

SPIDERS: selection of spectroscopic targets using AGN candidates detected in all-sky X-ray surveys

T. Dwelly,^{1*} M. Salvato,¹ A. Merloni,¹ M. Brusa,^{1,2,3} J. Buchner,^{4,5} S. F. Anderson,⁶ Th. Boller,¹ W. N. Brandt,^{7,8,9} T. Budavári,^{10,11,12} N. Clerc,^{1,13,14} D. Coffey,¹ A. Del Moro,¹ A. Georgakakis,¹ P. J. Green,¹⁵ C. Jin,¹ M.-L. Menzel,¹ A. D. Myers,¹⁶ K. Nandra,¹ R. C. Nichol,¹⁷ J. Ridl,¹ A. D. Schwobe¹⁸ and T. Simm¹

Affiliations are listed at the end of the paper

Accepted 2017 April 5. Received 2017 March 31; in original form 2017 January 31

ABSTRACT

SPIDERS (SPectroscopic IDentification of eROSITA Sources) is a Sloan Digital Sky Survey IV (SDSS-IV) survey running in parallel to the Extended Baryon Oscillation Spectroscopic Survey (eBOSS) cosmology project. SPIDERS will obtain optical spectroscopy for large numbers of X-ray-selected active galactic nuclei (AGN) and galaxy cluster members detected in wide-area *eROSITA*, *XMM-Newton* and *ROSAT* surveys. We describe the methods used to choose spectroscopic targets for two sub-programmes of SPIDERS X-ray selected AGN candidates detected in the *ROSAT* All Sky and the *XMM-Newton* Slew surveys. We have exploited a Bayesian cross-matching algorithm, guided by priors based on mid-IR colour-magnitude information from the *Wide-field Infrared Survey Explorer* survey, to select the most probable optical counterpart to each X-ray detection. We empirically demonstrate the high fidelity of our counterpart selection method using a reference sample of bright well-localized X-ray sources collated from *XMM-Newton*, *Chandra* and *Swift-XRT* serendipitous catalogues, and also by examining blank-sky locations. We describe the down-selection steps which resulted in the final set of SPIDERS-AGN targets put forward for spectroscopy within the eBOSS/TDSS/SPIDERS survey, and present catalogues of these targets. We also present catalogues of $\sim 12\,000$ *ROSAT* and ~ 1500 *XMM-Newton* Slew survey sources that have existing optical spectroscopy from SDSS-DR12, including the results of our visual inspections. On completion of the SPIDERS programme, we expect to have collected homogeneous spectroscopic redshift information over a footprint of ~ 7500 deg² for >85 per cent of the *ROSAT* and *XMM-Newton* Slew survey sources having optical counterparts in the magnitude range $17 < r < 22.5$, producing a large and highly complete sample of bright X-ray-selected AGN suitable for statistical studies of AGN evolution and clustering.

Key words: surveys – galaxies: active – quasars: general – galaxies: Seyfert – cosmology: observations – X-rays: galaxies.

1 INTRODUCTION

X-ray emission is a signpost of accretion of matter on to the supermassive black holes that seed the whole population of massive galaxies and may strongly influence their formation and subsequent evolution. X-ray selected samples of active galactic nuclei (AGN) are particularly powerful because, compared to UV/optical/mid-IR selection methods, X-ray selection is much less susceptible to (but

not completely immune from) the deleterious effects of obscuration by intervening material and the dilution of AGN light by the host galaxy (e.g. see the recent review by Brandt & Hasinger 2005, and references therein).

Samples of X-ray selected AGN have been relatively small compared to the purely optically selected AGN available from large area optical surveys such as the Sloan Digital Sky Survey (SDSS; York et al. 2000). Although there are many X-ray survey fields currently under active study, the spectroscopic completeness is typically low, or the survey extents are typically small (Brandt & Hasinger 2005). As a result, even amongst the most intensively

* E-mail: dwelly@mpe.mpg.de

studied fields, the number of X-ray AGN with spectroscopic identifications does not exceed a few thousand sources per field, as for example, in the XBoötes/the AGN and Galaxy Evolution Survey (Kochanek et al. 2012), and in the Baryon Oscillation Spectroscopic Survey (BOSS) ancillary project within the Northern XMM-XXL field (Menzel et al. 2016). It is therefore not surprising that our understanding of black hole growth across cosmic time lags significantly behind investigations of galaxy evolution. Indeed, the physical conditions under which AGN are fuelled likely depend on a number of parameters, such as host-galaxy stellar mass or position within the cosmic web. Disentangling the relative significance of those factors requires large samples to account for the potentially large intrinsic scatter of relations and co-variances between parameters of interest.

Efforts to obtain complete redshift information for all-sky X-ray samples, the exemplar being the Röntgen Satellite All-Sky Survey (RASS; Voges et al. 1999, 2000; Böller et al. 2016), have been hampered by the typically rather poor positional accuracy of the X-ray detections. For example, the mean and 95th percentile of the 1σ error radii are ~ 20 and ~ 35 arcsec, respectively for sources in the RASS catalogue. Historically, this has made selection of the correct optical counterparts difficult because, even at the relatively shallow depths of currently available wide-area imaging (e.g. the $r \sim 22.5$ mag limit reached by SDSS imaging; Aihara et al. 2011), there are already many possible optical counterparts found within the error circle of each all-sky X-ray source. Despite these difficulties, several groups have reported the results of cross-matching the RASS catalogues to counterparts found in wide-area optical and near-IR surveys (e.g. Véron-Cetty et al. 2004; Mickaelian et al. 2006; Parejko et al. 2008; Haakonsen & Rutledge 2009; Greiner & Richter 2015). However, these studies have often relied on some degree of human interaction in the cross-matching process, or have been limited to only the bright end of the population, both of which are undesirable features when compiling well-understood and complete samples.

Even where the cross-matching hurdle has been overcome, the follow up of tens of thousands of X-ray sources with single object spectrographs requires a prohibitively large telescope time allocation. The largest RASS follow-up programmes to date have, by necessity, focused on the optically bright part of the X-ray source population, and have been carried out as a small component of large-scale galaxy redshift surveys exploiting wide-field highly multiplexed fibre-fed spectrographs. For example, Anderson et al. (2007, hereafter A07) report spectroscopically identified counterparts for ~ 7000 RASS sources covering 5740 deg^2 (i.e. the spectroscopic footprint of the SDSS 7th data release, DR7, Abazajian et al. 2009). Their sample comprises 6224 broad line AGN (BLAGN), 515 emission line galaxies (ELGs) and 266 BL Lacs, based on visual inspection (VI) of the spectra. Another large sample was presented by Mahony et al. (2010), who report reliable spectroscopic identifications for 1715 RASS bright catalogue sources covering $17\,046 \text{ deg}^2$ obtained as part of the 6-degree-Field Galaxy Survey (6dFGS; Jones et al. 2004, 2009). The RASS-6dFGS sample has a 90 per cent redshift success rate at $b_j = 17.5$ (Vega) but has a rapidly declining success rate towards fainter fluxes, and so is dominated by optically bright objects. Unfortunately, these existing large spectroscopic samples are still incomplete, since a large fraction of the X-ray detections still lack a spectroscopically measured counterpart, and inhomogeneous, since a variety of criteria or supporting data have been used to select the counterparts to X-ray sources and to determine which of those counterparts receive spectroscopic follow up.

SPIDERS (**SP**ectroscopic **ID**entification of **eROSITA** Sources) is an observational programme within the SDSS-IV project (Blanton et al. 2017) which seeks to improve upon the aforementioned situation. SPIDERS will run for up to 6 yr (2014–2020) alongside the Extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016) and Time Domain Spectroscopic Survey (TDSS; Morganson et al. 2015) projects. The primary goal of SPIDERS is to obtain extensive, homogeneous and complete spectroscopic follow-up of extragalactic X-ray sources, both point-like and extended, using data from X-ray satellites and over the SDSS extragalactic imaging footprint. SPIDERS naturally splits into two main components; an AGN programme and a galaxy clusters programme (the latter is described by Clerc et al. 2016). The SPIDERS-AGN programme has been designed to collect $\sim 40\,000$ spectra of X-ray AGN, and to bring population studies of accreting supermassive black holes to a new level of accuracy. First demonstrations of the SPIDERS science applications, based on the BOSS follow-up of X-ray selected AGN in the Northern XMM-XXL field, were presented by Menzel et al. (2016) and Liu et al. (2016). SPIDERS will target X-ray sources detected by the forthcoming all-sky X-ray survey to be carried out by *eROSITA* (extended **R**öntgen Survey with an **I**maging **T**elescope **A**rray; Merloni et al. 2012; Predehl et al. 2016). However, we have started the SPIDERS project in advance of the *eROSITA* launch (scheduled for 2018). We present in this paper the first phase of the SPIDERS-AGN programme (a.k.a. ‘Tier-0’) which exploits existing (pre-*eROSITA*) all-sky X-ray source catalogues to explore the bright end of the X-ray AGN population.

An initial goal of the SPIDERS project is to obtain highly complete and reliable identifications for the optical counterparts to all RASS sources (from both the bright and faint catalogues; Voges et al. 1999, 2000), that fall within the eBOSS survey footprint and that have possible counterparts with magnitudes within the accessible range ($17 < i < 22.5$). In addition, SPIDERS will obtain redshifts for sources detected in the *XMM-Newton* Slew Survey (XMMSL; Saxton et al. 2008), which covers a very wide sky area (around 2/3 of the full sky). The XMMSL is a factor of a few shallower than the RASS but has the advantage of being sensitive over a broader and harder energy range (0.2–12 versus 0.1–2.4 keV).

As we discuss in detail later, the mid-IR, specifically the all-sky survey performed by the *Wide-field Infrared Survey Explorer* (*WISE*; Wright et al. 2010), is the vital stepping-stone that allows us to correctly select optical counterparts to the bright X-ray selected AGN in the RASS and XMMSL surveys. Studies made using the *Infrared Astronomical Satellite* (*IRAS*; Neugebauer et al. 1984), *Spitzer* (Werner et al. 2004) and *WISE* observatories have shown that AGN activity is almost always associated with mid-IR emission (e.g. Elvis et al. 1994; Stern et al. 2005, 2012; Assef et al. 2013). Indeed, the spectral energy distribution of X-ray selected AGN is characterized by a tight correlation between near-IR and X-ray flux (e.g. Mainieri et al. 2002; Brusa et al. 2005; Civano et al. 2012; Marchesi et al. 2016), particularly when high-spatial resolution mid-IR and hard-X-ray measurements are available (Gandhi et al. 2009; Asmus et al. 2014). The tendency of luminous AGN to stand out from other astronomical populations in the mid-IR has been extensively exploited to separate them from the field galaxy population (e.g. Lacy et al. 2004; Stern et al. 2005; Hickox et al. 2007; Donley et al. 2012; Assef et al. 2013; Mateos et al. 2013). In this work, we show how the combination of *WISE* imaging data and a Bayesian cross-matching algorithm (full details of which will be presented by Salvato et al., in preparation), can be used effectively to overcome most of the issues listed above, which have so far hampered the

realization of highly complete follow-up programmes for the RASS and XMMSL surveys.

The paper is laid out as follows. In Section 2, we describe the data sets used in this study. In Section 3 we describe the details of the cross-matching process, the selection of targets for spectroscopy within SPIDERS and the process by which we have visually inspected SDSS-DR12 spectra associated with our X-ray samples. In Section 4, we present several independent tests of the fidelity of our cross-matching and target selection schemes. In Section 5, we discuss the properties of the $>13\,000$ RASS and XMMSL sources with existing SDSS-DR12 spectral identifications. In Section 6, we compare our sample to that of A07 and describe our expectations for the completed SPIDERS-AGN programme.

Throughout this paper we express magnitudes in their native systems: AB magnitudes for SDSS (Fukugita et al. 1996) and Vega magnitudes for *WISE* (Assef et al. 2013). In order to allow direct comparison with existing works from the X-ray survey literature, we adopt a flat Λ cold dark matter cosmology with $h = H_0/[100 \text{ km s}^{-1} \text{ Mpc}^{-1}] = 0.7$; $\Omega_M = 0.3$; $\Omega_\Lambda = 0.7$, broadly consistent with the most recent *Wilkinson Microwave Anisotropy Probe* and *Planck* determinations (e.g. Hinshaw et al. 2013; Planck Collaboration XIII 2016). The data products released in this work can be obtained from the MPE X-ray surveys website (<http://www.mpe.mpg.de/XraySurveys>).

2 PREPARATION OF INPUT DATA SETS

In this section, we detail the steps taken to collate and prepare the input data sets that have been used to produce the final lists of SPIDERS targets. In all cases, we have only considered targets that lie within the area defined by the $10\,778 \text{ deg}^2$ of the SDSS-BOSS imaging footprint.¹ This footprint is a superset of the area that will be considered for the eBOSS observations. The BOSS imaging footprint consists of two large contiguous regions; 70 per cent of the total area is in the North Galactic Cap (NGC), and the remainder is in the South Galactic Cap (SGC), see e.g. Fig. 1. It is expected that, after 6 yr of operations, eBOSS will have observed approximately 7500 deg^2 within this footprint. The spatial filtering was carried out using the POLYID tool from the MANGLE² software suite (Hamilton 1993; Hamilton & Tegmark 2004; Swanson et al. 2008).

2.1 ROSAT All Sky Survey catalogue

The Röntgen Satellite (*ROSAT*; Trümper 1982) was used to carry out a 6-month-long scanning sky survey (RASS) in 1990–1991, covering around 99.7 per cent of the entire sky. Despite the many advanced X-ray observatories that have been launched since *ROSAT*, none has had a larger survey ‘grasp’ (used here to mean the product of telescope collecting area and field of view) in the soft X-rays, and so the RASS remains the most sensitive all-sky survey in the soft X-ray band (0.1–2.4 keV).

Two first-generation RASS source catalogues were produced, the Bright Source Catalogue (BSC; Voges et al. 1999), containing 18 806 X-ray sources (detected with $>0.05 \text{ ct s}^{-1}$, at least 15 X-ray counts and a minimum detection likelihood of 15), and the Faint Source Catalogue (FSC; Voges et al. 2000) which contains 105 924 sources (detected with at least six X-ray counts and a minimum detection likelihood of 6.5). We constructed a parent sample of

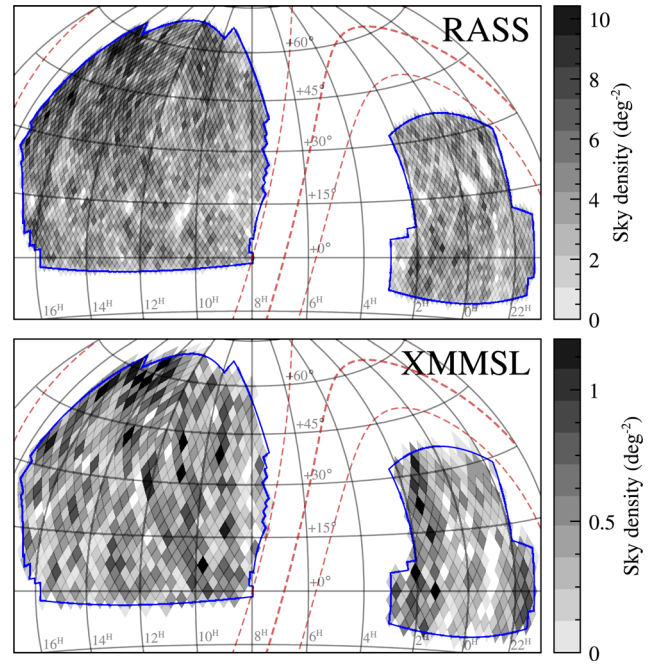


Figure 1. Sky distribution of the X-ray source samples considered in this paper. *Top panel:* RASS source density map gridded with a HEALPix pixelization ($\text{NSIDE} = 32$, $3.36 \text{ deg}^2 \text{ pix}^{-1}$) and displayed with an Equatorial Hammer–Aitoff projection. The density variations (and holes) are primarily due to the variations in effective exposure time and background count rate during the *ROSAT* survey. *Bottom panel:* Same for the XMMSL sources, but shown with a coarser pixel scale ($\text{NSIDE} = 16$, $13.4 \text{ deg}^2 \text{ pix}^{-1}$). Note the differences in the grey-scale ranges. The (solid blue) line indicates the perimeter of the BOSS imaging footprint. The Galactic plane is indicated (dashed red lines at $b = 0, \pm 15 \text{ deg}$).

32 408 X-ray sources from the concatenation of all RASS-BSC and RASS-FSC sources located within the BOSS imaging footprint. A small number (17) of RASS detections with undefined positional errors (likely to be detection algorithm artefacts) were then removed, leaving 32 391 sources. The median positional uncertainty (1σ radius, including a 6 arcsec systematic error) of the remaining RASS sources is 17 arcsec, and 95 per cent have uncertainties smaller than 34 arcsec. No attempt was made at this stage to filter the RASS catalogue any further, for example, by detection likelihood. We discuss the frequency and impact of spurious X-ray detections later in Section 4.3. The sky distribution of the RASS sample is shown in Fig. 1. The mean sky density of sources is 3.0 deg^{-2} , but their distribution is far from uniform, due primarily to the uneven sensitivity limit of the *ROSAT* all-sky survey.

We describe in Appendix A1 our method to estimate unabsorbed X-ray fluxes from the RASS count rates (i.e. correcting for the photoelectric absorption due to the Galactic column density in the direction of the source). The unabsorbed 0.1–2.4 keV flux distribution of the RASS sources is presented in Fig. 2. The distribution is strongly peaked with a median of $5.2 \times 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$, and with 86 per cent of the sources lying within ± 0.5 dex of this value.

We note that Boller et al. (2016) have recently presented the second revision of the *ROSAT* X-ray Survey catalogue (2RXS), using additional *ROSAT* survey data, improved source detection algorithms and improved source characterizations. Unfortunately the 2RXS catalogue was released after SPIDERS-AGN targets had been submitted for observation, and so we do not consider it further here. An associated work by several of us Salvato et al. (in

¹ http://www.sdss3.org/dr9/algorithms/boss_tiling.php

² <http://space.mit.edu/~molly/mangle/>

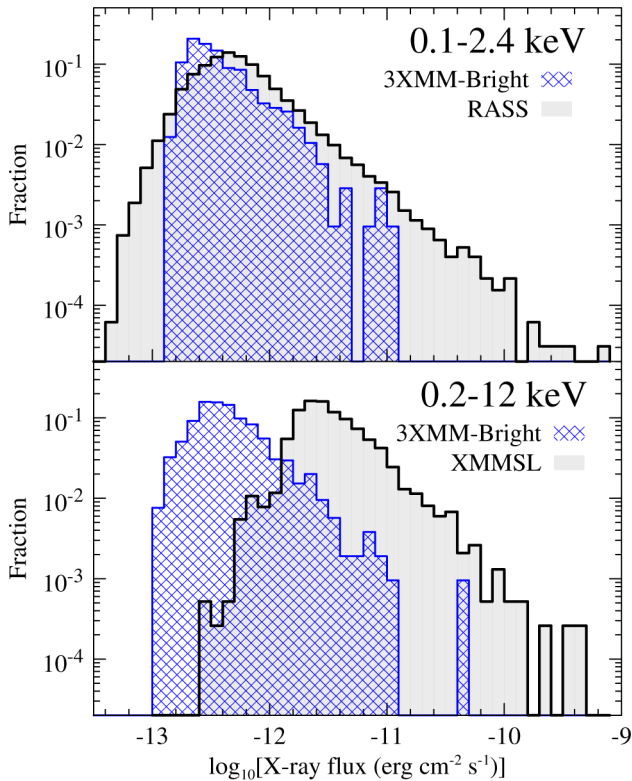


Figure 2. The X-ray flux distributions (corrected for Galactic absorption) of the RASS sample (0.1–2.4 keV band, upper panel) and the XMMSL sample (0.2–12 keV band, lower panel). The flux distributions of the 3XMM-Bright reference sample in the same bands are shown for comparison (see Section 2.6). The 3XMM-Bright fluxes have been converted from their native system and corrected for Galactic absorption, as described in Appendix A2.

preparation) will present *WISE* associations for 2RXS and XMMSL sources covering the entire extragalactic sky, using similar cross-matching techniques to those presented here.

2.2 *XMM-Newton* Slew survey catalogue (XMMSL)

As part of the normal operations of *XMM-Newton*, data are accumulated by the European Photon Imaging Camera pn detector during slews between pointed observations (Saxton et al. 2008). As of early 2014, the *XMM-Newton* slew observations have covered 65 per cent of the sky at least once.³ The XMMSL catalogue is built from sources detected in at least one of the following energy bands: soft (0.2–2 keV), hard (2–12 keV) and full (0.2–12 keV).

Our starting point is the ‘Clean’ version of the XMMSL catalogue release 1.6 (2014 February 26),⁴ from which we select the 4325 X-ray detections that fall inside the BOSS imaging footprint. However, some of the catalogue entries are multiple detections of the same X-ray source, whereas we require a list of unique sources. Therefore, where groups of XMMSL detections have the same UNIQUE_SRCNAME we choose only the best spatially constrained (i.e. having the smallest value of RADEC_ERR). Additionally, we choose only the best spatially constrained detection from any pairs of sources that lie within a 30 arcsec radius circle. These filtering

steps leave a catalogue of 3843 unique XMMSL sources.⁵ The sky distribution of the XMMSL sample is shown in Fig. 1. The rate of spurious X-ray detections in the ‘Clean’ version of the XMMSL catalogue is estimated to be 4 per cent in the full energy band.⁶ The median statistical positional uncertainty (1σ radius, excluding systematic errors) of the unique XMMSL sources is 4.0 arcsec, and 95 per cent have statistical uncertainties smaller than 15.4 arcsec.

The XMMSL catalogue includes flux estimates for each source, calculated from the observed count rates using a simple linear energy conversion factor (ECF) that assumes a $\Gamma = 1.7$ power-law spectrum absorbed by a fixed Galactic column density of $3 \times 10^{20} \text{ cm}^{-2}$. The ratio of unabsorbed to absorbed flux over the 0.2–12 keV band for this particular spectral model is 1.10 (calculated using WebPIMMS⁷), and we have therefore multiplied the catalogue fluxes by this factor to correct for Galactic absorption. Given the large statistical uncertainties on the fluxes of the XMMSL sources, and the relatively weak sensitivity of the *XMM-Newton* bandpass to moderate levels of Galactic absorption, we have applied a global correction to the XMMSL catalogue fluxes rather than attempting to correct for the Galactic column density local to each source. The unabsorbed 0.2–12 keV flux distribution of the XMMSL sources is presented in Fig. 2. The distribution is strongly peaked with a median of $3.0 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1}$, and with 89 per cent of the sources lying within ± 0.5 dex of this value.

We note that 1934/3843 (50 per cent) of the XMMSL sources have at least one RASS detection within 60 arcsec, and so the two X-ray samples used in our work are far from independent. However, since the XMMSL and RASS samples have very different characteristics and were selected in very different ways (telescopes, energy ranges, detection routines etc.), it is most convenient to treat these two samples separately.

2.3 AllWISE candidate counterpart catalogue

The *WISE* mission (Wright et al. 2010) carried out an all-sky survey in the 3.4, 4.6, 12 and 22 μm bands (we denote magnitudes measured in these bands as [W1], [W2], [W3] and [W4], respectively). The ‘AllWISE’ catalogue was released in late 2013 (Cutri 2013), and incorporates all data obtained by *WISE* in its original mission phase, including a second season of observations in the two shorter wavelength bands. The survey scan pattern results in inhomogeneous coverage (deeper at the Ecliptic poles than at the Ecliptic Equator). Throughout this work we only use the `w[1234]mpro` magnitude estimates from the AllWISE catalogue, which are appropriate for point-like sources. For 95 per cent of the extragalactic sky ($|b| > 15$ deg), the 5σ point-source magnitude limits in the [W1], [W2], [W3] and [W4] bands are at least as deep as 17.6, 16.1, 11.5 and 7.9 mag (Vega), respectively. We note that, for every source detected above 5σ in at least one band, the AllWISE catalogue contains a measurement of brightness in all four *WISE* photometric

⁵ We note here that an earlier version of our sample selection algorithm accidentally excluded all XMMSL sources with multiple X-ray detections. This affects ~ 10 per cent of the XMMSL sources lying in the first 4010 deg² of eBOSS plates (i.e. the tiling ‘chunks’ named eboss 1, 2, 3, 4, 5, 9 and 16). We corrected the target selection error before submitting SPIDERS targets for the remainder of the eBOSS sky. In the following sections, unless otherwise noted, all quoted statistics refer to the correctly filtered XMMSL catalogue. We indicate in the supplied catalogues (see Appendix E) which XMMSL sources were affected by this error.

⁶ <http://www.cosmos.esa.int/web/xmm-newton/xmmsl1d-ug>

⁷ <http://heasarc.gsfc.nasa.gov/cgi-bin/Tools/w3pimms/w3pimms.pl>

³ <http://www.cosmos.esa.int/web/xmm-newton/xsa#download>

⁴ http://nxs.esac.esa.int/catalogues/xmmsl1d_clean.fits.gz

bands, even where such measurements fall well below the nominal detection limits. At the faint limit of the AllWISE catalogue, the vast majority of sources have their highest SNR detection in the shortest wavelength *WISE* band.

We have restricted our search for AllWISE counterparts to within a 60 arcsec radius of the X-ray position of each member of the RASS sample described in Section 2.1. No down-selection is carried out on this AllWISE sample apart from removing the few duplicate entries that fall inside the search radii around two or more adjacent RASS sources. This procedure results in a catalogue of 450 409 unique potential AllWISE counterparts (a mean of 13.9 potential counterparts per RASS source).

An identical procedure was carried out to build a list of potential AllWISE counterparts to the XMMSL sample described in Section 2.2, resulting in a catalogue of 46 389 unique potential AllWISE counterparts (a mean of 13.4 potential counterparts per XMMSL source). Note that although the positions of XMMSL catalogue sources are typically much better determined than those of the RASS sources, a small fraction do have large (>30 arcsec) positional uncertainties. We have therefore used a relatively large search radius of 60 arcsec when compiling this initial list of potential AllWISE counterparts.

2.4 SDSS photometric counterpart catalogue

As part of the first three phases of the SDSS, a wide-area multi-band (*ugriz*) imaging survey was carried out using the 2.5-m telescope at the Apache Point Observatory, New Mexico (Gunn et al. 2006; Aihara et al. 2011). The most recent photometric catalogue derived from these imaging data was released as part of SDSS-DR13 (SDSS Collaboration et al. 2016). The nominal 95 per cent completeness limits of the SDSS imaging are 22.0, 22.2, 22.2, 21.3 and 20.5 mag in the *ugriz* bands, respectively.⁸ Starting with the SDSS-DR13 catalogue (specifically we considered primary objects from the ‘Datasweep’ files⁹), we down-selected a sample of potential optical counterparts to X-ray sources by searching within 65 arcsec of the X-ray positions of all members of the RASS and XMMSL samples described in Sections 2.1 and 2.2. The mean sky density of SDSS photometric objects lying near the X-ray sources in our sample is $3.8 \times 10^4 \text{ deg}^{-2}$. Note that we adopt a slightly larger search radius here compared to that used to select potential AllWISE counterparts. This is in order to prevent the effects of astrometric uncertainties scattering potential SDSS counterparts to AllWISE sources outside our considered search area.

2.5 SDSS spectroscopic counterpart catalogue

Since its first light in 1998, and up to its 12th data release (DR12), the SDSS project has obtained over five million astronomical spectra, with various spectrographs and within a number of different survey programmes (Alam et al. 2015).

For practical reasons, we have used two versions of the spectroscopic catalogues associated with SDSS-DR12: (i) an early prototype version of the DR12 spectroscopic catalogue and (ii) the full official DR12 spectroscopic catalogue. At the time when we created the SPIDERS target catalogues (2014 March–April), the final version of the BOSS spectroscopic catalogue was not yet completely finalized (and did not include the spectra collected as part of

SEQUELS, a few BOSS main-survey plates, and several plates associated with ancillary programmes). However, the combination of a pre-release version of the BOSS spectroscopic catalogue, joined with the official SDSS-DR8 spectroscopic catalogue included more than 95 per cent of the spectra that would go on to form the final official DR12 spectroscopic catalogue. To form a ‘proto-DR12’ catalogue, we first filtered the DR8 and pre-release BOSS catalogues to include only spectra with `SPECPRIMARY=1` (i.e. for each object with multiple spectroscopic observations, we only considered the best quality one, see Alam et al. 2015 for a description of how this choice is made) and have `ZWARNING=0` (i.e. we discarded spectra for which the BOSS analysis pipeline identified potential issues, including those spectra for which the fitting routine could not determine a reliable redshift). The combination of the filtered DR8 and pre-release BOSS catalogues (containing 1674 844 and 1857 100 sources, respectively) were considered when we decided which potential SPIDERS targets already had good-quality SDSS spectra available (and hence should be excluded from the list of potential targets in SDSS-IV). Use of this incomplete reference catalogue for target selection is acceptable, as its only impact is a slight decrease in observing efficiency, due to a small fraction (estimated to be less than 1 per cent) of SPIDERS AGN targets with existing BOSS spectra being re-observed during the eBOSS/TDSS/SPIDERS project.

However, as part of this paper we also release samples of RASS and XMMSL sources with existing SDSS spectra. Therefore, for the sake of clarity and repeatability, and to allow reference to fully documented data products, we prefer in these cases to base our samples on the officially released SDSS-DR12 spectroscopic sample, which contains all spectra collected by the SDSS and BOSS optical spectrographs (Smee et al. 2013) in the MJD range [51578:56837].¹⁰ In cases where multiple SDSS-DR12 spectra are available for a photometric object, we consider only the ‘best’ spectrum (i.e. the one flagged with `SPECPRIMARY=1`). Since we visually inspect all spectra matched to our X-ray source samples (see Section 3.5), we have not filtered the DR12 spectra on the basis of pipeline redshift warning flags.

2.6 3XMM-Bright reference catalogue

Within the Bayesian cross-matching framework (described in Section 3), the complex task of identifying reliable counterparts of RASS (and, to a lesser extent, XMMSL) sources benefits dramatically from the availability of a sub-sample of well-characterized X-ray sources at similar flux levels, with good enough positional accuracy to make the cross-matching exercise non-problematic. To this end, we have exploited the 3XMM serendipitous source catalogue (DR4),¹¹ which is derived from over 7427 *XMM-Newton* pointings performed up to 2012 December, covers a total of 794 deg², and contains 372 728 unique detections (for a description of the 3XMM programme, and a more recent revision of the catalogue, see Rosen et al. 2016). The median flux of 3XMM sources in the soft (0.2–2 keV) energy band is $\sim 6 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$, around two orders of magnitude fainter than the flux limit of the RASS. However, the 3XMM catalogue does include a significant tail of bright sources, which overlaps with the flux ranges spanned by the RASS and XMMSL samples. Therefore, we have exploited the X-ray bright end of the 3XMM catalogue to provide a reference catalogue of well-understood X-ray bright sources over the BOSS

⁸ <http://www.sdss.org/dr13/scope>

⁹ <http://www.sdss.org/dr13/imaging/catalogs/>

¹⁰ <http://data.sdss3.org/sas/dr12/sdss/spectro/redux/specObj-dr12.fits>

¹¹ http://nxs.a.esac.esa.int/catalogues/3XMM_DR4cat_slim_v1.0.fits.gz

imaging footprint. This (near) ‘truth’ sample is used to derive the priors that inform the Bayesian cross-matching process (section 3).

From the 3XMM parent sample, we select sources that meet all of the following criteria: (i) lie inside the BOSS imaging footprint, (ii) have X-ray fluxes $\geq 10^{-13}$ erg cm $^{-2}$ s $^{-1}$ in the 0.2–2 keV energy band (where $F_{0.2-2\text{keV}}$ is calculated as the unweighted sum over the individual flux measurements in the standard 3XMM energy bands: 0.2–0.5, 0.5–1 and 1–2 keV), (iii) have a very high detection likelihood ($SC_DET_ML > 50$), (iv) have low likelihood of being extended in the X-ray ($SC_EXT_ML < 8$), (v) have no warning flags set ($SC_SUM_FLAG = 0$), (vi) are not associated with observations of Solar system objects and (vii) are not in *XMM-Newton* observations associated with poor astrometry (this final cut was applied retrospectively).¹²

These cuts result in a high quality reference sample of 1049 bright, well measured, point-like X-ray sources (which we will hereafter call the 3XMM-Bright sample). The mean and 95th percentile of the 1σ error radii for the 3XMM-Bright sources are 0.6 and 1.4 arcsec, respectively, making unambiguous cross-correlation with multi-wavelength catalogues relatively simple (see Section 3.2 below). Fig. 2 demonstrates that the 3XMM-Bright sources have a similar range of X-ray fluxes to sources in the RASS sample, but are approximately 10 times fainter than sources in the XMMSL sample.¹³

3 SELECTING COUNTERPARTS TO X-RAY SOURCES

We summarize in this section the Bayesian cross-matching method used to associate RASS and XMMSL sources with optical/IR counterparts. This method expands upon the techniques introduced by Budavári & Szalay (2008) and Rots & Budavári (2011). We refer the reader to Salvato et al. (in preparation) for a full and generalized description of the Bayesian cross-matching method, a description of ‘NWAY’, the PYTHON implementation of the algorithm that we have used within this work, and a comparison to the likelihood ratio technique (Sutherland & Saunders 1992). Given that the RASS and XMMSL surveys are far from the confusion limit, we make the simplifying assumption that each X-ray detection is dominated by a single X-ray source having up to one counterpart at longer wavelengths. We briefly reiterate below (Section 3.1) those formulae that are pertinent to the case considered here, that is, where we just wish to find the best counterpart to each of a sample of X-ray sources from a catalogue of potential counterparts.

¹² Our first version of the 3XMM-Bright catalogue included an additional eight sources, all of which had no AllWISE counterparts within 3 arcsec. These eight sources were located close to each other within the ‘XMM-XXL North’ field and were detected within two consecutive *XMM-Newton* mosaic-mode observations (targets ‘XXLn074’ and ‘XXLn094’). We have visually inspected independently processed images derived from these *XMM-Newton* data sets, and find no bright X-ray sources at the catalogued positions of the eight 3XMM sources (N. Clerc, private communication). Furthermore, the 3XMM POSCOROK flag was set to FALSE for 7/8 of these detections, indicating that the pipeline had been unable to correct the X-ray astrometry against external optical/IR catalogues. Therefore, we removed these objects from the 3XMM-Bright sample.

¹³ See Appendix A2 for details of the method used to estimate the unabsorbed fluxes of the 3XMM-Bright sample in the RASS and XMMSL energy bands (0.1–2.4 and 0.2–12 keV, respectively).

3.1 Bayesian posterior probability of associations

The posterior probability for a cross-match between an X-ray source, i , and a potential counterpart j , is given by,

$$P_{\text{post}} = \left[1 + \Pi(\mathbf{x}_j) B_{ij} \frac{1 - P}{P} \right]^{-1}, \quad (1)$$

where $\Pi(\mathbf{x}_j)$ is a prior (or ‘bias’) term dependent on the location of counterpart j in some parameter space \mathbf{x} (see below), B_{ij} is the Bayes factor for the geometric association, given by,

$$B_{ij} = \frac{2}{\sigma_i^2 + \sigma_j^2} \exp \left[-\frac{\psi_{ij}^2}{2(\sigma_i^2 + \sigma_j^2)} \right], \quad (2)$$

where ψ_{ij} is the angular separation (in radians), and σ_i , σ_j are the respective positional uncertainties for sources i and j . P is a normalizing factor that takes account of the mean sky density of potential counterparts (ρ_s , units deg $^{-2}$), and the expected fraction of X-ray sources that have a true counterpart, η_x ,

$$P = \frac{\eta_x}{4\pi(180/\pi)^2 \rho_s} = 2.4241 \times 10^{-5} \frac{\eta_x}{\rho_s}. \quad (3)$$

The value of the prior, $\Pi(\mathbf{x})$, at some location in N -dimensional measurement space, \mathbf{x} , is given by the ratio $f_{\text{xray}}(\mathbf{x})/f_{\text{all}}(\mathbf{x})$, where $f_{\text{xray}}(\mathbf{x})$ is the probability density function of true counterparts to X-ray sources (normalized such that $\int f_{\text{xray}}(\mathbf{x}) d\mathbf{x} = 1$ over the parameter range of interest), and $f_{\text{all}}(\mathbf{x})$ is the probability density function of all potential counterparts, similarly normalized. In general, $\Pi(\mathbf{x})$ can be used to encode some or all of our prior knowledge of the distribution of the measurable properties of true counterparts to X-ray sources (e.g. magnitudes, colours). Setting $\Pi \equiv 1$ reduces equation (1) to the standard unweighted form (e.g. Rots & Budavári 2011).

In practice, we will not know $\Pi(\mathbf{x})$ exactly (as to do so would require that we had already measured the multi-wavelength properties of the X-ray sample), but we can estimate it, $\Pi'(\mathbf{x}) \sim \Pi(\mathbf{x})$, using a training sample of well-measured X-ray sources which we expect to be representative of objects in our main X-ray sample. In Section 3.2, we describe how we have computed the $\Pi'(\mathbf{x})$ used to select counterparts to RASS and XMMSL sources.

We note that the P_{post} statistic measures the probability of association of individual pairs of sources independently of the presence of other possible pairings. So, for example, although P_{post} cannot tell us the probability that at least one of the possible counterparts to an X-ray source is the correct one, it can be used to choose the most probably of these counterparts. A more complete treatment of such cases is implemented within the soon to be released version of the *NWAY* code (Salvato et al., in preparation). See also the treatment of multiple potential counterparts by Pineau et al. (2011).

3.2 Bayesian priors derived from a bright X-ray reference sample

We investigated the properties of the 3XMM-Bright sample searching for simple combinations of parameters in which the X-ray sources stand out clearly from the general field population. This is complicated by the heterogeneous nature of the bright X-ray source population. For example, Zickgraf et al. (2003) found that the optically bright end of the RASS-BSC catalogue is associated with a mix of AGN, galaxies, galaxy clusters, M stars, white dwarfs, K stars, F-G stars and cataclysmic variables.

High luminosity AGN (i.e. QSOs) typically outshine their host galaxies in the optical–mid-IR bands, but for lower luminosity AGN

Table 1. Summary of input catalogues after application of preparation and filtering steps described in the text. Note that the RASS, XMMSL and SDSS-spectroscopy catalogues have been filtered to include only objects lying inside the 10 788.3 deg² BOSS imaging footprint. The AllWISE and SDSS imaging catalogues have been filtered to only include objects lying within the indicated radii of RASS and XMMSL sources.

Catalogue name	Waveband	Number of objects	Mean density (deg ⁻²)	Notes	Section
RASS	0.1–2.4 keV	32 408	3.0	‘1RXS’. Combined BSC+FSC.	2.1
XMMSL	0.2–12 keV	3843	0.36	Release 1.6, 2014 February 26. Unique sources.	2.2
AllWISE	3.4–22 μm	450 600 51 725	1.6 × 10 ⁴	<60 arcsec from RASS sources <60 arcsec from XMMSL sources	2.3
SDSS photometry (DR13)	<i>ugriz</i>	1081 654 126 606	3.8 × 10 ⁴	<65 arcsec from RASS sources <65 arcsec from XMMSL sources	2.4
SDSS spectroscopy (proto-DR12)	380–920 nm (SDSS) 365–1040 nm (BOSS)	3316 373	310	All SDSS-I, II, most SDSS-III SPECPRIMARY = 1, ZWARNING = 0	2.5
SDSS spectroscopy (official-DR12)	380–920 nm (SDSS) 365–1040 nm (BOSS)	3658 581	340	All SDSS-I, II, III in BOSS footprint SPECPRIMARY = 1	2.5

(Seyferts), the host-galaxy emission may match or exceed the AGN emission. In addition, extinction along the line of sight could also mask any AGN signature in the UV/optical bands, and even into the near-IR bands as the extinction increases. However, mid-IR colours of bright AGN are less susceptible to these effects, and promise to provide a more universal tool for AGN identification.

The release of the AllWISE catalogue in 2013 November (Cutri 2013) spurred us to examine whether mid-IR imaging information could assist in optical counterpart selection for the RASS and XMMSL sources. The vast majority of the X-ray bright AGN detected in RASS and XMMSL are expected to be bright and red in the two shortest wavelength *WISE* channels (i.e. 3.4 and 4.6 μm). Therefore, RASS-selected AGN should stand out from field stars and galaxies in the mid-IR. In addition, counterparts to the non-AGN ‘contaminants’ in the X-ray sample (i.e. Galactic stars and bright nearby galaxies) are also expected to stand out (they are likely to be significantly brighter in the [W1] and [W2] bands than the bulk of the AllWISE field population). Furthermore, the sky density of field sources in the AllWISE catalogue is less than half that in the SDSS imaging catalogue (see Table 1), which dramatically reduces the rate of false identification w.r.t. a purely optical-based counterpart selection scheme. The useful dynamic range of the AllWISE catalogue is somewhat larger than that of the SDSS catalogue (as expected, given that *WISE* was designed as a true-all sky surveyor, whereas the SDSS imager was primarily designed to target faint galaxies and QSOs). In particular, the treatment for X-ray sources associated with very bright stars is simpler with AllWISE; point sources start to saturate at [W2] < 6.7 mag in *WISE* (Cutri 2013), but saturate at $r < 14.1$ mag in SDSS imaging (Gunn et al. 1998).

We start with the 3XMM-Bright catalogue described in Section 2.6, and search for counterparts in the AllWISE catalogue, using a simple cone search. We find that 1000/1049 (95.2 per cent) of the 3XMM-Bright sources had exactly one AllWISE counterpart within 3 arcsec of the X-ray position. We denote these 1000 matches as the ‘3XMM-Bright+AllWISE’ sample. The mean sky density of AllWISE field sources local to the 3XMM-Bright+AllWISE sources is 15 300 objects deg⁻² (estimated by measuring the density of AllWISE sources within 1 arcmin radius control regions placed at a distance 6 arcmin from each X-ray source). Therefore, we would naïvely expect 0.033 field sources to fall within the 3 arcsec radius circle we searched within for each 3XMM-Bright source. However, the typical AllWISE counterparts to the 3XMM-Bright sources are much brighter than typical field sources (the median [W2] magnitude of the 3XMM-Bright+AllWISE matches is 13.2 mag compared

to 16.6 mag for the AllWISE field population), and so an average AllWISE field source would be overwhelmed by the average AllWISE counterpart to a 3XMM-Bright source (see also section 5 of Broos et al. 2011). We have empirically estimated the rate at which random associations are both close enough and bright enough to contaminate the 3XMM-Bright+AllWISE sample: for each 3XMM-Bright+AllWISE source we measured the probability of a field AllWISE source brighter (in [W2]) than the true AllWISE counterpart to lie within any randomly placed 3 arcsec circle. We tested 100 randomized locations per 3XMM-Bright+AllWISE source for a total of 10⁵ samples, and find an almost negligible overall contamination rate of just 0.12 per cent. We note that using a matching radius larger than 3 arcsec would result in a higher completeness (fraction of X-ray sources with AllWISE counterparts), but would lead to more contamination from chance aligned field sources, and would also increase the number of cases where there is more than one potential AllWISE counterpart per 3XMM-Bright source (making unambiguous associations difficult). In Appendix B, we discuss in more detail the nature of the 49/1049 3XMM-Bright sources which do not have AllWISE counterparts within 3 arcsec, and conclude that only 3/49 are likely to be associated with genuinely mid-IR faint sources, the remaining cases are the result of various problems with either *XMM-Newton* or AllWISE photometry, e.g. blending/confusion, presence of saturated bright stars, and poor astrometry at the edge of the *XMM-Newton* field of view. Therefore, we are confident that our reference sample is not artificially excluding a significant part of the bright X-ray population.

We have used the following recipe to compute a pixelized map of the prior, $\Pi(x)$, in the [W2],[W1 – W2] colour–magnitude space, using as input the training sample (3XMM-Bright+AllWISE sources) and a ‘field’ sample (consisting of AllWISE objects lying within 1 arcmin radius control regions placed at a distance of 6 arcmin from each of the sources in the 3XMM-Bright+AllWISE sample): (i) compute 2D histograms from each of the training and field samples, covering the parameter interval $0 \leq [W2] \leq 25$, $-5 \leq [W1 - W2] \leq 5$, with pixel steps of 0.25 mag in [W2] and 0.1 mag in [W1 – W2], (ii) smooth each 2D histogram with a 2D Gaussian kernel having $\sigma_{[W2]} = 0.5$ and $\sigma_{[W1-W2]} = 0.2$, (iii) normalize each 2D histogram such that the sum of the pixel values over the considered range ($0 \leq [W2] \leq 25$, $-5 \leq [W1 - W2] \leq 5$) equals unity, (iv) threshold each 2D histogram such that no pixel has a value smaller than one over the total number of objects in the field sample. The chosen smoothing kernel is a fair compromise between the desire to reduce the shot noise of the training sample (pushing to larger

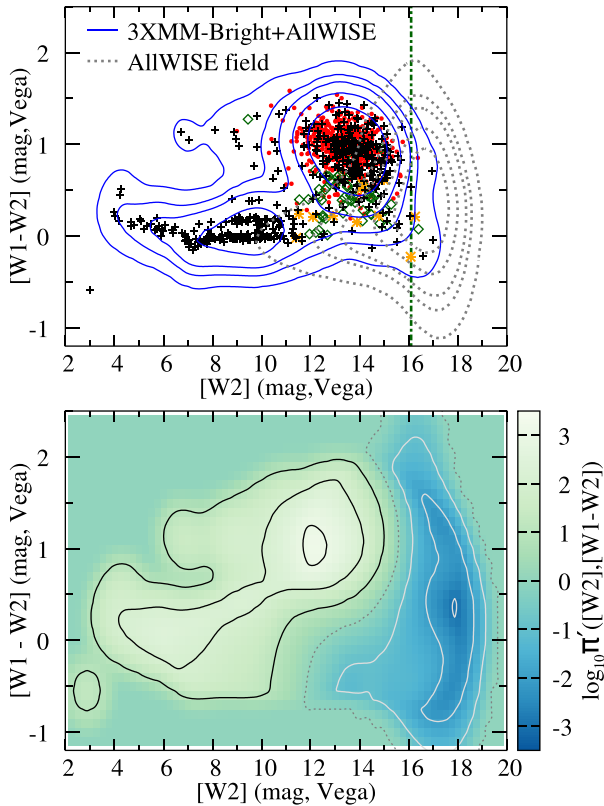


Figure 3. *Top panel:* Distribution of 3XMM-Bright+AllWISE and AllWISE field sources in AllWISE colour-magnitude space. 3XMM-Bright+AllWISE sources with SDSS-DR12 spectroscopic classifications are shown with red filled points (CLASS=QSO), green open diamonds (CLASS=GALAXY) and orange asterisks (CLASS=STAR). 3XMM-Bright+AllWISE sources lacking spectroscopic classifications are shown with small black + signs. Contours contain 99, 95, 90, 75 and 50 per cent of the smoothed density of each catalogue. The vertical line (green dot-dashed) shows the nominal 5σ detection limit for the [W2] band. *Bottom panel:* Map of the prior $\Pi'([W2], [W1 - W2])$, illustrating the weighting given to potential counterparts as a function of location in [W2], [W1 - W2] colour-magnitude space (see Section 3.1). Potential counterparts lying in locations with light (green) colours are up-weighted, and those lying in regions having darker (blue) colours are down-weighted. Contours are drawn at $\log_{10}\Pi'([W2], [W1 - W2]) = 3, 2, 1, 0, -1, -2, -3$.

kernels), versus the desire to retain the separation between the distributions of the X-ray reference sample and of field sources (pushing to smaller kernels). The parameter interval over which we have defined the prior contains more than $1-10^{-6}$ of the smoothed PDF for the 3XMM-Bright+AllWISE sample. The distribution of the 3XMM-Bright+AllWISE sample in AllWISE colour-magnitude space is presented in the upper panel of Fig. 3, illustrating clearly how these bright X-ray selected sources stand out from the field population in this measurement space. The ratio of the density distribution of our training (the 3XMM-Bright+AllWISE sources) and field samples constitutes our Bayesian prior $\Pi'([W1], [W1 - W2])$, which is also shown in the bottom panel of Fig. 3, illustrating which parts of colour-magnitude space are up- and down-weighted by the cross-matching routine.

In order to get a sense of the mix of sources that make up the 3XMM-Bright+AllWISE sample, we have cross-matched to the SDSS-DR12 spectroscopic catalogue, searching within a 2 arcsec radius from the AllWISE position. Of the 557 3XMM-Bright+AllWISE sources with SDSS spectra and reliable BOSS pipeline automated classifications, 490 are classified as QSO, 56 as

GALAXY (13 of which have a sub-class indicating AGN activity) and 11 as STAR. The selection function for the sub-sample of 3XMM-Bright+AllWISE sources with spectra is difficult to determine, but is certainly highly incomplete at the bright and faint ends. Therefore, we should not make quantitative predictions from these fractions. However, we can at least say with some confidence that for the range of X-ray fluxes probed by the RASS and XMMSL samples, the spectroscopic samples collected by SPIDERS will be dominated by QSOs, but with significant minorities of both normal galaxies and Galactic stars. Note that of the 3XMM-Bright+AllWISE sources with no spectral classification in SDSS-DR12, many lie in the part of colour-magnitude space expected to be occupied by Galactic stars and very bright nearby galaxies ($[W2] < 11$, $[W1 - W2] < 0.3$, see upper panel of Fig. 3).

3.3 Selection of SPIDERS_RASS_AGN targets

Armed with the priors described in the previous section, we can now proceed to the identification of the RASS sources. We start with the AllWISE catalogue described in Section 2.3. All except 11/32 391 (0.03 per cent) of the RASS sources in our sample have at least one potential AllWISE counterpart lying within 1 arcmin. For each RASS source, we used `nway` (version 1.0) to calculate the posterior probability of it being associated with each of the possible AllWISE counterparts within 1 arcmin, taking account of the X-ray position and its uncertainty, the AllWISE position, the number density of the sources, the [W2] magnitude and the [W1 - W2] colour. At the time of generation of the SPIDERS_RASS_AGN target catalogue, we made a decision to adopt a fixed radial positional uncertainty of 0.3 arcsec (1σ) for all AllWISE sources. In retrospect this was an unnecessary and sub-optimal choice, and it would have been much more correct to use the positional uncertainties tabulated in the AllWISE catalogue. However, we do not expect that this has had a significant effect on our target selection given that the X-ray positional uncertainties are almost always much larger than those of AllWISE sources, and so dominate the denominators of equation (2).

In an ideal world, we would obtain spectra for all potential counterparts for each X-ray source above some minimum P_{post} . However, the combination of the BOSS spectrograph plug-plate fibre-collision constraint (minimum fibre separation > 62 arcsec), the single-pass survey strategy over the main eBOSS/TDSS/SPIDERS footprint (Dawson et al. 2016), and the limited fibre-budget allocated to the SPIDERS-AGN programme, mean that we only attempt to target a single potential counterpart per RASS source. Therefore, for each RASS source, only the ‘best’ AllWISE counterpart (i.e. the one having the highest posterior probability) was considered in the following steps.

We find that for 30 855/32 391 (95.3 per cent) of RASS sources we have a best matching AllWISE counterpart with posterior probability $P_{\text{post}} \geq 0.01$ (see Section 4.2 for a discussion of the P_{post} threshold below which the sample becomes significantly contaminated by interlopers). RASS sources with best AllWISE counterparts having $P_{\text{post}} < 0.01$ were not considered further. The best matching AllWISE counterparts have a median P_{post} of 0.86 compared to 0.025 for the second best counterparts. In 82 per cent of cases, the P_{post} for the best match is more than twice that for the second best match indicating a very secure choice. In a small fraction of cases (4 per cent), the second best match has a P_{post} within 10 per cent of the best match, and for such cases we cannot differentiate significantly between the best and second best AllWISE counterparts. See Section 4.3 for a discussion of how the fraction of RASS sources lacking any $P_{\text{post}} \geq 0.01$ AllWISE counterparts depends on X-ray detection

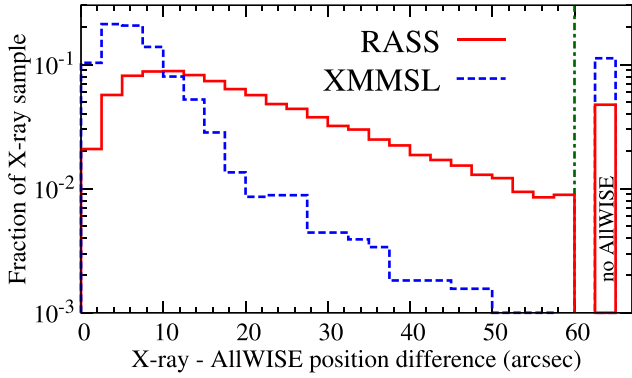


Figure 4. Distribution of separations between X-ray positions and AllWISE positions for the RASS+AllWISE (red solid line) and XMMSL+AllWISE (blue dashed lines) samples. X-ray sources without any valid AllWISE counterpart (i.e. no counterparts with $P_{\text{post}} \geq 0.01$) are represented in the rightmost bin. The vertical line (green dot-dashed) shows the maximum search radius considered.

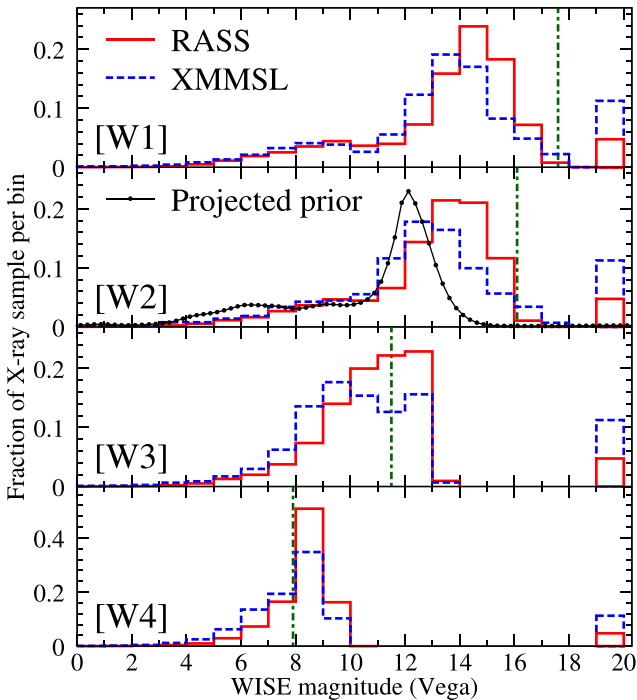


Figure 5. Distribution of mid-IR brightness in the *WISE* ([W1], [W2], [W3] and [W4]) bands for the RASS (red solid line) and XMMSL (blue dashed lines) samples, as measured with AllWISE point-source photometry. X-ray sources with no AllWISE counterpart are represented in the faintest magnitude bin. Vertical lines (green dot-dashed) show the nominal 5σ point source detection limits (valid for >95 per cent of the extragalactic sky) in each *WISE* band. In the second ([W2]) panel, we also show (black line with points) an arbitrarily scaled projection of $\Pi'([W1], [W1 - W2])$ (averaged over the $[W1 - W2]$ axis).

likelihood. The distribution of position differences between the RASS sources and their best matching AllWISE counterparts is shown in Fig. 4. The magnitude distributions of the best matching AllWISE counterparts are shown in Fig. 5.

The best AllWISE counterpart for each RASS source was then matched to an optical counterpart from the SDSS-DR13 photometric catalogue. The brightest (as measured by modelMag_r) object within 1.5 arcsec of the AllWISE position is chosen as the

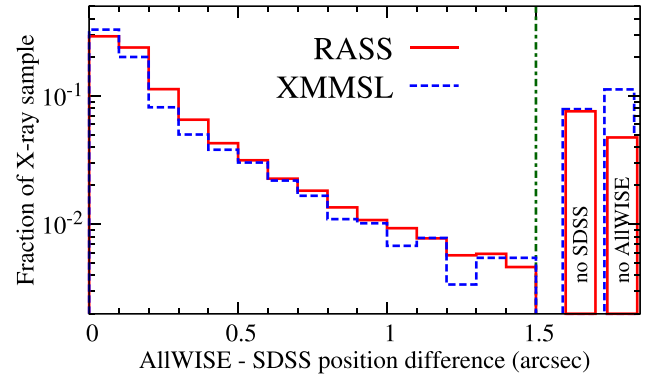


Figure 6. Distribution of separations between AllWISE positions and SDSS photometric positions for the RASS and XMMSL sources with AllWISE counterparts. For comparison, the fraction of each X-ray sample without a valid AllWISE counterpart and the fraction of X-ray+AllWISE sources without an optical counterpart are represented in the two rightmost bins. The vertical line (green dot-dashed) shows the maximum search radius considered.

most probable optical association. At the time of generating the SPIDERS targeting catalogues (early 2014), this choice seemed appropriate. However, later analysis suggests that better choices could have been made. In Section 4.4, we discuss how our simplistic approach to associating AllWISE sources with optical counterparts has impacted the target selection. We will make improvements to this step in future SPIDERS-AGN studies, taking into account the positional uncertainties and sky densities of the mid-IR and optical populations. We find that 28 515/30 855 (92.4 per cent) of the RASS+AllWISE sources have at least one SDSS-DR13 photometric optical counterpart within 1.5 arcsec of the AllWISE position. The distribution of AllWISE–SDSS positional offsets is shown in Fig. 6, and the magnitude distribution is shown in Fig. 7.

We then further down-select the 28 515 remaining potential RASS+AllWISE+SDSS targets to get to our final list of SPIDERS objects to be forwarded to the eBOSS tiling team. First, we removed previously spectroscopically identified sources by matching against a list of good-quality primary SDSS-*proto*-DR12 spectra (see Section 2.5). There are 11 643 RASS+AllWISE+SDSS sources with at least one reliable SDSS-*proto*-DR12 spectrum (within 1 arcsec of the optical photometric catalogue position), and thus, which are not considered for targeting. The majority of these (77 per cent) have a pipeline classification of $\text{CLASS}=\text{QSO}$. The properties of the spectroscopically identified sources are discussed in more detail in Section 5.1.

We then remove RASS+AllWISE+SDSS objects brighter than the nominal magnitude limit of eBOSS (i.e. $\text{fiber2Mag}_i < 17.0$, where fiber2Mag_i is a measure of the expected flux from the object that would be enclosed within a 2 arcsec diameter fibre under average seeing conditions). This bright source cut removes 7092 objects. To improve our robustness against imperfect modelling of very bright (and possibly saturated) objects, we also remove a small number (319) of objects which escape the fiber2Mag_i cut but which have $\text{modelMag}_i < 16.0$. These bright limits are motivated by the desire to avoid the on-chip spectra of very bright stars overwhelming the spectra of their neighbours, which can be many magnitudes fainter. We also remove the 283 very faint targets ($\text{fiber2Mag}_i > 22.5$, for which we do not expect to be able to obtain useful spectra with a 2.5 m-class telescope), and the 150

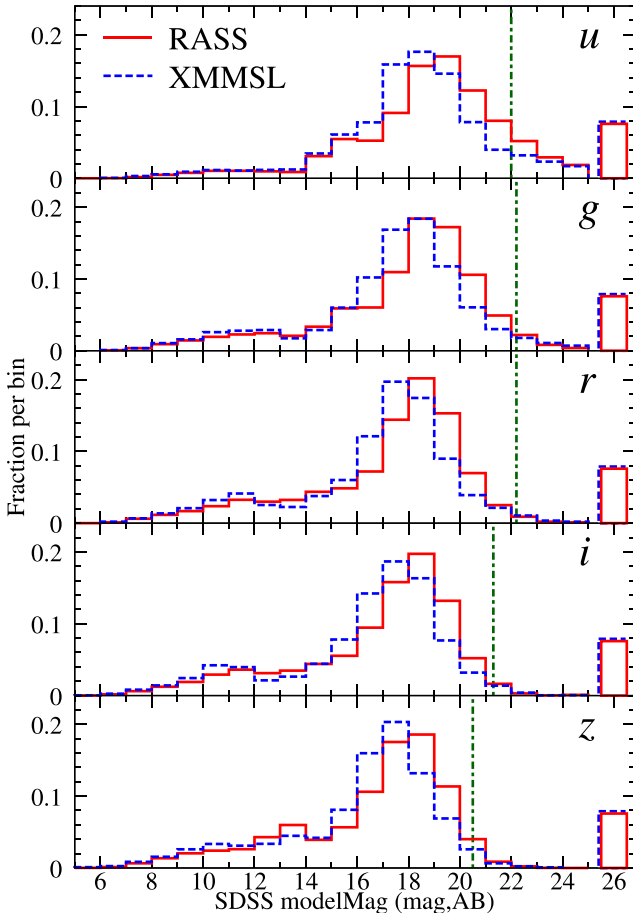


Figure 7. Distribution of optical brightness in the SDSS *ugriz* filters for the RASS and XMMSL sources with AllWISE counterparts, as measured by the SDSS-DR13 photometric catalogue *modelMag* parameter. X-ray+AllWISE sources without an optical counterpart are represented in the faintest magnitude bin. Vertical lines (green dot-dashed) show the nominal 95 per cent completeness limits in each SDSS filter.

targets which fall within the BOSS-DR10 bright star mask.¹⁴ The final SPIDERS_RASS_AGN target list contains 9028 candidate targets over the full BOSS targeting footprint (area 10 778 deg²), a density of 0.84 targets deg⁻².

The matching and filtering steps described above are summarized in a flow diagram, see Fig. 8. The format of the catalogue of SPIDERS_RASS_AGN targets is described in Appendix E3. In Fig. 9, we show the sky distribution of SPIDERS_RASS_AGN targets over the sky. Note that the SGC contains a higher density of targets than the NGC because only a small part of this footprint was targeted during SDSS-I/II, and so a smaller fraction of the RASS sources have existing identifications. In Figs 10, 11 and 12, we show the distributions of the RASS sample in X-ray flux, *r*-band model magnitude and [W2] magnitude at various stages during the cross-matching and down-selection steps described above.

3.4 Selection of SPIDERS_XMMSL_AGN targets

The process of selecting counterparts to the unique XMMSL sources in the BOSS footprint followed closely that carried out for RASS

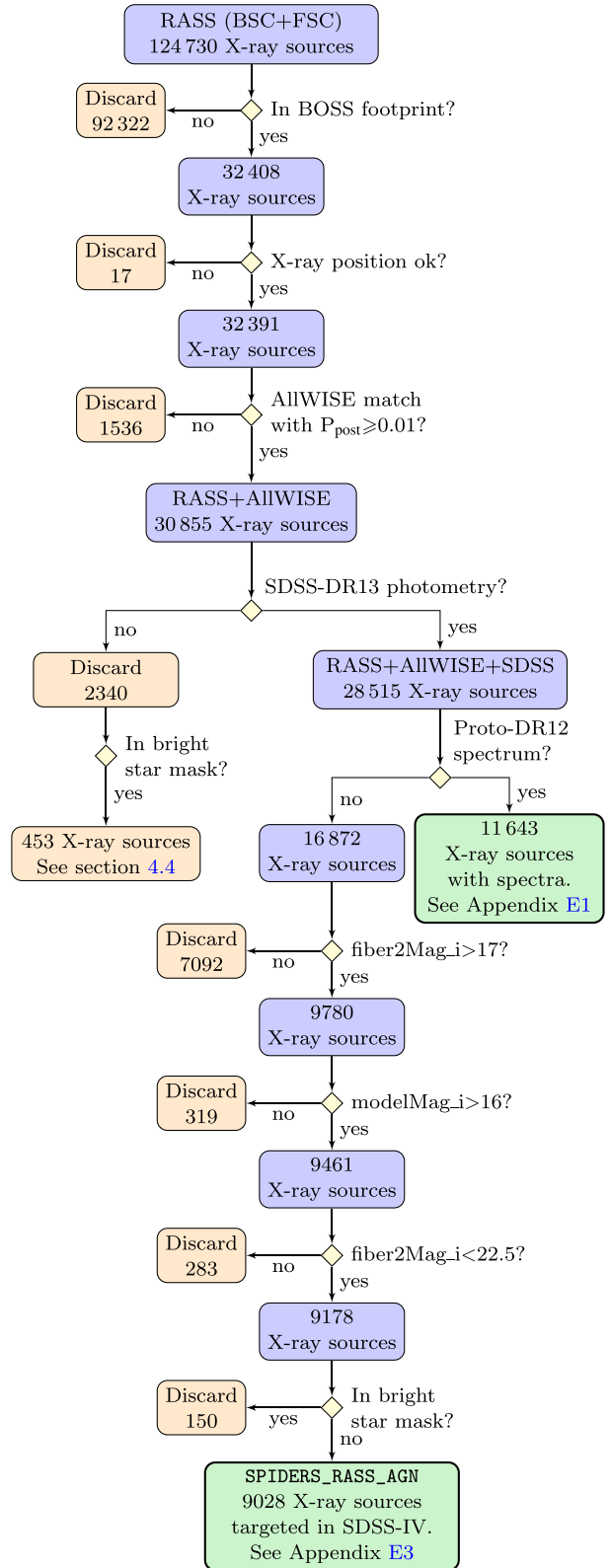


Figure 8. A schematic representation of the decision tree which leads to the selection of SPIDERS_RASS_AGN targets. See Section 3.3 for details of the selection steps.

¹⁴ http://data.sdss3.org/sas/dr10/boos/lss/reject_mask/bright_star_mask_pix.ply

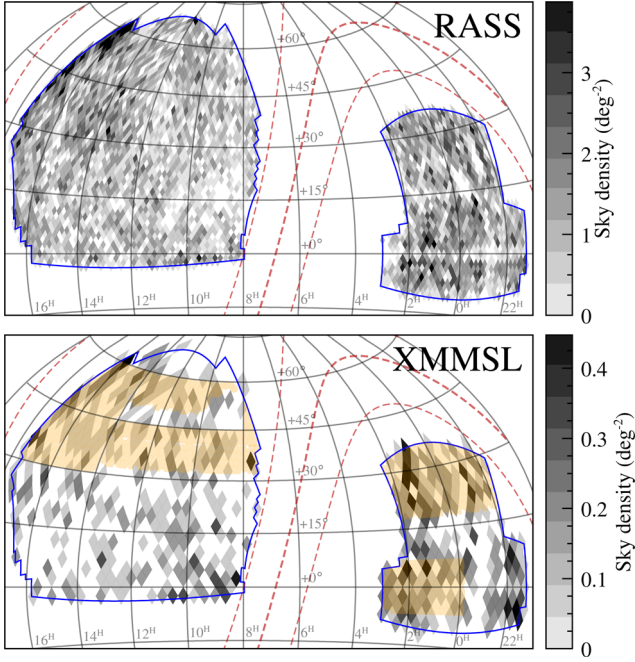


Figure 9. *Upper panel:* Sky distribution of SPIDERS_RASS_AGN sources put forward for targeting in eBOSS, displayed with a HEALPix pixelization (NSIDE = 32, 1 pixel = 3.36 deg²). *Lower panel:* Same for SPIDERS_XMMSL_AGN sources put forward for targeting in eBOSS, but shown on a coarser pixel scale (NSIDE = 16, 1 pixel = 13.4 deg²). The (solid blue) line indicates the perimeter of the BOSS imaging footprint. The Galactic plane is indicated (dashed red lines at $b = 0, \pm 15$ deg). In the lower panel, the (orange) shading indicates the 4010 deg² area over which XMMSL sources with multiple X-ray detections were accidentally excluded during the targeting process (see Section 3.4). The (orange) shaded region is also the sky area for which we compile statistics for the overlap of SPIDERS-AGN targets with other eBOSS/TDSS/SPIDERS target classes (see Appendix C).

sources (see Section 3.3). For each XMMSL source (all 3843 of which have at least one potential AllWISE counterpart lying within 1 arcmin), we used N_{WAY} to calculate the posterior probability of it being associated with each of the AllWISE counterparts lying within 1 arcmin. We adopted the same priors on the $[W2], [W1 - W2]$ plane that were used for the RASS sample. In doing this, we have ignored the significant differences between the X-ray flux distribution of the XMMSL sources and the fluxes of the 3XMM-Bright reference sample (see Fig. 2). Using a prior derived from a better matched training sample would have been preferable, but we chose not to do so because of the very small number of 3XMM-Bright+AllWISE sources above the bright X-ray fluxes probed by the XMMSL survey. We find that for 3411/3843 (88.8 per cent) of XMMSL sources we have a best matching AllWISE counterpart with posterior probability $P_{\text{post}} \geq 0.01$. The 431 XMMSL sources without at least one AllWISE counterpart above this threshold were not considered further. The 3XMM-Bright+AllWISE reference sample suggests that at the flux limit of the XMMSL survey, virtually all persistent X-ray sources should be detected in the AllWISE survey. However, the fraction of XMMSL sources lacking AllWISE counterparts (11 per cent) is nearly three times larger than the 4 per cent of full-band X-ray detections in the ‘Clean’ XMMSL catalogue that are expected to be entirely spurious (see Section 2.2). This disparity implies that there is either a residual incompleteness in our cross-matching routine, and/or that the XMMSL catalogue contains a small but significant fraction of transient X-ray sources, which may not have persistent mid-IR counterparts above the AllWISE detection limits. We note that if we use a set of control X-ray positions (placed 6 arcmin away from each true X-ray position), and rerun the cross-matching algorithm, the fraction of sources without any valid AllWISE counterpart increases to 24 per cent of cases. The distribution of position differences between the XMMSL sources and their best matching AllWISE counterparts is shown in Fig. 4. The magnitude distributions of the best matching AllWISE counterparts are shown

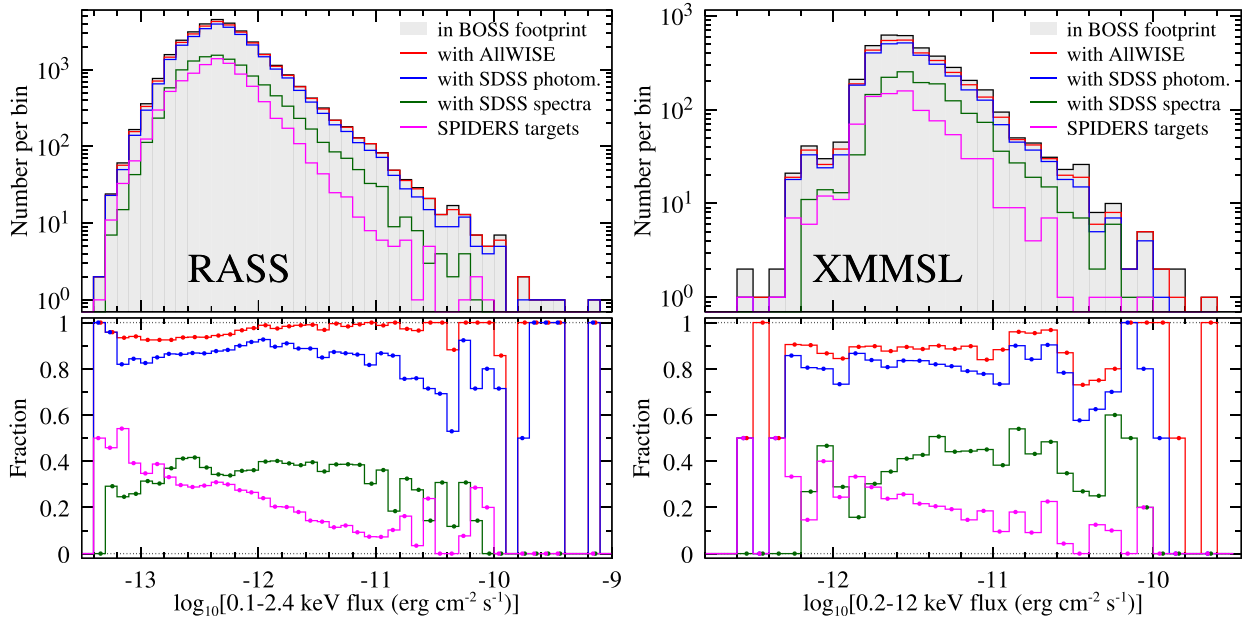


Figure 10. *First panel:* The X-ray flux distribution of RASS sources at various stages of the cross-matching and down-selection process described in Section 3.3. From top to bottom, the curves show the X-ray flux distributions for the following subsets: all RASS sources within the BOSS imaging footprint (thick black line), RASS+AllWISE sources (red), RASS+AllWISE+SDSS sources (blue), RASS+AllWISE+SDSS+z sources (green) and SPIDERS_RASS_AGN sources (magenta). *Second panel:* The same information as the upper panel, but shown as a ratio. The curves have been divided by the total number of RASS sources in each flux bin. *Third and fourth panels:* The same information shown for the XMMSL sample.

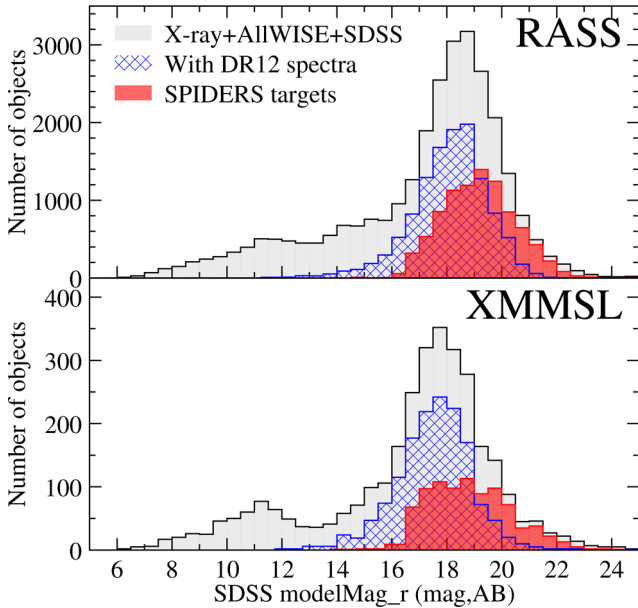


Figure 11. The SDSS r -band magnitude distribution of the optical counterparts to X-ray sources (grey histograms), those with spectroscopy available in SDSS-DR12 (blue hatched histograms) and of the SPIDERS targets to be observed during SDSS-IV (red histograms). The upper panel shows RASS sources, and the lower panel shows XMMSL sources.

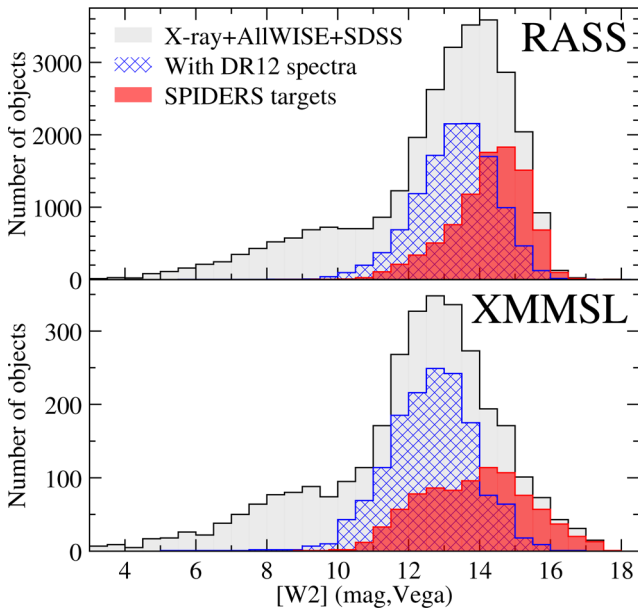


Figure 12. The [W2] magnitude distribution of the AllWISE counterparts of X-ray sources (grey histograms), those with spectroscopy available in SDSS-DR12 (blue hatched histograms) and of the SPIDERS targets to be observed during SDSS-IV (red histograms). The upper panel shows RASS sources, and the lower panel shows XMMSL sources.

in Fig. 5. The ‘best’ AllWISE’ counterpart was then matched to the SDSS-DR13 photometric catalogue, and 3142/3411 (92.1 per cent) of the XMMSL+AllWISE sources have at least one optical counterpart within 1.5 arcsec of the AllWISE position (see Figs 6 and 7). We then further down-selected from this list of 3142 optical counterparts to reach our final list of SPIDERS_XMMSL_AGN targets for observation within SDSS-IV. Matching against the SDSS-DR12 spectral catalogue removes 1411 previously spectroscopi-

cally identified X-ray sources. The properties of the spectroscopically identified XMMSL sources are discussed in more detail in Section 5.1. For the same reasons described in Section 3.3, we have removed from consideration all optical counterparts brighter than the nominal magnitude limit of eBOSS (i.e. $\text{fiber2Mag}_i < 17.0$, 746 sources), very bright objects ($\text{modelMag}_i < 16.0$, 59 sources), very faint targets ($\text{fiber2Mag}_i > 22.5$, 32 sources) and targets that fall within the BOSS bright star mask (21 targets). The final SPIDERS_XMMSL_AGN target list contains 873 candidate targets over the full BOSS targeting footprint ($0.081 \text{ targets deg}^{-2}$).

The matching and filtering steps described above are summarized in a flow diagram, see Fig. 13. The format of the catalogue of SPIDERS_XMMSL_AGN targets is described in Appendix E4. In Fig. 9, we show the sky distribution of SPIDERS_XMMSL_AGN targets over the sky. In Figs 10, 11 and 12, we show the distributions of the XMMSL sample in X-ray flux, r -band magnitude and [W2] magnitude at various stages during the cross-matching and down-selection steps described above.

3.5 X-ray sources with SDSS-DR12 spectra and visual inspections

As noted in Sections 3.3 and 3.4, our selection of targets for SPIDERS used a prototype version of the SDSS-DR12 spectroscopic sample to remove from consideration any objects having existing spectroscopy. In order to form a clean sample, we have repeated the cross-matching process of X-ray+AllWISE+SDSS objects to the official SDSS-DR12 catalogue. Using a simple 1 arcsec radial search around the SDSS-DR13 photometric position, we find that SDSS-DR12 contains spectra for 11 913 RASS and 1482 XMMSL sources. Fig. 14 shows the sky distribution of the X-ray sources with spectroscopic identifications in SDSS-DR12. Note the banded structure in the SGC region, which is imprinted by the narrow sky coverage of SDSS observations prior to the start of the BOSS project.

Although the BOSS redshift fitting algorithm (Bolton et al. 2012) has been shown to be accurate and robust for the galaxies and QSOs that have been targeted in previous iterations of the SDSS project, the routine has not yet been validated for X-ray selected AGN that dominate the SPIDERS AGN samples. Therefore, we have carried out a programme of visual inspection (VI) in order to check the accuracy of the pipeline redshifts and spectral classes for SPIDERS targets.

We first collated the existing VI information from several prior works, namely the DR12Q quasar catalogue (P aris et al. 2017; we consider all inspections from the superset and supplementary catalogues having $z_{\text{PERSON_CONF}} = 3$), the DR7Q quasar catalogue (Schneider et al. 2010), the VIs (of BLAGN, NLAGN and BL Lac counterparts to RASS sources) from A07 and the visually inspected Blazar/BL Lac catalogue of Plotkin et al. (2010). Of the 11913 RASS sources with SDSS-DR12 spectra, secure visually inspected redshifts were already available for 8848 sources: 1875 from DR12Q, 161 from Plotkin et al. (2010), 6196 from A07 and 4981 from DR7Q. Likewise, for the 1482 XMMSL sources with SDSS-DR12 spectra, VI information was already available for 1046 sources: 214 from DR12Q, 53 from Plotkin et al. (2010), 546 from A07 and 596 from DR7Q. We have visually inspected all of the remaining spectra (3325 unique spectra, associated with 3065 RASS and 436 XMMSL sources). Following our experience with the BOSS/XMM-XXL sample (Menzel et al. 2016), we divided the spectra into higher and lower risk categories. Higher risk spectra were those which met any of the following criteria: a pipeline

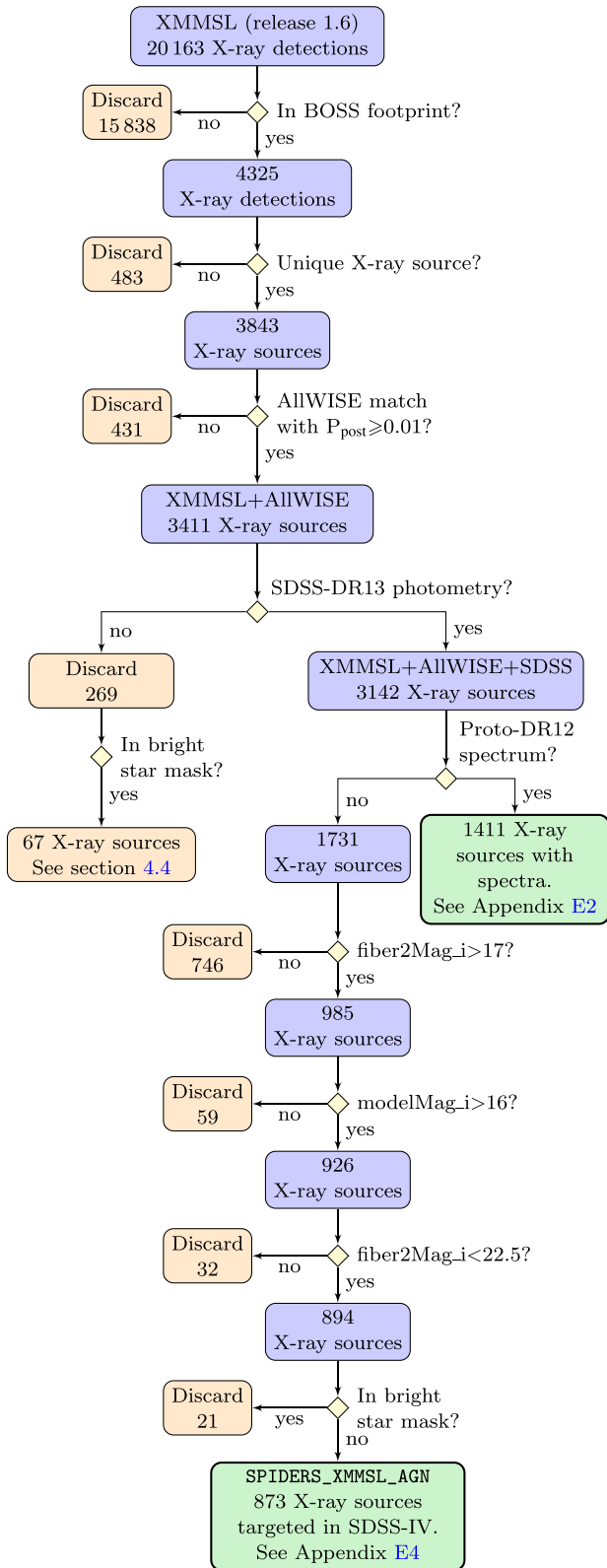


Figure 13. A schematic representation of the decision tree which leads to the selection of SPIDERS_XMMSL_AGN targets for observation within SDSS-IV. See Section 3.4 for more details.

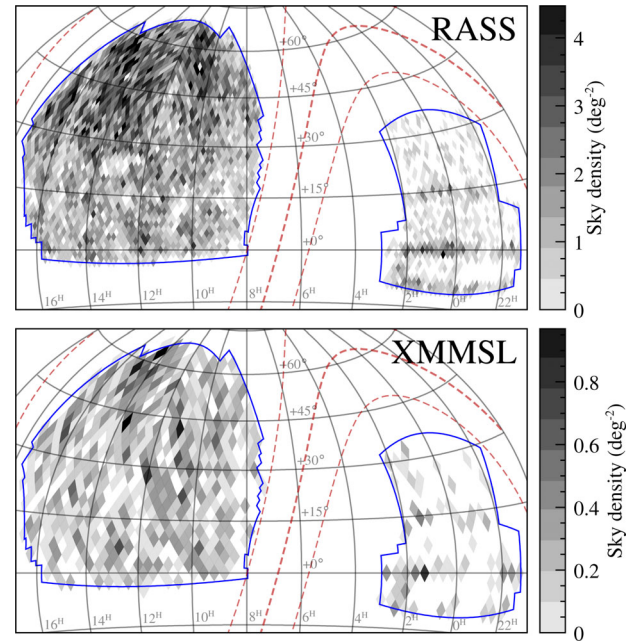


Figure 14. Upper panel: Sky distribution of RASS sources with spectroscopic identifications in SDSS-DR12, displayed with a HEALPix pixelization (NSIDE = 32, 1 pixel = 3.36 deg²). Lower panel: Same for XMMSL sources, but shown on a coarser pixel scale (NSIDE = 16, 1 pixel = 13.4 deg²). The (solid blue) line indicates the perimeter of the BOSS imaging footprint. The Galactic plane is indicated (dashed red lines at $b = 0, \pm 15$ deg).

redshift outside the range $0.01 \leq z \leq 1.0$, pipeline redshift uncertainty greater than 0.01, any redshift fitting warning flag ($ZWARNING > 0$), an r -band model magnitude fainter than 20.0 or median per-pixel SNR outside the interval [2:50]. Please see Bolton et al. (2012) for a definition of the $ZWARNING$ flag. There were 569 spectra within the higher risk category, which were examined on average by three visual inspectors. The remaining 2756 lower risk spectra were each examined by at least one visual inspector.

3.5.1 Visual inspection tools and consolidation

VI was performed using the web-browser-based SPECCY¹⁵ tool, developed (by us) to enable rapid inspection of large numbers of SPIDERS spectra. The SPECCY tool presents users with a plot of flux versus wavelength for a single observed SDSS spectrum, together with a number of functions to aid the determination of these parameters, including zooming/panning to regions of interest, and box-car smoothing of the observed spectrum. The user may overlay the observed spectrum with: the observer-frame wavelengths of common emission/absorption lines, a template spectrum (by default the best-fitting template found by the eBOSS pipeline is shown), a scaled version of the sky background spectrum, the statistical error spectrum or the residual (data-model) spectrum. The user can adjust the displayed redshift of the template and emission/absorption lines. The user submits the following information per spectrum: (i) a ‘visual’ redshift measurement; (ii) a redshift confidence flag (3 = highly secure, 2 = uncertain, 1 = poor/unusable, 0 = insufficient data), (iii) a classification [we used only six classes: QSO (including BLAGN), Broad Absorption Line QSO, Galaxy

¹⁵ <https://gitlab.rzg.mpg.de/tdwelly/speccy>

(including NLAGN), Blazar, Star, None] and (iv) freehand comments for problematic/unusual spectra.

We collected more than 5000 VIs from a cadre of 12 inspectors (assembled from within the coauthor list of this paper), and we then consolidated cases where we had multiple inspections per spectrum. We defined a standard decision tree based on the relative agreement of the multiple inspectors in three categories: redshift, redshift confidence and classification. In nearly all cases, inspections from multiple inspectors are consistent, and re-inspection/manual-reconciliation (carried out by A. Merloni) was deemed necessary in only a small fraction of cases, mostly due to discrepant redshift assignments. We give a single ‘best’ redshift and spectral classification, plus a merged confidence flag for each inspected spectrum. Where VI information was available from multiple works, we have used the following order of precedence: (i) our own VIs, (ii) DR12Q, (iii) Plotkin et al. (2010), (iv) A07 and finally (v) DR7Q.

In a small number of cases (93) the SDSS-DR12 pipeline redshift estimates disagree by more than 1 per cent ($\frac{|z_{VI} - z_{DR12}|}{1+z_{VI}} > 0.01$) from those determined through VI. These pipeline failures tend to be ‘catastrophic’, and in 70 per cent of cases they overestimate the true redshifts. As might be expected, a large fraction (45/93, 48 per cent) of the pipeline redshift failures are related to sources visually classified as Blazar/BL Lac, i.e. having a strong and relatively featureless continuum. Despite having higher than average SNR spectra, the VI process determined high confidence redshifts only for 13/45 (29 per cent) of the Blazar/BL Lacs. Another important failure mode, accounting for ~ 18 per cent of redshift failures, occurs when the pipeline misidentifies the Mg II emission line as Ly α . For most of the redshift failures (57/93, 61 per cent), the pipeline ZWARNING flag is set to a non-zero value, indicating that the algorithm itself has identified a problem with the spectrum and/or the redshift fitting process.

4 ASSESSING THE FIDELITY OF OUR TARGET SELECTION METHOD

We have undertaken a number of independent tests designed to assess the fidelity of the steps which we have used to select SPIDERS-AGN targets for observation in SDSS-IV. These tests include (i) checking the X-ray \rightarrow AllWISE association step for a sub-sample of RASS and XMMSL sources which appear in the *XMM-Newton*, *Chandra* and *Swift-XRT* serendipitous catalogues, (ii) using blank field populations to estimate the rate of spurious X-ray \rightarrow AllWISE associations over the whole RASS and XMMSL samples, (iii) evaluating the fraction of spurious X-ray detections in the RASS sample and (iv) quantifying the success rate of the association of AllWISE with SDSS-DR13 photometric sources.

4.1 Verification of our X-ray–AllWISE association method using a bright X-ray reference sample

In order to estimate the reliability of our X-ray–mid-IR–optical cross-matching technique, we require an independent catalogue of bright X-ray sources that have well-determined positions, and that have similar X-ray fluxes to the RASS and XMMSL samples. We have formed a reference sample by selecting a set of well measured bright X-ray sources from the 3XMM catalogue (Rosen et al. 2016), the *Chandra* Source Catalogue (CSC; Evans et al. 2010) and the *Swift* X-ray Telescope Point Source catalogue (1SXPS; Evans et al. 2014). In the following, we describe how we built this reference catalogue (selected with a particular emphasis on astrometric accuracy, which in practice requires high signal-to-

noise detections), and how we have used it to measure the reliability of the cross-matching process.

CSC sources: Starting from the full CSC (v1.1) catalogue¹⁶ we selected the sub-sample of 1818 sources that are bright ($0.2\text{--}2$ keV flux $> 5 \times 10^{-14}$ erg cm $^{-2}$ s $^{-1}$), high-quality (detection significance > 5 ; error ellipse major axis < 3 arcsec), point-like, unconfused, unsaturated and that lie within the BOSS imaging footprint.

3XMM sources: We re-used the catalogue of 1049 bright high quality point-like 3XMM-DR4 sources previously described in Section 2.6.

1SXPS sources: Starting from the full 1SXPS catalogue¹⁷ we selected the sub-sample of 2142 non-Gamma Ray Burst sources that are bright ($0.3\text{--}2$ keV flux $> 5 \times 10^{-14}$ erg cm $^{-2}$ s $^{-1}$), high-quality (> 40 net counts; 90 per cent error radius < 5 arcsec; detection flags < 8) and that lie within the BOSS imaging footprint.

The bright sub-samples of the CSC, 3XMM and 1SXPS reference samples were merged to form a single catalogue of bright astrometric reference sources. Where multiple detections of single X-ray sources were found in more than one catalogue (within a matching radius of 10 arcsec) we adopted a simple hierarchical approach, prioritizing *Chandra* detections over *XMM-Newton* detections, and *XMM-Newton* over *Swift* (even if in some cases this leads to a small loss of positional information). We have manually tidied the catalogue to deal with a handful of degenerate cases (two lensed QSOs, a non-nuclear source in NGC 4051, and a case where two detections of a source made in overlapping observations have not been correctly associated with each other). The merged (and tidied) reference catalogue contains 4752 unique X-ray sources (1813 with best detection in CSC, 962 from 3XMM and 1977 from 1SXPS), and is presented in Appendix E6.

We matched the X-ray sources to the nearest counterpart in the AllWISE catalogue (Cutri 2013), limiting our search to within a 5 arcsec radius of the X-ray position for CSC and 3XMM sources, and within 10 arcsec for 1SXPS sources. We found 4524/4752 of the merged X-ray reference sample had a matching AllWISE counterpart. In only 2.7 per cent of these cases was there more than one potential AllWISE counterpart within the search radius (mostly 1SXPS sources). The median position difference between X-ray and nearest AllWISE positions is 0.9 arcsec, and 90 per cent are separated by less than 3 arcsec. For the X-ray sources having detections in more than one of the CSC, 3XMM and 1SXPS reference samples, we have verified that these simple criteria consistently select the same ‘best’ AllWISE counterpart for each independent X-ray detection. We therefore make the assumption, given the high quality of the X-ray positional information for this reference sample, that these X-ray \rightarrow AllWISE associations are secure.

Finally, we have matched the merged X-ray reference catalogue to the RASS and XMMSL samples described in Sections 3.3 and 3.4, respectively. Using a search radius of 60 arcsec (and using X-ray positional information only), we find 1381 and 421 counterparts to the reference catalogue in the RASS and XMMSL samples, respectively. We calculate the fraction of cases in which, starting with the RASS or XMMSL positional information, we have chosen exactly the same AllWISE counterpart that was chosen for the X-ray reference source. The success rates for RASS and XMMSL sources are comparable, with $1314/1381 = 95.1 \pm 0.6$ per cent of RASS sources successfully matched and $404/421 = 96.0 \pm 1.0$ per cent of XMMSL sources (naïve 1σ binomial confidence interval). When

¹⁶ <http://cxc.cfa.harvard.edu/csc/>

¹⁷ <http://www.swift.ac.uk/1SXPS/docs.php>

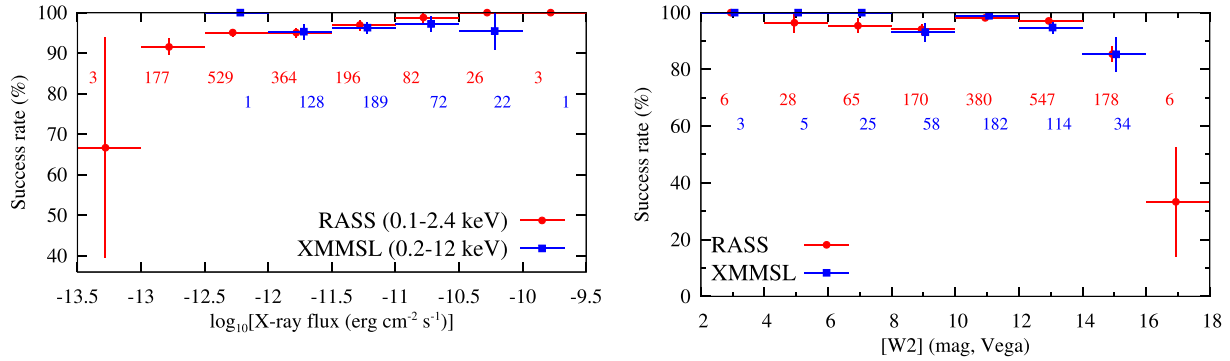


Figure 15. The success rate of the Bayesian cross-matching routine, defined as the rate at which it selects the correct AllWISE counterparts for RASS and XMMSL sources, as determined from the astrometric reference sample. *Upper panel:* Success rate as a function of X-ray flux (expressed in the native band-passes for each sample). *Lower panel:* Success rate as a function of the [W2] magnitude of the correct counterpart. The number of X-ray sources in each bin is indicated. The vertical error bars show the (naïve) binomial uncertainty. Markers are slightly offset horizontally for clarity.

only considering RASS sources matched to astrometric reference sources from the CSC, the success rate is slightly lower (92.5 ± 1.7 per cent), which is possibly due to the lower median X-ray flux of the CSC sources. There is a slight trend of decreasing success rate towards fainter X-ray fluxes, see Fig. 15. Except for the very faintest X-ray flux bin (containing only three RASS sources), the success rate is always >90 per cent. As might be expected, there is a stronger trend of decreasing success rate towards fainter [W2] magnitudes. The success rate in the $14 < [W2] < 16$ mag range is ~ 85 per cent for both RASS and XMMSL sources. Approximately 10 per cent of the X-ray samples have counterparts in this magnitude range.

4.2 Assessment of the rate of incorrect X-ray–IR associations using blank field samples

The Bayesian cross-matching algorithm implemented in *NWAY* reports the posterior probability (P_{post}) of any two objects being associated, and hence for each X-ray source gives a ranked list of potential counterparts. However, the Bayesian matching algorithm does not tell us about the rate of false positives, so we carry out an empirical check to determine the rate of spurious interlopers as a function of posterior probability (P_{post}). We have retrieved a ‘control’ catalogue of AllWISE sources within 1 arcmin radius of the positions of the RASS sources after the X-ray positions have been offset by +0.1 deg in Declination (using locations near the real RASS positions is preferred over randomly distributed positions in order to ensure that the offset catalogue samples approximately the same distribution of Galactic latitudes and AllWISE exposure depth as the RASS sources). For each offset position, we follow the method described in Section 3.1 to select the AllWISE source having the largest posterior probability of being associated with the offset RASS position. The distributions of P_{post} for the RASS and offset samples are shown in Fig. 16 (upper panel). We count up the cases where the P_{post} for the best association at the offset position is greater than the P_{post} for the best association at the actual RASS position. The lower panel of Fig. 16 shows the rate of such cases, and allows us to make an estimate of the fraction of the RASS+AllWISE associations that are spurious.¹⁸ For the pur-

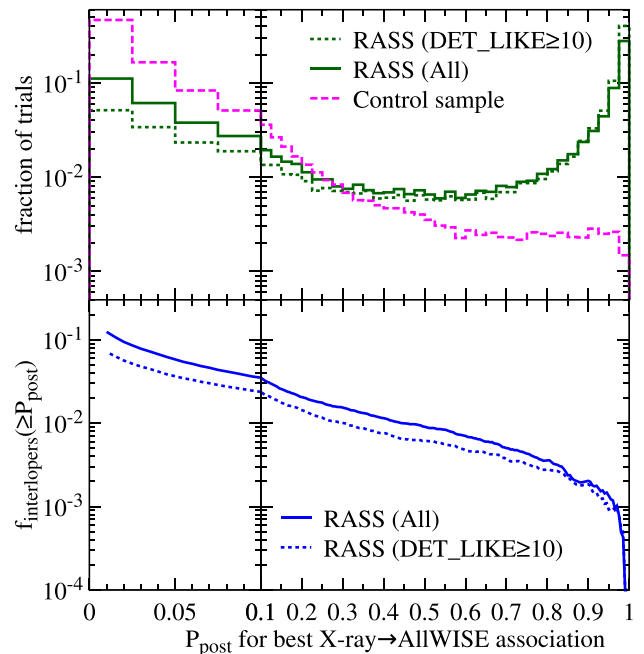


Figure 16. *Upper Panel:* The distributions of P_{post} for the most probable AllWISE associations for each RASS detection (green solid histogram), and for a control sample of nearby locations (magenta dashed histogram, offset by +0.1 deg in Declination from the RASS positions). *Lower Panel:* The expected rate of interlopers in the RASS+AllWISE sample above a given minimum threshold in P_{post} . $f_{\text{interlopers}}(\geq P_{\text{post}})$ is the cumulative fraction of RASS sources for which the most probable X-ray→AllWISE association at the offset position has a higher P_{post} than the most probable association at the actual RASS position, and represents an estimate of the fraction of spurious associations in the sample. In each panel, the dotted curves show the equivalent distributions for just the subset of RASS sources having high-confidence X-ray detections ($\text{DET_LIKE} \geq 10$). Note the change in x -axis scale at $P_{\text{post}} = 0.1$.

poses of choosing SPIDERS targets for spectroscopic follow up in SDSS-IV, we have made a cut on the posterior probability at the nominally rather low level of $P_{\text{post}} \geq 0.01$. However, the cumulative curve in the lower panel of Fig. 16 demonstrates that adopting even this low probability threshold, the total fraction of spurious associations within the RASS+AllWISE sample is at a reasonably low level (12.4 per cent of the total, equivalent to ~ 3800 spurious associations within the RASS+AllWISE sample). We repeated this

¹⁸ We did not exclude the 0.4 per cent of control positions that happen to lie close to neighbouring RASS sources, and so slightly overestimate the spurious fraction.

exercise for the subset of RASS sources with higher X-ray detection likelihood ($\text{DET_LIKE} \geq 10$), i.e. a sub-sample that should be only slightly contaminated with spurious X-ray detections. For this sub-sample, there is a smaller fraction of associations having low P_{post} , and hence a lower overall rate of spurious associations (7 per cent for $P_{\text{post}} \geq 0.01$). This suggests that at least part of the lowest probability RASS+AllWISE associations are due to spurious X-ray detections, as we discuss in more detail below. If a very pure sample is a priority, then filtering the RASS+AllWISE sample on the basis of a minimum value of P_{post} is possible. For example, in order to achieve a spurious association fraction < 3 per cent, then one should apply a cut of $P_{\text{post}} \geq 0.123$ to the full RASS+AllWISE sample ($P_{\text{post}} \geq 0.071$ for the subset of RASS sources with $\text{DET_LIKE} \geq 10$).

4.3 An evaluation of spurious X-ray detections within the RASS sample

Another source of uncertainty in our targeting procedure is the incidence of spurious detections in the RASS catalogue. As mentioned above, the RASS-FSC contains sources detected down to a relatively low confidence level (detection likelihood, $\text{DET_LIKE} > 6.5$). Boller et al. (2016) carried out simulations of the detection procedure used to produce the 2RXS catalogue. They found that the fraction of spurious detections, averaged over the whole sky and above a detection likelihood (named EXT_LIKE in the 2RXS catalogue) threshold of 6.5, could be as high as ~ 30 per cent. The spurious fraction is strongly dependent on the detection likelihood threshold adopted, and drops to ~ 2 per cent at $\text{DET_LIKE} > 10$. Unfortunately, equivalent simulations are not available for 1RXS, so we are unable to make a direct estimate of the spurious fraction in the 1RXS catalogue. However, since both 1RXS and 2RXS are based on almost the same X-ray data sets, and share many of the same source detection procedures, we assume that the spurious fraction in our 1RXS sample is comparable to that in the 2RXS catalogue.

We investigate the rate of contamination by spurious sources further by examining the properties of the RASS sample as a function of DET_LIKE . Fig. 17 shows the distribution in X-ray detection likelihood for the RASS sample, and the sub-samples with AllWISE and SDSS-DR13 photometric counterparts. Clearly, the vast majority of RASS sources not matched to any AllWISE counterpart

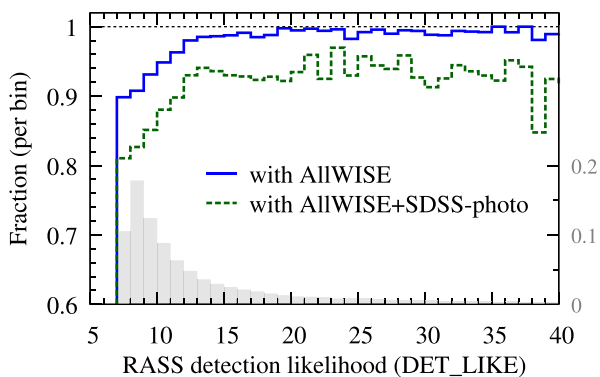


Figure 17. Fraction of RASS sources that are associated with longer wavelength counterparts as a function of X-ray detection likelihood DET_LIKE . The solid blue line shows the fraction of RASS sources having a best AllWISE counterpart with $P_{\text{post}} \geq 0.01$, the green dashed line shows the fraction with both AllWISE and SDSS-DR13 photometric counterparts. The shaded grey histogram shows the relative fraction of the RASS sample falling in each bin of DET_LIKE (see right-hand scale).

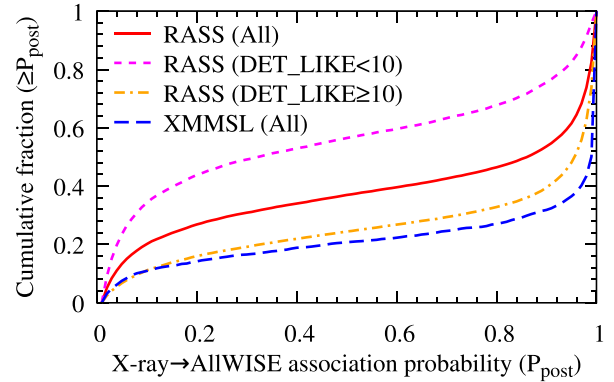


Figure 18. Cumulative distribution of the posterior association probability P_{post} for X-ray \rightarrow AllWISE matches. We show the distributions for the full RASS+AllWISE sample (red solid curve), as well as for the high (orange dot-dashed curve) and low (magenta short-dashed curve) detection likelihood subsets (split at a detection likelihood of $\text{DET_LIKE} = 10$). For comparison, we also show the distribution for the full XMMSL+AllWISE sample (blue long-dashed curve).

have low detection likelihoods, and we should expect about half of them to be spurious detections in the RASS catalogue. The fraction of RASS+AllWISE sources lacking optical photometric counterparts appears to be relatively independent of the X-ray detection likelihood.

We can also verify the reliability of the source identifications, by looking at the posterior probability P_{post} distribution of those RASS sources that are matched to AllWISE counterparts. Fig. 18 shows the cumulative (normalized) posterior probability distribution for RASS+AllWISE sources, and the breakup of the same distribution as a function of X-ray detection likelihood. Clearly, the population of sources with lower X-ray detection likelihoods have a much lower average posterior probability of being associated to their WISE counterparts. It can also be seen that for high detection likelihood RASS sources ($\text{DET_LIKE} \geq 10$) the cumulative distribution of P_{post} is very similar to that for XMMSL sources, which are expected to have only a ~ 4 per cent contamination by spurious X-ray detections (see also the online documentation¹⁹ Saxton et al. 2008).

4.4 An evaluation of the method used to associate AllWISE sources with optical counterparts

The *NWAY* tool (Salvato et al. in preparation) can compute association probabilities for sets of targets located in two or more catalogues (e.g. X-ray, IR and optical), with any number of priors defined for each possible combination of wavebands (also see the recent work by Pineau et al. 2017). However, in order to avoid imprinting complex biases into our target selection for SPIDERS-AGN, we decided to instead use the rather simple method of choosing the brightest *r*-band counterpart in the SDSS-DR13 photometric data base lying within 1.5 arcsec of the AllWISE position. For a non-negligible fraction (~ 8 per cent) of the X-ray+AllWISE sources, no optical counterpart is found within this small search radius. Fig. 17 demonstrates that the fraction of X-ray+AllWISE sources lacking optical counterparts does not strongly depend on X-ray detection likelihood (which roughly scales with X-ray brightness).

In Fig. 19, we show the [W1] magnitude distribution of X-ray-AllWISE sources lacking optical counterparts. We use the

¹⁹ <http://www.cosmos.esa.int/web/xmm-newton/xmmsl1d-ug>

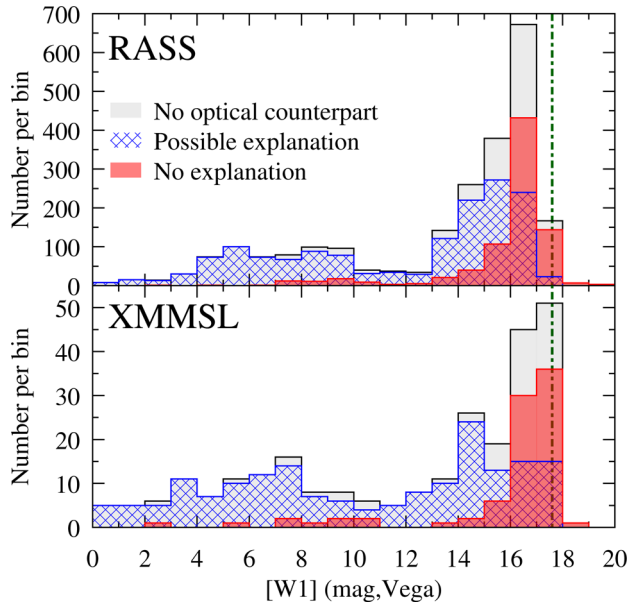


Figure 19. Distribution of AllWISE [W1] magnitude for the RASS (upper panel) and XMMSL (lower panel) sources that have AllWISE counterparts but for which we do not find SDSS-DR13 photometric counterparts within 1.5 arcsec of the AllWISE coordinates (grey shaded histogram). The hashed blue regions indicate the subset of sources for which we have a possible explanation (they either lie within the SDSS bright star mask, a bad SDSS imaging field or inside the disc of a large bright galaxy, have at least one AllWISE warning flag set, or have an SDSS-DR13 photometric counterpart within 5 arcsec of the AllWISE position). The red shaded histogram indicates the remaining sources where we have no immediate explanation for the lack of an optical counterpart. The vertical line (green dot-dashed) shows the nominal [W1] 5σ point source detection limit.

[W1] band here because at the faint end of the distribution ($[W1] > 15$ mag), more than 90 per cent of objects have their highest SNR detection in this band. Some of the cases at the bright end of the [W1] distribution can be attributed to very bright stars, for which the imaging was heavily saturated, leading to catalogue incompleteness and/or degraded IR/optical positions. Indeed, 453 of the RASS and 67 of the XMMSL sources that lack optical counterparts lie within the BOSS-DR10 bright star rejection mask. For 64 RASS and 12 XMMSL sources, the AllWISE positions fall within the optical extent of bright galaxies (de Vaucouleurs et al. 1992). In these cases, we might reasonably expect the AllWISE and SDSS positions to differ. A few additional cases (32 RASS and 6 XMMSL sources) fall within SDSS fields having bad photometry.²⁰ For 17 per cent of the RASS sources lacking optical counterparts (23 per cent for XMMSL), the AllWISE `cc_flags` column is non-zero in either the [W1] or [W2] bands (indicating potential contaminated or spurious detections; Cutri 2013), much higher than the rate of flagged detections (~ 6 per cent) in the X-ray+AllWISE samples for which we do find SDSS counterparts.

After filtering out all these cases, we are still left with 1689 RASS and 173 XMMSL sources lacking optical counterparts. The most obvious remaining explanation is that our adopted search radius was too small. Indeed, if we expand the search radius to 5 arcsec around the AllWISE position then we find counterparts to a further 833 RASS and 81 XMMSL sources. This leaves 856 RASS and

92 XMMSL sources (~ 3 per cent of the X-ray+AllWISE parent samples) where we cannot immediately explain the lack of an optical counterpart. Apart from a few very bright objects (likely to be high proper motion stars or stars that saturate the SDSS imaging), nearly all of the unexplained cases lie at the faint end of the AllWISE magnitude distribution (see Fig. 19). It is possible that some fraction of these faint objects could have such red optical–IR colours that they are undetected in the SDSS imaging. We will carry out a more thorough analysis of such objects in a future SPIDERS-AGN data release, including an inspection of deeper imaging.

5 RESULTS

The main goal of this work is to provide a complete account of the process by which we chose candidate X-ray AGN targets for observation in the shallower tiers of the SDSS-IV/SPIDERS spectroscopic survey. This serves the dual purpose of enabling further scientific investigation by documenting exactly the selection process, as well as providing large and homogeneous catalogues of bright X-ray sources with reliable multi-wavelength identification (in a well-defined statistical sense) and spectroscopic information. In this section, we outline the salient features of these catalogues, and we provide an outlook of possible applications of the SPIDERS-AGN samples.

We describe the catalogue format for spectroscopically identified RASS and XMMSL sources in Appendices E1 and E2.

5.1 Properties of X-ray sources with existing spectra in SDSS-DR12

Of the X-ray sources with SDSS-DR12 spectra, there are 11788/11913 RASS and 1469/1482 XMMSL sources having counterparts with VI confidence levels of three or equivalent (i.e. secure redshifts), and pipeline redshifts that agree well with the VI redshifts (i.e. $|z_{\text{pipe}} - z_{\text{VI}}| < 0.01[1 + z_{\text{VI}}]$), or that are identified as Blazar/BL Lac by VI, and have a VI confidence level of 2 or greater.

5.1.1 X-ray luminosity distribution

For each of the RASS and XMMSL sources with secure redshifts, we calculate the rest-frame X-ray luminosities as follows. We use a K -correction term that is appropriate for the same simple power-law spectral models that were assumed when converting from observed count rate to flux. Specifically, we compute,

$$L_{[E'_{\min}:E'_{\max}]} = 4\pi d_L^2(z) K_{\text{corr}} F_{[E_{\min}:E_{\max}]}, \quad (4)$$

where $L_{[E'_{\min}:E'_{\max}]}$ is the luminosity in the rest-frame $[E'_{\min} : E'_{\max}]$ energy interval, $d_L(z)$ is the luminosity distance based on the visually inspected spectroscopic redshift (see Section 3.5) and $F_{[E_{\min}:E_{\max}]}$ is the Galactic absorption corrected flux in the observed $[E_{\min} : E_{\max}]$ energy band. The K -correction term, including bandpass conversion, is given by

$$K_{\text{corr}}(E'_{\min}, E'_{\max}, E_{\min}, E_{\max}, \Gamma, z) = (1+z)^{\Gamma-2} \frac{E_{\max}^{2-\Gamma} - E_{\min}^{2-\Gamma}}{E'_{\max}{}^{2-\Gamma} - E'_{\min}{}^{2-\Gamma}}, \quad (5)$$

where Γ is the assumed spectral index ($\Gamma = 2.4$ for RASS sources and $\Gamma = 1.7$ for XMMSL sources).

For RASS sources, we calculate the 0.5–2 keV luminosity from the Galactic absorption corrected 0.1–2.4 keV flux, assuming

²⁰ http://data.sdss3.org/sas/dr10/boos/lss/reject_mask/badfield_mask_postprocess_pixs8.py

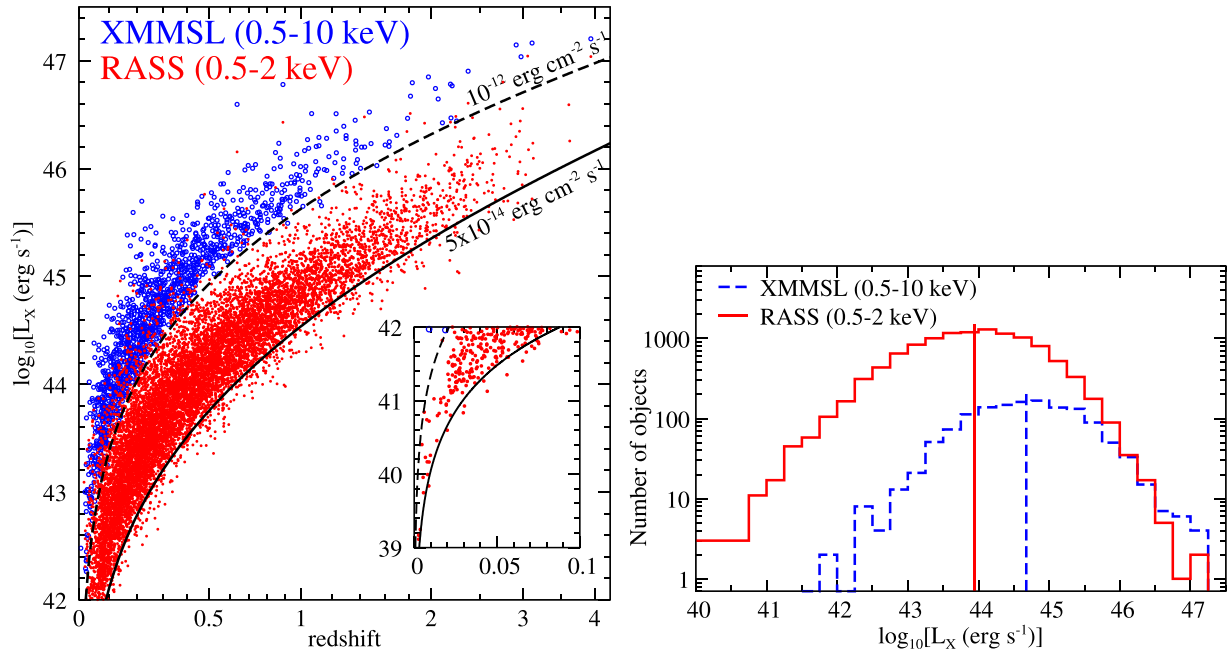


Figure 20. *Upper panel:* Distribution of spectroscopically identified RASS and XMMSL sources with high confidence redshift measurements in the redshift – X-ray-luminosity plane. X-ray luminosities (L_X) are given in the 0.5–2 keV band for RASS sources (red dots) and in the 0.5–10 keV band for XMMSL sources (blue open circles). The luminosities corresponding to two fiducial flux limits are indicated with labelled black curves. An inset shows the few X-ray sources falling below the lower luminosity limit of the main plot. *Lower panel:* A histogram showing the X-ray luminosity distribution of the high confidence extragalactic RASS and XMMSL samples. Vertical lines show the median X-ray luminosities of each sample.

$\Gamma = 2.4$. Similarly, for the XMMSL sources we assume $\Gamma = 1.7$ to calculate the 0.5–10 keV luminosity from the Galactic absorption corrected 0.2–12 keV flux. The distribution of the spectroscopically identified RASS and XMMSL sources in the redshift–X-ray luminosity plane is shown in Fig. 20. The median X-ray luminosities of the spectroscopically identified RASS and XMMSL samples are $10^{43.9}$ and $10^{44.7}$ erg s^{-1} , respectively. Both samples contain significant numbers of sources (1177 for RASS and 473 for XMMSL) at extremely high X-ray luminosities ($>10^{45}$ erg s^{-1}).

In Figs 21 and 22, we illustrate the partition of the X-ray samples by spectral class as a function of X-ray luminosity and redshift, respectively.

5.1.2 Spectral classifications

In Table 2, we give a breakdown of the BOSS pipeline spectral classifications (or VI classification for Blazar/BL Lac) for the counterparts to the X-ray sources with confident spectral classifications. As expected, the vast majority (79 per cent for RASS and 84 per cent for XMMSL) of these objects have optical spectra classified as QSO, or as GALAXY but with some signature of AGN emission (i.e. SUBCLASS = AGN, AGN BROADLINE, BROADLINE, STARBURST BROADLINE or STARFORMING BROADLINE). Blazar/BL Lacs identified through VIs constitute 2 and 5 per cent of the RASS and XMMSL samples, respectively. Another 2–3 per cent of the spectra are classified as GALAXY with signs of ongoing star-formation activity (SUBCLASS = STARBURST and STARFORMING). It is feasible that AGN emission features in these star-forming galaxies (SFGs) could be overwhelmed by the emission lines powered by star formation. Alternatively some or all of the X-ray emission could be powered by star-formation activity. Only a small fraction (1–2 per cent) of the X-ray sources are classified as STAR.

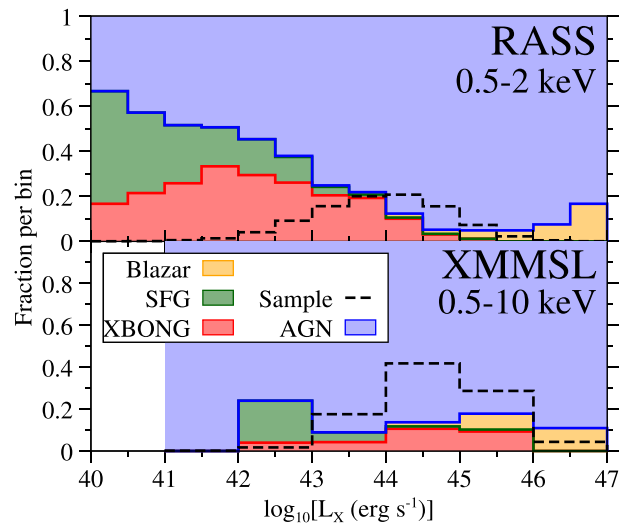


Figure 21. The relative frequency of spectroscopic classifications of RASS (upper panel) and XMMSL sources (lower panel) in bins of X-ray luminosity. In each bin, the colours indicate the fraction of objects that are classed in five broad categories: AGN (including QSOs, narrow line AGN and any objects exhibiting broad emission lines, blue shaded region), Blazars (including objects classed as BL Lac, yellow), SFGs (including objects classed as star-forming or starburst galaxies, green), Galactic stars (magenta) and XBONG (objects classed as Galaxy with no subclass, red). The relative number of objects in the parent sample falling in each bin is indicated by the black dashed histogram.

However, virtually all of the X-ray sources having optically bright but unobserved counterparts (unobserved because they are brighter than the SDSS spectroscopic limit) are also likely to be Galactic stars.

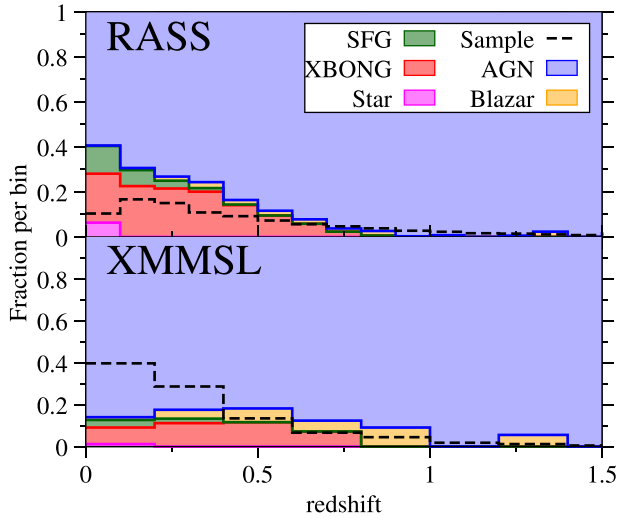


Figure 22. The relative frequency of spectroscopic classifications of RASS (upper panel) and XMMSL sources (lower panel) in bins of redshift. Same colour scheme as Fig. 21.

Table 2. The frequency of spectroscopic classifications within the samples of RASS and XMMSL targets with spectra in SDSS-DR12. These statistics only include the objects for which we either have (i) high confidence VIs and pipeline redshifts that agree with the VI redshifts or (ii) where VI has classified the object as Blazar/BL Lac with reasonable confidence. Except for Blazar/BL Lacs, the CLASS and SUBCLASS parameters are taken directly from the BOSS pipeline outputs, described in full by Bolton et al. (2012).

CLASS	SUBCLASS	N_{RASS}	N_{XMMSL}
Blazar/BL Lac		259	70
	No subclass	1639	123
	AGN	190	31
	AGN BROADLINE	50	8
	BROADLINE	98	10
GALAXY	STARBURST	77	5
	STARBURST BROADLINE	8	4
	STARFORMING	297	25
	STARFORMING BROADLINE	18	4
	Sub-total	2518	258
QSO	No subclass	246	8
	AGN	18	1
	AGN BROADLINE	104	27
	BROADLINE	6305	731
	STARBURST	19	1
STAR	STARBURST BROADLINE	2133	391
	STARFORMING	7	0
	STARFORMING BROADLINE	93	13
	Sub-total	9012	1189
	Other	CV	27
F5		10	0
F9		22	1
K1		12	0
K3		10	0
K7		13	1
M2		11	0
M3		12	0
M4		33	0
Other subclass		77	4
Sub-total	258	22	
Total		11788	1469

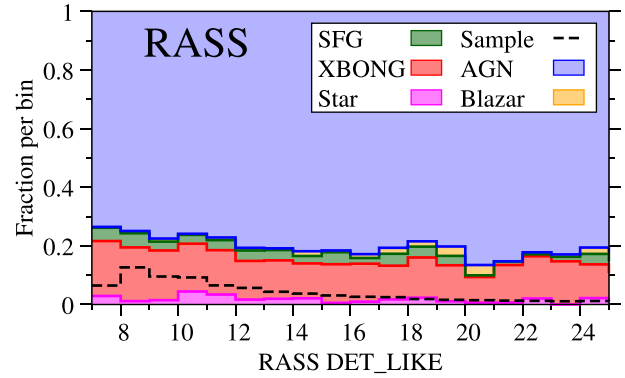


Figure 23. The relative frequency of spectroscopic classifications of RASS sources in bins of the X-ray detection likelihood. Same colour scheme as Fig. 21.

A significant minority of the X-ray sources (14 per cent for RASS and 8 per cent for XMMSL) have spectra classified by the BOSS pipeline as GALAXY, with no subclass, i.e. optically quiescent galaxies (commonly referred to as ‘X-ray bright, optically normal galaxies’, XBONGs Comastri et al. 2002). Many examples of XBONGs have been identified in previous X-ray surveys (e.g. Barger et al. 2001; Georgantopoulos & Georgakakis 2005; Page et al. 2006; Civano et al. 2007; Trump et al. 2009; Menzel et al. 2016, and references within) with the optical dullness of such galaxies most often attributed to either dilution of a ‘normal’ AGN signature by the host-galaxy light, or to objects with intrinsically weak AGN emission signatures (Trump et al. 2009). Alternatively, when the X-ray and optical measurements are well separated in time, then strong luminosity variations could potentially be another reason for the lack of a clearly observed optical AGN signature. It is also possible for weak AGN signatures to remain hidden within the noise of low SNR and/or low resolution spectra; as typically obtained by spectroscopic follow-up programmes where the primary goal is just to obtain redshift measurements for as many X-ray sources as possible.

In previous studies, XBONGs have typically been found in much smaller quantities than seen in our two X-ray samples (c.f. rates of 2.5 per cent in the XMM-COSMOS sample; Cappelluti et al. 2009; Trump et al. 2009, and 4 per cent in the XMM-XXL survey; Menzel et al. 2016). We would also not expect XBONGs to appear in the high X-ray luminosity part of our samples (see Fig. 21), where the mismatch between X-ray and optical properties is most marked. In order to explain these apparent discrepancies, we have investigated the SPIDERS XBONG sample in more detail.

One potential explanation for the lack of clear spectroscopic signatures of AGN activity appearing in the spectra of these objects is that such features (e.g. broad $H\alpha$, $H\beta$; strong $[O\ III]$) are redshifted out of the observed wavelength range. However, we can see from Fig. 22 that the SPIDERS XBONGs all lie at low redshifts. Indeed, most (>75 per cent) of the apparent XBONGs lie at $z < 0.4$, where the $H\alpha$ line is still within the red limits of the SDSS and BOSS spectrographs (920 and 1040 nm, respectively; Smee et al. 2013); all of the apparent XBONGs lie at redshifts where the $H\beta$ and $[O\ III]$ lines are comfortably within the wavelength range of the observed spectra.

Fig. 23 shows that there is no obvious trend for the fraction of quiescent galaxies to increase at low X-ray detection likelihoods, so we cannot immediately attribute the apparent excess of quiescent galaxies to spurious X-ray detections. However, Fig. 24

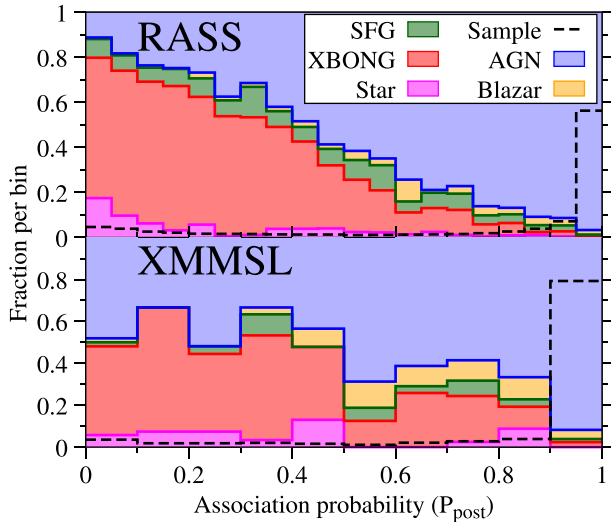


Figure 24. The relative frequency of spectroscopic classifications of RASS (upper panel) and XMMSL sources (lower panel) in bins of the X-ray→AllWISE association posterior probability (P_{post}). Same colour scheme as Fig. 21.

illustrates that these quiescent galaxies form a significant fraction of the X-ray→AllWISE associations with small values of P_{post} , and so could potentially be contaminated by a higher than average rate of incorrect associations. Note that the BOSS pipeline can assign an empty subclass to galaxies having only weak emission lines, including objects with weak evidence of AGN emission (such features were not uniformly searched for as part of the VI process). Therefore, not all of the SPIDERS-AGN sources associated with quiescent galaxies will be completely passive.

In order to keep our input X-ray samples as complete as possible, we have not filtered the RASS and XMMSL catalogues to exclude objects with significantly extended X-ray emission; in any case, the significant detection of source extent for marginally resolved objects close to the detection limit is extremely difficult. Therefore, it is possible, in some fraction of cases, that we have falsely attributed X-ray emission to AGN activity, whereas it is actually emission from the intra-cluster medium of a galaxy cluster. In order to test this hypothesis, we examined the rate at which the RASS and XMMSL sources that are associated with spectroscopically quiescent galaxies also appear within the optically selected ‘red-sequence Matched-filter Probabilistic Percolation’ (redMaPPer) catalogue of candidate cluster member galaxies presented by Rykoff et al. (2014). We cross-matched the redMaPPer catalogue to the optical counterparts of RASS and XMMSL sources having SDSS-DR12 spectra, using the optical positions and a matching radius of 1 arcsec. We only consider redMaPPer galaxies that have cluster membership probabilities of at least 50 per cent (and which lie in clusters of ‘richness’ ≥ 20). We find that a significant fraction (~ 35 per cent) of the XBONGs are matched to redMaPPer cluster member galaxies (623/1639 and 38/123 of the XBONG counterparts to RASS and XMMSL sources, respectively). As a comparison, the frequency with which redMaPPer galaxies are matched to the non-XBONG counterparts to RASS and XMMSL sources is much lower, of order 1 per cent. It is therefore likely that some, but not all of the apparent XBONGs in our spectroscopic samples are actually due to X-ray detections of galaxy clusters. We note that only 2 per cent of the SPIDERS-AGN targets put forward for observation within eBOSS/TDSS/SPIDERS are also selected as galaxy

cluster targets by Clerc et al. (2016). In Appendix C, we present a full breakdown of the overlap of SPIDERS-AGN targets with other eBOSS/TDSS/SPIDERS target classes.

6 DISCUSSION

6.1 Comparison with Anderson et al. (2007)

Our spectroscopically identified sample of RASS sources described in Section 5.1 is 70 per cent larger than that presented by A07. This is due both to the smaller spectroscopic survey footprint that was available to A07 (5740 deg² versus the 10 788 deg² considered here), and also due to the increased sky density of spectroscopic identifications in our sample. In the common area, our catalogue contains 32 per cent more RASS sources with spectra than the catalogue of A07.

We find that 6444 of the RASS sources with unambiguous spectroscopic counterparts presented by A07 also appear in our catalogue of RASS sources with AllWISE+SDSS-DR13 photometric counterparts. For 95.6 per cent of these X-ray sources, we choose exactly the same optical counterpart (within 0.5 arcsec) as chosen by A07. For most (207/281) of the disagreements, we choose a counterpart that is closer to the RASS position than that chosen by A07. This is not unexpected since A07 have considered all potential spectroscopic counterparts lying anywhere within a 1 arcmin radius circle centred on the X-ray position (essentially a flat prior), whereas the Bayesian cross-matching algorithm smoothly down-weights potential counterparts lying far from the nominal X-ray position. The two samples have similar bulk properties, e.g. the median *r*-band magnitude is 18.2 for our spectroscopically identified RASS sample compared to 18.3 for the subset that also appears in A07. Similarly, for extragalactic sources, the median X-ray luminosities are 10^{43.9} and 10^{44.0} erg s⁻¹ (0.5–2 keV), and the median redshifts are 0.36 and 0.41 for our spectroscopically identified RASS sample and the A07 subset, respectively.

6.2 Predictions for the SPIDERS-AGN Tier-0 programme

The combination of existing (SDSS-DR12) redshift measurements and the forthcoming SDSS-IV/SPIDERS data set will allow us to build a highly complete sample of X-ray selected AGN. We can make a confident prediction for the characteristics of the RASS and XMMSL samples at the end of the SPIDERS programme.

Scaling down from the full BOSS footprint (10 778 deg²), we expect ~ 22500 RASS sources and ~ 2670 XMMSL sources to lie within the 7500 deg² predicted to be covered by the final eBOSS survey footprint. Including both the existing SDSS-DR12 spectra and spectra expected to be collected during SDSS-IV, we estimate that ~ 63 per cent of RASS and XMMSL sources in the eBOSS footprint will have spectroscopic redshift information available at the completion of the SPIDERS observations. This does not take into account the large fraction (around 25 per cent) of RASS and XMMSL sources that are associated with optical counterparts (mostly stars) brighter than the bright limit of the BOSS spectrograph ($\text{fiber2Mag}_i < 17$). Discounting these optically very bright sources, the spectroscopic coverage rises to 88 per cent. The success rate for reliably measuring redshifts/classifications from SDSS spectra of counterparts to RASS and XMMSL sources is very high, even for the faint end of the population (~ 97 per cent). Therefore we estimate that, at least for $\text{fiber2Mag}_i \geq 17$, the fraction of the RASS and XMMSL samples with reliable

spectroscopic redshift measurements will be ~ 85 per cent by the end of the survey.

In summary, the combination of our novel source identification method, and systematic spectroscopic follow-up with SDSS will result, after completion of the SPIDERS-AGN Tier-0 programme, in highly complete samples of X-ray selected AGN, probing in a unique way, the portion of the luminosity-redshift plane populated by the most luminous accreting black holes.

7 CONCLUSIONS

We have presented here identification of RASS and XMMSL point-like sources with longer wavelength counterparts in SDSS and *WISE* catalogues with a novel Bayesian cross-matching algorithm that allows priors from multiple catalogues to be considered. In particular, we have shown how the location of sources in the *WISE* $[W1 - W2]$ colour-magnitude diagram provides a very efficient way to identify correctly (with 88 per cent purity overall, and 93 per cent for the subset having high confidence X-ray detections) the mid-IR counterparts to the bright X-ray sources detected in the RASS and XMMSL surveys.

We have then used our new identifications to inform the targeting for the ongoing SPIDERS survey, a sub-programme of the SDSS-IV project, and we present the list of SPIDERS-AGN targets submitted over the entire BOSS imaging footprint. In addition, we present samples of 11 913 RASS+AllWISE and 1482 XMMSL sources that we have associated with spectra published in the SDSS-DR12 catalogue, already one of the largest uniformly selected spectroscopic samples of X-ray AGN ever compiled.

Based on such samples, and on the progress of the ongoing eBOSS/TDSS/SPIDERS survey, we estimate that, at the end of the SDSS-IV programme, combining spectra from all generations of SDSS surveys, we will compile a highly complete (~ 85 per cent) sample of good-quality optical spectra for more than 15 000 bright X-ray AGN selected both in the soft (*ROSAT*) and broad (*XMM-Newton*) X-ray bands, over an area of ~ 7500 deg².

ACKNOWLEDGEMENTS

We are very grateful to the referee, F. Carrera, for his detailed comments and suggestions that have substantially improved this work.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation and the U.S. Department of Energy Office of Science. The SDSS-III web site is www.sdss3.org. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington and Yale University.

Funding for the SDSS-IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science

and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University and Yale University.

This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, and *NEOWISE*, which is a project of the Jet Propulsion Laboratory/California Institute of Technology. *WISE* and *NEOWISE* are funded by the National Aeronautics and Space Administration. This research has made use of the NASA/IPAC Infrared Science Archive, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This research has made use of ‘Aladin sky atlas’ developed at CDS, Strasbourg Observatory, France (Bonnarel et al. 2000; Boch & Fernique 2014). This research has made use of data obtained from the *Chandra* Source Catalog, provided by the *Chandra* X-ray Center (CXC) as part of the *Chandra* Data Archive. This research has made extensive use of the *STILTS* toolset (Taylor 2006). MB acknowledges support from the FP7 Career Integration Grant ‘eEASy: supermassive black holes through cosmic time: from current surveys to eROSITA-Euclid Synergies’ (CIG 321913).

REFERENCES

- Abazajian K. N. et al., 2009, *ApJS*, 182, 543
 Aihara H. et al., 2011, *ApJS*, 193, 29
 Alam S. et al., 2015, *ApJS*, 219, 12
 Anderson S. F. et al., 2007, *AJ*, 133, 313 (A07)
 Asmus D., Hönig S. F., Gandhi P., Smette A., Duschl W. J., 2014, *MNRAS*, 439, 1648
 Assef R. J. et al., 2013, *ApJ*, 772, 26
 Barger A. J., Cowie L. L., Mushotzky R. F., Richards E. A., 2001, *AJ*, 121, 662
 Blanton M. R. et al., 2017, preprint ([arXiv:1703.00052](https://arxiv.org/abs/1703.00052))
 Boch T., Fernique P., 2014, in Manset N., Forshay P., eds, *ASP Conf. Ser. Vol. 485, Astronomical Data Analysis Software and Systems XXIII*. Astron. Soc. Pac., San Francisco, p. 277
 Bogdanov S., Grindlay J. E., 2009, *ApJ*, 703, 1557
 Boller T., Freyberg M. J., Trümper J., Haberl F., Voges W., Nandra K., 2016, *A&A*, 588, A103

- Bolton A. S. et al., 2012, *AJ*, 144, 144
 Bonnarel F. et al., 2000, *A&AS*, 143, 33
 Brandt W. N., Hasinger G., 2005, *ARA&A*, 43, 827
 Broos P. S. et al., 2011, *ApJS*, 194, 2
 Brusa M. et al., 2005, *A&A*, 432, 69
 Budavári T., Szalay A. S., 2008, *ApJ*, 679, 301
 Cappelluti N. et al., 2009, *A&A*, 497, 635
 Civano F. et al., 2007, *A&A*, 476, 1223
 Civano F. et al., 2012, *ApJS*, 201, 30
 Clerc N. et al., 2016, *MNRAS*, 463, 4490
 Comastri A. et al., 2002, *ApJ*, 571, 771
 Cutri R. M. et al., 2013, *VizieR Online Data Catalog*, 2328
 de Vaucouleurs G., de Vaucouleurs A., Corwin H. G., Jr, Buta R. J., Paturel G., Fouque P., 1992, *VizieR Online Data Catalog*, 7137
 Dawson K. S. et al., 2016, *AJ*, 151, 44
 Donley J. L. et al., 2012, *ApJ*, 748, 142
 Elvis M. et al., 1994, *ApJS*, 95, 1
 Evans I. N. et al., 2010, *ApJS*, 189, 37
 Evans P. A. et al., 2014, *ApJS*, 210, 8
 Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, *AJ*, 111, 1748
 Gandhi P., Horst H., Smette A., Hönig S., Comastri A., Gilli R., Vignali C., Duschl W., 2009, *A&A*, 502, 457
 Georgakakis A., Nandra K., 2011, *MNRAS*, 414, 992
 Georgantopoulos I., Georgakakis A., 2005, *MNRAS*, 358, 131
 Greiner J., Richter G. A., 2015, *A&A*, 575, A42
 Gunn J. E. et al., 1998, *AJ*, 116, 3040
 Gunn J. E. et al., 2006, *AJ*, 131, 2332
 Haakonsen C. B., Rutledge R. E., 2009, *ApJS*, 184, 138
 Hamilton A. J. S., 1993, *ApJ*, 417, 19
 Hamilton A. J. S., Tegmark M., 2004, *MNRAS*, 349, 115
 Hartmann D., Burton W. B., 1997, *Atlas of Galactic Neutral Hydrogen*. Cambridge Univ. Press, Cambridge
 Hickox R. C. et al., 2007, *ApJ*, 671, 1365
 Hinshaw G. et al., 2013, *ApJS*, 208, 19
 Hudelot P. et al., 2012, *VizieR Online Data Catalog*, 2317
 Høg E. et al., 2000, *A&A*, 355, L27
 Jones D. H. et al., 2004, *MNRAS*, 355, 747
 Jones D. H. et al., 2009, *MNRAS*, 399, 683
 Kochanek C. S. et al., 2012, *ApJS*, 200, 8
 Lacy M. et al., 2004, *ApJS*, 154, 166
 Liu Z. et al., 2016, *MNRAS*, 459, 1602
 Mahony E. K., Croom S. M., Boyle B. J., Edge A. C., Mauch T., Sadler E. M., 2010, *MNRAS*, 401, 1151
 Mainieri V., Bergeron J., Hasinger G., Lehmann I., Rosati P., Schmidt M., Szokoly G., Della Ceca R., 2002, *A&A*, 393, 425
 Marchesi S. et al., 2016, *ApJ*, 817, 34
 Mateos S., Saxton R. D., Read A. M., Sembay S., 2009, *A&A*, 496, 879
 Mateos S., Alonso-Herrero A., Carrera F. J., Blain A., Severgnini P., Caccianiga A., Ruiz A., 2013, *MNRAS*, 434, 941
 Menzel M.-L. et al., 2016, *MNRAS*, 457, 110
 Merloni A. et al., 2012, preprint ([arXiv:1209.3114](https://arxiv.org/abs/1209.3114))
 Mickaelian A. M., Hovhannissyan L. R., Engels D., Hagen H.-J., Voges W., 2006, *A&A*, 449, 425
 Morganson E. et al., 2015, *ApJ*, 806, 244
 Myers A. D. et al., 2015, *ApJS*, 221, 27
 Neugebauer G. et al., 1984, *ApJ*, 278, L1
 Page M. J. et al., 2006, *MNRAS*, 369, 156
 Parejko J. K., Constantin A., Vogeley M. S., Hoyle F., 2008, *AJ*, 135, 10
 Pâris I. et al., 2017, *A&A*, 597, A79
 Pence W. D., Chiappetti L., Page C. G., Shaw R. A., Stobie E., 2010, *A&A*, 524, A42
 Pineau F.-X., Motch C., Carrera F., Della Ceca R., Derrière S., Michel L., Schwobe A., Watson M. G., 2011, *A&A*, 527, A126
 Pineau F.-X. et al., 2017, *A&A*, 597, A89
 Planck Collaboration XIII, 2016, *A&A*, 594, A13
 Plotkin R. M. et al., 2010, *AJ*, 139, 390
 Predehl P. et al., 2016, in den Herder J. A., Takahashi T., Bautz M., eds, *Proc. SPIE Conf. Ser. Vol. 9905, Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray*. SPIE, Bellingham, 99051K
 Rosen S. R. et al., 2016, *A&A*, 590, A1
 Rots A. H., Budavári T., 2011, *ApJS*, 192, 8
 Rykoff E. S. et al., 2014, *ApJ*, 785, 104
 Saxton R. D., Read A. M., Esquej P., Freyberg M. J., Altieri B., Bermejo D., 2008, *A&A*, 480, 611
 Schneider D. P. et al., 2010, *AJ*, 139, 2360
 SDSS Collaboration et al., 2016, preprint ([arXiv:1608.02013](https://arxiv.org/abs/1608.02013))
 Smee S. A. et al., 2013, *AJ*, 146, 32
 Stern D. et al., 2005, *ApJ*, 631, 163
 Stern D. et al., 2012, *ApJ*, 753, 30
 Sutherland W., Saunders W., 1992, *MNRAS*, 259, 413
 Swanson M. E. C., Tegmark M., Hamilton A. J. S., Hill J. C., 2008, *MNRAS*, 387, 1391
 Taylor M. B., 2006, in Gabriel C., Arviset C., Ponz D., Enrique S., eds, *ASP Conf. Ser. Vol. 351, Astronomical Data Analysis Software and Systems XV*. Astron. Soc. Pac., San Francisco, p. 666
 Trump J. R. et al., 2009, *ApJ*, 706, 797
 Trümper J., 1982, *Adv. Space Res.*, 2, 241
 Véron-Cetty M.-P., Véron P., 2010, *A&A*, 518, A10
 Véron-Cetty M.-P. et al., 2004, *A&A*, 414, 487
 Voges W. et al., 1999, *A&A*, 349, 389
 Voges W. et al., 2000, *VizieR Online Data Catalog*, 9029
 Werner M. W. et al., 2004, *ApJS*, 154, 1
 Wright E. L. et al., 2010, *AJ*, 140, 1868
 York D. G. et al., 2000, *AJ*, 120, 1579
 Zickgraf F.-J., Engels D., Hagen H.-J., Reimers D., Voges W., 2003, *A&A*, 406, 535

SUPPORTING INFORMATION

Supplementary data are available at [MNRAS](https://www.mnras.org) online.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: ESTIMATING X-RAY FLUXES

A1 Estimating X-ray fluxes for RASS sources

A full X-ray spectral analysis of each RASS source is not feasible. Therefore, in order to convert the observed instrumental *ROSAT* count rates (which are collected over the 0.1–2.4 keV range) into physical fluxes we assume a simple spectral model of a power law absorbed by the full Galactic column of neutral material. This simplifying assumption, valid for the extragalactic AGN in which we are most interested, does however result in overestimated unabsorbed fluxes, for any Galactic sources that lie in front of a significant fraction of the total Galactic column.

We estimate the Galactic column density in the direction of each RASS source using the map of Galactic N_{H} provided by the NASA LAMBDA team,²¹ which in the region of interest (Dec. > -30 deg) is based on data from the Leiden/Dwingeloo HI Survey (Hartmann & Burton 1997). The RASS sources in our sample have a range of Galactic column density in the range $10^{19.6} \leq N_{\text{H}} \leq 10^{21.1} \text{ cm}^{-2}$, with median $N_{\text{H}} = 10^{20.36} \text{ cm}^{-2}$, see the bottom panel of Fig. A1.

To predict the instrumental count rates expected per unit flux for the above-mentioned spectral model, we use the *XSPEC* tool combined with the *ROSAT* PSPC-C (Position Sensitive Proportional

²¹ http://lambda.gsfc.nasa.gov/product/foreground/combnh_map.cfm

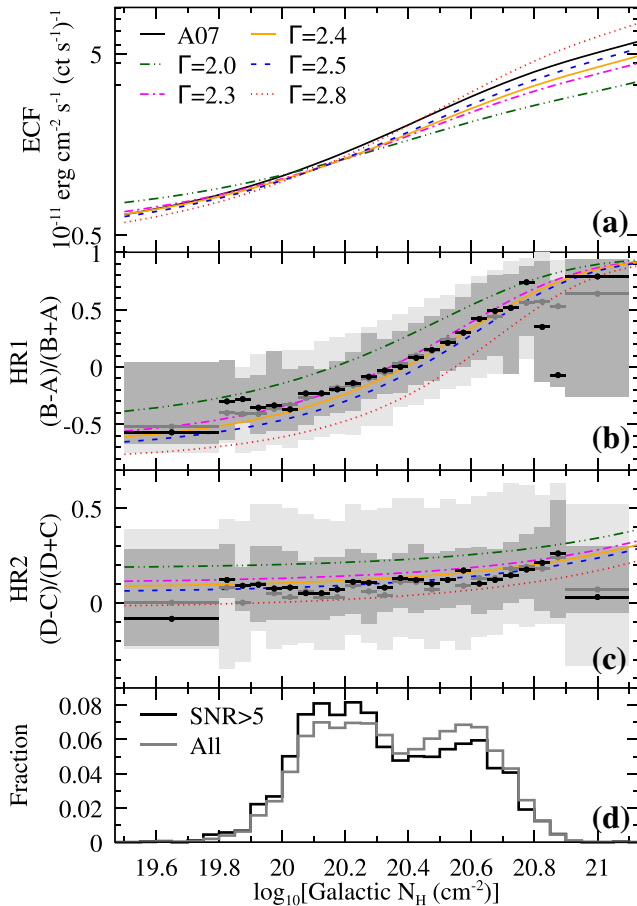


Figure A1. *Panel a:* ECFs for the *ROSAT* PSPC-C, instrument assuming a simple spectral model of a power law absorbed by the Galactic column. Also shown is the ECF assumed by A07. *Panel b:* Distribution of observed X-ray hardness ratio HR1 (see text for description) for RASS sources in the BOSS imaging footprint as a function of Galactic column. The grey points with horizontal bars and the light grey shaded areas show the median and inter-quartile range for all the RASS sources found in each bin of Galactic N_{H} . The black points with horizontal bars and darker shaded area show the same but for just those RASS sources detected with $\text{SNR} \geq 5$. The curves show the values expected for the absorbed power-law spectral model using the same colour-code indicated in panel a. *Panel c:* Same as panel b, but for the X-ray hardness ratio measure ‘HR2’. *Panel d:* Distribution of Galactic column density for RASS sources in the BOSS footprint (grey histogram). Also shown is the distribution for those RASS sources detected with $\text{SNR} \geq 5$.

Counter-C) on-axis response matrix,²² which is appropriate for RASS observations made prior to 1991 January 25. We calculated the count rates and model fluxes over a grid covering the range $10^{19.5} \leq N_{\text{H}} \leq 10^{22} \text{ cm}^{-2}$, and for spectral slopes in the range $1.0 \leq \Gamma \leq 3.0$. From the *XSPEC* outputs, we then calculate the multiplicative ECFs, which convert from observed count rate to energy flux, and also the ratios of count rates expected between different detector energy ranges (i.e. hardness ratios). Voges et al. (1999, 2000) provide two hardness ratios for the RASS sources; HR1 is calculated between energy bands ‘A’ (PSPC-C channels 11–41, ~ 0.1 – 0.4 keV) and ‘B’ (channels 52–201, ~ 0.5 – 2 keV), and HR2 is calculated between energy bands ‘C’ (channels 52–90, ~ 0.5 – 0.9 keV) and ‘D’ (chan-

nels 91–201, ~ 0.9 – 2 keV). The hardness ratios are computed as $\text{HR} = (H - S)/(H + S)$, where H and S are the vignetting-corrected count rates in the harder and softer bands, respectively.

We compare the predicted model hardness ratios (HR1, HR2) with the observed distributions, binned in small steps of Galactic N_{H} , see Fig. A1. The RASS sources have a wide scatter of hardness ratios, presumably the result of some intrinsic range of spectral shapes within the source population that is additionally broadened by measurement errors. The smaller scatter in HR1 and HR2 for RASS sources with $\text{SNR} > 5$ indicates that at least some of this scatter is likely to be due to measurement errors. However, a trend of increasing median (and inter-quartile range) of the hardness ratios with Galactic N_{H} can be clearly seen for both the full source sample and for the $\text{SNR} > 5$ subset. Over the majority of the N_{H} range, the median of the HR2 distribution is well traced by a spectral model having $\Gamma = 2.5$. At Galactic columns smaller than $N_{\text{H}} \sim 10^{20.5} \text{ cm}^{-2}$, the distribution of RASS sources in HR1 appears to favour a slightly harder spectrum with $\Gamma \sim 2.3$, hinting at mild spectral curvature over the 0.1 – 2 keV energy range. Therefore, in order to convert the RASS instrumental count rates and errors to fluxes, we have adopted the intermediate spectral index of $\Gamma = 2.4$.

Note that A07 assumed a slightly softer slope of $\Gamma = 2.5$ to convert from RASS count rates to fluxes. In the top panel of Fig. A1, we show the ECF derived using the A07 recipe in comparison to the ECFs derived using *XSPEC*. The A07 ECF is always equal to or larger than the ECF adopted here, with the ratio increasing from ~ 1 at Galactic $N_{\text{H}} = 10^{19.5} \text{ cm}^{-2}$ to ~ 1.2 at $N_{\text{H}} = 10^{21} \text{ cm}^{-2}$, and then declining at larger values of N_{H} . At the median Galactic N_{H} of the RASS sources ($10^{20.36} \text{ cm}^{-2}$), the ECF of A07 is 12 per cent larger than that used in our work.

A2 Estimating unabsorbed X-ray fluxes for 3XMM-Bright sources

The flux estimates supplied in the 3XMM catalogue were calculated assuming a simple power-law spectrum with photon-index $\Gamma = 1.7$ absorbed by a fixed neutral column density of $N_{\text{H}} = 2 \times 10^{20} \text{ cm}^{-2}$ (Mateos et al. 2009). These fluxes have not been corrected for Galactic absorption.

We have adopted the following method to estimate the unabsorbed fluxes of the 3XMM sources in the 0.1 – 2.4 keV band to allow direct comparison with the fluxes of RASS sources. We first make the assumption that the 3XMM catalogue fluxes represent a good estimate of the absorbed energy flux in each of the relatively narrow energy bands over which they were calculated: 0.2 – 0.5 keV (band 1), 0.5 – 1 keV (band 2) and 1 – 2 keV (band 3). We then use a look-up table to correct the flux in each of these three energy bands by an amount appropriate for the Galactic column density local to the direction of each source (Hartmann & Burton 1997). The look-up table was calculated using *XSPEC* and assumes an intrinsic spectral model of a power law having photon index $\Gamma = 2.4$. The unabsorbed flux in the 0.2 – 0.5 keV band is then extrapolated to 0.1 – 0.5 and the 1 – 2 keV band is extrapolated to 1 – 2.4 keV using correction factors appropriate for a spectral model of a power law having photon index $\Gamma = 2.4$. Specifically, for a power law this bandwidth correction factor is given by $F_{[E'_{\min}:E'_{\max}]} / F_{[E_{\min}:E_{\max}]} = (E'_{\max}{}^{2-\Gamma} - E'_{\min}{}^{2-\Gamma}) / (E_{\max}{}^{2-\Gamma} - E_{\min}{}^{2-\Gamma})$, where $[E'_{\min}:E'_{\max}]$ and $[E_{\min}:E_{\max}]$ give the energy bounds of the new and original energy bands, respectively. Finally, our estimate of the unabsorbed 0.1 – 2.4 keV flux is given by the unweighted sum of the unabsorbed flux estimates in the 0.1 – 0.5 , 0.5 – 1 and 1 – 2.4 keV bands.

²² http://legacy.gsfc.nasa.gov/caldb/data/rosat/pspc/cpf/matrices/pspc_gain1_256.rmf

APPENDIX B: ANALYSIS OF 3XMM-Bright SOURCES WITHOUT COUNTERPARTS IN ALLWISE

In order to determine if the few 3XMM-Bright sources lacking mid-IR counterparts represent a population of X-ray bright but IR faint emitters, we have visually inspected the *XMM-Newton*, SDSS and AllWISE images at the X-ray locations of each of the 49/1049 3XMM-Bright sources that do not have at least one AllWISE counterpart within 3 arcsec. The reasons for this can be categorized as follows. (i) In 19/49 cases, the lack of a counterpart in the AllWISE catalogue is due to pairs or groups of astrophysical objects being blended in the ~ 6 arcsec resolution AllWISE images. In many of these cases, the correct counterpart to the X-ray source (as determined from inspection of higher angular resolution SDSS optical images) is lost in the wings of the PSF of a bright star. (ii) In 15/49 cases, the X-ray source lies very close to the edge of the *XMM-Newton* field of view, and hence the X-ray positions are less well constrained than the nominal positional errors in the catalogue would suggest. For most (10/15) of these high off-axis-angle sources, a potential AllWISE counterpart lies within 5 arcsec of the X-ray position, and for the remainder an AllWISE counterpart lies within 15 arcsec. (iii) In 9/49 cases, the X-ray position lies within the optical disc of a nearby well resolved galaxy. (iv) In 3/49 cases, the X-ray source can be associated with a high proper motion star, which has moved more than 3 arcsec from the X-ray position in the time that passed between the *XMM-Newton* observation and the AllWISE observations. (v) Finally, in just 3/49 cases, the 3XMM-Bright source can be associated with a faint optical counterpart that in the mid-IR is below the detection limit of the AllWISE survey. One of these (3XMM J003027.4+045139) is the pulsar PSR J0030+0451 (Bogdanov & Grindlay 2009), one (3XMM J133935.6-003132) is associated with a faint counterpart ($i' = 20.9 \pm 0.1$, SDSS J133935.59-003132.7) in the SDSS imag-

ing, and one (3XMM J220221.4+015330) has no optical counterpart visible in the SDSS imaging. Fortunately, the latter source lies within the footprint of the CFHTLS-Wide survey, and we find that there is a single catalogued detection within 3 arcsec of the 3XMM-Bright position: an extremely faint object with $i = 24.06 \pm 0.16$ (CFHTLS 1411_033310; Hudelot et al. 2012). We conclude that mid-IR-faint objects do not form a large fraction of the X-ray source population at the bright X-ray fluxes probed by the 3XMM-Bright sample.

APPENDIX C: SDSS TARGETING EFFICIENCY AND OVERLAP OF SPIDERS-AGN TARGETS WITH OTHER eBOSS/TDSS/SPIDERS TARGET TYPES

Within the eBOSS/TDSS/SPIDERS project, a heterogeneous mix of targets from a variety of classes share the focal plane during each observation. Designing a near-optimal set of SDSS plug-plates based on these target lists (‘tiling’ in the SDSS parlance) is a non-trivial task, described in full by Dawson et al. (2016). In summary, the method is to first assign each target category into one of several ranked collision groups. Each group is tiled in sequence, taking account of the results of previous tiling rounds. Within each collision group, a priority ranking scheme is used to determine which target receives a fibre in cases of collisions. SPIDERS targets are considered within the first round of fibre assignments, and as they form the numerically smallest target type and have the highest scientific requirement for completeness, they are given the highest priorities within that round. Therefore, SPIDERS targets can potentially miss out on being assigned a fibre only if they are within 62 arcsec of another SPIDERS target (or if they are flagged as belonging to one of the eBOSS cosmology samples, see below). Within SPIDERS, we adopt an internal target priority ranking, again following a

Table C1. The numbers of SPIDERS_RASS_AGN and SPIDERS_XMMSL_AGN targets that are associated with other eBOSS/TDSS/SPIDERS target types, calculated for the first 4009.6 deg² of eBOSS. The ‘ET1’ and ‘ET2’ columns give the index of the bits that identify this target class in the EBOSS_TARGET1 and EBOSS_TARGET2 bitmaps, respectively. The N_{RASS} and N_{XMMSL} columns give the respective numbers of SPIDERS_RASS_AGN and SPIDERS_XMMSL_AGN targets that overlap with the given target class. Values in bold are the self matches.

eBOSS target flag	ET1	ET2	N_{RASS}	N_{XMMSL}
LRG1_WISE	1	n/a	9	1
QSO1_VAR_S82	9	n/a	6	0
QSO1_EBOSS_CORE	10	n/a	705	47
QSO1_PTF	11	n/a	270	15
QSO1_EBOSS_FIRST	14	n/a	96	11
TDSS_TARGET	30	any	467	48
TDSS_FES_DE	30	21	15	6
TDSS_FES_NQHISN	30	23	3	0
TDSS_FES_VARBAL	30	25	1	0
TDSS_B	30	26	161	18
TDSS_FES_HYPQSO	30	27	11	2
TDSS_FES_HYPSTAR	30	28	5	1
TDSS_CP	30	31	274	22
SPIDERS_RASS_AGN	31	0	4057	120
SPIDERS_XMMSL_AGN	31	4	120	376
SPIDERS_RASS_CLUS	31	1	93	12
SPIDERS_XCLASS_CLUS	31	5	3	3
Any non-SPIDERS target			1030	82
Unique targets			2852	200
Total			4057	376

scheme wherein rarer populations are tiled with higher priority. The SPIDERS targets are ranked as follows (from highest to lowest priority): (i) Brightest Cluster Galaxies from the SPIDERS-Clusters survey, (ii) SPIDERS_XMMSL_AGN targets, (iii) SPIDERS_RASS_AGN targets and (iv) SPIDERS-Clusters member galaxies. We note that, in order to avoid imprinting a bias in the selection function of eBOSS main samples, any SPIDERS targets that are also selected by one of the eBOSS cosmology target selection algorithms (see below) are tiled with the same priority as targets from those other programmes (i.e. at a lower priority than for unique SPIDERS targets).

At the time of writing (early 2017), a substantial fraction of the tiling process for the eBOSS/TDSS/SPIDERS survey has been carried out, covering 4010 deg² of sky (the tiling ‘chunks’ named internally as *eboss1--5*, *eboss9* and *eboss16*). In this initial sky area, indicated in Fig. 9, we find that 3971/4057 of SPIDERS_RASS_AGN targets and 369/376 of SPIDERS_XMMSL_AGN targets are assigned a fibre; an impressive targeting efficiency of 98 per cent.

Within eBOSS/TDSS/SPIDERS there are a number of independent routes through which an object can be selected to receive a fibre, and so some targets are included in more than one target selection scheme. In Table C1, we list the eBOSS/TDSS/SPIDERS target categories that overlap with the SPIDERS_RASS_AGN and SPIDERS_XMMSL_AGN targets in the first 4010 deg² of tiled eBOSS sky. We find that 25 per cent of SPIDERS_RASS_AGN targets and 22 per cent of SPIDERS_XMMSL_AGN targets overlap with at least one other non-SPIDERS target class. We can see that for SPIDERS_RASS_AGN targets, the most frequent overlap (17 per cent) is with the core eBOSS QSO target sample (*QSO1_EBOSS_CORE*), which is targeting optical+mid-IR selected objects in the redshift range $0.9 < z < 2.2$ (Myers et al. 2015). There is also some overlap with the TDSS project (12 per cent of SPIDERS_RASS_AGN targets), see Morganson et al. (2015) for more details. For SPIDERS_XMMSL_AGN, the most frequent overlap, as would be expected, is with SPIDERS_RASS_AGN targets (32 per cent), although there is also a significant overlap with the eBOSS QSOs and TDSS targets. Note that the entries for XMMSL targets in Table C1 should be considered indicative only, as for this sky area we used a version of the SPIDERS_XMMSL_AGN catalogue which mistakenly excluded sources having multiple X-ray detections (see Section 2.2).

APPENDIX D: SEQUELS: SPIDERS-AGN TARGETS IN THE eBOSS PILOT SURVEY

The Sloan Extended QUasar, ELG and LRG Survey (SEQUELS) was a pilot programme designed to demonstrate the target selection for the main eBOSS/TDSS/SPIDERS survey (Alam et al. 2015; Dawson et al. 2016). The spectroscopic observations for SEQUELS were carried out in the MJD range [56660:57166] and include a total of 117 good-quality plates, covering a sky area of 471.9 deg² within the NGC. A summary of the target selection processes for all categories of SEQUELS target types is given in section A.3 of Alam et al. (2015), here we provide a detailed description of the process by which SPIDERS-AGN targets were selected for SEQUELS. The latter differs substantially from the two-step RASS→AllWISE→SDSS method by which SPIDERS-AGN targets were selected for observation in the main part of SDSS-IV (as described in Section 3.3). Note that the XMMSL catalogue was not considered when choosing targets for SEQUELS.

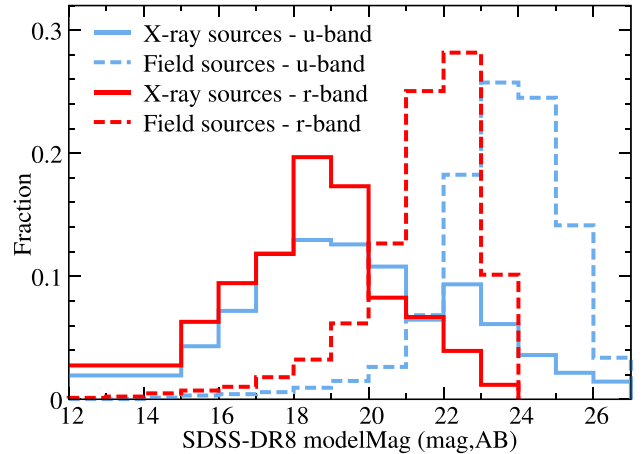


Figure D1. A graphical representation of the Bayesian priors that were used when choosing optical counterparts to RASS sources for the SEQUELS survey. For each filter considered (*u*: light blue lines, *r*: red lines), we show the magnitude distribution of the X-ray reference sample (solid) and the field population (dashed). The ratio of the reference curve to the field curve gives the prior for any potential candidate.

The total time available for SEQUELS observations was not well known at the time of selecting targets (2013), and so all contributing teams were asked to provide sufficient targets to cover an optimistically large footprint; defined by the 813 deg² of the BOSS imaging footprint lying within the bounds $120 \leq \text{RA} \leq 210$ deg and $+45 \leq \text{Dec.} \leq +60$ deg. There are 3049 RASS (BSC+FSC) sources lying inside this region, 3042 of which have at least one optical counterpart within 1 arcmin in the SDSS-DR8 photometric catalogue (Aihara et al. 2011). We removed from further consideration any RASS sources with X-ray positions lying within 30 arcsec radius of any bright star in the Tycho-II catalogue (Høg et al. 2000, 276 X-ray sources), or any object in the AGN catalogue of Véron-Cetty & Véron (2010, 307 X-ray sources), or any object having a pipeline classification of QSO in the SDSS-DR11 spectroscopic catalogue (Alam et al. 2015, 1174 X-ray sources). After these filtering steps, we are left with 1563 X-ray sources.

We used a prototype version of *nWAY* (Salvato et al. in preparation) to choose the most probable optical counterpart for each remaining RASS source. We instructed *nWAY* to take account of the X-ray positions and their uncertainties, the positions and model *u*, *r* magnitudes of the candidate optical counterparts, plus pre-defined priors on the expected distribution of counterparts in *u* and *r*. To derive the priors for the SEQUELS SPIDERS-AGN target selection, we first computed the *u*- and *r*-band magnitude distributions of a sample of well-understood, X-ray bright *XMM-Newton*-detected sources matched to the SDSS-DR7 catalogue taken from Georgakakis & Nandra (2011), and then divided these distributions by the *u*- and *r*-band magnitude distributions of all SDSS-DR8 photometric objects in the sky area considered.²³ Our adopted priors for SEQUELS are illustrated in Fig. D1. After running *nWAY*, which produces a ranked list of potential counterparts for each remaining X-ray source, we further down selected the sample to form a list of potential spectroscopic targets. For each RASS source, we

²³ Before reaching the decision to use *u*- and *r*-band magnitude priors, we first experimented with a number of alternative combinations of priors, both in magnitude and colour. However, none of the tested combinations performed as well as the *u*- and *r*-band magnitude priors finally adopted here.

considered only the ‘best’ optical counterpart (the one having the highest posterior probability as determined by N_{WAY}). Many (408/1563) of the best optical counterparts already had spectroscopic identifications/redshifts in the SDSS-DR11 spectroscopic catalogue (Alam et al. 2015, 80 with $CLASS=QSO$; 77 with $CLASS=STAR$; 240 with $CLASS=GALAXY$) and so were removed from further consideration. Of the remaining 1155 X-ray sources, we retained only the 630 having counterparts inside the magnitude range accessible by ~ 1 hour exposures with the BOSS spectrograph ($17 \leq r \leq 22$ mag; 455 were too bright, and 30 were too faint). Finally, we discarded three of the remaining RASS sources which had poorly determined X-ray positions and uncertainties. The filtering steps left 627 optical counterparts to be put forward for targeting within the SEQUELS programme. A catalogue of these targets is presented in Appendix E5.

At the plate design stage (after collation of all categories of SEQUELS targets), 599/627 of our targets were allocated a fibre within at least one of the 222 plates designed for SEQUELS. There were 307 RASS-selected AGN targets of which lie within the 117 plates that were actually manufactured and observed before the conclusion of the SEQUELS programme. The SDSS spectra associated with these targets have the $SPIDERS_RASS_AGN$ flag set within the $EBOSS_TARGET0$ bitmask in the SDSS-DR13 spectroscopic data base (SDSS Collaboration et al. 2016).

APPENDIX E: DETAILED DESCRIPTION OF SPIDERS-AGN CATALOGUES

In this section, we describe the format and content of the catalogues released as part of this paper. All of these catalogues are supplied in the Flexible Image Transport System (FITS) file format (Pence et al. 2010) and are distributed via the MPE X-ray surveys website (http://www.mpe.mpg.de/XraySurveys/SPIDERS/SPIDERS_AGN). The SPIDERS-AGN target catalogues are also hosted (albeit, in a slightly different format to that described below) on the SDSS-DR13 website (<https://data.sdss.org/sas/dr13/eboss/spiders/target>), and their content is described on the corresponding ‘data model’ webpage (https://data.sdss.org/datamodel/files/SPIDERS_TARGET). For each column of each catalogue, we give the column name, the FITS format code and a description of the column’s meaning, including units where appropriate.

E1 Catalogue of RASS sources with spectroscopy in SDSS-DR12

See Section 3.5 for a description of how this catalogue was created. The RASS+SDSS-DR12 catalogue contains 11913 entries, with the following columns:

- $RASS_*$ X-ray source properties derived from the RASS source catalogues (Voges et al. 1999, 2000)
- $RASS_ID$ (16A) RASS source identifier
- $RASS_BF_FLAG$ (1A) Flag indicating if this X-ray source is taken from the RASS-BSC (‘B’) or RASS-FSC (‘F’)
- $RASS_RA$ (1D) Right Ascension of X-ray source (deg, J2000)
- $RASS_DEC$ (1D) Declination of X-ray source (deg, J2000)
- $RASS_RADEC_ERR$ (1E) 1σ uncertainty on X-ray position (arcsec)
- $RASS_SRC_RATE$ (1E) Observed X-ray count rate in the 0.1–2.4 keV band ($ct\ s^{-1}$)
- $RASS_SRC_RATE_ERR$ (1E) Statistical uncertainty on X-ray count rate ($ct\ s^{-1}$)

- $RASS_BACK_RATE$ (1E) Estimated background count rate at source location ($ct\ s^{-1}\ arcmin^{-2}$)
- $RASS_TEXP$ (1E) Effective exposure time at source location (s)
- $RASS_HR1$ (1E) First hardness ratio measure
- $RASS_HR1_ERR$ (1E) Uncertainty on first hardness ratio measure
- $RASS_HR2$ (1E) Second hardness ratio measure
- $RASS_HR2_ERR$ (1E) Uncertainty on second hardness ratio measure
- $RASS_EXT_LIKE$ (1E) Likelihood of X-ray source being extended
- $RASS_DET_LIKE$ (1E) Likelihood of X-ray source detection
- $RASS_ECF_GAMMA2p4$ (1E) Multiplicative energy conversion factor used to convert from observed count rate to unabsorbed flux assuming a power-law spectrum with slope $\Gamma=2.4$ ($[erg\ cm^{-2}\ s^{-1}]/[ct\ s^{-1}]$)
- $RASS_SRC_FLUX$ (1E) Unabsorbed flux in the 0.1–2.4 keV band corrected for full Galactic column ($erg\ cm^{-2}\ s^{-1}$)
- $RASS_SRC_FLUX_ERR$ (1E) Statistical uncertainty on unabsorbed flux ($erg\ cm^{-2}\ s^{-1}$)
- $RASS_GAL_LONG$ (1D) Galactic longitude of X-ray source (deg)
- $RASS_GAL_LAT$ (1D) Galactic latitude of X-ray source (deg)
- $RASS_LOGGALNH$ (1E) Logarithm of Galactic column density in direction of X-ray source ($\log_{10}[cm^{-2}]$)
- $RASS_ALLW_DIST$ (1D) Distance between X-ray position and position of best AllWISE counterpart (arcsec)
- $RASS_ALLW_BIAS$ (1D) Bayesian prior (or bias) factor, $\Pi(x)$, of the best AllWISE counterpart, derived from its location in $[W2]-[W1 - W2]$ parameter space
- $RASS_ALLW_BFPOST$ (1D) Bayesian posterior probability of the best X-ray \rightarrow AllWISE association, before considering bias factor
- $RASS_ALLW_POST$ (1D) Bayesian posterior probability of the best X-ray \rightarrow AllWISE association, including the bias factor
- $ALLW_*$ Properties of the best AllWISE counterpart to the X-ray source (all taken from Cutri 2013)
- $ALLW_ra$ (1D) The Right Ascension of the AllWISE source (deg, J2000)
- $ALLW_dec$ (1D) The Declination of the AllWISE source (deg, J2000)
- $ALLW_siga$ (1E) The component of the positional uncertainty of the AllWISE source parallel to Right Ascension (arcsec)
- $ALLW_sigdec$ (1E) The component of the positional uncertainty of the AllWISE source parallel to Declination (arcsec)
- $ALLW_w1mpro$ (1E) The $[W1]$ magnitude of the AllWISE source (mag, Vega)
- $ALLW_w1sigmpro$ (1E) The $[W1]$ magnitude uncertainty of the AllWISE source (mag, Vega)
- $ALLW_w2mpro$ (1E) The $[W2]$ magnitude of the AllWISE source (mag, Vega)
- $ALLW_w2sigmpro$ (1E) The $[W2]$ magnitude uncertainty of the AllWISE source (mag, Vega)
- $ALLW_w3mpro$ (1E) The $[W3]$ magnitude of the AllWISE source (mag, Vega)
- $ALLW_w3sigmpro$ (1E) The $[W3]$ magnitude uncertainty of the AllWISE source (mag, Vega)
- $ALLW_w4mpro$ (1E) The $[W4]$ magnitude of the AllWISE source (mag, Vega)

- ALLW_w4sigmpro (1E) The [W4] magnitude uncertainty of the AllWISE source (mag, Vega)
- ALLW_W1_W2 (1D) The [W1 – W2] colour of the AllWISE source (mag, Vega)
- ALLW_cc_flags (4A) A set of ‘contamination and confusion’ quality flags for the AllWISE source
- ALLW_r_2mass (1E) The distance from the AllWISE source to the nearest 2MASS counterpart (arcsec)
- ALLW_pa_2mass (1E) The position angle of the nearest 2MASS counterpart from the AllWISE source (deg)
- ALLW_n_2mass (1J) The number of 2MASS counterparts within 3 arcsec of the AllWISE source
- ALLW_j_m_2mass (1E) The *J*-band magnitude of the nearest 2MASS counterpart to the AllWISE source (mag, Vega)
- ALLW_j_msig_2mass (1E) The *J*-band magnitude uncertainty of the nearest 2MASS counterpart (mag, Vega)
- ALLW_h_m_2mass (1E) The *H*-band magnitude of the nearest 2MASS counterpart (mag, Vega)
- ALLW_h_msig_2mass (1E) The *H*-band magnitude uncertainty of the nearest 2MASS counterpart (mag, Vega)
- ALLW_k_m_2mass (1E) The *K_s*-band magnitude of the 2MASS counterpart (mag, Vega)
- ALLW_k_msig_2mass (1E) The *K_s*-band magnitude uncertainty of the nearest 2MASS counterpart (mag, Vega)
- SDSS_* Properties of the best SDSS-DR13 photometric counterpart to the AllWISE source (taken from SDSS Collaboration et al. 2016)
- SDSS_RUN (1J) Element of the standard five-part SDSS photometric source identification descriptor
- SDSS_RERUN (3A) see above
- SDSS_CAMCOL (1J) see above
- SDSS_FIELD (1J) see above
- SDSS_ID (1J) see above
- SDSS_RA (1D) The Right Ascension of the SDSS-DR13 photometric counterpart (deg, J2000)
- SDSS_DEC (1D) The Declination of the SDSS-DR13 photometric counterpart (deg, J2000)
- SDSS_MODEL_MAG_u (1E) The ‘model magnitude’ of the optical counterpart to the X-ray source in the *u* band, assuming a 2 arcsec diameter fibre (mag, AB)
- SDSS_MODEL_MAG_g (1E) as above, but for the *g* band
- SDSS_MODEL_MAG_r (1E) as above, but for the *r* band
- SDSS_MODEL_MAG_i (1E) as above, but for the *i* band
- SDSS_MODEL_MAG_z (1E) as above, but for the *z* band
- SDSS_FIBER2MAG_u (1E) The ‘fiber magnitude’ of the optical counterpart to the X-ray source in the *u* band, assuming a 2 arcsec diameter fibre (mag, AB)
- SDSS_FIBER2MAG_g (1E) as above, but for the *g* band
- SDSS_FIBER2MAG_r (1E) as above, but for the *r* band
- SDSS_FIBER2MAG_i (1E) as above, but for the *i* band
- SDSS_FIBER2MAG_z (1E) as above, but for the *z* band
- DR12_* Properties extracted from the SDSS-DR12 spectroscopic catalogue (Alam et al. 2015)
- DR12_SURVEY (6A) The SDSS survey in which this spectrum was obtained
- DR12_PLATE (1J) The SDSS spectroscopic plate index
- DR12_MJD (1J) The date on which the last spectroscopic data for the plate were obtained (MJD, days)
- DR12_FIBERID (1J) The index of the spectroscopic fibre through which this spectrum was obtained
- DR12_PLUG_RA (1D) The Right Ascension of the spectroscopic fibre position (deg, J2000)
- DR12_PLUG_DEC (1D) The Declination of the spectroscopic fibre position (deg, J2000)
- DIST_PHOT_PLUG (1E) The distance between the SDSS-DR13 photometric catalogue position and the spectroscopic fibre position (arcsec)
- DR12_Z (1E) The best-fitting redshift solution computed by the SDSS-DR12 spectroscopic pipeline
- DR12_Z_ERR (1E) Uncertainty on the best redshift solution computed by the SDSS-DR12 spectroscopic pipeline
- DR12_ZWARNING (1J) Flag that is set to >0 when the pipeline has encountered problems during the spectral reduction and redshift fitting process
- DR12_SN_MEDIAN_ALL (1E) The median SNR (per pixel) of the spectrum
- DR12_RCHI2 (1E) The reduced χ^2 of the best redshift solution
- DR12_CLASS (6A) The broad spectral classification computed by the SDSS-DR12 spectroscopic pipeline
- DR12_SUBCLASS (21A) The detailed spectral classification computed by the SDSS-DR12 spectroscopic pipeline
- DR12_NSPECOBS (1J) Number of SDSS-DR12 spectra that are associated with this object
- DR12_RUN2D (6A) Version code for the lower level SDSS spectral reduction pipeline used to process this spectrum
- DR12_RUN1D (6A) Version code for the higher SDSS spectral reduction software used to process this spectrum
- DR7Q_* VI information for this spectrum extracted from the SDSS-DR7Q quasar catalogue (Schneider et al. 2010)
- DR7Q_MEMBER (1B) Flag, set to 1 if a VI for this spectrum was part of the SDSS-DR7Q catalogue
- DR7Q_Z_VI (1E) VI redshift from SDSS-DR7Q
- DR12Q_* VI information for this spectrum extracted from the SDSS-DR12Q quasar catalogue (Pâris et al. 2017)
- DR12Q_MEMBER (1B) Flag, set to 1 if a VI for this spectrum was part of the SDSS-DR12Q catalogue
- DR12Q_Z_VI (1E) VI redshift from SDSS-DR12Q
- DR12Q_CLASS_PERSON (1J) VI spectral classification from SDSS-DR12Q
- DR12Q_Z_CONF_PERSON (1J) VI redshift confidence from SDSS-DR12Q
- DR12Q_ORIGIN (8A) Whether this spectrum appeared in the SDSS-DR12Q main catalogue, the superset catalogue or the supplementary bad spectrum catalogue
- A07_* VI information for this spectrum extracted from Anderson et al. (2007)
- A07_MEMBER (1B) Flag, set to 1 if a VI for this spectrum was part of the A07 catalogue
- A07_Z (1E) VI redshift from A07
- A07_CLASS (6A) VI spectral classification from A07
- A07_FXcor (1E) X-ray flux estimate (0.1–2.4 keV) from A07 ($\text{erg cm}^{-2} \text{s}^{-1}$)
- A07_logLX (1E) X-ray luminosity estimate (0.1–2.4 keV) from A07 ($\log_{10}[\text{erg s}^{-1}]$)
- P10_* VI information for this spectrum extracted from Plotkin et al. (2010)
- P10_MEMBER (1B) Flag, set to 1 if a VI for this spectrum was part of the Plotkin et al. (2010) catalogue
- P10_Z (1E) VI redshift from Plotkin et al. (2010)
- P10_ZSP (1A) VI spectral classification from Plotkin et al. (2010)
- P10_CONF (1A) Original VI confidence measure from Plotkin et al. (2010)

- P10_Z_CONF (1J) Derived VI confidence measure from Plotkin et al. (2010)
- D17_* VI information for this spectrum produced by the SPIDERS team
- D17_MEMBER (1B) Flag, set to 1 if at least one VI was carried out by the SPIDERS team
- D17_NINSPECTORS (1J) Number of SPIDERS visual inspectors who examined the spectrum
- D17_Z (1E) Redshift derived from the VIs carried out by the SPIDERS team
- D17_Z_CONF (1J) Redshift confidence derived from the VIs carried out by the SPIDERS team
- D17_CLASS (8A) Spectral class derived from the VIs carried out by the SPIDERS team
- D17_RECONCILED (1J) Flag, set to 1 if manual-reconciliation of the SPIDERS team VIs was carried out
- VI_BEST (7A) The origin of the ‘best’ VI (one of DR7Q, DR12Q, AO7, P10 or SPIDERS)
- Z_BEST (1E) The best VI redshift
- Z_CONF_BEST (1J) The confidence of the best VI redshift
- Z_DIFF (1E) The difference between the SDSS-DR12 pipeline redshift and the best VI redshift
- CLASS_BEST (8A) The best VI spectral classification
- DL_BEST (1E) The luminosity distance derived from the best VI redshift (Mpc)
- RASS_LOGLX (1E) Estimate of the X-ray luminosity (corrected for Galactic absorption) in the rest-frame 0.1–2.4 keV band ($\log_{10}[\text{erg s}^{-1}]$)
- RASS_LOGLX_ERR (1E) Statistical uncertainty on the X-ray luminosity ($\log_{10}[\text{erg s}^{-1}]$)
- RASS_LOGLX_05_2 (1E) Estimate of the X-ray luminosity (corrected for Galactic absorption) in the rest-frame 0.5–2 keV band ($\log_{10}[\text{erg s}^{-1}]$)

E2 Catalogue of XMMSL sources with spectroscopy in SDSS-DR12

See Section 3.5 for a description of how this catalogue was created. The RASS+SDSS-DR12 catalogue contains 1482 entries, with the following columns:

- XMMSL_* X-ray source properties derived from the XMMSL source catalogue (Saxton et al. 2008)
- XMMSL_UNIQUE_SRCNAME (40A) XMMSL source identifier
- XMMSL_OBSID (10A) *XMM-Newton* observation identifier of the slew in which the source was detected
- XMMSL_RA (1D) Right Ascension of X-ray source (deg, J2000)
- XMMSL_DEC (1D) Declination of X-ray source (deg, J2000)
- XMMSL_RADEC_ERR (1D) 1σ uncertainty on X-ray position (arcsec)
- XMMSL_DATE_OBS (19A) Date of start of *XMM-Newton* slew (YYYY-MM-DDThh:mm:ss)
- XMMSL_DATE_END (19A) Date of end of *XMM-Newton* slew (YYYY-MM-DDThh:mm:ss)
- XMMSL_SCTS_FULL (1E) Background-subtracted source counts in the 0.2–12 keV energy band (cts)
- XMMSL_SCTS_FULL_ERR (1E) Statistical uncertainty on above (cts)
- XMMSL_SCTS_SOFT (1E) Same as above but for the 0.2–2 keV band

- XMMSL_SCTS_SOFT_ERR (1E) Same as above but for the 0.2–2 keV band
- XMMSL_SCTS_HARD (1E) Same as above but for the 2–12 keV band
- XMMSL_SCTS_HARD_ERR (1E) Same as above but for the 2–12 keV band
- XMMSL_BG_MAP_FULL (1E) Estimated number of background counts within source extraction aperture in the 0.2–12 keV band (cts)
- XMMSL_BG_MAP_HARD (1E) Same as above but for the 0.2–2 keV band
- XMMSL_BG_MAP_SOFT (1E) Same as above but for the 2–12 keV band
- XMMSL_EXP_MAP_FULL (1E) Effective exposure time in the 0.2–12 keV band (s)
- XMMSL_EXP_MAP_SOFT (1E) Same as above but for the 0.2–2 keV band
- XMMSL_EXP_MAP_HARD (1E) Same as above but for the 2–12 keV band
- XMMSL_RATE_FULL (1E) Background-subtracted and vignetting-corrected count rate in the 0.2–12 keV energy band (ct s^{-1})
- XMMSL_RATE_FULL_ERR (1E) Statistical uncertainty on above (ct s^{-1})
- XMMSL_RATE_SOFT (1E) Same as above but for the 0.2–2 keV band
- XMMSL_RATE_SOFT_ERR (1E) Same as above but for the 0.2–2 keV band
- XMMSL_RATE_HARD (1E) Same as above but for the 2–12 keV band
- XMMSL_RATE_HARD_ERR (1E) Same as above but for the 2–12 keV band
- XMMSL_FLUX_FULL (1E) Estimated X-ray flux (corrected for full Galactic column) in the 0.2–12 keV energy band ($\text{erg cm}^{-2} \text{s}^{-1}$)
- XMMSL_FLUX_FULL_ERR (1E) Statistical uncertainty on above ($\text{erg cm}^{-2} \text{s}^{-1}$)
- XMMSL_FLUX_SOFT (1E) Same as above but for the 0.2–2 keV band
- XMMSL_FLUX_SOFT_ERR (1E) Same as above but for the 0.2–2 keV band
- XMMSL_FLUX_HARD (1E) Same as above but for the 2–12 keV band
- XMMSL_FLUX_HARD_ERR (1E) Same as above but for the 2–12 keV band
- XMMSL_HR1 (1E) Hardness ratio compute between the 0.2–2 keV (soft) and 2–12 keV (hard) energy bands
- XMMSL_HR1_ERR (1E) Statistical uncertainty on above
- XMMSL_RASS_DIST (1E) Distance to nearest RASS source (arcsec)
- XMMSL_GAL_LONG (1D) Galactic longitude of X-ray source (deg)
- XMMSL_GAL_LAT (1D) Galactic latitude of X-ray source (deg)
- XMMSL_LOGGALNH (1E) Logarithm of Galactic column density in direction of X-ray source ($\log_{10}[\text{cm}^{-2}]$)
- XMMSL_ALLW_DIST (1D) Distance between X-ray position and position of best AllWISE counterpart (arcsec)
- XMMSL_ALLW_BIAS (1D) Bayesian prior (or bias) factor, $\Pi(x)$, of the best AllWISE counterpart, derived from its location in $[W2]-[W1 - W2]$ parameter space

- XMMSL_ALLW_BFPOST (1D) Bayesian posterior probability of the best X-ray \rightarrow AllWISE association, before considering bias factor
- XMMSL_ALLW_POST (1D) Bayesian posterior probability of the best X-ray \rightarrow AllWISE association, including the bias factor
- ALLW_ra (1D) See Section E1
- ALLW_dec (1D) See Section E1
- ALLW_sigra (1E) See Section E1
- ALLW_sigdec (1E) See Section E1
- ALLW_w1mpro (1E) See Section E1
- ALLW_w1sigmpro (1E) See Section E1
- ALLW_w2mpro (1E) See Section E1
- ALLW_w2sigmpro (1E) See Section E1
- ALLW_w3mpro (1E) See Section E1
- ALLW_w3sigmpro (1E) See Section E1
- ALLW_w4mpro (1E) See Section E1
- ALLW_w4sigmpro (1E) See Section E1
- ALLW_W1_W2 (1D) See Section E1
- ALLW_cc_flags (4A) See Section E1
- ALLW_r_2mass (1E) See Section E1
- ALLW_pa_2mass (1E) See Section E1
- ALLW_n_2mass (1J) See Section E1
- ALLW_j_m_2mass (1E) See Section E1
- ALLW_j_msig_2mass (1E) See Section E1
- ALLW_h_m_2mass (1E) See Section E1
- ALLW_h_msig_2mass (1E) See Section E1
- ALLW_k_m_2mass (1E) See Section E1
- ALLW_k_msig_2mass (1E) See Section E1
- SDSS_RUN (1J) See Section E1
- SDSS_RERUN (3A) See Section E1
- SDSS_CAMCOL (1J) See Section E1
- SDSS_FIELD (1J) See Section E1
- SDSS_ID (1J) See Section E1
- SDSS_RA (1D) See Section E1
- SDSS_DEC (1D) See Section E1
- SDSS_MODEL_MAG_u (1E) See Section E1
- SDSS_MODEL_MAG_g (1E) See Section E1
- SDSS_MODEL_MAG_r (1E) See Section E1
- SDSS_MODEL_MAG_i (1E) See Section E1
- SDSS_MODEL_MAG_z (1E) See Section E1
- SDSS_FIBER2MAG_u (1E) See Section E1
- SDSS_FIBER2MAG_g (1E) See Section E1
- SDSS_FIBER2MAG_r (1E) See Section E1
- SDSS_FIBER2MAG_i (1E) See Section E1
- SDSS_FIBER2MAG_z (1E) See Section E1
- DR12_SURVEY (6A) See Section E1
- DR12_PLATE (1J) See Section E1
- DR12_MJD (1J) See Section E1
- DR12_FIBERID (1J) See Section E1
- DR12_PLUG_RA (1D) See Section E1
- DR12_PLUG_DEC (1D) See Section E1
- DIST_PHOT_PLUG (1E) See Section E1
- DR12_Z (1E) See Section E1
- DR12_Z_ERR (1E) See Section E1
- DR12_ZWARNING (1J) See Section E1
- DR12_SN_MEDIAN_ALL (1E) See Section E1
- DR12_RCHI2 (1E) See Section E1
- DR12_CLASS (6A) See Section E1
- DR12_SUBCLASS (21A) See Section E1
- DR12_NSPECOBS (1J) See Section E1
- DR12_RUN2D (6A) See Section E1
- DR12_RUN1D (6A) See Section E1
- DR7Q_MEMBER (1B) See Section E1
- DR7Q_Z_VI (1E) See Section E1
- DR12Q_MEMBER (1B) See Section E1
- DR12Q_Z_VI (1E) See Section E1
- DR12Q_CLASS_PERSON (1J) See Section E1
- DR12Q_Z_CONF_PERSON (1J) See Section E1
- DR12Q_ORIGIN (8A) See Section E1
- A07_MEMBER (1B) See Section E1
- A07_Z (1E) See Section E1
- A07_CLASS (6A) See Section E1
- A07_FXcor (1E) See Section E1
- A07_logLX (1E) See Section E1
- P10_MEMBER (1B) See Section E1
- P10_Z (1E) See Section E1
- P10_ZSP (1A) See Section E1
- P10_CONF (1A) See Section E1
- P10_Z_CONF (1J) See Section E1
- D17_MEMBER (1B) See Section E1
- D17_NINSPECTORS (1J) See Section E1
- D17_Z (1E) See Section E1
- D17_Z_CONF (1J) See Section E1
- D17_CLASS (8A) See Section E1
- D17_RECONCILED (1J) See Section E1
- VI_BEST (7A) See Section E1
- Z_BEST (1E) See Section E1
- Z_CONF_BEST (1J) See Section E1
- Z_DIFF (1E) See Section E1
- CLASS_BEST (8A) See Section E1
- DL_BEST (1E) See Section E1
- XMMSL_LOGLX_FULL (1E) Estimate of the X-ray luminosity (corrected for Galactic absorption) in the rest-frame 0.2–12 keV band ($\log_{10}[\text{erg s}^{-1}]$)
- XMMSL_LOGLX_FULL_ERR (1E) Statistical uncertainty on the X-ray luminosity in the rest-frame 0.2–12 keV band ($\log_{10}[\text{erg s}^{-1}]$)
- XMMSL_LOGLX_SOFT (1E) As above, but for the rest-frame 0.2–2 keV band
- XMMSL_LOGLX_SOFT_ERR (1E) As above, but for the rest-frame 0.2–2 keV band
- XMMSL_LOGLX_HARD (1E) As above, but for the rest-frame 2–12 keV band
- XMMSL_LOGLX_HARD_ERR (1E) As above, but for the rest-frame 2–12 keV band
- XMMSL_LOGLX_05_2 (1E) As above, but for the rest-frame 0.5–2 keV band
- XMMSL_LOGLX_05_2_ERR (1E) As above, but for the rest-frame 0.5–2 keV band
- XMMSL_LOGLX_05_10 (1E) As above, but for the rest-frame 0.5–10 keV band
- XMMSL_LOGLX_05_10_ERR (1E) As above, but for the rest-frame 0.5–10 keV band

E3 Catalogue of RASS sources (SPIDERS_RASS_AGN) to be targeted in eBOSS/TDSS/SPIDERS

See Section 3.3 for a description of how this catalogue was created. The SPIDERS_RASS_AGN catalogue contains 9028 entries, with the following columns:

- RA (1D) The Right Ascension of the SDSS-DR13 photometric counterpart to the X-ray source (deg, J2000; SDSS Collaboration et al. 2016)

- DEC (1D) The Declination of the SDSS-DR13 photometric counterpart to the X-ray source (deg, J2000; SDSS Collaboration et al. 2016)

- FIBER2MAG (5E) The ‘fiber magnitude’ of the optical counterpart to the X-ray source in the *ugriz* bands, assuming a 2 arcsec diameter fibre (mag, AB; SDSS Collaboration et al. 2016)

- RUN (1J) Element of the standard five-part SDSS-DR13 source identification descriptor for the optical counterpart to the X-ray source (SDSS Collaboration et al. 2016)

- RERUN (3A) See above

- CAMCOL (1J) See above

- FIELD (1J) See above

- ID (1J) See above

- PRIORITY (1J) Priority assigned to this target for the purposes of the eBOSS/TDSS/SPIDERS tiling process (0 = highest priority, 99 = lowest priority)

- XRAY_SRC_NAME (16A) Source identifier from the RASS catalogue (Voges et al. 1999, 2000)

- XRAY_RA (1D) The Right Ascension of the X-ray source (deg, J2000; Voges et al. 1999, 2000)

- XRAY_DEC (1D) The Declination of the X-ray source (deg, J2000; Voges et al. 1999, 2000)

- XRAY_RADEC_ERR (1E) The positional uncertainty of the X-ray source (arcsec; Voges et al. 1999, 2000)

- XRAY_FLUX (1D) The X-ray flux of the source ($\text{erg cm}^{-2} \text{s}^{-1}$), corrected for Galactic absorption, in the 0.1–2.4 keV band (see Section A1)

- XRAY_DET_ML (1D) The detection likelihood for the X-ray source (Voges et al. 1999, 2000)

- BAYES_POSTERIOR_PROB (1E) The Bayesian posterior probability (P_{post}) of the association between the X-ray and AllWISE source

- ALLWISE_DESIGNATION (19A) The source identifier for the AllWISE counterpart to the X-ray source (Cutri 2013)

- ALLWISE_RA (1D) The Right Ascension of the AllWISE counterpart to the X-ray source (deg, J2000; Cutri 2013)

- ALLWISE_DEC (1D) The Declination of the AllWISE counterpart to the X-ray source (deg, J2000; Cutri 2013)

- ALLWISE_W2MPRO (1E) The [W2] magnitude of the AllWISE counterpart to the X-ray source (mag, Vega; Cutri 2013)

- ALLWISE_W1_W2 (1E) The [W1 – W2] colour of the AllWISE counterpart to the X-ray source (mag, Vega)

Note that this catalogue also formed part of the SDSS-DR13, and can be downloaded from https://data.sdss.org/sas/dr13/eboss/spiders/target/spiderstargetAGN-SPIDERS_RASS_AGN-v2.1.fits.

E4 Catalogue of XMMSL sources (SPIDERS_XMMSL_AGN) to be targeted in eBOSS/TDSS/SPIDERS

See Section 3.4 for a description of how this catalogue was created. The SPIDERS_XMMSL_AGN catalogue contains 873 entries, with the following columns:

- RA (1D) See Section E3

- DEC (1D) See Section E3

- FIBER2MAG (5E) See Section E3

- RUN (1J) See Section E3

- RERUN (3A) See Section E3

- CAMCOL (1J) See Section E3

- FIELD (1J) See Section E3

- ID (1J) See Section E3

- PRIORITY (1J) See Section E3

- XRAY_SRC_NAME (24A) Source identifier from the XMMSL catalogue (column UNIQUE_SRCNAME in official XMMSL catalogue)

- XRAY_OBSID (10A) The *XMM-Newton* Observation Identifier of the observation in which this detection was made

- XRAY_RA (1D) The Right Ascension of the X-ray source (deg, J2000)

- XRAY_DEC (1D) The Declination of the X-ray source (deg, J2000)

- XRAY_RADEC_ERR (1E) The positional uncertainty of the X-ray source (arcsec)

- BAYES_POSTERIOR_PROB (1E) See Section E3

- ALLWISE_DESIGNATION (19A) See Section E3

- ALLWISE_RA (1D) See Section E3

- ALLWISE_DEC (1D) See Section E3

- ALLWISE_W2MPRO (1E) See Section E3

- ALLWISE_W1_W2 (1E) See Section E3

- MISSING_IN_V3p1 (1B) Set to 1 if this source was excluded in an earlier version of the catalogue (see Section 2.2 for details)

Note that an older version of this catalogue (v3.1, containing 819 entries, see caveats in section 2.2) formed part of the SDSS-DR13, and can be downloaded from https://data.sdss.org/sas/dr13/eboss/spiders/target/spiderstargetAGN-SPIDERS_XMMSL_AGN-v3.1.fits. The revised version of the catalogue presented here, which fixes these problems, is only available from the MPE X-ray surveys website: http://www.mpe.mpg.de/XraySurveys/SPIDERS/SPIDERS_AGN.

E5 Catalogue of RASS sources targeted in SEQUELS

See Section D for a description of how this catalogue was created. This catalogue contains 627 entries, with the following columns:

- RA (1D) The Right Ascension of the SDSS-DR8 photometric counterpart to the X-ray source (deg, J2000; Aihara et al. 2011)

- DEC (1D) The Declination of the SDSS-DR8 photometric counterpart to the X-ray source (deg, J2000; Aihara et al. 2011)

- FIBER2MAG (5E) The ‘fiber magnitude’ of the optical counterpart to the X-ray source in the *ugriz* bands, assuming a 2 arcsec diameter fibre (mag, AB; Aihara et al. 2011)

- RUN (1J) Element of the standard five-part SDSS-DR8 source identification descriptor for the optical counterpart to the X-ray source (Aihara et al. 2011)

- RERUN (3A) see above

- CAMCOL (1J) see above

- FIELD (1J) see above

- ID (1J) see above

- PRIORITY (1J) See Section E3

- XRAY_SRC_NAME (16A) See Section E3

- XRAY_RA (1D) See Section E3

- XRAY_DEC (1D) See Section E3

- XRAY_RADEC_ERR (1E) See Section E3

- XRAY_FLUX (1D) See Section E3

- XRAY_DET_ML (1D) See Section E3

- BAYES_POSTERIOR_PROB (1E) The Bayesian posterior probability (P_{post}) of the association between the X-ray and optical source.

This catalogue formed part of the SDSS-DR13, and so can also be downloaded from https://data.sdss.org/sas/dr13/eboss/spiders/target/spiderstargetSequelsAGN-SPIDERS_RASS_AGN-v1.1.fits.

E6 Catalogue of bright X-ray sources used as an astrometric reference sample

For completeness, we provide here the astrometric reference catalogue of bright X-ray sources (from *Chandra*, *XMM-Newton* and *Swift* serendipitous catalogues) described in Section 4.1. This catalogue contains 4752 entries with the following columns:

- XBEST_CAT (5A) A code (one of ‘CSC’, ‘3XMM’ or ‘SXPS’) identifying the catalogue from which the ‘best’ X-ray properties of this source were derived
- XBEST_RA (1D) The ‘best’ X-ray determined Right Ascension of this source (deg, J2000)
- XBEST_DEC (1D) The ‘best’ X-ray determined Declination of this source (deg, J2000)
- XBEST_FLUX_SOFT (1D) An estimate of the X-ray flux in the ‘soft’ energy band (0.2–2 keV for CSC and 3XMM, 0.3–2 keV for 1SXPS; all with units $\text{erg cm}^{-2} \text{s}^{-1}$)
- XBEST_FLUX_FULL (1D) An estimate of the X-ray flux in the ‘full’ energy band (0.2–7 keV for CSC, 0.2–12 keV for 3XMM, 0.3–10 keV for 1SXPS; all with units $\text{erg cm}^{-2} \text{s}^{-1}$)
- XBEST_POSERR (1D) The positional uncertainty in the form given by the parent catalogue (error ellipse half-major axis for CSC, 1σ error radius for 3XMM and 90 per cent containment radius for 1SXPS; all with units of arcsec)
- CSC_name (20A) Name of this source (if and) as it appears in the CSC catalogue (Evans et al. 2010)
- XMM_NAME (21A) Name of this source (if and) as it appears in the 3XMM catalogue (Rosen et al. 2016)
- SXPS_Name (22A) Name of this source (if and) as it appears in the 1SXPS catalogue (Evans et al. 2014)
- XBEST_NWISE (1J) Number of potential AllWISE associations within the search radius (i.e. within 5 arcsec for CSC and 3XMM, and within 10 arcsec for 1SXPS)
- DIST_XBEST_ALLW (1D) Distance from the best X-ray position to the nearest AllWISE counterpart (arcsec)
- ALLW_designation (20A) The source identifier for the nearest AllWISE counterpart to the X-ray source (Cutri 2013)
- ALLW_ra (1D) The Right Ascension of the nearest AllWISE counterpart to the X-ray source (deg, J2000; Cutri 2013)
- ALLW_dec (1D) The Declination of the nearest AllWISE counterpart to the X-ray source (deg, J2000; Cutri 2013)
- ALLW_w1mpro (1D) The [W1] magnitude of the nearest AllWISE counterpart to the X-ray source (mag, Vega; Cutri 2013)
- ALLW_w1sigmpro (1D) The [W1] magnitude uncertainty of the nearest AllWISE counterpart to the X-ray source (mag, Vega; Cutri 2013)

- ALLW_w2mpro (1D) The [W2] magnitude of the nearest AllWISE counterpart to the X-ray source (mag, Vega; Cutri 2013)
- ALLW_w2sigmpro (1D) The [W2] magnitude uncertainty of the nearest AllWISE counterpart to the X-ray source (mag, Vega; Cutri 2013)
- ALLW_W1_W2 (1D) The [W2] magnitude uncertainty of the nearest AllWISE counterpart to the X-ray source (mag, Vega; Cutri 2013)

¹Max-Planck-Institut für extraterrestrische Physik, Giessenbachstr, D-85748 Garching, Germany

²Dipartimento di Fisica e Astronomia, Università di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italy

³INAF-Osservatorio Astronomico di Bologna, via Ranzani 1, I-40127 Bologna, Italy

⁴Millennium Institute of Astrophysics, Vicuña, MacKenna 4860, 7820436 Macul, Santiago, Chile

⁵Pontificia Universidad Católica de Chile, Instituto de Astrofísica, Casilla 306, Santiago 22, Chile

⁶Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195, USA

⁷Department of Astronomy and Astrophysics, 525 Davey Lab, The Pennsylvania State University, University Park, PA 16802, USA

⁸Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA

⁹Department of Physics, 104 Davey Lab, The Pennsylvania State University, University Park, PA 16802, USA

¹⁰Department of Applied Mathematics and Statistics, Johns Hopkins University, 3400 N. Charles St, MD 21218, USA

¹¹Department of Computer Science, Johns Hopkins University, 3400 N. Charles St, MD 21218, USA

¹²Department of Physics and Astronomy, Johns Hopkins University, 3400 N. Charles St, Baltimore, MD 21218, USA

¹³CNRS, IRAP, 9 Av. colonel Roche, BP 44346, F-31028 Toulouse cedex 4, France

¹⁴Université de Toulouse, UPS-OMP, IRAP, Toulouse, France

¹⁵Harvard-Smithsonian Center for Astrophysics, 60 Garden St, MS #20, Cambridge, MA 02138, USA

¹⁶Department of Physics and Astronomy, University of Wyoming, Laramie, WY 82071, USA

¹⁷Institute of Cosmology and Gravitation, Dennis Sciama Building, University of Portsmouth, Portsmouth PO1 3FX, UK

¹⁸Leibniz-Institut für Astrophysik Potsdam, An der Sternwarte 16, D-14482 Potsdam, Germany

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.