

Data Reduction Software of the X-shooter Spectrograph

P. Goldoni^a, F. Royer^b, P. François^b, M. Horrobin^c, G. Blanc^a,
J. Vernet^d, A. Modigliani^d, J. Larsen^d

^aAPC/UMR 7164, 11 Place Marcelin Berthelot, Paris, France;

^bGEPI/UMR 8111, Observatoire de Paris-Meudon, 5 Place Jules Janssen, Meudon, France;

^cUniversity of Amsterdam, Kruislaan 403, Amsterdam, the Netherlands;

^dESO, Karl Schwarzschild Str. 2, Garching bei München, Germany

ABSTRACT

We present the Data Reduction Software (DRS) being developed at APC, Paris Observatory, Amsterdam University and ESO for the X-shooter échelle spectrograph. X-shooter is the first VLT second generation instrument, expected to be operational in 2008. The DRS will be fully integrated in the ESO VLT system and it will use the ESO Common Pipeline Library. We discuss the data reduction related to slit and IFU observations. X-shooter data have two main characteristics, on the one hand the exceptionally wide band (0.3-2.4 μm) covered in a single exposure, and on the other hand the spectral format with highly curved orders and tilted lines. The reduction process is described and the critical issues related to the above characteristics, notably the sky subtraction, the optimal extraction, and the construction of 1D/2D/3D output products, are addressed. Some aspects of the spectrophotometric calibration are also discussed.

Keywords: Spectroscopy, Data Reduction, Pipeline

1. INTRODUCTION

1.1. Framework of the DRS project

X-shooter¹ is a single object spectrograph which will be installed at the Cassegrain focus of one of the UTs of the VLT in 2008. It will be the first second generation instrument to be installed at VLT. The consortium that builds X-shooter comprises the Copenhagen University Observatory (DK), INAF (Istituto Nazionale d'Astrofisica) (IT), the GEPI-Observatoire de Paris-Meudon and the Laboratoire d'Astroparticule et Cosmologie (APC)(FR), ASTRON and the Universities of Amsterdam and Nijmegen (NL) and ESO. The aggressive development schedule and the contemporary development of other second generation instruments made a large external hardware and financial contribution necessary.

The Data Reduction Software (DRS) is developed by APC, GEPI and the University of Amsterdam. The DRS is designed to be a collection of data reduction recipes based on the ESO Common Pipeline Library (CPL) operating in the VLT DFS environment. Therefore the usual Data Reduction tools of the VLT environment such as the Data Organizer, the Reduction Block Scheduler, EsoRex and Gasgano will be used in conjunction with the X-shooter pipeline. In this scheme, the consortium will be able to develop the recipes in a modular way according to a well established template and to focus on instrument specific algorithms. Basic operations of input/output and image manipulations will be performed using CPL subroutines.

Further author information: (Send correspondence to P. Goldoni)

P. Goldoni: E-mail: paolo.goldoni@apc.univ-paris7.fr, Telephone: +33-1 69 08 55 25

1.2. Instrument overview

The X-shooter is a high-efficiency spectrograph with a spectral resolution in the range 4000-10000, depending on wavelength and on slit width. X-shooter consists of a central structure (the backbone) which supports three prism-cross-dispersed échelle spectrographs optimized for the UV-Blue, Visible and Near-IR wavelength ranges respectively. The backbone contains the calibration and acquisition units, an Integral Field Unit (IFU) which can be inserted in the light path and reformats a field of $1.8'' \times 4''$ into a long slit of $0'.6 \times 12''$, two dichroics to split the beam among the three arms and relay optics to feed the entrance slits of the spectrographs. The spectral format is fixed.

- X-shooter has *three different arms*, covering different spectral ranges:
 - UVB: starting at 300 nm. The upper limit is determined by the first dichroic which reflects the UVB light. The crossing wavelength is 558 nm.
 - VIS: starting at 558 nm, the high pass wavelength cut being set by the first dichroic. The upper limit is determined by the second dichroic which reflects the Visual-Red light. The crossing wavelength is 1020 nm.
 - NIR: starting at 1020 nm, the high pass wavelength cut being set by the second dichroic. The upper limit is determined by the cut off of the detector and the atmospheric absorption.
- X-shooter has *two observing configurations*:
 - SLIT: in this case the atmospheric dispersion is compensated by ADCs in UVB and VIS arms. In NIR, the effect is not significant for $Z_D < 60^\circ$.
 - IFU: in the telescope focal plane the IFU pattern covers a field of $1.8'' \times 4''$ and reformats it to a virtual slit image composed by three aligned slices of $0'.6$ width, $4''$ each long. All other components are as in the SLIT case. The IFU is located before the ADCs in the light path. Flux will then be lost at the entrance of the IFU in VIS, and mostly in UVB, before the light pass through the ADC.

Slits of different sizes can be inserted in the light path. The possibility to use several slits with width ranging from $0'.3$ to $1'.5$ is currently foreseen.

1.3. Specific Features

The wide spectral band covered by X-shooter implies that different observing techniques will be used depending on the scientific objective of the observation and on the source intensity along the spectral band.

The key issue, as far as the NIR is concerned, is the background subtraction. The infrared sky, at wavelengths shorter than $2\mu\text{m}$, is dominated by the emission from the hydroxyl molecule OH. This emission is highly structured, spatially and spectrally and exhibits intensity variations of up to 10% on time scales of about 5 to 15 minutes.² At wavelengths longer than $2\mu\text{m}$, observations are generally strongly affected by thermal background radiation from the sky and the telescope, with strongly increasing background when going to longer wavelengths. Accurate sky subtraction is essential and this requires special attention and procedures. Therefore *three different observing modes* are available:

- staring mode: the science target is observed at a fixed position on the slit. The sky is estimated from the object frames themselves, on both sides of the object. This requires that the object is bright enough to be localized on the frame, and not too extended, to allow the estimation of the sky. This technique is the standard one for observations in the optical range. An accurate sky background subtraction in this mode is crucial to properly reduce data in the NIR where the sky contribution is predominant. This is described in Sect. 3.2.

- nodding mode: the classical technique in IR spectroscopy is to observe object(s) at two or more positions along the slit. These positions are usually called A and B and pairs of observations (AB or BA) are repeated into sequences. The B position is shifted with respect to A by an amount δ along the slit. The sky is removed at the first order by subtracting one frame from the other and thus obtaining two difference frames: A-B and B-A. The B-A frame is then shifted by δ and the resulting frame (B-A)' is summed to the (A-B) one. The extraction is then performed on the resulting frame (A-B) + (B-A)'. This process is sometimes called double pass sky subtraction
- on-off mode: when the object is extended (i.e. covering a large fraction of the entrance slit), the sky has to be sampled away from the object. This may result in a loss of efficiency for the observations, which can amount to 50% of the observing time if the sky has to be sampled as frequently as the object.

To each of the observing mode corresponds a dedicated science reduction recipe. The three arms will be reduced separately, using the same recipe, which will behave according to the arm. The sky emission lines present in the UVB and VIS spectral ranges can also be removed using standard nodding observation techniques. The spectra from the three arms will be merged into a single spectrum only for quick look purposes.

These three observing mode will be available in SLIT observing configuration while in IFU observing configuration only staring mode and on-off mode will be available because the IFU is too short to allow nodding in the cross-dispersion direction.

The highly curved and tilted orders require a special effort to perform wavelength calibration and optimal extraction, details are in Sect. 3.1 and 3.4.

Spectrophotometric calibration of such an instrument is a challenge due to the very wide spectral range and to the small number of suitable standard stars in NIR. We briefly describe our strategy in Sect. 3.5.

When the science target is a point source, and with a signal per pixel high enough to be localized and extracted, a 1D spectrum will be produced. In general however, the observing configurations of X-shooter, SLIT and IFU, allow the combination of spatial and spectral informations and the production of, respectively, 2D and 3D spectroscopic data. These 2D/3D products will be the default output of science reduction (see Sect. 3.6).

2. DRS OVERVIEW

The X-shooter DRS will be fully integrated in the VLT-DFS pipeline and will make use of libraries and reduction tools developed at ESO. More precisely, the DRS reduction recipes will be written in standard ANSI C using the ESO Common Pipeline Library. The CPL is presently used by ESO as the main tool for the development of data reduction pipelines for VLT instruments.

To support the development of X-shooter DRS by the consortium, ESO provides a set of functions and recipes converted from the existing UVES pipeline, so that from the definition of algorithms to the validation of the recipes it will be possible to rely on the UVES experience.

X-shooter calibration, and especially the wavelength calibration will depend on a physical model, i.e. a numerical model of the instrument based on the engineering data.³ An optimized comparison between the physical model and the calibration data will allow to obtain the instrument configuration with much more precision than with the usual methods (e.g. factor of five improvement in STIS⁴). The use of the physical model in wavelength calibration is detailed in Sect. 3.1.

2.1. Inputs

The DRS pipeline will receive in input the raw data to be processed and the associated calibration data:

- Raw images from the VLT instrument software
- The calibration data
Calibration data and instrument parameters are extracted from the calibration database. Two types of calibration data exist:

- master calibrations which are generated by the pipeline either by processing of daytime calibrations or attached calibrations (TBC), examples are flat fields and wavelengths solutions,
- ancillary data like standard star tables, reference line catalogs, atmospheric extinction tables, physical model parameters and other reference tables required for data reduction.

2.2. Outputs

The DRS pipeline will produce the following data:

- flux and wavelength calibrated 2D (and 1D when relevant) spectra with associated errors,
- three-dimensional data cubes containing the position and the wavelength information of the observation (in case of IFU observation),
- QC parameters and performance monitoring values. The QC parameters are divided in QC0 parameters for processing control and QC1 parameters to check of the instrument performances with respect to the nominal ones.

Throughout all the reduction, each image will be associated with an error image and a bad pixel map. The error images and bad pixel maps will be propagated through each step. Likewise for the extracted spectra, so that the final result will be associated to a realistic estimation of the errors. The final outputs will be produced for each arm of X-shooter. A single merged spectrum from 0.3 to 2.4 μm will be created to allow some QC and a quick look of the spectrum.

2.3. Data Reduction Scheme

2.3.1. Calibration

The X-shooter recipes for computation of master bias, dark and flat-field frames, and order definition, will be identical to the ones used in the UVES pipeline.

Wavelength map and IFU geometry: these recipes are very specific to X-shooter and are further described in Sect. 3.

- the wavelength map is obtained with a multi-pinhole arc lamp exposure,
- a full slit arc lamp exposure allows to adjust the wavelength map and to correct from shifts along the dispersion relation,
- the shifts between the slitlets of the IFU pattern are determined using the observation of a star at different positions in the $1''.8 \times 4''$ field of the IFU on the sky.

Telluric and spectrophotometric standard stars will also be observed to mask the position of telluric absorption lines and calibrate in flux.

2.3.2. Science reduction

The reduction of science frames will strongly depend on the observing mode: staring, nod-and-jitter, on-off. The flowcharts in Fig. 1 respectively stand for the reduction scheme in staring mode (a) and on-off mode (b).

- Bias and dark subtraction only occur in staring mode. In nod-and-jitter (respectively on-off) bias and dark contributions are removed when on and off frames (respectively A and B frames) are subtracted one from the other.
- The inter-order background is fitted and subtracted.
- The frame is divided by the flat-field.

- The object is localized on the slit. The localization is needed to perform the sky subtraction, the CRH detection and the optimal extraction.
- The sky subtraction in the staring mode is carried out following the method described in Sect. 3.2. In the other observational modes, it is carried out at the beginning of the reduction by subtracting the different input frames.
- The cosmic ray hits (CRH) removal can be done in different ways: either by combining different similar exposures, or by detecting them in single exposures (Sect. 3.3).
- From here, two different processings lead to 1D or 2D/3D products.

3. DETAILED DESCRIPTION

3.1. Wavelength Calibration

In the case of X-shooter, the shape of the orders is strongly curved with a parabolic shape, and the tilt of the spectral lines is large (see Table 1), therefore the wavelength cannot be expressed only as a function of position in the dispersion direction, but depends also on the cross-dispersion position. The position (X, Y) on the detector as a function of the wavelength, order number and position on the slit (λ, n, s) will be mapped using a multi-pinhole calibration exposure (see Fig. 2). The multi-pinhole frame is illuminated by an arc lamp masking the entrance slit. There will be seven $0''.5$ holes evenly spaced along the $12''$ slit.

The mapping calibration recipe will use the physical model of the instrument to provide a guess solution of the line positions on the calibration frame. Each of these spectral lines will be fitted by 2-dimensional Gaussians and the resulting positions (X, Y) will be adjusted by a global polynomial fit on the whole detector.

This bi-dimensional dispersion relation (spatial and spectral directions) will be used to get the wavelength of specified pixel positions (for sky subtraction and optimal extraction), and to correct for optical distortions and resample the frames into a (λ, s) grid (see Sec. 3.6 and Fig. 6 in the case of the IFU).

Table 1. Average line tilt and height of échelle orders from the X-shooter optical design, for the three spectral arms of the instrument.

Arm	Height of central order	Average tilt
UVB	130 pix	-6.4 pix
VIS	150 pix	-10.9 pix
NIR	200 pix	-5.8 pix

3.2. Sky subtraction in staring mode

Like any other instrument which operates in the red and near-IR spectral range, correcting for sky background emission is a crucial step in the reduction process. In the case of X-shooter it is complicated by the following:

- The curvature of the orders on the detector.
- The change of plate scale in the direction of the slit.
- Unknown position and shape of the objects in the slit.

The change of plate scale is the unique problem for X-shooter sky subtraction. The change is small, at worst 0.8%. But, as the photon noise on a bright sky emission line can be less than 0.3%, not correcting for the change will lead to systematic errors.

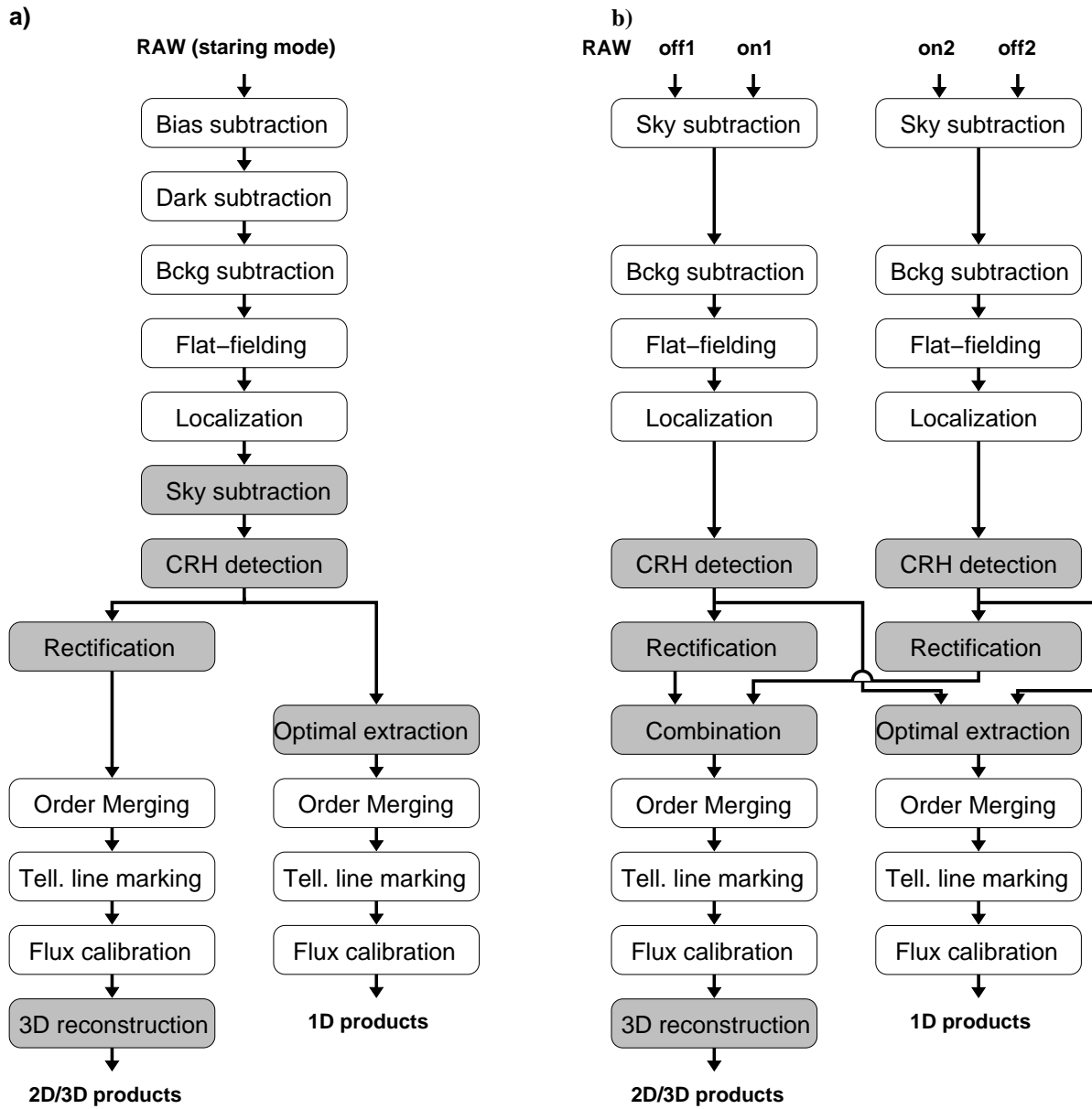


Figure 1. Flowcharts of the reduction of science frames. The reduction steps in gray boxes are further detailed in Sect. 3. **a)** Staring mode; **b)** On-off mode.

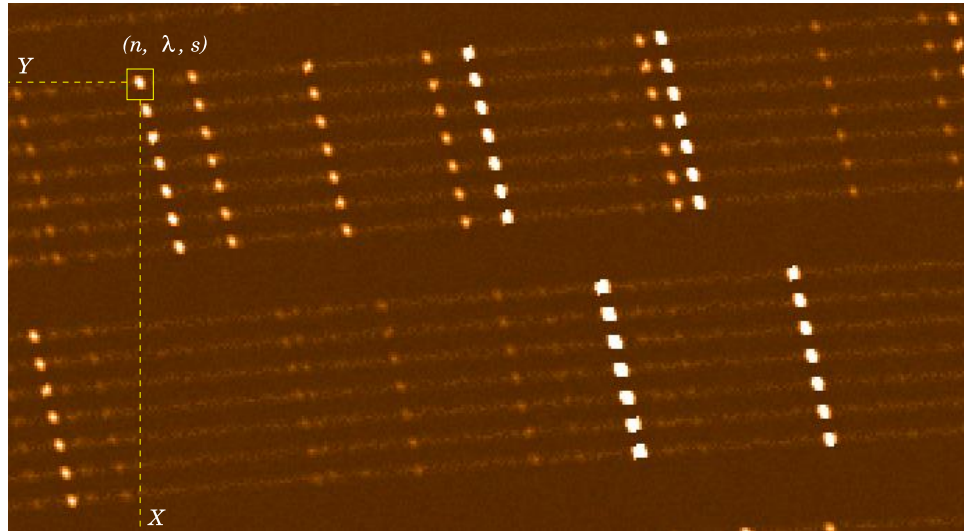


Figure 2. Example of a raw multi-pinhole calibration frame (simulated for the VIS arm in X-shooter). A zoomed part of the detector is shown. For a given line (in the box), with known wavelength (λ), order number (n) and position on the slit (s), the position on the detector (X, Y) is fitted using a 2-dimensional Gaussian.

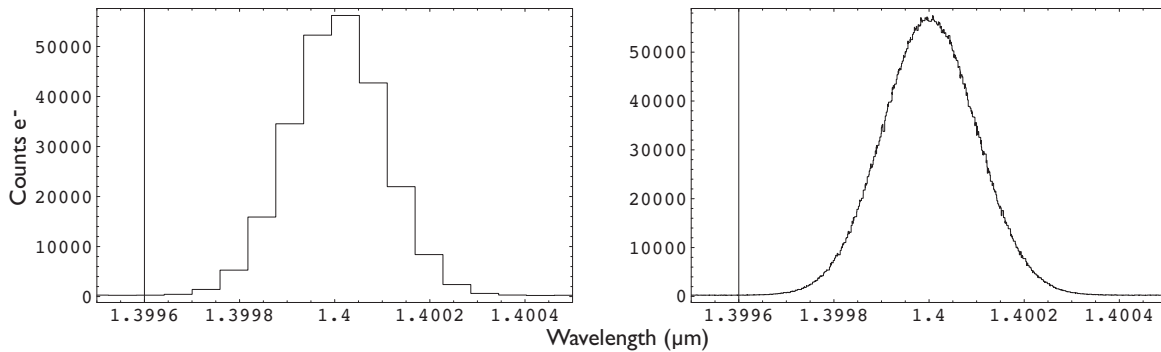


Figure 3. A comparison of a single row of data (left) and a sampling of all data rows (right).

3.2.1. Super sampled Bézier fitting.

Super sampling is a recently developed technique for fitting sky background (see the SDSS spectral pipeline or Kelson⁵). Its main advantage is that it can be applied to the raw detector format, no resampling is needed, so it is immune to interpolation effects such as change in PSF or noise correlation.

The technique relies on the fact that the spectral lines are tilted, so each pixel samples a slightly different part of wavelength space. An accurate wavelength map is used to re-bin the raw 2D data into a 1D data set with the wavelength of each pixel as the x-axis and the pixel value in the y-axis. (see fig 3).

This super-sampled data is then fitted with a generic function. Normally a continuous Bézier spline is used, as this requires no a priori knowledge of the line function. The difference between the fit and the data is then compared to the variance of the data, and outlying points are rejected, the process then repeats iteratively. As we are making a fit in wavelength space, the rejection process will remove both cosmic ray hits and object flux from the sky fit (although it is not possible to distinguish their contributions). The sky fit is then calculated for, and subtracted from, the original 2D frame. Therefore the final output frame consists of: object flux, cosmic rays, the difference noise of the sky flux and the errors from the flat field. Fig 4 shows this fitting method applied to simulated X-shooter data.

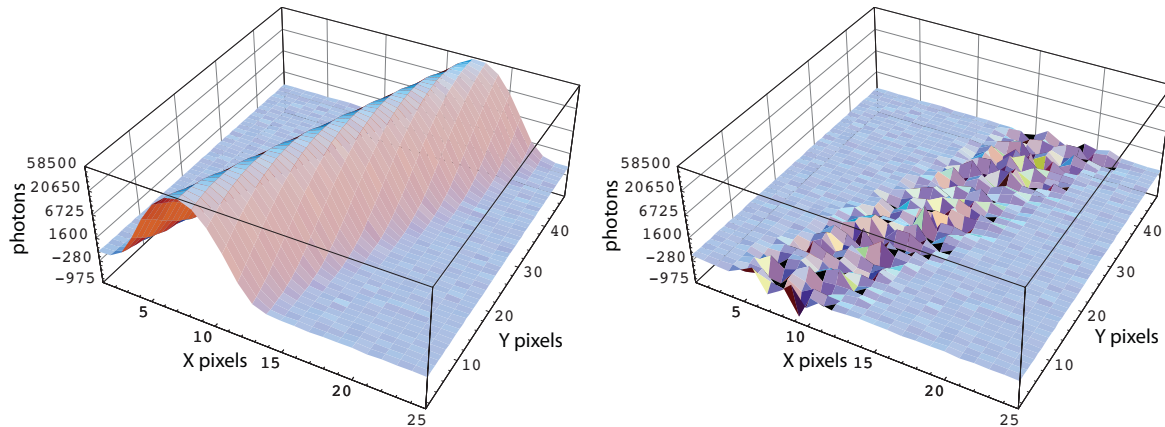


Figure 4. The original input data (left) and the residuals after the sky model is subtracted (right). Note that the photon axis has a logarithmic scale.

In order to cope with the change of plate scale we will have to make a slight change to the normal super-sampling method. If we consider a classical Gaussian line shape, the change of plate scale can be considered as a change in its σ with the distance along the slit, s . If we take the Logarithm of the Gaussian and multiply by σ^2 we can separate out the terms in λ and s , see equation 1. This $\text{Log} \times \sigma^2$ space allows us to apply super-sampling methods to the data and simultaneously account for the change of plate scale.

$$\sigma^2 \times \text{Log}(A \times e^{-\frac{(\lambda-\mu)^2}{2\sigma(s)^2}}) = -\frac{(\lambda-\mu)^2}{2} + \text{Log}A \times \sigma(s)^2 \quad (1)$$

3.3. Cosmic rays subtraction for single images

It is foreseen that in certain cases (e.g. staring exposures of faint objects, UVB-driven science observations) long single exposures will be taken. Moreover in case of seeing variation between the exposures or of image shifting by a non integer number of pixels, CR rejection on single images may be required. For this purpose we plan to use the van Dokkum⁶ method, based on Laplacian edge detection. Its advantage is that it is very efficient and adapted to undersampled data which is important for the infrared arm of X-shooter which suffers from mild undersampling.

For spectroscopic images it requires a preliminary subtraction of sky lines whose strong gradients are identified by the algorithm as cosmic rays, therefore in the flowcharts in Fig. 1 it has been placed after the sky subtraction. In a recent paper⁷ it has been identified as the most efficient ($\sim 85-90\%$ of cosmic rays hit identified) between several different methods.

On the other hand it is computationally intensive, which may be a problem during routine analysis. We performed some preliminary tests on a Dell Linux PC (PIV 2.8 GHz, 1 GB RAM), using an IDL routine made available by P. van Dokkum on his web page. On a $2k \times 2k$ simulated image of X-shooter VIS arm, one iteration takes about 70s, so for the 3 to 4 iteration typically needed, the total computing time would be 210-280s. We estimate that a factor at least 3 should be gained from optimization of the code and use of faster machines resulting in a total estimated computing time (for two $2k \times 2k$ CCDs and one $1k \times 1k$ detector) of $\leq 160-200$ s. Therefore this method will be feasible for long exposures.

3.4. Optimal Extraction

Optimal extraction⁸ is a method to extract cross-dispersed spectra delivering the maximum possible signal-to-noise ratio while simultaneously preserving a good spectrophotometric accuracy. This is accomplished by adding nonuniform weights to the pixels in the extraction sum. The original method,⁸ however, is only suited for

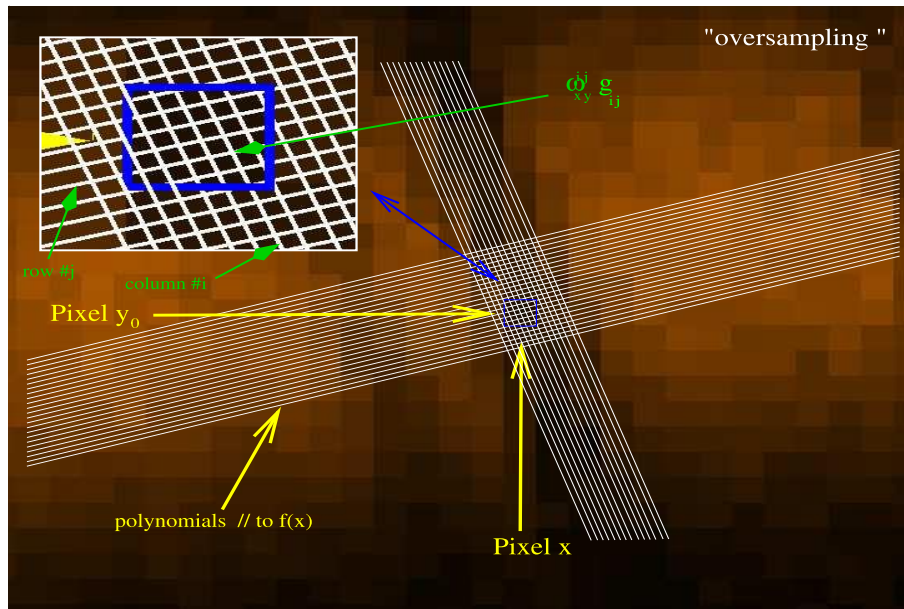


Figure 5. Description of the oversampling method. The inclined horizontal lines represent the order (the spectrum) while the inclined vertical lines represent the line (the slit). In the upper left corner we show a zoom over the pixel (x, y_0) which is subsampled in several subpixels. The grid inside the detector pixel (x, y_0) defines the locations where the weights ω_{xy}^{ij} have to be computed for that pixel.

straight orders and lines and therefore it is not applicable to X-shooter. The optimal extraction function used for X-shooter will be based on the mathematical method described by Piskunov & Valenti.⁹

The main idea of the Piskunov & Valenti method is to obtain a precise determination of the cross-order and spatial profile shapes by subsampling them using several subpixels (virtual resampling). This technique allows to improve the knowledge of the cross-order and spatial profiles (see Fig. 5) and therefore allows to describe the spectral shape with greater precision.

This method has been successfully applied to the pipeline of ESPaDOnS, a spectrograph installed at the Canada France Hawaii Telescope. (see <http://www.ast.obs-mip.fr/projets/espadons/espadons.html> for a description of the project). The spectral format of the EsPaDons frames is similar to the X-shooter one (i.e. curved orders with tilted lines).

With the knowledge of this spatial and spectral decomposition, we can therefore apply the method of optimal extraction.

3.5. Spectrophotometric Calibration

A reasonably accurate spectrophotometric capability is one of X-shooter requirements, the required spectrophotometric accuracy is $\leq 10\%$ for zenithal distances up to 50° . While this constraint may not appear to be very strong, two main issues make this requirement difficult:

- Spectrophotometric calibration in NIR is difficult due to the lack of suitable standard stars
- The very large waveband makes very difficult to calibrate with a single standard star

To address the first problem, the X-shooter consortium will organize observing campaigns to observe standard stars with the ISAAC/SOFI spectrograph at ESO in order to produce a catalog adapted for X-shooter. The catalog should be ready for X-shooter commissioning.

The flux ratio between $0.3 \mu\text{m}$ and $2.4 \mu\text{m}$ is close to 3000 for a 20,000 K blackbody. This flux ratio will cause saturation in the UVB arm and an insufficient signal to noise in the NIR arm. To overcome this problem, the

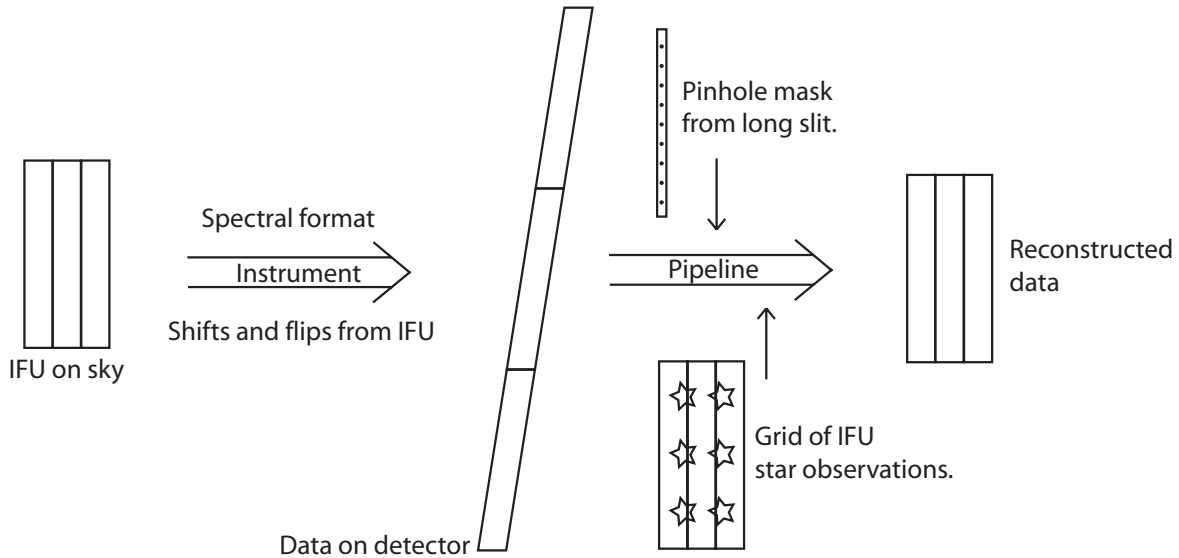


Figure 6. A schematic diagram of the image reconstruction process for the IFU.

observation of spectrophotometric standard stars will be split in two. Hot stars will be observed with the UVB *and* VIS arm while cool stars will be observed with the VIS *and* NIR arm. The UVB-VIS observations will be performed in staring mode while the VIS-NIR observations will be performed in nodding mode. By using this strategy a complete flux calibration will be obtained across the whole wavelength range. The observing time in excess needed is minimal however a caveat of this procedure is that two standard stars have to be observable during the same night.

3.6. Image Reconstruction

IFU image reconstruction is always a non-trivial process, this is particularly true for X-shooter as it has no built in system for calibrating the spatial coordinates of the IFU.

In order to align the IFU with the instrument coordinate system, we make use of the fact that the central slice is actually a “straight through” position, i.e. it contains no extra optics when compared to the normal slits. This section of the IFU is calibrated using a pinhole mask, which can be placed in the same focal plane as the IFU unit. The two outer slices are then calibrated relative to the central slice on sky, a star is observed over a grid of positions on the IFU and the offsets between the slices are calculated. Mechanical simulations of the IFU unit indicate that the flexures and distortions in the mirror systems of the 2 outer slices will be so small that they will not be detectable.

Using this calibration of detector pixels to sky coordinates, the 3 slices are then independently rectified to a common grid. This rectification is only performed in the direction along the slices, as they are assumed to be perfectly aligned across the IFU. The rectification process uses standard interpolation kernels. Conversion from the X, Y plane to the s, λ plane of the final data is performed in a single step. See figure 6 for a schematic diagram of this process.

The X-shooter IFU does not allow for the use of atmospheric dispersion correctors, so all atmospheric dispersion will be visible in the final data. This means that there will be data at the edges of the slices that does not cover the whole wavelength range, the default behaviour of the pipeline will be to remove this data from the final data set.

3.7. Project Status

The design of the X-shooter DRS has been recently finalized, the coding in C of the simplest functions has already begun. As explained in the introduction, we are already making use, wherever possible of CPL functions.

Concerning the more complex algorithms described in Section 3, we will test them using appropriate tools such as IDL or Mathematica. To test them we will use simple simulated images and, starting mid-2006, the images produced by the physical model. Currently the first working scientific analysis pipeline is expected to be ready at the end of 2007 for the Preliminary Acceptance in Europe. The DRS project should be completed at end of 2008.

ACKNOWLEDGMENTS

PG, FR, PF and GB thank the Programme Astroparticule for its support.

REFERENCES

1. S. D'Odorico et al., "X-shooter, a UV to K band, intermediate resolution VLT spectrograph," *SPIE Proceedings* **this volume**, 2006.
2. S. K. Ramsay, C. M. Mountain, and T. R. Geballe, "Non-thermal emission in the atmosphere above Mauna Kea," *MNRAS* **259**, p. 751, 1992.
3. P. Ballester and M. Rosa, "Modeling echelle spectrograph," *APJS* **126**, p. 563, 1997.
4. P. Kerber, F. Bristow and M. Rosa, "STIS calibration enhancement (STIS-CE): Dispersion solutions based on a physical instrument model," *Proceedings of 2005 HST Calibration Workshop*, 2005.
5. D. D. Kelson, "Optimal techniques in two-dimensional spectroscopy: Background subtraction for the 21st century," *PASP* **115**, p. 688, June 2003.
6. P. G. van Dokkum, "Cosmic-ray rejection by laplacian edge detection," *PASP* **113**, p. 1420, Nov. 2001.
7. C. L. Farage and K. A. Pimblet, "Evaluation of cosmic ray rejection algorithms on single-shot exposures," *PASA* **22**, p. 24, 2005.
8. K. Horne, "An optimal extraction algorithm for ccd spectroscopy," *PASP* **98**, p. 609, June 1986.
9. N. E. Piskunov and J. A. Valenti, "New algorithms for reducing cross-dispersed échelle spectra," *A&A* **385**, p. 1095, Apr. 2002.