Research in Astronomy and Astrophysics, 25:025018 (15pp), 2025 February
https://doi.org/10.1088/1674-4527/ada349
CSTR: 32081.14.RAA.ada349

BYSpec: An Automatic Data Reduction Package for BFOSC and YFOSC Spectroscopic Data

Zi-Chong Zhang[1,2,3], Jun-Bo Zhang[4], Ju-Jia Zhang[5,6], De-Yang Song[1,2], Jing Chen[1,2], Ming-Yi Ding[1,2], Nan Zhou[1,2,3], Liang Wang[1,2,3], and Kai Zhang[1,2,3]

[1] Nanjing Institute of Astronomical and Optics & Technology, Chinese Academy of Sciences, Nanjing 210042, China; liangwang@niaot.ac.cn
[2] CAS Key Laboratory of Astronomical Optics & Technology, Nanjing Institute of Astronomical Optics & Technology, Nanjing 210042, China
[3] University of Chinese Academy of Sciences, Beijing 100049, China
[4] CAS Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China
[5] Yunnan Observatories, Chinese Academy of Sciences, Kunming 650216, China
[6] International Centre of Supernovae, Yunnan Key Laboratory, Kunming 650216, China
Received 2024 November 22; revised 2024 December 14; accepted 2024 December 24; published 2025 February 12

Abstract

BFOSC and YFOSC are the most frequently used instruments in the Xinglong 2.16 m telescope and Lijiang 2.4 m telescope, respectively. We developed a software package named "BYSpec" (BFOSC and YFOSC Spectra Reduction Package) dedicated to automatically reducing the long-slit and echelle spectra obtained by these two instruments. The package supports bias and flat-fielding correction, order location, background subtraction, automatic wavelength calibration, and absolute flux calibration. The optimal extraction method maximizes the signal-to-noise ratio and removes most of the cosmic rays imprinted in the spectra. A comparison with the 1D spectra reduced with IRAF verifies the reliability of the results. This open-source software is publicly available to the community.

Key words: methods: data analysis – instrumentation: spectrographs – techniques: image processing

1. Introduction

Slit spectrographs are common in many large telescopes. They usually have a simple structure, high throughput, and are easy to use. Low-resolution slit spectrographs offer a cost-effective and efficient method for capturing the spectra of faint celestial objects. Most of them use grisms as dispersion elements, and the spectral resolutions are typically below 1000. Long slit spectrographs can simultaneously feed the light from a scientific interested object and a nearby object as the reference to correct for the atmospheric transmission variations as well as a better sky emission correction. On the other hand, the spectra of different positions of an extended source (such as galaxies and nebulae) along the slit can be obtained at the same time with long-slit spectrographs. Echelle grism, together with cross dispersers, can provide a spectral resolution of >1000, yet the throughput is generally lower. Both long slit and echelle spectrographs play essential roles in various research fields in modern astrophysics.

The light collected by a spectrograph is usually recorded by two-dimensional detectors such as CCDs. Integrating the photons of the targets along the cross-dispersion direction yields one-dimensional spectra. However, effects such as detector bias, dark current, and quantum efficiency of the individual pixels affect the fluxes. Photons from the sky background, scattering light, and cosmic rays also contaminate science objects. Additionally, the wavelength of each pixel on the one-dimensional spectra needs to be calibrated, which can be performed by comparing the spectra with that of an arc lamp full of emission lines whose wavelengths are already known. The spectra reduction procedure usually refers to the above processes from the raw images to the wavelength and sometimes flux-calibrated one-dimensional spectra (λ versus F_λ), which are ready for scientific analysis. Due to its complexity, it takes significant effort for an even well-trained observer.

The Beijing-Faint Object Spectrograph and Camera (BFOSC) mounted on the Cassegrain focus of the 2.16 m telescope (Fan et al. 2016) in Xinglong Observatory is a multi-purpose instrument with imaging and spectroscopy modes. Its optical and mechanical designs are large inherited from the EFOSC (ESO Faint Object Spectrograph and Camera; Buzzoni et al. 1984) on the 3.6 m NTT telescope. In the past decade, it has been the most frequently used instrument in this telescope. According to the time allocation table, 231 or 87.5% of the observing nights are allocated to BFOSC users in the semester from 2024 September to 2025 June, and a large fraction of these nights are used for spectroscopy observations, especially with "long-slit" and "echelle" modes. In the long-slit mode, users can select one of the nine switchable grisms (designated "G3"—"G12") as the disperser to get the reciprocal linear dispersion of 80–800 Å mm^{-1}, corresponding to the spectral resolution of 1.2–12 Å pixel^{-1}, or $\lambda/\Delta\lambda \simeq$ 240–800 for the slit width of 2.″3. Only the first diffraction order ($m = 1$) is used in this mode. The slit has a length of 9.′4 on the sky, which facilitates the dispersion of the sky background over a large area, enables simultaneous observation of a reference star for differential spectrophotometry, and allows the acquisition of spectra from multiple positions of extended sources. While in echelle mode, an

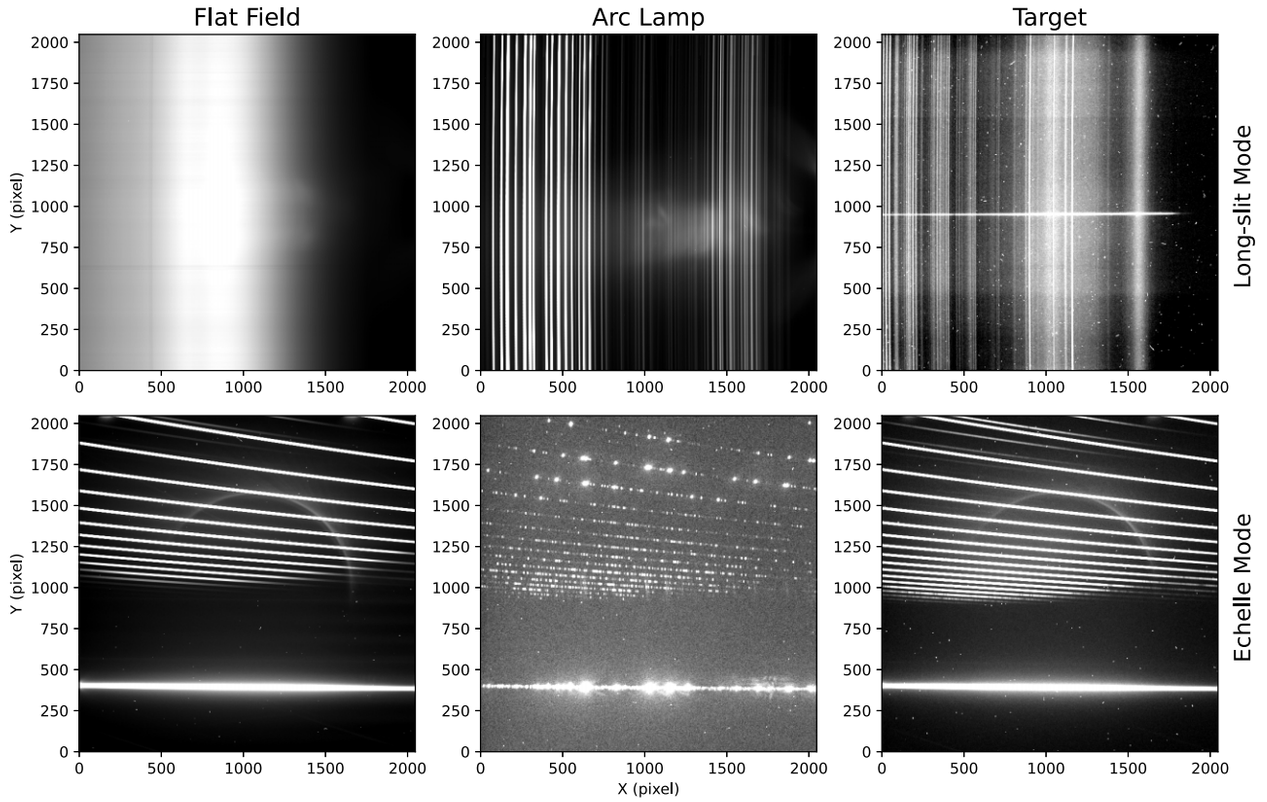

Figure 1. An overview of raw images for different objects (flat-fielding, arc lamp, and target) taken by BFOSC with long-slit (upper row) and echelle modes (bottom row). For both modes, the *x*-axis is the dispersion direction and the *y*-axis is the slit direction.

echelle grating ("E9") is inserted into the optical path as the main disperser to provide a higher spectral resolution of 0.4–1.0 Å pixel^{-1}. In this mode, more than 10 diffraction orders can be used and they are arranged along the slit direction by using one of the three grisms (G10—G12) as the cross-disperser. A short slit (3.″0) is used to prevent spectral overlapping and only a single source can be observed in each exposure. The wavelength coverage and dispersion for different modes of BFOSC can be found in Table 1 of Fan et al. (2016). Similarly, the Yunnan Faint Object Spectrograph and Camera (YFOSC) is the primary instrument on the 2.4 m telescope (Wang et al. 2019) in Lijiang, Yunnan Observatory. It shares almost the same structure as BFOSC, thus their spectra formats resemble each other and can be reduced in the same way. Many efforts have been made to construct a standard manner to analyze the massive spectra data collected by these two instruments, including tutorials written by staff astronomers in Xinglong Observatory using IRAF (Image Reduction and Analysis Facility), and PyFOSC (Fu 2020), a Python package for long-slit spectra reduction but also relies on IRAF.

We developed a Python software package named "BYSpec" (BFOSC and YFOSC Spectra Reduction Package)[7] to perform the standard reduction procedures from raw images to wavelength-calibrated, "science-ready" one-dimensional spectra. In this paper, we introduce the algorithms adopted by the software and its performance.

2. Data Overview

Before presenting the software package, we give a brief overview of the spectral images taken with various modes and their differences. Figure 1 shows the raw images taken by BFOSC, where the panels in the upper row show the flat-fielding, arc lamp, and scientific target respectively taken in the long-slit mode, and the lower row shows the same objects but in the echelle mode. Throughout this paper, the *X*-axis is designated as the dispersion direction, while the *Y*-axis represents the slit direction or cross-order direction.

In the long-slit mode, the flat-fielding is generated by a tungsten halogen lamp through the grism, therefore it is a dispersed "spectroscopic" flat rather than the widely-used "white-light" flats in photometry observations. The counts in the short-wavelength region, corresponding to the right part of the CCD detector, are significantly lower than those in the long-wavelength region. This discrepancy is attributed to the low flux of the tungsten-halogen

[7] The software is publicly available at https://github.com/wangleon/byspec.

lamp in the short-wavelength range. The arc lamp (in this case, FeAr) uniformly illuminates the entire slit, producing a series of emission lines across the image. For a point source target, its spectrum appears as a horizontal stripe at the center of the image, with all vertical lines representing sky emission lines.

In the echelle mode, the diffraction orders produced by the echelle grating are aligned along the slit direction, beginning at $y \simeq 900$. The bright stripe observed at $y \simeq 400$ corresponds to the 0th order of the cross-disperser, where all diffraction orders generated by the main disperser overlap, rendering it unusable. The flat-fielding image has two primary purposes: determining the positions of the diffraction orders on the CCD images and correcting for pixel-to-pixel response variations within the spectral region. However, pixels between the echelle orders are challenging to correct using only the spectroscopic flat in the same mode, as they receive insufficient illumination. In both modes, the photometric white flat is not always available. As a result, we implemented the flat-field correction in BYSpec relying only on the spectroscopic flats within the corresponding mode.

3. Reduction Procedures

BYSpec package supports the long-slit and echelle spectra taken by BFOSC and YFOSC. The completed workflow is shown in Figure 2.

Our software reduces the raw images night-by-night, regardless of multiple modes (long-slit, echelle) or configurations (gratings, echelle grisms) used in the same night. For an image taken under a specific mode and configuration, the program automatically looks for the corresponding calibration files (bias, flat-fielding, and arc lamp) taken on the same night and executes the reduction processes.

3.1. Generating Observation Logs

The first step is to scan all the raw image files in an observation night and generate a machine-readable observation log file in ASCII format. For BFOSC and YFOSC, the observing information (image type, target name, exposure time, starting time, etc.) is stored differently. Most of them are in the filenames for BFOSC, while these are saved as keywords in the FITS headers. BYSpec scans the raw file directory, extracts the filenames and FITS keywords, and creates a standardized log file. An example of an observation log for BFOSC is shown in Figure 3. Then BYSpec classifies all the images into five categories, i.e., bias, flat, arc lamp, flux standard stars, and scientific targets (see Figure 2) according to the records in the observation log file.

3.2. Bias Subtraction and Overscan Correction

The bias refers to an offset value added to each pixel's counts so that the readout number is always positive. Multiple bias frames with the same pixel binning, readout speeds, gain, and zero exposure time are often taken and averaged to suppress the readout noise and measure this offset. The averaged "master" bias frames for BFOSC is shown in Figure 4, where spatial patterns exist in both frames. Since no obvious pattern along the vertical direction is seen, BYSpec calculates the averaged values along the vertical direction and takes the smoothed profiles as the bias values subtracted from all the other frames. As bias is mode and configuration-independent, long-slit and echelle modes can share the same bias frames. For CCDs with overscan regions, such as YFOSC, overscan correction will be introduced in Section 4.1.

3.3. Flat-fielding Correction in Long-slit Mode

The CCD pixels have slightly different responses to photons. Dust, interference fringes, and other optical issues can cause non-uniform illumination across the field of view. Like many other spectrographs, BFOSC and YFOSC use halogen tungsten lamps with continuous spectra to get dispersed flat-fielding frames. However, due to the different optical paths, flat-fielding frames under long-slit and echelle modes are very different. The tungsten lamps in long-slit mode can illuminate almost the whole slit, and therefore the entire CCD is illuminated, as shown in the upper left panel of Figure 5. The fluxes in the short-wavelength region are weak because of the low temperature of the tungsten lamp, and the vertical dark stripes are the telluric absorption bands. Some horizontal dark stripes appear in every image and are probably caused by the dust in the slit.

For the long-slit mode, BYSpec performs a customized third-order 2D Savitzky–Golay filter (Savitzky & Golay 1964) over the flat-fielding image to get a smoothed image. By setting the length of the filter window to 150 and the order of polynomials to 3, it effectively removes the response variations of adjacent pixels while leaving the overall energy distributions of the tungsten lamp. Then, the pixel-to-pixel variation response, represented as the sensitivity map (shown in the upper right panel of Figure 5) is derived by dividing the "master" flat by the smoothed image. Following bias correction, all images with scientific targets and arc lamps are divided by this sensitivity map.

3.4. Order Location and Flat-fielding Correction in the Echelle Mode

The light from the tungsten lamp in the echelle mode is dispersed by the cross-disperser and only illuminates the pixels in the echelle orders (see the lower left panel of Figure 5). As the echelle mode uses a short slit, the widths of orders are only a few pixels. Owing to this difference, BYSpec treats flat-fielding frames of long-slit and echelle modes differently.

In the echelle mode, all images are trimmed to exclude the regions containing the 0th order generated by the cross-disperser. This can avoid erroneous identification as a normal diffraction order. Subsequently, BYSpec coadds the spectroscopic flats as the "master" flat-fielding, which is used to locate the order positions.

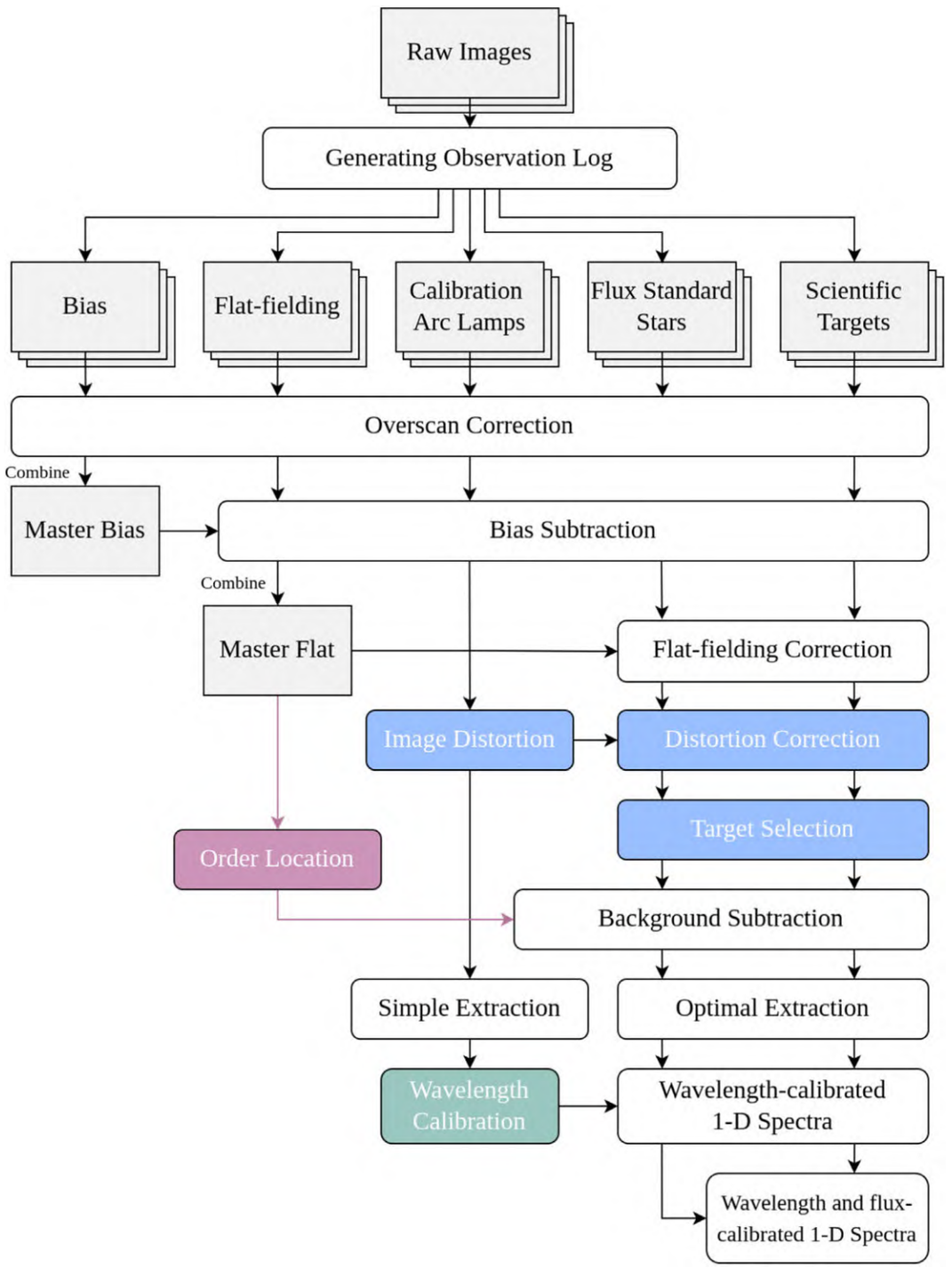

Figure 2. Data reduction workflow of BYSpec. Gray boxes represent images; white boxes are common procedures shared by both long-slit and echelle modes. Purple and blue boxes are steps dedicated to echelle and long-slit modes, respectively, and green boxes represent the steps included in both modes but have different methods.

frameid	fileid	datatype	object	exptime	dateobs	mode	config	slit	filter	binning	gain	rdnoise	q99	observer
[1]	202304180001	BIAS	BIAS	0.0	2023-04-18T11:26:40					1x1	2.2	7.8	308	XingLong
[2]	202304180002	BIAS	BIAS	0.0	2023-04-18T11:26:48					1x1	2.2	7.8	309	XingLong
[3]	202304180003	BIAS	BIAS	0.0	2023-04-18T11:26:56					1x1	2.2	7.8	308	XingLong
[4]	202304180004	BIAS	BIAS	0.0	2023-04-18T11:27:04					1x1	2.2	7.8	309	XingLong
[5]	202304180005	BIAS	BIAS	0.0	2023-04-18T11:27:13					1x1	2.2	7.8	309	XingLong
[6]	202304180006	BIAS	BIAS	0.0	2023-04-18T11:27:21					1x1	2.2	7.8	308	XingLong
[7]	202304180007	BIAS	BIAS	0.0	2023-04-18T11:27:29					1x1	2.2	7.8	310	XingLong
[8]	202304180008	BIAS	BIAS	0.0	2023-04-18T11:27:38					1x1	2.2	7.8	309	XingLong
[9]	202304180009	BIAS	BIAS	0.0	2023-04-18T11:27:46					1x1	2.2	7.8	309	XingLong
[10]	202304180010	BIAS	BIAS	0.0	2023-04-18T11:27:54					1x1	2.2	7.8	308	XingLong
[45]	202304180045	SLITTARGET	SN2023eoc	30.0	2023-04-18T19:14:37			1.8"		1x1	2.2	7.8	419	ZhangJunbo
[46]	202304180046	SPECLTARGET	SN2023eoc	3300.0	2023-04-18T19:15:57	longslit	G4	1.8"	385LP	1x1	2.2	7.8	594	ZhangJunbo
[47]	202304180047	SLITTARGET	BD+332642	25.0	2023-04-18T20:15:48			1.8"		1x1	2.2	7.8	491	ZhangJunbo
[48]	202304180048	SPECLTARGET	BD+332642	600.0	2023-04-18T20:16:54	longslit	G4	1.8"	385LP	1x1	2.2	7.8	440	ZhangJunbo
[55]	202304180055	SPECLFLAT	FLAT	70.0	2023-04-18T21:07:16	longslit	G4	1.8"	385LP	1x1	2.2	7.8	23072	XingLong
[56]	202304180056	SPECLFLAT	FLAT	70.0	2023-04-18T21:08:34	longslit	G4	1.8"	385LP	1x1	2.2	7.8	23251	XingLong
[57]	202304180057	SPECLFLAT	FLAT	70.0	2023-04-18T21:09:51	longslit	G4	1.8"	385LP	1x1	2.2	7.8	23067	XingLong
[58]	202304180058	SPECLFLAT	FLAT	70.0	2023-04-18T21:11:08	longslit	G4	1.8"	385LP	1x1	2.2	7.8	23003	XingLong
[59]	202304180059	SPECLFLAT	FLAT	70.0	2023-04-18T21:12:25	longslit	G4	1.8"	385LP	1x1	2.2	7.8	23003	XingLong
[60]	202304180060	SPECLLAMP	FeAr	30.0	2023-04-18T21:17:56	longslit	G4	1.8"	385LP	1x1	2.2	7.8	3850	XingLong
[61]	202304180061	SPECLLAMP	FeAr	300.0	2023-04-18T21:18:43	longslit	G4	1.8"	385LP	1x1	2.2	7.8	41645	XingLong

Figure 3. An example of generated observation log of BFOSC in 2023 April 18. The file is stored in `Astropy` ASCII table format, which is both machine and human-readable.

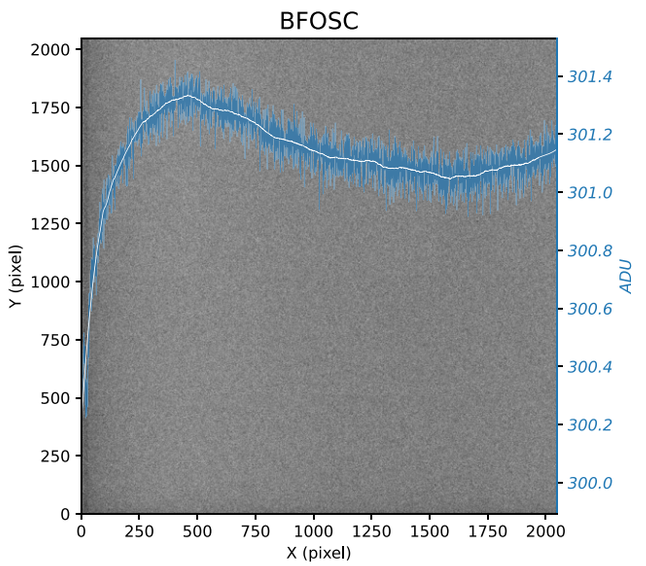

Figure 4. Co-added master bias frame of BFOSC. Overplotted cyan lines are averaged values along the *Y*-axis and the white curves are smoothed values. The white lines are used to subtract from other images.

Then, the vertical profiles from different columns are shifted with second-order polynomials relative to the profile of the central column and stacked together to obtain a "coadded" profile, as shown in the left panel of Figure 6. Subsequently, a peak detection algorithm is performed on the coadded profile to find their locations (y_1, y_2, ...y_n), representing the y coordinates of intersection points of the orders with the central column. Then these coordinates are converted to the y positions of each column by reversing the second-order polynomials. Finally, the coordinates (x, y) of each order are fitted by third-order polynomials, and the coefficients are saved. All the profile manipulations are performed on a logarithmic scale, mitigating the significant brightness differences between the orders located in the upper (long-wavelength) and lower (short-wavelength) parts of the image. This approach ensures that the detection process is almost equally sensitive to both short- and long-wavelength orders, with no orders omitted, thereby maximizing the wavelength coverage.

Interestingly, our methods can locate some faint "ghost" and "secondary" orders between the normal diffraction orders, which are commonly seen in many other echelle spectrographs (e.g., Cretignier et al. 2021). Ghost orders are caused by spurious reflections between the optical elements and the detector surfaces. They are usually 1 to 2 order-of-magnitude weaker than the normal orders, and their positions have different spatial distributions. On the other hand, cross-disperser separates the different echelle orders using the first diffraction order ($m_\times = 1$), while the short wavelength of the secondary order ($m_\times = 2$) contaminates the long wavelength orders. We call them "abnormal" orders, contrary to the "normal" ($m_\times = 1$) orders. If the abnormal orders intersect with the normal ones, the extracted 1D spectra will be overestimated by a small amount. Fortunately, none of the BFOSC and YFOSC echelle orders are intersected by abnormal orders, as shown in the right panel of Figure 6. However, the detected abnormal orders are still blocked when determining the background level using the pixels between the normal orders.

After finding the order locations in the echelle mode, BYSpec extracts the 1D spectra of the flat lamp by directly summing up ± 5 pixels above and below the order positions. Then the one-dimensional spectra of different orders are smoothed by fitting lower-order polynomials and extended to 2D images with the widths of only 10 pixels of each order. The sensitivity map is constructed by dividing the 2D flat-fielding into the extended, smoothed images. All pixels outside the 10 pixels are set to unity.

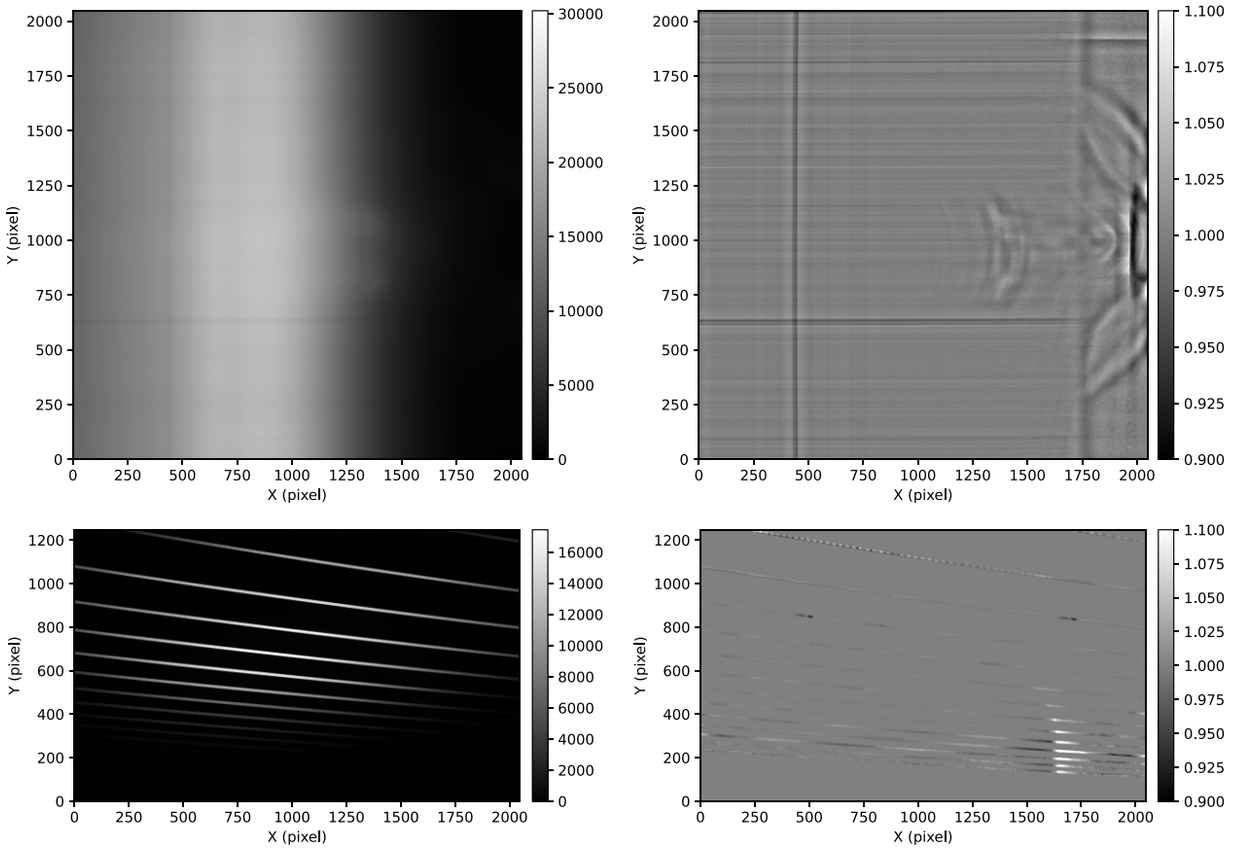

Figure 5. The coadded flat-fielding images of BFOSC in long-slit mode (upper left) and echelle mode (lower left). The right panels show the sensitivity maps after smoothing the flat-fielding of two modes.

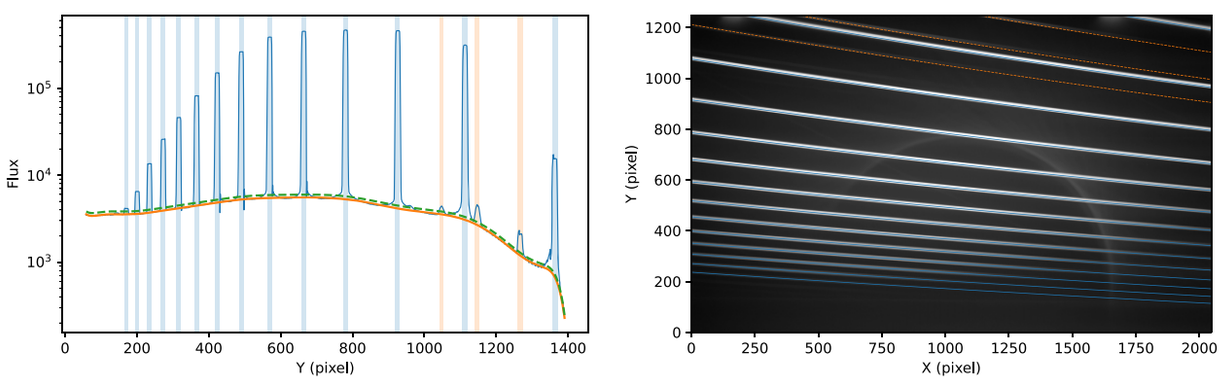

Figure 6. Order detection in the `E9+G10` echelle mode of BFOSC. The left panel shows the coadded cross-order profiles (blue solid curves) and the detected normal orders (filled blue stripes) as well as the abnormal orders (filled orange stripes). The orange solid line and the green dashed line represented the fitting background and $+3\sigma$ above the background, which is used as the threshold of the peak detection algorithm. The right panel shows the coadded flat-fielding image in the logarithm scale. Detected normal orders and abnormal orders are overplotted as solid blue and dashed orange lines, respectively.

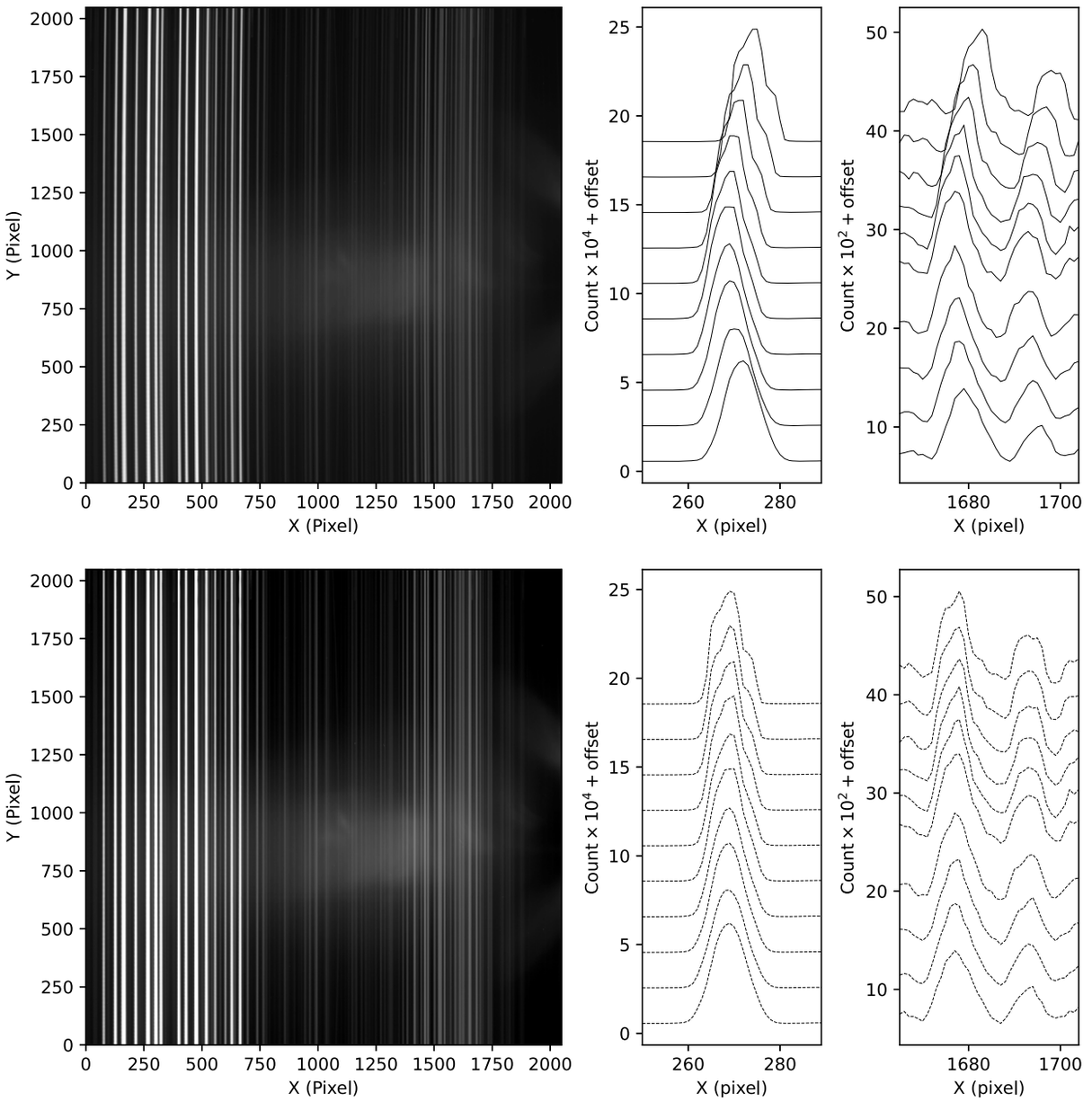

Figure 7. An image of FeAr arc lamp (upper left panel) of BFOSC shows "smile distortion," where the emission lines curve toward the right side, as shown in the upper right panels, where cross profiles of two selected regions along the *X*-axis (dispersion direction) every 200 rows are plotted with offsets. Obviously, the emission lines are not straight but curve toward one side along the dispersion direction. The lower left panel shows the image after the distortion correction, where the emission lines (dashed lines in the lower right panels) of different rows are aligned along the slit direction.

3.5. Distortion Correction in the Long-slit Mode

The spectral lines imaged by BFOSC in the long-slit mode appear to curve across the detectors. We show this effect by illuminating the slit with an FeAr arc lamp as shown in Figure 7. This is commonly seen in imaging spectrographs and is known as the spectral curvature or "smile distortion." This effect is caused by the optical aberrations in the optical system and curved focal plane. The curvature makes the relation between pixel and wavelength vary in the different positions along the slit, which further introduces systematic errors in the wavelength calibration if they are not properly taken into account.

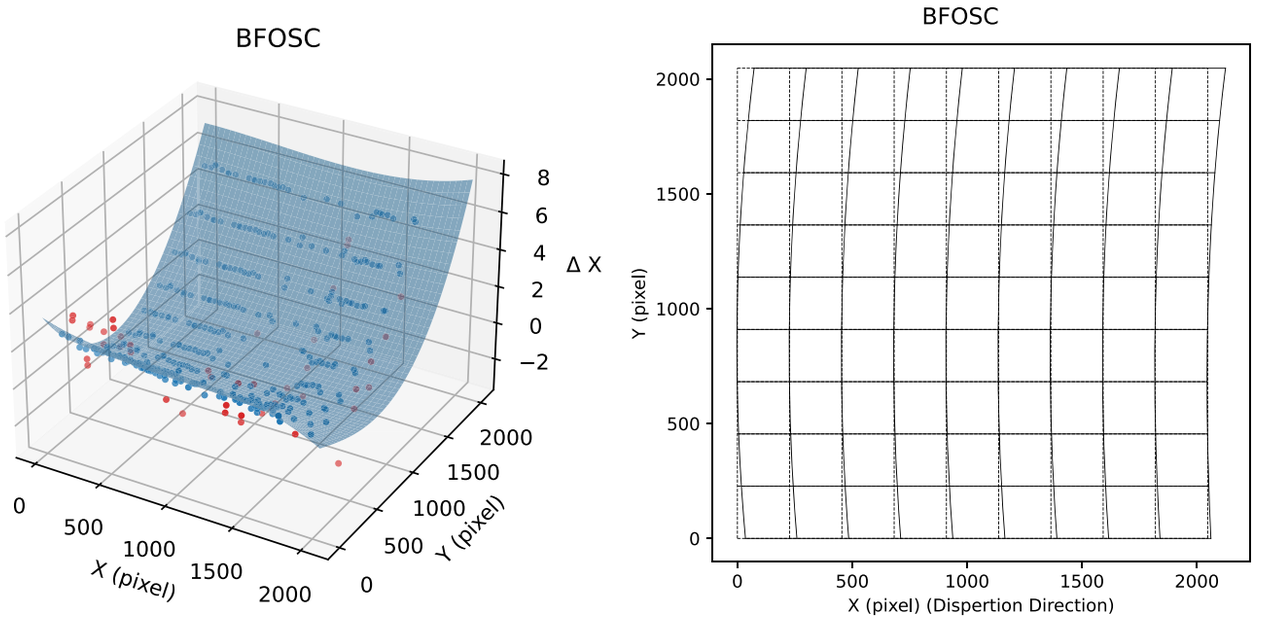

Figure 8. The left panel shows the bivariate polynomial fitting (blue surfaces) used in smile distortion correction for BFOSC. Blue solid points represent the relative shift along the dispersion direction of the peaks of emissions lines in the arc lamp, and red points are outliers above or below 3σ. The right panel shows the smile distortion as recovered from this fitting. The dashed lines are the ideal grids without any distortion, and the solid lines are the images of the grids affected by the smile distortion. The distortion (Δx) is exaggerated by 10 times for better visualization.

Table 1
Source of built-in Line List in BYSpec

Species	Number of Lines	Wavelength Range (nm)	Source
Argon	501	300–1050	Norlén (1973)
Helium	10	380–710	Martin (1960)
Neon	43	490–890	Saloman & Sansonetti (2004)
Iron	77	340–870	Crosswhite (1975)
Thorium	3762	300–10600	Palmer & Engleman (1983)

BYSpec uses the long-slit arc lamp spectra to correct this distortion. First, BYSpec extracts the spectra of the arc lamp (usually FeAr) within ± 10 pixels of the central row along the slit direction and fits the emission lines using the generalized Gaussian function $F(x) = A\exp(-(|x - c|/\alpha)^{\beta}) + b$, where x is the coordinates, and A, c, b, α and β are free parameters. Subsequently, BYSpec scans along the slit direction (y-axis for BFOSC data) with a step of 200 pixels and fits the emission lines in the same manner. For the same emission lines, the shift along the dispersion direction relative to the referenced spectra (Δx_y, where $y = \pm 200$, ± 400...) are fitted with a low-order bivariate polynomial ($\Delta x_y = f(x, y)$), as shown in the left panel of Figure 8. Since the shift along the dispersion direction is a function of both dispersion direction (x) and slit direction (y), nonlinear smile distortion at different positions across the image can be taken into account and corrected.

The right panel of Figure 8 shows a schematic view of the smile distortion observed in BFOSC. It is clearly shown that the distortion (Δx) is nonlinear in both dispersion and slit directions. The maximum deviation from center to edge is $\sim$7 pixels for BFOSC. Ignoring this effect could lead to significant systematic errors in wavelength calibration.

3.6. Wavelength Calibration

In long-slit mode, after the smile distortion, wavelength calibration is performed by fitting the central positions of emission lines versus their known wavelengths by polynomials. Both BFOSC and YFOSC have multiple choices for arc lamps, including FeAr and HeNe lamps. Therefore, we integrate wavelength tables of various elements in BYSpec. The sources have been listed in Table 1.

In the wavelength calibration, BYSpec uses the generalized Gaussian functions plus constant background levels to fit the profiles within a specific window (usually ± 7 pixels) centering at each line in the arc lamps. In the long-slit mode, the spectral resolution is only a few hundred, some lines cannot be fully resolved and exhibit significant deviation from symmetric shapes. For such cases, we use multiple generalized Gaussian functions together with a constant background to fit the profiles of multiple lines simultaneously. Figure 9 shows examples of single-line and

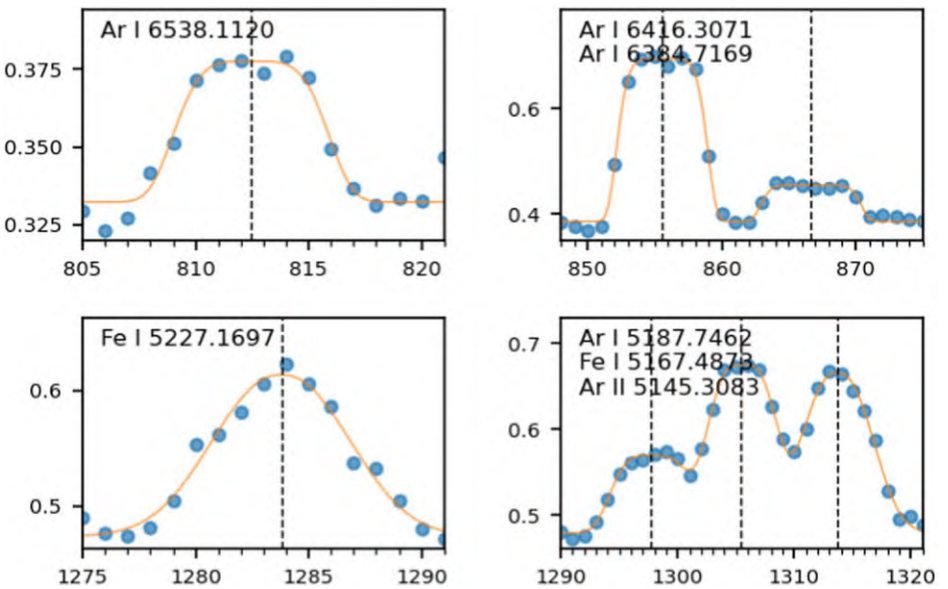

Figure 9. Examples of single-line and multiple-line fitting for FeAr arc lamp in G4 long-slit mode of BFOSC. Blue points denote the extracted fluxes and orange curves are fitted profiles using generalized Gaussian functions. Fitted line centers are plotted with dashed vertical lines. The fluxes are normalized to 1×10^3.

multiple-line fitting. For the long-slit mode, this method increases the number of useful lines by up to twice and hence significantly reduces the standard deviations on λ_i versus x_i.

In the echelle mode, as the spectral resolution is higher, most of the emission lines listed in the line lists are fully resolved, therefore only single line fitting is used, and the wavelength calibration is performed order-by-order, i.e., fitting the relations of x_i versus λ_i in the same order. We tried to use the global fitting as proposed by Baranne et al. (1996), but got a much larger standard deviation, probably due to the complicated optical distortion, even taking account of the smile distortion as we corrected in the long-slit mode. This reminds us that the image processing procedures of the two different modes should not be the same, as the light beam passes through different optical elements.

We noticed that wavelength identification is time-consuming work even for an expert. On the other hand, the wavelength coverage only shifts a small amount (usually less than 10 pixels) on the detectors over the long term. Therefore, we uploaded the identified wavelengths together with the arc lamp spectra of BFOSC and YFOSC in different modes to the Amazon AWS server as templates. When BYSpec is running, the spectra of the same arc lamp, using the same modes (echelle or long-slit), and the same configuration (including the grating used) will be downloaded automatically. The shift of emission lines in the CCD pixels between the user yielded and the template we uploaded is calculated using cross-correlation functions (CCFs), and the wavelengths of the new arc lamp spectra will be identified automatically. We have collected wavelength templates from various nights under different configurations and modes for both BFOSC and YFOSC. If multiple templates are found, BYSpec will use the one that most closest to the date of observation. Ultimately, the wavelength solutions in long-slit mode and echelle mode are shown in Figures 10 and 11 respectively.

3.7. Background Subtraction for Single Point Source

Background subtraction is essential in spectral reduction as the photons from the celestial targets are contaminated by the sky background and the stray light inside the instrument. As most of the targets observed by BFOSC and YFOSC are faint, the fluxes in some spectral regions are dominated by sky-light emissions (e.g., [O I] 557.7 nm). For point sources like stars observed in long-slit mode, we implemented an automatic method to determine the background, as shown in Figure 12. Two areas besides the observed targets are selected as the windows using an iterative polynomial fitting along the slit direction. The counts inside the windows are averaged as the background spectra and are subsequently subtracted from the spectra of the observed target.

In echelle mode, the background lights of targets are determined using the pixels between the diffraction orders, including the faint "abnormal orders," as mentioned in Section 3.4. Pixels within certain pixel ranges (±8 pixels for BFOSC) to the detected orders are blocked, and the fluxes of the rest pixels are fitted with iterative third-order polynomials column-by-column. Finally, the determined background lights are subtracted from every science image.

3.8. Target Selection and Optimal Extraction

The spectrum of a point source imaged on the CCD plane is not straight along the spatial (slit) direction. We scan the images along

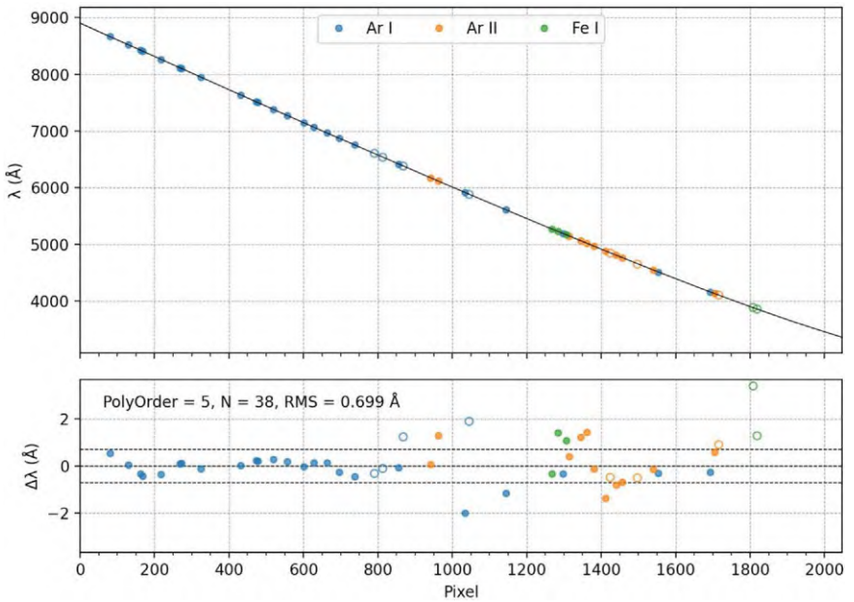

Figure 10. The black solid line in the upper panel shows the fitted wavelength solution (i.e., x vs. λ_x for each pixel x along the dispersion direction) of long-slit mode `G4` for BFOSC using a FeAr arc lamp. The colored points are the emission lines of various species used to fit the wavelength solution, and the open points represent the outliers outside $\pm 3\sigma$ or the line intensity is too low to get a reliable central position. The order of fitted polynomial is set to 5 and the residuals are plotted in the lower panel. There are a total of 38 lines used and the rms is $\sim$0.7 Å, corresponding to $\sim 1/4$ pixel, or $1/20$ of the FWHM of the resolving element.

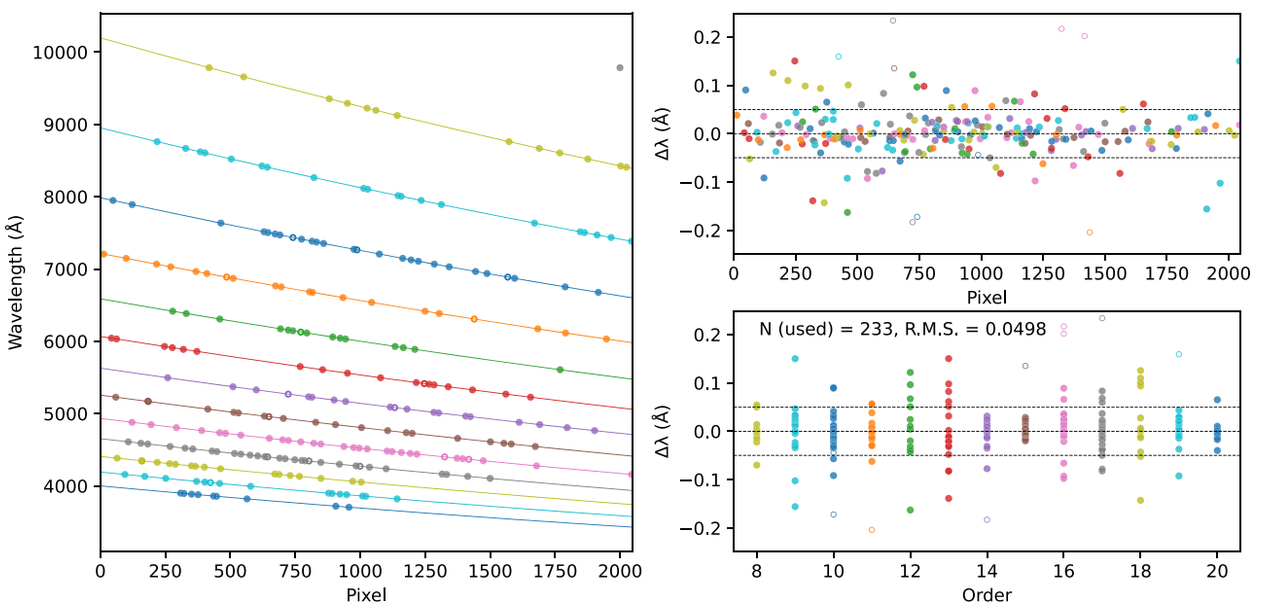

Figure 11. Wavelength solution (left) for echelle mode (`E9+G10`) for BFOSC using a FeAr arc lamp. The left panel shows the fitted wavelength solution (colored solid lines) of each echelle order, together with the emission lines used (colored points) in the fitting. The upper right panel shows the residuals along the dispersion direction, and the lower right panel shows the residuals in different echelle orders. Lines in the same echelle order are plotted in the same color. The polynomial fitting is performed for each echelle order independently, and the total rms is $\sim$0.05 Å, or $1/10$ pixel, or $1/50$ of the FWHM of the resolving element.

the dispersion direction and use Gaussian functions to fit the profiles along the slit direction to find the peaks. Then polynomials are used to describe the CCD positions of the target spectra. The result of target selection is shown in Figure 13.

Direct summing up the counts along the slit direction within several pixels above and below the polynomial yield the flux at this wavelength, i.e., a point on the one-dimensional spectra. However, this method does not consider the different noises of the

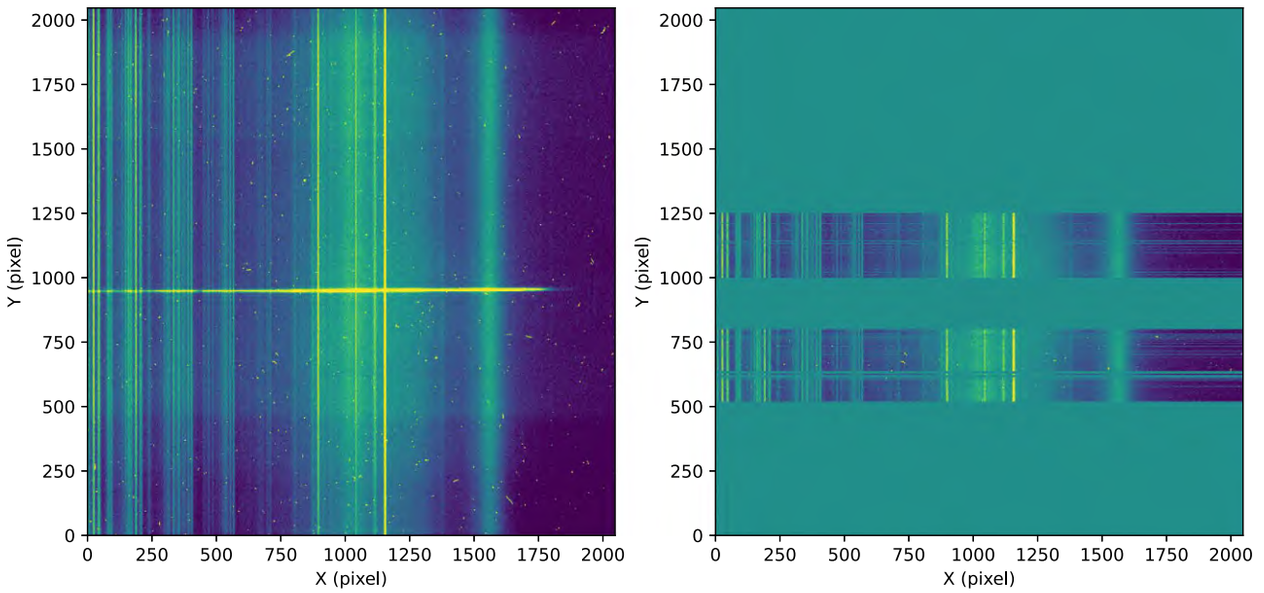

Figure 12. The bias-subtracted, flat-fielding, and distortion-corrected image (left) taken with long-slit mode using BFOSC. The bright horizontal stripe is the dispersed spectra of a point-source target, and the vertical stripes are the sky emission lines. The selected regions used for background subtraction are plotted in the right panel with the same color scale. The pixels outside the background regions are set to zero.

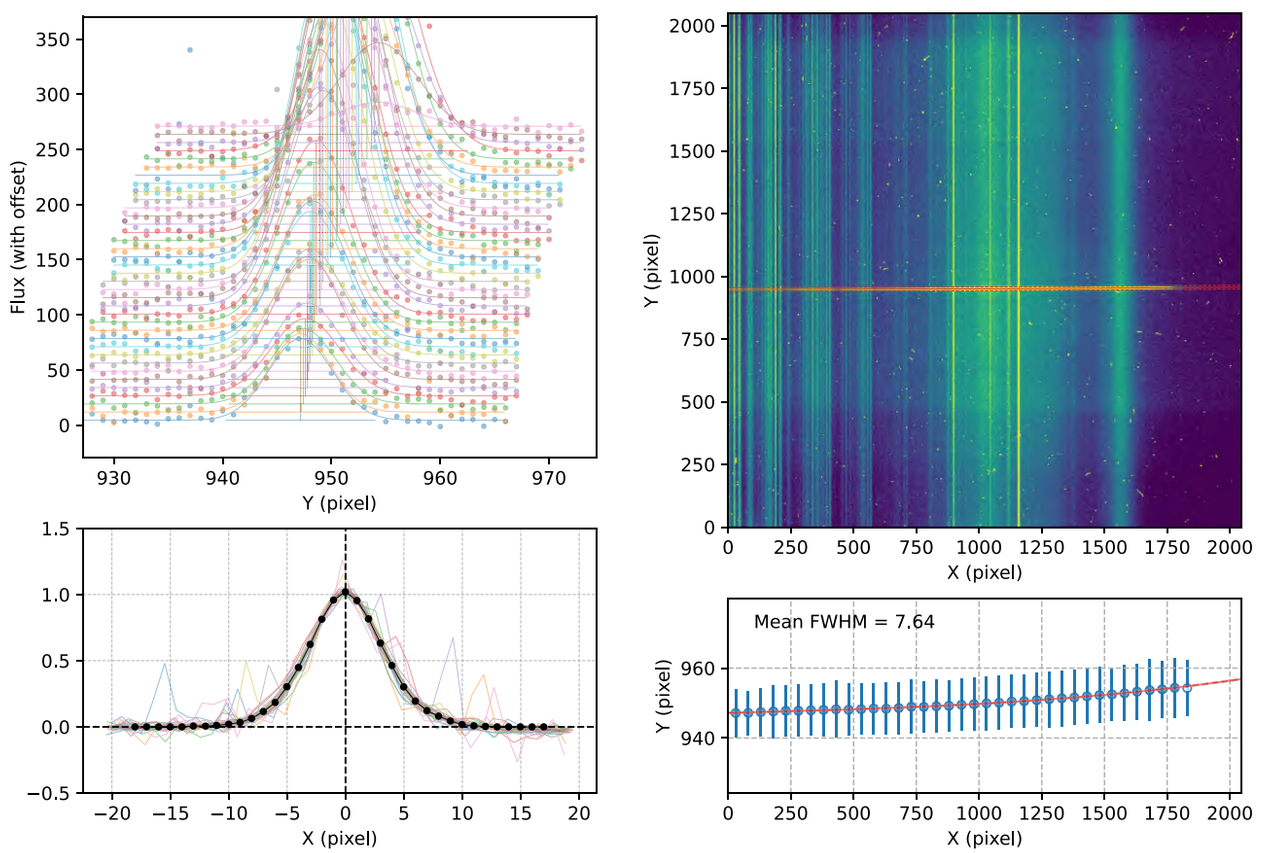

Figure 13. Target location and optimal extraction for long-slit mode of BFOSC. The upper left panel shows the profile fitting along the slit direction at different x, and the normalized profiles are plotted in the lower left panel. The solid black curve is the calculated averaged profile used in the optimal extraction. The upper right panel shows the bias-subtracted, flat-fielding, and distortion-corrected image, which is ready for subsequent background correction and 1D spectral extraction. The extraction aperture is plotted in orange. The target positions (described with a third-order polynomial) and FWHMs of profiles are plotted in the lower right panel.

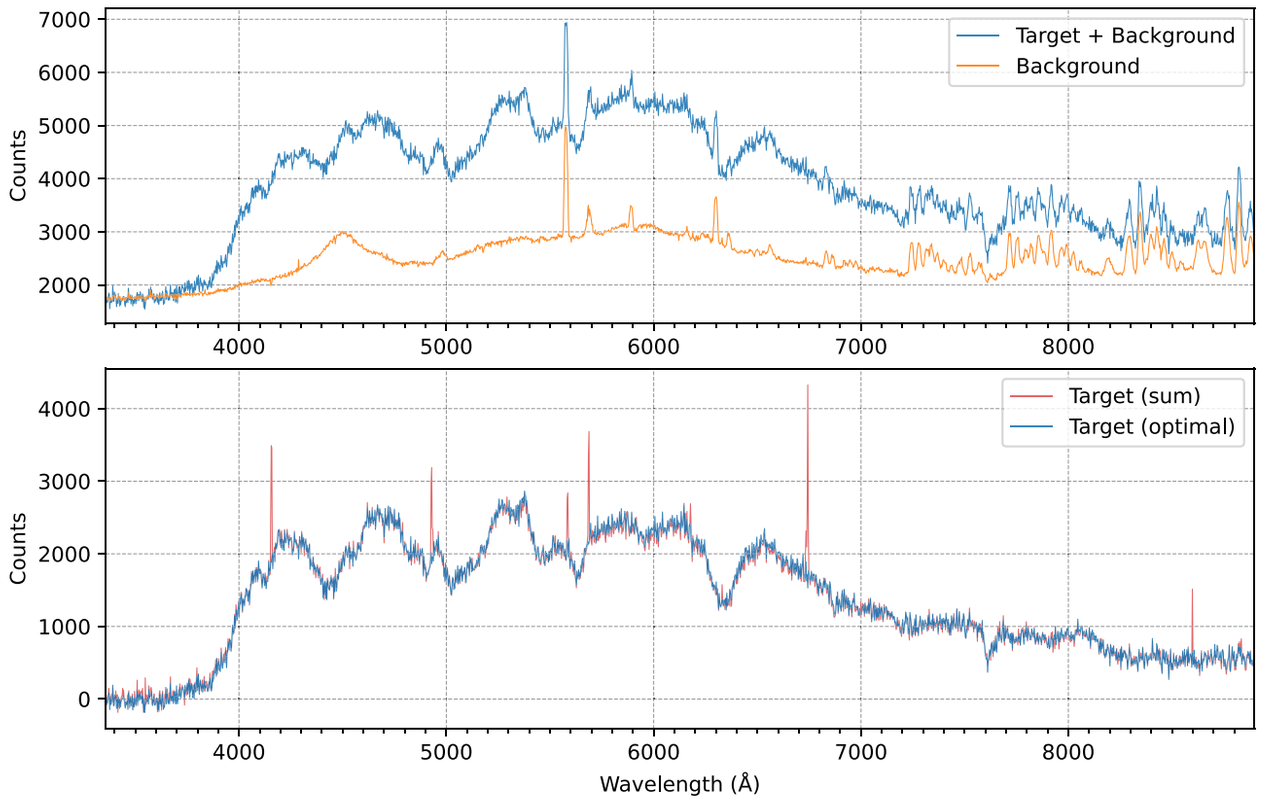

Figure 14. The blue solid line in the upper panel depicts the extracted spectrum using the image before the background subtraction, and the orange line is the spectrum of the background normalized to the width that equals the target extraction aperture. The resulting spectra after background subtraction with both sum-up and optimal extraction methods are plotted in the lower panel.

pixels. The optimal extraction method proposed by Horne (1986) has been widely used (e.g., Piskunov & Valenti 2002). In their approach, pixels along the slit direction are fitted by spatial profiles and are assigned different weights according to the photon noise and read-out noise. Cosmic rays can be identified and removed by placing an upper threshold on the residuals of the fitting. We adopted both direct summing and optimal extraction methods for long-slit and echelle spectra in BYSpec. The comparison is shown in Figure 14, where the optimal extraction removes all the six cosmic rays falling inside the extraction aperture. The target of this spectrum is a supernova (SN2023eoc) observed on 2023 April 18. Also plotted are the extracted spectra with and without background subtraction. It is clearly shown that the sky emission lines (e.g., the oxygen lines at 557.7 nm and 630.0 nm) are perfectly removed in the resulting spectrum.

3.9. Flux Calibration

BYSpec can execute the absolute flux calibration for long-slit mode if a spectroscopic standard star is observed on the same night. The absolute flux calibration aims to convert the ADU counts in the resulting spectra to physical units (e.g., erg s^{-1} cm^{-2} $Å^{-1}$), by taking account of atmospheric extinction of the observing sites, and differences on airmass and exposure time. BYSpec supports all spectroscopic standard stars listed in Hamuy et al. (1992, 1994) and Oke (1990). The SEDs are stored in the AWS cloud and can be downloaded automatically whenever needed, although some of them are too faint for 2 m telescopes.

For the echelle mode, absolute flux calibration is quite challenging. This mode requires correction for the blaze function to attain an accurate one-dimensional spectrum of the scientific target. Near the blaze angle, the spectral intensity reaches a maximum due to the grating characteristics, rather than the actual observed target. To nullify this effect, a blaze function correction is essential. The flat-fielding's ideal one-dimensional spectrum should exhibit uniform intensity, but blaze-induced intensity variations occur. Hence, dividing the extracted scientific target's one-dimensional spectrum by the flat-fielding spectrum accomplishes the blaze function correction.

4. Performance and Comparison

4.1. Adaptability and Compatibility

The YFOSC data closely resembles BFOSC data, with the primary difference being the use of a larger CCD in YFOSC. Additionally, in YFOSC, the slit direction aligns with the

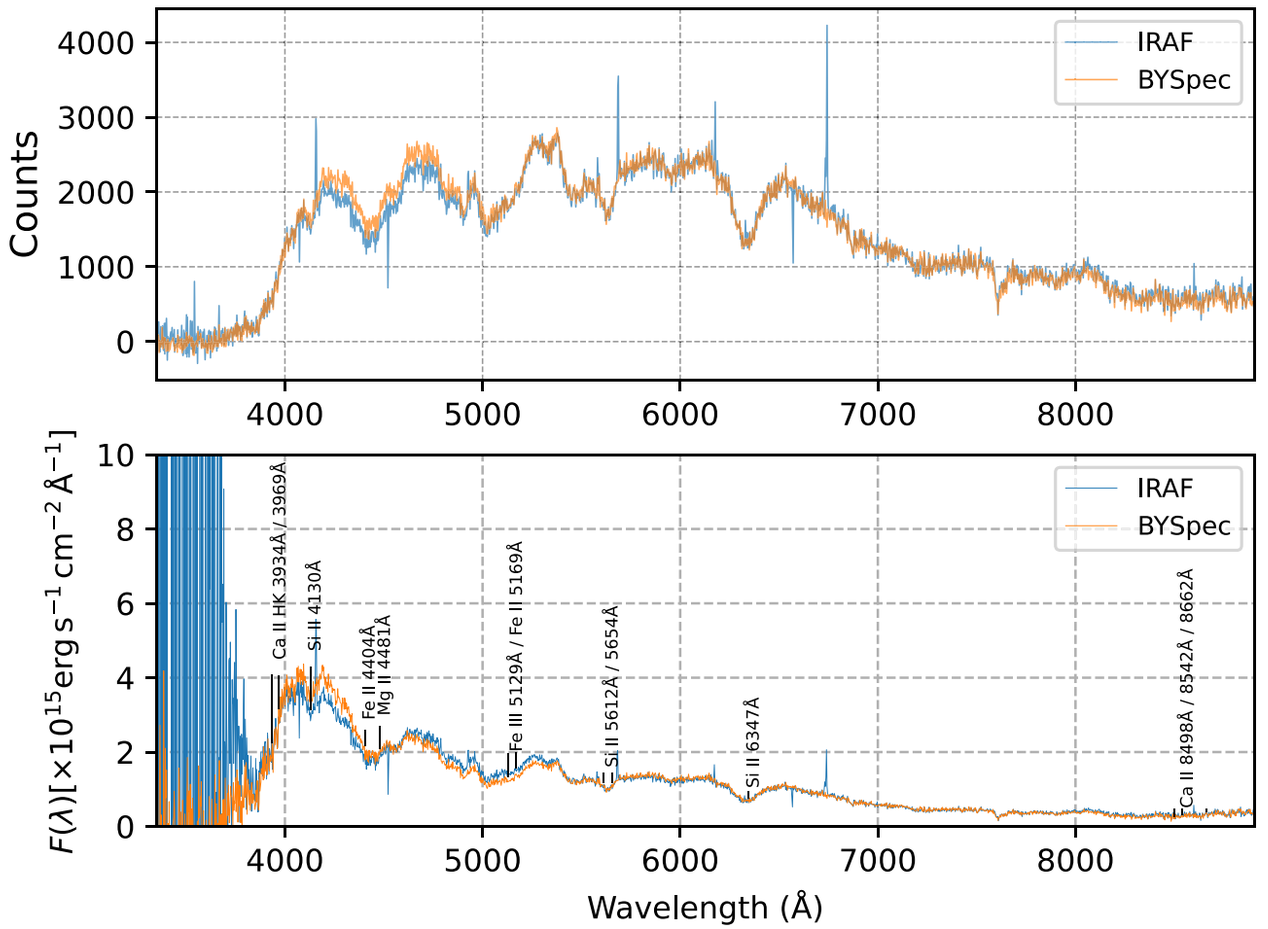

Figure 15. Top: Comparison of wavelength calibrated 1D long-slit spectra returned by IRAF (blue line) and BYSpec (orange line). Bottom: Comparison of the absolute flux-calibrated spectra. The target is a supernova (SN2023eoc), and some apparent absorption features are annotated.

X-axis, and the dispersion direction aligns with the *Y*-axis. These differences can be addressed by trimming and rotating the images accordingly. The YFOSC detector includes overscan regions, which facilitate monitoring variations in bias and readout noise over time. In BYSpec, these overscan regions are averaged along the readout direction and subtracted from the light-exposed pixels prior to bias correction. Apart from these adjustments, the data reduction procedures are identical to those used for BFOSC.

4.2. Autonomous Execution

Traditional data reduction with IRAF relies on a command-line interface that requires substantial user input at various stages, such as file classification, order location, background subtraction, and wavelength calibration. While IRAF includes its own scripting language (CL), its functionality is limited compared to modern programming languages like Python. To address these limitations, flexible and modern tools such as PyRAF have been developed, integrating IRAF tasks while offering enhanced features and usability. Our software, BYSpec, can automatically execute all essential data reduction steps, including overscan, bias and flat-fielding correction, order location, background subtraction, wavelength calibration, and flux calibration for BFSOC and YFOSC spectroscopic data. The calibration processes, which rely on external templates such as wavelength and flux calibration, are facilitated by downloading predefined files from web servers. The principle is to minimize human intervention and ensure a consistent data product for these two instruments. This is particularly convenient for observers using both BFOSC and YFOSC in the same research project.

4.3. Comparison with Other Packages

To validate the reliability of BYSpec, we reduced the same data set with IRAF and compared the results in Figure 15. The procedures with IRAF follow the standard manner, including

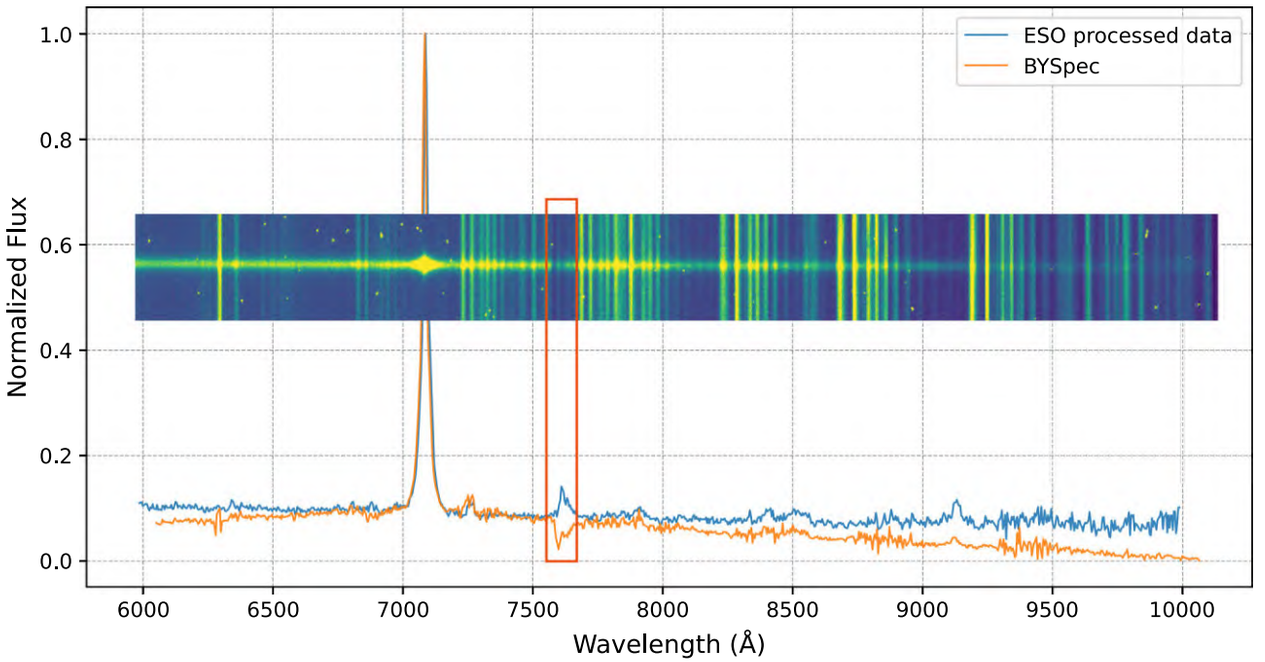

Figure 16. Comparison of reduced 1D spectra of EFOSC with BYSpec (orange) and processed data from ESO Science Portal (blue). A portion of the raw image is overplotted and the dispersion axis is generally aligned with the 1D spectra. Due to the nonlinear relation of pixel i vs. λ_i, small offsets can be seen on both edges.

bias subtraction, flat-fielding correction, background subtraction, and wavelength calibration. We got consistent results, and the one-dimensional spectra returned by BYSpec successfully removed the cosmic rays falling in the extraction aperture and the background regions, therefore the apparent outliers both above and below the IRAF spectra do not exist in our results. We also compared the spectra after the absolute flux calibration in Figure 15, and the difference is also insignificant. The large fluctuations below 4000 Å are due to the poor signal-to-noise ratios of both the target star and spectroscopic standard star. The reduction procedure using IRAF requires experience and the spectra output might be slightly different for different users. However, we can still conclude that BYSpec gets consistent one-dimensional spectra products with widely used software like IRAF.

We also adopted BYSpec to other similar instruments like EFOSC on the ESO 3.58 m New Technology Telescope (NTT) in La Silla. We downloaded the raw data of SN2016ezh taken in the long-slit mode on 2021 September 4, from ESO data archive.[8] We reduced the data following the same procedures as we set for BFSOC and YFOSC, and compared our results with the spectra downloaded from ESO Archive Science Portal[9] in Figure 16. It is shown that our result generally fits the spectra of ESO's pipeline, except for the telluric A band around 7600 Å, where ESO's pipeline exhibits a bump while our result has the absorption feature. We carefully checked the raw image and are confident with our result.

[8] https://archive.eso.org/eso/eso_archive_main.html
[9] https://archive.eso.org/scienceportal/home

4.4. Known Issues and Limitations

BYSpec is not perfect and has some issues. First, BYSpec cannot suppress the ring-like pattern that exists in the background of flat-fielding of BFOSC under the echelle mode, as shown in the right panel of Figure 6. The ring is probably caused by the stray light from the telescope or the optical elements inside the spectrograph. It causes an overestimation of pixel response as shown in the lower right panel of Figure 5. However, as the same pattern also appears in the target images and causes an overestimation of fluxes, this effect is somehow canceled. But we still remind the users to be cautious about any abnormal fluctuations at $x \sim 1600$ in orders below 5000 Å when using the BFOSC spectra returned by BYSpec. Second, an important advantage of long-slit spectrographs is that the spectra at different positions of an extended source (e.g., nebulae and galaxies) can be recorded simultaneously. However, BYSpec currently only supports single-point sources. We will continue to develop BYSpec for extended sources.

5. Conclusion

In this paper, we introduce our newly developed spectroscopic data reduction software called BYSpec. It can be used to

reduce the long-slit and echelle spectra of BFOSC on the Xinglong 2.16 m telescope and YFOSC on the Lijiang 2.4 m telescope. The data reduction procedures include overscan and bias correction, flat-fielding correction, distortion correction, background subtraction, one-dimensional spectral extraction, wavelength calibration, and flux calibration. The output of BYSpec is the wavelength-calibrated and flux-calibrated (if applicable) one-dimensional spectra, which are ready for scientific usage. The reliability of our software is validated by comparing the results with other software like IRAF. BYSpec is open-source, easy to use, and can be executed fully automatically if no complicated extraction for extended sources is required. Our software has built-in template wavelength and flux calibration files in the cloud server and therefore reduces the human interaction which requires extensive experience. As FOSC-like instruments are widely used in many telescopes, we will adopt the software for more similar instruments.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under grant No. U2031144. We acknowledge the support of the staff from the Xinglong 2.16 m and Lijiang 2.4 m telescopes. This work is partially supported by the Open Project Program of the Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences. J.J.Z. is supported by the National Key R&D Program of China with No. 2021YFA1600404, the National Natural Science Foundation of China (12173082), the Yunnan Fundamental Research Projects (grant 202201AT070069), the Top-notch Young Talents Program of Yunnan Province, the Light of West China Program provided by the Chinese Academy of Sciences, and the International Centre of Supernovae, Yunnan Key Laboratory (No. 202302AN360001).

ORCID iDs

Zi-Chong Zhang https://orcid.org/0009-0006-0874-3273
Jun-Bo Zhang https://orcid.org/0000-0002-3395-6178
Ju-Jia Zhang https://orcid.org/0000-0002-8296-2590
Jing Chen https://orcid.org/0000-0001-8869-653X
Ming-Yi Ding https://orcid.org/0000-0001-6898-7620
Liang Wang https://orcid.org/0000-0003-3603-1901

References

Baranne, A., Queloz, D., Mayor, M., et al. 1996, A&AS, 119, 373
Buzzoni, B., Delabre, B., Dekker, H., et al. 1984, Msngrr, 38, 9
Cretignier, M., Dumusque, X., Hara, N. C., & Pepe, F. 2021, A&A, 653, A43
Crosswhite, H. M. 1975, The Iron-neon Hollow-cathode Spectrum (Washington, DC: National Bureau of Standards (NBS))
Fan, Z., Wang, H., Jiang, X., et al. 2016, PASP, 128, 115005
Fu, Y., 2020 PyFOSC: A Pipeline Toolbox for BFOSC/YFOSC Long-slit Spectroscopy Data Reduction, v1.1.0, Zenodo doi:0.5281/zenodo.10967240
Hamuy, M., Suntzeff, N. B., Heathcote, S. R., et al. 1994, PASP, 106, 566
Hamuy, M., Walker, A. R., Suntzeff, N. B., et al. 1992, PASP, 104, 533
Horne, K. 1986, PASP, 98, 609
Martin, W. C. 1960, J Res Natl Bur Stand A Phys Chem, 1960, 19
Norlén, G. 1973, PhyS, 8, 249
Oke, J. B. 1990, AJ, 99, 1621
Palmer, B. A., & Engleman, R. 1983, Atlas of the Thorium Spectrum (Los Alamos, LA : National Laboratory)
Piskunov, N. E., & Valenti, J. A. 2002, A&A, 385, 1095
Saloman, E. B., & Sansonetti, C. J. 2004, JPCRD, 33, 1113
Savitzky, A., & Golay, M. J. E. 1964, AnaCh, 36, 1627
Wang, C.-J., Bai, J.-M., Fan, Y.-F., et al. 2019, RAA, 19, 149