PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

The XRISM science data center: optimizing the scientific return from a unique x-ray observatory

Loewenstein, Michael, Hill, Robert, Holland, Matthew, Miller, Eric, Yaqoob, Tahir, et al.

Michael Loewenstein, Robert S. Hill, Matthew P. Holland, Eric D. Miller, Tahir Yaqoob, Trisha F. Doyle, Patricia L. Hall, Efrem Braun, Chris Baluta, Koji Mukai, Yukikatsu Terada, Makoto Tashiro, Hiromitsu Takahashi, Masayoshi Nobukawa, Tsunefumi Mizuno, Takayuki Tamura, Shin'ichiro Uno, Shin Watanabe, Ken Ebisawa, Satoshi Eguchi, Yasushi Fukazawa, Ryo Iizuka, Satoru Katsuda, Takao Kitaguchi, Aya Kubota, Shinya Nakashima, Kazuhiro Nakazawa, Hirokazu Odaka, Masanori Ohno, Naomi Ota, Rie Sato, Yasuharu Sugawara, Megumi Shidatsu, Tsubasa Tamba, Yuichi Terashima, Yohko Tsuboi, Yuusuke Uchida, Hideki Uchiyama, Shigeo Yamauchi, "The XRISM science data center: optimizing the scientific return from a unique x-ray observatory," Proc. SPIE 11444, Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray, 114445D (13 December 2020); doi: 10.1117/12.2560840

SPIE.

Event: SPIE Astronomical Telescopes + Instrumentation, 2020, Online Only

The XRISM Science Data Center: Optimizing the scientific return from a unique X-ray observatory

Michael Loewenstein[a,b], Robert S. Hill[b,c], Matthew P. Holland[b], Eric D. Miller[d], Tahir Yaqoob[b,e], Trisha F. Doyle[b,f], Patricia L. Hall[b,f], Efrem Braun[b,c], Chris Baluta[b], Koji Mukai[b,e], Yukikatsu Terada[g,h], Makoto Tashiro[g,h], Hiromitsu Takahashi[i], Masayoshi Nobukawa[j], Tsunefumi Mizuno[i], Takayuki Tamura[h], Shin'ichiro Uno[k], Shin Watanabe[h], Ken Ebisawa[h], Satoshi Eguchi[l], Yasushi Fukazawa[i], Katsuhiro Hayashi[h], Ryo Iizuka[h], Satoru Katsuda[g], Takao Kitaguchi[m], Aya Kubota[n], Shinya Nakashima[h], Kazuhiro Nakazawa[o], Hirokazu Odaka[p], Masanori Ohno[i], Naomi Ota[q], Rie Sato[h], Yasuharu Sugawara[h], Megumi Shidatsu[r], Tsubasa Tamba[p], Atsushi Tanimoto[p], Yuichi Terashima[r], Yohko Tsuboi[s], Yuusuke Uchida[i], Hideki Uchiyama[t], and Shigeo Yamauchi[q]

[a]University of Maryland, Department of Astronomy, College Park, MD, USA
[b]NASA, Goddard Space Flight Center, Greenbelt, Maryland, USA
[c]ADNET Systems, Inc., 6720B Rockledge Drive, Bethesda, MD, USA
[d]Massachusetts Institute of Technology, Kavli Institute for Astrophysics and Space Research, Cambridge, MA, USA
[e]University of Maryland Baltimore County, Center for Space Sciences and Technology, Baltimore, MD, USA
[f]INNOVIM, LLC, 6401 Golden Triangle Dr, Greenbelt, MD, USA
[g]Saitama University, Graduate School of Science and Engineering, Sakura, Saitama, Japan
[h]Japan Aerospace Exploration Agency, Institute of Space and Astronautical Science, Sagamihara, Kanagawa, Japan
[i]Hiroshima University, School of Science, Higashi-Hiroshima, Japan
[j]Nara University of Education, Department of Teacher Training and School Education, Nara, Japan
[k]Nihon Fukushi University, Faculty of Health Sciences, Handa, Aichi, Japan
[l]Fukuoka University
[m]RIKEN Nishina Center, Wako, Saitama, Japan
[n]Shibaura Institute of Technology, Department of Electronic Information Systems, Saitama-shi, Saitama, Japan
[o]Nagoya University, Department of Physics, Nagoya, Aichi, Japan
[p]The University of Tokyo, Department of Physics, Tokyo, Japan
[q]Nara Women's University, Department of Physics, Nara, Japan
[r]Ehime University, Department of Physics, Matsuyama, Ehime, Japan
[s]Chuo University, Department of Physics, Bunkyo, Tokyo, Japan
[t]Shizuoka University, Faculty of Education, Shizuoka, Japan

ABSTRACT

The X-Ray Imaging and Spectroscopy Mission, XRISM, is currently scheduled to launch in 2022 with the objective of building on the brief, but significant, successes of the ASTRO-H (Hitomi) mission in solving outstanding

Further author information: (Send correspondence to M.L.)
M.L.: E-mail: michael.loewenstein-1@nasa.gov

Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray, edited by Jan-Willem A. den Herder
Shouleh Nikzad, Kazuhiro Nakazawa, Proc. of SPIE Vol. 11444, 114445D · © 2020 SPIE
CCC code: 0277-786X/20/$21 · doi: 10.1117/12.2560840

astrophysical questions using high resolution X-ray spectroscopy. The XRISM Science Operations Team (SOT) consists of the JAXA-led Science Operations Center (SOC) and NASA-led Science Data Center (SDC), which work together to optimize the scientific output from the Resolve high-resolution spectrometer and the Xtend wide-field imager through planning and scheduling of observations, processing and distribution of data, development and distribution of software tools and the calibration database (CaldB), support of ground and in-flight calibration, and support of XRISM users in their scientific investigations of the energetic universe. Here, we summarize the roles and responsibilities of the SDC and its current status and future plans. The Resolve instrument poses particular challenges due to its unprecedented combination of high spectral resolution and throughput, broad spectral coverage, and relatively small field-of-view and large pixel-size. We highlight those challenges and how they are being met.

Keywords: XRISM, Science Operations, X-ray Astronomy, Astronomical Software

1. INTRODUCTION AND OVERVIEW

The X-ray Imaging and Spectroscopy Mission, XRISM[1] (see, also, Tashiro et al., these proceedings) is a mission with international collaboration intended to recover most of the science capabilities of the short-lived Hitomi satellite.[2] XRISM is scheduled for launch in 2022. The scientific objective of XRISM is to investigate celestial sources in the soft X-ray bandpass (∼0.3–12 keV). XRISM has two instruments, the Resolve Soft X-ray Spectrometer, with 5–7 eV energy resolution and a $3' \times 3'$ Field-of-View (FoV) on 36 pixels; and the Xtend Soft X-ray Imager, with a $38' \times 38'$ FoV. X-rays are focused onto both detectors by identical lightweight X-Ray Mirror Assembly (XMA) telescopes with angular resolution of $1.7'$ or better. In addition to the Resolve microcalorimeter array, subsystems of particular importance for data analysis are the Resolve anti-coincidence (antico) detector used for identifying cosmic ray events, and the Filter Wheel (FW) and Modulated X-Ray Sources (MXS) mounted above the Resolve cryogenic dewar aperture assembly.[3] The MXS provide a means of pixel-by-pixel gain calibration without contamination of astronomical source data.

The charge of the XRISM Science Operations Team (SOT) is to optimize the science return from the mission and to support the mission PI in science operations. The SOT consists of the Science Operations Center (SOC) at JAXA (Terada et al., these proceedings) and the Science Data Center (SDC) at NASA. The responsibilities of the SOC include planning of spacecraft operations needed to fulfill approved observations, managing the observation database, support of prospective and approved Japanese (and some other) users of XRISM, and operating the data "pre-pipeline" (see below).

The SDC develops and operates the data processing pipeline; develops observation scheduling, pipeline, post-pipeline, and data analysis software; maintains the XRISM Calibration Database (CalDB); and supports the High Energy Astrophysics Science Archive Research Center (HEASARC) and its European and Japanese counterparts in providing general user support and Guest Observer Facility services. Software and CalDB releases to the XRISM science team occur on a 6–12 month cadence. As the launch of XRISM approaches, software will be included in the HEASoft software package that includes standard software for many high-energy astrophysics missions and runs on many commonly-used operating systems,* and the XRISM CalDB files will be ingested by the HEASARC calibration database system.† Software version control is maintained by the use of git repositories and a dedicated server. Design and requirement documents, user handbooks, and as-built guides are configuration-controlled using the NASA/Goddard Space Flight Center Instrument Projects Division Technical Data Management System. Issue tracking is realized using NASA/Goddard Space Flight Center Engineering and Technology Division Jira software.

In this contribution, we present the status and future plans of these aspects of SDC operations, with an emphasis on updates and improvements with respect to what was done for Hitomi.[4]

*https://heasarc.gsfc.nasa.gov/docs/software/lheasoft/
†https://heasarc.gsfc.nasa.gov/docs/heasarc/caldb/caldb_intro.html

2. DATA PROCESSING AND HANDLING

The pre-pipeline (Terada et al, thess proceedings) operated by the SOC converts spacecraft science and housekeeping (HK) telemetry stored in the form of Raw Package Telemetry (RPT) files into First FITS Files (FFF) that conform to the HEASARC/OGIP FITS standard.[5] The pre-pipeline also creates FITS files required for time assignment and FITS files that characterize the spacecraft orbit and attitude. These data are divided into distinct sequences (or "OBSID") that represent observations defined to span the time from the start of the slew from the previous target to that of the following target. Single observations may be split into multiple sequences for reasons of file size or instrument configuration. Time assignment (including time-tagging of X-ray events, HK data, and attitude and orbit files) is conducted in the pre-pipeline.

The SDC continues data processing from this stage by operating the XRISM processing pipeline, which calibrates and cleans the FFF data to produce higher level event files and to extract higher-level data products. The results are transferred to two archives: the NASA HEASARC, and the JAXA Data ARchive and Transmission System (DARTS).

The flow of data processing and transfer is illustrated in Fig. 1 and discussed in more detail below.

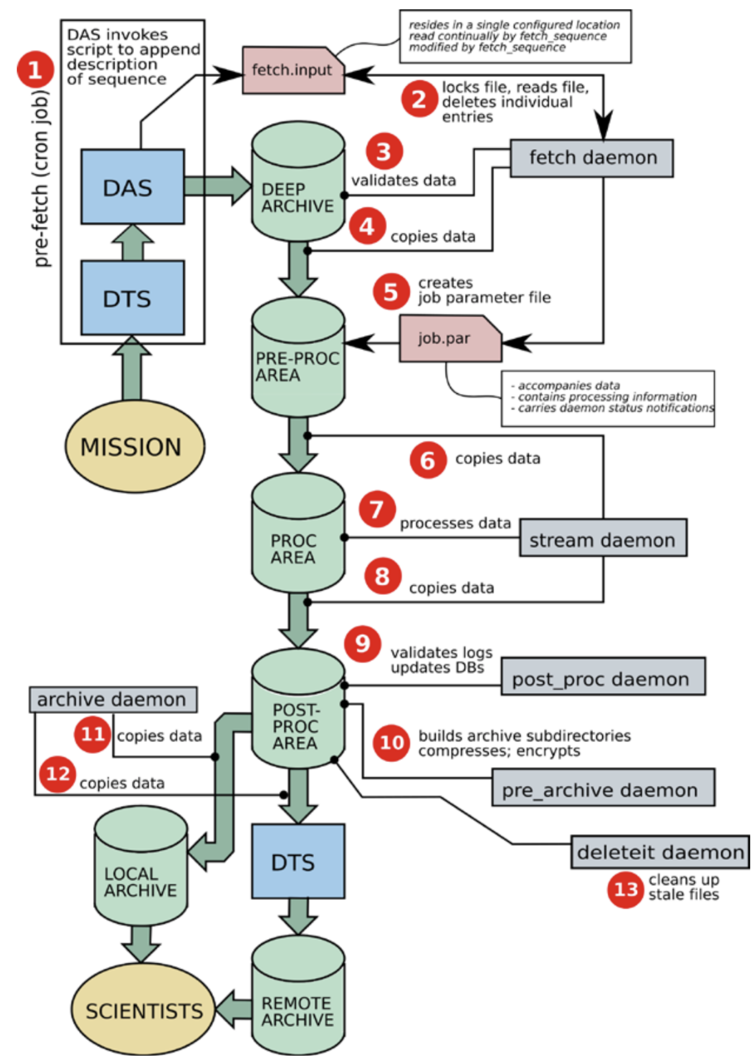

Figure 1. Data flow and processing steps applied by the SDC starting with the acquisition of FFF, and concluding with data delivery to scientific investigators. Intermediate steps where the data is copied for backup storage are also indicated, but not discussed further.

2.1 Pipeline Processing of XRISM Data

Following acquisition (Sec. 2.3) and validation (Step 1–4 in Fig. 1), the automated data pipeline prepares (Step 5) and processes (Step 6) the pre-pipeline data on a sequence-by-sequence basis, independently for Resolve and Xtend, in stages devoted to (1) calibration, (2) screening, and (3) preview product creation. Prior to these stages it applies an aberration correction to the attitude file, calculates the average on-source pointing direction, and derives orbital quantities needed for screening the data - such as those based on times of passage through the South Atlantic Anomaly, satellite elevation and orientation with respect to the Sun and Earth, and magnetic cutoff rigidity. These quantities, and others derived from general and instrument-specific HK file extensions, are stored in "extended housekeeping" and "make-filter" FITS files created in the pipeline that are queried during the screening stage.

In the calibration step, the processing pipeline populates (or creates, if necessary) columns in the FITS file needed to define higher level spatial coordinates on the celestial sphere, photon energy, and event quality whilst preserving all of the original telemetry information contained in the FFF files. Good time intervals (GTI) based on attitude, orbit, and HK information are defined at this step, and written to separate GTI extensions to event files. The GTI are applied to the event files, along with additional filtering based on event quality, to create final cleaned event files, with the unfiltered files maintained in order to facilitate re-screening. A standard set of preview products - light curves, images, spectra, and spectral response files - are extracted or constructed as part of the final data processing step. In addition to data associated with particular observations, trend data files used to monitor the performance and health of the XRISM instruments and subsystems are created. Trend files may be virtually identical to the original FFF, or may be subject to (full or partial) calibration, (possibly custom) screening, and product extraction. Antico and MXS processing are included in the Resolve component of the pipeline. Science files are organized into a standard directory structure for each sequence, consisting of cleaned, unfiltered, and product sub-directories for each detector, and an auxiliary sub-directory with instrument-independent HK and other general files. The trend data is, instead, organized into top-level directories defined first by month, then by instrument (or as instrument-independent), and finally by trend data type. Interface control documents specify file naming conventions, FITS file formats, and file disposition; and, define column and header keyword names and data types.

Pipeline processing of XRISM data is largely identical to that for Hitomi, aside from the removal of content associated with the two high-energy Hitomi instruments that have no XRISM counterpart. Most updates consist of (1) changes to code that deletes, adds, or updates FITS header in order to reconcile the output files with requirements documentation specifications, (2) minor bug fixes that do not affect the content of output event files, (3) rectification of filename assignment, and (4) changes associated with updates to software tasks and calibration files (such as new, or re-defined, input parameters or CalDB files/extensions) used in the pipeline.

The most substantial revision involves the assignment of energy to Resolve X-ray events. The uncalibrated pulse height amplitude (PHA) for each event that is stored in the Resolve FFF event files is calculated by the on-board digital pulse processing, and is converted to a photon energy by a sequence of pipeline tasks using one of the following three types of on-board calibration source. (1) Mn K-shell X-rays from a radioactive ^{55}Fe source are collimated to constantly illuminate a dedicated pixel offset from the main array pixels so that it is outside of the aperture (this replaces a non-active pixel on the array in the detector electronics). (2) The MXS illuminates the entire detector array, producing very rapidly pulsed X-rays with a low duty cycle to enable simultaneous, nearly continuous, energy gain calibration (using Cr and Cu K-shell lines) for each pixel without significantly interfering with the data from the observation. The MXS may be operated in either *direct* mode wherein irradiation of a primary target that directly illuminates the array produces strong emission features from K-shell transitions of Cu (Kα at 8.04, Kβ at 8.90 keV) and Cr (Kα at 5.41, Kβ at 5.94 keV), or in *indirect* mode wherein a secondary target produces weaker fluorescent features from Al (Kα at 1.49, Kβ at 1.56 keV) and Mg (Kα at 1.25 keV, Kβ at 1.30 keV).[6] (3) Finally, primarily as a backup, there is a ring of ^{55}Fe sources on the Filter Wheel that may be rotated into the optical path.[6] For Hitomi, the pipeline always applied the "calibration pixel" method (1) with an additional adjustment to account for pixel-to-pixel variation. Other, intermediate, "gain history", files were created when the relevant data were present (note that no observations with the MXS activated were conducted for Hitomi) to allow for energy re-calibration by investigators. For XRISM, energy calibration using the MXS, expected to be the most accurate of the three methods, takes precedence (which drove some re-arrangement of pipeline code). The gain method decision tree is illustrated in Fig. 2. All possible gain history files continue to be constructed. Comprehensive tests of the flow and application of MXS data in the pipeline are planned.

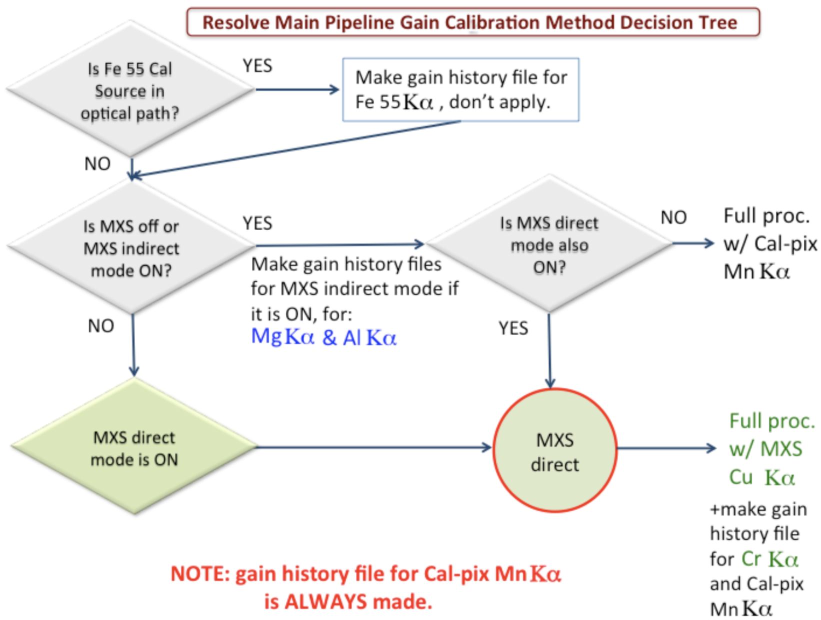

Figure 2. Pipeline procedure for calibrating the XRISM Resolve gain, and making gain history files, for the different calibration methods and lines.

2.2 Scheduling and Metadata

The term metadata refers here to information that characterizes an observation, from the proposal process through planning, observation, processing, and delivery to archives and XRISM investigators. As with Hitomi, and other previous X-ray missions such as ASCA, RXTE, Swift and Suzaku, these are maintained in an observational database (ODB). ODB entries are utilized in, and updated during, observation planning and scheduling (see below), and pre-pipeline and pipeline processing. The ODB is composed of proposal and observation components in relational database (rdb) format. The former includes information from accepted proposals, such as Principle Investigator (PI), proposal subject category; and, approved exposure time, pointing coordinates, instrument configuration, and any observational constraints for each source. The observation component reproduces the proposal data, and captures planned and as-flown data handling, attitude and orbit control, and instrument settings. In addition this component includes pre-pipeline and pipeline processing information (i.e., software and CalDB version numbers), as well as dates for data distribution to the PI and data archives, and the date when the data is due to become publicly available. The pipeline updates the ODB (Fig. 1 Step 9) and extracts ODB information to construct a master observation database file for each sequence, which is ingested by the HEASARC and DARTS for use in accessing the XRISM data archives by general users. These are delivered to the archives as a table in Transportable Database Aggregate Table (TDAT) format.

The XRISM Observation Scheduling Software (XOSS) system will be based on the system developed for Hitomi. The software builds on the systems used successfully for the ASCA and Suzaku missions. Schedules are constructed based upon observation efficiencies (where the efficiency is defined as the ratio of "good" on-target observation time to total observation time). Both satellite orbital constraints (SAA passages, target occultations, etc.) and operational constraints, as well as observation-specific constraints are calculated to define good (and bad) time intervals. While the scheduling software is GUI-based, it is designed to be flexible. Thus it provides a toolkit that may be used to calculate specific schedule-related information for an individual target, or about the spacecraft itself. The system interacts with the ODB in two directions, obtaining target metadata from, and registering target schedule information to, the ODB.

2.3 Data Transfer and Archiving

Pre-pipeline-processed and ODB data are transferred from the SOC at JAXA to the NASA SDC, and pipeline-processed data from the SDC to the HEASARC and DARTS, via xDTS (Steps 1, 11, and 12 in Fig. 1, labeled as "DTS") – the XRISM-modified version of the legacy Data Transfer System used for SWIFT and other missions. The xDTS system uses the sftp protocol for executing transfers and acknowledgments between server and receiver. Each site has its own identifier and each dataset its own unique label. The data are staged in a configurable location at the sending site (either the SOC or SDC). At the receiving site (the SDC, or the archives in Japan or the U.S.) data are checked for proper file transmission. A unique label transmitted with the data determines the subsequent routing.

Each database transmission includes the entire table, and not only the latest record updates. The database table precedes the data, to avoid having "orphan" sequences that cannot be processed or searched for in the archive. xDTS uses key pair authentication to allow the sender or the receiver to transfer the data using sftp. The receiving site polls a directory on the sending site for outgoing metadata messages indicating pending transfers, and copies them to its own storage for processing. The receiver then grabs the data files indicated by the metadata. ODB polling occurs at least once a day whether or not pre-pipeline data from a new observation is available. The master observation database files are transferred to the HEASARC and DARTS via xDTS, as is information that enables the HEASARC to notify the PI of the availability of the data and - for proprietary data - inform them of the PGP key needed to decrypt the data. The flow of data initiated via xDTS is illustrated in Fig. 3.

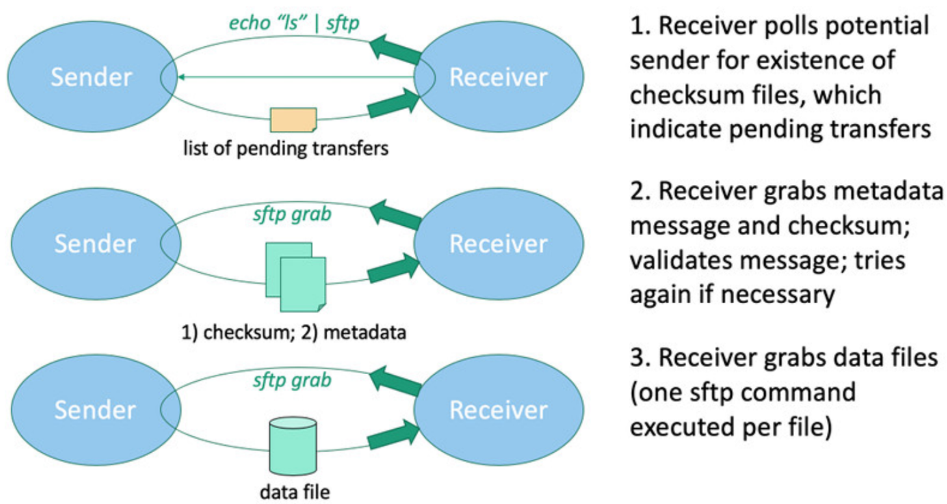

Figure 3. Data flow in xDTS.

xDTS supports the execution of pre-defined scripts based on the data type given at send-time. The Data Archive System (DAS, Step 1 and, also, 10 in Fig. 1) is one such script; it takes the staging directory and the transferred files as arguments, and then copies them into a destination directory, preserving directory structure and ensuring that file integrity is maintained.

Partially-integrated testing of the pipeline, initiated from JAXA/ISAS using xDTS and mock XRISM ingest FFF sequences prepared by the SDC, proceeding through receipt of the sequence and its flow through the pipeline daemon processes, and completed by a pulling of the pipeline output to JAXA/ISAS using xDTS, has been successfully conducted. Extensions of these tests will proceed as additional segments of the system become operational and integrated.

3. SOFTWARE AND CALIBRATION DATABASE FOR DATA PROCESSING AND ANALYSIS

XRISM software is configured for installation within the HEASoft environment, so that it may be included in the HEASoft software package around the time of the launch of XRISM. The calibration information is stored in files, in FITS format, and interfaces with the software via routines that query the Calibration Database (CalDB) metadata. The calibration files are ingested into the HEASARC, which stores and indexes datasets associated with the calibration of high energy astronomical instrumentation. Software and CalDB releases (or "Builds") to the XRISM science team occur on a 6-12 month cadence; three builds have been released, with a fourth in development. The Hitomi software tasks and CalDB files are described in detail in the HEASARC Hitomi Data Analysis Guide,[‡] and Description of the Hitomi Calibration Files,[§] respectively. Here, we describe some of the significant changes to tasks that were, are being, or are planned to be, implemented. Note that revisions to pipeline-related tasks and calibration files require parallel changes to the processing part of the pipeline, as well as to the pipeline processing tools that are included as part of the XRISM software and enable users to reprocess sequences. Care is taken to maintain synchronicity between these.

Since much of the XRISM software is similar to that developed for Hitomi, a dual-build software architecture is implemented in order to facilitate re-use of Hitomi code – with all tasks subsumed into a high-level library split into Hitomi, XRISM, and common code libraries (Fig. 4). All modifications developed by the SDC are limited to XRISM-related functionality.

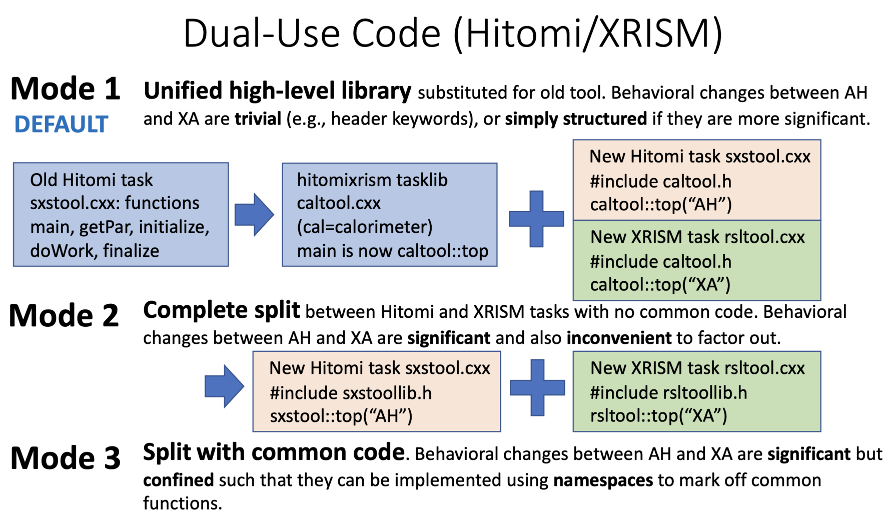

Figure 4. Hitomi/XRISM dual software build architecture.

There are over 60 active software tasks of relevance to XRISM data processing, data analysis, and observation simulation. Roughly 40% of these are mission independent, but may require XRISM-specific CalDB or other support files (i.e., for simulations). Another ∼ 20% are applicable to both Xtend and Resolve, ∼ 25% to Resolve only, and ∼ 15% to Xtend only. The SDC released Build 1, which introduced the dual build but was functionally equivalent to existing Hitomi Tools, in December 2018 to the XRISM science team. Build 2, released September 2019, included modifications to Resolve MXS, energy scale, and detector efficiency handling (software and CalDB). Build 3, released June 2020, focused on Xtend and Resolve redistribution matrix file (RMF) generation and Resolve event flagging, as well as Resolve identification of associations of secondary and primary grade (see below for definitions) events. Build 4, currently under development, will introduce greatly expanded capabilities for simulating the passage of X-rays through XRISM (and similar) optics and update the Xtend CCD charge transfer inefficiency correction algorithm used in energy assignment. Many of these changes are driven by new requirements introduced through interaction with the the Resolve (which includes XMA)

[‡] https://heasarc.gsfc.nasa.gov/docs/hitomi/analysis/hitomi_analysis_guide_20171214.pdf
[§] https://heasarc.gsfc.nasa.gov/docs/hitomi/calib/caldb_doc/ah_sct_caldb_v20170630.pdf

and Xtend instrument teams, as well as with members of the larger XRISM Science Team. Some of the more significant modifications are described in what follows.

3.1 Resolve Software Tools and CalDB

3.1.1 MXS Tasks and CalDB

The MXS timing task, *rslmxsgti*, and its driver script, *rslmxstime*, use Resolve HK information to determine whether pulsed MXS X-rays are illuminating the Resolve microcalorimeter array at any given time. This information is used to construct time intervals when the event file is to be used for pixel-dependent energy scale calibration and, inversely, to create cleaned event lists when the observation is mostly free from MXS contamination. The software must account for relative time offsets between the start of the pulse and when an LED is switched on or off, as well as that between the MXS and spacecraft clocks; and, a steeply declining - but non-negligible - post-pulse afterglow. An amplitude threshold parameter was added to these tasks that is used to define the effective afterglow duration. The new formalism accounts for possible differences between the LED start and stop offsets, for how these vary with LED current, and for the dependencies of the afterglow duration time on the threshold and pulse length (not including afterglow). The revisions were validated using specially-tailored HK test files, and provisional CalDB data provided by the Resolve instrument team that describes these newly-considered dependencies of pulse offsets and afterglow. Additional software testing and CalDB updates, and possible code updates, will be driven by future ground testing that includes the MXS subsystem.

3.1.2 Resolve Event Flagging

The *rslflagpix* task sets a 16-bit STATUS column in Resolve FITS event files, identifying Resolve events that may be affected by antico and MXS event coincidence, temporal proximity to other events, and several types of crosstalk. Several revisions to the *rslflagpix* tool have been made to improve the flexibility of the tool, fix some minor bugs, and implement changes recommended by the Resolve instrument team. These include the addition of an option to select separate energy threshold requirements for the setting of different flags, the introduction of a new pulse height ratio threshold for flagging electrical crosstalk, and rectification of the energy conditions for potential crosstalk with X-rays from the collimated ^{55}Fe calibration-pixel source. In addition, the task now supports an update to the CalDB data that relaxes the assumption that the two redundant output channels of the anti-coincidence detector have identical relative time-tagging with respect to the microcalorimeter array.

3.1.3 Resolve Energy Scale

Software modifications were made to implement a unified calibrated energy range for Resolve events. Hitomi SXS events were calibrated by default up to a high-energy limit of ~ 16 keV, with special processing using the *sxsextend* tool required to expand beyond this to energies where calibration requirements are less stringent - but sensitivity of the detector + mirror system non-negligible for some possible applications. The energy channel of the extended energy scale had been recorded in a separate event file column from that of the standard scale, requiring the user to specify which scale to use. For XRISM a different approach is adopted in which a single unified set of calibrated event energies is created with a default high-energy limit of 30 keV. The *rslpha2pi* task, which assigns energy to X-ray events for Resolve, was also modified to automatically handle the possibility of a modified gain for the detector system operating in cryogen-free mode (i.e., when mechanical cryo-coolers are exclusively used), which may be implemented following the exhaustion of liquid Helium in the Resolve Dewar.

3.1.4 New Grade and Pixel Dependencies

All Resolve events are assigned both a pixel number and a resolution grade by on-board processing. Grades are determined by the combination of the time interval to the proximate event in the same pixel (**H**igh-, **M**id-, or **L**o-res), and the proximate *preceding* event in the same pixel (**p**rimary, or **s**econdary). The separations that define these, in units of ms and (80 μsec) pulse samples, are shown in Fig. 5. A number of Resolve calibration tasks are being generalized to account for known, anticipated, or prospective variations in detector performance as a function of grade or pixel that were not accounted for by Hitomi software.

Unlike the case for Hitomi, two distinct absorber thicknesses, which, along with their filling factor, determine the quantum efficiency (QE), are represented in the XRISM Resolve pixel array. New QE CalDB files are

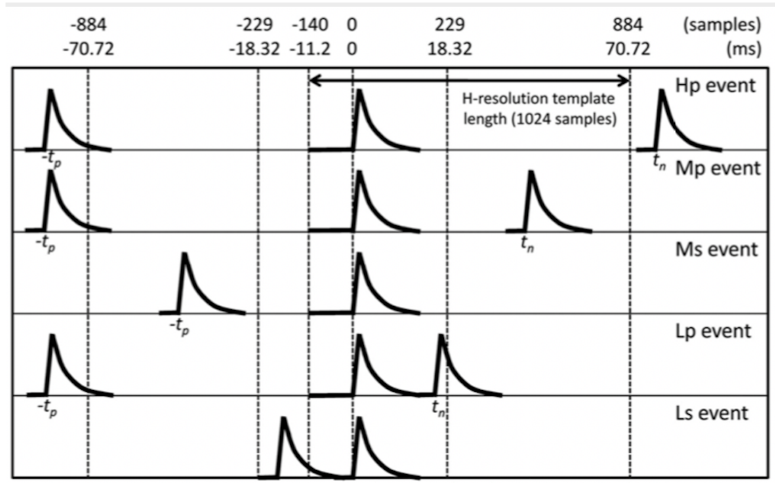

Figure 5. Schematic view of grade definitions for pulses arriving at time t=0.[7] The times tp and tn denote the previous and next pulse, respectively.

now formatted and structured to encapsulate arbitrary pixel-to-pixel variations, and software tasks and library subroutines that account for the effect of QE on total effective area and exposure maps used for flat-fielding images were updated accordingly.

The CalDB file that stores various time-related quantities was updated to generalize the coefficients that, for a given Resolve pulse-shape, convert times when an event triggers the on-board detection processing chain to times when the photon actually arrives at the detector. The code that applies these coefficients was updated accordingly, allowing for distinct corrections for primary and secondary events, in addition to the previously supported variations by pixel and whether the event was of high-, mid-, or low-resolution.

Finally, there are new pixel and grade dependencies to the parameters that describe the Resolve Line Spread Function (LSF), which are described in the following subsection.

3.1.5 Resolve Redistribution Matrix File Generation

The Redistribution Matrix File (RMF) tabulates the probability that an input photon of a given energy will be detected in a given output PI channel; that is, the response of the instrument to photons. The *rslrmf* task, and its driver script, *rslmkrmf*, calculate the RMF for a selection of Resolve pixels and resolution grades. The task uses the LSF parameters stored in multiple extensions of the RMF parameters CalDB file to calculate an RMF for each pixel and each grade, which may then be weighted and combined based on a given distribution over pixel and grade. These weighting factors may be obtained from a Resolve event file to construct an RMF file matched to the spectrum extracted from that event file.

The LSF is assumed to consist of (1) a Gaussian core (the dominant component) with a pixel-dependent and grade-dependent FWHM, (2) a pixel-dependent low-energy exponential tail due to energy loss at the surface of the absorbers, (3) a pixel-dependent extended low-energy electron loss continuum (ELC), (4) pixel-dependent discrete escape peaks from Hg and Te fluorescence in the absorber, and (5) a pixel-dependent Si Kα fluorescence emission line. In Hitomi software, only the Gaussian core was considered pixel-dependent. Revisions to the software handling and application of the CalDB data have been implemented to realize these generalizations. Only the core component is grade-dependent, and is identical for primary and secondary events, i.e. dependent only on whether the event is H, M (Mp/Ms), or L (Lp/Ls). However, the software now accommodates separate calibration data for Mp and Ms events while retaining the option to apply Mp calibration to Ms events. Finally, the ELC is no longer (necessarily) a constant function of energy, but consists of tabulated functions for a set of input energies that span the full Resolve input range up to 30 keV. Again, this is reflected in changes to the structure and content of the RMF parameters CalDB file, and also in additional code required to read and utilize these data.

The RMF generator now supports the option to split the output RMF into two files, or one file with two extensions, with separate input energy grids. This new feature is driven by the fine spectral resolution of Resolve, which, along with the broad sensitivity, results in a large number (60,000) of energy channels. One component of the split response contains only the ELC response, and the other contains the remaining response components. Since the structure of standard RMF files allows for efficient storage of sparse matrices,¶ by defining a relatively coarse grid for the dense ELC response and a standard grid for the sparse remaining components, the composite matrix file size may be greatly reduced whilst retaining the necessary high energy resolution. This is illustrated in Fig. 6 for an ELC binning factor of 8, which reduces the size of the RMF FITS file by a factor of ~ 15. Tests demonstrate that spectral fitting run times are reduced by a factor of ~ 5 with no loss of fitting precision. This allows for faster handling by spectral fitting software. The Xspec package‖ was updated to automatically support such multiple-extension response files.

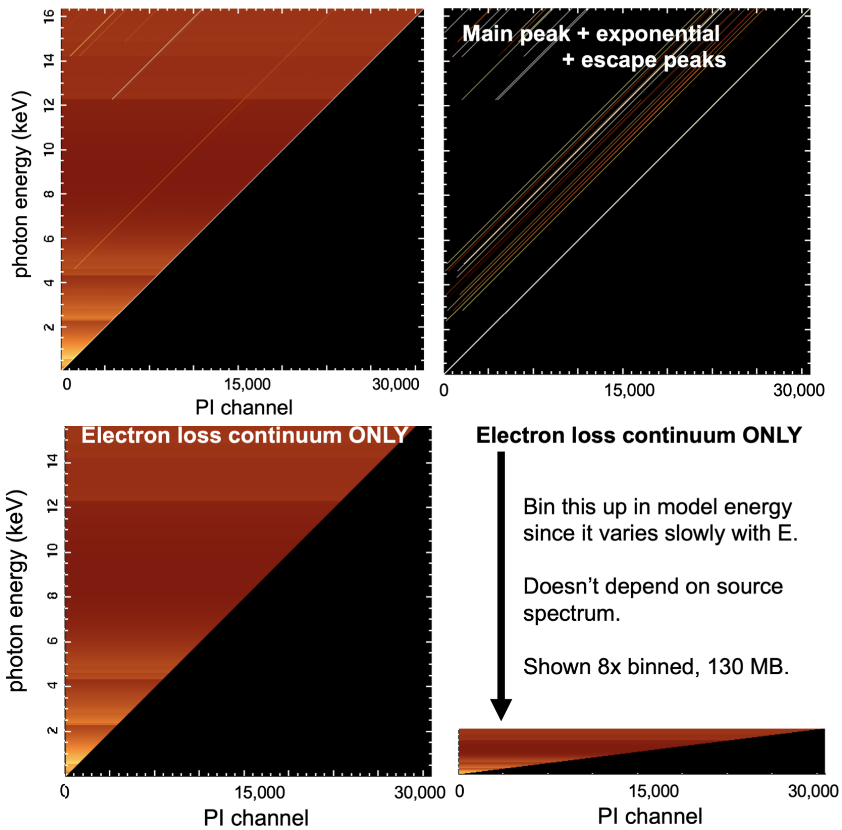

Figure 6. Illustration of the split RMF formalism, showing heat maps of spectral re-distribution matrix from input energy to 0.5 eV output energy ("PI") channel. Top left: all components; top right: main Gaussian peak plus escapes peaks - components with narrow width, but sparse coverage; bottom left: slowly varying ELC filling the matrix up to the corresponding channel value; bottom right: ELC binned by a factor of 8.

3.2 Xtend Tasks and CalDB

Unlike Resolve, the Xtend CCDs have several "clocking" modes to increase the dynamic range of target flux that can be observed without suffering pile-up. These modes must all be supported by the software and CalDB, and calibration can vary between modes. The supported modes are expected to be similar to SXI: Full Window mode, in which the entire CCD is read out after a 4-s integration time; 1/8 Window mode, in which a full-width, 1/8th-height region centered on the aimpoint is read out every 0.5 s; Full Window Burst mode, in which a Full

¶https://heasarc.gsfc.nasa.gov/docs/heasarc/caldb/docs/memos/cal_gen_92_002/cal_gen_92_002.html
‖https://heasarc.gsfc.nasa.gov/docs/xanadu/xspec/index.html

Window integration is shortened to 0.1 s followed by 3.9 s of dead time; and 1/8 Window Burst, in which the 1/8 Window integration is 0.1 s followed by 0.4 s of dead time. Full Window mode is supported on all four Xtend CCDs, but the other clocking modes are only allowed on the CCD pair that includes the aimpoint.

Since Xtend is closely based on Hitomi's SXI, the bulk of the software tools and CalDB formats employed for SXI can be reused for XRISM. There are two changes to the SXI code and CalDB formats due to alterations in correction algorithms, described below.

3.2.1 Xtend Redistribution Matrix File Generation

The Suzaku XIS instrument was plagued throughout its mission by an issue that caused spectral residuals of up to 10% near the Si-K edge at 1.839 keV. This issue has recently been resolved by introducing a discontinuity in the energy dependence of the primary response peak at that energy;[8] this is explained qualitatively by the sudden change in attenuation length of X-rays across this energy, and the resulting change in charge drift distance through the depleted bulk of the CCD. Since this effect is common to all CCDs, and since the SXI and Xtend RMF algorithms are based on the XIS method, the Xtend Instrument Team updated the algorithm to introduce an energy dependence in the central PI channel value of the main response peak in the Xtend RMF generation task *xtdrmf*. The effective PI value of the main Gaussian peak as a function of energy is now included in the Xtend RMF parameters CalDB file. Updates to the *xtdrmf* code that extract the new information and utilize it to improve the response fidelity near the Si K edge were implemented and tested in Build 3.

3.2.2 Xtend Charge Transfer Inefficiency and Energy Assignment

The Xtend *xtdpi* task applies a series of pixel-based calibration corrections to each Xtend event, determines the event grade, sums the pixel values to obtain a summed event pulse height, and applies an overall gain correction to convert pulse height to a value proportional to the incoming photon energy. One set of minor changes being implemented for Build 4 is designed to facilitate diagnostic analysis of calibration data, and correct some problematic behavior for non-standard sets of parameters.

The *xtdpi* task includes an algorithm for calculating charge transfer inefficiency (CTI), which depends on the rate at which charge is transferred in parallel columns in the CCD. At the end of an exposure, the charge is transferred quickly from the Imaging Area (IA) to the Frame Store (FS). The CTI parameters used by *xtdpi* were expanded to account for differences in the CTI characteristics between the IA and FS. This enables the use of a common set of probabilities for different windowing modes.

Finally, columns were added to several Xtend CalDB files, and code implemented to utilize these, to mitigate possible divergences in functions used to calculate the charge trail, CTI, and grade offset correction functions.

3.3 XMA Tasks and CalDB

The multi-mission *xrtraytrace* task is a standalone raytracing code that simulates the passage of photons through a thin-foil-type X-ray telescope, from the aperture through to the focal plane. It is used to calculate the XRISM mirror effective area, and hence is a critical component of data analysis. Significant enhancements to the raytracing software suite and associated CalDB files are underway, and a powerful new driver tool *xmasim* is being developed that will introduces new capabilities and enable the calculation of statistical errors on key quantities calculated in raytracing simulations.

The enhancements to *xrtraytrace* include (1) accounting for a new inner foil structure in the XRISM telescopes that is designed to reduce stray light impact on the innermost secondary mirror foil (Fig. 7), and (2) complex modeling of the thermal shield (including non-uniformity) to improve the accuracy in the calculation of the loss in effective area due to the thermal shield.

The *xmasim* tool is a sophisticated driver of *xrtraytrace* that provides (1) the ability to include objects such as the Resolve gate valve and filters placed in the optical path between the telescope body and the focal plane, (2) the ability to construct a series of images of the unconverged beam above the focal plane for use in investigative studies (Fig. 8), (3) a new type of compact output file that will be used to pre-calculate a library grid over off-axis and azimuthal angles of energy-dependent PSF images, (4) inclusion of detector effects, (5) parallelization to enable distributed computation for very high statistics runs, and (6) calculation of statistical errors on the effective area.

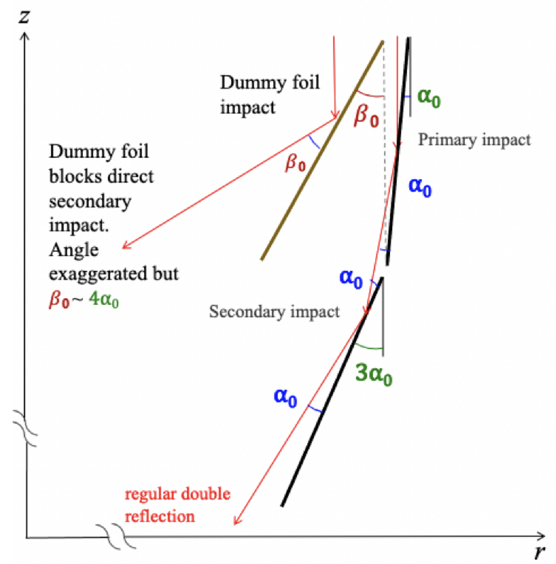

Figure 7. Illustration of effect of new inner ("dummy") foil structure in the XRISM telescopes designed to reduce stray light impact on the innermost secondary mirror foil. Updates to *xrtraytrace* are designed to account for this.

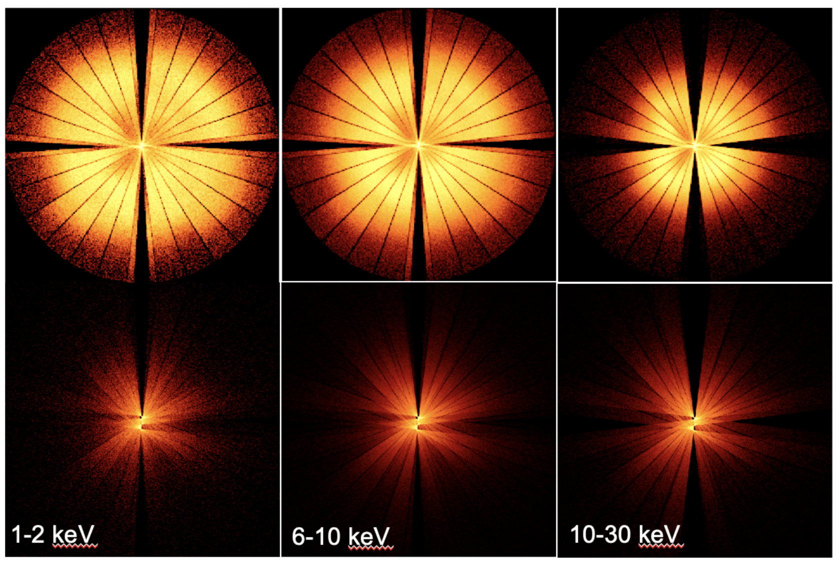

Figure 8. Photon distribution at the height of the Resolve Filter Wheel Be filter, approximately 0.9 m above the focal plane (top) which may be calculated using *xmasim*, compared to that at the focal plane (bottom). Images for three distinct energy bands are shown.

Parallel enhancements are also planned for the tool that generates the XRISM auxiliary response function (ARF) - which calculates the effective area for a specific observation and region in the sky, including Resolve and Xtend detector efficiencies. The updates introduce options for distributed computation of the effective area, with statistical errors, and on-the-fly raytracing of the Resolve filters and gate valve.

Compilation of a high statistics Resolve Point Spread Function (PSF) library using *xmasim* will enable spectroscopy that accounts for cross-contamination of spatial regions due to the PSF spatial redistribution. Additional work will include effects of dust scattering on the PSF and effective area. A graphical overview of the suite of raytracing tools is shown in (Fig. 9).

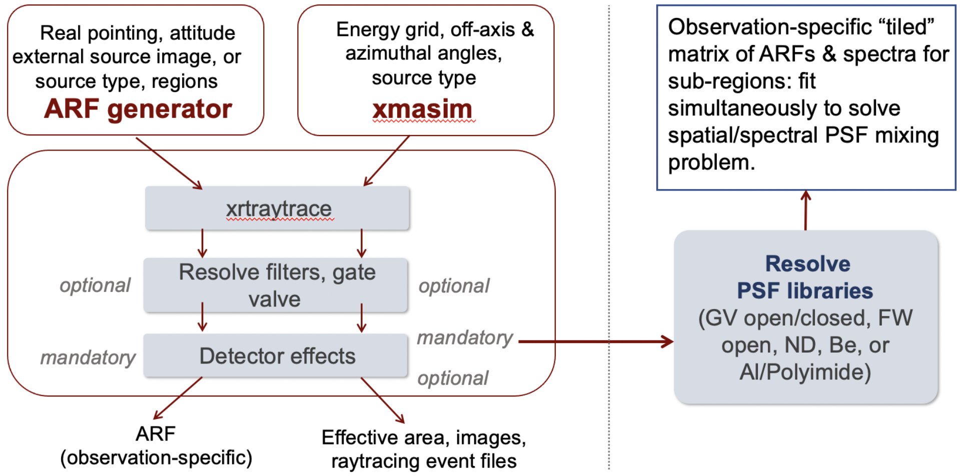

Figure 9. Graphical overview of the Suite of raytracing tools.

3.4 Non X-ray Background Estimation

The Non X-ray Background (NXB) due to interaction of high-energy particles (and γ-ray photons) is a limiting source of systematic error for Xtend science for faint point sources and low surface-brightness regions of extended sources. The SDC is working with the SOC and XRISM Instrument Teams in planning NXB database generation, and in providing flexible software tools and recipes to access it, construct spectra, and correctly apply it in spectral fitting.

These stakeholders have agreed to a preliminary set of requirements as to how NXB data will be accumulated, processed, indexed, structured according to instrument modes and configurations, and accessed. This will involve updating of the Hitomi algorithm and code to generate XRISM NXB spectra, and development of new tools used to construct (and re-construct) the database and provide access to to it by users of the tools. Work on these tasks will commence after the release of Build 4.

3.5 X-Ray Spectral Line IDentifier and Explorer

The X-ray Spectral Line IDentifier and Explorer (XSLIDE) is designed to be a simple and intuitive graphical user interface that allows astronomers inside and outside the field of high-resolution X-ray spectroscopy to interact with XRISM data. Users of XSLIDE may conduct quick browsing and surveying of proprietary or publicly available XRISM spectra as a prelude to application of more rigorous spectral fitting tools, and as a convenient source of displaying and characterizing spectra for purposes such as writing proposals or giving presentations.

XSLIDE (1) reads in and plots X-ray count-spectra and effective area files; (2) fits the underlying continuum of the fluxed spectra and plots the fit; (3) detects and indicates peaks (both positive and negative, i.e. associated with emission and absorption spectral lines); (4) compares the peaks with atomic lines catalogued in a user-selected database of known transitions and identifies matches; and, (5) performs simple plasma diagnostics based on a selection of line ratios. The current implementation, including the GUI, is realized in Python, using object-oriented principles. Analysis is separated from the interface to facilitate scripting and development.

XSLIDE is currently in the prototype phase (Fig 10). It has been revised from a previous version made for Hitomi, with improvements that optimize the code and reduce computation time, and the addition of unit tests.

Its final design, deployment as a desktop and/or web application, method of integration with data archives, and complexity of available diagnostics are currently under consideration by the SDC with input from the XRISM Science Team and other experts in high resolution X-ray spectroscopy. Prototypes of browser-based versions of XSLIDE are being constructed as part of this effort.

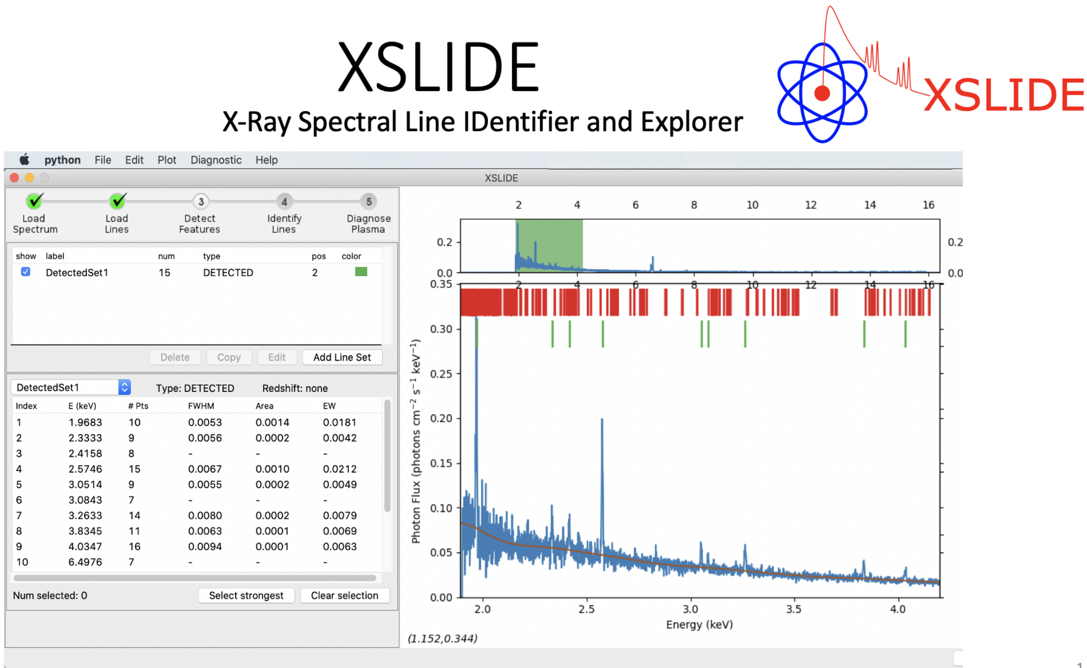

Figure 10. XSLIDE prototype interface.

4. SOFTWARE DEVELOPMENT AND TESTING PROTOCOLS

XRISM software development proceeds along the following steps. (1) At the *prototyping* stage, SDC scientists provide SDC software developers with change-request documentation that specifies new requirements, demonstrates (often via a tested software prototype) proof-of-concept, and includes validation and verification requirements methodology. (2) At the *development* stage, code and documentation are revised and developers conduct tests to verify that the updated software meets the requirements. (3) At the *reconciliation* stage, nearly-complete updates are independently validated and any outstanding issues are iteratively identified and resolved by developers and scientists.

Doxygen** is used to generate HTML and PDF versions of XRISM software descriptions from annotated source code. These will be distributed as part of software releases, allowing users to explore the structure of the dual build, and details of software design and algorithms.

The complete spectrum of readiness tests directed to running the pipeline on real data and analyzing data products includes several nested classes, which are listed here in order of granularity from the widest to the narrowest scope:

1. Integrated pipeline tests
   (a) End-to-end tests (Fig. 1)
   (b) Standalone tests of the processing script
2. Comprehensive Functional Tests (CFT) of tools
   (a) Long tests
      Multi-step tests of multiple tools using custom scripts
      Single-tool unit tests with realistic input file
   (b) Limited Functional Tests (LFT)
      Short unit tests of tools
      Short unit tests of libraries

** https://www.doxygen.nl/index.html

3. Ticketed tests
    (a) Bug fix verification
    (b) New feature verification

Comprehensive Functional Tests (CFT) and Limited Functional Tests (LFT) are shared outside the pipeline development team. CFT function as acceptance tests for the XRISM tools when they are integrated into the publicly available HEASoft software package, and LFT (similar to Hitomi unit tests) are available to end users to verify the XRISM software installation. Ticketed tests, which follow implementation of requests, may become permanent by promotion to CFT. The CFT are especially significant, because they are the most fundamental class of tests defining successful implementation of the XRISM SDC software. The individual tests in the class are controlled via a test matrix. Each test is part of a soundness proof for the tool being tested. The functionality of each tool is defined in documentation that takes the form of help files, which in practice define the capabilities, inputs, and outputs for each tool in scientific terms. The test matrix for a tool associates each test with a publicly asserted behavior of the tool. The relationship between test classes is illustrated as a "test stack" in Fig. 11.

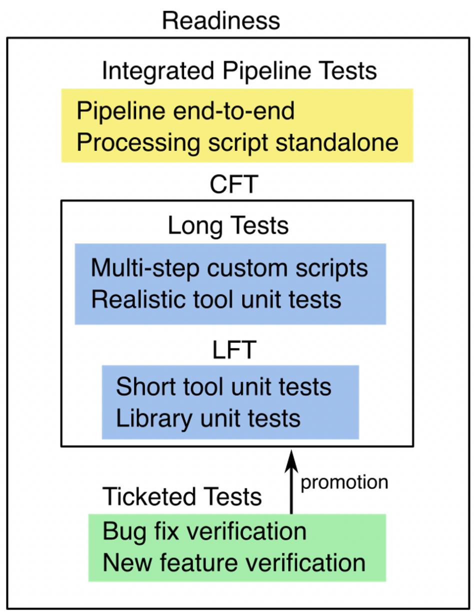

Figure 11. XRISM pipeline test stack illustrating the relationship between test classes.

Validation, as well as being a practical means to quickly clarify all manner of data system interface issues, is facilitated by construction of a XRISM "mini-mission" archive, currently consisting of converted archival Hitomi flight data plus data from GSE (Ground-test Support Equipment) testing of the Resolve and Xtend instruments. Since the provided ground data sets are based on detector-only testing, they lack the full complement of files expected to compose the content associated with an observation. These are, therefore, anchored to a basis Hitomi OBSID from which files such as those with orbit, attitude, and spacecraft HK data are adopted. Mini-mission data sets are shared with the rest of the XRISM Science Team for the purposes of observation planning, and software testing and training. Conversion procedures include transformation in accord with the re-definition of epoch (which was 2014 for Hitomi, but is 2019 for XRISM), and file format and naming conventions laid out in the XRISM FITS Definition and Archive documents. As needed, and when practical, re-assignment of quantities to synchronize the data with updates to software and CalDB files is performed. As a result, the mini-mission archive evolves in step with the SDC Build releases. The conversion is also applied to FFF-like data that represent pre-pipeline output. These may be ingested by the pipeline for testing, and serve as living examples of the FFF definition documentation. Observation database (ODB) spreadsheets may concurrently be set up for each data set. Place-holder values are used where needed, but the layout and general information included in the ODB adheres to the latest XRISM template.

5. THE SDC AND USER SUPPORT

XRISM user support in the U.S. will be directed by the Guest Observer Facility (GOF) in close collaboration with the HEASARC. HEASARC will host the U.S. XRISM archive, and provides shared infrastructure for software and calibration data for multiple missions that will include XRISM. GOF activities hosted by the HEASARC include (1) response to Help Desk tickets, (2) supporting maintenance of the mission website, and (3) providing user analysis guides and/or science threads. The GOF will also provide material in support of, and facilitate the review of, Guest Investigator proposals.

The SDC has provided spectral response and NXB files, as well as accompanying instructions, to the Science Team to assist with simulations needed for Performance and Verification Phase observation, and in-flight calibration planning (Miller et al., these proceedings). The SDC will continue to provide XRISM-specific support files for simulations conducted with the multi-mission *heasim* tool, and plans to update the *rslbranch* tool with added functionality for bright sources.

6. ACKNOWLEDGMENTS

The work of the SDC is supported through continual interaction with the XRISM Resolve and Xtend instrument teams, and benefits from feedback and guidance from the XRISM Science Management Office, NASA/GSFC XRISM Project Office, and XRISM Science Team. We gratefully acknowledge Michael Witthoeft for constructing a working XSLIDE prototype, and for his advice on its improvement; and, to Lorella Angelini for laying down the foundation of the SDC through her co-leadership of the the Hitomi Software and Calibration team.

REFERENCES

[1] Tashiro, M., Maejima, H., Toda, K., et al., "Concept of the X-ray Astronomy Recovery Mission," in [*Space Telescopes and Instrumentation 2018: Ultraviolet to Gamma Ray*], den Herder, J.-W. A., Nikzad, S., and Nakazawa, K., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **10699**, 1069922 (July 2018).

[2] Takahashi, T., Kokubun, M., Mitsuda, K., et al., "Hitomi (ASTRO-H) X-ray Astronomy Satellite," *Journal of Astronomical Telescopes, Instruments, and Systems* **4**, 021402 (Apr. 2018).

[3] Leutenegger, M. A., Audard, M., Boyce, K. R., et al., "In-flight verification of the calibration and performance of the ASTRO-H (Hitomi) Soft X-ray Spectrometer," *Journal of Astronomical Telescopes, Instruments, and Systems* **4**, 021407 (Apr. 2018).

[4] Angelini, L., Terada, Y., Dutka, M., et al., "Astro-H/Hitomi data analysis, processing, and archive," *Journal of Astronomical Telescopes, Instruments, and Systems* **4**, 011207 (Jan. 2018).

[5] Hanisch, R. J., Farris, A., Greisen, E. W., et al., "Definition of the Flexible Image Transport System (FITS)," *Astromomy & Astrophysics* **376**, 359–380 (Sept. 2001).

[6] De Vries, C. P., Yamasaki, N., den Herder, J., et al., "Calibration sources and filters of the soft x-ray spectrometer instrument on the Hitomi spacecraft," *Journal of Astronomical Telescopes, Instruments, and Systems* **4**, 021404 (Apr. 2018).

[7] Ishisaki, Y., Yamada, S., Seta, H., et al., "In-flight performance of pulse-processing system of the ASTRO-H/Hitomi soft x-ray spectrometer," *Journal of Astronomical Telescopes, Instruments, and Systems* **4**, 011217 (Jan. 2018).

[8] Okazaki, K. et al., "The spectral response of X-ray CCDs in the energy band around Si-K edge: a solution to the Si-K edge problem for the XIS onboard Suzaku," in [*High Energy, Optical, and Infrared Detectors for Astronomy VIII*], Holland, A. D. and Beletic, J., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **10709**, 107091F (July 2018).