Paris 2008 Workshop : two strategies for future large X-ray surveys of AGN with XMM-Newton

O. Garcet ; V. Borkowski ; E. Gosset ; P.G. Sprimont ; J. Surdej

Institute of Astrophysics and Geophysics, University of Liege, Belgium

In this short document we present two possible strategies to define the most-suited X-ray surveys capable to test the AGN unified scheme based upon the study of the evolution of the AGN clustering properties, as a function of the redshift, X-ray luminosity, and the X-ray classification. To this end, we will first compute and compare the two-point angular correlation function obtained from a soft X-ray selected sample of AGN (mostly type 1) and from a hard X-ray selected sample of AGN (mostly type 2). Then for each of these two samples, obtain the correlation function in both luminosity and redshift bins. Finally we plan to compute the spatial correlation function in order to directly derive the spatial correlation length.

The challenge is now to first define a sky area with enough X-ray sources in order to settle down the issue to a sufficiently high significance threshold (let us say 4-σ or more) and therefore draw robust conclusions. That is, obtain the correlation function with sufficiently small error bars.

To estimate the sky area needed concerning angular correlation studies, we start from the results of Gandhi et al. (2006). Concerning the soft band, they have fitted a power law model to the ACF. If we use the same binning size and fit parameters (θ_0 and γ), we have estimated that the survey size has to be about 5 times larger in order to have a 4-σ significance in the first 3 bins, up to an angular distance of 230 arcsec. This means that the requested sky area would have around 27 deg^2 if covered with 10 ks X-ray pointings. In order to have an estimation of the redshift distribution of soft X-ray selected AGN, we have used the values of Hasinger et al. (2005). They have found that about 31% have $z < 0.2$; 33% have $0.2 < z < 0.8$ and 35% have $0.8 < z < 3.2$. Therefore, in order to compute the ACF in these 3 redshift bins and still have the same 4-σ level, we would need to have a sky area again 3 times as large, that is around 80 deg^2. Note that according to the log(N)-log(S) of these soft X-ray sources, if we request deeper pointings of 20 ks, the expected number of X-ray sources twice fainter would be increased by about 1.6. Therefore, observing with 20 ks X-ray pointings, a sky area of 50 deg^2 is required.

Concerning the hard X-ray band, no clustering signal has been reported by Poshak et al. (2006). However, only keeping the hard X-ray sources which have an $HR > -0.2$ in the deeper G pointings (20ks), they have fitted a power law to the ACF. Doing the same exercise as for the soft X-ray sources, but here up to 270 arcsec, we find that a sky area of 18 deg^2 covered by 20 ks X-ray pointings would be needed.

The advantage here is that we request the 4-σ level only up to around 300 arcsec at most, for both the soft and hard samples. Therefore, these two large sky areas need not be contiguous.

Second, the sky area also has to be large enough in order to reduce the cosmic variance as much as possible to justify the fair sample hypothesis. As a quick check, the homogeneity scale of the universe is around 100 Mpc. At a median redshift of $z = 1$, this corresponds to an angular scale around 12500 arcsec which is about 3.5 deg. Therefore, even if a single contiguous X-ray survey layout is not needed, each patch would need a sky area of at least

5-10 deg^2. The sky areas quoted above (18 and 50 deg^2) are much larger than this threshold.

A third interesting configuration for the survey layout would consist of two sufficiently large X-ray surveys, but which would lie in very different directions in order to tackle isotropy issues.

Finally, for spatial correlation studies, our starting point is the recent paper of Yang et al. (2006). Their sample contains about 250 identified X-ray AGN over a sky area of 0.4 deg^2 in the CLASXS and detected in the [2-8] keV band. They have studied the evolution of the redshift-space correlation function in 4 redshift bins from $z = 0.1$ to $z = 3$. In order to tell whether there is a clustering evolution as a function of redshift at the 4-σ level, we estimate that the size of the sample has to be multiplied by 10-20. Therefore the needed number of identified X-ray sources is around 2500-5000.

With a 20 ks X-ray pointing, the limiting flux F_{2-10} is around $\sim 5 \times 10^{-15}$ cgs. At this flux limit, the sky density of hard X-ray sources is around 240 deg^{-2} and the density of type II AGN (defined as an X-ray source which has $HR > -0.2$) is around 100 deg^{-2}. Therefore, in order to have 5000 type II AGN, we need 50 deg^2.

For all of the above reasons, we suggest to design an X-ray survey which consists of 20 ks pointings over an area of 50 deg^2.

A) Extension of an already existing X-ray survey

The first strategy is to start from already existing X-ray surveys and extend it up to 50 deg^2. The main argument here is to use already observed multi-λ data (both photometric and spectroscopic), as the Swire, CFHTLS, SDSS,... in order to save huge amounts of observing time.

As a first suggestion, the future X-ray survey could completely cover the 49 deg^2 of the Swire, spread over 6 areas, with an already existing X-ray survey in each of the 6 areas. Note that the XMM-LSS (presently 10 deg^2) is in one of these 6 areas and already completely covers its Swire field. Moreover, UKIDSS already partially covers these 6 Swire areas. Proceeding this way, it would not require much observing time to have CFHT photometry over these 49 deg^2. Also note that the field W1, which is part of the CFHTLS, is on the XMM-LSS survey and has already been observed.

With an X-ray survey of 50 deg^2, we have shown that we will already be able to compute the angular correlation function with a 4-σ level in both the soft and X-ray bands, as a function of the hardness ratio and also in several redshift bins, using the photometric redshifts. Unfortunately, these 6 Swire areas are almost not covered by spectroscopic surveys like the SDSS or the 2dF QSO Redshift survey.

This kind of layout will also open the road to the estimation of the spatial correlation function for both soft and hard X-ray selected AGN over both a large and not so shallow sky area.

Indeed, up to now, the spatial correlation function of AGN has been based on samples with at most 300-400 X-ray selected AGN and mostly in the soft band (e.g Mullis et al. (2004) ; Gilli et al. (2005) ; Yang et al. (2006)). In order to be able to compare the spatial correlation function of type I and type II AGN at a threshold of at least 4-σ, we will need a very deep spectroscopic follow-up (e.g $R \sim 23.5$) using 8m class telescopes over the planned sky area of 50 deg^2 : Assuming that $\frac{F_X}{F_R} \sim 10$ for hard X-ray selected AGN, we will end up with a density of 190 identified hard X-ray sources per deg^2, among which 73 are type II AGN. Therefore, over 50 deg^2, we will have around 9300 identified hard X-ray selected AGN, among which 3600 are type II AGN.

Using Vimos on the VLT, an optical spectrum with $\frac{S}{N} \sim 3-4$ of a source which has $R = 23.5$ is obtained in about 2h. 16 Vimos pointings are needed to cover 1 deg^2. Therefore 800 vimos pointings are required, which amounts to about 1600h of spectroscopic follow-up.

It can be shown that basically if we multiply the sample size by x, the error bar on the correlation function will be decreased by x. Therefore, going from 400 to 9300 X-ray selected AGN, the error bars on the spatial correlation function would be divided by 23, which is highly significant. Note that for the type II AGN alone, the error bars will be divided by 9.

The total X-ray survey over the 6 Swire fields is well suited to tackle isotropic issues, as these 6 Swire fields are in very different region of the sky.

The important thing to keep in mind in X-ray surveys is that multi-λ data are mandatory along with the X-ray data, because it is almost impossible to identify the source based upon the X-ray observations alone, due to the poor photon statistics.

This first alternative is therefore very attractive because it saves a huge amount of follow-up observing time allocated to obtain these mandatory multi-λ data.

B) Define an entirely new X-ray survey design

The other strategy is to start from scratch and define an entirely new X-ray layout. Basically, this second alternative would try to answer the same questions as the first strategy. As we have seen before, a sky area around 50 deg^2 covered with 20 ks X-ray pointings is required to tackle these issues, for both angular and spatial clustering analysis. The interest of the second alternative is that we do not have any location constraint from pre existing multi-λ data, and that we will be able to have an entirely contiguous sky area of 50 deg^2. This is very helpful for example to study the evolution of the correlation length as a function of the survey size.

For this second strategy, we will also need a very deep spectroscopic follow-up using 8m class telescopes (e.g $R \sim 23.5$) over the planned sky area of 50 deg^2 : Assuming that $\frac{F_X}{F_R} \sim 10$ for hard X-ray selected AGN, we will end up with a density of 190 identified hard X-ray sources per deg^2, among which 73 are type II AGN. Therefore, we will have around 3600 type II AGN within the 50 deg^2.

Using Vimos on the VLT, an optical spectrum with $\frac{S}{N} \sim 3-4$ of a source which has $R = 23.5$ is obtained in about 2h. 16 Vimos pointings are needed to cover 1 deg^2. Therefore 800 vimos pointings are required, which amounts to about 1600h of spectroscopic follow-up

We would like to stress again that the X-ray data alone are almost useless to do science. Therefore, the spectroscopic follow-up is an unavoidable part of the project and we have to really think about its feasibility.

This second alternative is more suited for spatial correlation studies over a large and contiguous sky area than the first one. Although the main drawback is the huge amount of required follow up observing time.

Therefore, as a final conclusion, we would like to stress that in order to make this project viable, a spectrograph similar to the 2dF coupled with an 8 meter class telescope is mandatory.