SOURCE MATCHING IN THE SDSS AND RASS: WHICH GALAXIES ARE REALLY X-RAY SOURCES?

John K. Parejko¹, Anca Constantin¹, Michael S. Vogeley¹, and Fiona Hoyle²

¹ Department of Physics, Drexel University, Philadelphia, PA 19104, USA; parejko@drexel.edu
² Department of Physics and Astronomy, Widener University, Chester, PA 19013, USA

Received 2007 June 19; accepted 2007 August 29; published 2007 November 29

ABSTRACT

The current view of galaxy formation holds that all massive galaxies harbor a massive black hole at their center, but that these black holes are not always in an actively accreting phase. X-ray emission is often used to identify accreting sources, but for galaxies that are not harboring quasars (low-luminosity active galaxies), the X-ray flux may be weak, or obscured by dust. To aid in the understanding of weakly accreting black holes in the local universe, a large sample of galaxies with X-ray detections is needed. We cross-match the ROSAT All Sky Survey (RASS) with galaxies from the Sloan Digital Sky Survey Data Release 4 (SDSS DR4) to create such a sample. Because of the high SDSS source density and large RASS positional errors, the cross-matched catalog is highly contaminated by random associations. We investigate the overlap of these surveys and provide a statistical test of the validity of RASS–SDSS galaxy cross-matches. The SDSS quasars provide a test of our cross-match validation scheme, as they have a very high fraction of true RASS matches. We find that the number of true matches between the SDSS main galaxy sample and the RASS is highly dependent on the optical spectral classification of the galaxy; essentially no star-forming galaxies are detected, while more than 0.6% of narrow-line Seyferts are detected in the RASS. Also, galaxies with ambiguous optical classification have a surprisingly high RASS detection fraction. This allows us to further constrain the SEDs of low-luminosity active galaxies. Our technique is quite general, and can be applied to any cross-matching between surveys with well-understood positional errors.

Key words: galaxies: active – quasars: general – X-rays: galaxies – X-rays: general

Online-only material: color figures, machine-readable table

1. INTRODUCTION

Distinguishing the processes that contribute to the emission from the centers of galaxies is vital to understanding the co-evolution of galaxies and their central black holes. Among nearby galaxies, a large fraction of central emission sources are of ambiguous nature (Ho et al. 1997); emission-lines in optical spectra of many galaxies seem to reflect a mix of behavior between bona-fide accretion (Seyfert-like) and active star formation (H II-like). In order to discriminate between the various possible ionicization mechanisms and penetrate the obscuring dust layers that encircle these sources, we need observations at multiple wavelengths. In particular, X-rays are less prone to dust absorption and thus can be used to distinguish between accretion sources and emission from young, hot stars. This can clarify the observed optical emission spectra and allow us to better describe the central accretion sources in low-luminosity active galactic nuclei (AGN).

For an accurate census of the local galactic population, one must study a statistically significant number of sources. The Sloan Digital Sky Survey (SDSS; York et al. 2000) provides the largest sample of galaxies with spectra which allow emission-line classification of central sources. The ROSAT All-Sky Survey (RASS; Voge et al. 1999, 2000) is the widest and deepest survey of the X-ray sky. The SDSS and RASS are well matched in terms of depth, but have quite different astrometry and spatial resolution. Previous studies matching a variety of SDSS and RASS sources include analyzing the X-ray properties of spectroscopically confirmed quasars (Anderson et al. 2003, 2007), generating an X-ray detected galaxy-cluster catalog (Popesso et al. 2004), searching for optically unidentified neutron stars (Agueros et al. 2006), and surveying the multi-wavelength properties of SDSS galaxies (Obrč et al. 2006).

Because broad-line quasars are expected to be strong X-ray sources, one would expect a large number of matches between the RASS and the SDSS for these objects. Anderson et al. (2003) characterized the RASS properties of spectroscopically identified broad-line quasars from the SDSS as well as some narrow-line sources. They qualitatively discuss the likelihood that a given RASS–SDSS match is a true match and include “normal” (non- or weakly emitting) galaxies as a comparison of what a weak correlation would look like. They study more than 1000 RASS–SDSS quasar/AGN and briefly discuss a few properties of the sample. Their sample reproduces the expected nonlinear optical/X-ray (2500 Å/2 keV) relationship among broad-line sources. The follow-up study, Anderson et al. (2007), examines ~7000 sources with similar results.

A different investigation involves identifying RASS sources with no obvious optical counterpart. For example, this is useful for finding optically dim neutron stars. Agüeros et al. (2006) identified all SDSS sources within four times the positional error of each RASS source. They then removed from their catalog any RASS source with an SDSS match which could have produced the X-ray flux. After removing objects with NED identifications, visually identified bad fields, and known galaxy-clusters, 11 RASS sources with no plausible SDSS optical counterpart remained. They claim that this number is consistent with the number of isolated neutron stars expected in the SDSS field. Studying poorly understood matching samples in this way can clarify whether the sample includes primarily true matches or primarily false matches.

A recent comparison of RASS and SDSS in a multi-wavelength study (Obrč et al. 2006) identified 267 RASS matches within 30′′ of SDSS DR1 main sample galaxies (Strauss et al. 2002; Abazajian et al. 2003). They list a false association fraction of ~9% (computed statistically based on the RASS...
source density) and also show the positions of their galaxies on an optical emission-line classification diagram (the BPT diagram; Baldwin et al. 1981). They did not investigate known bad matches (as in Agüeros et al. 2006), nor did they elaborate on the positions of the RASS detected galaxies on their BPT diagram.

The ROSAT All Sky Survey was produced from data acquired in ROSAT’s scanning mode, but ROSAT also performed many individual targeted observations, resulting in several pointed catalogs. These catalogs were generated from serendipitous source discoveries made during individual targeted observations. Because of this, they contain a large number of sources in very small fields scattered over the sky with highly varying exposure durations, making source upper limits difficult to compute. Previous studies (e.g., Suchkov et al. 2006) have examined the properties of SDSS quasars found in these catalogs.

Schulte-Ladbeck et al. (2005) looked at star-forming galaxies in the SDSS DR1 and matched them to several different ROSAT catalogs, including the RASS. Their final results involve 14 star-forming galaxies which they claim to be X-ray sources (four of which were previously studied). We were not able to determine exactly which catalog they used in their published results. However, only three of the galaxies listed in their Table 1 (IC 2233, NGC 4030, and NGC 4900) have RASS counterparts. For these galaxies, which are very nearby and thus large in projected size, the RASS error circle encompasses multiple SDSS sources which could also produce the measured X-ray flux, e.g., bright stars, H II regions, and the galactic core. Therefore, it is difficult to discern whether these galaxies represent X-ray bright star-forming galaxies or not. Some star-forming galaxies are expected to be X-ray emitters, but whether these galaxies are actually detected in RASS remains to be seen.

The XMM-Newton and Chandra X-ray satellites both provide much improved pointing, resolution, and depth over ROSAT, but their fields of view are quite small. Both have produced serendipitous source catalogs similar to the ROSAT pointed catalogs mentioned above. The initial XMM serendipitous source catalog was compared with the USNO A2.0 optical catalog (Georgakakis et al. 2006) to find 46 optically identified non-AGN galaxies with substantial X-ray flux. Hornschemeier et al. (2005) matched serendipitous source detections in Chandra with SDSS DR2 (Abazajian et al. 2004) to find 42 X-ray-emitting galaxies of a variety of types. The XMM-slew survey (Freyberg et al. 2005) aims to solve the field of view and uniformity problems by taking data during spacecraft slews between targets. It will produce an all-sky map of equivalent depth to RASS, with more than six times better resolution and pointing accuracy, in roughly six years.

In this paper, we investigate the accuracy of matching RASS sources with SDSS galaxies. In Section 2 we describe the data sets used in this study, including the systematics of selecting an appropriate galaxy sample from SDSS. The details of the cross-matching procedure and the statistical methods are described in Section 3 and the final matched data sets, separated by galaxy spectroscopic class, are detailed in Section 4. We find that a RASS–SDSS galaxy match cannot be trusted to represent the galaxy’s true X-ray flux without first identifying the galaxy’s spectral type. Section 5 provides a preliminary analysis of the new XMM-slew catalog and shows its utility in clarifying the presence of X-ray sources in galaxies.

2. DATA

2.1. SDSS

Our focus is on the main galaxy spectroscopic sample which includes all galaxies with Petrosian r magnitudes brighter than 17.77 with the exception of those not observed due to fiber collision. Because of the size of the fiber-plugs, spectroscopic targets for a single plate must be separated by at least 55″. This was more of a problem in DR1 and DR2 before overlapping plates and follow-up observations filled in many of the missing objects. The complete SDSS spectroscopic catalog includes more galaxies with spectra than just the main galaxy sample. We restrict ourselves to the main galaxy sample to avoid sample bias; some SDSS objects were selected for spectroscopy due to their proximity to FIRST radio sources (Becker et al. 1995) and/or RASS X-ray sources. See the Appendix for details on our SDSS source selection process, and the importance of using the main galaxy sample in cross-matching studies.

SDSS studies at MPA/JHU produced a catalog3 of secondary source products generated from the SDSS spectroscopic data (see Brinchmann et al. 2004, for the DR2 catalog paper). This catalog includes simultaneous measurements of the emission and absorption line profiles. The complete catalog includes all objects in SDSS (regardless of magnitude) that are spectroscopically identified as galaxies; sources with emission-line widths greater than 1000 km s−1 are not included (thus all objects identified as Seyferts and LINERs in this paper are type 2 objects). We restrict ourselves to the intersection of the MPA/JHU catalog and the SDSS DR4 main galaxy sample described above.

2.1.1. SDSS Galaxies

We classify galaxies based on their optical emission-line properties. Galaxies showing at least a 2σ detection of flux in the emission-features Hα, Hβ, [O iii], [N ii], [S ii], and [O i] are classified as emission-line galaxies, while those that show some but not all of these lines are called “unclassifiable” galaxies. The strong line emitters are further separated into

3 Data catalogs from SDSS studies are available from MPA/JHU http://www.mpa-garching.mpg.de/SDSS/.
of roughly 25 RASS varies between upper-limits. The average integration time per target in the RASS because we would eventually like to compute source together covering 92% of the sky. We restrict ourselves to the ′′ plus a 6 systematic error (1σ used to avoid systematic error) of 10–20 ′′ resulted in an astrometric positional error (1σ statistical error plus a 6′′ systematic error) of 10–20′′ (Figure 2). We show in Section 3.3 that the 6′′ systematic error is likely overestimated; 3′′ is likely more correct.

3.3. XMM-Newton Slew Survey

The RASS catalog is the current best compromise between width and depth for X-ray data, but it has limitations, as noted above. To produce an improved catalog, the X-ray Multi Mirror satellite (XMM-Newton) is collecting X-ray counts during slews between targeted observations. The first release of the XMM-Newton Slew Survey (XMM-slew; Freyberg et al. 2005) covers 6240 square degrees of sky, in narrow north–south slews, using the EPIC-pn CCD because of its large detector area, fast read-out rate and high sensitivity to hard X-rays. Although the average exposure time is only ~10 s for any given source, the large mirror area and sensitive detector make it nearly as deep as the RASS in the soft band (0.2–2 keV), and deeper and wider than any previous survey in the hard band (2–12 keV). The quoted 8′′ positional error along the slew direction is dominated by the accuracy of the attitude reconstruction. The EPIC-pn resolution of 4′′ is roughly a factor of 6 better than the RASS resolution; thus XMM-slew can resolve many of the confused RASS sources.

Two XMM-slew catalogs were released, a “total” catalog containing all detected sources, and a “clean” catalog with known bad sources removed and a higher detection threshold. We examine the clean sample in this study; it contains 2713 sources with detections in at least one band.
3. CROSS-MATCHING

Cross-matching two surveys is simple enough: count all objects separated by less than some threshold distance (in our case, 60″) as possible matches. But the validity of such a match depends on the differing sky coverage, sensitivity, positional accuracy, and spatial resolution of the two matched surveys. These differences lead to matches due to purely random associations, multiple cross-matches for single sources, and erroneous flux measurements due to contributions from multiple sources. For example, the ROSAT PSPC is more than an order of magnitude worse than the SDSS in both resolution and astrometry, and the SDSS source density is much higher. Understanding the RASS–SDSS galaxy sample is particularly difficult for sources that are not necessarily expected to be strong X-ray emitters, such as spectroscopically identified low-luminosity narrow-line AGN, passive or starburst galaxies.

In this section, we attempt to quantify the true and random components of RASS–SDSS cross-matches.

Figure 3 illustrates an example of the issues faced in matching RASS and SDSS. Here a RASS source overlaps two spectroscopically identified SDSS galaxies and is not centered on either of them. One of the galaxies hosts a quasar and thus is the likely source of the X-ray flux, while the other is identified as a star-forming galaxy (H II-type optical spectrum) and thus is expected to contribute little to the X-ray flux. If the quasar were unidentified—because it had no spectrum taken—the star-forming galaxy could have been considered the X-ray source. Another problem is that the center of the X-ray source does not coincide with any of the optical sources. This could be simply due to the astrometric errors in the RASS catalog (Figure 2), or to contributions to the total X-ray emission from the other quasar at the top of the image. This example is not singular: There are many such confusing matches in the RASS–SDSS galaxy sample because of the high SDSS source density. Also, this RASS source is relatively bright, and thus has better centroiding (positional error given as 8′′) than most RASS sources and was particularly easy to catch.

3.1. Obvious X-ray Emitters

When an SDSS object is the actual source of the RASS X-rays (a true match), the distance between the X-ray and optical source positions should be small. Some obvious choices for true matches are quasars and quasar candidates. To qualitatively assess whether these “obvious” choices are correct, we plot the distribution of distances between the center of the RASS and SDSS sources—the source-separation histogram—for these particular systems in Figure 4. The upper panel includes the following RASS-matched SDSS sources: spectroscopically identified quasars, sources with $u - g < 0.6$ (quasar candidates), a subset of the quasar candidates restricted to $15.5 < u < 21$ and “random match” between galaxies and the RASS, as described in the following section. The lower panel shows the $u$-magnitude versus source-separation distribution for blue sources. Note the clustering of points at small source separations for $15.5 < u < 21$, suggesting that these are true RASS–SDSS matches.

From the upper panel, spectroscopically identified quasars show an obvious peak at small source separations. “Blue” objects (all SDSS sources with $u - g < 0.5$), which include some objects in the “quasar” sample, have a peak at small separations as well as a prominent tail. The “blue2” sample (subset of “blue” with $15.5 < u < 21$) has a much smaller tail, suggesting a smaller fraction of incorrect matches. Out of these samples, spectroscopically identified quasars appear to represent the most reliable RASS–SDSS cross-match, with the fewest points with large separations.

Another possibility for “obvious” X-ray sources would be bright stars, such as those from the Tycho star catalog. Figure 7 of Voge et al. (1999) plots the source separation of RASS and Tycho sources, but the text does not describe the shape of the
distribution, beyond mentioning “chance coincidences” beyond 40”. Guillout et al. (1999) also plot the RASS/Tycho source separation for matches out to 150”, with an estimate of the random component. They also give a “best fit of a log normal to the distribution of the expected physical matches,” but no total model distribution. No rationale for the choice of a log-normal is given, nor do they provide a goodness-of-fit value. As bright stars can produce a strong off-axis component in the PSPC (possibly producing large source separations), we will not use them in our analysis.

3.2. Purely Random Matches

Incorrect cross-matches between catalogs are due to random associations between optical and X-ray sources. Previous work estimated the random contamination by comparing the source density of the two catalogs, which works well for samples with a small random contamination fraction. We model these incorrect matches by generating “offset” SDSS object catalogs and matching them to the RASS. We produced ten such offset catalogs each from the SDSS galaxy and quasar catalogs by offsetting all objects (either galaxies or quasars, respectively) from their true R.A. and decl. by a fixed amount in a fixed direction, with a different offset and direction for each offset catalog to reduce systematic effects. The maximum offset was 1° in R.A. and decl. This procedure preserves the on-sky source distribution of the SDSS, while moving sources far away from their original RASS associations. When these catalogs are matched to the RASS, the result is a linearly increasing source-separation histogram, \( dN/dr \propto r \); as the radius increases, more sources fall within the matching circle. We compare these random catalogs with our galaxy or quasar RASS matches to determine the fractional contamination by purely random associations.

3.3. Confirming Quasars

X-ray source positional measurements have independent, normally distributed errors in both planar components. This is analogous to darts thrown at a small target. The precision of each throw is known, but individual throws may have different precisions. The distribution of dart–target distances is given by a Rayleigh distribution having a probability density function (PDF),

\[
P(r) \propto \frac{r e^{-r^2/2\sigma^2}}{\sigma^2},
\]

with scale parameter, \( \sigma \), and separation distance, \( r \). In the case of X-ray measurements, the positional precision, \( \sigma \), is affected by the X-ray flux (reliability of centroiding depends on the number of X-rays) and the pointing accuracy and resolution of the measuring apparatus. The precision of each RASS source measurement is listed in the catalog as the positional error (Figure 2).

We reproduce the source-separation histogram for RASS–SDSS quasar matches by simulating X-ray source measurements using the corresponding RASS positional errors plus a small random component. Because the RASS positional errors are dependent on the X-ray flux, we use the positional errors from the RASS–SDSS quasar matched catalog. For each such RASS source, we generate a Rayleigh distribution with the positional error of that source as the scale parameter, \( \sigma \). The sum of the probability distribution function from each source gives our “simulated true match” curve. This PDF is the parent distribution for the true matches between RASS and SDSS quasars. Random associations between RASS and SDSS quasars have a linearly increasing source-separation histogram, as shown above. A linear combination of these two distributions (simulation PDF and random straight-line) should reproduce the observed RASS–SDSS quasar source-separation histogram.

We show the quasar source-separation histogram, simulated true match curve, and random component in Figure 5. The simulation curve, which does not include the random component, matches the actual quasar source-separation histogram very well except at the tail end. Combining the simulation and random components via a \( \chi^2 \)-minimization on the amplitude of each component yields an excellent fit. The total fit is not shown in Figure 5 because it would be completely masked by the data. This fit has a \( \chi^2 \) per degree of freedom of 0.68, indicating that the above model is a good choice. However, the distributions match only if the RASS positional errors are all reduced by 3°, implying that the quoted 6” systematic offset was overestimated.

The thin upper curve in Figure 5 gives the “true matching fraction” for RASS–SDSS quasar matches (percent, right axis). This is the number of true matches (simulation curve) divided by the total fit (simulation+random) at that radius. Note that at 30°,
about 90% of the RASS–SDSS quasars matches are legitimate. We also find that at 60″ there is ~6% total contamination to the RASS–SDSS quasar catalog. This agrees with the estimate from Anderson et al. (2007) of ~5% contamination for their sample.

3.4. Galaxies

Matching SDSS main sample galaxies to the RASS results in 3169 total matches. In contrast to quasars, the RASS–SDSS galaxy source-separation histogram rises quickly, but is then relatively flat out to 60″, as seen in Figure 6. This suggests that while some galaxies are detected as X-ray sources, a large fraction are simply random associations. We model the RASS–SDSS galaxy source-separation histogram following the procedure outlined for quasars above. In this case, the positional errors are those of the RASS–SDSS galaxy matched catalog. Figure 6 compares this model with the actual histogram. Note that the simulated true match distribution is somewhat wider than the equivalent quasar curve, as RASS sources associated with galaxies have a lower mean flux and thus have larger positional errors. The $\chi^2$ per degree of freedom of the total fit (simulated + random) is 1.18 for galaxies.

To reduce the effect of source confusion in our RASS–SDSS galaxy sample, we remove from our matched galaxy catalog RASS sources that are also positionally matched with likely X-ray emitters. Our method is similar to that employed by Agüeros et al. (2006) who removed RASS sources that overlapped with spectroscopically identified quasars, blue point sources (potential quasars), bright objects (ROSAT contaminant), and sources with a quasar-like X-ray/optical spectral slope. Our requirements are more relaxed, as our aim is not to eliminate all obvious X-ray sources, but rather to identify X-ray counterparts of galaxies. Thus, we only remove RASS sources from our matched galaxy catalog that are close to the most reliable RASS cross-matches: within 40″ of an SDSS quasar or within 30″ of an object in the “blue2” list described above. Also, if two SDSS galaxies match to one RASS source, we take only the nearest match. This reduces the sample to 1876 galaxies, with many obviously incorrect matches removed, such as the “match” shown in Figure 3. This “cleaned” catalog improves the $\chi^2$ of the simulation + random fit to 0.96 and is the catalog employed in the analysis that follows.

4. RASS DETECTIONS BY GALAXY SPECTROSCOPIC CLASS

One would expect galaxies with different optical spectroscopic classes to produce different X-ray fluxes and thus to have different matching fractions. Obrič et al. (2006) list the RASS matching fractions for SDSS galaxies showing no emission as well as AGN, star-forming, and unknown emission-line galaxies. They also plot their RASS matches on an emission-line classification diagram analogous to the left-most plot in Figure 1. However, they do not discuss random matches, nor do they remove known invalid matches (e.g. quasars). Thus, their sample includes many SDSS galaxies which are unlikely to be true matches to RASS sources. To investigate the connection between RASS detection likelihood and optical spectroscopic class, we separate the cleaned RASS–SDSS galaxy catalog into subclasses as described in section 2.1.2. For each of these subclasses, we simulate the source-separation histogram via their corresponding RASS positional errors and linear random components as before, and list the $\chi^2$ of the fits in Table 1.

Figure 7 compares the actual and simulated distributions for the different galaxy classes. The left plot shows the four different types of classified emission-line galaxies, while the right plot shows the unclassified and passive galaxies. The thin red curves show the true matching fraction at a given radius. Note the high true matching fraction for galaxies with potentially significant optical emission from a central accretion source: the Seyfert, LINER, and transition objects. Also note the relatively high true matching fraction for unclassified emission and passive galaxies. Galaxies with their optical emission dominated by star formation have a very small true matching fraction; though there are a large number of RASS–SDSS matches for H II and unclassifiable galaxies, most of those matches are purely random associations.

We list the detection fractions for the various spectral classes in Table 2, including quasars for comparison. This detection fraction is the integrated simulation curve divided by the total number of galaxies in that class. Note the relatively high detection fraction for galaxies with AGN-dominated optical emission, including the transition objects. The large X-ray detection fraction for unclassified emission sources (defined

---

**Table 1**

| CLASS     | All | Passive | Unclassifiable | Emission | Unc. Emission | H II | Transition | LINER | Seyfert |
|-----------|-----|---------|----------------|----------|---------------|-----|------------|-------|--------|
| $\chi^2$  | 1.01| 1.24    | 1.14           | 0.47     | 0.87          | 1.53| 1.47       | 0.47  | 1.00   |

**Table 2**

| CLASS       | Quasar | All | Passive | Unclassifiable | Emission | Unc. Emission | H II | Transition | LINER | Seyfert |
|-------------|--------|-----|---------|----------------|----------|---------------|-----|------------|-------|--------|
| Detection   | 8.3    | 0.12| 0.41    | 0.11           | 0.11     | 0.28          | 0.004| 0.19       | 0.41  | 0.66   |

---

Figure 6. Reconstructed ROSAT–SDSS galaxy source-separation with ~36% true matches + ~64% random matches out to 60″. The thin curve from the upper left to the lower right, cutting through the histogram, is the true matching fraction (percent, right axis). There is very good agreement between the total simulation curve and the actual distribution.
in Section 2.1.2) suggests that many of these objects harbor obscured accretion.

The number of passive, unclassifiable galaxy, and LINER matches to RASS are slightly underpredicted by the model at moderate radii (20–40′). Visual inspection of these galaxies confirms that some of them are in or near clusters, which would produce an X-ray source near to, but not coincident with, the galaxy. We do not have a cluster catalog to remove these “contaminants” but a visual tally shows that between half and two-thirds of the RASS-matched passive galaxies may be contaminated by the presence of a galaxy-cluster. However, some of these galaxies appear to be field galaxies, and thus we may be finding X-ray bright, optically normal galaxies (XBONGs; see Georgantopoulos & Georgakakis 2005). We plan to examine these objects in more detail in future work.

We list the RASS–SDSS source-separation radii at various fixed matching fractions in Table 3 for all the objects discussed in this paper. Note that at no radius do H II galaxies show even a 50% true matching fraction. The true matching fraction for star-forming galaxies is extremely low because such galaxies do not produce X-rays at a level detectable by the RASS and/or because the X-rays they produce are completely obscured by dust. Because nearly all RASS–SDSS star-forming galaxy matches are due to random associations, no claims can be made about X-ray-emitting star-forming galaxies from these data alone. The XMM-Slew survey, XMM-Newton serendipitous source catalog, and the Swift BAT catalog all observe at higher X-ray energies (less attenuated by dust), and so could help clarify the X-ray emission properties of these galaxies.

For comparison with previous studies, we give the cumulative true matching fraction at fixed radii in Table 4. These values are computed from the ratio of the integrals of the simulated and total curves in Figure 7, in contrast with the previous table, derived from the point-wise ratios. Again, note that H II galaxies have a very small cumulative true-match fraction, even at small radii. All other matched sub-samples, except for the unclassifiable galaxies, contain more than 85% true RASS matches below 20′.

Finally, we have included a supplementary online catalog, containing the sources examined in this study, with their RASS matching likelihoods. A catalog stub explaining the columns is given in Table 5. The objects in the catalog are grouped by spectroscopic class, and ordered by increasing match likelihood within each class. This catalog may be used for follow-up studies, or for comparison with other catalogs matched to SDSS.

5. FUTURE DIRECTIONS: XMM-SLEW

For comparison with the RASS, we have matched the XMM-Slew clean catalog (first release) to both SDSS galaxies and quasars. Figure 8 plots the source-separation histogram for these sources. The total number of matches is quite small.
Table 5
Catalog Stub for the Supplementary Online Catalog

| SDSS_name | SDSS_SpecObjID | RASS_name | RASS_delta | RASS_posErr | match_likelihood | galaxy_class |
|-----------|----------------|-----------|------------|-------------|-----------------|--------------|
| SDSS J134156.14+032052.7 | 149123710048010240 | 1RXS J134152.5+032028 | 59.751 | 23.0 | 0.0745 | transition |
| SDSS J122528.17+634851.7 | 169109687071932416 | 1RXSJ122525.9+634949 | 59.669 | 28.0 | 0.0749 | transition |

Notes. SDSS_SpecObjID is the unique identifier for the exact SDSS spectrum; RASS_delta is the separation between the RASS and SDSS source positions in arcseconds; RASS_posErr is the positional error of that RASS source in arcseconds; match_likelihood is the probability that a given RASS–SDSS match is a real match, from the curves in Figure 7.

We have examined the matching statistics between the ROSAT All Sky Survey and the SDSS main galaxy sample. Our technique—simulating the RASS–SDSS source-separation via the RASS positional errors plus a linear random component—can reproduce the measured source-separations for RASS–SDSS quasar matches as well as RASS–SDSS galaxies and subclassifications of galaxies. We find that the likelihood of a given cross-match match being a true match depends strongly on the optical spectral classification of a given galaxy. We find that essentially no optically classified star-forming galaxy has a true RASS counterpart, while LINERs, Seyfert 2s and transition and unclassified emission galaxies do have reliable X-ray detections. We also find a surprising number of galaxies lacking optical emission lines which appear to be detected in the RASS. A complete, low-redshift SDSS galaxy-cluster catalog could be used to clarify these XBONG candidates.

| Source Separation (arcsec) | Number |
|---------------------------|--------|
| 0                         | 5      |
| 10                        | 15     |
| 20                        | 20     |
| 30                        | 10     |
| 40                        | 5      |
| 50                        | 5      |
| 60                        | 0      |

Figure 8. XMM-slew versus SDSS cross matches for galaxies and quasars. The ~8” XMM-slew positional errors are readily visible in the quasar matches.

due to the small number of XMM-slew sources and the small overlap area between the surveys. Because of the nature of the XMM-slew survey, we cannot perform the same analysis as above; the narrow width of the slew strips is too small for a reliable random fraction to be determined, yet. From the source-separation histogram, 20” appears to be a reliable cut-off for true matches. Accepting only those matches within this radius results in 38 galaxy matches and 115 quasar matches to XMM-slew.

A coverage map for the XMM-slew data is not yet available, so it is not possible to determine the percentage of ROSAT detections that are non-detections in XMM-slew. However, among the 38 “reliable” matches are member(s) of each galaxy class described above. Most of these XMM-slew detections are in the soft band (0.2–2 keV), but there are a few galaxies with a detected hard X-ray flux. An example is shown in Figure 9: an interacting pair of galaxies optically classified as a transition and a Seyfert. The transition galaxy shows a hard X-ray flux and substantial radio point source, while the Seyfert is unidentified in hard and soft X-rays and shows a ~2σ detection in the FIRST catalog. There is no RASS source at this location. We plan to follow up on this intriguing pair to better understand their emission properties and spectral shape.

We have examined the matching statistics between the ROSAT All Sky Survey and the SDSS main galaxy sample. Our technique—simulating the RASS–SDSS source-separation via the RASS positional errors plus a linear random component—can reproduce the measured source-separations for RASS–SDSS quasar matches as well as RASS–SDSS galaxies and subclassifications of galaxies. We find that the likelihood of a given cross-match match being a true match depends strongly on the optical spectral classification of a given galaxy. We find that essentially no optically classified star-forming galaxy has a true RASS counterpart, while LINERs, Seyfert 2s and transition and unclassified emission galaxies do have reliable X-ray detections. We also find a surprising number of galaxies lacking optical emission lines which appear to be detected in the RASS. A complete, low-redshift SDSS galaxy-cluster catalog could be used to clarify these XBONG candidates.

Figure 9. Interacting galaxy pair not detected in soft X-rays. The central-source optical classifications are marked. The SDSS optically classified transition galaxy is found in the XMM-slew survey at only hard (2–12 keV) energies. The stripe on the left of the GALEX image is due to detector edge effects. The pair is UGC 08327. (A color version of this figure is available in the online journal)
Figure 10. SDSS galaxies with RASS matches within 1". Note the sharp drop at 30" within the sample of objects spectroscopically classified as galaxies compared to the main galaxy sample (all galaxies with $P_{	ext{petro}} < 17.77$). Some galaxies which are not in the main sample were targeted specifically because they were within 30" of a RASS source.

Our technique can be applied to any cross-matching between two surveys. The only requirement is that the positional errors of each measurement be known; no arbitrary fitting parameters are needed. By comparing the observed source-separation histogram with a linear combination of the probability distribution functions computed from the positional errors and a random matched catalog, a "true matching fraction" can be determined for any two matched catalogs. This is not limited to X-rays: As a test, we were also able to reproduce the source-separation histogram for a matched catalog of SDSS spectroscopic stars and GALEX UV sources. The technique works best for catalogs containing mostly point sources, as centroiding extended sources can be difficult and the centers of sources may be wavelength dependent.

We are grateful to Jenny Greene, Alina Badus, and Danny Pan for providing useful comments on earlier drafts of this paper, and to the anonymous referee for several helpful additions. We acknowledge support from NASA grant NAG5-12243 and NSF grant AST-0507647.

Funding for the Sloan Digital Sky Survey (SDSS) and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, and the Max-Planck-Institut für extraterrestrische Physik (MPE) at Garching, Germany.

APPENDIX

SDSS GALAXY SELECTION

The main galaxy sample does not contain all the galaxies with spectra: a galaxy could also have a spectrum taken if it is within 2" of a FIRST radio source or within the error-circle (10–30") of a RASS source. Luminous red galaxies are selected for follow-up spectra based on their position in the (g-r, r-i, i) color–color–magnitude cube. Spectra are also taken for a variety of serendipitous sources including low surface-brightness galaxies. These other sources are all dimmer than 17.77 in the r-band, and biased toward AGN and star-forming galaxies. The systematics of these serendipitous sources are poorly understood.

The primary method for downloading large data sets from SDSS is CasJobs. To extract the main galaxy sample from SDSS CasJobs, use the SpecObj parameter ObjType and select those objects classified as "galaxy". This includes all objects that were targeted for spectroscopy because they met the main galaxy sample criterion. This classification is before the spectra were taken, and is thus a uniform sample. A more naive selection might be to take all objects spectroscopically classified as galaxies: those with SpecObj parameter SpecClass listed as "galaxy". However, this sample includes all objects with a galaxy-like spectrum, which includes objects targeted for the above reasons in addition to the main galaxy sample.

In Figure 10 we show the source-separation histogram for these two different samples. The "photometric" sample is the main galaxy sample used in this study. The spectroscopic sample, with a peak at 30", includes objects specifically targeted because they were near a RASS source. The fiber-selection process allocates spare spectroscopic fibers to sources within 30" of a RASS source. These objects, having SpecObj parameter ObjType classifications "ROSAT_A", "ROSAT_B", "ROSAT_C", or "ROSAT_D", account for roughly 2% of all objects with spectra in SDSS. Stoughton et al. (2002) claim over half of these ROSAT-based targets turn out to be quasars or AGN. This results in a factor of two increase in potential matches at matching radii below 30". This is why a statistical analysis of RASS matches to SDSS must stick with the main galaxy sample; the other sources were selected non-uniformly, and though they may result in odd and interesting spectra, they produce a strong bias in X-ray matching properties.

REFERENCES

Abazajian, K., et al. 2003, AJ, 126, 2081
Abazajian, K., et al. 2004, AJ, 128, 502
Adelman-McCarthy, J. K., et al. 2006, ApJS, 162, 38
Ag{"{u}}eros, M., et al. 2006, AJ, 131, 1740
Anderson, S. F., et al. 2003, AJ, 126, 2209
Anderson, S. F., et al. 2007, AJ, 133, 313
Baldwin, J., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559
Blanton, M. R., Lin, H., Lupton, R. H., Maley, F. M., Young, N., Zehavi, I., & Loveday, J. 2003, AJ, 125, 2276

5 SDSS CasJobs is available from http://casjobs.sdss.org/CasJobs/.
