We thank the referee for the helpful comments. In the revised draft, we have removed Sec. 4, and moved the simulation discussion to Sec. 3.3. The discussion of LogN-LogS is now moved to Sec. 3.6. The discussion of NWAY in Appendix B is moved to Sec. 5.3 as per referee's suggestion. We have also included an additional appendix C with a table of all photometric redshifts we calculated for the 4.5 deg^2 field with forced-photometry catalog. Throughout the revised draft, we have highlighted the changes in boldface, except for typo corrections and several updates on various numbers due to the small modifications made to the X-ray source catalog.

Reviewer's Comments:

This paper presents the X-ray point sources catalog of the XMM-SERVS survey on the XMM-LSS field. This is an important work and the catalog comprises of a large sample of AGN. It is a nice work and it needs to be published but there are several points that need to be addressed before the publication.
The two main points are: establish a likelihood threshold for the X-ray source reliability, perform a classification of the X-ray sources using the spectroscopic and photometric information.

Section 3 and 4:

I overall do not agree with the sequence of the sections and the work flow. Choosing a detml value just because it was used in the XMM-COSMOS survey is not the right approach as each survey has its own properties. As it is now, it looks like the authors chose 2 values of detml and then did all the work. Instead, the simulations should go first, then the desired completeness is chosen (99%, 99.5% or whatever) and as a consequence the detml values corresponding to the chosen completeness are found. Moreover, it could be possible that different detml values are found for different bands.
I think the content of section 4 on the simulations should go much earlier and should be the driver of the detml threshold choice. If this is what the authors have done already, then it is just matter to restructure the paper, otherwise, it might be needed to find the new detml and change the results on number of sources and so on.

We thank the referee for suggesting this approach, as the previous spurious fractions quoted for the detml=6 catalog did leave the impression that the reliability of the catalog is limited (e.g., spurious fraction at the hard band is ~2%). In the revised draft we are now using a single set of detml values that were determined to have a constant spurious fraction, 1%. We have also moved part of the simulation section to before where the main X-ray catalog was described. Note that the new catalog has 5242 sources with only 10 sources less than the previous detml > 6 catalog. This is because only a very small number of sources had detml values smaller than the new thresholds in all three bands.

Originally, the additional detml=10.8 threshold was chosen in addition to detml=6 because the spurious fraction at detml=6 is high in the hard band (~2%). With the new set of detml thresholds inspired by the referee, the spurious fraction for this catalog is well-defined and is comparable, if not better, than previous XMM surveys. Therefore, we consider that it is no longer necessary to include an additional catalog with an arbitrarily higher detml. We believe using a single set of detml values would be more straightforward for the readers and the community to use the catalog without confusion. In addition, we have also calculated the reliability for sources at each band according to their detml values for the users to decided whether they want to cut the catalog at a reliability threshold of their choosing.

Section 5:

To move a step forward of just presenting an X-ray catalog, the authors performed multiwavelength classification of the sources and provide also spectroscopic and photometric redshifts. I understand that it is not easy to classify the spectra as these come from different surveys and instruments, but at least a classification dividing the sources in broad line and narrow line + absorption line is needed. At the fluxes covered by this survey, the AGN population is still dominated by broad line AGN, therefore it is important to have such a classification also for the photometric redshifts. Moreover, I am suggesting to use the method of Salvato et al. 2009 used for the XMM-COSMOS survey to compute the photometric redshift. I am suggesting this method as it properly address the issue of photometric redshift for point and extended sources (in optical not X-ray band) and how to treat the sources when having broad line and non broad line AGN.
In the end, it would be nice to have a general classification of the sources, using both the results from the spectroscopic and photometric classification.

First, in the previous draft, we have included the basic spectroscopic classifications in the main X-ray source table, where we obtained the broad-line flags from publicly available catalogs. We have updated the text in Sec. 4.6 (Sec. 5.3 of the original draft) to point the readers to the specific column where the broad-line/non-broad-line classification is provided for sources with spectroscopic coverage. Additional text has also been added to Sec. 4.6 to describe the details of the spectroscopic classification. Notably XMM-LSS will be covered by both VLT MOONS and SUBARU-PFS surveys in the near future, which will provide additional spectroscopic information of the X-ray sources presented in our paper.

As for the photometric classification and methods for computing photo-z and identifying broad-line sources using optical morphologies, we were unable to use the Salvato et al. 2009 method due to several reasons. The Salvato et al. 2009 method requires high-quality HST morphology information which is not available in the XMM-LSS field. We have attempted using the morphological flags (classification_extendedness) from the HSC survey. For the spectroscopically confirmed type 1 AGNs, we find that none of them are classified as a point source in all 5 HSC bands. For the sources that were not targeted in previous spectroscopic campaigns, only ~ 10—15% of them are classified as ``extended” in at least one of the HSC bands. This is expected as the sources without specz are usually fainter, higher redshift sources, and with the limited spatial resolution of ground-based telescopes, the morphological information from the HSC data reduction pipeline is unreliable. Additionally, without the morphological information, the empirical X-ray flux cut that Salvato et al. (2009, 2011) used for identifying broad-line AGN candidates is not suitable for our sample. We have also estimated the expected fraction of broad-line AGNs in the previous draft (3rd paragraph of Sec. 5.4 of the original draft), and we expect that only a small fraction of sources without spectroscopic coverage would be broad-line AGNs.

We also ran our photoz codes with an additional broad-line quasar template on the 352 point sources not targeted by spectroscopic observations. These sources are identified as a point source in all five HSC bands. We find that the photometric redshift uncertainties for ~80% of these sources are higher than the results obtained without the broad-line quasar template, and only 23% of these sources would have high-quality photometric redshifts (Qz < 1.0, e.g., Sec. 5.5 of the draft). Due to these reasons, we do not consider the Salvato et al. (2009, 2011) methods to be suitable for our catalog. Without the crucial optical variability and high-resolution HST morphological information, we cannot provide reliable photometric classification for our sources. We do agree with the referee that it is important to investigate the photometric and spectroscopic properties of our sources, and we will present detailed analyses on the SED and X-ray spectral properties in future works.

Minor comments, some of them are also related to the detml choice that should be addressed.
—Introduction, page 2 line 30: specify XMM COSMOS when talking about COSMOS

Done

— End of section 2.1: Need more information about the use of the Chandra data from the Chandra Source Catalog. I know this is not subject of this paper but as this data are used a lot, some extra information is really necessary.

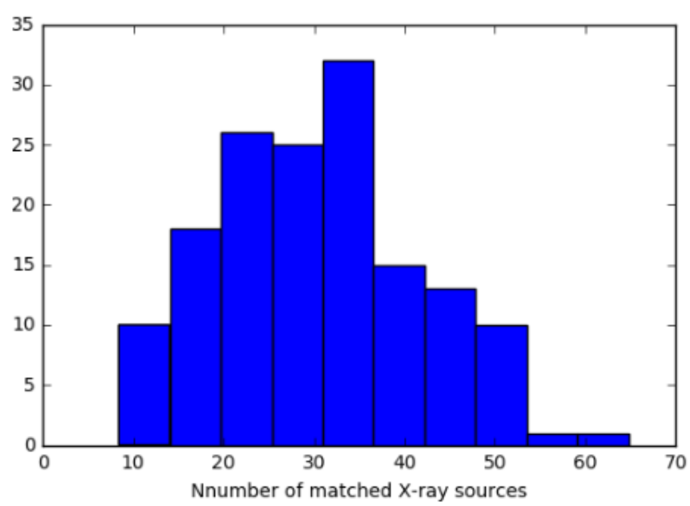
Additional information is now provided.

— Figure 3 left: very very hard to see. I would suggest to invert the color map and maybe use a log scale. End of section 3.1

Color map inverted.

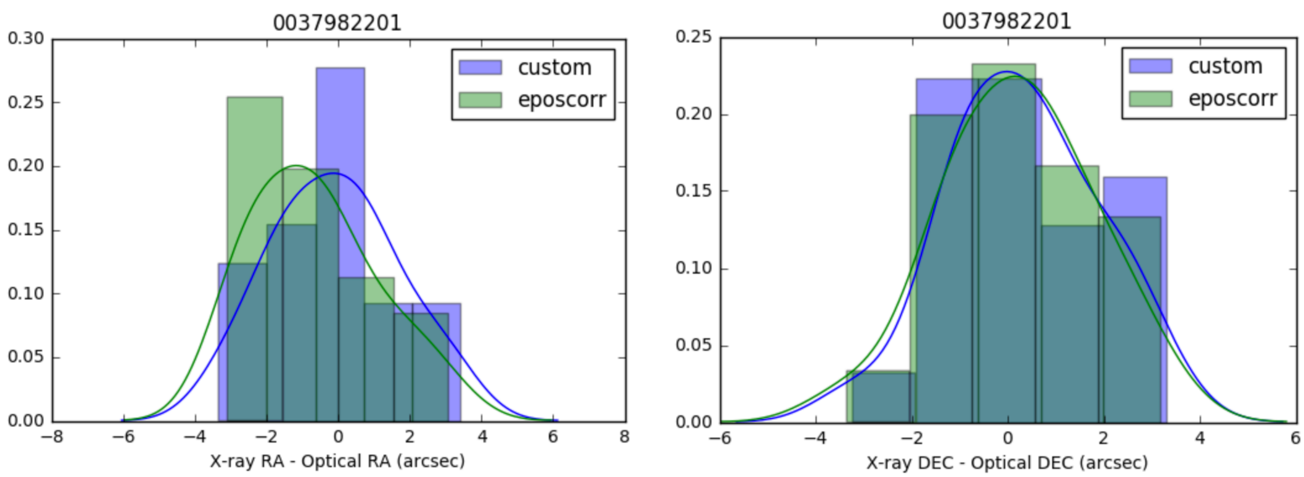
— "The median number of X-ray sources in an ObsID with only one HSC counterpart within 3" is 32." How does the distribution of this look like? I would like to see it.

The distribution is shown below



— "We visually inspect these ObsIDs and conclude that our approach does improve the alignments between the optical and corrected X-ray images." How does the distribution of the offsets look like? I would like to see it. The visual inspection is not enough to see the improvement.

The ``visual inspection'' described in the previous draft was not clear, we thank the referee for pointing this out. By ``visual inspection'', we meant visually inspecting the RA/DEC offset distributions, not visually inspecting the real images. For instance, the obsid with the most significant difference between the results from XMMSAS eposcorr task and our own method is 0037980301. We show the distributions of the X-ray and optical offsets for this obsID below. The text has been updated in the revised draft.



Left: X-ray to optical offsets in RA. Right: X-ray to optical offsets in DEC. The offsets computed using eposcorr are shown as the green histograms, while the offsets computed using our custom method were shown as the blue histograms. The associated kernel density estimations are also shown as the green and blue curves.

Section 3.3:
— "For the 5228 X-ray sources detected in the full-band during the second-pass source-searching process, a total of 2085
X-ray sources are …" Please provide the numbers for the sample at detml>10.8 which is more reliable. How does this compare with XMM-COSMOS?

— "Using the angular separations between the 2085 X-ray sources and their unique optical counterparts, we derive an empirical relation…" How does this change using only the sources with detml>10.8?

We acknowledge that the number 2085 quoted in the previous draft was incorrect, as it was the number for a previous internal version of X-ray catalog from a 4.5 deg^2 subregion. The number is now updated to 2434. The best-fitting parameters are slightly different.

Although we are no longer using a different detml > 10.8 threshold in the revised draft, we summarize the information for the detml > 10.8 catalog here. For the full-band catalog, there are 2366 sources with a 18 < *i* < 23 HSC counterpart within 3”, and the best-fit slope and intercept for Eq. 1 in the draft are almost exactly the same (~ 0.001 and 0.00001 difference in slope and intercept, respectively).

For XMM-COSMOS, we use the matching results from Brusa et al. (2010), and found that the number of X-ray sources with a 18 < *i* < 23 optical counterpart within 3” is 991.

Section 3.4
— About the energy conversion factors: How are you accounting for the differences due to fields observed in different epochs? I assume you need to find a value weighted on exposure time to account for differences in the effective area and so on.

We have re-calculated ECFs for different epochs, and we have switched to using the exposure-time-weighed ECFs for each source in the revised draft.

— What is the completeness and reliability of the sample at detml=6? how many spurious sources are you expecting at this level?

Since we have switched to using the spurious fraction based detml values, the expected spurious source count is 1%, i.e., ~52 spurious sources. As for completeness, our simulation suggests that < 0.45% of the sources are not detected due to source blending. This was described in the last paragraph of Sec. 4.1 in the previous draft (which is now Sec. 3.3 in the revised draft).

— Please recompute the HR using BEHR.

For XMM-Newton, the X-ray counts are determined based on combining those from all three EPIC cameras. In this case, BEHR cannot properly account for the different background levels across the different instruments, therefore we cannot use the BEHR method commonly adopted for other X-ray instruments such as Chandra. The HR values provided in this work are for the purpose of first-order estimates on the spectral shapes. Detailed X-ray spectral analyses for our sources will be presented in a separate article.

Figure 10 right and figure 11: please plot the curves at both detml>6 and >10.8

See response to major comments. We are no longer using two different detml values, therefore we do not plot two sets of curves. The figures are updated with the new detml values.

Section 4
- See general comment above.
- Again, here, the lgN-lgS should be recomputed with both detml>6 and >10.8

logN-logS is not sensitive to the choice of detml other than how far the measurements could go down to the faint-end of flux distributions, and we are no longer using two different detml thresholds.

Section 5
Figure 15: the cutouts need to be labeled.

Done.

Section 5.1

- The results for the likelihood ratio need to be presented for both samples detml>6 and >10.

We are no longer using two detml values in the revised draft. As we stated earlier in the response to the main comments, including an additional, arbitrary detml value could cause unnecessary confusion, and there is no qualitative difference between the matching results of the detml > 6 catalog and the detml > 10.8catalog.

Section 5.3 and 5.4
The paper is completely missing a discussion on the spectroscopic classification which needs to be discussed. I understand the work was done by different surveys and different instruments but redshifts and classification are such an important piece of an X-ray sample. Moreover, the same redshifts/spectra are used for training the photo-z.

See responses to the major comments. Also, our photometric redshifts were computed with EASY, which is a template-fitting code that does not require a training set.

The photometric redshift are computed using EAZY, but I would suggest to use the method of Salvato et al. (2009) which was applied to the XMM-COSMOS survey, which flux limit is consistent with this survey. Point sources and extended sources were there treated differently by Salvato et al., therefore I suggest to use their method. Moreover, given the good multiwavelength data available for the entire survey, I would suggest to compute photometric redshifts for all the sources.

In the revised draft, we have extended our photoz coverage to the 4.5 deg^2 region covered by both VIDEO and SERVS. Also see the response to the major comments on the choice of photoz estimation method. Figure 18 is also updated in the revised draft as now we have combined the cumulative distributions of the spec-z-only and the spec-z+photo-z samples in the the same panel.

Appendix
I appreciate the comparison published between the NWAY and MLE results. I think this part should be included in the paper as this is mainly a catalog paper, therefore it is not out of context to have this comparison. Is there a difference in X-ray properties for those sources where NWAY and MLE point to a different counterpart? It would be interesting to add this.

Appendix B is now moved to Sec. 5. Generally speaking, X-ray sources with different NWAY and MLE counterparts are fainter, which is expected as their larger positional uncertainties lead to higher numbers of counterpart candidates. Therefore, they are more likely to have different NWAY and MLE results. Additional text is included in the revised draft.