A low redshift X-ray survey for normal galaxies.

1. Motivation: Observations at X-ray wavelengths are a unique diagnostic of galaxies, providing information that is complementary to that inferred from other wavebands. The binary stellar population (high and low mass X-ray binaries) and the hot gas component of external galaxies, the metal enrichment of the inter-stellar and intergalactic medium, are all subjects of X-ray astronomy. Until recently however, X-ray observations of normal galaxies (i.e. non-AGN) were restricted to a selected number of X-ray bright systems in the local Universe (< 100Mpc). While these data are useful for constraining the nature of the X-ray emission in galaxies, they cannot be used for statistical studies, e.g. luminosity function. This has been changed by surveys with the *XMM-Newton* and the *Chandra* missions, which have detected the first ever X-ray selected normal galaxy samples, at low and moderate redshifts (Hornschemeier et al. 2003; Georgakakis et al. 2004; Kim et al. 2006). These samples have been used to estimate the number density of X-ray selected galaxies on the sky (Bauer et al. 2004; Georgakakis et al. 2007), quantify their X-ray luminosity function (XLF; Norman et al. 2004; Georgantopoulos et al. 2005; Ptak et al. 2007), calibrate the relation between X-ray luminosity and star-formation rate (Georgakakis et al. 2006) and assess the significance of the diffuse hot gas component of galaxies in the enrichment of the intra-cluster medium (Hornschemeier et al. 2006).

These studies show that normal galaxies are a fast evolving population at redshifts $z < 1$ and that they will outnumber AGN at faint X-ray fluxes, $f_X < 10^{-17}$erg/s/cm^2 (Fig. 1). Future ultra deep X-ray surveys will therefore be dominated by galaxies at high redshift, $z > 1$, opening the prospect of studying their evolution close to, or past the peak of the star-formation rate density of the Universe. There are suggestions that the evolution of galaxies at X-ray wavelengths is going to be different from that determined at other wavebands, depending on the characteristic timescales of high and low mass X-ray binaries (Ghosh & White 2001).

In order to fully exploit the potential of high redshift surveys, we first need to improve our understanding of X-ray selected normal galaxies in the nearby Universe, $z \approx 0.1$. A large X-ray survey at low redshift will provide tight constraints on the statistical properties of galaxies, such as the XLF, which can then be combined with high redshift surveys for evolutionary studies. Currently, the largest sample of X-ray selected normal galaxies at redshifts $z \approx 0.1$ covers about 15deg^2 and has only 67 sources (Georgakakis et al. 2006), highlighting the need for larger surveys (see Fig 2). Another major source of uncertainty in normal galaxy studies is associated with the sample selection. Because X-ray surveys are dominated by AGN, the identification of normal galaxies is not straightforward. This involves either rigid cutoffs to a pre-defined set of parameters (Bauer et al. 2004; Georgakakis et al. 2007; see Fig. 3), or the estimation of the probability that a source with a particular set of parameters is a galaxy (Norman et al. 2004; Ptak et al. 2007). Both approaches rely heavily on apriori knowledge of the range of the parameter space occupied by normal galaxies. Low redshift surveys can provide useful input on this issue. High signal–to–noise observations are feasible for nearby, apparently bright sources, thereby allowing a reliable separation of galaxies from AGN without the need to apply priors to the X-ray source properties.

The issues raised above can be addressed by performing an XMM X-ray survey in the region of the Sloan Digital Sky Survey (SDSS; Stoughton et al. 2002). The SDSS is currently the largest photometric and spectroscopic sample of galaxies at low redshift, covering about 1/3 of the sky. The strongest feature of the SDSS is the moderate resolution and high signal-to-noise optical spectroscopy for galaxies at $z \approx 0.1$. Such high quality optical spectroscopy is an efficient and highly reliable way of identifying non-AGN among X-ray sources (see Fig. 3).

2. Survey setup: It is proposed that the XMM observations are carried out in the area of the SDSS. This is the largest optical survey of the nearby Universe (≈ 0.1), providing high quality optical photometry and spectroscopy over about

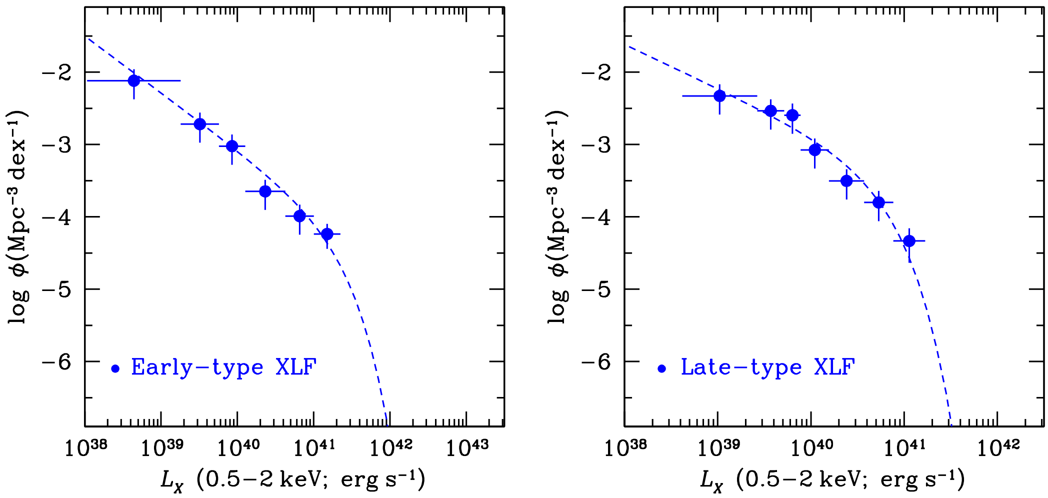

Figure 2: The XLF for early (left panel) and late (right panel) type X-ray selected normal galaxies. In both panels the points and the dashed (blue) lines correspond to the XLF estimates and best-fits presented by Georgakakis et al. (2006). These authors compiled the largest sample of X-ray selected normal galaxies at redshifts $z < 0.2$, by combining shallow wide-area surveys (Needles in the Haystack Survey, 1XMM) with deep pencil-beam samples (Chandra Deep Field North and South) to get a wide X-ray luminosity baseline. The final sample consists of 67 normal galaxies, of which 33 have emission line optical spectra (late types) and 34 show absorption optical lines only (early types). AGN contamination was controlled by selecting as galaxy candidates X-ray sources with (i) low X-ray–to–optical flux ratio, $\log f_X/f_{opt} < -2$, (ii) soft X-ray spectral properties, $N_H < 10^{22}\mathrm{cm}^{-2}$ for $\Gamma = 1.9$, (iii) optical spectra that did not show AGN signatures and (iv) X-ray luminosity $L_X(0.5-2\mathrm{keV}) < 10^{42}\mathrm{ergs}^{-1}$. Although these XLFs are the best current estimates at low redshift, there are issues that need to be addressed. Firstly, the small number of sources is a limitation in the determination of the XLF. Secondly, cosmic variance is a concern at the faint end of the XLF, as the majority of the sources below $L_X \approx 10^{39}$erg/s are mainly from the pencil-beam Chandra Deep Fields. Finally, the $\log f_X/f_{opt}$ and L_X cutoffs adopted for the selection of normal galaxies guarantee minimal contamination by AGN ($< 15\%$; Georgakakis et al. 2007), but at the expense of possibly missing X-ray luminous galaxies, such as starbursts or massive ellipticals.

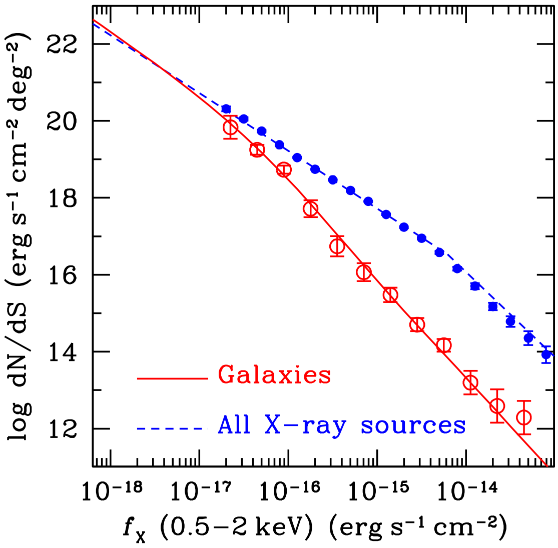

Figure 1: Differential X-ray number counts for normal galaxies (open red circles) and the entire X-ray point source population (filled blue circles). The latter is estimated by combining the Chandra Deep Field North and South with the Extended Chandra Deep Field South and the Chandra survey of the AEGIS (Georgakakis et al. in preparation). The dashed (blue) line passing through the points is the best fit double power-law relation. The galaxy dN/dS is from Georgakakis et al. (2007). The continuous (red) line is estimated by integrating the X-ray luminosity function of Georgakakis et al. (2006) assuming luminosity evolution of the form $(1+z)^{2.4}$. Below about 10^{-17}ergs^{-1}cm^{-2} galaxies become the dominant population.

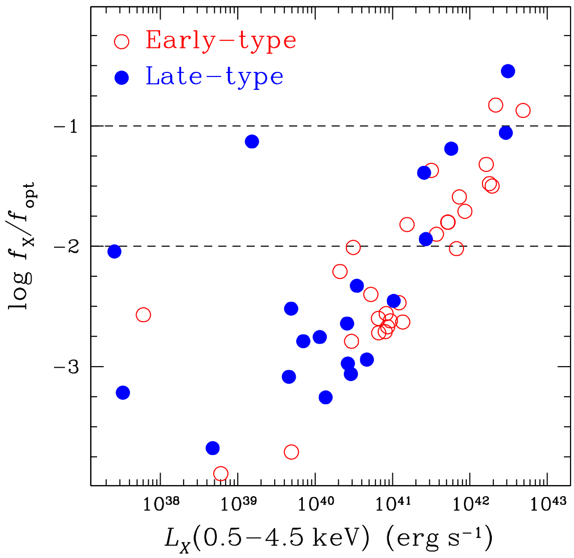

Figure 3: X-ray–to–optical flux ratio, $\log f_X/f_{opt}$, plotted as a function of X-ray luminosity in the 0.5-4.5keV spectral band. The points are X-ray selected normal galaxies identified by cross-correlating the 2XMMp catalogue with the SDSS main spectroscopic sample. The open (red) circles are normal galaxies with absorption line optical spectra, the filled (blue) circles are star-forming systems. The high quality SDSS spectra can be used to separate normal galaxies from AGN reliably, thereby providing clean and unbiased galaxy samples, independent of e.g. $\log f_X/f_{opt}$ and L_X. This is a significant improvement over previous studies that applied priors to these parameters (e.g. $\log f_X/f_{opt} < -2$ and $L_X < 10^{42}\,\mathrm{erg\,s^{-1}}$) to minimise AGN contamination, in the expense of missing luminous normal galaxies. The plot indeed shows that a sizeable fraction of X-ray selected normal galaxies, about 50%, have either $\log f_X/f_{opt} > -2$ or $L_X > 10^{42}\,\mathrm{erg\,s^{-1}}$.

10 000deg^2 as well as advanced data products, such as star-formation rates, stellar masses and local galaxy density measurements.

The main spectroscopic sample of the SDSS includes galaxies brighter than $r = 17.7$ with a redshift distribution that extends out to $z \approx 0.2$. To take advantage of the SDSS optical spectroscopy it is proposed to limit the normal galaxy sample to that redshift. The area of the proposed XMM survey is set by the requirement to provide a large and unbiased sample of galaxies over a wide luminosity range so that the XLF can be constrained to a high degree of accuracy. In particular, it is required that the Poisson uncertainty in individual XLF bins is smaller than $\delta N/N = 20\%$ over the interval $L_X = 10^{38} - 10^{42}\,\mathrm{erg\,s^{-1}}$. This essentially sets the area to about 200 deg^2 and the exposure time to abour 5 ks per pointing. This is demonstrated in Fig. 4, which plots as a function of L_X the expected number of galaxies at $z < 0.2$ and the relative Poisson uncertainty in the determination of the XLF.

3. Choice of field: It has been argued above that a 200 deg^2 XMM survey in the SDSS area is needed, with an exposure time of 5 ks per pointing. An obvious field choice for such a survey is the part of the SDSS that is going to be observed spectroscopically by AAT/AOmega as part of the Galaxy And Mass Assembly (GAMA) survey. This optical survey consists of 3 equatorial SDSS fields covering a total area of 200 deg^2 (see Fig. 5). In addition to the existing SDSS observations, GAMA will provide optical spectroscopy for *all* galaxies to $r = 19.8$ mag about, 2 mag fainter than the SDSS main spectroscopic sample. The multiwavelength data in this field are presented in Table 1. The wealth of observations in this patch of the sky guarantee minimal requirements in terms of follow-up observations.

The XMM Survey Science Centre has constructed a serendipitous catalogue of X-ray sources (2XMM) using archival XMM observations that could potentially be used for the proposed study. The overlap between 2XMM and SDSS is about 100 deg^2. However, about 25 per cent of these fields are targeting extended sources, such as clusters or nearby galaxies, resulting in high background and bright effective point source detection limit. These fields are therefore unsuitable for the proposed study and limit the usable overlap between SDSS and 2XMM to about 75 deg^2. Such a serendipitous survey is useful for most of goals of the proposed study and can be used in addition to the proposed observations, reducing the requested time by 30%. The penalty in this approach is that one looses the GAMA spectroscopic survey data.

A clear gap in the multiwavelength coverage of Table 1 is the mid- and far-IR. This is unfortunate as the mid-IR is proposed as a promising wavelength regime for selecting normal galaxies against AGN and the only one available to future ultra-deep surveys at high-z, where high quality spectroscopy is impossible to obtain. However, calibration and refinement of this selection method is required. One way to address this issue is to survey with XMM SWIRE fields that overlap with the SDSS. There are 3 such fields, ELAIS-N1/N2 and Lockman Hole, covering a total area of 24 deg^2. The penalty here is that again the GAMA survey does not overlap with any of the SWIRE fields. Fig. 6 demonstrates that the SWIRE flux limits are sufficient to detect a large fraction of the X-ray star-forming galaxies at $z \approx 0.1$, at least in the IRAC and MIPS-24μm bands and possibly at the MIPS 70 and 160μm.

In summary, it is proposed that an XMM survey is performed over 100deg^2 in the GAMA area and 24 deg^2 in the SWIRE ELAIS-N1/N2 and Lockman Hole fields. This will be complemented with 75 deg^2 of archival data in the SDSS to provide a 200 deg^2 X-ray survey of the low-z Universe, $z < 0.2$. Assuming a setup time for XMM of 4 ks per pointing and 0.2 deg^2 FOV it is estimated that the proposed program requires 5.6Ms.

4. Specific science aims:

(i) Constrain the local XLF of normal galaxies to an accuracy of better than 20% over 3 dex in X-ray luminosity using a sample that is free from selection biases. An accurate estimate of the local XLF is essential for evolution studies and also to predict the expected number of normal galaxies in future ultra deep surveys.

(ii) Explore the mid-IR SED of X-ray selected star-forming galaxies. The $L_X - L_{IR}$ correlation for star-forming galaxies is proposed as an efficient and reliable method for selecting normal galaxies

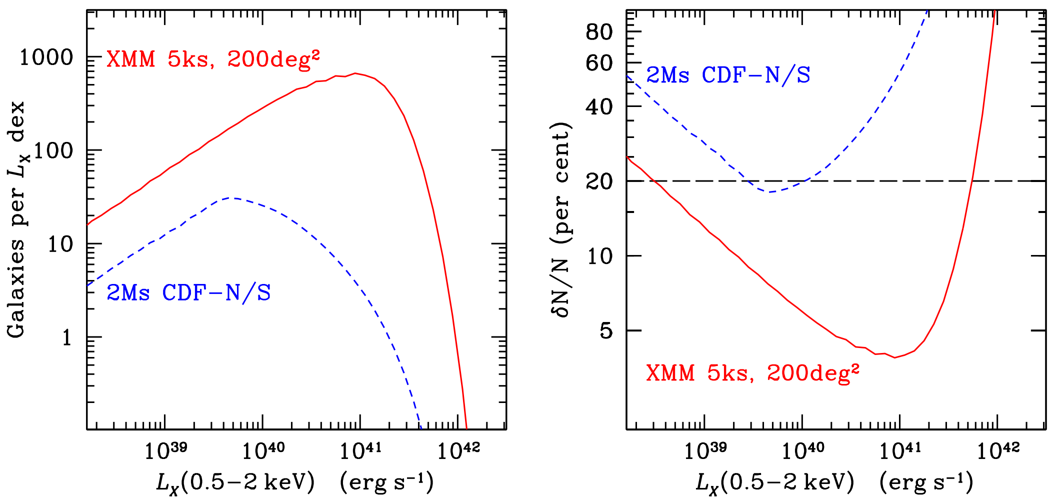

Figure 4: **Left panel:** Number of X-ray selected normal galaxies at $z < 0.2$ as a function of X-ray luminosity for a 5 ks XMM survey over 200 $\deg^2$ (red continuous) and the combined sample of the 2Ms Chandra Deep Field North and South (blue dashed). These curves are estimated by integrating the Georgakakis et al. (2006) XLF taking into account the sensitivity curves of the different observational setups. The 2Ms sample is more efficient in finding galaxies, i.e. higher number of sources per square degree. A wide area survey with XMM-Newton however, will detect a factor of at least 5 more galaxies at $z < 0.2$ at any luminosity. **Right panel:** Relative per cent Poisson uncertainty of the XLF as a function of luminosity. The proposed survey will provide accuracy of better than 20% at both faint and bright luminosities, $10^{38} - 10^{42}\,\mathrm{erg\,s^{-1}}$.

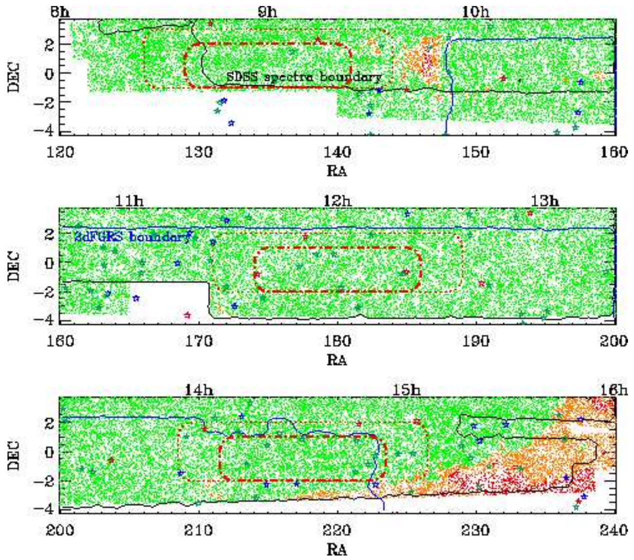

Figure 5: SDSS regions that will be targeted by AOmega as part of the GAMA spectroscopic survey to $r = 19.8$, 2 mag deeper than the SDSS main spectroscopic sample. The points represent SDSS galaxies in a narrow magnitude range. Green points represent areas where Galactic extinction in $r < 0.2$; orange points for $0.2 - 0.4$; and red points for > 0.4. The black and blue lines define the boundaries for spectroscopy from SDSS and 2dFGRS, respectively. The stars represent stars from the bright star catalogue. The thick dash-and-dotted red lines represent our initial 3x12deg^2 target regions.

Table 1: Multi-wavelength coverage of the equatorial 200deg^2 SDSS strip where the Galaxy And Mass Assembly (GAMA) spectroscopic survey is going to be carried out. The spectroscopic data from both the SDSS and the GAMA and the high quality optical and near-IR photometry make this strip ideal for a shallow wide area XMM survey of the nearby Universe.

Survey/Mission name	waveband	flux/magnitude limit
SDSS-photometry	$ugriz$	$r \approx 22.5$ mag
SDSS-spectroscopy	3500-9000Å	$r < 17.7$ mag
GAMA-spectroscopy	3500-9000Å	$r < 19.8$ mag
UKIDSS-LAS	$YJHK$	$K \approx 18.5$ mag
FIRST	1.4 GHz	$S_{1.4\mathrm{GHz}} \approx 1$ mJy

against AGN, especially in future ultra-deep surveys targeting the high redshift Universe, where the high quality optical spectroscopy of the SDSS is impossible to obtain. The proposed study will allow us to explore whether the $L_X - L_{IR}$ correlation is a good alternative to optical spectroscopy for finding normal galaxies and also what is the expected level of AGN contamination.

(iii) Constrain the kpc and Mpc scale environment of normal galaxies at $z < 0.2$. It is expected that in high density environments, like groups or clusters, processes like ram pressure will strip the hot gas component of galaxies, which will appear X-ray faint for their optical luminosity. We therefore expect a dependence of local density on $\log f_X/f_{opt}$.

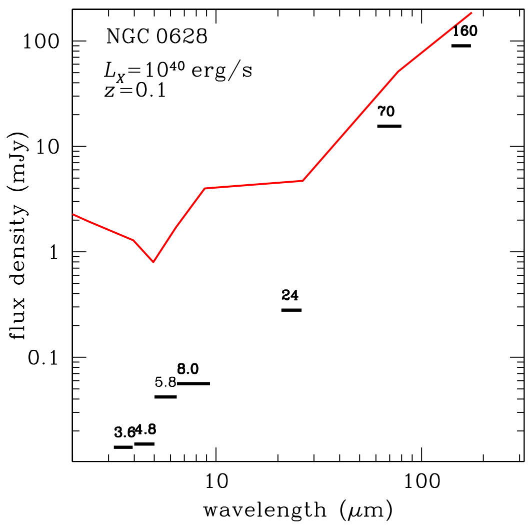

Figure 6: Expected mid-IR flux for a moderate star-forming galaxy with $L_X(0.5-2\,\mathrm{keV}) = 10^{40}\,\mathrm{erg\,s^{-1}}$. For the conversion from X-ray to infrared luminosity we adopt the $L_X - L_{FIR}$ correlation of Ranalli et al. (2003). The L_{FIR} is converted to $12\mu m$ luminosity using the relation of Takeuchi et al. (2005). The infared SED adopted here is that of the Sc galaxy NGC 0608 (Dale et al. 2007) observed as part by the SINGS program. The flux limits of the SWIRE survey in the IRAC and MIPS bands are shown with the horizontal bars. We expect to detect a large fraction of the X-ray selected star-forming galaxies with $z \approx 0.1$ to the IRAC and MIPS-$24\mu m$ flux limits, even in the case of low X-ray luminosities $L_X(0.5-2\,\mathrm{keV}) > 10^{39}\,\mathrm{erg\,s^{-1}}$. The most luminous of these systems ($L_X(0.5-2\,\mathrm{keV}) = 10^{40}\,\mathrm{erg\,s^{-1}}$) will also be detected at longer wavelengths. The IR SED is believed to be the only tool for selecting star-forming galaxies against AGN in future ulltadeep X-ray surveys of the high-z Universe, where high quality optical spectroscopy is difficult or impossible to obtain. The proposed study will refine this normal galaxy selection method.