



# DRAGraces: A Pipeline for the GRACES High-resolution Spectrograph at Gemini\*

André-Nicolas Chené<sup>1</sup> , Shunyan Mao<sup>2</sup>, Michael Lundquist<sup>3</sup>

(Gemini collaboration),

Eder Martioli<sup>4,5,7</sup>

(OPERA collaboration),

and

Jeffrey L. Carlin<sup>6</sup>

(Usability testing)

<sup>1</sup> Gemini Observatory/NSF's NOIRLab, 670 N. A'ohoku Place, Hilo, HI 96720, USA; [andrenicolas.chene@gmail.com](mailto:andrenicolas.chene@gmail.com)<sup>2</sup> Department of Physics and Astronomy, University of Victoria, PO Box 1700 STN CSC, Victoria, BC V8W 2Y2, Canada<sup>3</sup> University of Arizona, Steward Observatory, 933 N. Cherry Avenue, Tucson, AZ 85719, USA<sup>4</sup> Institut d'Astrophysique de Paris, UMR7095 CNRS, Université Pierre & Marie Curie, 98bis boulevard Arago, F-75014 Paris, France<sup>5</sup> Laboratório Nacional de Astrofísica, Rua Estados Unidos 154, 37504-364, Itajubá—MG, Brazil<sup>6</sup> Vera C. Rubin Observatory, Legacy Survey of Space and Time (LSST), 950 N. Cherry Avenue, Tucson, AZ 85719, USA

Received 2020 July 3; revised 2020 December 10; accepted 2020 December 14; published 2021 February 9

## Abstract

This paper describes the software Data Reduction and Analysis for GRACES (DRAGRACES), which is a pipeline reducing spectra from Gemini Remote Access to the CFHT ESPaDOnS Spectrograph (GRACES) at the Gemini North Telescope. The code is written in the IDL language. It is designed to find all the GRACES frames in a given directory, automatically determine the list of bias, flat, arc, and science frames, and perform the whole reduction and extraction within a few minutes. We compare the output from DRAGRACES with that of the Open source Pipeline for ESPaDOnS Reduction and Analysis (OPERA), a pipeline developed at the Canada–France–Hawaii Telescope (CFHT) that also can extract GRACES spectra. Both pipelines were developed completely independently, yet they give very similar extracted spectra. They both have their advantages and disadvantages. For instance, DRAGRACES is more straightforward and easy to use and is less likely to be derailed by a parameter that needs to be tweaked, while OPERA offers a more careful extraction that can be significantly superior when the highest resolution is required and when the signal-to-noise ratio is low. One should compare both before deciding which one to use for their science. Yet, both pipelines deliver a fairly comparable resolution power ( $R \sim 52.8k$  and  $36.6k$  for DRAGRACES and  $R \sim 58k$  and  $40k$  for OPERA in high- and low-resolution spectral modes, respectively), wavelength solution, and signal-to-noise ratio per resolution element.

*Unified Astronomy Thesaurus concepts:* [Software documentation \(1869\)](#); [Astronomy data reduction \(1861\)](#); [High resolution spectroscopy \(2096\)](#)

## 1. Introduction

Gemini Remote Access to CFHT ESPaDOnS Spectrograph (GRACES) is the high-resolution spectroscopy capability available at the Gemini North facility (Chené et al. 2014). It is the result of a cooperation between the Canada–France–Hawaii Telescope (CFHT), Gemini, and NRC Herzberg Astronomy and Astrophysics Research Centre (NRC Herzberg, HAA; Canada). It combines the large collecting area of the Gemini North telescope with the high-resolving power and high efficiency of the Echelle spectropolarimetric device for the observation of stars (ESPaDOnS) spectrograph at CFHT, to deliver high-resolution spectroscopy across the optical region. This is achieved through a 270 m fiber optic feed from the Gemini North telescope to ESPaDOnS. The on-sky diameter of the fiber is  $1''.2$ . GRACES achieves a maximum resolution power of  $R \sim 60,000$  between 400 and 1000 nm, with throughput redward of 600 nm higher than 10%, comparable to those of currently available high-resolution spectrographs in 8–10 m class telescopes. The Gemini Multi-Object Spectrograph

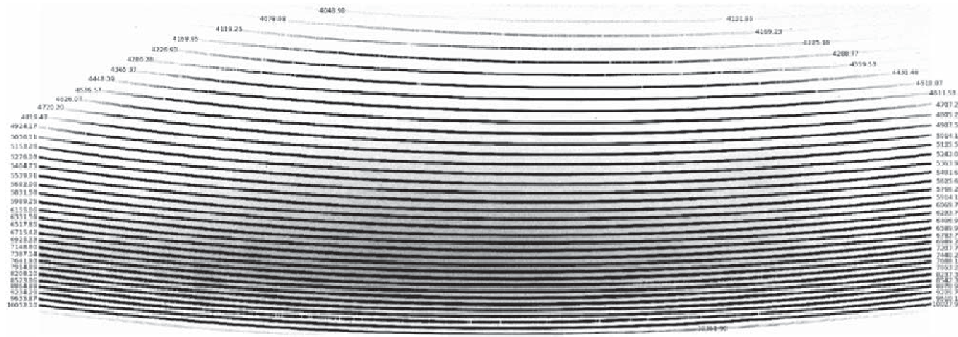
(GMOS-N) holds the fiber injector on the Gemini side, and is used to acquire targets. The GMOS-N On-Instrument Wave Front Sensor is primarily used for guiding, but the peripheral wavefront sensors (PWFSSs) could also be chosen under certain circumstances. On the CFHT side of the fiber, GRACES has its own bench with its own image slicer inside of the ESPaDOnS enclosure that feeds the light into the spectrograph.

GRACES has been offered at Gemini since the 2015B semester, and is planned to remain in operation until at least 2022. It is a visiting instrument, which means that it is not fully integrated into the Gemini system like a facility instrument would be. Among other things, Gemini does not support any data reduction software for GRACES. It depends on the Open source Pipeline for ESPaDOnS Reduction and Analysis (OPERA), developed at CFHT (Martioli et al. 2012), which is run at the end of each of the GRACES observing runs.

This paper describes the basic functionality of an alternative GRACES data reduction pipeline, the Data Reduction and Analysis for GRACES (DRAGRACES), originally developed as a safety backup during commissioning in case OPERA would not work with the GRACES data as expected, contained in the v1.3.1 release. As with most software, DRAGRACES' development is ongoing and features are being added regularly.

\* First released on 2016 September 20.

<sup>7</sup> CFHT Corporation, 65-1238 Mamalahoa Hwy, Kamuela, HI 96743, USA.



**Figure 1.** This figure shows the raw 2D spectrum of the A3 star HIP 57258. It also is the GRACES first light! On each side the starting and ending wavelength for each of the orders is written. A 90   rotation was applied to the frame, so the rows are now vertical and column horizontal.

DRAGRACES is an open-source code, written in the IDL<sup>8</sup> language, that is being developed under the git version-control system on GitHub: <https://github.com/AndreNicolasChene/DRAGRACES/releases/tag/v1.3.1>. The code uses the IDL Astronomy Library (astrolib; Landsman 1993), as well as the imdisp.pro function for figures displaying 2D images. Please note that contrary to what the affiliation of the first authors would suggest, the Gemini observatory is not involved with DRAGRACES development and maintenance.

## 2. About GRACES

### 2.1. Instrument’s Brief Description

GRACES consists of three components: (1) an injection module sending the light from the Gemini telescope into the GRACES fibers, (2) two 270 m long GRACES fibers, and (3) a receiver unit responsible for injecting the light from the fibers into the ESPaDOnS spectrograph at the CFHT. A complete description of the GRACES components is presented by Anthony et al. (2012). The receiver unit contains a bench holding the optics, an image slicer, a dekker blocking unwanted light from the slicer, a shutter, and a pickoff mirror sending the light to the spectrograph. The bench can rotate to switch between two slicing modes. One mode slices the image of the two fibers into two parts each, giving a lower spectral resolution mode ( $R \sim 40k$ ) called the two-fiber spectral mode. That mode is used when the sky is observed simultaneously with the target. The other mode slices the image of the one fiber in four, giving a higher spectral resolution mode ( $R \sim 60k$ ) called the one-fiber spectral mode.

Because spectra are obtained through most of the same optics and recorded on the same detector as the ESPaDOnS spectrograph, the data format is very comparable to that of ESPaDOnS at the CFHT (Manset & Donati 2003). The main differences are the position of the traces on the detector, the shape of the slices, and the telescope information (in the standard Gemini format) in the headers (see the next section).

### 2.2. Data Description and Access

GRACES raw and reduced data can be accessed from the Gemini Observatory Archive (GOA) search form.<sup>9</sup> Each frame is recorded in a single extension fits file. An example of a stellar spectrum is presented in Figure 1. Some scattered light is affecting the data, peaking somewhere around 7800   . Light contamination correction is covered in Section 3.9.

The GRACES data are similar, but not identical to those of ESPaDOnS. First of all, GRACES data have merely any signal in orders bluer than the 56th order. Also, the orders are not on the same pixels on the chip. At the central row of the detector, the difference in pixel of the center of the orders between the ESPaDOnS and GRACES spectra can be fairly well described with the parabola  $40 + 0.16x + 0.032x^2$ , though this information is not used in the pipeline. The width of the orders is the same for the two instruments, but the shape is not identical. One can compare the shape from ESPaDOnS shown in Martioli et al. (2012) with the shape of GRACES in Chen   et al. (2014). Finally, the header for GRACES follows the Gemini standards<sup>10</sup> and the header for ESPaDOnS follows the CFHT standards (Figure 2).

The required calibration frames are five flat-field (quartz halogen lamp) exposures, three arc (Th–Ar lamp) exposures, and ten bias exposures. Calibrations are taken in the evening before and in the morning after the science observations were taken. There is a maximum of one change of spectral mode per night. If the spectral mode stayed the same all night long, the calibrations that were taken the closest in time with the science data should be used. If the spectral mode did change during the night, the calibrations with the mode corresponding to that of the science data must be used. Note that there are strict rules as part of Gemini’s routine operation to guarantee that all the calibrations are always taken in the right modes for every night. Would any calibration go missing due to a mistake or unforeseen circumstances, all the science data of the night would not pass the quality assessment and the corresponding observations would be put back into the queue to be rescheduled.

### 2.3. Motivation

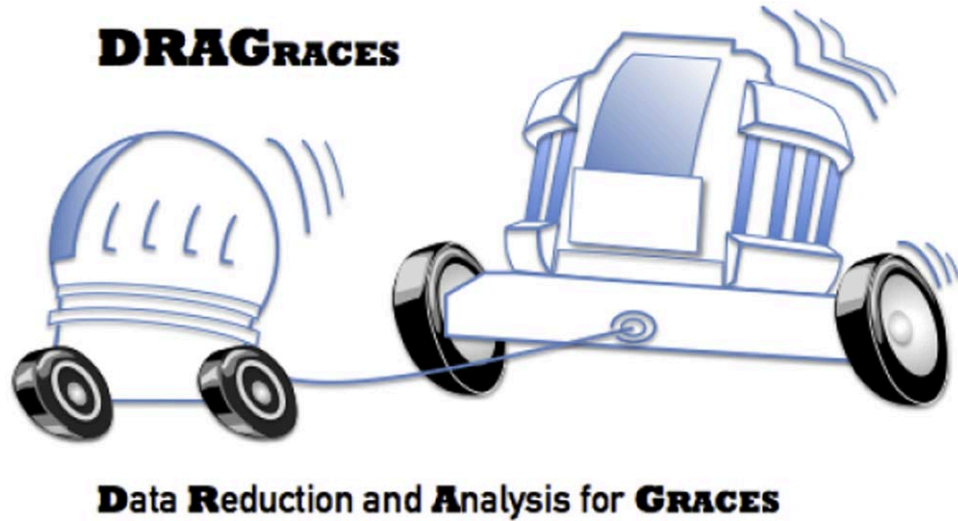
GRACES was a quick and low-cost option to allow high-resolution spectroscopy to the Gemini Observatory. It of course takes advantage of the exceptional efficiency of the ESPaDOnS spectrograph, but also of the unprecedented quality of the 270 m long fiber between the telescope and the instrument. GRACES is also a demonstrator of the possibility to leverage the use of a shared instrument between more than one observatory without having to move the instrument.

DRAGRACES was first developed as a quick look tool for nighttime operations and was later expanded as a full pipeline. It was also developed as a backup solution before it was known whether OPERA would be usable to extract GRACES data. Since

<sup>8</sup> Interactive Data Language.

<sup>9</sup> <https://archive.gemini.edu/searchform>

<sup>10</sup> <http://www.gemini.edu/observing/phase-iii>



**Figure 2.** DRAGRACES logo features both the CFHT and Gemini North domes, linked together by the GRACES fiber. The CFHT dome is pictured as the driving force, illustrating the central role that the ESPaDOnS spectrograph and the CFHT support team play in GRACES. Drag races are also popular events on the Big Island of Hawai‘i.

OPERA ended up being successful in doing so, the Gemini Observatory delivers extracted data obtained using the OPERA pipeline (Martoli et al. 2012) and DRAGRACES became more of an alternative for those who would prefer using it (e.g., Carlin et al. 2018). But more importantly, DRAGRACES has the advantage of being under development still, as opposed to OPERA for which all development have been discontinued. This means that DRAGRACES can be continue to improve and evolve based on the users’ feedback. Since neither OPERA nor DRAGRACES have been extensively tested using a wide range of data type and scenarios, comparing the pipelines’ output is a safe way to proceed, and a potential way to improve DRAGRACES’ performance. DRAGRACES may eventually be adopted as the default pipeline at Gemini, not only improve the analysis of future projects but also of the archival data.

### 3. The Pipeline

The pipeline will process all the files it can find in a given directory. It can identify the different read modes and spectral modes, and process them separately. In summary, the steps are:

1. sorting the fits files based on observation type (bias, flat, arc, and object) as defined in the OBSTYPE header keyword;
2. creating the master bias, master ThAr, and master flat-field frames;
3. finding the trace of the orders on the master flat-field frame;
4. determining the slit tilt;
5. normalizing the flat field in 2D;
6. identifying the arc lines and calculating the wavelength solution in 1D;
7. correcting for background illumination in 2D;
8. extracting the science spectra; and
9. wavelength calibrating the science spectra.

#### 3.1. Accessing the Source Code

The pipeline is open source, version controlled, easy to use, and available to the community.<sup>11</sup>

#### 3.2. Assumptions

1. All the necessary files are in the same directory.
2. The spectrograph is stable enough, so the trace of the orders will always have the same width and the same shape.
3. The spectrograph is stable enough, so the orders will not move significantly on the detector within  $\sim 10$  hr of observations.
4. The spectrograph is stable enough, so the orders will never move by more than a few pixels with respect to where they were during the 2015 commissioning.

These assumptions are supported by many years of calibration data. The only occasion where the position of the traces shifted on the detector was after several components needed to be replaced due to a lightning strike in 2018 August. DRAGRACES was updated to use the right hard-coded values in pixels to search for the traces depending whether the data are from before or after the lightning event.

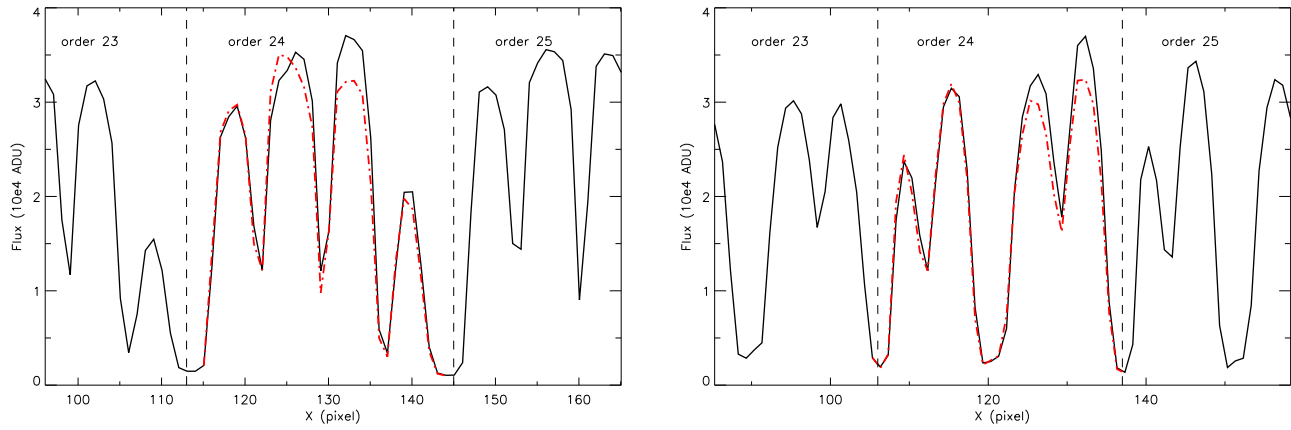
#### 3.3. List Creation

DRAGRACES automatically creates lists of files. It begins by looking into the raw data directory collecting all of the FITS files with a file name that start with an “N,” followed by eight digits (for the UT date), then the letter “G,” and four digits (for the file number), e.g., N20170611G0005.fits, which is the standard Gemini GRACES format.

All the frames are sorted based on the header keyword “OBSTYPE,” which can either be “BIAS,” “FLAT,” “ARC,” or “OBJECT.” The bias frames are divided into slow (“READNOISE” =  $2.9 e^-$ ) and normal (“READNOISE” =  $4.2 e^-$ ) readmode. The other frames are divided into one-fiber (“GSLIPOS” = “FOURSLICE”) and two-fiber (“GSLIPOS” = “TWO SLICE”) spectral modes. Of course, the lists of bias, arcs, or flats from the user is used instead of the automated ones, when provided.

If the data for more than one night are in the same directory, it is recommended to reduce each of the nights using the UT dates option. Otherwise, all the calibration files present will be combined and applied to all of the object frames.

<sup>11</sup> The source code and necessary files are accessible through GitHub: <https://github.com/AndreNicolasChene/DRAGRACES>.



**Figure 3.** Cut of the flat-field image through row 2004, centered on the order of 24 for the one-fiber spectral mode (left panel, observed on 2017 June 18) and the two-fiber spectral mode (right panel, observed on 2017 June 11). The vertical dashed lines mark the limits of the order of 24, and the red dashed-dotted line is the order’s profile at row 2304 (i.e., the middle of the detector) shifted to center of the selected trace.

Sometimes the calibrations obtained in the evening are in the same mode as the ones obtained in the morning. In such case, it is recommended to use only the set that is the closest in time to the spectra to process.

If the user wants to impose lists of bias, flat, and/or arc frames, the file names of the lists can be entered as an optional input (see `dg.pro` help). The lists have to be in the ASCII format with the fits files listed as a single column (one list per calibration type).

### 3.4. The Master Calibration Frames

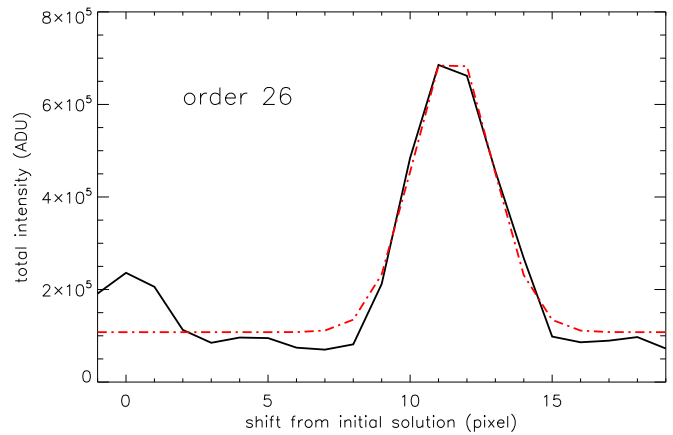
Function: `overs_corr`

Bias, arc, and flat frames are overscan corrected. The bias frames are combined first using a simple average. The resulting bias master frame is used to correct each individual arc and flat frame before they are themselves combined into their respective master frames. Note that there are no cosmic-ray rejection methods implemented at the moment. A visual inspection of the calibration frames is recommended before including them for the reduction.

### 3.5. Finding the Orders

Function: `find_trace`

The order search starts in the middle of the detector, i.e., on the detector’s row 2304, using the master flat-field frame. At that row, it is assumed that the centers of the orders are near values that are hard-coded in the pipeline. The level of stability of the instrument allows the assumption that those values are within a few pixels from the current centroids. The current centroids are determined using a cross-correlation technique to compare each order’s profile (as observed on row 2304, i.e., the middle of the detector) with a model, which is assumed to be a sum of four Gaussian functions with a sigma of two pixels and separated by a fixed number of pixels (see an example in Figure 3). Determining those centroids for each order within one pixel accuracy is sufficient, since the goal is to isolate the profile of the trace at the middle row of the detector. For each order, the profile at row 2304 replaces the four Gaussian models in the cross-correlation that determines the centroid of the trace at every row. The measured centroids along the detector as a function of the X pixels (vertical) are fitted with a polynomial of an order of 4, and the coefficients are saved in a file to use later during the extraction.



**Figure 4.** Convolution of the order of 26 of the observed ThAr with a comb function that consists of delta peaks with an amplitude of 1 at each wavelength where a spectra line is expected. The result is a weighted-mean line profile. The red dashed-dotted line is the result of the Gaussian fit. Note that the line is wide due to the smearing caused by the slit tilt.

### 3.6. Determining the Slit Tilt

Functions: `reduce`, `extract`, `find_lines`

The slit tilt is introduced by the instrument, with the intention of preventing any overlap of the redder orders on the chip. While this is practical for optical considerations, it causes the need for an extra step before moving on with the reduction.

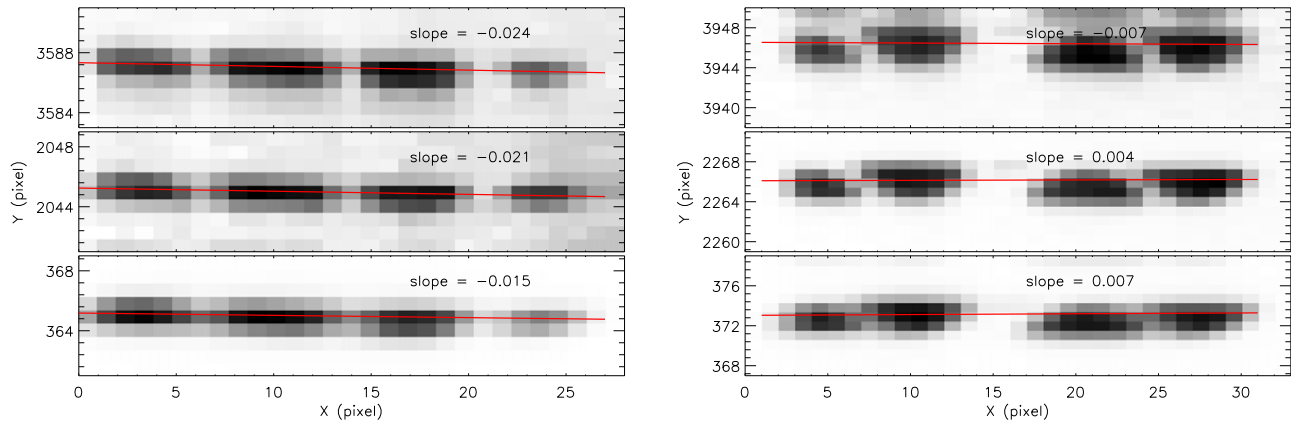
To determine the slit tilt, the pipeline measures the slope of tens of identified ThAr spectral lines per order on the arc master frame along each order.

#### 3.6.1. First Crude Wavelength Solution and Line Identification

The code contains a first estimate of the wavelength solution for each order. It was originally obtained using the function `IDENTIFY` from IRAF.<sup>12</sup> on an extracted ThAr frame observed during GRACES commissioning. At the first pass, this estimated solution is applied to a 1D arc master frame, extracted using a simple sum of all the columns in each order. The crude (i.e., not exact, but close) wavelength solution is assumed to be a shift of a few pixels from the initial one.

<sup>12</sup> IRAF is distributed by the National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.





**Figure 5.** Close-up view of three selected ThAr lines on the 2D frame in both the one-fiber (left) and two-fiber (right) spectral modes. The lines were chosen to be located as far away from one another as possible along the order. The red slopes are the results of the linear fit for each line.

To determine and apply the shift for each order, the pipeline

1. creates a comb function that has a delta peak with an amplitude of one at each wavelength where a line is expected, following the ThAr line list given by Palmer & Engleman (1983);
2. masks the saturated lines on the ThAr spectrum and all the pixels affected next to them;
3. convolves the observed ThAr spectrum with the comb function (the result is a single weighted-mean profile, see Figure 4);
4. fits a Gaussian function on the resulting profile to determine the centroid (see Figure 4); and
5. shifts the first estimate of the wavelength solution by  $-1$  times the value of the centroid determined by the fit.

### 3.6.2. Measure of the Tilt

Thanks to the procedure that determines the crude wavelength solution, we now have the position in pixels of many of the ThAr lines. The script selects a stamp from the 2D ThAr master spectrum centered on each identified line with an amplitude higher than 50 ADUs to avoid small lines, and lower than 65,000 ADUs to avoid saturated lines. The stamp is as wide as the order’s trace, and as high as four times the expected FWHM of the spectral lines. To determine the slit tilt, the script does a cross-correlation of the corresponding line profile from the 1D arc spectrum with each column in the stamp. This determines the center of light across the spectral line. The tilt is defined as the slope of a robust linear fit of the measured centers. The tilt is determined for the whole order at once (see Figure 5). Because there is a clear trend in the tilt determined as a function of wavelength (or Y pixel along the chip), a robust linear fit is done to determine the tilt to correct for a function of the Y pixel for each order.

A final wavelength solution is found later (see Section 3.8), but for now only the slit tilt value for each order is required.

### 3.7. Flat-field Normalization

Function: `reduce`

Now that the trace and the slit tilt are known for all of the orders (see Figure 6), the pipeline can normalize the flat field. Each order (and each fiber independently per order in the two-fiber spectral mode) is rectified and converted into a straight stripe.

#### 3.7.1. Orders Rectification

At each line, a bilinear interpolation is used to get the flux within an order that goes along a slope corresponding to the slit tilt, and centered on the order’s trace. The interpolated vector becomes the corresponding line of the rectified order. Science frames are later rectified in the exact same way.

#### 3.7.2. Normalization

The pipeline fits a polynomial of the order of 6 to each of the columns of the rectified flat field. Each column is divided by its fit. The resulting normalized flat-field frame is a mosaic of all the rectified orders aligned next to each other, and all the values are around 1. The edges of the orders are suffering from some sampling effects caused by the rectification (from the interpolation), and are later excluded from extraction.

### 3.8. Final Line Identification

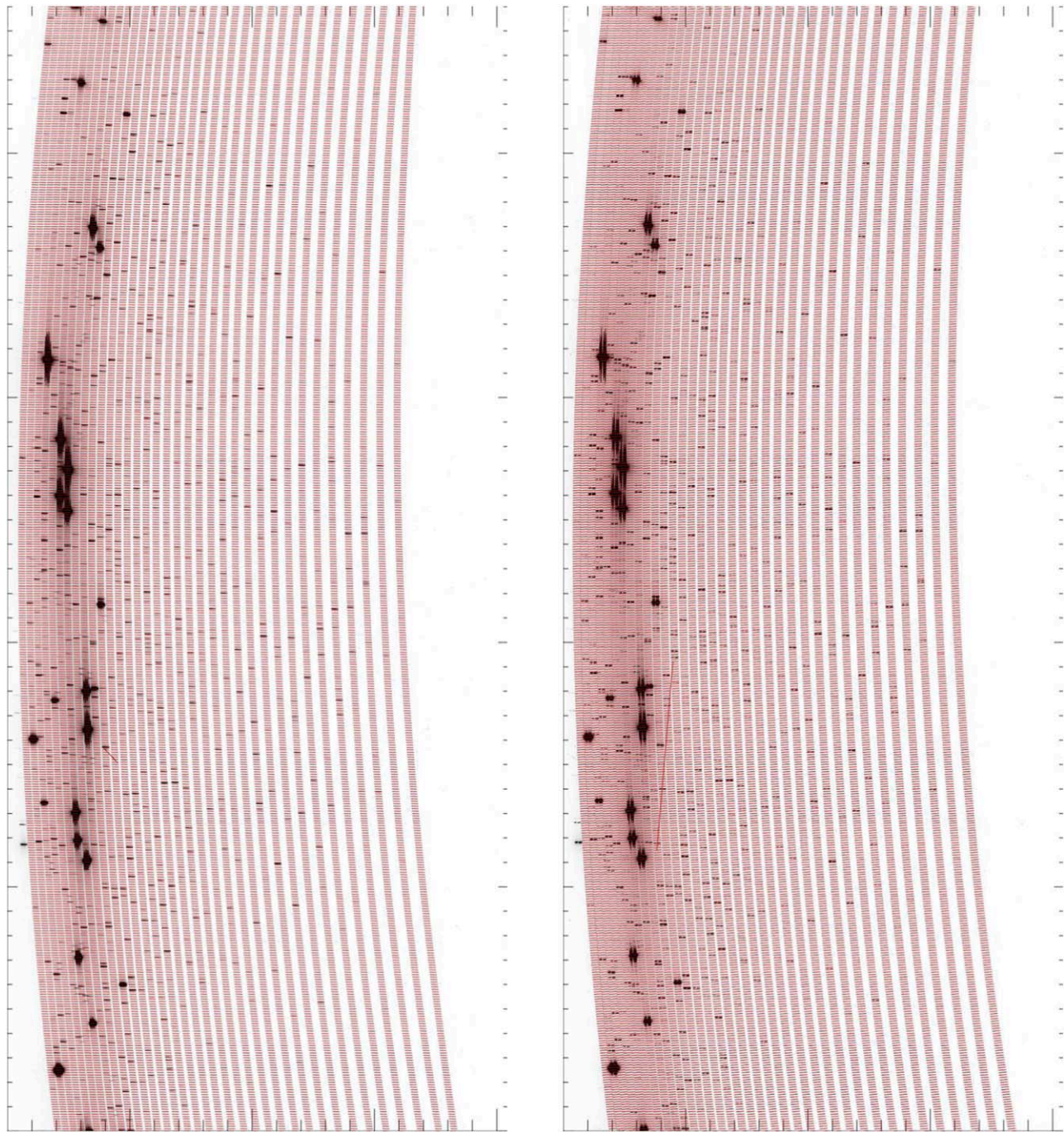
Functions: `reduce`, `extract`, `find_lines`, `wavel_sol`

The final arc master spectrum is rectified as described in Section 3.7.2, and extracted using a simple sum of the columns in each order/fiber. There is no clipping or cosmic-ray rejection method implemented at this point. The last 300 pixels are rejected, since this region is severely affected by vignetting. This final extraction is giving the best resolution, since the tilt is corrected before the spectrum is extracted. The crude wavelength solution of the new 1D ThAr spectrum is applied again as described in Section 3.6. It is used to find again the spectral lines, which are fitted individually assuming a Gaussian profile. This gives the real positions in pixels for each ThAr line identified from the ThAr line list (Palmer & Engleman 1983).

### 3.9. Illumination Correction

Function: `illumcor`

The orders are well separated, even on the reddest side of the spectrum (i.e., down to the order of 22). Indeed, there are at least a few pixels between orders where the flux should be down to  $\sim 0$  ADU. However, there is flux detected in these regions toward the reddest orders, and a stronger signal in the spectrum gives a stronger effect. To correct for this effect, the pipeline fits and subtracts the flux measured in the area

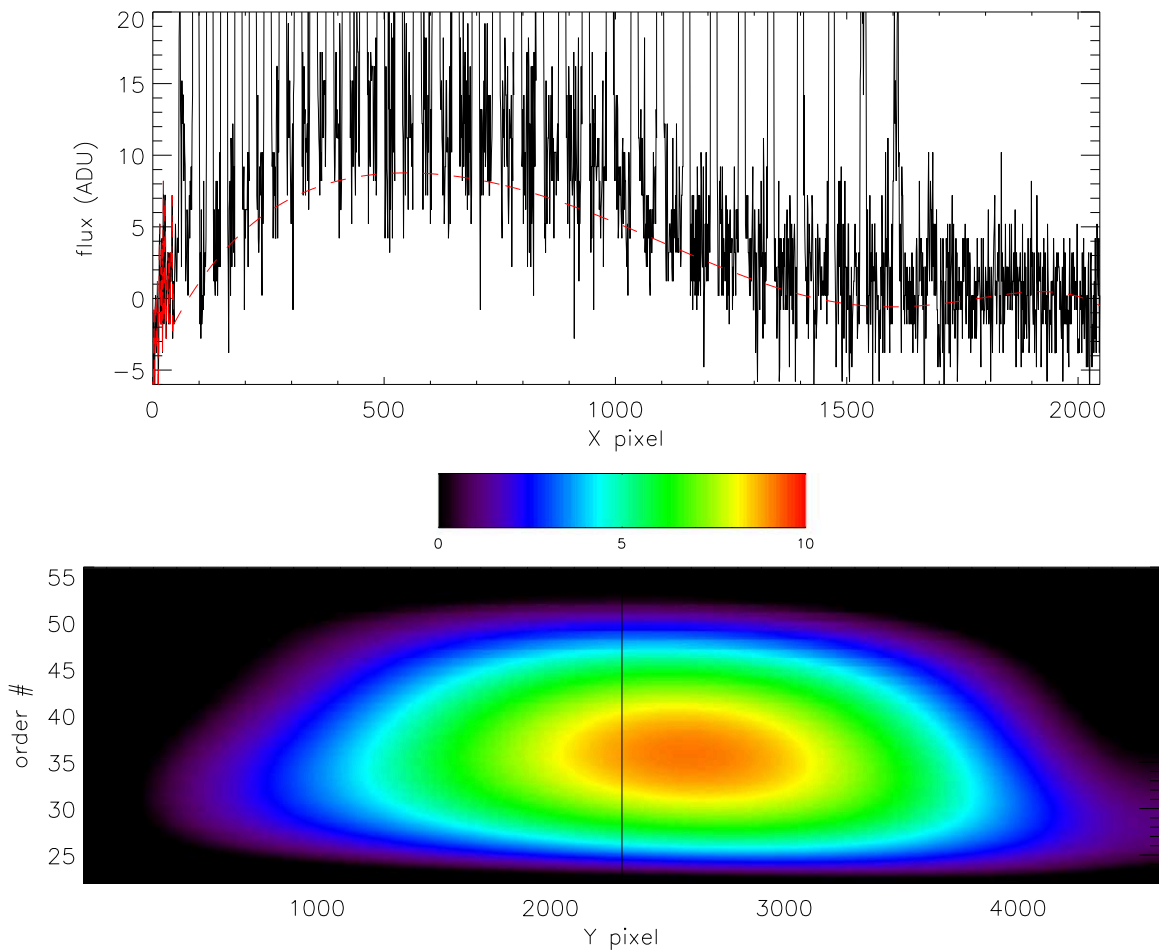


**Figure 6.** Examples of ThAr spectra (one-fiber spectral mode to the left and two-fiber spectral mode to the right) and the orders traced in red. Each red line has the slope corresponding to the slit tilt at the corresponding wavelength.

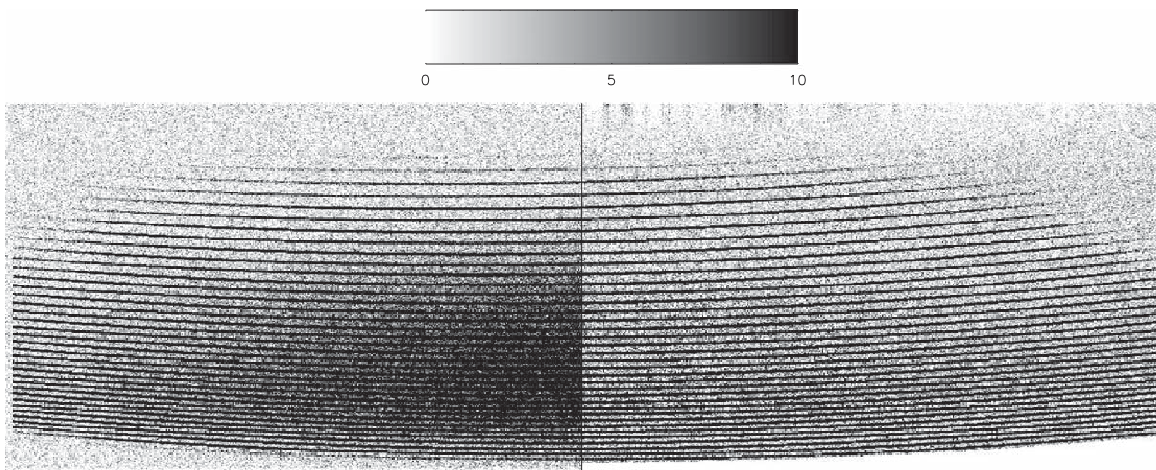
between the orders on each of the science exposures (unless the user specifically chooses to avoid it) using the following steps.

1. The level of illumination between each order is measured on the 2D frame. This is done line by line, and by looking at each interval one by one. If the interval is 3 pixels wide or less, the illumination is assumed to be the minimum value. If it is wider, it is assumed to be the minimum of a polynomial fit of the order of 3 of the interval (to avoid strong dependence on the noise). All the values are recorded in a matrix that contains the background illumination value for each order in each line.
2. The background illumination values along the  $Y$ -axis (lines) are smoothed using a 40 pixels bin.
3. Line by line, a polynomial of the order of 6 is fit to estimate the background illumination values within the orders (see the top panel of Figure 7), and all of the results are stacked in a 2D image.
4. The traces are rectified and the slit tilt is corrected, just as described in Section 3.7.1.
5. For each order, a surface using a polynomial of the order of 5 is fit and the final background illumination frame is created to subtract from each of the reduced science frames (see the bottom panel of Figure 7).





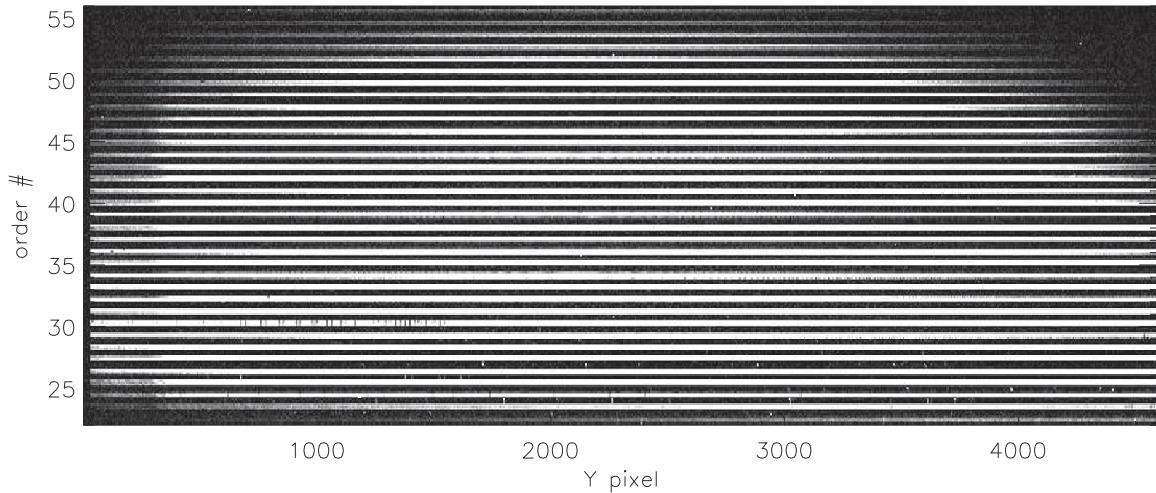
**Figure 7.** Top: cut of the 2D flat-field frame across the orders in the middle of the detector. The red dashed line is the fit to the flux measured in the area between orders. Bottom: 2D surface of the illumination. The black vertical line marks where the cut plotted on the top panel was extracted.



**Figure 8.** Sample 2D frame of the spectrum of BD+28 4211. The left half is the raw frame, while the right half was corrected for contaminating background illumination. The function *illumcor* does not correct the 2D spectrum before the orders are rectified, and this figure was created to illustrate the importance of removing the light contamination.

Note that bluer wavelengths (around and past X pixel 1500) are left out of the fit, and the background often diverges. It was decided to leave that section out, so the polynomial fit would be less constrained while trying to reproduce the background at redder wavelengths. Also, this causes no real losses, since there is no more useful signal in that region.

The function *illumcor* corrects for any contaminating light source, such as stray or scattered light. Such a correction becomes even more important as the exposure time is long. Figure 8 shows an example of a nonrectified 2D science frame with (right) and without (left) background illumination correction (Figure 9).



**Figure 9.** Reduced 2D science spectrum. All the orders are now rectified and stacked next to one another.

### 3.10. Final Extraction

Functions: `overs_corr`, `reduce`, `illumcorr`, `extract`

At this point, the pipeline is ready to reduce the science frames and extract the spectra.

#### 3.10.1. Primary Reduction

Each science frame is reduced using the following steps:

1. overscan correction,
2. bias subtraction,
3. order rectification and slit tilt correction, as described in Section 3.7.1,
4. background illumination correction, as described in Section 3.9, and
5. division by the normalized flat obtained as described in Section 3.7.

#### 3.10.2. Spectral Extraction and Wavelength Calibration

Each order is extracted using a simple sum of the lines. There are no cosmic-ray rejection or any clipping implemented at the moment.

For each extracted order, the pipeline calculates a wavelength solution using a polynomial fit of the order of 4 on the pixel position of the ThAr lines (identified as described in Section 3.8) as a function of central wavelength. To apply this solution, the extracted spectrum is first linearized by resampling each order using a spline function. The step is chosen to be determined by the highest resolution measured within the order of the wavelength range.

It is possible to skip the wavelength solution, if the user prefers to do a separate calibration. In this case, the arc spectrum and the science spectra are simply extracted and saved.

The spectra are not normalized by the continuum or flux calibrated. Spectra of the same target are not combined into a single spectrum, but instead they are recorded in separate fits files. The wavelength solution is not corrected using telluric lines and is not adjusted to correct for heliocentric motion. In brief, it is a

simple sum of the flux. Any correction, combination, calibration, or normalization needs to be done separately.

## 4. Organization of the Reduced Data

The reduced spectra are multi-extension fits files. The file name is the same as the original file, with the prefix “ext\_.” Each extension has one section of the spectrum, corresponding to one of the orders. For the one-fiber spectral mode, there are 35 extensions, with each order from 22 to 57 on each extension. For the two-fiber spectral mode, there are twice as many extensions, as for each order, there is the sky and the target spectra which are each on a different extension. Each extension contains a unique wavelength solution for the corresponding order, using the keywords CRVAL1, CRPIX1, and CD1\_1.

In the two-fiber spectral mode, since the target and the sky are recorded in separate extensions, one needs to perform the final sky subtraction outside of DRAGRACES. Note that because the sky fiber has a smaller transmission than the target fiber, one has to apply a factor 1.25 to the sky spectrum before subtracting it from the target fiber signal.

## 5. Comparison with OPERA

It is quite surprising, yet reassuring, to compare the extracted spectra from OPERA and DRAGRACES and see how similar they are. Yet, there are some differences, most of which can be explained by some of the inherent differences between the two pipelines.

The fundamental difference between the two pipelines relies on the method for extracting the fluxes. OPERA uses a tilted rectangular aperture which is oversampled in a subpixel grid, where they apply either the subpixel sum or the optimal extraction method (Horne 1986), while DRAGRACES uses the simple sum of pixels in a rectified row. The use of optimal extraction explains why OPERA is more robust to bad pixels and cosmic rays and the oversampling explains the slightly higher resolution obtained by OPERA in the on-slice mode. On the other hand the main advantage of DRAGRACES is how straightforward it is to download and run it, without further installation required, and how easy it can be for anyone with an



**Table 1**  
Comparison between DRAGRACES and OPERA

Feature	DRAGRACES	OPERA
Language	IDL	C/C++ and Python or Make file
Open source	yes	yes
License type	MIT	GPL 3.0
Repository for download	yes <sup>a</sup>	yes <sup>b</sup>
Online documentation	yes <sup>c</sup>	yes <sup>d</sup>
Identify files for reduction	yes	yes
Interactive	no	no
Run automatically	yes	yes
Stack calibration data	yes	yes
Correct for overscan	yes	no
Subtract bias	yes	yes
Correct for scattered light	yes, by global polynomial fit	no
Perform flat-fielding	yes, on 2D frame	yes, on 1D spectrum
Correct for blaze	no	yes
Account for slit tilt	yes	yes
Method to compensate slit tilt	interpolation/rectification	use a tilted aperture
Allow variable slit tilt	yes	yes, but constant tilt per order
Extraction aperture shape	rectified rectangle ( $1 \times n$ pixels)	oversampled tilted rectangle ( $m \times n$ pixels)
Allow oversampling <sup>e</sup>	no	yes
Extraction method	sum	sum or optimal extraction <sup>f</sup>
Calculate flux errors	no	yes
Robust to cosmic rays	no	yes
Wavelength calibration	yes	yes
Line detection method	convolution with a comb function	cross-correlation with 2D inst. profile
Line measurement method	Gaussian fit	multiple Gaussian fit
Line rejection method	sigma-clipping	sigma-clipping
Th–Ar atlas	Palmer & Engleman (1983)	Lovis & Pepe (2007)
$R = \lambda/\Delta\lambda$ (1-fiber)	$52.8 \pm 0.1$	$56.7 \pm 0.7$
$R = \lambda/\Delta\lambda$ (2-fiber)	$36.3 \pm 0.3$	$35.7 \pm 0.9$
RV accuracy	$\sim 500 \text{ m s}^{-1}$	$\sim 100 \text{ m s}^{-1}$
Provide RV correction	no	yes, heliocentric
Continuum normalization	no	yes
Products in FITS format	yes	yes
Development status	ongoing	interrupted

#### Notes.

<sup>a</sup> <https://github.com/AndreNicolasChene/DRAGRACES>.

<sup>b</sup> <https://github.com/CFHT/OPERA>.

<sup>c</sup> <http://drforum.gemini.edu/topic/graces-pipeline-dragraces/>.

<sup>d</sup> <http://wiki.lna.br/wiki/espectro>.

<sup>e</sup> Oversampling meaning sampling spectral elements at steps smaller than 1 pixel.

<sup>f</sup> Optimal algorithm by Home (1986).

IDL license to customize it. Table 1 presents a summary of the features to compare the two pipelines.

#### 5.1. Flux, Signal-to-noise Ratio, and Wavelength Solution

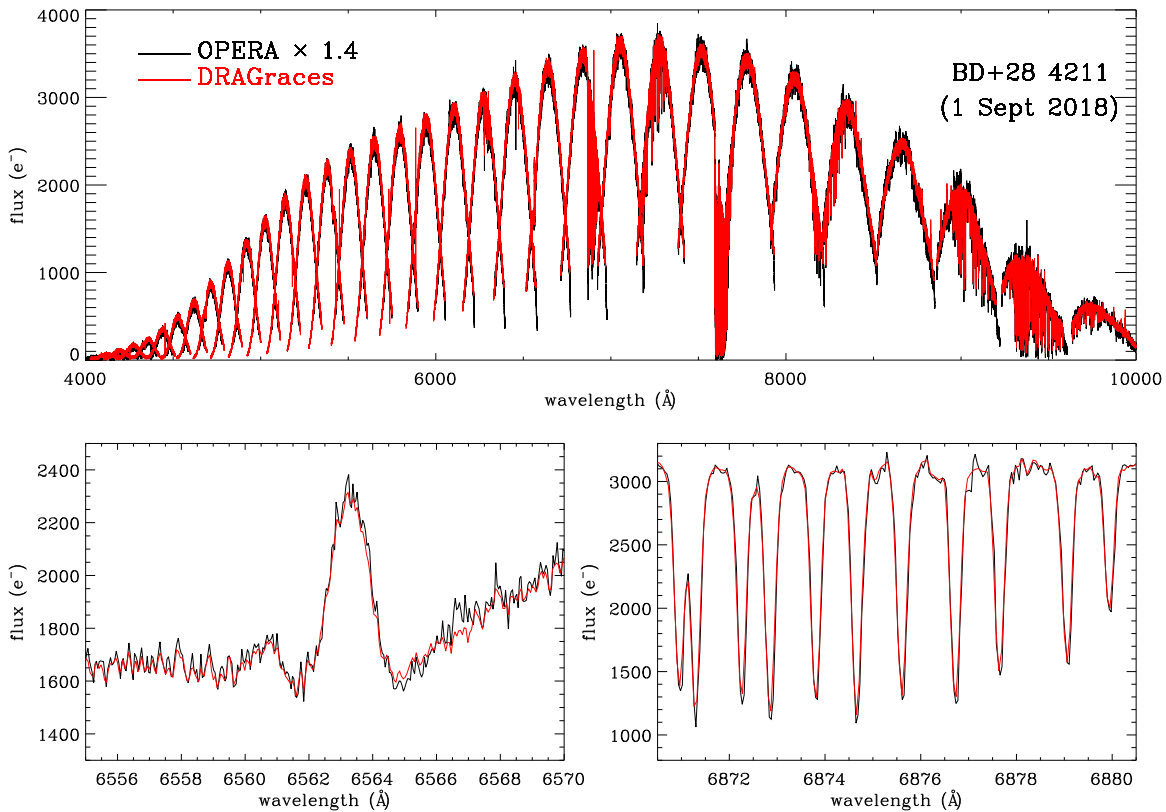
At first look, the spectrum extracted with the OPERA spectrum has a lower flux. But this is just a result of the resampling performed by OPERA to optimize the spectral resolution. The resampling gives a wavelength step equivalent to  $\sim 0.7$  pixel (the exact value is not public), which shows that the spectrum has to be multiplied by a factor 1.4 to match the flux of the spectrum produced by DRAGRACES (see Figure 10).

Spectra extracted with OPERA and DRAGRACES have the same signal. A close-up on the H $\alpha$  line (Figure 10) shows that the two extractions give the same spectrum, for the most part. Because the DRAGRACES sampling is more coarse, it has a higher signal-to-noise ratio per wavelength step. Yet, we can recognize the same noise pattern in the two spectra.

Finally, the wavelength solutions are identical for the two pipelines. A close-up to any narrow telluric line shows that there is no apparent shift between the two spectra. This is reassuring, since the wavelength solution seems to be independent to the assumptions used in both codes.

#### 5.2. Resolution

We extracted a ThAr spectrum using the two pipelines to compare the spectral resolutions they achieve. On the extracted spectra, we selected all the lines with a signal-to-noise ratio superior to 10, rejected lines that are within one resolution element from each other, and rejected lines that are not identified in the ThAr line list (Palmer & Engleman 1983). For each selected line, we used a Monte Carlo optimizer and a Gaussian model to fit the profile, while trying different wavelength interval cutoffs to get the best result. The results are plotted in Figures 11 and 12. In the one-fiber spectral



**Figure 10.** Comparison of one spectrum of the star BD+28 4211 observed on 2018 September 1, and reduced with both OPERA and DRAGRACES. The observations were taken in the two-fiber spectral mode, and the plotted spectrum is that of the star. The spectrum was not flux calibrated, not normalized, not corrected for any instrumental shift in wavelength, and was not corrected for the heliocentric velocity. The spectrum extracted with OPERA was multiplied by a factor 1.4 to match the DRAGRACES extraction. This difference comes from the different wavelength sampling between the two extractions. The top panel show the whole spectrum as extracted. The bottom panels show a close-up on the  $H\alpha$  line (left) and telluric lines (right).

mode, the results are coming from the science fiber (usually pointed at the target). In the two-fiber spectral mode, there is one result for each fiber, where one is usually pointed at the target and the other is usually pointed at the sky during nighttime observations. The FWHM in nanometers (nm) of the selected lines as a function of wavelength is plotted in red (left panels) and the resolution power  $R = \Delta\lambda/\lambda$  is plotted in blue (right panels).

#### 5.2.1. DRAGRACES Spectral Resolution

The mean resolution achieved by DRAGRACES in the one-fiber spectral mode is just over  $R \sim 52.8 \pm 0.1k$ . This is below the advertised  $R \sim 60k$  for GRACES.<sup>13</sup> However, this should be expected as the GRACES specs were determined using the pipeline OPERA, which is optimizing the spectral resolution by resampling the resolution elements to a subpixel size. The same is not performed by DRAGRACES. In the two-fiber spectral mode, though, the achieved resolution power is  $R \sim 36.3 \pm 0.3k$ , barely missing the advertised  $R \sim 40k$  for GRACES. Since the resolution elements in the two-fiber spectral mode are bigger, the need to optimize the spectral resolution is smaller.

One may see that in the one-fiber spectral mode, even if the resolution remains stable all along the spectrum, it varies linearly within each order. The difference between the blue and the red end is as big as  $10k$ . This can be explained by a tilt that was left between the beam and the detector while optimizing

the position of the receiver unit in the instrument as a trade-off between optimizing throughput and minimizing vignetting. Consequently, the focus is not perfectly uniform along a given order.

#### 5.2.2. OPERA Spectral Resolution

In the one-fiber spectral mode, OPERA achieves an average resolution that is superior to that of DRAGRACES. The measured resolution is near  $R \sim 58k$ .

In the two-fiber spectral mode both OPERA and DRAGRACES deliver a very comparable resolution power ( $R \sim 36k$ ).

#### 5.3. Cosmetics

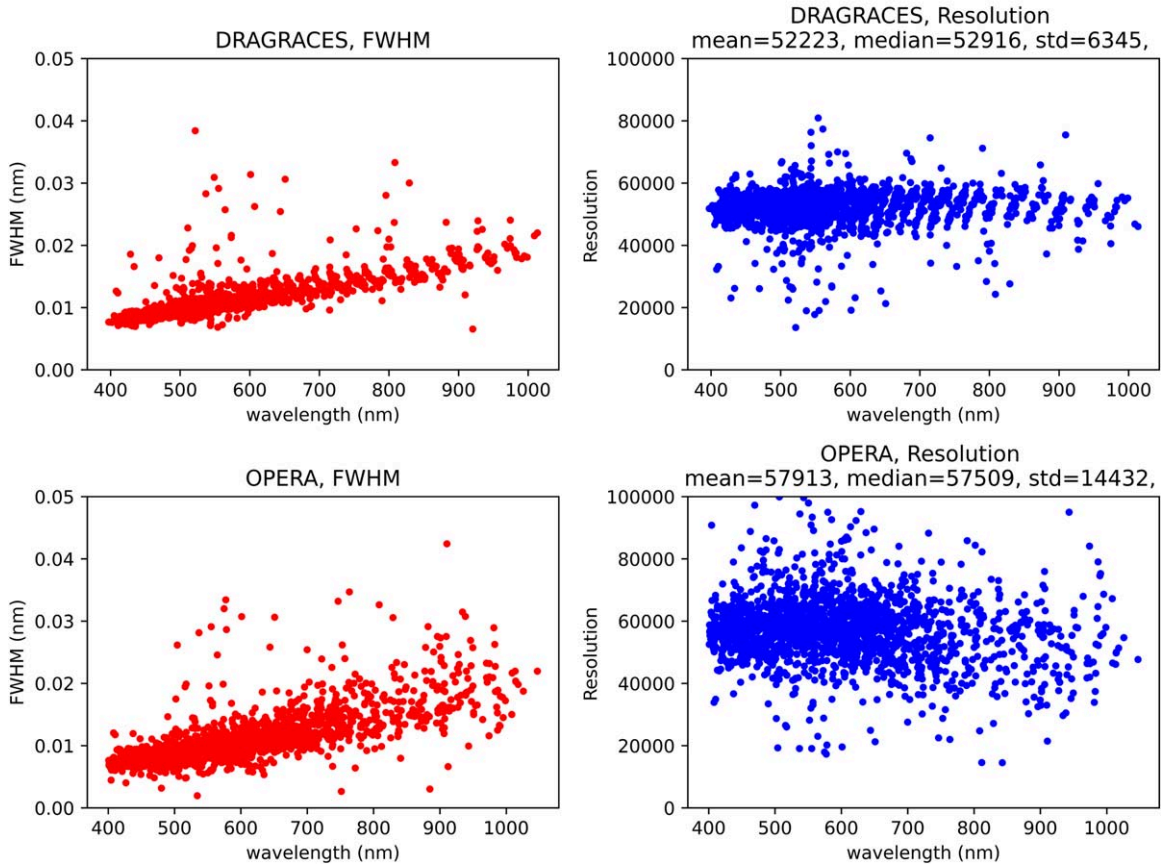
For the most part DRAGRACES and OPERA give very comparable, if not the same, results. However, depending on the case, one can be more affected by cosmetic issues than the other.

From users' feedback, we know that spectra extracted with OPERA can occasionally suffer from strong artificial patterns in the continuum that do not appear in the spectra extracted with DRAGRACES. Those can very likely be corrected by a user who knows how to tune the OPERA parameters, but this requires a fairly high level of expertise with the pipeline. However, extraction from DRAGRACES is very basic, and can sometimes fall short with some users' expectations.

One can note the installation of OPERA used at the Gemini Observatory alters the profile of strong and narrow emission lines,

<sup>13</sup> See GRACES webpage: [www.gemini.edu/GRACES](http://www.gemini.edu/GRACES).

obs info: year=2015, num=71



**Figure 11.** FWHM (red, left) and resolution power  $R$  (blue, right) of the emission lines selected in a ThAr spectrum extracted using DRAGRACES (top) and OPERA (bottom). The ThAr spectrum was observed on 2015 August 7 in the one-fiber spectral mode.

like ThAr or nebular lines. Indeed, while narrow lines with amplitudes lower than  $6000 \text{ e}^-$  (flux as extracted from the chip, not normalized by the exposure time) have the same profile in the OPERA and GRAGRACES extraction, the OPERA extraction of stronger lines gives profiles that diverge from the expected one. Those profiles have two components, i.e., an artificially narrow peak ( $\text{FWHM} \sim 1$  pixel) and a low amplitude plateau (see Figure 13). The most likely cause for this profile alteration is the choice of parameters used for the extraction. Yet, the fine tuning of the OPERA parameters is beyond the scope for this paper. This profile issue is at the origin of the bigger scatter in the FWHM and  $R$  values found in the spectrum extracted using OPERA compared to using DRAGRACES (see Figures 11 and 12).

We cannot determine thoroughly in which cases one pipeline is more appropriate than the other. This would require extensive testing with a wide variety of data types, which was not performed with either of the two pipelines. This is why it is highly recommended to always compare the two extractions before choosing which pipeline to use.

## 6. Conclusions

DRAGRACES is a pipeline written in IDL to reduce data from the spectrograph GRACES. Its strengths are that (1) it is easy to download and use, (2) it requires very low effort to get extracted spectra, as everything is automated, including making file lists, (3) it takes only a few minutes to extract a spectrum,

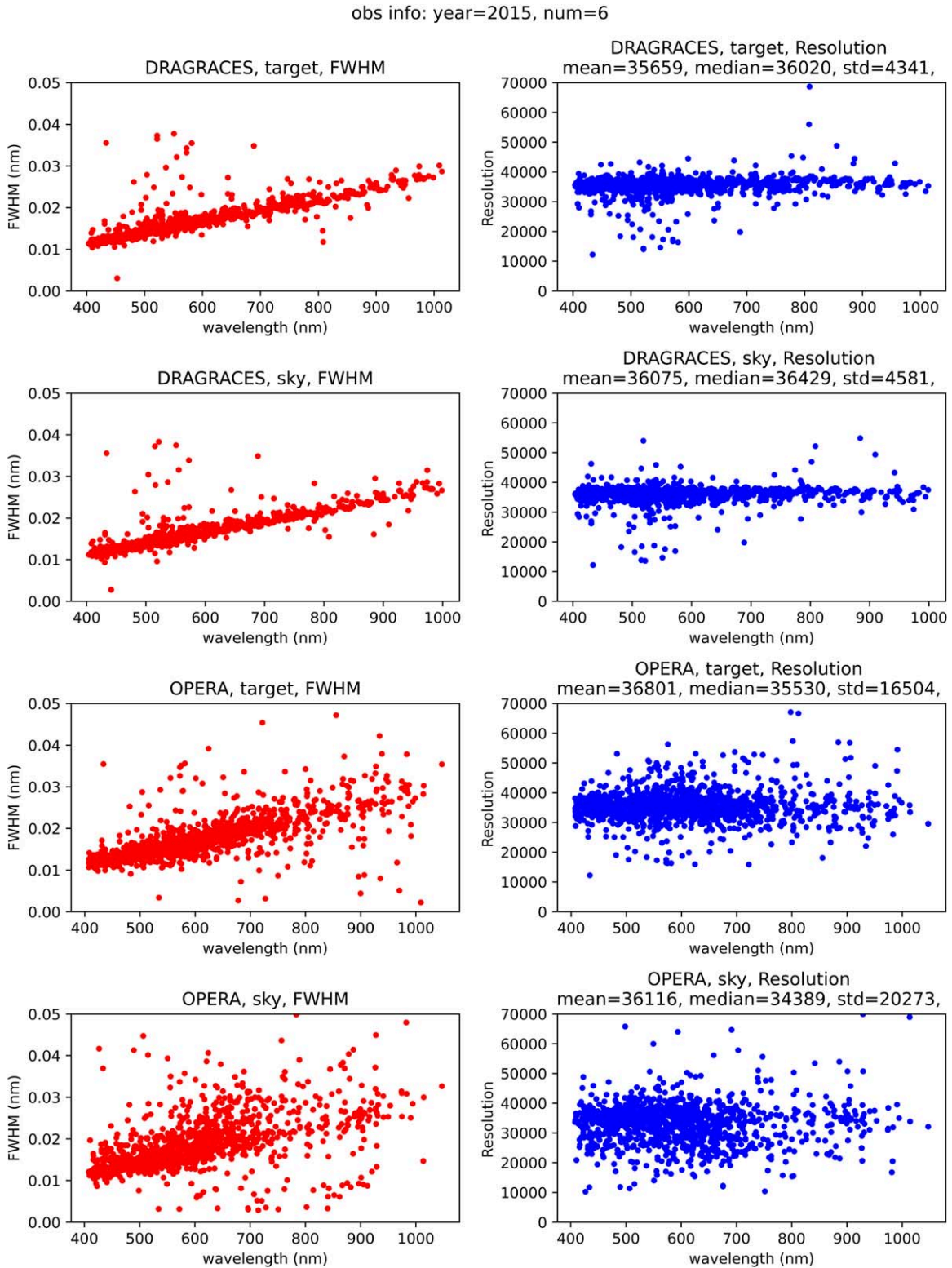
(4) the options are straightforward, and (5) users have the freedom to modify the code to add additional capabilities. However, its weaknesses are that (1) the reduction and extraction are basic, with no optimization, no cosmic-ray rejection, and no weighting applied; (2) there are many steps involving interpolation, which cause the errors to be further correlated; (3) no error array is provided (at present); (4) flux calibration, normalization, and/or wavelength solution corrections have to be performed outside of DRAGRACES; and (5) one needs an IDL license to run it. On the other hand:

1. the reduction and extraction are basic, with no optimization, no cosmic-ray rejection, no weighting applied
2. there are many steps involving interpolation, which causing errors to be further correlated
3. there are no error array provided (for the moment)
4. flux calibration, normalization and/or wavelength solution corrections have to be performed outside of DRAGRACES
5. one needs an IDL license to run it.

Future development to improve the noted weaknesses may be planned, assuming GRACES will still be in use for a long period of time.

Neither OPERA or DRAGRACES underwent any sort of thorough testing. That can be a concern as the quality of the results is not well determined. However, considering that both pipelines have been developed completely independently, that

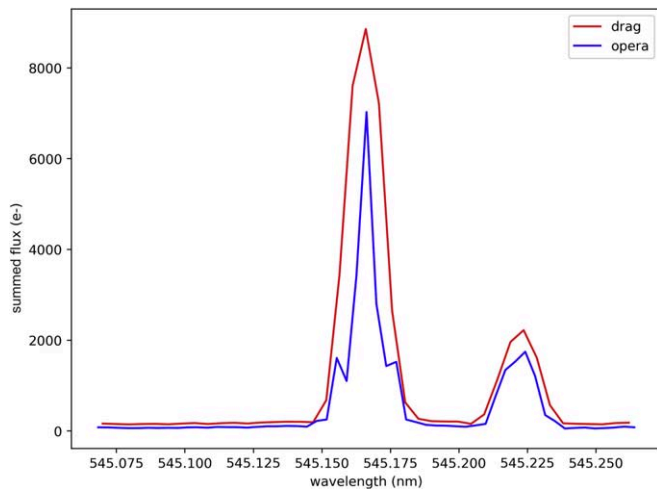




**Figure 12.** Same as 11. The ThAr spectrum was observed on 2015 June 6 in the two-fiber spectral mode. Both the results for the target and the sky fibers are plotted.

they are based on different principles, and that they produce very similar results is in many ways very reassuring. Each of those two pipelines have their advantages and disadvantages, and one should check both before choosing which one they will

base their research on. Yet, regardless of that choice, we are now confident that the extracted spectra are sensible, and the science results are not going to be significantly affected by the way the data were reduced.



**Figure 13.** Examples of ThAr line profiles obtained using DRAGRACES (red) and OPERA (blue) illustrating how the OPERA extraction provided by Gemini alters strong, narrow lines. For lines with moderate amplitudes, the profiles are identical. For stronger lines, the profile changes from nearly Gaussian to a sharp peak on top of two low shoulders. It is probably caused by parameters in the OPERA make file that need to be adjusted in the installation used at the Gemini Observatory.

This work was supported by the international Gemini Observatory, a program of NSF’s NOIRLab, which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National

Science Foundation, on behalf of the Gemini partnership of Argentina, Brazil, Canada, Chile, the Republic of Korea, and the United States of America. J.L.C. acknowledges support from NSF grant AST-1816196.

*Facility:* Gemini.

*Software:* astrolib (Landsman 1993), imdisp (written by Liam E. Gumley).

### ORCID iDs

Andr -Nicolas Chen   <https://orcid.org/0000-0002-1115-6559>

Michael Lundquist  <https://orcid.org/0000-0001-9589-3793>

Eder Martioli  <https://orcid.org/0000-0002-5084-168X>

Jeffrey L. Carlin  <https://orcid.org/0000-0002-3936-9628>

### References

- Anthony, A., Hill, A., Pazder, J., et al. 2012, *Proc. SPIE*, **8446**, 84464
- Carlin, J. L., Sheffield, A. A., Cunha, K., & Smith, V. V. 2018, *ApJ*, **859L**, 10
- Chen , A.-N., Padzer, J., Barrick, G., et al. 2014, *Proc. SPIE*, **9151**, 915147
- Horne, K. 1986, *PASP*, **98**, 609
- Landsman, W. B. 1993, in ASP Conf. Ser. 52, *Astronomical Data Analysis Software and Systems II*, ed. R. J. Hanisch, R. J. V. Brissenden, J. Barnes et al. (San Francisco, CA: ASP), **246**
- Lovis, C., & Pepe, F. 2007, *yCat*, *J/A+A*, **468/1115**, 0
- Manset, N., & Donati, J.-F. 2003, *Proc. SPIE*, **4843**, 425
- Martioli, E., Teeple, D., Manset, N., et al. 2012, *Proc. SPIE*, **8451**, 84512
- Palmer, B. A., & Engleman, R., Jr. 1983, *Atlas of the Thorium Spectrum* (Los Alamos, NM: Los Alamos National Laboratory)