sources detected in each AXIS simulation, and logarithm of the 0.5–2 keV luminosity corresponding to the flux at which 20% of the survey area is covered, see Table 2, at different high-z thresholds.

	Gilli et al. (2007) mock										Vito et al. (2014) mock				
AXIS Survey	$z>3$		$z>4$		$z>5$		$z>6$		$z>7$		$z>3$	$z>4$	$z>5$	$z>6$	$z>7$
	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$n_{\rm src}$	$n_{\rm src}$	$n_{\rm src}$	$n_{\rm src}$
Deep	127	41.1	27	41.3	7	41.5	2	41.7	1	41.9	343	63	15	8	3
Intermediate	1066	41.8	200	42.1	61	42.3	11	42.5	2	42.6	1948	389	103	29	8
Wide	5997	42.9	966	43.2	127	43.4	16	43.6	11	43.7	6690	1291	321	63	23

Notes. The number of detections are computed rescaling the numbers obtained in the simulations to the survey proposed area (2.5 deg^2 for the 300 ks, intermediate-area survey and 50 deg^2 for the 15 ks, wide-area survey). On the left, we report the results obtained using the Gilli et al. (2007) mock, while on the right we report the results obtained using the Vito et al. (2014) mock (more details in the text).

Table 4. Number of expected high-redshift detections in the deep and wide tier of the ATHENA–WFI survey simulation, and logarithm of the 0.5–2 keV luminosity limit (computed using the flux limit at 20% of the survey area).

	Gilli et al. (2007) mock										Vito et al. (2014) mock				
ATHENA survey	$z>3$		$z>4$		$z>5$		$z>6$		$z>7$		$z>3$	$z>4$	$z>5$	$z>6$	$z>7$
	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$n_{\rm src}$	$n_{\rm src}$	$n_{\rm src}$	$n_{\rm src}$
Deep	1744	42.1	326	42.4	68	42.6	15	42.8	6	42.9	2667	534	147	38	15
Wide	7344	42.5	1350	42.8	259	43.0	54	43.2	22	43.3	8878	1831	416	85	23

5.2.2. Constraints on theoretical models of early black hole accretion

In Fig. 11 we plot the AGN 0.5–2 keV luminosity as a function of redshift for the AXIS and ATHENA $z>3$ samples. While ATHENA will collect more sources, AXIS is expected to sample luminosities ∼1 dex fainter up to redshift ∼8, once again highlighting the complementarity between the two instruments.

In Fig. 11 we also report the evolution with redshift of the X-ray luminosity of two idealized BHs growing to $\mathrm{Log}(M_{\rm BH}/M_{\odot})=9.1$ at $z=6$ through continuous accretion. Such large SMBHs are commonly found to power luminous

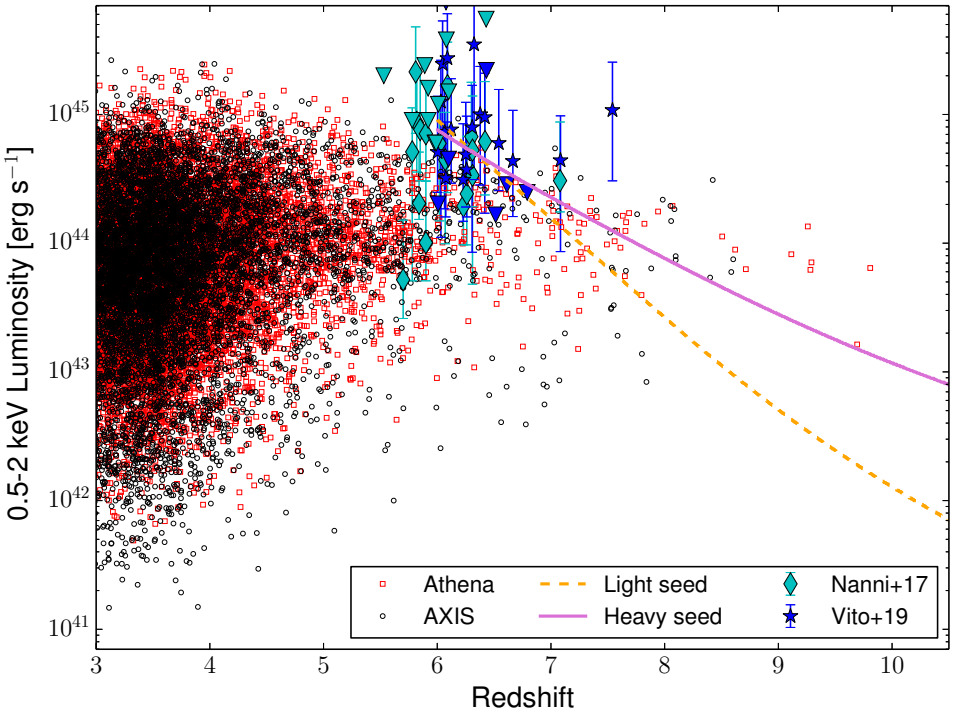

Fig. 11. 0.5–2 keV luminosity as a function of redshift for the $z > 3$ sources detected in the simulated AXIS (black circles) and ATHENA (red squares) surveys. 1σ uncertainties are computed using the Gehrels (1986) equations. The $z > 5.5$ AGNs detected in the X-rays by currently available facilities are also plot for comparison (Vito et al. 2019, blue stars; Nanni et al. 2017, cyan diamonds). Two different models of black hole seed accretion are also shown (light SMBH seed: dashed yellow line; heavy SMBH seed: solid magenta line; see the text for more details).

($L_X \sim 10^{45}$ erg s^{-1}) quasars at $z = 6$ in current wide-area optical surveys, such as SDSS (Fan et al. 2006) and PanSTARRS (Bañados et al. 2016).

In the general case of an accreting BH radiating at a given fraction $\lambda \equiv L_{\rm bol}/L_E$ of its Eddington luminosity L_E and with constant efficiency ϵ, the BH mass grows as $M(t) = M_{\rm seed} e^{t/t_{\rm Sal}}$, where $M_{\rm seed}$ is the BH seed mass and $t_{\rm Sal}$ is the Salpeter e-folding time: $t_{\rm Sal} = 50\ {\rm Myr}\ \left(\frac{9\epsilon}{1-\epsilon}\right)\lambda^{-1}$. Under these assumptions, the QSO bolometric luminosity also grows exponentially, that is, $L_{\rm bol}(t) = L_{\rm bol,0} e^{t/t_{\rm Sal}}$, where $L_{\rm bol,0}$ is the bolometric luminosity at the beginning of the accretion. We assume $\epsilon = 0.1$ and converted from bolometric to 0.5-2 keV luminosities using the recent AGN bolometric corrections of Duras et al. (2020). We rescale the 2-10 keV band luminosities considered by Duras et al. (2020) to the 0.5-2 keV band using a photon index of $\Gamma = 1.9$.

We consider two different seed masses that bracket the range proposed by theory: (i) a light seed with mass $M_{\rm seed} = 10^2\ M_\odot$, similar to those expected from the remnants of the first, PopIII stars, and (ii) a heavy seed with $M_{\rm seed} = 10^5\ M_\odot$. Under favorable environmental conditions, such massive seeds may form through the direct collapse of large, pristine gas clouds (see Inayoshi et al. 2020, for a recent review on the formation of early BHs). In order to produce a $10^9\ M_\odot$ BH by $z = 6$, light seeds must be continuously accreting at their Eddington limit ($\lambda = 1.0$) for ~820 Myr (i.e., since $z_{\rm start} = 30$; yellow dashed curve). Trivially, heavy seeds are able to produce the same $10^9\ M_\odot$ masses at $z = 6$ by accreting over a shorter period of time (~670 Myr; $z_{\rm start} = 16$) and with lower Eddington ratios ($\lambda = 0.7$; magenta solid curve). Clearly, these growth models are oversimplified, but, as shown in Fig. 11, in principle both ATHENA and AXIS surveys would have the sensitivity to detect the progenitors of SDSS QSOs. Based on our simulations, these surveys are expected to detect ~80 AGNs at $z \geq 7$. Among this high-z AGN population, the two instruments might be able to track the progenitors of SDSS QSOs up to $z \sim 8$ ($z \sim 9$) if they grow from light (heavy) seeds.

In Fig. 12, we show the AXIS and ATHENA X-ray luminosity functions at $z \sim 6.5$ and $z \sim 7.5$. As it can be seen, the two surveys would nicely complement each other, with AXIS reaching ~1 dex lower luminosities (down to $\log(L_{2-10\,\rm keV} \sim 42$ at $z = 6.5$) and ATHENA achieving a better source statistic at 2–10 keV luminosities $\gtrsim 5 \times 10^{43}$ erg s^{-1}.

In Fig. 12, we also plot the XLF predictions from the Vito et al. (2014) model, that we used as a reference for our mock: consequently, the simulated XLFs closely match the Vito et al. (2014) one. In the same figure, we show the range of predictions of several different models of SMBH early accretion, both hydro-dynamical (namely, EAGLE, Crain et al. 2015; Schaye et al. 2015; McAlpine et al. 2016, Horizon-AGN, Dubois et al. 2014, Illustris, Vogelsberger et al. 2014, and MassiveBlackII, Khandai et al. 2015) and semi-analytical (GALFORM, Cole et al. 2000; Lacey et al. 2016, L-Galaxies Guo et al. 2011; Henriques et al. 2015, MERAXES Mutch et al. 2016; Qin et al. 2017, and SHARK Lagos et al. 2018). The range of predictions has been taken from Amarantidis et al. (2019), to which we refer for a complete description of the models (see also Ni et al. 2020, for another theoretical prediction).

As it can be seen, the basically nonexistent observational evidence at $z > 7$ is reflected in the large discrepancy (~3 dex in the luminosity range 10^{43}–10^{44} erg s^{-1}) between the predictions of the theoretical models. This issue would be significantly addressed by the launch of AXIS and ATHENA: in fact, while these two instruments are not expected to directly detect the first BH seeds, they would allow us to constrain the AGN $z > 7$ XLF with uncertainties <50% over a wide range of luminosities. This would enable an unprecedented tuning of the theoretical SMBH accretion models, allowing us to rule out many combinations of parameters.

5.3. Heavily obscured black hole accretion in the early Universe

Based on all models of AGN population synthesis (see, e.g., Gilli et al. 2007; Ballantyne et al. 2011; Ananna et al. 2019) and

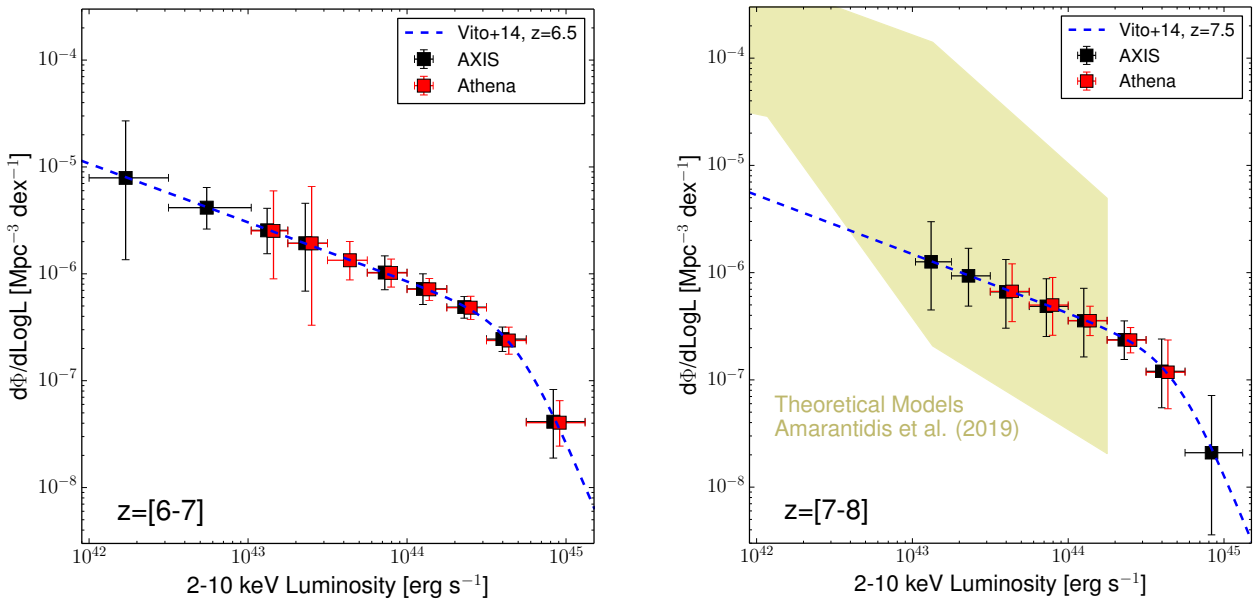

Fig. 12. Expected X-ray luminosity function at $z \sim 6.5$ (*left*) and $z \sim 7.5$ (*right*), of AXIS (black) and ATHENA (red), respectively. The Vito et al. (2014) XLFs at the same redshifts are plotted as a blue dashed line and a solid cyan line, respectively. In the right panel, the range of predictions from the hydro-dynamical simulations and semi-analytical models discussed in Amarantidis et al. (2019, khaki area) is also shown for comparison.

Table 5. Expected number of Compton thick AGN (i.e., sources having column density $N_{\rm H} > 10^{24}\,{\rm cm}^{-2}$) in each of the planned AXIS and ATHENA surveys.

	AXIS						ATHENA					
	Whole sample				>100 net counts		Whole sample				>100 net counts	
Survey	All		$z > 3$		All	$z > 3$	All		$z > 3$		All	$z > 3$
	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$n_{\rm src}$	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$L_{\rm lim}$	$n_{\rm src}$	$n_{\rm src}$
Deep	850	37.0	62	41.6	452	31	4476	38.1	218	43.8	2821	131
Intermediate	4396	37.8	268	42.3	335	10	–	–	–	–	–	–
Wide	6113	38.9	172	44.5	69	0	7236	38.5	236	44.2	594	0

Notes. We report both the overall results and the number of sources for which more than 100 0.5–7 keV net counts are detected, thus making possible a reliable X-ray spectral analysis. The logarithm of the 0.5–2 keV luminosity is computed using the spectral model described in Sect. 5.3 and a 0.5–2 keV flux corresponding to the one at which 20% of the survey area is covered (see Table 2).

on the available observational evidence, it is well established that most of the mass accretion in SMBHs takes place in an obscured phase, where the column density of the material surrounding the SMBH on pc scales is $N_{\rm H} > 10^{22}\,{\rm cm}^{-2}$. Indeed, a large part of AGNs are expected to be Compton thick (CT-), that is, with column density $N_{\rm H} > 10^{24}\,{\rm cm}^{-2}$: as shown in Fig. 1, right panel, in the mock we used in this work ∼45% of the AGNs are Compton thick (see also, e.g., Georgakakis et al. 2013).

While, based on this information, we expect to detect a large number of CT-AGNs within the AXIS surveys, for a proper characterization of these objects the ∼5–10 0.5–7 keV net counts needed for a detection are not enough. Instead, one needs to detect >100 net counts to perform a X-ray spectral analysis and constrain with reliable uncertainties the obscuring material column density.

In Table 5 we report the expected number of CT-AGNs detected in each of the three AXIS surveys, both in the overall sample and at $z > 3$: the results are derived using the Vito et al. (2014) high-z mock catalog. While we expect to detect most CT-AGNs in the wide area survey, only the intermediate and the deep X-ray surveys would allow us to properly characterize a significant number of CT sources through X-ray spectroscopy. Overall, the AXIS surveys are expected to contain ∼850 CT-AGNs with >100 net counts in the 0.5–7 keV band, ∼40 of which at $z > 3$: in currently available *Chandra* and *XMM-Newton* surveys, fewer than five $z > 3$ CT-AGN are detected with >100 net counts in the 0.5–7 keV band (see, e.g., Brightman et al. 2014; Lanzuisi et al. 2018; Corral et al. 2019). Similar surveys would then allow a full X-ray characterization of heavily obscured sources at the peak of the black hole accretion history.

In Table 5 we also report the number of Compton thick AGNs detected in the two proposed ATHENA surveys. The large ATHENA grasp, combined with its planned deep exposure, will be particularly handy to find and characterize large numbers of these otherwise elusive objects. In particular, we expect to detect ∼130 $z > 3$ CT-AGN with more than 100 0.5–7 keV net counts in the two ATHENA surveys. While AXIS is expected to detect fewer high-z CT-AGN with >100 0.5–7 keV net counts, it would also generally sample a population of obscured sources intrinsically less luminous than those detected by ATHENA.

As we mentioned earlier in the text, and as shown in Fig. 3, in heavily obscured sources the >10 keV flux predicted by the physically motivated `borus02` model is 30–60% fainter than the one predicted by the `pexmon` one. Consequently, we expect that the number of CT sources reported in Table 5 might be somewhat over-estimated.

More quantitatively, and focusing at first on the high-z subsample, the observed 0.5–2 keV (2–7 keV) band corresponds, at $z = 3$, to the 2–8 keV (8–28 keV) rest-frame one. In these energy ranges, the ratio between the `borus02` and `pexmon` fluxes are $r_{0.5-2} = 1.07$ and $r_{2-10} = 0.47$. As a consequence, we expect that the choice of `pexmon` instead of `borus02` does not affect the number of CT-AGN detections in the 0.5–2 keV band in our simulation. To instead check how significant is the effect in the 2–10 keV band, we rescale by r_{2-10} the 2–10 keV fluxes of the $z > 3$ CT sources detected in our simulated surveys. We find that in the AXIS intermediate and wide surveys ∼10–15% of the CT-AGNs originally detected would have a flux fainter than the one corresponding to 20% of the area covered by the survey (see Table 2) and would likely be missed. In the AXIS deep field, the fraction of missed objects is even smaller, being ∼5%. In ATHENA we expect to miss ∼10% of the $z > 3$ CT-AGNs in the deep survey, and ∼15% in the wide one.

Finally, we note that, at $z \sim 1$ (i.e., the average redshift of the CT-AGNs detected in AXIS and ATHENA, and the redshift where the bulk of CT-AGNs would be detected) the ratio between the rest-frame `borus02` and `pexmon` 0.5–7 keV flux is ∼ 1. We therefore do not expect the overall number of CT detections to change significantly.

In Fig. 13 we show the AXIS spectrum of a Compton thick AGNs at $z = 3.5$ simulated using `borus02`, as it would be observed in the AXIS deep, 5 Ms survey. It is worth pointing out, however, that `borus02` does not take into account off-nuclear absorption, caused by the interstellar medium in the host galaxy, which can be significant in high-z AGNs (see, e.g., Circosta et al. 2019; D'Amato et al. 2020; Ni et al. 2020).

The source has column density $\mathrm{Log}(N_{\rm H} = 24.5$ and 0.5–2 keV rest-frame, absorption corrected luminosity $\mathrm{Log}(L_{0.5-2} = 44.5)$. Such a luminosity corresponds to an observed flux in the same band $f_{0.5-2} \sim 2 \times 10^{-16}$, and would lead to the detection of ∼250 net counts in the 0.5–7 keV band in a 5 Ms AXIS observations. With such a count statistic, it would be possible to measure the line-of-sight column density with 90% confidence uncertainties ≤30%, as shown in the inset of Fig. 13. Based on our simulations, the deep and intermediate AXIS surveys should detect ∼20 CT-AGN at $z > 3$ at least as bright as the one shown in Fig. 13.

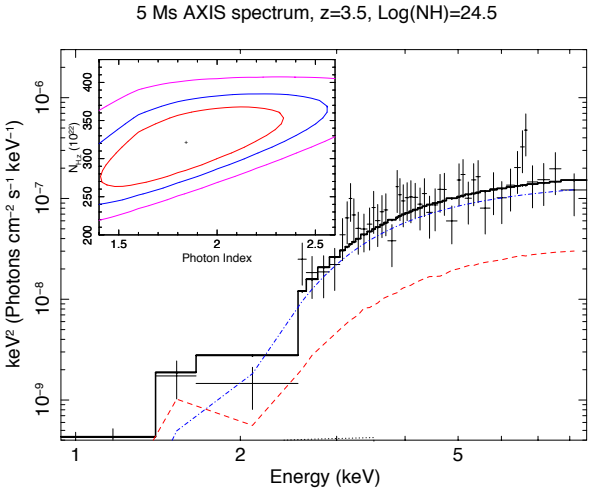

Fig. 13. Spectrum of a $z = 3.5$ CT-AGN with $\mathrm{Log}(N_{\rm H}) = 24.5$ and 0.5–2 keV intrinsic luminosity $\mathrm{Log}(L_{0.5-2}) = 44.5$, as it would be detected in the AXIS deep, 5 Ms survey. The spectrum is rebinned with a minimum of 5 counts per bin. In the insets, we show the 68, 90 and 99% confidence level contours for the power law photon index Γ and the line-of-sight column density $N_{\rm H}$.

6. Summary and conclusions

In this paper, we presented new mock catalogs of active galactic nuclei, nonactive galaxies and clusters of galaxies for the simulations of X-ray surveys. The mocks are made available online also in the `SIXTE` format, which makes them ready-to-use for any type of simulation.

All the mocks are derived from the most up-to-date observational evidence and its extrapolations to fluxes not yet sampled by current X-ray facilities. In particular, two different AGN mocks have been derived using the Gilli et al. (2007) AGN population synthesis model and, at redshift $z > 3$, the Vito et al. (2014) X-ray luminosity function. These mocks reach 0.5–2 keV luminosities $L_{0.5-2} = 10^{40}$ erg s^{-1} and fluxes in the same band $f_{0.5-2} = 10^{-20}$ erg cm^{-2} s^{-1}, that is, way below the flux and luminosity limits of current X-ray facilities. Therefore, our mocks can be easily used both to simulate X-ray surveys with current facilities, such as *Chandra*, *XMM-Newton*, or the eROSITA all-sky mission, and to make predictions for future instruments, such as the forthcoming ATHENA mission, or the AXIS and *Lynx* mission concepts.

In particular, in this work we used our mocks to simulate a set of surveys with ATHENA and the proposed AXIS probe. We find that these future, next generation surveys may transform our knowledge of the deep X-ray Universe. Some examples are as follows.

1. As shown in Table 2, AXIS would lead to the detection of over 275 000 X-ray sources at a $>3\,\sigma$ significance level. About 80% of these objects are expected to be AGNs, while the remaining 20% (i.e., ∼50 000 sources), are expected to be nonactive galaxies. Based on a comparison with currently available X-ray surveys, ∼90% of these objects would be detected in the X-rays for the first time.
2. The combination of ATHENA and AXIS would be strategical to improve our knowledge of the high-z redshift Universe, which has not been significantly explored by current X-ray facilities. Based on our simulations, ATHENA and AXIS are expected to detect ∼20 000 $z > 3$ AGNs, that is, a factor ∼60 more than those detected with current X-ray facilities. This would allow us to investigate for the first time with unprecedented statistics the history of SMBH accretion before the peak of AGN activity. The two instruments would also detect $\gtrsim$70 $z > 7$ AGNs, and would make it possible to constrain the AGN X-ray luminosity function at $z \sim 7$ down to luminosities $L_{2-10\,\rm keV} \sim 10^{43}$ erg s^{-1}. This would enable an unprecedented tuning of theoretical SMBH accretion models.
3. ATHENA and AXIS are expected to detect thousands of heavily obscured AGNs at high redshift: these are the objects which are thought to accrete more efficiently, but their large column density has made them invisible to current X-ray facilities. For a subsample of over 150 of these $z > 3$ CT-AGNs, the two instruments would detect more than 100 net counts, thus allowing us to reliably constrain the obscuring material column density. This will enable an accurate census of the AGN population before the peak of AGN activity at $z \sim 2$–3.

Finally, we expect that the proposed surveys would enable a significant amount of "serendipitous science". More in detail, the planned combination of area and depth will enable the high-significance detection of many different targets of interest, such as tidal disruption events, high-z star-forming galaxies, low-z non-active galaxies, and many others.

Acknowledgements. We thank the anonymous referee for the useful suggestions, that helped in significantly improving the paper, and Stephanie LaMassa for her helpful explanation of the Stripe 82X survey technical layout. We acknowledge financial contribution from the agreement ASI-INAF n. 2017-14-H.O. S.E. acknowledges financial contribution from the contract ASI-INAF Athena 2015-046-R.0 and from INAF "Call per interventi aggiuntivi a sostegno della ricerca di main stream di INAF".

References

Aird, J., Comastri, A., Brusa, M., et al. 2013, ArXiv e-prints [arXiv:1306.2325]
Aird, J., Coil, A. L., Georgakakis, A., et al. 2015, MNRAS, 451, 1892
Amarantidis, S., Afonso, J., Messias, H., et al. 2019, MNRAS, 485, 2694
Ananna, T. T., Salvato, M., LaMassa, S., et al. 2017, ApJ, 850, 66
Ananna, T. T., Treister, E., Urry, C. M., et al. 2019, ApJ, 871, 240
Arnaud, K. A. 1996, in Astronomical Data Analysis Software and Systems V, eds. G. H. Jacoby, & J. Barnes, ASP Conf. Ser., 101, 17
Bañados, E., Venemans, B. P., Decarli, R., et al. 2016, ApJS, 227, 11
Baldi, A., Ettori, S., Molendi, S., et al. 2012, A&A, 537, A142
Balestra, I., Tozzi, P., Ettori, S., et al. 2007, A&A, 462, 429
Ballantyne, D. R., Draper, A. R., Madsen, K. K., Rigby, J. R., & Treister, E. 2011, ApJ, 736, 56
Baloković, M., Brightman, M., Harrison, F. A., et al. 2018, ApJ, 854, 42
Barger, A. J., Cowie, L. L., Bauer, F. E., & González-López, J. 2019, ApJ, 887, 23
Barret, D., Lam Trong, T., den Herder, J. W., et al. 2016, in Proc. SPIE, SPIE Conf. Ser., 9905, 99052F
Bi, S., Feng, H., & Ho, L.C. 2020, ApJ, 900, 124
Brightman, M., Nandra, K., Salvato, M., et al. 2014, MNRAS, 443, 1999
Broos, P., Townsley, L., Getman, K., & Bauer, F. 2012, AE: ACIS Extract
Buchner, J., Georgakakis, A., Nandra, K., et al. 2015, ApJ, 802, 89
Capak, P. L., Riechers, D., Scoville, N. Z., et al. 2011, Nature, 470, 233
Cappelluti, N., Li, Y., Ricarte, A., et al. 2017, ApJ, 837, 19
Chiappetti, L., Fotopoulou, S., Lidman, C., et al. 2018, A&A, 620, A12
Circosta, C., Vignali, C., Gilli, R., et al. 2019, A&A, 623, A172
Civano, F., Marchesi, S., Comastri, A., et al. 2016, ApJ, 819, 62
Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168
Comastri, A., Ranalli, P., Iwasawa, K., et al. 2011, A&A, 526, L9
Comparat, J., Merloni, A., Salvato, M., et al. 2019, MNRAS, 487, 2005
Corral, A., Georgantopoulos, I., Akylas, A., & Ranalli, P. 2019, A&A, 629, A133
Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937
Cucchetti, E., Pointecouteau, E., Peille, P., et al. 2018, A&A, 620, A173
D'Amato, Q., Gilli, R., Vignali, C., et al. 2020, A&A, 636, A37
Dauser, T., Falkner, S., Lorenz, M., et al. 2019, A&A, 630, A66
Despali, G., Giocoli, C., Angulo, R. E., et al. 2016, MNRAS, 456, 2486
Diemer, B. 2018, ApJS, 239, 35
Donley, J. L., Rieke, G. H., Pérez-González, P. G., & Barro, G. 2008, ApJ, 687, 111
Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, ApJ, 748, 142
Dubois, Y., Pichon, C., Welker, C., et al. 2014, MNRAS, 444, 1453
Duras, F., Bongiorno, A., Ricci, F., et al. 2020, A&A, 636, A73
Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, AJ, 132, 117
Finoguenov, A., Tanaka, M., Cooper, M., et al. 2015, A&A, 576, A130
Gehrels, N. 1986, ApJ, 303, 336
Georgakakis, A., Carrera, F., Lanzuisi, G., et al. 2013, ArXiv e-prints [arXiv:1306.2328]
Georgantopoulos, I., Comastri, A., Vignali, C., et al. 2013, A&A, 555, A43
Gilli, R., Norman, C., Vignali, C., et al. 2014, A&A, 562, A67
Gilli, R., Comastri, A., & Hasinger, G. 2007, A&A, 463, 79
Gilli, R., Su, J., Norman, C., et al. 2011, ApJ, 730, L28
Goulding, A. D., Forman, W. R., Hickox, R. C., et al. 2012, ApJS, 202, 6
Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, MNRAS, 413, 101
Hasinger, G., Miyaji, T., & Schmidt, M. 2005, A&A, 441, 417
Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, MNRAS, 451, 2663
Hodges-Kluck, E. J., Gallo, E., Seth, A., Greene, J., & Baldassare, V. 2020, ApJ, 898, 106
Inayoshi, K., Ichikawa, K., & Ho, L. C. 2020, ApJ, 894, 141
Käfer, F., Finoguenov, A., Eckert, D., et al. 2020, A&A, 634, A8
Khandai, N., Di Matteo, T., Croft, R., et al. 2015, MNRAS, 450, 1349
Kocevski, D. D., Hasinger, G., Brightman, M., et al. 2018, ApJS, 236, 48
Lacey, C. G., Baugh, C. M., Frenk, C. S., et al. 2016, MNRAS, 462, 3854
Lagos, C. D. P., Tobar, R. J., Robotham, A. S. G., et al. 2018, MNRAS, 481, 3573
LaMassa, S. M., Urry, C. M., Glikman, E., et al. 2013a, MNRAS, 432, 1351
LaMassa, S. M., Urry, C. M., Cappelluti, N., et al. 2013b, MNRAS, 436, 3581
LaMassa, S. M., Urry, C. M., Cappelluti, N., et al. 2016, ApJ, 817, 172
Lanzuisi, G., Ranalli, P., Georgantopoulos, I., et al. 2015, A&A, 573, A137
Lanzuisi, G., Civano, F., Marchesi, S., et al. 2018, MNRAS, 480, 2578
Leauthaud, A., Finoguenov, A., Kneib, J.-P., et al. 2010, ApJ, 709, 97
Lehmer, B. D., Alexander, D. M., Chapman, S. C., et al. 2009, MNRAS, 400, 299
Lehmer, B. D., Basu-Zych, A. R., Mineo, S., et al. 2016, ApJ, 825, 7
Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, ApJS, 228, 2
Marchesi, S., Civano, F., Elvis, M., et al. 2016a, ApJ, 817, 34
Marchesi, S., Civano, F., Salvato, M., et al. 2016b, ApJ, 827, 150
Marchesi, S., Lanzuisi, G., Civano, F., et al. 2016c, ApJ, 830, 100
Marchesi, S., Ajello, M., Marcotulli, L., et al. 2018, ApJ, 854, 49
Marchesi, S., Ajello, M., Zhao, X., et al. 2019, ApJ, 872, 8
Maughan, B. J., Jones, C., Forman, W., & Van Speybroeck, L. 2008, ApJS, 174, 117
McAlpine, S., Helly, J. C., Schaller, M., et al. 2016, Astron. Comput., 15, 72
Meidinger, N., Barbera, M., Emberger, V., et al. 2017, in Proc. SPIE, SPIE Conf. Ser., 10397, 103970V
Merloni, A., Predehl, P., Becker, W., et al. 2012, ArXiv e-prints [arXiv:1209.3114]
Miyaji, T., Hasinger, G., Salvato, M., et al. 2015, ApJ, 804, 104
Murphy, K. D., & Yaqoob, T. 2009, MNRAS, 397, 1549
Murray, S. S., Kenter, A., Forman, W. R., et al. 2005, ApJS, 161, 1
Mushotzky, R. F., Aird, J., Barger, A. J., et al. 2019, BAAS, 51, 107
Mutch, S. J., Geil, P. M., Poole, G. B., et al. 2016, MNRAS, 462, 250
Nandra, K., Barret, D., & Barcons, X. 2013, ArXiv e-prints [arXiv:1306.2307]
Nandra, K., Laird, E. S., Aird, J. A., et al. 2015, ApJS, 220, 10
Nandra, K., O'Neill, P. M., George, I. M., & Reeves, J. N. 2007, MNRAS, 382, 194
Nanni, R., Vignali, C., Gilli, R., & Moretti, A. 2017, A&A, 603, A128
Nanni, R., Gilli, R., Vignali, C., et al. 2020, A&A, 637, A52
Ni, Y., Di Matteo, T., Gilli, R., et al. 2020, MNRAS, 495, 2135
Pierre, M., Pacaud, F., Adami, C., et al. 2016, A&A, 592, A1
Planck Collaboration XVI. 2014, A&A, 571, A16
Pratt, G. W., Arnaud, M., Biviano, A., et al. 2019, Space Sci. Rev., 215, 25
Qin, Y., Mutch, S. J., Poole, G. B., et al. 2017, MNRAS, 472, 2009
Ranalli, P., Comastri, A., & Setti, G. 2005, A&A, 440, 23
Reichert, A., Böhringer, H., Fassbender, R., & Mühlegger, M. 2011, A&A, 535, A4
Ricci, C., Ueda, Y., Koss, M. J., et al. 2015, ApJ, 815, L13
Rosati, P., Borgani, S., & Norman, C. 2002, ARA&A, 40, 539
Sazonov, S., & Khabibullin, I. 2017, MNRAS, 468, 2249
Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521
Schmidt, M., Schneider, D. P., & Gunn, J. E. 1995, AJ, 110, 68
Stern, D., Assef, R. J., Benford, D. J., et al. 2012, ApJ, 753, 30
The Lynx Team 2018, ArXiv e-prints [arXiv:1809.09642]
Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709
Treister, E., Schawinski, K., Volonteri, M., & Natarajan, P. 2013, ApJ, 778, 130
Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, ApJ, 786, 104
Vito, F., Gilli, R., Vignali, C., et al. 2014, MNRAS, 445, 3557
Vito, F., Brandt, W. N., Yang, G., et al. 2018, MNRAS, 473, 2378
Vito, F., Brandt, W. N., Bauer, F. E., et al. 2019, A&A, 630, A118
Vogelsberger, M., Genel, S., Springel, V., et al. 2014, MNRAS, 444, 1518
Wang, F., Fan, X., Yang, J., et al. 2017, ApJ, 839, 27
Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, ApJS, 195, 10
Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2016, ApJS, 224, 15