Mon. Not. R. Astron. Soc. **000**, 1–2 (2015) Printed 9 September 2015 (MN LATEX style file v2.2)

A new Cosmic Origin Spectrograph Reduction Pipeline: PyCOS

Cameron J. Liang⋆ and Hsiao-Wen Chen

Department of Astronomy & Astrophysics, and Kavli Institute for Cosmological Physics, University of Chicago, Chicago IL 60637

9 September 2015

ABSTRACT

We present a new and freely available suite of codes PyCOS fully developed in Python. PyCOS combines extracted 1D spectra and correct wavelength-dependent offset between exposurs in the Cosmic Origin Spectrograph onboard the *Hubble Space Telescope* (*HST*). PyCOS is designed to have no dependencies other than the standard Numpy and Scipy packages.

Key words: cosmology:theory – galaxies:halos – simulations:feedback

1 INTRODUCTION

The Cosmic Origin Spectrograph (COS; Green et al. 2012) has revolutionized many areas in our understandings of the Universe, ranging from low redshift IGM, the UV ionization background [references] and most recently the circumgalactic medium (Tumlinson et al. 2011; Werk et al. 2014; Liang & Chen 2014; Bordoloi et al. 2014; Johnson et al. 2015).

However, many of our understanding through the UV spectra observed by COS relies on the sensitive information on the absorption line. Therefore, a correct co-addition of exposures of the same object is critical.

Traditionally, the simplest method is to apply a constant shift to the entire 1D spectrum to align with a reference exposure. The

2 REDUCTION WORKFLOW

The pipeline is written in mind to reduce spectra of UV QSOs.

2.1 Extraction of 1D spectra

The first step in the pipeline is to extract 1D spectra for coaddition in the later steps. The correct extraction of the spectra are crucial. Note that most objects (e.g., QSOs) by COS have medium to low signal-to-noise (SNR). In this cases, the uncertainty of the observed flux can be Poisson-limited. Assumption of gaussian uncertainty will generally under-estimate the true uncertainty. To correctly accout for the low counts of UV photons, we compute asymmetric uncertainty as laid out in [References]. We compute the asymmetric uncertainty ($n + dn_+$ and $n - dn_-$) encompassing 68.27% of the area in a Poisson distribution given n_i photons in a wavelength bin λ_i.

The input fits files xxxxx_x1dsum.fits contains wavelength, flux, error, net count rate, exposure time and gcount rate (see COS manual Handbook). Note that the flux has been calibrated through STScI pipeline. These files are intermediate products co-added from the STScI pipeline, combining exposures from a single setting of the instrument (e.g., same central wavelength, aperture combining all FPPOS (grating offset position); see COS manual Handbook). If the net count is less than 135, the asymmetric error calculation is applied, otherwise a gaussian error is assumed. The number 135 is chosen to be large enough so the gaussian approximation is good to 8% fractional error relative to the Poisson. The number of photon used in the calculation is the gross count rate multipled by the exposure time ($n_g = dn_g/dt \times \tau_{exp}$). In addition to the asymmetric error, the standard error from STScI reduction pipeline is recorded. The output intermediate files (xxxx.x1dsum) include follow five arrays: wavelength, flux, error (gassuain), net_count, df_plus, df_minus (which comes from the Poisson distribution).

2.2 Alignment of 1D spectra

One can divide the issues into two problems: (1) a relative wavelength offset between exposures and (2) an absolute elongation and/or compression of the spectra.

With the first issue, the intermediate extracted spectra is the wavelength dependent offset between different exposures (possibly caused by thermal "breathing" of the detecters), hence the motivation of PyCOS . In general, there exists mis-alignment, elongation, and compression of spectra. A coaddition that neglects such effects will introduce large error on the absorption analysis, such as the shape of the absorption line profiles (i.e., b paramber), column density and number of absorbing components. These problematic absorption lines will propagate to a very different interpretation of the physical properties of the system in question.

To account for relative wavelength dependent offset $\Delta\lambda(\lambda)$ between different exposures, one can use absorption lines as based-

⋆ E-mail:jwliang@oddjob.uchicago.edu

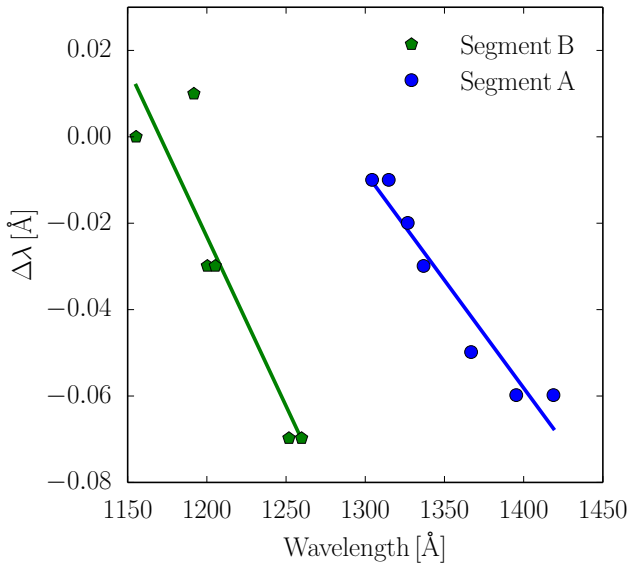

Figure 1. An example of showing the wavelength dependent fits of PKS0405. Blue circles and green pentagon represent alignment points relative to a reference exposure, obtained from `PyCOS` for segment A and B, respectively. The lines are best fit applied to the spectrum for the correction.

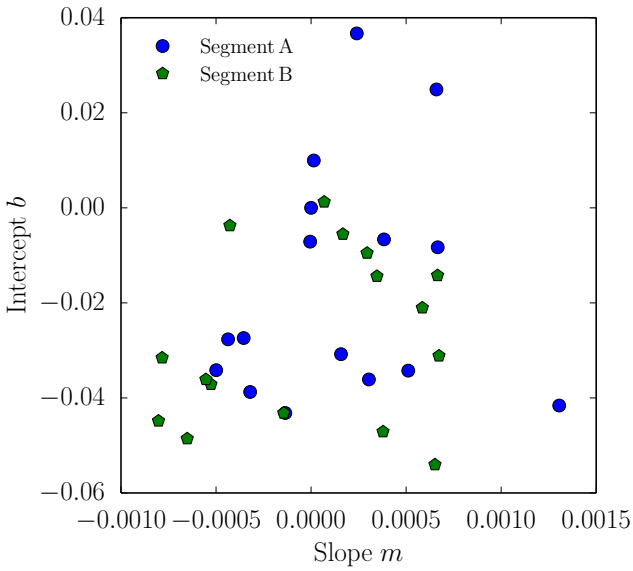

Figure 2. The distribution of the slopes and intercepts from the best-fit lines of $\Delta\lambda(\lambda)$. It shows the diversity of the lines needed to align and correct the wavelength dependent offsets. Explicitly, the parameters are from a pivoted fit: $\Delta\lambda = m(\lambda - \bar{\lambda}) + b$, where $\bar{\lambda}$ is the median wavelength of the range to be fitted.

lines. `PyCOS` is flexible and designed for the users to choose what absorption lines to use for alignment between exposures. The goodness of alignment is parametrized by χ^2 both quantitatively and visually to ensure robustness (as shown in Fig x). For each alignment of an absorption between the reference spectrum, and the align spectrum, the χ^2 is computed as:

$$\chi^2 = \sum_i \frac{(f_{\rm ref,i} - f_{\rm alg,i})^2}{df^2_{\rm ref,i} + df^2_{\rm alg,i}} \quad (1)$$

Note that one can in principle use any absorption lines present in the spectra (e.g., Milk-Way and Non- Milky-Way lines) since the coaddition is of the same object. It is recommended to use non-saturated lines since they will give the best estimates for χ^2. This process preserves the relative velocity structure of the absorptions lines in the spectra. The Milky-Way lines are not explicitly forced to have $z = 0$, which can be beneficial for study of the Milky-Way absorptions.

There are two segments in an exposure by each grating (a and b). For each segment, a polynomial is fit to the offset data points created from the alignment process, $\Delta\lambda = f(\lambda)$. From Fig 1, one can see that $f(\lambda)$ can be quite noisy. We recommend a conservative linear fit but higher orders are available. We warn users that extrapolation from an high order polynomial can introduce large error near the end of the spectra. The data points from the alignment process $\{\Delta\lambda_i, \lambda_i\}$ are recorded for each exposure, and name of such files is in the format fits_xxxxx.x1dsum if one desires to re-fit later. Fig 2 shows the collection of best-fit parameters from a linear fit to all of the alignment of exposures against a reference. The best fit line is $\Delta\lambda = m(\lambda - \bar{\lambda}) + b$.

2.3 Absolute Calibration - final rectification

In general, there exists elongation or compression in the spectra, even after wavelength dependent offsets between exposures. A final rectification is needed for an absolute calibration with the assumption that Milky-Way lines have a systemic velocity $v = 0\,\mathrm{km\ s^{-1}}$. This rectification is applied independently to each grating, since the distortions can be different.

2.4 Co-addition

To combine different exposures of the same astrophysical object, each exposure is weighted by the median SNR. In other words, each exposure carries only weight, where the weight is the squared of the median signal-to-noise ratio, $\langle S/N \rangle^2$. Users can choose other options for co-addition (pixel by pixel weighting scheme or no weighting at all).

ACKNOWLEDGMENTS

We thank ...
CL is supported by NESSF ..

REFERENCES

Bordoloi, R., Tumlinson, J., Werk, J. K., et al. 2014, ApJ, 796, 136
Green, J. C., Froning, C. S., Osterman, S., et al. 2012, ApJ, 744, 60
Johnson, S. D., Chen, H.-W., & Mulchaey, J. S. 2015, MNRAS, 449, 3263
Liang, C. J., & Chen, H.-W. 2014, MNRAS, 445, 2061
Tumlinson, J., Thom, C., Werk, J. K., et al. 2011, Science, 334, 948
Werk, J. K., Prochaska, J. X., Tumlinson, J., et al. 2014, ApJ, 792, 8

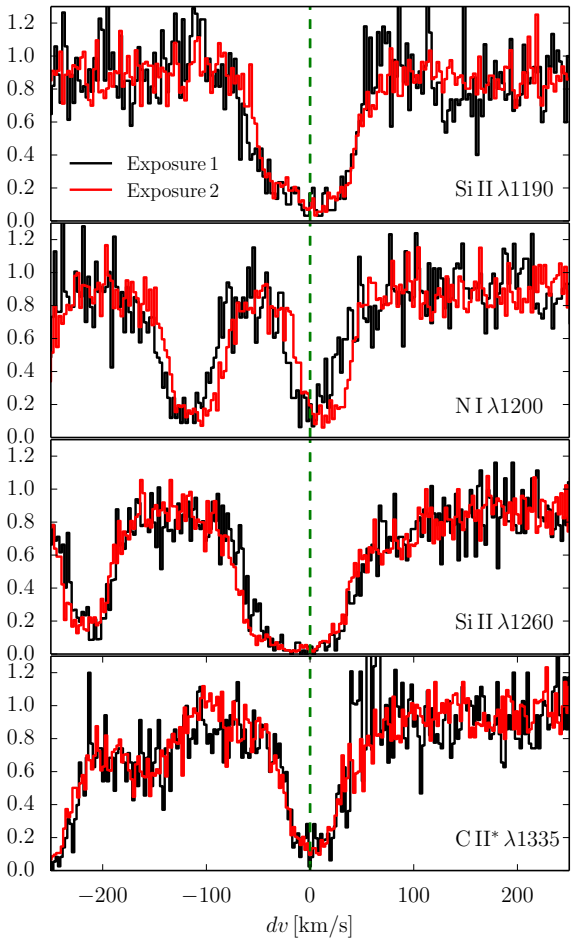

Figure 3. Comparison of two different combined x1dsum 1D spectra (black and red) of the PKS0405-123 before relative calibration. The comparison shows not all absorption lines are simultaneously aligned, implying a wavelength dependent offset.