Keck DEIMOS Archive of High Redshift Candidate Galaxies

Contributed Datasets Description

These contributed datasets represent one of the final products of the NASA ADAP project "Automated Reductions, Analysis, and Dissemination of the Keck DEIMOS Data Archive: Science-Ready Spectra in Key Fields" (PI: B. Holden). The other primary outcome of this project is "The Keck/DEIMOS Stellar Archive". The overarching goal of the proposal was to provide the scientific community with science-ready spectra obtained with the Keck DEIMOS spectrograph, enabling a wide range of astrophysical investigations.

The datasets described here are drawn from the Keck Observatory Archive (KOA) and consist of deep DEIMOS observations targeting high-redshift galaxy candidates. They were reduced using the PypeIt software package, which produced science-ready two- and one-dimensional spectra, following the workflow described below.

Data Selection, Retrieval, and Organization

The selection of datasets was carried out by searching the KOA archive for publicly available DEIMOS observations whose proposal titles explicitly referenced studies of high-redshift galaxy candidates. To further refine the target selection and enhance the scientific readiness of the datasets, we cross-matched the KOA-selected targets with an extensive photometric catalog. This catalog — kindly provided by Gabriel Brammer and not yet fully public — is part of the Complete Hubble Archive for Galaxy Evolution (CHArGE) project, which aims to uniformly process a large volume of archival Hubble Space Telescope data (and Spitzer/IRAC data when available), including imaging and slitless spectroscopy relevant to the study of distant galaxies.

From an initial sample of nearly 500 DEIMOS slitmasks associated with high-redshift galaxy studies, after excluding incomplete datasets — defined as science frames lacking corresponding calibration

frames, or calibration frames without associated science frames - and performing the HST-based cross-matching, we identified 323 unique slitmasks drawn from 81 observing programs, led by 23 principal investigators and spanning semesters from 2002B through 2020A.
In total, we reduced 1,928 science frames, corresponding to 21,698 individual slits. These observations span four of the five DEIMOS gratings, the 600ZD, 830G, 900ZD, 1200G.

All raw data files that satisfied the selection criteria, including both science and calibration frames, were downloaded from the KOA using the Python Interface to KOA Data, PyKOA. Once downloaded, the data were stored in a S3 cloud environment provided by the National Research Platform's Nautilus Cluster. A set of custom scripts, available in an associated GitHub repository, was then used to reorganize the data, identify complete datasets, perform single-frame reductions, and execute additional processing steps such as 1D and 2D coaddition and flux calibration.

The `adap_reorg_setup.py` script organizes the raw data by slitmask name using the following directory structure: <SLITMASK_NAME>/<GRATING_DISPANGLE_FILTER>/<DATE>/. During this organizational step, only complete datasets are retained for reduction. A dataset is considered complete if all required calibration files are available within a three-day temporal window. The following specific requirements were imposed.

- Arc Frames:
  - Either a single arc frame or a combination of arc frames are used to ensure that all available arc lamps are included. If all lamps are present in one frame, that frame alone is used, even if additional arc frames exist. If the lamps are distributed across multiple frames, we combine frames as needed, prioritizing those that contain the largest number of lamps.
  - If more than one frame satisfy the requirement above, preference is given to arc frames with exposure times between 1 and 15 seconds and elevations near 45°.
- Flat-Field Frames:
  - Three quartz lamp flat-field frames with exposure times between 5 and 30 seconds, again prioritizing observations taken at elevations close to 45°.

Bias and dark frames are not included in the reduction process. Instead, the overscan regions of the detectors are used to subtract the bias level.

Once the data are organized, the PypeIt's `pypeit_setup` script is executed to generate the configuration file (the *.pypeit* file), which defines the parameters guiding the reduction process (see more details here).

Single-Frame Data Reduction with PypeIt

Each individual science frame is reduced to a one-dimensional, wavelength-calibrated spectrum using PypeIt version 1.17.2. This automated pipeline performs wavelength calibration, flat-field correction, two-dimensional sky subtraction, object detection, and spectral extraction (see more details here).

PypeIt processes the eight DEIMOS detectors as four mosaic images, where each red/blue pair of detectors is reduced together. The blue detector image is interpolated onto the pixel coordinate system of the red detector using a fifth-order affine transformation that accounts for spatial and spectral shifts. While this interpolation slightly enlarges cosmic rays and other detector artifacts in the blue detector, resulting in larger masked regions compared to the red detector, it significantly improves the wavelength solution and object tracing by solving for both detectors simultaneously. This is now a standard procedure for DEIMOS reductions in PypeIt.

We employ the PypeIt "full template" wavelength calibration algorithm. To estimate the approximate wavelength range, PypeIt first cross-correlates the observed arc spectrum with an archived, calibrated arc spectrum corresponding to the relevant grating. Individual arc lines are then identified, and a sixth-order Legendre polynomial is fit to derive the final wavelength solution for each slit. A subsequent correction is applied to account for spectral flexure between the afternoon arc exposures and the nighttime science observations, using the extracted sky spectrum.

For each slit, sky subtraction and object detection are performed iteratively. An initial sky subtraction uses all pixels in the slit to estimate and subtract the sky background. The resulting sky-subtracted image is collapsed along the wavelength axis, and a preliminary object detection is carried out using a signal-to-noise (S/N) threshold of 5. A second sky subtraction is then performed with

object pixels masked, followed by a second object detection pass. Objects are traced and extracted using both boxcar extraction — with a boxcar width of $2''$ — and optimal extraction following the Horne algorithm. For object detected with S/N ≥ 4, the optimal extraction performs a b-spline fit to the object profile, otherwise the code will use a simple Gaussian fit with full width at half maximum (FWHM) determined from the object finding. A final sky subtraction is performed after the second object detection step.

The extracted spectra include both targeted program objects and serendipitously detected sources. To assign sky coordinates and slitmask design identifiers, PypeIt matches slit positions to the input DEIMOS slitmask design information. A detected object is classified as a program target if its measured distance from the left edge of the slit is within $0.5''$ of the predicted position. Objects falling outside this tolerance are classified as serendipitous detections, and their sky coordinates are computed as offsets from the slit center.

If a programmed target is not detected above the specified S/N threshold, an extraction is forced at the predicted position. In such cases, the object's FWHM for the optimal extraction is set to the average FWHM of all detected objects within the same slitmask.

To ensure data quality, we applied a set of criteria to vet each slit's reduction. Slits failing any of the following conditions are classified as having poor reductions and are excluded from subsequent analysis steps:

- The slit is flagged as having failed one or more calibration or extraction steps.
- The wavelength solution has $RMS > 0.4$ pixels.
- The identified arc lines cover less than 60% of the spectral coverage.
- The computed spectral flexure correction exceeds 10 pixels, except in specific cases where visual inspection indicated that large flexure shifts were justified.
- The distribution of χ residuals deviates significantly from a Gaussian profile, specifically exhibiting a standard deviation less than 0.6 or greater than 1.6.

Two-Dimensional and One-Dimensional Coaddition

As a subsequent processing step, spectra from multiple exposures (when available) are grouped and coadded to produce both 2D and 1D combined spectra. For 2D coaddition, a configuration file (the *.coadd2d* file) is first generated to define the coaddition parameters. The PypeIt's `pypeit_coadd_2dspec` script is then executed, performing an optimally weighted coaddition in which the squared S/N ratios of the extracted objects are used as weights.

The coaddition is performed on a slitmask-by-slitmask basis, excluding slits previously identified as having poor reductions. These excluded slits are explicitly listed under the `exclude_slits` parameter in the *.coadd2d* file. The primary advantage of 2D coaddition is that visual inspection of the combined 2D spectra facilitates the identification of faint emission features from high-redshift galaxy candidates and enables the coaddition of serendipitous detections, which is not possible during purely 1D coaddition (described below). The `pypeit_coadd_2dspec` script, after performing the 2d coaddition, extracts also one-dimensional spectra from the coadded frames using the same extraction procedures applied to the single-frame reductions. These 1D spectra are not flux calibrated.

The purely 1D coaddition of the targeted object is performed by PypeIt's `pypeit_collate_1d` script. This tool groups spectra by object based on sky coordinates and coadds all matching spectra using a weighted average of the optimally extracted spectra, or boxcar extractions when optimal extractions are unavailable. Spectra associated with slits deemed to have poor reductions are excluded from the coaddition, with details recorded in the *collate_ warnings.txt* file.

Prior to coaddition, `pypeit_collate_1d` applies a relative flux calibration to each individual 1D spectrum using archival sensitivity functions created for each DEIMOS grating (see details here). These relatively flux-calibrated spectra are expressed in units of $[10^{-17}\ erg/s/cm^2/\mathring{A}]$. While their overall normalization may differ from an absolute calibration, they have the correct spectral shape across the wavelength range.

Notes

A detailed description of the data products provided in these contributed datasets can be found in the accompanying README file. Note that the reduction, coaddition, and fluxing processes were performed in an automated way - supplemented by a semi-automated vetting procedure - therefore, the users are strongly cautioned that these data should not be used for scientific analysis without additional, independent vetting.

For questions regarding these contributed datasets, please contact Bradford Holden <holden@ucolick.org>, Debora Pelliccia <dpelliccia@ucolick.org>, or Dusty Reichwein <dusty@ucolick.org>.