

DAXA: Traversing the X-ray desert by Democratising Archival X-ray Astronomy

David J. Turner^{1,2}, Jessica E. Pilling², Megan Donahue¹, Paul A. Giles², Kathy Romer², Agrim Gupta¹, Toby Wallage², and Ray Wang¹

¹ Department of Physics and Astronomy, Michigan State University, Lansing, Michigan, USA

² Department of Physics and Astronomy, University of Sussex, Brighton, UK

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#)
- [Repository](#)
- [Archive](#)

Editor: [Open Journals](#)

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

We introduce a new, open-source, Python module for the acquisition and processing of archival data from many X-ray telescopes, Democratising Archival X-ray Astronomy (hereafter referred to as DAXA). The aim of DAXA is to provide a unified, easy-to-use, Python interface to the disparate X-ray telescope data archives and their processing tools. We provide this interface for the majority of X-ray telescopes launched within the last 30 years. This module enables much greater access to X-ray data for non-specialists, while preserving low-level control of processing for X-ray experts. It is useful for identifying relevant observations of a single object of interest, but it excels at creating and managing multi-mission datasets for serendipitous or targeted studies of large samples of X-ray emitting objects. Once relevant observations are identified, the raw data can be downloaded (and optionally processed) through DAXA, or pre-processed event lists, images, and exposure maps can be downloaded if they are available. With a decade-long ‘X-ray desert’ potentially on the horizon, archival data will take on even greater importance, and enhanced access to those archives will be vital to the continuation of X-ray astronomy.

Statement of need

X-ray observations provide a powerful view of some of the most extreme processes in the Universe, and have had a profound impact on our understanding of many types of astrophysical objects. Every sub-field of astronomy, astrophysics, and cosmology has benefited significantly from X-ray coverage over the last three decades; the observation of X-ray cavities in galaxy clusters caused by central AGN helped to shed light on the cooling-flow problem (McNamara et al., 2001); further X-ray observations allowed for the measurement of spatially-resolved entropy in hundreds of clusters, dramatically increasing understanding of cooling and heating processes in their cores; quasi-periodic eruptions (QPE) from active galactic nuclei (AGN) were discovered (Miniutti et al., 2019); the high-energy view of young stars gave insights into their magnetic fields and stellar winds (Getman et al., 2005; Güdel et al., 2007); calibrating mis-centering for galaxy cluster weak-lensing studies helped constrain cosmological parameters (Zhang et al., 2019); and X-rays even helped probe the irradiation of exoplanets (Poppenhaeger et al., 2021). Indeed, X-ray telescopes have created many entirely new fields of study; they provided the first evidence of X-ray sources outside the solar system (Giacconi et al., 1962); discovered the first widely accepted black hole, and launched the study of supernova remnants (Bowyer et al., 1965); and found ionized, volume-filling, gas within the Coma galaxy cluster (the intra-cluster medium) (Kellogg et al., 1971), with the implication that clusters were more than collections of galaxies. These non-exhaustive lists make evident the importance of X-ray observations to the astronomy, astrophysics, and cosmology communities.

The current workhorse X-ray observatories, *XMM* (Jansen et al., 2001) and *Chandra* are ageing, with *Chandra* in particular experiencing a decline in low-energy sensitivity that might limit science cases (other telescopes are online but are more specialised); these missions cannot last forever. If we are to enter an X-ray desert, where the astrophysics community has only limited access to new X-ray observations from specialised missions like *Swift* (Gehrels et al., 2004), *NuSTAR* (Harrison et al., 2013), and *XRISM* (XRISM Science Team, 2020), then archival data (and serendipitous studies) take on an even greater value than they already hold. DAXA is part of an ecosystem of open-source software designed around the concept of enabling serendipitous studies of X-ray emitting objects, and can download and prepare X-ray observations for use with tools like 'X-ray: Generate and Analyse' (XGA; D. J. Turner et al. (2022)). X-ray observations are uniquely well suited for the kind of archival study facilitated by DAXA and XGA, as they generally record the time, position, and energy of each individual photon impacting the detector (true for all missions currently implemented in DAXA); this means that we can create images, lightcurves, and spectra for any object detected within the field-of-view, even if it was not the target. With this software, we enable the maximum use of existing X-ray archives, to traverse the X-ray desert and ensure that we are fully prepared for future X-ray telescopes such as *Athena* (Nandra et al., 2013) and *Lynx* (Gaskin et al., 2019). Having easy access to the whole X-ray observation history of an object can provide valuable astrophysical context at little extra cost.

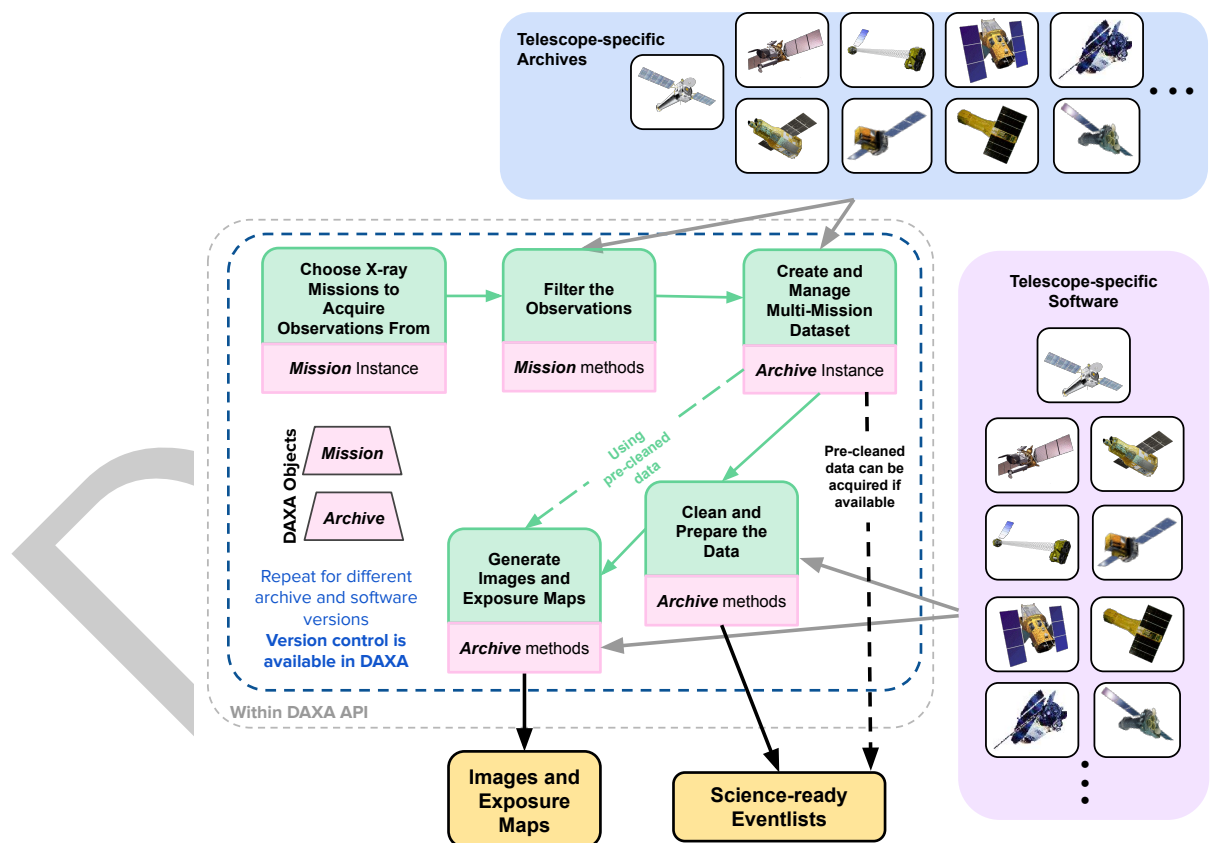


Figure 1: A flowchart showing a brief overview of the DAXA workflow. We indicate the different ways that DAXA can be used to access, process, and use archival X-ray data.

As such, X-ray data should be made as accessible as possible, both for X-ray experts and non-specialists who may face barriers to entry; X-ray data can be particularly intimidating to those astronomers who have not used it before, though their research may benefit from a high-energy view. Difficulty of use undermines the open-source nature of X-ray astronomy

65 data, which organisations such as the European Space Agency (ESA) and the High Energy
66 Astrophysics Science Archive Research Center (HEASARC) have gone to great lengths to build.
67 This may limit the reach and scientific impact of X-ray telescopes; we should seek to maximise
68 the user of X-ray data, both to support X-ray astronomy through the 'X-ray desert', and to
69 persuade funding bodies of the great need for further X-ray telescopes.

70 We build on ESA and HEASARC's success and make the data more accessible by providing
71 a normalised interface to different backend software packages and datasets, allowing for the
72 easy processing of X-ray data to a scientifically useful state; this is in addition to the ability to
73 download pre-processed data from many of the data archives. Through DAXA, most X-ray
74 observatory archives are accessible through a single unified interface available in a programming
75 language that is ubiquitous in astronomy (Python); locally searching for data relevant to
76 a particular sample gives us the opportunity to better record and share the exact search
77 parameters, through a Jupyter notebook for instance.

78 Features

79 DAXA contains two types of Python class: mission classes and the archive class (see [Figure 1](#)
80 for a schematic of the structure of the module). Mission classes directly represent a telescope
81 or survey (for instance there are separate classes for pointed and survey observations taken by
82 *ROSAT* ([Truemper, 1982](#)), as the characteristics of the data are quite different), and exist
83 to provide a Python interface with the current telescope observation database. Such mission
84 classes allow the user to easily identify data relevant to their objects of interest with various
85 filtering methods (it is also possible to download the entire archive of a telescope); these
86 include filtering on spatial position (determining whether a coordinate of interest is within
87 the field-of-view), filtering on the time of the observation (also filtering on whether a specific
88 coordinate was observed at a specific time, for whole samples with different coordinates and
89 times of interest), and filtering on specific observation identifiers (ObsIDs) if they are already
90 known. Each mission class has some knowledge of the characteristics of the telescope it
91 represents (such as the field-of-view) to make observation filtering easier. The user can also
92 select a subset of instruments, if the telescope has more than one, to exclude any that may
93 not contribute to their analysis.

94 Once a set of relevant observations have been identified, for either a single mission or a set of
95 missions, a DAXA data archive can be declared. When a user declares a DAXA archive, the
96 selected data are automatically downloaded from the various telescope datasets, and then
97 ingested and organised so that they can be managed through the DAXA interface. We have
98 also implemented user-friendly, multi-threaded, data preparation and cleaning routines for
99 some telescopes (*XMM* and *eROSITA* in particular, though more will be added); fine control
100 of the parameters that configure these processes is retained, but default behaviours can be
101 used if the user is unfamiliar with the minutiae of X-ray data preparation. Another key benefit
102 of reducing data with DAXA is the easy access to data logs through its interface, in case
103 of suspected problems during the reduction processes. The module is also capable of safely
104 handling processing failures, recording at which processing step the failure occurred for a
105 particular ObsID.

106 All of this information is retained permanently, not just while the initial DAXA processes are
107 running. Any DAXA archive can be loaded back in after the initial processing, once again
108 providing access to the stored logs, and processing information. At this point the archives can
109 also be updated, either by searching for new data from the existing missions, adding data from
110 a different mission, or re-processing specific observations to achieve more scientifically useful
111 data. Any such change will be recorded in the archive history, so that the data archive can have
112 a specific version that refers to its exact state at any given time; this version can be referred
113 to in published work using the data archive. Each data archive is also capable of creating a file
114 that other DAXA users can import, and which will recreate the data archive by downloading
115 the same data, and processing it in the same way; this renders making fully processed, and

large, X-ray data files available with a piece of research unnecessary. This feature in particular can be used to further one of the tenets of open-source science - reproducibility.

Existing software packages

There are no direct analogues to our module, though we must acknowledge the many pieces of software (and data archives), that greatly facilitated the development of DAXA. Data access is made possible primarily by the HEASARC data archive, though the Astroquery (Ginsburg et al., 2019) module is also used. HEASARC provides an online interface to query their data archive, which has similar functionality to some of the filtering methods of mission classes in DAXA (though we provide slightly more functionality in that regard), and they provide Python SQL examples to access the data, but none of the data management and cleaning functionality that we include.

DAXA also builds on the various telescope-specific software packages to perform data preparation and cleaning. Particularly important are the *XMM* Science Analysis System (SAS; Gabriel et al. (2004)) and the complementary extended SAS (eSAS; Snowden & Kuntz (2011)) packages, which allow us to provide simple Python interfaces to the complex, multi-step, processes that are required to prepare raw *XMM* data for scientific use. The analogous *eROSITA* Science Analysis Software System (eSASS; Predehl et al. (2021)) must also be mentioned, as it provides the tools needed to reduce and prepare *eROSITA* data. In this vein we must also acknowledge the HEASoft package, which is almost ubiquitous in X-ray data analyses, and is used by both SAS and eSASS.

Another related software package is the other module in our open-source X-ray astronomy ecosystem, X-ray: Generate and Analyse (XGA; D. J. Turner et al. (2022)) - it exists to analyse large samples of sources using large sets of X-ray data. DAXA is designed to go hand-in-hand with XGA, as it will build and manage the kind of dataset required for XGA to attain maximum usefulness. We emphasise that such datasets do not *have* to be analysed with XGA however.

We have created a one-stop-shop for downloading and processing archival X-ray data, making it more accessible and user-friendly, particularly for non-specialists. DAXA is greater than the sum of its parts, but is only possible because of the existing software packages it builds upon; we hope that it only enhances the value that astrophysicists derive from the other software we have mentioned.

Research projects using DAXA

DAXA has been useful for several research projects by different collaborations, and we anticipate that the number of research projects benefiting from its features will increase significantly. The primary use to which DAXA has been put is to assemble the multi-mission X-ray dataset (*XMM*, *Chandra*, *eROSITA*, *Swift*, and *ROSAT*) for the X-ray follow-up component of the Local Volume Complete Cluster Survey (LoVoCCS; Fu et al. (2022), David J. Turner, Donahue, Evrard, et al. (2024, in prep.)). It was used to identify the relevant observations, download, and process them, as well as to organise the significant number of files and make it easier for the dataset to be served to the X-ray community. Construction and administration of such large, complicated, multi-mission datasets is rendered quick and easy.

The X-ray Cluster Survey (XCS; formerly known as the *XMM* Cluster Survey) collaboration now uses DAXA to create and manage their processed X-ray archive; particularly useful is DAXA's support for telescopes other than *XMM*, which has allowed the serendipitous science undertaken by XCS to expand to the use of different telescopes. These telescopes are complementary to *XMM*, and also increase the sky coverage, which in turn increases the likelihood that an object of interest has an accessible X-ray observation.

As DAXA now supports XCS, it has contributed to a research project that has measured X-ray

properties (spectral, time-series, and photometric) for every LOFAR source that falls on an *XMM* observation (a DAXA generated dataset was used for an XGA analysis). This kind of bulk analysis is trivial when our software packages are utilised, and will result in a comprehensive catalogue that is invaluable to the radio astronomy community.

Finally, DAXA has been used to identify *XMM* and *Chandra* (alongside other telescopes, though they play only a supporting role) observations of a series of galaxy groups that appear in the foreground of UV bright quasars (David J. Turner, Donahue, & Stocke, 2024, in prep.). Absorption features that indicate the presence of Oxygen VI were identified in the spectra of several of the quasars, and the data that DAXA identified and retrieved allowed for an exploration of the hot-gas properties of these groups.

Future Work

The most significant new features implemented in DAXA will be new mission classes added when new X-ray telescope archives become available, or one of the existing missions that we have not yet implemented is added (for instance *XMM* observations taken whilst slewing). We will also seek to include support for more telescope-specific cleaning methods taken from their backend software; additionally we wish to implement our own generic processing and cleaning techniques where possible, applicable to multiple missions. We also aim to include source detection capabilities; specifically techniques that are generally applicable to multiple missions whilst taking into account instrument-specific effects.

Acknowledgements

DT and MD are grateful for support from the National Aeronautic and Space Administration Astrophysics Data Analysis Program (NASA80NSSC22K0476).

KR and PG acknowledge support from the UK Science and Technology Facilities Council via grants ST/T000473/1 and ST/X001040/1. JP acknowledges support from the UK Science and Technology Facilities Council via grants ST/X508822/1.

DT is grateful to Amanda Witwer for comments on the structure and content of this paper.

References

- Bowyer, S., Byram, E. T., Chubb, T. A., & Friedman, H. (1965). Cosmic X-ray Sources. *Science*, 147(3656), 394–398. <https://doi.org/10.1126/science.147.3656.394>
- Fu, S., Dell'Antonio, I., Chary, R.-R., Clowe, D., Cooper, M. C., Donahue, M., Evrard, A., Lacy, M., Lauer, T., Liu, B., McCleary, J., Meneghetti, M., Miyatake, H., Montes, M., Natarajan, P., Ntampaka, M., Pierpaoli, E., Postman, M., Sohn, J., ... Wilson, G. (2022). LoVoCCS. I. Survey Introduction, Data Processing Pipeline, and Early Science Results. 933(1), 84. <https://doi.org/10.3847/1538-4357/ac68e8>
- Gabriel, C., Denby, M., Fyfe, D. J., Hoar, J., Ibarra, A., Ojero, E., Osborne, J., Saxton, R. D., Lammers, U., & Vacanti, G. (2004). The XMM-Newton SAS - Distributed Development and Maintenance of a Large Science Analysis System: A Critical Analysis. In F. Ochsenbein, M. G. Allen, & D. Egret (Eds.), *Astronomical data analysis software and systems (ADASS) XIII* (Vol. 314, p. 759).
- Gaskin, J. A., Swartz, D. A., Vikhlinin, A., Özel, F., Gelmis, K. E., Arenberg, J. W., Bandler, S. R., Bautz, M. W., Civitani, M. M., Dominguez, A., Eckart, M. E., Falcone, A. D., Figueroa-Feliciano, E., Freeman, M. D., Günther, H. M., Havey, K. A., Heilmann, R. K., Kilaru, K., Kraft, R. P., ... Zuhone, J. A. (2019). Lynx X-Ray Observatory: an

- overview. *Journal of Astronomical Telescopes, Instruments, and Systems*, 5, 021001.
<https://doi.org/10.1117/1.JATIS.5.2.021001>
- Gehrels, N., Chincarini, G., Giommi, P., Mason, K. O., Nousek, J. A., Wells, A. A., White, N. E., Barthelmy, S. D., Burrows, D. N., Cominsky, L. R., Hurley, K. C., Marshall, F. E., Mészáros, P., Roming, P. W. A., Angelini, L., Barbier, L. M., Belloni, T., Campana, S., Caraveo, P. A., ... Zhang, W. W. (2004). The Swift Gamma-Ray Burst Mission. *611*(2), 1005–1020. <https://doi.org/10.1086/422091>
- Getman, K. V., Flaccomio, E., Broos, P. S., Grosso, N., Tsujimoto, M., Townsley, L., Garmire, G. P., Kastner, J., Li, J., Harnden, Jr., F. R., Wolk, S., Murray, S. S., Lada, C. J., Muench, A. A., McCaughrean, M. J., Meeus, G., Damiani, F., Micela, G., Sciortino, S., ... Feigelson, E. D. (2005). Chandra Orion Ultradeep Project: Observations and Source Lists. *160*(2), 319–352. <https://doi.org/10.1086/432092>
- Giacconi, R., Gursky, H., Paolini, F. R., & Rossi, B. B. (1962). Evidence for x Rays From Sources Outside the Solar System. *9*(11), 439–443. <https://doi.org/10.1103/PhysRevLett.9.439>
- Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., Cowperthwaite, P. S., Craig, M. W., Deil, C., Guillochon, J., Guzman, G., Liedtke, S., Lian Lim, P., Lockhart, K. E., Mommert, M., Morris, B. M., Norman, H., Parikh, M., Persson, M. V., Robitaille, T. P., Segovia, J.-C., Singer, L. P., ... a subset of the astropy collaboration. (2019). astroquery: An Astronomical Web-querying Package in Python. *157*, 98. <https://doi.org/10.3847/1538-3881/aafc33>
- Güdel, M., Briggs, K. R., Arzner, K., Audard, M., Bouvier, J., Feigelson, E. D., Franciosini, E., Glauser, A., Grosso, N., Micela, G., Monin, J.-L., Montmerle, T., Padgett, D. L., Palla, F., Pillitteri, I., Rebull, L., Scelsi, L., Silva, B., Skinner, S. L., ... Telleschi, A. (2007). The XMM-Newton extended survey of the Taurus molecular cloud (XEST). *468*(2), 353–377. <https://doi.org/10.1051/0004-6361:20065724>
- Harrison, F. A., Craig, W. W., Christensen, F. E., Hailey, C. J., Zhang, W. W., Boggs, S. E., Stern, D., Cook, W. R., Forster, K., Giommi, P., Grefenstette, B. W., Kim, Y., Kitaguchi, T., Koglin, J. E., Madsen, K. K., Mao, P. H., Miyasaka, H., Mori, K., Perri, M., ... Urry, C. M. (2013). The Nuclear Spectroscopic Telescope Array (NuSTAR) High-energy X-Ray Mission. *770*(2), 103. <https://doi.org/10.1088/0004-637X/770/2/103>
- Jansen, F., Lumb, D., Altieri, B., Clavel, J., Ehle, M., Erd, C., Gabriel, C., Guainazzi, M., Gondoin, P., Much, R., Munoz, R., Santos, M., Schartel, N., Texier, D., & Vacanti, G. (2001). XMM-Newton observatory. I. The spacecraft and operations. *365*, L1–L6. <https://doi.org/10.1051/0004-6361:20000036>
- Kellogg, E., Gursky, H., Leong, C., Schreier, E., Tananbaum, H., & Giacconi, R. (1971). X-Ray Observations of the Virgo Cluster, NGC 5128, and 3c 273 from the UHURU Satellite. *165*, L49. <https://doi.org/10.1086/180714>
- McNamara, B. R., Wise, M. W., Nulsen, P. E. J., David, L. P., Carilli, C. L., Sarazin, C. L., O'Dea, C. P., Houck, J., Donahue, M., Baum, S., Voit, M., O'Connell, R. W., & Koekemoer, A. (2001). Discovery of Ghost Cavities in the X-Ray Atmosphere of Abell 2597. *562*(2), L149–L152. <https://doi.org/10.1086/338326>
- Miniutti, G., Saxton, R. D., Giustini, M., Alexander, K. D., Fender, R. P., Heywood, I., Monageng, I., Coriat, M., Tzioumis, A. K., Read, A. M., Knigge, C., Gandhi, P., Pretorius, M. L., & Agís-González, B. (2019). Nine-hour X-ray quasi-periodic eruptions from a low-mass black hole galactic nucleus. *573*(7774), 381–384. <https://doi.org/10.1038/s41586-019-1556-x>
- Nandra, K., Barret, D., Barcons, X., Fabian, A., den Herder, J.-W., Piro, L., Watson, M., Adami, C., Aird, J., Afonso, J. M., Alexander, D., Argiroffi, C., Amati, L., Arnaud, M., Atteia, J.-L., Audard, M., Badenes, C., Ballet, J., Ballo, L., ... Zhuravleva, I. (2013). The Hot and Energetic Universe: A White Paper presenting the science theme motivating the

- 255 Athena+ mission. *arXiv e-Prints*, arXiv:1306.2307. <https://arxiv.org/abs/1306.2307>
- 256 Poppenhaeger, K., Ketzer, L., & Mallonn, M. (2021). X-ray irradiation and evaporation of
257 the four young planets around V1298 Tau. *500*(4), 4560–4572. [https://doi.org/10.1093/](https://doi.org/10.1093/mnras/staa1462)
258 [mnras/staa1462](https://doi.org/10.1093/mnras/staa1462)
- 259 Predehl, P., Andritschke, R., Arefiev, V., Babyshkin, V., Batanov, O., Becker, W., Böhringer,
260 H., Bogomolov, A., Boller, T., Borm, K., Bornemann, W., Bräuninger, H., Brüggner, M.,
261 Brunner, H., Brusa, M., Bulbul, E., Buntov, M., Burwitz, V., Burkert, W., ... Yaroshenko,
262 V. (2021). The eROSITA X-ray telescope on SRG. *647*, A1. [https://doi.org/10.1051/](https://doi.org/10.1051/0004-6361/202039313)
263 [0004-6361/202039313](https://doi.org/10.1051/0004-6361/202039313)
- 264 Snowden, S. L., & Kuntz, K. D. (2011). *Cookbook for analysis procedures for XMM-newton*
265 *EPIC MOS observations of extended objects and the diffuse background*.
- 266 Truemper, J. (1982). The ROSAT mission. *Advances in Space Research*, 2(4), 241–249.
267 [https://doi.org/10.1016/0273-1177\(82\)90070-9](https://doi.org/10.1016/0273-1177(82)90070-9)
- 268 Turner, David J., Donahue, M., Evrard, A., Fu, S., Dell'Antonio, I., Chary, R.-R., Clowe, D.,
269 Cooper, M. C., Lacy, M., Lauer, T., Liu, B., McCleary, J., Meneghetti, M., Miyatake, H.,
270 Montes, M., Natarajan, P., Ntampaka, M., Pierpaoli, E., Postman, M., ... Wilson, G. (2024,
271 in prep.). *X-LoVoCCS. I. Initial X-ray properties of LoVoCCS galaxy clusters*.
- 272 Turner, David J., Donahue, M., & Stocke, J. (2024, in prep.). *An X-ray view of galaxy*
273 *groups in the line of sight of UV-bright quasars - comparing X-ray properties to absorption*
274 *detection of warm-hot gas*.
- 275 Turner, D. J., Giles, P. A., Romer, A. K., & Korbina, V. (2022). XGA: A module for
276 the large-scale scientific exploitation of archival X-ray astronomy data. *arXiv e-Prints*,
277 arXiv:2202.01236. <https://arxiv.org/abs/2202.01236>
- 278 XRISM Science Team. (2020). Science with the X-ray Imaging and Spectroscopy Mission
279 (XRISM). *arXiv e-Prints*, arXiv:2003.04962. <https://arxiv.org/abs/2003.04962>
- 280 Zhang, Y., Jeltema, T., Hollowood, D. L., Everett, S., Rozo, E., Farahi, A., Bermeo, A.,
281 Bhargava, S., Giles, P., Romer, A. K., Wilkinson, R., Rykoff, E. S., Mantz, A., Diehl,
282 H. T., Evrard, A. E., Stern, C., Gruen, D., von der Linden, A., Splettstoesser, M., ...
283 DES Collaboration. (2019). Dark Energy Survey Year 1 results: calibration of cluster
284 mis-centring in the redMaPPer catalogues. *487*(2), 2578–2593. [https://doi.org/10.1093/](https://doi.org/10.1093/mnras/stz1361)
285 [mnras/stz1361](https://doi.org/10.1093/mnras/stz1361)