

# AGN and QSOs in the eROSITA All-Sky Survey

## Part II: Studies of large-scale structure

Alexander Kolodzig<sup>1</sup>, Marat Gilfanov<sup>1,2</sup>, Gert Hütsi<sup>1,3</sup>, Rashid Sunyaev<sup>1,2</sup>

<sup>1</sup> Max-Planck-Institut für Astrophysik (MPA), Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany, e-mail: alex@mpa-garching.mpg.de

<sup>2</sup> Space Research Institute (IKI), Russian Academy of Sciences, Profsoyuznaya ul. 84/32, Moscow, 117997 Russia

<sup>3</sup> Tartu Observatory, Tõravere 61602, Estonia

Received ? 2013; accepted ? 2013

### ABSTRACT

The four year X-ray all-sky survey (eRASS) of eROSITA telescope aboard the Spektrum-Roentgen-Gamma satellite will detect  $\sim 3$  million AGN with a median redshift of  $z \approx 1$  and typical luminosity of  $L_{0.5-2.0\text{keV}} \sim 10^{44} \text{ erg s}^{-1}$ . We show that this unprecedented AGN sample, complemented with redshift information, will supply us with outstanding opportunities for large-scale structure research. For the first time, detailed redshift and luminosity resolved studies of the bias factor for X-ray selected AGN will become possible. The eRASS AGN sample will not only improve the redshift and luminosity resolution of these studies but will also expand their luminosity range beyond  $L_{0.5-2.0\text{keV}} \sim 10^{44} \text{ erg s}^{-1}$ , thus making possible direct comparison of clustering properties of luminous X-ray AGN and optical quasars. These studies will dramatically improve our understanding of AGN environment, triggering mechanisms, growth of super-massive black holes and their co-evolution with dark matter halos.

The eROSITA AGN sample will become a powerful cosmological probe. It will make possible detection of baryonic acoustic oscillations (BAOs) for the first time with X-ray selected AGN. With the data from entire extragalactic sky, BAO will be detected at a  $\geq 10\sigma$  confidence level in the full redshift range and with  $\sim 8\sigma$  confidence in the  $0.8 < z < 2.0$  range, currently uncovered by any existing BAO surveys. In order to exploit the full potential of the eRASS AGN sample, photometric and spectroscopic surveys of large area and a sufficient depth will be needed.

**Key words.** Surveys – X-rays: general – Quasars: general – Galaxies: active – Cosmology: Large-Scale Structure of Universe

## 1. Introduction

Large-scale structure (LSS) studies are established as an important tool for studies in two major areas of astrophysics: cosmology and galaxy evolution. A key of their success is the increasing number of surveys at different wavelengths with increasing depths and sky coverages. In X-rays, many deep, extragalactic surveys have been performed in the last decade (Brandt & Hasinger 2005). However, in comparison to other wavelengths, X-ray surveys with a large sky coverage and with sufficient depth are still rare. The previous X-ray all-sky survey was performed by ROSAT<sup>1</sup> (Truemper 1993; Voges et al. 1999) more than two decades ago. Its successor with a  $\sim 30$  times better sensitivity will be the four year long all-sky survey (eRASS) of the eROSITA<sup>2</sup> telescope (Predehl et al. 2010), to be launched aboard the Russian Spektrum-Roentgen-Gamma satellite<sup>3</sup> in 2014.

The major science goals of the eROSITA mission are to study cosmology with clusters of galaxies and active galactic nuclei (AGN) and to constrain the nature of dark matter (DM) and dark energy. For a comprehensive description of the eROSITA mission see the science book of eROSITA (Merloni et al. 2012).

In this work we explore the potential of studying LSS with the AGN sample to be detected in eRASS. We focus on two important aspects of LSS studies: the clustering strength (repre-

sented by the linear bias factor, Sect. 3) and the baryonic acoustic oscillations (BAOs, Sect. 4). For measuring the former quantity redshift accuracy of photometric surveys is sufficient, therefore bias studies can be successfully conducted during and soon after the time eRASS is finished. The BAO measurements, on the other hand, will be much more difficult to accomplish as the spectroscopic redshift accuracy over large sky areas will be required. Note that sufficient redshift accuracy can be also provided by high quality narrow-band multi-filter photometric surveys.

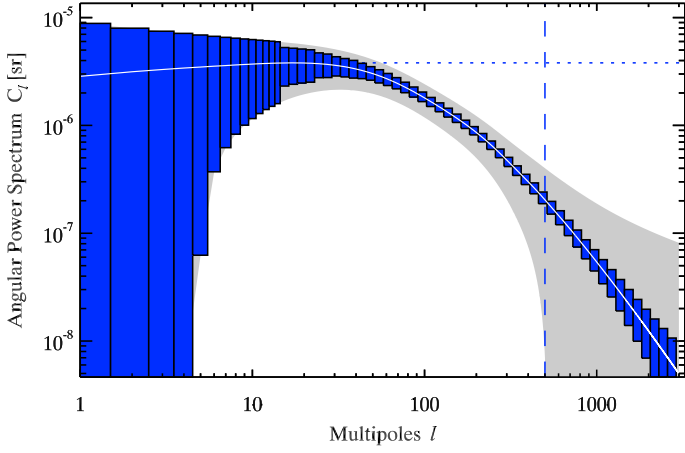
In our previous work (Kolodzig et al. 2012), we have studied the statistical properties of the AGN sample of eRASS and will adopt these results here. In the current work, we focus on the AGN detected in the soft energy band (0.5 – 2.0 keV) and on the extragalactic sky ( $|b| > 10^\circ$ ,  $\sim 34 \text{ } 100 \text{ deg}^2$ ). In the following calculations we will assume the four year average sensitivity of  $1.1 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$  from Kolodzig et al. (2012, Table 1).

Optical followup surveys will be needed to provide identification and redshift information to a desired accuracy for all eRASS AGN. Current optical surveys are not sufficient in size and/or depth. A sensitivity of  $I \sim 22.5 \text{ mag}$  ( $R \sim 23.0 \text{ mag}$ ) is required to detect at least 95 % of the eRASS AGN (Kolodzig et al. 2012). Many photometric and spectroscopic surveys with different parameters have been proposed or are being already in construction (e.g. Merloni et al. (2012)). For the purpose of our investigation we will assume that redshifts are available for all eRASS AGN. We will explore the effects of redshift-errors in the forthcoming paper (Hütsi et al. 2013, in prep.).

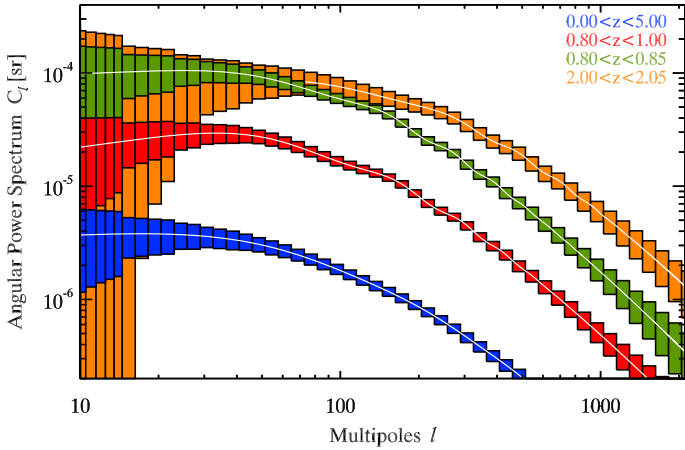
<sup>1</sup> <http://www2011.mpe.mpg.de/xray/wave/rosat/>

<sup>2</sup> <http://www.mpe.mpg.de/eROSITA>

<sup>3</sup> <http://hea.iki.rssi.ru/SRG>



**Fig. 1.** The angular power spectrum of the full eRASS AGN sample (soft band, four years) for the extragalactic sky ( $f_{\text{sky}} \approx 0.83$ ) and  $0 < z < 5$ . The gray shaded area and the blue histogram show the  $1\sigma$  uncertainty region (Eq. 4) without and with  $\ell$ -binning, respectively. The horizontal dotted line shows the level of shot noise, which was already subtracted from the angular power spectrum. For multipoles above the vertical dashed line (representing  $l_{\text{max}} \approx 500$ ) our assumption of a linear clustering starts to break down. Therefore, we do not consider these multipoles in our further calculations.



**Fig. 2.** Same as Fig. 1 but additionally with the angular power spectra for various narrow and broad redshift ranges (see Sect. 2.2).

We assume for this work a flat  $\Lambda$ CDM cosmology with the following parameters:  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  ( $h = 0.70$ ),  $\Omega_m = 0.30$  ( $\Omega_\Lambda = 0.70$ ),  $\Omega_b = 0.05$ ,  $\sigma_8 = 0.8$ . Luminosities are given for the soft energy band (0.5 – 2.0 keV) and we use the decimal logarithm throughout the paper.

## 2. Angular power spectrum

The common tool to study LSS is to measure and analyze the clustering of objects (such as AGN) with the 2-point correlation function (2pCF) or the power spectrum (PS) (e.g. Peebles 1980). The two methods, 2pCF and PS, have their benefits and disadvantages (e.g. Wall & Jenkins 2012) but should contain in the end the same information about the LSS. In this work we will use the angular power spectrum  $C_\ell$  to characterize clustering properties of objects. In order to predict the power spectra which will be measured with eRASS AGN, we will rely on the

model for AGN clustering of Hütsi et al. (2012), in particular, we will use their model II. The details of our calculations are summarized below.

We calculate the angular power spectrum as follows:

$$C_\ell = \frac{2}{\pi} \int P(k) [W_\ell(k)]^2 k^2 dk \quad (1)$$

where the projection kernel is

$$W_\ell(k) = \int_{z_{\text{min}}}^{z_{\text{max}}} f(z) g(z) b(M_{\text{eff}}, z) j_\ell(k r(z)) dz \quad (2)$$

Hereby,  $P(k)$  is the  $z = 0$  3D linear power spectrum, for which we will use fitting formulae of Eisenstein & Hu (1998),  $f(z)$  is the normalized radial selection function,  $g(z)$  is the linear growth function (e.g. Dodelson 2003),  $b(M_{\text{eff}}, z)$  is the AGN linear clustering bias factor and  $j_\ell$  are the spherical Bessel functions of order  $\ell$ , where  $r(z)$  is the co-moving distance to redshift  $z$  (e.g. Hogg 1999).

The radial selection function is defined as the (normalized) differential redshift distribution of AGN, which we calculate with the AGN X-ray luminosity function (XLF),  $\phi(\log L, z)$ , of Hasinger et al. (2005)<sup>4</sup>. It is the only quantity that contains the information about eRASS, since it depends on the survey sensitivity ( $S$ ) as following:

$$\frac{dN}{dz}(S, z) = \frac{dV(z)}{dz} \int_{\log L_{\text{min}}(S, z)}^{\log L_{\text{max}}} \phi(\log L, z) d \log L \quad (3)$$

Here,  $\frac{dV(z)}{dz} [\text{Mpc}^3 \text{ sr}^{-1}]$  is the co-moving volume element and  $L_{\text{min}}(S, z) = 4\pi S d_L^2(z)$ , where  $d_L(z)$  is the luminosity distance (e.g. Hogg 1999).

The AGN linear clustering bias factor,  $b(M_{\text{eff}}, z)$ , is computed with the analytical model of Sheth et al. (2001) by assuming an effective mass  $M_{\text{eff}}$  of the DM halo (DMH) where the AGN reside. Based on recent observations (e.g. Alleinato et al. 2011; Krumpel et al. 2012; Mountrichas et al. 2013), we assume an effective mass of  $M_{\text{eff}} = 2 \times 10^{13} h^{-1} M_\odot$ .

In our work, we only focus on the linear clustering regime. Therefore we restrict our calculations to the spatial co-moving scales larger than  $k_{\text{max}} \approx 0.2 h \text{ Mpc}^{-1}$ , corresponding to wavelengths larger than  $\sim 30 h^{-1} \text{ Mpc}$ . The associated multipole number is  $\ell_{\text{max}}(\bar{z}) = k_{\text{max}} r(\bar{z})$  and depends on the median redshift  $\bar{z}$  of the considered redshift bin. At the median redshift of eRASS AGN sample this is  $\ell_{\text{max}}(\bar{z} \sim 1) \approx 500$ . Thus, for our calculations we do not consider  $C_\ell$  at multipoles higher than  $\ell_{\text{max}}$ .

For simplicity, we do not take linear redshift space distortions (RSD) (Kaiser 1987) into account. Since the signal-to-noise ratio (S/N) in our angular power spectra is rather poor at small multipoles (see Figs. 1 and 2), where linear RSD become most significant, we do not expect that our results would change significantly if we would consider them in our calculations.

### 2.1. Uncertainties

The variance of the  $C_\ell$  can be well approximated with:

$$\sigma_{C_\ell}^2 = \frac{2}{(2\ell + 1) f_{\text{sky}}} \left( C_\ell + \frac{1}{N} \right)^2 \quad (4)$$

<sup>4</sup> See Kolodzig et al. (2012) for details.

assuming Gaussian statistics of the matter fluctuations ( $\ell \lesssim \ell_{\max}$ ). Hereby,  $f_{\text{sky}}$  is the sky fraction, which takes into account the effective loss of modes due to partial sky coverage, and  $N$  is the AGN surface number density [ $\text{sr}^{-1}$ ], which is computed with the AGN XLF and the survey sensitivity of eRASS (See Kolodzig et al. 2012). The first term ( $C_\ell$ ) in the brackets represents the cosmic variance and becomes important at large scales (small  $\ell$ ). The second term, the shot noise ( $N^{-1}$ ), takes into account that we are using a discrete tracer (AGN) and becomes dominant at small scales (large  $\ell$ ), where  $N^{-1} \gg C_\ell$  (see e.g. Fig. 1). In order to minimize the uncertainty in  $C_\ell$  both high sky coverage and large number density of objects are needed.

## 2.2. Results

In Fig. 1, we show the expected angular power spectrum of the full eRASS AGN sample after four years for the entire extragalactic sky. By introducing redshift information (available from other surveys), the angular power spectrum becomes a relevant tool for LSS studies. In Fig. 2, we can see how its amplitude increases with a decreasing size of the redshift bin and also oscillations (see Sect. 4) in the angular power spectrum become more prominent. We can see from the two angular power spectra of  $0.80 < z < 0.85$  and  $2.00 < z < 2.05$  (with same redshift bin size) in Fig. 2, that the turnover of the spectrum and the positions of the oscillations depend on the redshift. The amplitudes are also different because the clustering strength increases with redshift (see Sect. 3). Since the redshift distribution of eRASS peaks around  $z \sim 0.8$ , the number density around  $z \sim 0.8$  is much higher than at  $z \sim 2.0$  and therefore the uncertainty of the angular power spectrum is smaller for  $0.80 < z < 0.85$  than for  $2.00 < z < 2.05$ .

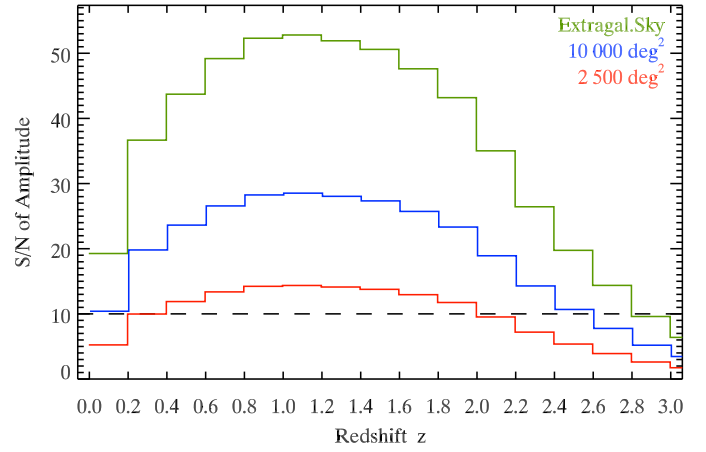
## 3. Linear bias factor

The linear bias factor  $b$  is an important parameter for clustering analysis of AGN. It connects the underlying DM distribution with the AGN population. Observationally, it has been so far a very challenging task to measure this connection with high accuracy, due to low statistics (e.g. Krumpel et al. 2010, 2012; Miyaji et al. 2011; Allevato et al. 2011, 2012; Koutoulidis et al. 2013; Mountrichas et al. 2013). With detailed knowledge of the behavior of the linear bias factor with the redshift and luminosity we will be able to improve our understanding of major questions, such as what is the environment of AGN, what are the major triggering mechanisms of AGN activity and how super massive black holes (SMBH) co-evolve with the DMH over cosmic time.

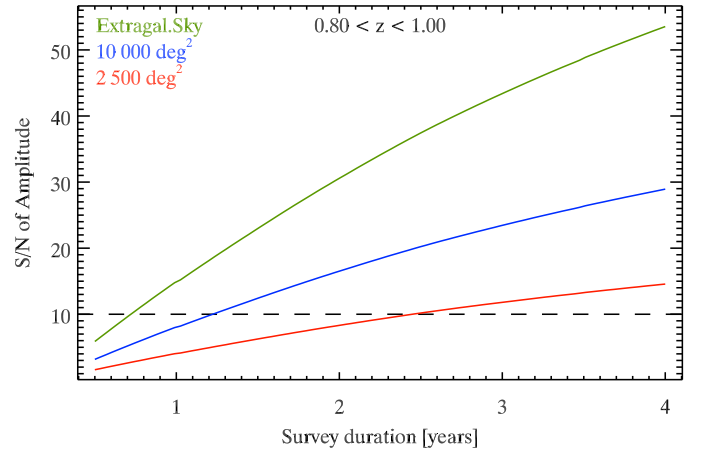
### 3.1. Method

The linear bias factor is measured by comparing the amplitudes of the observed PS of tracer objects and of the theoretical PS of the DM, under the assumptions of a certain cosmology. Since the PS amplitude of tracer objects is proportional to the square of the linear bias factor ( $\propto b^2$ ), its uncertainty directly reflects the uncertainty of measuring the latter. Knowing the amplitude ( $A$ ) of our angular power spectrum we are able to estimate this uncertainty. The S/N for measuring the normalization of the power spectrum  $C_\ell$  assuming that its shape is known, is given by:

$$\frac{S}{N} = \frac{A}{\delta A} = \sqrt{\sum_{\ell=1}^{\ell_{\max}(z)} \left( \frac{C_\ell}{\delta C_\ell} \right)^2} \quad (5)$$



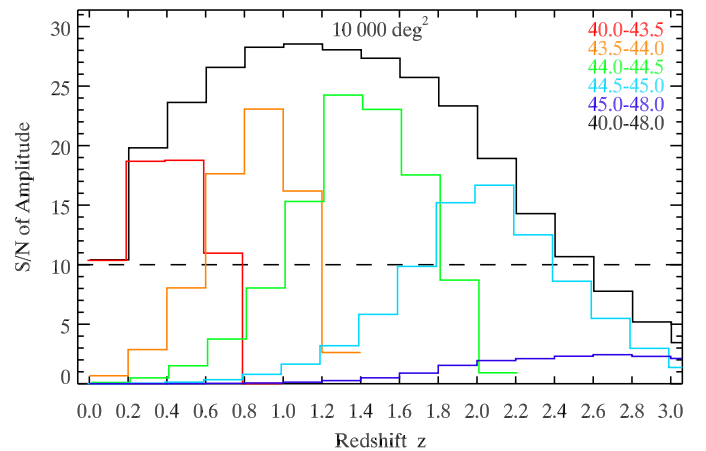
**Fig. 3.** Signal-to-noise ratio of the amplitude of the angular power spectrum (Eq. 5) as a function of the redshift for different sky fractions. A  $\Delta z = 0.2$  binning is assumed.



**Fig. 4.** Signal-to-noise ratio of the amplitude of the angular power spectrum as a function of the survey duration for the redshift bin  $0.8 < z < 1.0$  at different sky fractions.

Here, we are assuming that all multipoles are independent.

We are using a redshift binning of  $\Delta z = 0.20$  for our calculation (see Fig. 3 and 5) but other bin sizes would also be applicable to demonstrate our results. In current observations



**Fig. 5.** Same as Fig. 3 but for different luminosity ranges (in units of  $\log(L[\text{erg s}^{-1}])$ ) and the sky coverage of  $10\,000 \text{ deg}^2$ .

(e.g. Allevato et al. 2011; Krumpke et al. 2012; Koutoulidis et al. 2013) a typically much larger bin size is used in order to achieve a reasonable S/N ratio for  $b$  in each redshift bin.

### 3.2. Results

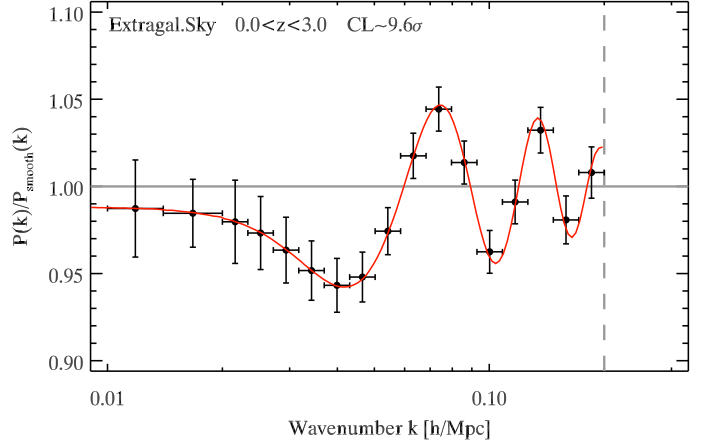
In Fig. 3 the achievable S/N of the power spectrum amplitude is shown as a function of redshift. The shape of the curves is dominated by the redshift distribution of AGN modified by the quadratic-like increase of the linear bias factor with redshift at constant DMH mass (e.g. Sheth et al. 2001). We can see that we are able to measure the amplitude to a high accuracy ( $< 10\%$ ) for a wide redshift range even with a rather small fraction of the sky (e.g.  $\sim 2500 \text{ deg}^2$ ).

The analysis of the linear bias factor can be performed before the entire four year long eRASS is completed, as we can see from Fig. 4. For a SDSS-like sky coverage of  $10000 \text{ deg}^2$  (blue curve) one can already work with the data of eRASS after only 1.5 years (three full sky scans) to study the evolution of the linear bias factor to an accuracy of better than  $10\%$  in the amplitude for the redshift bin  $0.8 < z < 1.0$ . For a sky region of  $2500 \text{ deg}^2$  it needs five full sky scans (2.5 years). For the neighboring redshift bins  $0.6 < z < 0.8$  and  $1.0 < z < 1.2$  the results are similar. The sensitivities used for this calculation are taken from Fig. 1 of Kolodzig et al. (2012).

Owing to the high S/N of the power spectrum amplitude, we will be able to separate the AGN in different luminosity groups. This is demonstrated in Fig. 5 for a sky coverage of  $10000 \text{ deg}^2$ . We can see that we will be able to achieve an accuracy of  $< 10\%$  for most luminosity groups for a certain redshift range. This means that it will be possible to perform redshift and luminosity resolved analysis of the linear bias factor with eRASS with high statistical accuracy. We note that in our calculation the difference in the S/N of the luminosity groups in Fig. 5 is driven only by the difference in the redshift distribution. Although there are some observational evidence for a correlation between the DMH mass and the luminosity of AGN/quasars (e.g. Krumpke et al. 2012; Richardson et al. 2012; Shen et al. 2012; Koutoulidis et al. 2013), it is still rather tentative. We therefore assumed that the source luminosity is independent on the DMH mass.

## 4. Baryonic acoustic oscillations

Acoustic peaks in the power spectra of matter and CMB radiation are one of the major probes to measure the kinematics of the Universe (e.g. Weinberg et al. 2012). They were predicted theoretically over four decades ago (Sunyaev & Zeldovich 1970; Peebles & Yu 1970) and now have become a standard tool of observational cosmology. Unlike acoustic peaks in the angular power spectrum of CMB, their amplitude in the matter power spectrum in the  $\Lambda$ CDM Universe is small. For this reason, only recently galaxy surveys have reached sufficient breadth and depth for the first convincing detection of BAO, achieved with the SDSS data (Cole et al. 2005; Eisenstein et al. 2005; Hütsi 2006; Tegmark et al. 2006). Since then BAO have been measured extensively up to the redshift  $z \sim 0.8$ , in particular with luminous red giant galaxies (LRGs) (e.g. Anderson et al. 2012). Above this redshift limit, BAO features were only found in the correlation function of the transmitted flux fraction in the Lyman- $\alpha$  forest of high-redshift quasars (Busca et al. 2013; Slosar et al. 2013), but have not yet been directly detected in distribution of galaxies.



**Fig. 6.** Baryonic acoustic oscillations in the power spectrum for the extragalactic sky in redshift range of  $0.0 < z < 3.0$ . At wavenumbers above the vertical dashed line (corresponding to  $0.2 h \text{ Mpc}^{-1}$ ) our assumption of a linear clustering starts to break down. The red curve shows the original input model for the BAOs.

For the so far uncharted redshift range from  $z \sim 0.8$  up to  $\sim 2.0$ , AGN, quasars and emission-line galaxies (ELGs) are proposed to be the best tracers to measure BAOs, however, currently existing surveys do not to achieve the required statistics for a proper detection (Sawangwit et al. 2012; Comparat et al. 2013). eRASS and the proposed SDSS-IV<sup>5</sup> survey program eBOSS<sup>6</sup> (2014-2020) will be the first surveys to change this situation. eRASS will achieve sufficiently high density of objects  $N \sim 40 \text{ deg}^{-2}$  in this redshift range and will have by far the largest sky coverage in comparison to eBOSS and all other proposed BAO surveys. This would enable one to push the redshift limit of BAO detections in the angular power spectra of galaxies far beyond the present limit of  $z \sim 0.8$ .

### 4.1. Method

By construction of the model, BAOs are included in the AGN clustering model of Hütsi et al. (2012), through the 3D linear power spectrum (Sect. 2). Oscillations can be noticed, for example, in Fig. 2, in the power spectra of objects selected in narrow redshift intervals. As the angular scales of acoustic peaks depend on the redshift, in the power spectra computed for broad redshift intervals BAO are smoothed out, due to superposition of signals coming from many different redshift slices.

Although analysis of the real data will be conducted in a much more elaborate way, for the purpose of this calculation we will use a simple method to estimate the amplitude and statistical significance of the BAO signal detection. We divide a broad redshift interval into narrow slices of the width  $\Delta z$  and for each slice compute the angular power spectrum,  $C_\ell(z)$ , and convert multipole number to wavenumber  $k = \ell/r(z)$  to obtain  $P(k, z)$ . These power spectra are co-added in the wavenumber space to obtain the total power spectrum  $P(k)$  of objects in the broad redshift interval. The so constructed power spectrum will have unsmoothed BAO features. To estimate their statistical significance, we also construct a model  $C_{\ell, \text{smooth}}(z)$  without acoustic peaks, by smoothing the matter transfer function, similar to how it was done in Eisenstein & Hu (1998). From this model we com-

<sup>5</sup> <http://www.sdss3.org/future/>

<sup>6</sup> [http://lamwvs.oamp.fr/cosmowiki/Project\\_eBoss](http://lamwvs.oamp.fr/cosmowiki/Project_eBoss)

pute the smoothed power spectrum in the wavenumber space,  $P_{\text{smooth}}(k)$ , not containing BAOs. To illustrate the amplitude of the BAO signal, one can plot the difference  $P(k) - P_{\text{smooth}}(k)$  or the ratio  $P(k)/P_{\text{smooth}}(k)$ .

By the analogy with Eq. (5), the S/N of the BAO detection in the eRASS sample can be computed as:

$$\frac{S}{N} = \sqrt{\sum_z \sum_{\ell}^{\ell_{\text{max}}(z)} \left( \frac{C_{\ell}(z) - C_{\ell,\text{smooth}}(z)}{\sigma_{C_{\ell}}} \right)^2} \quad (6)$$

where the outer summation is performed over the redshift slices and the variance  $\sigma_{C_{\ell}}^2$  is calculated from Eq. 4.

The result of this calculation depends on the choice of the thickness of the redshift slice  $\Delta z$ . For too large values of the  $\Delta z$  BAO will be smeared out, as discussed above (cf. Fig. 2). On the other hand, for too small values of the  $\Delta z$ , at which the thickness of the redshift slice becomes somewhat smaller than the comoving linear scale of acoustic oscillations, the cross-spectra between different redshift slices will need to be taken into account in computing the  $P(k)$ . For the purpose of these calculations we chose  $\Delta z = 0.05$ . The corresponding thickness of the redshift slice at  $z \sim 1$  is approximately equal to the comoving linear scale of the first BAO peak. Note that the omission of the cross-spectra in our calculation leads to some underestimation of the confidence level of BAO detection.

#### 4.2. Results

In Fig. 6 and 7 we show the ratio  $P(k)/P_{\text{smooth}}(k)$  along with its uncertainties computed as described above. As these plots demonstrate, with the whole eRASS AGN sample for the extragalactic sky we should be able to detect the BAOs with a confidence level (CL) of  $\sim 10\sigma$  (Fig. 6). For the currently unexplored redshift range of  $0.8 - 2.0$  the confidence level of  $\sim 8\sigma$  will be achieved, which can be seen in the top panel of Fig. 7. Decreasing the sky area to  $20\,000 \text{ deg}^2$  or  $10\,000 \text{ deg}^2$  (see middle panel of Fig. 7) we obtain  $\sim 6\sigma$  and  $\sim 4\sigma$ , respectively. In Fig. 8 we show for different redshift ranges how the confidence level of the BAO detection depends on the sky coverage. The curves follow a  $f_{\text{sky}}^{-0.5}$ -dependence, as it is expected from Eq. (4).

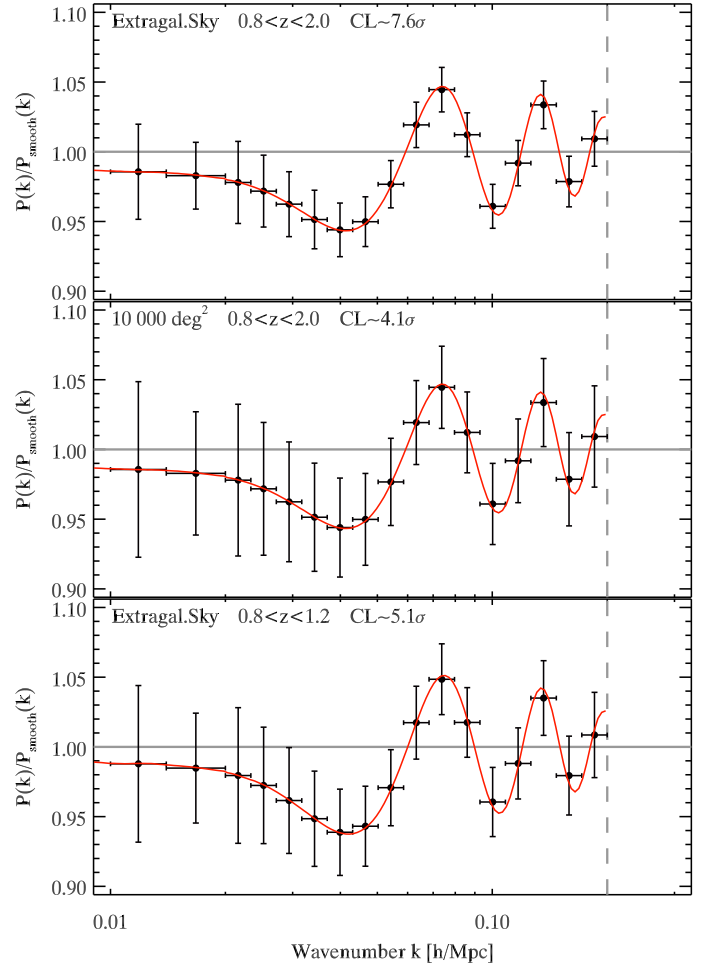
As Fig. 8 shows, for the redshift ranges  $0.0 - 0.8$ ,  $0.8 - 1.2$  and  $1.2 - 2.0$  the achievable confidence levels are rather similar, therefore the power spectra ratio shown in the bottom panel of Fig. 7 is representative for all three redshift intervals. Comparing the upper and bottom panels of Fig. 7, one can see how the BAO signal depends on the redshift range, whereas the top and middle panels show the degradation due to reduced sky coverage.

## 5. Discussion and conclusions

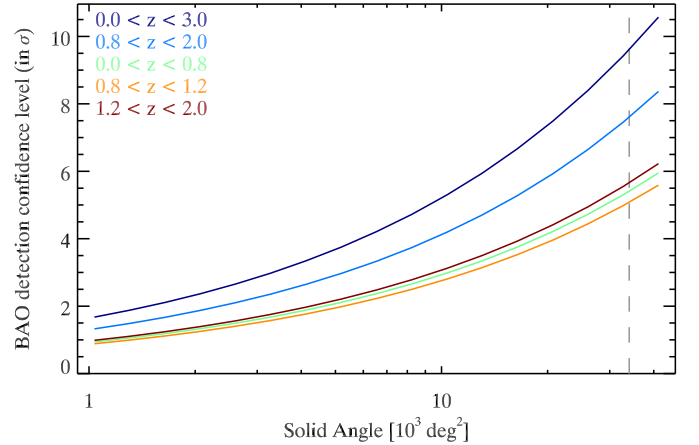
### 5.1. Linear bias factor

Measurement of the linear bias factor provides a simple and direct method to estimate the average mass of DMHs hosting a given sub-population of AGN. With the eRASS, these measurements will become possible to unprecedented detail. The dramatic improvement of the redshift and luminosity resolution of DMH mass measurements will have a great impact on our understanding of the environment of AGN, AGN triggering mechanisms and SMBH co-evolution with the DMH.

An observational evidence of a correlation between DMH mass and the luminosity of AGN has already been found but uncertainties are still large and AGN luminosities available for



**Fig. 7.** Same as Fig. 6 but for different redshift ranges and sky coverages.



**Fig. 8.** The confidence level of a BAO detection as a function of sky coverage for different redshift ranges (see Sect. 4.1 for further explanations). The vertical gray dashed line shows the area of the extragalactic sky.

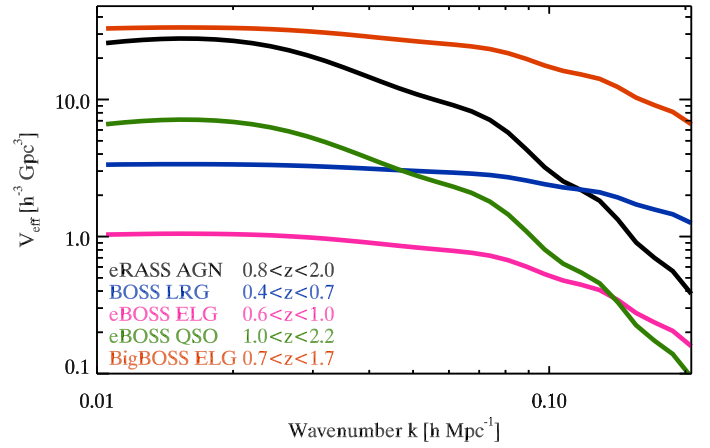
these studies are typically limited by  $L \sim 10^{44} \text{ erg s}^{-1}$  (e.g. Allevalo et al. 2011; Krumpel et al. 2012; Mountrichas et al. 2013; Koutoulidis et al. 2013). Koutoulidis et al. (2013) compared their results from clustering studies of AGN in four extragalactic X-ray surveys of different depth and coverage (CDFN, CDFS, COSMOS and AEGIS) with theoretical predictions of

Fanidakis et al. (2012). Their goal was to determine, from the AGN bias factor measurements, the dominant SMBH growth mode for AGN of different luminosities – either through galaxy mergers and/or disk instabilities or through accretion of hot gas from the halo of the galaxy. However, uncertainties of bias factor measurements and of DMH mass estimation were too large to clearly disentangle the dominant growth mode as a function of luminosity. In particular, the luminosity range of objects available for their analysis,  $L \sim 10^{42-44} \text{ erg s}^{-1}$ , was too narrow to challenge the prediction of Fanidakis et al. (2012) that luminous galaxies with  $L > 10^{44} \text{ erg s}^{-1}$  reside in DMHs of moderate mass of  $\sim 10^{12} M_{\odot}$ . For the same reason, no direct comparison was possible with the results of the optical quasar surveys (Alexander & Hickox 2012). To overcome this limitation, Allevato et al. (2011) studied broad line (BL) AGN from the COSMOS survey and found for their luminosity bin  $L \approx 10^{43-46} \text{ erg s}^{-1}$  a significantly higher DMH mass than inferred from quasar studies, suggesting that for broad line AGN major merger may not be the dominant triggering mechanism, in reasonable agreement with recent simulations (Draper & Ballantyne 2012; Hirschmann et al. 2012). However, the large width of the luminosity bin required to accumulate sufficient statistics did not allow them to draw a firm conclusion.

As illustrated by Fig. 5, the eRASS AGN sample will not only dramatically improve statistics but will also expand the luminosity range beyond  $L \sim 10^{44} \text{ erg s}^{-1}$ , to the luminosity domain characteristic of quasars. Thus, the eRASS data will not only increase the redshift and luminosity resolution of DMH mass estimations of AGN, but will open possibilities for detailed comparison of clustering properties of luminous AGN and optical quasars. Another aspect of bias measurements with eRASS, determining their uniqueness is that they are based on the X-ray selected AGN sample and will cover a very broad SMBH mass range, broader than that in AGN samples produced by optical/IR or radio surveys (e.g. Hickox et al. 2009).

The growth rate of SMBHs over time can be measured from the XLF of AGN (e.g. Aird et al. 2010) and eRASS will improve the accuracy and redshift resolution of these studies tremendously (e.g. Kolodzig et al. 2012). Combined with clustering bias data, these measurements will be placed in a broader context and connected with DMH properties, and will thus provide new insights on the co-evolution of SMBHs with their DMHs and will also help to investigate the dominant triggering mechanisms of AGN activity.

The AGN clustering model used in this paper and, correspondingly, calculations of the AGN linear bias factor ignored the internal structure of DMHs, i.e. they were restricted to the scales larger than the size of a typical DMH. Expressed in the language of the halo occupation distribution (HOD) formalism, these calculations operated with the population averaged halo occupation numbers. The angular resolution of the eRosita telescope,  $\sim 30''$  FOV averaged HEW, is sufficient to resolve sub-halo linear scales. Clustering measurements on small scales will permit to obtain a detailed picture of how AGN are distributed within a DMH (e.g. to measure fractions of central and satellite AGN) and how the HOD depend on the DMH mass and redshift, and AGN luminosity. Extrapolating results of XMM-COSMOS data analysis by Richardson et al. (2013) we may expect that high accuracy determination of the HOD parameters will be easily achieved with eRASS data, which will be able to address all these questions, advancing our understanding of AGN clustering on small scales and their HOD.



**Fig. 9.** The effective volumes of BAO surveys listed in Table 1 as a function of the wavenumber. Effective volumes are computed for redshift ranges indicated in the plot.

## 5.2. BAO

The BAO detection beyond redshift  $\sim 0.8$  will be a very significant milestone for the direct measurement of the kinematics of the Universe. eRASS will be able to map this uncharted redshift region with a sufficiently high AGN number density to measure BAOs with a high statistical significance (see Fig. 8). For a proper forecast of how these measurements would improve our constraints on cosmological parameters, Markov Chain Monte Carlo (MCMC) simulations (e.g. Lewis & Bridle 2002) and/or Fisher matrix calculations (e.g. Tegmark et al. 1997) have to be made, which is beyond the focus of our work. Sawangwit et al. (2012) have performed a MCMC simulation for QSOs and demonstrate that a  $3 - 4\sigma$  BAO detection (of a  $3000 \text{ deg}^2$  QSO survey with  $N = 80 \text{ deg}^{-2}$ ) for  $1.0 < z < 2.2$  can already significantly reduce the uncertainties. Although the survey parameters of eRASS differ (much larger sky coverage but smaller source density for the same redshift region,  $\sim 40 \text{ deg}^{-2}$ ), the results of Sawangwit et al. (2012) can give one an idea of how eRASS AGN sample will improve the accuracy of cosmological parameters determination.

Our calculation are limited to the linear regime, and do not take non-linear structure growth into account, which would somewhat smear out the BAO signal. This will lead to a decrease of their detection significance (e.g. Eisenstein et al. 2007). However, with BAO reconstruction methods one will be able to correct for this effect to some extent (e.g. Padmanabhan et al. 2012; Anderson et al. 2012). We should also note, that our confidence level estimation of the BAO detection are rather conservative as they neglects information contained in the cross-spectra. This will counter balance the negative effect of BAO smearing, as it will be demonstrated in a forthcoming paper (Hütsi et al. 2013, in prep.) where we will study BAO forecasts in a broader context and with accurate account of these effects.

### 5.2.1. Comparison with BAO surveys

We now compare the potential of the eRASS AGN sample with dedicated BAO surveys in the optical band. For the latter, we consider the already completed BOSS CMASS survey (Anderson et al. 2012), the planned eBOSS survey which implementation starts in 2014 and will continue till 2020 and the future BigBOSS survey (Schlegel et al. 2011) anticipated in the

**Table 1.** Parameters of BAO surveys.

Survey	Tracer Object	$b_{\text{tr}}$ ( $z = 0$ )	Redshift Range	$\Omega_{\text{survey}}$ [ $10^3 \text{ deg}^2$ ]	$\mathcal{N}$ [ $\text{deg}^{-2}$ ]	$\langle n \rangle$ [ $10^{-4} \text{ h}^3 \text{ Mpc}^{-3}$ ]	$V_{\text{eff}}$ [ $\text{h}^{-3} \text{ Gpc}^3$ ] ( $k = 0.07 h \text{ Mpc}^{-1}$ )	Implem. Date	Ref.
eRASS	AGN	$\sim 1.33$	0.8–2.0	$\sim 34.1$	$\sim 40$	$\sim 0.12$	$\sim 7.8$	2014–2018	
BOSS	LRG	$\sim 2.00$	0.4–0.7	$\sim 10.0$	$\sim 80$	$\sim 2.3$	$\sim 2.9$	finished	[1]
eBOSS	ELG	$\sim 1.00$	0.6–1.0	$\sim 1.5$	$\sim 180$	$\sim 2.1$	$\sim 0.5$	2014–2020	[2]
eBOSS	QSO	$\sim 1.20$	1.0–2.2	$\sim 7.5$	$\sim 90$	$\sim 0.21$	$\sim 1.7$	2014–2020	[2]
BigBOSS	ELG	$\sim 0.84$	0.7–1.7	$\sim 14.0$	$\sim 1730$	$\sim 6.3$	$\sim 24.0$	> 2020	[3]

**Notes.**  $b_{\text{tr}}$  - bias factor of the tracer object;  $\Omega_{\text{survey}}$  - solid angle covered by the survey  $\mathcal{N}$  - surface number density;  $\langle n \rangle$  - average volume number density;  $V_{\text{eff}}$  - effective survey volume at the first BAO peak  
Ref.: [1] Anderson et al. (2012) and Dawson et al. (2013); [2] eBOSS team, priv. comm.; [3] Schlegel et al. (2011)

2020 timeframe. The Table 1 summarizes key parameters of these surveys relevant for the BAO studies.

A quantity often used to estimate the statistical performance of a galaxy clustering survey is its effective volume (e.g. Eisenstein et al. 2005):

$$V_{\text{eff}}(k) = \Omega_{\text{survey}} \int_{z_{\text{min}}}^{z_{\text{max}}} \left[ \frac{n(z) P_{\text{tr}}(k, z)}{n(z) P_{\text{tr}}(k, z) + 1} \right]^2 \frac{dV(z)}{dz d\Omega} dz \quad (7)$$

where  $\Omega_{\text{survey}}$  is the solid angle covered by the survey,  $P_{\text{tr}}(k, z) = b_{\text{tr}}(z) (g(z)/g(0)) P(k)$  is the power spectrum of objects used as LSS tracer,  $b_{\text{tr}}(z)$  is their redshift-dependent bias factor and  $n(z)$  is their redshift distribution [ $\text{h}^3 \text{ Mpc}^{-3}$ ]. Other quantities are defined in the context of Eqs. (1)–(3). For optical surveys we will assume that  $b_{\text{tr}}(z)g(z) = \text{constant}$ , therefore we only need to compute  $P_{\text{tr}}(k, 0)$ . The  $n(z)$  dependences for optical surveys were taken from references listed in Table 1.

The results of these calculations are plotted in Fig. 9 where we show the effective volumes of different surveys as a function of the wavenumber. Their values at the first BAO peak are listed in the respective column of Table 1. In these calculations integration in Eq. (7) was performed over the optimal redshift range of each survey, as listed in Table 1. For eRASS we used  $z = 0.8 - 2.0$  range in order to emphasize its strength in this uncharted redshift region. The eRASS effective volume for the full redshift range is  $\sim 70\%$  larger for the first BAO peak.

The result of effective volume calculations obviously depends on the assumptions regarding values and redshift dependences of comoving density, bias and the growth factor, which are not always precisely known, especially for the future surveys. Nevertheless, these curves should give a reasonably accurate comparison of how good different surveys are in measuring the power spectrum at different scales (note that the uncertainty of the power spectrum is proportional to  $V_{\text{eff}}^{-0.5}$ ).

We can see from Fig. 9 that the effective volumes of eRASS AGN and eBOSS QSO samples fall most rapidly towards smaller scales in comparison with other surveys. This is a consequence of the lower volume density of X-ray selected AGN and optical QSOs (Table 1). For the same reason, the statistical errors in the eRASS AGN power spectrum are dominated by the shot noise, but the high sky coverage of eRASS keeps them small. As one can see from the figure, eRASS is more competitive at larger scales, upto the second BAO peak, where its sensitivity becomes comparable to BOSS. It should be noted however, that the BOSS AGN sample covers a relatively low redshift domain,  $z \lesssim 0.7$ , whereas all other surveys presented in Fig. 9 are aimed at significantly larger redshifts,  $z \gtrsim 0.7$  (Table 1). Around the first peak the effective volume of eRASS is a factor of  $\sim 2 - 3$  larger

than for BOSS but a few times smaller than BigBOSS (Table 1). On the other hand, eRASS exceeds eBOSS at all wavenumbers. This would still be the case if one considers subset of the eRASS sample covering only  $\sim 1/3$  of the extragalactic sky.

In the conclusion of this section we note that it is remarkable that the statistical strength of eRASS for BAO studies is competitive with that of dedicated BAO surveys, even though eRASS was never designed for this purpose. Potentially, the eRASS AGN sample will become the best sample for BAO studies beyond redshift  $z \gtrsim 0.8$  until arrival of BigBOSS in the end of this decade. However, this potential will not be realized without comprehensive redshift measurements.

### 5.3. Redshift data

We assumed so far that redshifts of all eRASS AGN are known. Now we briefly outline requirements to the redshift data imposed by the science topics discussed above.

The linear bias as well as luminosity function studies do not demand high accuracy of the redshift determination. Indeed, values of the order of  $\delta z \sim 0.1 - 0.2$  should be sufficient, unless analysis with much higher redshift resolution is required. In principle, such accuracy can be provided by photometric surveys. However one would need to investigate the impact of large fraction of catastrophic errors, from which AGN redshift determination based on the standard photometric filter sets are known to suffer (Salvato et al. 2011). Of particular importance are redshift and luminosity trends in catastrophic errors. These problems will be considered in the forthcoming paper (Hütsi et al. 2013, in prep.). Provided that they are properly addressed, optical photometric surveys of a moderate depth of  $I \gtrsim 22.5 \text{ mag}$  (Kolodzig et al. 2012) and with the sky coverage exceeding  $\gtrsim 2500 \text{ deg}^2$  (Fig. 3) would already produce first significant results. An existing survey with such parameters is SDSS. Its depth would allow detection of  $\approx 80\%$  of eRASS AGN (Kolodzig et al. 2012) and with its sky coverage of  $\sim 14500 \text{ deg}^2$  one should be able to conduct high accuracy measurements of the linear bias factor. Among other, on-going surveys, the Pan-STARRS PS1  $3\pi$  survey (Chambers & the Pan-STARRS Team 2006) fulfills the necessary depth and sky coverage criteria.

BAO studies, on the other hand, require a much higher redshift accuracy of the order of  $\delta z \sim 0.01$ . Such an accuracy can be only achieved in spectroscopic surveys or in high quality narrow-band multi-filter photometric surveys. For example, for a  $4\sigma$  detection of BAOs in the redshift range  $0.8 < z < 1.2$ , a spectroscopical survey is needed of the depth of  $I > 22.5 \text{ mag}$  (Kolodzig et al. 2012) and sky coverage of at least  $\sim 20000 \text{ deg}^2$  (Fig. 8). Promising candidates are the pro-

posed 4MOST (de Jong et al. 2012) and WEAVE (Dalton et al. 2012) surveys, which would cover a large fraction of the sky with a multi-object spectrograph in the southern and northern hemisphere respectively. An important caveat is that the angular resolution of eRASS (FOV averaged HEW of  $\sim 30''$ ) is insufficient to provide accurate positions of X-ray sources for spectroscopic follow-ups with multi-object spectrographs. Therefore additional photometric surveys (for example Pan-STARRS PS1  $3\pi$ ) will be needed in order to refine source locations to the required accuracy.

Finally, we note that we excluded from consideration a number of observational effects and factors, such as source confusion and source detection incompleteness, positional accuracy, telescope vignetting and non-uniformities in the survey exposure, and a number of others. These factors and effects are well known in X-ray astronomy and data analysis methods and techniques exist to properly address them in the course of data reduction.

## 6. Summary

We have explored the potential of eRosita all-sky survey for large scale structure studies and have shown that eRASS with its  $\approx 3$  million AGN sample will supply us with outstanding opportunities for detailed LSS research. Our results are based on our previous work (Kolodzig et al. 2012), where we investigated statistical properties of AGN in eRASS, and the AGN clustering model of Hütsi et al. (2012).

We have demonstrated that with eRASS, the linear bias factor of AGN can be studied to unprecedented accuracy and detail. Its redshift evolution can be investigated with an accuracy of better than  $\sim 10\%$  using data from the sky patches of  $\sim 2500 \text{ deg}^2$ . Using the data from sky area of  $\geq 10000 \text{ deg}^2$ , statistically accurate redshift and luminosity resolved studies will become possible for the first time. Bias factor studies will yield meaningful results much before the full 4 years survey will be completed. The eRASS AGN sample will not only improve the redshift and luminosity resolution of bias studies but will also expand their luminosity range beyond  $L_{0.5-2.0 \text{ keV}} \sim 10^{44} \text{ erg s}^{-1}$ , thus making possible direct comparison of clustering properties of luminous X-ray AGN and optical quasars. These studies will dramatically improve our understanding of the AGN environment, triggering mechanisms, growth of super-massive black holes and their co-evolution with dark matter halos. The photometric redshift accuracy should be sufficient for the bias factor studies, although the impact of large fraction of catastrophic errors typical for standard broad band filter sets yet needs to be investigated (Hütsi et al. 2013, in prep.).

For the first time with the X-ray selected AGN, eRASS will be able to detect BAO with high statistical significance of  $\sim 10\sigma$ . Moreover, it will push the redshift limit of BAO detections far beyond the current limit of  $z \sim 0.8$ . The accuracy of BAO investigation in this uncharted redshift range will exceed that to be achieved by eBOSS, planned in the same timeframe and will be only superseded by BigBOSS proposed for implementation beyond 2020. Until then, eRASS AGN can potentially become the best sample for BAO studies beyond  $z \geq 0.8$ . However, in order for this potential to be realized and exploited, spectroscopic quality redshifts for large fraction of the sky will be required.

*Acknowledgements.* We thank Mirko Krumpe, Viola Allevalo and Andrea Merloni for useful discussions. A. Kolodzig acknowledges support from and participation in the International Max-Planck Research School (IMPRS) on Astrophysics at the Ludwig-Maximilians University of Munich (LMU).

## References

- Aird, J., Nandra, K., Laird, E. S., et al. 2010, *MNRAS*, 401, 2531  
 Alexander, D. M. & Hickox, R. C. 2012, *New A Rev.*, 56, 93  
 Allevalo, V., Finoguenov, A., Cappelluti, N., et al. 2011, *ApJ*, 736, 99  
 Allevalo, V., Finoguenov, A., Hasinger, G., et al. 2012, *ApJ*, 758, 47  
 Anderson, L., Aubourg, E., Bailey, S., et al. 2012, *MNRAS*, 427, 3435  
 Brandt, W. N. & Hasinger, G. 2005, *ARA&A*, 43, 827  
 Busca, N. G., Delubac, T., Rich, J., et al. 2013, *A&A*, 552, A96  
 Chambers, K. & the Pan-STARRS Team. 2006, in *The Advanced Maui Optical and Space Surveillance Technologies Conference*  
 Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, *MNRAS*, 362, 505  
 Comparat, J., Kneib, J.-P., Escoffier, S., et al. 2013, *MNRAS*, 428, 1498  
 Dalton, G., Trager, S. C., Abrams, D. C., et al. 2012, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 8446, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*  
 Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, 145, 10  
 de Jong, R. S., Bellido-Tirado, O., Chiappini, C., et al. 2012, *ArXiv e-prints*, 1206.6885  
 Dodelson, S. 2003, *Modern cosmology*  
 Draper, A. R. & Ballantyne, D. R. 2012, *ApJ*, 751, 72  
 Eisenstein, D. J. & Hu, W. 1998, *ApJ*, 496, 605  
 Eisenstein, D. J., Seo, H.-J., Sirko, E., & Spergel, D. N. 2007, *ApJ*, 664, 675  
 Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, *ApJ*, 633, 560  
 Fanidakis, N., Baugh, C. M., Benson, A. J., et al. 2012, *MNRAS*, 419, 2797  
 Hasinger, G., Miyaji, T., & Schmidt, M. 2005, *A&A*, 441, 417  
 Hickox, R. C., Jones, C., Forman, W. R., et al. 2009, *ApJ*, 696, 891  
 Hirschmann, M., Somerville, R. S., Naab, T., & Burkert, A. 2012, *MNRAS*, 426, 237  
 Hogg, D. W. 1999, *ArXiv Astrophysics e-prints*, 9905116  
 Hütsi, G. 2006, *A&A*, 449, 891  
 Hütsi, G., Gilfanov, M., Kolodzig, A., & Sunyaev, R. 2013, in prep.  
 Hütsi, G., Gilfanov, M., & Sunyaev, R. 2012, *A&A*, 547, A21  
 Kaiser, N. 1987, *MNRAS*, 227, 1  
 Kolodzig, A., Gilfanov, M., Sunyaev, R., Sazonov, S., & Brusa, M. 2012, *ArXiv e-prints*, 1212.2151  
 Koutoulidis, L., Plionis, M., Georgantopoulos, I., & Fanidakis, N. 2013, *MNRAS*, 428, 1382  
 Krumpe, M., Miyaji, T., & Coil, A. L. 2010, *ApJ*, 713, 558  
 Krumpe, M., Miyaji, T., Coil, A. L., & Aceves, H. 2012, *ApJ*, 746, 1  
 Lewis, A. & Bridle, S. 2002, *Phys. Rev. D*, 66, 103511  
 Merloni, A., Predehl, P., Becker, W., et al. 2012, *ArXiv e-prints*, 1209.3114  
 Miyaji, T., Krumpe, M., Coil, A. L., & Aceves, H. 2011, *ApJ*, 726, 83  
 Mountrichas, G., Georgakakis, A., Finoguenov, A., et al. 2013, *MNRAS*, 430, 661  
 Padmanabhan, N., Xu, X., Eisenstein, D. J., et al. 2012, *MNRAS*, 427, 2132  
 Peebles, P. J. E. 1980, *The large-scale structure of the universe*  
 Peebles, P. J. E. & Yu, J. T. 1970, *ApJ*, 162, 815  
 Predehl, P., Andritschke, R., Böhringer, H., et al. 2010, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 7732, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*  
 Richardson, J., Zheng, Z., Chatterjee, S., Nagai, D., & Shen, Y. 2012, *ApJ*, 755, 30  
 Richardson, J. W., Chatterjee, S., Zheng, Z., Myers, A., & Hickox, R. C. 2013, *ArXiv e-prints*, 1303.2942  
 Salvato, M., Ilbert, O., Hasinger, G., et al. 2011, *ApJ*, 742, 61  
 Sawangwit, U., Shanks, T., Croom, S. M., et al. 2012, *MNRAS*, 420, 1916  
 Schlegel, D., Abdalla, F., Abraham, T., et al. 2011, *ArXiv e-prints*, 1106.1706  
 Shen, Y., McBride, C. K., White, M., et al. 2012, *ArXiv e-prints*, 1212.4526  
 Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *MNRAS*, 323, 1  
 Slosar, A., Iršič, V., Kirkby, D., et al. 2013, *J. Cosmology Astropart. Phys.*, 4, 26  
 Sunyaev, R. A. & Zeldovich, Y. B. 1970, *Ap&SS*, 7, 3  
 Tegmark, M., Eisenstein, D. J., Strauss, M. A., et al. 2006, *Phys. Rev. D*, 74, 123507  
 Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, *ApJ*, 480, 22  
 Truemper, J. 1993, *Science*, 260, 1769  
 Voges, W., Aschenbach, B., Boller, T., et al. 1999, *VizieR Online Data Catalog*, 9010, 0  
 Wall, J. V. & Jenkins, C. R. 2012, *Practical Statistics for Astronomers*  
 Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2012, *ArXiv e-prints*, 1201.2434