EUROPEAN SOUTHERN OBSERVATORY

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral

Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VERY LARGE TELESCOPE

HARPS Pipeline User Manual

ESO-331895

Issue 3.3.12

Date 2025-07-02

Prepared:	A. Modigliani, D. Sosnowska, C. Lovis	2025-07-02	
	Name	Date	Signature
Approved:	P.Ballester		
	Name	Date	Signature
Released:	M. Peron		
	Name	Date	Signature

This page was intentionally left blank

Change record

Issue/Rev.	Date	Section/Parag. affected	Reason/Initiation/Documents/Remarks
3.0.0	31/01/2023	all	First public release,pipeline version 3.0.0
3.2.0	15/05/2024	-	Coordinated public release update
3.3.0	22/11/2024	-	public release update
3.3.6	28/02/2025	-	telluric correction, update
3.3.12	02/07/2025	-	public release update

This page was intentionally left blank

Contents

1 Introduction

1.1 Purpose

The ESPRESSO pipeline adapted for HARPS is a subsystem of the *VLT Data Flow System* (DFS). Through the rest of the document it will be called HARPS pipeline. Its target user is both the user community and ESO *Data Products Department* (DPD) in the generation of master calibration data, in the reduction of scientific exposures, and in the data quality control. It should also serve as a quick look tool for *Paranal Science Operations* (PSO). Additionally, the HARPS pipeline recipes are made public to the user community, to allow a more personalised processing of the data from the instrument. The HARPS pipeline was developed with the declared objective of providing precise Radial Velocity (RV) data and accurate wavelength calibration. The purpose of this document is to describe a typical HARPS data reduction sequence with the HARPS pipeline.

This manual is a complete description of the data reduction recipes implemented by the the HARPS pipeline, reflecting the status of the HARPS pipeline as of Jul 02 2025 (version 3.3.12).

1.2 Acknowledgments

The HARPS pipeline has been designed, implemented and developed by the Geneva Observatory. We are particularly grateful to the responsibles for the data reduction: Christophe Lovis, Danuta Sownowska and Alex Segovia for their contributions and support.

Andrea Modigliani, from ESO, provided the Reflex workflow, DFS support and most of the documentation. We thank also Pedro Figueira, Andrea Mehner and Richard Anderson for providing useful feedback to improve the reflex workflow and documentation.

1.3 Scope

This document describes the HARPS pipeline used at ESO-Garching and ESO-Paranal for the purpose of data assessment and data quality control.

The examples on running individual pipeline recipes in this manual use the esorex command and manually created list of input files. Several interfaces to automatically organise the data, create the list of input files and execute the pipeline recipes in the proper sequence are available, see the ESO pipeline page for details.

Please note that the use of Gasgano as a GUI for processing data is deprecated. Its use is no longer recommended and the related section in this manual, as well as support for Gasgano as a data processing GUI application in general will be dropped entirely in a future release.

Updated versions of the present document may be found on [1]. For general information about the current instrument pipelines status we remind the user of [2]. Quality control information are at [3].

Additional information on the Common Pipeline Library (CPL) and ESOREX can be found respectively at [4], [5]. A description of the instrument is in [6]. The HARPS instrument user manual is in [7]. The HARPS Reflex tutorial is in [8].

1.4 Reference documents

[1] HARPS Pipeline Users' Manual   VLT-MAN-ESO-19500-XXXX
                              www.eso.org/pipelines
[2] Current pipeline status
www.eso.org/observing/dfo/quality/pipeline-status.html
[3] ESO-Data Flow Operation home page   http://www.eso.org/observing/dfo/quality/
[4] CPL home page   www.eso.org/cpl
[5] ESOREX home page   www.eso.org/cpl/esorex.html
[6] HARPS home page
www.eso.org/sci/facilities/lasilla/instruments/harps
[7] VLT HARPS User Manual   VLT-MAN-ESO-14700-3517
www.eso.org/sci/facilities/paranal/instruments/espresso/doc.html
[8] Reflex HARPS Tutorial   www.eso.org/pipelines

2 Overview

In collaboration with instrument consortia, the Pipeline Systems Department (PSD) of the Software Developement Division is implementing data reduction pipelines for the most commonly used VLT/VLTI instrument modes. These data reduction pipelines have three main purposes:

Data quality control: pipelines are used to produce the quantitative information necessary to monitor instrument performance.

Master calibration product creation: pipelines are used to produce master calibration products (*e.g.*, combined bias frames, super-flats, wavelength dispersion solutions).

Science product creation: using pipeline-generated master calibration products, science products are produced for the supported instrument modes (*e.g.*, combined ISAAC jitter stacks; bias-corrected, flat-fielded FORS images, wavelength-calibrated UVES spectra). The accuracy of the science products is limited by the quality of the available master calibration products and by the algorithmic implementation of the pipelines themselves. In particular, adopted automatic reduction strategies may not be suitable or optimal for all scientific goals.

Instrument pipelines consist of a set of data processing modules that can be called from the command line, from the automatic data management tools available on Paranal or from Gasgano.

ESO offers two front-end applications for launching pipeline recipes, *Gasgano* and *EsoRex* [5] both included in the pipeline distribution (see Appendix A, page 61). These applications can also be downloaded separately from `www.eso.org/gasgano` and `www.eso.org/cpl/esorex.html`. GASGANO is a Data File Organiser developed and maintained by ESO to help its user community to manage and organise in a systematic way the astronomical data observed and produced by all VLT compliant telescopes, i.e. by telescopes which are being operated through Observation Blocks. The tool also supports FITS files which are not generated by those telescopes with limited functionality. For further information, please, refer to the latest `Gasgano User's Manual`. In the special case of the HARPS pipeline, which uses a lot of RAM, we decided not to provide Gasgano support.

Recent pipeline improvements are listed in Section 3. The HARPS instrument is described in Section 4. In section 5 we list known data reduction problems for the HARPS pipeline and possible solutions. The different types of HARPS raw frames and auxilliary data are described in Sections 6, and 7. An overview of the data reduction, the input data, and the recipes involved in the calibration cascade is provided in section 8. More details on inputs, products, quality control measured quantities, and controlling parameters of each recipe is given in section 9. A brief introduction to the usage of the available reduction recipes using EsoRex is presented in Section 10. More detailed descriptions of the data reduction algorithms used by the individual pipeline recipes can be found in Section 11.

In Appendix A the installation of the HARPS pipeline recipes is described.

3 Recent improvements

This release of the HARPS pipeline (version 3.3.12)

- allows to correct the tellurics due to H2O molecules.

- In case of `OBJ_DARK` sets sky correction has to be 'none'.

- If `DPR.TYPE = 'STAR,NONE,<sp_type>'`, the value of `fibre_B` has to be set to DARK. If not, the wavelength matrix can't be found.

Telluric correction for water can now be activated for HARPS. Due to only faint water bands that can be accessed within the HARPS wavelength range, the telluric correction Model can often diverge from physical values when weather conditions are really dry and telluric lines are weak and thus can strongly be contaminated by stellar lines (this will depend on a star by star basis due to spectral line content, BERV and systemic velocity). Although this divergence will not affect the derived RV by more than 10 cm/s, the telluric corrected spectra can be significantly altered. The quality control of the telluric correction have been set to 0 (not valid) as soon as the pressure fitted by the telluric model is significantly larger that the maximum ambiant pressure seen in La Silla (maximum pressure 780, value set at 850). Telluric corrected spectra with a bad quality control should be used with caution, and in any case, users should check the quality of the spectra. A good spectral range to do so is the region 5940 - 5952 Angstrom where the most significant H2O telluric lines are present in the HARPS wavelength range. The best is to plot in this spectral range the non corrected and corrected spectra (`S2D_BLAZE_A.fits` and `S2D_BLAZE_TELL_CORR_A.fits`), along with the fitted tellurique model (`S2D_TELL_SPECTRUM_A.fits`). We note that the telluric correction on ESPRESSO, based on the same algorithm, do not show those divergences because much stronger H2O telluric lines can be accessed in the redder part of the instrument wavelength coverage.

For the rest it is very similar to release 3.2.0 for what concerns HARPS data reduction. We remind below the major improvements with respect to previous public release:

- Added support of the EGGS mode for HARPS. The EGGS mode for the HARPS instrument was validated and all the recipes are operational. Some tuning of the reduction and recipes parameters might still be needed.

- Adaptation of the wavelength solution with the LFC for HARPS The `wave_LFC` reduction was validated for HARPS. Some tuning of the reduction and recipe parameters might still be needed.

- All required static calibrations are now delivered with the pipeline.

4 HARPS Instrument Description

HARPS is a fibre-fed cross dispersed echelle spectrograph located in the coude-west room in the building of the 3.6m telescope. Its science driver is the search for extrasolar planets using the radial velocity method with an accuracy of 1m/sec. This accuracy will be reached with the Simultaneous THAr Reference method.

The following sections describe the main components of the instrument in the sequence the light passes through them. Finally we give a quick overview of the instrument operations.

The La Silla 3.6m telescope is equatorially mounted; HARPS uses its Cassegrain focus.

4.1 Cassegrain Fibre Adapter

To connect the fibre link to the telescope, the La Silla Engineering Department designed and manufactured a new fibre adapter for the Cassegrain focus. The HARPS Cassegrain Fibre Adapter (HCFA) fulfils a number of functions: It allows the remotely controlled exchange of the fibres for HARPS and the CES and provides both fibre feeds with an atmospheric dispersion compensator (ADC) and the possibility to use the telescope's guide camera for guiding on the respective fibre entrance. For HARPS, there is also a neutral density filter and the feed for the calibration fibre which carries the light from a separate calibration unit. This is a crucial part of the calibration concept of HARPS. As a unique feature, HARPS will offer the observer two options for precise wavelength calibration: the default Simultaneous Thorium reference method and the use of the Iodine absorption cell. This Iodine cell is also mounted in the HCFA and can be moved in and out of the telescope beam under remote control.
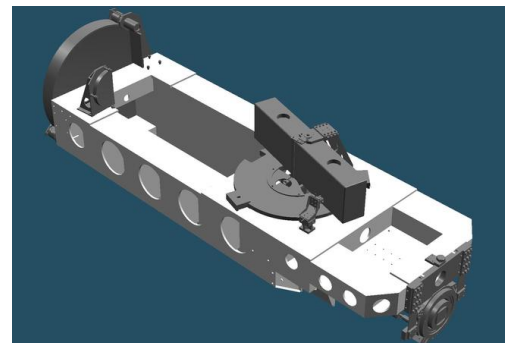
(a) (b) c

Figure 4.1.1: A view of the HARPS Cassegrain fibre adapter (a), the vacuum vessel (b) and the spectrograph (c).

4.2 Fibre link, image scrambler

Strictly speaking HARPS is a distributed system, and one of the most important components connecting its various parts is the observation fibre link. Its purpose is to feed the spectrograph down in the telescope building with (a) the star light collected at the Cassegrain focus and with (b) either the ThAr spectrum for the simultaneous

calibration or with light from the night sky for better sky subtraction. For HARPS we chose two 70 micron fibres (type FVP made by Polymicro), corresponding to 1 arcsecond on the sky, and put it in a "shower tube" made of sturdy steel mesh for mechanical protection. The total length of the fibres from the Cassegrain focus to the spectrograph entrance (by way of the declination bearing, telescope fork, northern telescope mount bearing and into the coude west room) is 38 metres.

To minimise focal ratio degradation, the light is coupled into the object and reference fibres by means of two microlens doublets per fibre. By projecting the image on the fibre input end, the telescope pupil is at infinity. This design combines an excellent image quality with easy, uncritical alignment.

A double image scrambler is located at the entrance of the object/reference fibres into the vacuum vessel of the spectrograph. In combination with the fibre feed, it serves to stabilise the spectrograph illumination: the object may move at the fibre entrance due to guiding errors or seeing, but the intensity distribution at the fibre exit, i.e. the spectrograph entrance, will not change. In addition the scrambler serves as the feedthrough for the fibres into the vacuum and it also houses, on the atmosphere side, the exposure shutter.

The light is finally led to the spectrograph entrance inside the vacuum vessel by means of two short (2 metres) pieces of fibre. The coupling to the spectrograph is again achieved by a pair of doublet microlenses per fibre.

The second fibre link leads, as mentioned before, from the calibration unit in the coude west room up to the Cassegrain focus. It is therefore also about 38 metres long, consisting of a pair of 300 micron core diameter fibres (type FVP made by Polymicro).

4.3 Vacuum Vessel with Spectrograph

The vacuum vessel has the purpose of protecting the spectrograph proper from temperature variations and from the effects of refractive index variations of air. This vessel has a volume of approximately 2m3. It is evacuated by means of a turbo molecular pump before the start of operations; we expect to repeat the regeneration of the vacuum about once or twice per month.

Since the long term stability of the spectrograph is of paramount importance for the success of the exoplanet search, the vacuum should be broken as seldom as possible. For this reason there are no moving functions inside the vacuum except the focussing mechanism of the camera. This will however be adjusted and locked before the vessel is finally closed.

The spectrograph itself is a cross dispersed echelle spectrograph, very similar to UVES at the VLT. It is a white pupil design with the grism cross disperser placed in the white pupil. The echelle grating, a copy of the UVES mosaic, is operated in quasi-Littrow condition. An f/2.1 parabolic mirror serves as collimator and is used in triple pass. A dioptric camera images the cross-dispersed spectra (one each from the object and reference fibres) side by side onto a mosaic of two 2kx4k EEV CCDs. 68 orders cover a spectral range from 378-691nm. All optical components are mounted on a stainless steel optical bench. The optical parameters are listed in the following table.

System	2 fibres (1" dia.), spectral range 378-691nm, collimated beam 208mm
Echelle grating	R4, 31.6 gr/mm blaze angle 75 degrees, mosaic 2x1 on Zerodur monolith 840x214x125mm, efficiency $> 65\%$ in the visible
Cross disperser grism	FK5 grism, 257.17 gr/mm blazed at 480nm, 240x230x50mm, T=73% (av)
Collimator mirror	Zerodur with protected silver coating, f=1560mm, used diameter 730mm, triple pass
Camera	all dioptric, 6 elements in 6 groups, f=728mm, f/3.3, $T > 85\%$
Spectral format	"upper" CCD (Jasmin): 89-114, 533-691nm "lower" CCD (Linda): 116-161, 378-530nm
Spectral resolution	RS=120,000 (measured)
Sampling per spectral element	4.1 px per FWHM
fibres A to B spectra separation	17.3 px
Order separation	Jasmin: order 89: 1.510mm = 100.7px, order 114: 0.940mm = 62.7px Linda: order 116: 0.910mm = 60.7px, order 161: 0.513mm = 34.2px
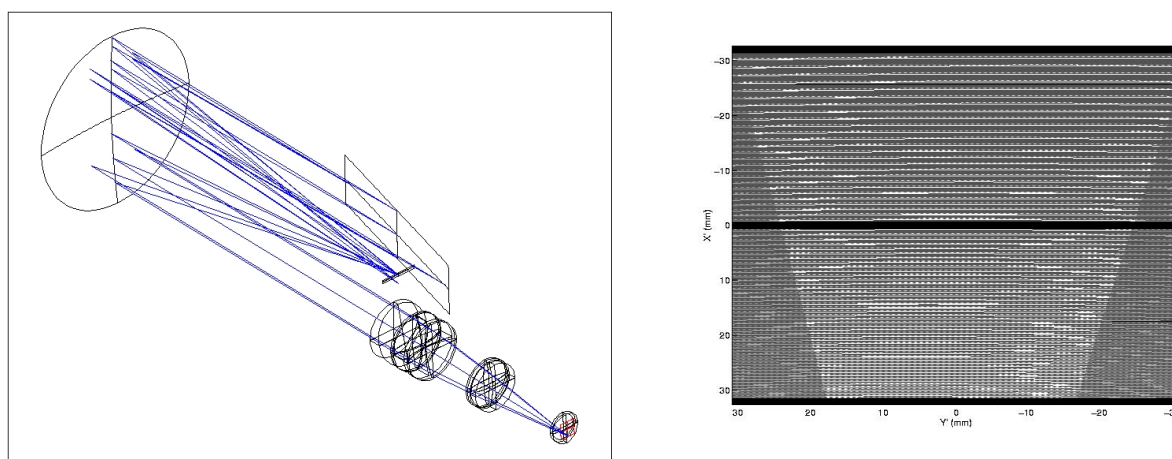
Figure 4.3.1: The HARPS optical scheme (left) and spectral format (right).

4.4 Detector System

HARPS employs a mosaic of two EEV type 44-82 CCDs (nicknamed Jasmin and Linda). The spectral format is thus 4096x4096 pixels (15 micron square) of which a field of 62.7x61.4mm is actually used at a sampling of 4 pixels per spectral element. The main properties of the chips is summarised in the following table.

type	EEV 44-82
spectral format	4096x4096 pixels (15 micron square); used: 62.7x61.4mm
Cosmetics	Science grade (grade 1)
Read-out modes	104 kpx/sec and 416 kpx/sec, one port per chip

As HARPS is a stationary instrument, its detectors are cooled by a continuous flow cryostat (CFC) of the current ESO standard design. A special feature of the HARPS cryostat is however the fact that the detector head is mounted to the spectrograph bench inside the vacuum vessel, with the actual CFC outside. Both are

connected by a stainless steel bellows which protects the detector high vacuum (10E-6 mbar) from the mere "emptiness" (10E-2 mbar) of the spectrograph vessel. The detector head window also serves as field lens of the camera optics.

As a consequence of this configuration, all cables connecting the detector head to the FIERA controller have to pass through vacuum feed throughs.

4.5 Miscellaneous Components

Calibration Unit: The HARPS calibration unit provides the instrument with light for wavelength and flatfield calibration. For this purpose it contains a set of hollow-cathode Thorium-Argon and halogen lamps which can be remotely switched on and off. A motorised exchange mechanism allows to position the calibration fibre in front of any desired lamp. The calibration fibre pair connects the calibration unit, which is located next to the climatised HARPS enclosure in the coude west room,with the fibre adapter at the Cassegrain focus

Exposure meter: In order to improve the observing efficiency by always applying the correct exposure time (according to the selected signal-to-noise ratio SNR) we fitted an exposure meter, again following the example of UVES. Two photon counters are used to separately measure the light coming from the object and reference fibres which is reflected off the gap on the echelle between the two gratings comprising the mosaic.

Iodine cell: An iodine cell is built into the HCFA to allow use of the Iodine self-calibration method. A second Iodine cell is included in the Calibration Unit so that a continuum light with the iodine spectrum in absorption can be projected into the reference fibre.

Thermal enclosure: In order to keep the spectrograph temperature as constant as possible it was decided to put the vacuum vessel in an additional thermal enclosure. This is a well insulated room on the coude floor of the 3.6m telescope building which is itself already temperature stabilised. Based on preliminary measurements we expect to keep the temperature variations of the spectrograph bench inside the vacuum vessel (the only place that really matters) below 0.1K, ideally at 0.01K!

5 Known problems

Several pipeline problems have been identified:

- **Slow execution time.**

  The code parallelisation with OpenMP does enable faster reduction on multi-core platforms. The speed improves significantly if data and reduction is done on very fast I/O disks, like SSDs, and processors with high clock speed.

- **Cosmic rays in raw calibrations.** There have been reported few cases in which the science pipeline crashes with no apparent reason, providing the following error message:

```
10:00:07 [WARNING] espdr_correct_flux: [tid=000] ESC[31mFlux correction not performed:
                   no flux template available for spectral type F5ESC[0m
10:00:07 [ INFO ] : [tid=000] Computing CCF for fibre A sky_sub for size_y = 170 ...
10:00:47 [ ERROR ] espdr_sci_red: [tid=000] espdr_compute_CCF failed for sky_sub:
                   Access beyond boundaries
10:00:47 [ ERROR ] cpl_errorstate_dump: [tid=000] Lost 11192 CPL error(s)
10:00:47 [ ERROR ] cpl_errorstate_dump_one_level: [tid=000] [11193/11212]
                   'Access beyond boundaries' (11) at cpl_image_get:cpl_image_io.c:747
```

  To our best knowledge, the problem is associated to the presence of cosmic rays on some raw calibrations (mainly wavelength calibrations), which led to a product that caused the failure in the science recipe. Future versions of the pipeline will be able to identify such issues and flag them appropriately.

  To overcome the issue, two solutions are currently available:

  - Remove the cosmic ray from the faulty calibrations with external tools before starting the data reduction.
  - Replace the calibrations with those from a previous or following day. It might be worth replacing all the calibrations (all WAVE, ORDERDEF, and FLAT types), even if not all are affected by cosmic rays, to avoid fibre misalignment between calibrations from different days.

- **dependency of results from operative systems**

  In some cases, it has been noticed a small variation on the radial velocity depending on the operative system of few cm/sec. This has to do with the intrinsic precision of the compilers. It is therefore recommended to use the same hardware when reducing the data.

For updated information we recommend the user to also read the Data Reduction F.A.Q. page: *http://www.eso.org/sci/data-processing/faq.html*.

6 Instrument Data Description

The HARPS instrument produces raw data in 3 different configurations or modes (HARPS, EGGS, and PO-LARIMETRY). we say it again, is a copy of the ESPRESSO pipeline). The EGGS mode for the HARPS instrument was validated and all the recipes are operational. Some tuning of the reduction and recipes parameters might still be needed. The mode is identified by the FITS keyword `HIERARCH ESO INS MODE`

To reduce a science frame, the following raw calibration frames are needed:

type	# frames	comments
ORDERDEF	2	each frame illuminates only one fiber, to trace orders
FLAT	5	to determine short and long scale responsiveness variations
ThoriumP	1	both fibres illuminated by ThAr , for wavelength calibration
STD,WAVE	1	fib A observes a spectrophotometric standard, fib B observes the ThAr or FP lamp to flux calibrate the observed science spectrum

It is also necessary to have a set of static calibration data (see next Section).

The following sections provides a brief description of each raw data type involved in the data reduction chain.

6.1 ORDERDEF frames

Two frames are acquired: ORDERDEF_A and ORDERDEF_B and used for order/slice definition, identification and tracing for fibre A and B. A continuum light source is used to illuminate fibre A (or B), while fibre B (or A) is dark (separate frames for different fibres are necessary for the automatic identification of orders or slices).

6.2 FLAT frames

A set of 5 frames are acquired, with the two fibres illuminated simultaneously. Because the reduction recipe, copied from the ESPRESSO pipeline, requires two sets, each of them with one fibre illuminated only (`FLAT_A`, and `FLAT_B`), it is necessary to provide to the recipe twice the same flat, one with flag `FLAT_A` and another time with flag `FLAT_B`.

6.3 Wavelength calibration frames

One wavelength calibration frame obtained illuminating both fibres with the ThAr lamp (or one fibre with ThAr lamp, and the other with Fabry-Perot lamp)

6.4 Contamination by simultaneous reference frames

- `CONTAM_FP`: Measurement of contamination light induced on fibre A by FP simultaneous reference on fibre B. Fibre A is dark, while FP light is injected into fibre B as in a science exposure with simultaneous reference (i.e. same flux level).

6.5 Fiber-to-fiber Relative Efficiency frames

`EFF_AB`: Relative efficiency of fibre B vs. fibre A as a function of wavelength. During twilight, the telescope is pointed at zenith and the skylight is injected into both fibres.

6.6 Spectrophotometric Calibration (FLUX) frames

`FLUX_CALIB`: A spectrophotometric standard star is observed, as in a science exposure, with sky.

6.7 SCIENCE frames

Science exposures are of 3 types:

type	# frames	comments
STAR,DARK	1	Fibre A observes a science target, fibre B nothing.
STAR,SKY	1	Fibre A observes a science target, fibre B the sky background.
STAR,WAVE	1	Fibre A observes a science target, fibre B obserecs the Thar or FP lamp.

All types are reduced by the same recipe and require the same set of calibrations.

7 Static Calibration Data

In the following section static calibration (called also ancillary) data required for ESPRESSO data reduction are listed. Static calibration data correspond to Images or Tables describing configuration setups, reference data, or characteristics of the instrument that are considered to be fixed. This sets them apart from the calibration previously described. For each of them we indicate the corresponding value of the HIERARCH ESO PRO CATG, in short its PRO.CATG, which has to be used to identify the frames listed in the *Set of Frames* (see Section 10.3.1, page 45).

7.1 CCD Geometry Table

These are the static CCD configuration tables describing the CCD geometry. The table contains the number and sizes of the detectors, outputs and prescan and overscan regions. There is one table for each of the supported detector binnigs. Its PRO.CATG is `CCD_GEOM`.

7.2 Instrument Configuration Table

These are the static instrument configuration tables providing the pipeline recipes with all necessary input parameters that are intimately linked to the instrument configuration being used. Its PRO.CATG is `INST_CONFIG`.

7.3 Wavelength Matrix Images

These are the wavelength calibration arrays (one per fibre) in S2D format, with the wavelength of each extracted pixel stored as data value. These are used as static input frames in the flat and wavelength (THAR/FP or FP/THAR) data reduction. Their PRO.CATG is `STATIC_WAVE_MATRIX_A/B`[1].

7.4 DLL Matrix Images

These are the pixel widths in wavelength calibration arrays (one per fibre) in S2D format, with the width in wavelength of each extracted pixel stored as data value. These are used as static input frames in the wavelength (THAR/FP or FP/THAR) and science data reduction. Their PRO.CATG is `STATIC_DLL_MATRIX_A/B`[2].

---

[1] As static WAVE and DLL matrix frames, have a large size, to reduce the size of the pipeline package, we decided to include them in the demo data, available http://www.eso.org/sci/software/pipelines/, and to deliver them to the user together with the raw data when retrieved with the calSelector tool. The FITS filename of the frame delivered with the demo data contains the instrument mode, the binning, the corresponding PRO.CATG and fibre information and a date that correspond to the date since the frame is valid. Raw data should be associated with the closest in time static frame, which is the default behavior in case of Reflex based data reduction.

[2] DLL stands for Delta LL, a short notation corresponding to Delta Wavelength.

7.5 THAR line tables

These are the static ThAr line tables (one per fibre) containing the wavelengths, approximate positions, flux and values of D - the FP mirrors distance, for the emission lines found in the ThAr extracted spectra (S2D). The column called "grouping" is used to save the mode of the corresponding FP peak. The table is not expected to change over the lifetime of the instrument owing to the high long-term stability of ESPRESSO. Their PRO.CATG is `REF_LINE_TABLE_A/B`.

7.6 Flux Standard Star tables

The Flux Standard Star Table lists all the stars (part of the calibration plan) for which we have a precise physical flux calibration [ergs/s/cm^2/A] on an even wavelength scale. The table contains absolute spectral energy distributions of a sample of spectrophotometric standard stars. It is a required input of the recipe `espdr_cal_flux` during response and efficiency estimation, where the observed photometric standard star spectrum, after proper rescaling by exposure time, gain and atmospheric extinction, is divided by the spectrum of the corresponding standard star from this catalog. Its PRO.CATG is `STD_TABLE`.

7.7 Extinction Table

This is the table containing the atmospheric extinction curve for Paranal, as provided in Patat et al. (2011), A&A 527, 91 (their Appendix B). This is used in the `espdr_cal_flux` to determine the efficiency and instrument response and in the `espdr_sci_red`, to flux calibrate the observed science object spectrum. Its PRO.CATG is `EXT_TABLE`.

7.8 Flux Template Table

This is the table containing observed spectral energy distributions of a sample of reference stars with different spectral types spanning late-F to early-M. Suitable stars are solar-metallicity dwarf stars of spectral types F to M, observed at high SNR and at low airmass. To build the flux template, the S2D flux of the star is summed in each spectral order and normalised to one at an arbitrary wavelength (e.g. 550 nm). The flux template is simply the integrated normalised flux in each spectral order. Flux Template table corresponds to the flux distribution as a function of wavelength for stars of different spectral types. This primordial flux is used to correct the effect of chromatic extinction of the atmosphere, composed of several factors, like Rayleight scattering and ozone + aerosols absorption. Its PRO.CATG is `FLUX_TEMPLATE`.

7.9 CCF Template Tables

The template tables, or CCF masks, are used by the cross-correlation process. They consist of a list of rest (RV=0) wavelengths for the spectral lines of interest, along with their relative depth (contrast) with respect to the continuum. Depths are used as weights in the CCF process since Doppler precision is proportional to line depth. Their PRO.CATG are `MASK_TABLE`.

These CCF masks are created from actual spectra of stars covering a range of sub-spectral types. Spectral lines are identified through an automatic procedure and included in the mask if they meet a number of criteria such as large enough depth and limited blending. The current pipeline version includes CCF masks for all M dwarfs down to M5 (individual fits files). The masks cover the full wavelength range of HARPS For QSO observations the CCF computation on the object spectrum is tuned off.

7.10 `PIXEL_GEOM_A-B` and `PIXEL_SIZE_A-B`

Calibrations that describe the geometry of the detector and the pixels to account for the detector stitching effect described in Coffinet et al. 2019, A&A, 629A, 27.

7.11 Bad and hot pixel masks

The ESPRESSO pipeline foresee the reduction of darks and bias frames to construct bad and hot pixel masks. In the case of HARPS, there are no regular bias and dark frames acquired, therefore the bad and hot pixel masks are static calibratins that are provided with the pipeline distribution. Their categories are: `HOT_PIXEL_MASK` and `BAD_PIXEL_MASK`

8 Data Reduction

We give below an overview of the global reduction cascade, starting from basic calibrations down to science reduction. The ESPRESSO association map is shown in Figure 8.0.0. The HARPS association map is similar, with the exception that BIAS, DARK, and LED_FF steps are missing.

- ORDERDEF frames (one per fibre) are used to identify and trace spectral orders or slices on the detector. Each order or slice is fitted in the cross-dospersion direction, with a gaussian every defined number of pixels. The gaussians peaks are fitted with a low degree polynomial, of which coefficients are saved in the recipe products.

- For each fibre, the order profile in cross-dispersion direction is found using high-SNR, co-added FLAT frames. Then, the orders are extracted from the FLAT frames using this profile, and the spectral flat-field is generated. The (extracted) blaze function is obtained through smoothing of the FLAT spectra and correction for the spectral energy distribution of the calibration lamp. The order extraction assumes that 1) main-dispersion direction runs approximately parallel to CCD rows/columns, and 2) slit image tilt is close to zero with respect to CCD columns/rows. In this case, order extraction becomes extremely simple and does not require wavelength calibration frames to track different positions along the slit. The ESPRESSO optical design makes this strategy possible (line tilt very close to zero), and even recommended owing to its simplicity. This method has been successfully applied to HARPS and other radial-velocity spectrographs.

- CONTAM frames are used to measure cross-fibre contamination on fibre A from the simultaneous reference on fibre B (ThAr lamp, laser comb or Fabry-Perot). Contamination frames are used during extraction in the science reduction.

- The relative efficiency of channels A and B as a function of wavelength is measured using EFF_AB frames, which are obtained through blue sky observations. The obtained relative efficiency is used to scale the subtraction on science exposures with simultaneous sky. Even though these are called fibre-to-fibre relative efficiency they, de facto, calibrate the whole Coude train as seen through fibre A and B. As such, they are UT dependent.

- The wavelength calibration for both fibres is determined using WAVE frames.

- FLUX_CALIB frames are used to compute the absolute efficiency of the instrument as a function of wavelength, using spectrophotometric standard stars. The efficiency is computed from the comparison between the observed spectrum and reference flux table.The efficiency curve is used in the science reduction to calibrate the science spectrum in flux. The precision of the flux calibration is generally low because of highly variable fibre losses due to seeing and the low frequency of standard star observations.

- Finally, SCIENCE frames are of two different sub-types depending on the source of light on the fibre B: sky or Fabry-Perot. Science reduction makes use of all calibration products listed above and generates extracted S2D spectra and merged, rebinned S1D spectra, together with S2D and S1D error and quality maps. Finally, the cross-correlation function (CCF) of the S2D spectrum is computed and the radial velocity is obtained from a Gaussian fit to the CCF.
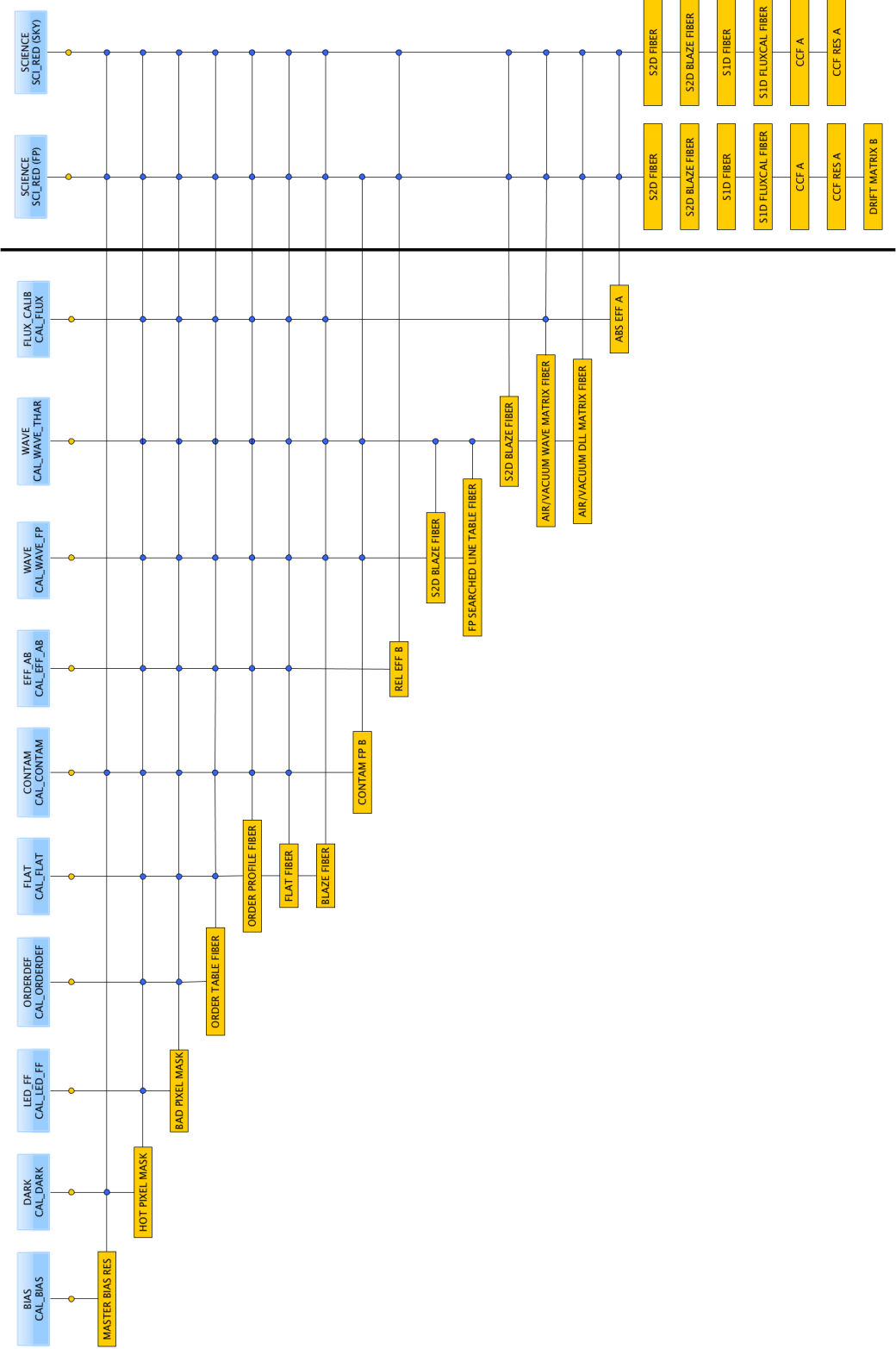
Figure 8.0.0: The cascade of the ESPRESSO pipeline recipes. For the calibration recipes, only the products used later in the reduction chain are listed. The science recipes have all the products listed.

9 Pipeline Recipes Interfaces

In this section we provide for each recipe [3] examples of the required input data (and their tags). In the following we assume that /path_file_raw/filename_raw.fits and /path_file_cdb/filename_cdb.fits are existing FITS files.

We also provide a list of the pipeline products for each recipe, indicating their default recipe name, the value of the FITS keyword HIERARCH ESO PRO CATG (in short PRO.CATG) and a short description.

The relevant keywords used to classify each frame are the following:

Association keyword	Information
HIERARCH ESO DPR TYPE	data type
HIERARCH ESO INS MODE	mode
HIERARCH ESO PRO CATG	product category

For each recipe we list in a table the input parameters (as they appear in the recipe configuration file), the corresponding aliases (the corresponding names to be eventually set on command line) and their default values. Also quality control parameters are listed. Those are stored in relevant pipeline products. More information on instrument quality control can be found on `www.eso.org/qc`.

The user may obtain brief description of the main input recipe parameters by typing

esorex -help recipe,

for example,

esorex -help `espdr_orderdef`

A possible esorex configuration parameter value is –suppress-prefix=FALSE, in which case all products will be renamed with a prefix settable by the parameter –output-prefix, defaulted to `out_`, and an increasing number, like (`out_0000.fits`, `out_0001.fits`, `out_000N.fits`) For this reason the table briefly describing the products contains also a first column indicating the product ID, which is the value of the product number (with minimum significant digits).

The pipeline performs several quality assesment on the data. The result is stored in a QC keyword that has value "QC <specifier> CHECK", where the <specifier> indicates the kind of check done. A check is successful if the value of the keyword is 1. Some recipe may do several tests. It is also created a keyword "QC <recipe> CHECK", where <recipe> indicates the recipe executed, that indicates the overall product of all checks.

The PRO.CATG chosen for the extracted spectra falls into two groups: 'S2D_*' spectra, that can be of type A or B to represent the object or the calibration fibersi, respectively. These files are 2D images where the rows (Y axis) contain the extracted spectrum of a given order or slice. We remind the reader that in order to keep compact the spectrograph the input light beam from the input slices is split in two slices. These are dispersed by the grating and imaged on the detectors (one for each instrument arm) on different orders. As the input beam is sliced in two component the order shows two components, adjacent one to another. These order slices are distinguishable in UHR and HR modes and overlap in the MR mode. There are 45 orders in the Blue and 40 in the red. In UHR and HR modes the pipeline extracts then two slices for each order, and in MR only one.

---

[3]We do not describe here two recipes: the recipe espdr_single_bias, used to reduce HARPS data, and the recipe espdr_wave_TH_drift, used for the wavelength calibration when are not available enough good quality FP, FP frames.

9.1 espdr_orderdef

9.1.1 Input

type	TAG	n	setting
raw	ORDERDEF_A	1	any
raw	ORDERDEF_B	1	any
ref	CCD_GEOM	+ 1	match
ref	INST_CONFIG	1	match
cdb	HOT_PIXEL_MASK	1	match
cdb	BAD_PIXEL_MASK	1	match

9.1.2 Output

ID	PRO.CATG	type	Note
0	ORDER_TABLE_A	cdb	Order table tracing fibre A
1	ORDER_TABLE_B	cdb	Order table tracing fibre B

9.1.3 Quality control

This recipe computes the following QC parameters (on the image extension n, order m, and read-out region i,j):

QC key name	description
QC EXTn ROXi ROYj MAX FLUX	max flux on raw image [ADU]
QC ORDERm EXTn POS	order position [pix]
QC ORDERm EXTn RES MIN	min residuals [pix]
QC ORDERm EXTn RES MAX	max residuals [pix]
QC ORDERm EXTn RES STDEV	rms residuals [pix]
QC ORDERm EXTn NB	physical order number
QC ORDER NB	number of orders in the image
QC EXTn ROXi ROYj MAX FLUX	max flux, raw image [ADU]
QC SATURATION CHECK	check on saturation
QC ORDERDEF ORDER CHECK	check on number of detected orders
QC ORDERDEF STDEV CHECK	check on rms residuals
QC ORDERDEF MIN CHECK	check on min residuals
QC ORDERDEF MAX CHECK	check on max residuals
QC ORDERDEF CHECK	overall check

9.1.4 Parameters

parameter name	description

`ovsc_sig_clip_method`	method for sigma clipping in OVSC, can be mean or median. [mean]
`ovsc_ksigma`	ksigma for sigma clipping in OVSC, must be between: 1.50 and 1000.00. [4.0]
`ovsc_max_iter`	maximal number of iterations in OVSC, must be between: 1 and 5000000. [10]

9.2 espdr_mflat

9.2.1 Input

type	TAG	n	bin
raw	FLAT_A	5...n	any
raw	FLAT_B	5...n	any
ref	CCD_GEOM	1	match
ref	INST_CONFIG	1	match
ref	STATIC_WAVE_MATRIX_A	1	match
ref	STATIC_WAVE_MATRIX_B	1	match
cdb	HOT_PIXEL_MASK	1	match
cdb	BAD_PIXEL_MASK	1	match
cdb	ORDER_TABLE_A	1	match
cdb	ORDER_TABLE_B	1	match

9.2.2 Output

ID	PRO.CATG	type	Note
0	FF_BACKGROUND_MAP_A	cdb	Background map for fibre A
1	FS2D_A	cdb	Order by order exracted spectrum for fibre A
2	BLAZE_A	cdb	Blaze for fibre A
3	FSPECTRUM_A	cdb	Order by order exracted spectrum for fibre A
4	ORDER_PROFILE_A	cdb	Order profile for fibre A
5	FF_BACKGROUND_MAP_B	cdb	Background map for fibre B
6	FS2D_B	cdb	Order by order exracted spectrum for fibre B
7	BLAZE_B	cdb	Blaze for fibre B
8	FSPECTRUM_B	cdb	Order by order exracted spectrum for fibre B
9	ORDER_PROFILE_B	cdb	Order profile for fibre B

9.2.3 Quality control

This recipe computes the following QC parameters (on the image extension n, order m, and read-out region i,j):

QC key name	description
QC EXTn ROXi ROYj MAX FLUX	max flux on raw image [ADU]
QC ORDERm FLAT RMS	FLAT RMS in the order

QC ORDERm SNR	SNR in the order
QC ORDERm COSMIC NB	cosmics number in the order
QC EXTn ROXi ROYj MAX FLUX	max flux, raw image [ADU]
QC EXTn BKGR MEAN	bkgr mean in the extension [e-]
QC EXTn BKGR MIN	bkgr min in the extension [e-]
QC EXTn BKGR MAX	bkgr max in the extension [e-]
QC SATURATION CHECK	quality check on saturation
QC FLAT RMS CHECK	quality check on rms
QC FLAT SNR CHECK	quality check on SNR
QC FLAT BKGR CHECK	background quality check
QC FLAT CHECK	overall quality check

9.2.4 Parameters

parameter name	description
ovsc_sig_clip_method	method for sigma clipping in OVSC, can be mean or median. [mean]
ovsc_ksigma	ksigma for sigma clipping in OVSC, must be between: 1.50 and 1000.00. [4.0]
ovsc_max_iter	maximal number of iterations in OVSC, must be between: 1 and 5000000. [10]
flat_sig_clip_method	method for sigma clipping in master FLAT, can be: mean or median. [median]
flat_ksigma	ksigma for sigma clipping in MFLAT, must be between: 1.50 and 1000.00. [10.0]
background_sw	Background measurement activation (on/off). [on]
bkgr_grid_size_x	Grid size in x used to calculate the background, between: 16 and 10000. [577]
bkgr_grid_size_y	Grid size in y used to calculate the background, between: 16 and 10000. [256]
flat_extraction_method	Method used to extract orders. [horne]
flat_extraction_ksigma	ksigma for extraction, must be between: -1.00 and 20.00. [-1.0]

9.3 espdr_wave_TH_drift

9.3.1 Input

type	TAG	n	setting
raw	THAR_THAR	1	any
ref	CCD_GEOM	1	match
ref	INST_CONFIG	1	match
cdb	HOT_PIXEL_MASK	1	match
cdb	BAD_PIXEL_MASK	1	match
cdb	ORDER_TABLE_A	1	match
cdb	ORDER_TABLE_B	1	match
cdb	ORDER_PROFILE_A	1	match
cdb	ORDER_PROFILE_B	1	match
cdb	FSPECTRUM_A	1	match
cdb	FSPECTRUM_B	1	match
cdb	WAVE_MATRIX_A	1	match

cdb	WAVE_MATRIX_B	1	match
cdb	DLL_MATRIX_A	1	match
cdb	DLL_MATRIX_B	1	match

9.3.2 Output

ID	PRO.CATG	type	Note
0	S2D_THAR_THAR_A	cdb	S2D Extracted spectrum for fibre A
1	S2D_THAR_THAR_B	cdb	S2D Extracted spectrum for fibre B
2	S2D_BLAZE_THAR_THAR_A	cdb	S2D Extracted blaze for fibre A
3	S2D_BLAZE_THAR_THAR_B	cdb	S2D Extracted blaze for fibre B
4	WAVE_MATRIX_THAR_THAR_A	cdb	Wave map (vacuum) for fibre A
5	WAVE_MATRIX_THAR_THAR_B	cdb	Wave map (vacuum) for fibre B
6	DLL_MATRIX_THAR_THAR_A	cdb	DLL map (vacuum) for fibre A
7	DLL_MATRIX_THAR_THAR_B	cdb	DLL map (vacuum) for fibre B
8	AIR_DLL_MATRIX_THAR_THAR_A	cdb	DLL map (air) for fibre A
9	AIR_DLL_MATRIX_THAR_THAR_B	cdb	DLL map (air) for fibre B
10	WAVE_TABLE_THAR_THAR_A	cdb	Table of coefficients of the wave sol fit for each order (fibre A)
11	WAVE_TABLE_THAR_THAR_B	cdb	Table of coefficients of the wave sol fit for each order (fibre B)

9.4 espdr_cal_contam

9.4.1 Input

type	TAG	n	setting
raw	RAW_CONTAM_FP	1	any
ref	CCD_GEOM	1	match
ref	INST_CONFIG	1	match
cdb	MASTER_BIAS_RES	1	match
cdb	HOT_PIXEL_MASK	1	match
cdb	BAD_PIXEL_MASK	1	match
cdb	ORDER_TABLE_A	1	match
cdb	ORDER_TABLE_B	1	match
cdb	ORDER_PROFILE_A	1	match
cdb	ORDER_PROFILE_B	1	match
cdb	FSPECTRUM_A	1	match
cdb	FSPECTRUM_B	1	match

9.4.2 Output

ID	PRO.CATG	type	Note

0	`CONTAM_S2D_A`	qc	S2D Extracted spectrum for fibre A
1	`CONTAM_S2D_B`	qc	S2D Extracted spectrum for fibre B
2	`CONTAM_FP`	cdb	Contamination of fibre B on fibre A

9.4.3   Quality control

This recipe computes the following QC parameters (on the image extension n, order m, and read-out region i,j):

QC key name	description
QC EXTn ROXi ROYj MAX FLUX	max flux, raw image
QC SATURATION CHECK	quality check on saturation
QC ORDERm MAX FLUX	S2D flux max, fibre A
QC CONTAM FLUX CHECK	quality check on flux
QC CONTAM CHECK	overall quality check

9.4.4   Parameters

parameter name	description
`ovsc_sig_clip_method`	method for sigma clipping in OVSC, can be mean or median. [mean]
`ovsc_ksigma`	ksigma for sigma clipping in OVSC, must be between: 1.50 and 1000.00. [4.0]
`ovsc_max_iter`	maximal number of iterations in OVSC, must be between: 1 and 5000000. [10]
`background_sw`	if on (default value) then background is subtracted, 1=on 0=off. [0]
`contam_bkgr_grid_size_x`	Grid size in x used to calculate the background, between: 128 and 512. [577]
`contam_bkgr_grid_size_y`	Grid size in y used to calculate the background, between: 128 and 512. [256]
`contam_extraction_method`	Method used to extract orders. [horne]
`contam_extraction_ksigma`	ksigma used to extract orders. [3.5]

9.5   espdr_cal_eff_ab

9.5.1   Input

type	TAG	n	setting
raw	`EFF_AB`	1	any
ref	`CCD_GEOM`	1	match
ref	`INST_CONFIG`	1	match
cdb	`HOT_PIXEL_MASK`	1	match
cdb	`BAD_PIXEL_MASK`	1	match
cdb	`ORDER_TABLE_A`	1	match
cdb	`ORDER_TABLE_B`	1	match
cdb	`ORDER_PROFILE_A`	1	match
cdb	`ORDER_PROFILE_B`	1	match

cdb	`FSPECTRUM_A`	1	match
cdb	`FSPECTRUM_B`	1	match

9.5.2 Output

ID	PRO.CATG	type	Note
0	`S2D_BLAZE_EFF_A`	qc	S2D Extracted blaze spectrum for fibre A
1	`S2D_BLAZE_EFF_B`	qc	S2D Extracted blaze spectrum for fibre B
2	`REL_EFF_B`	cdb	spectrum describing the trasmission ratio fibre B/fibre A in pixel units.

9.5.3 Quality control

This recipe computes the following QC parameters (on the image extension n, order m, and read-out region i,j):

QC key name	description
QC EXTn ROXi ROYj MAX FLUX	max flux, raw image [ADU]
QC SATURATION CHECK	check on saturation
QC EXTn BKGR MEAN	bkgr mean in the extension [e-]
QC EXTn BKGR MIN	bkgr min in the extension [e-]
QC EXTn BKGR MAX	bkgr max in the extension [e-]
QC EFFAB BKGR CHECK	check on background level
QC ORDERm SNR	SNR in the order m
QC ORDERm COSMIC NB	cosmics number in the order
QC EFFAB SNR CHECK	check on SNR
QC REL EFF MIN	min relative efficiency
QC REL EFF MAX	max relative efficiency
QC REL EFF MIN CHECK	check on computation min efficiency
QC REL EFF MAX CHECK	check on computation max efficiency
QC EFFAB CHECK	overall check on effab products

9.5.4 Parameters

parameter name	description
`ovsc_sig_clip_method`	method for sigma clipping in OVSC, can be mean or median. [mean]
`ovsc_ksigma`	ksigma for sigma clipping in OVSC, must be between: 1.50 and 1000.00. [4.0]
`ovsc_max_iter`	maximal number of iterations in OVSC, must be between: 1 and 5000000. [10]
`eff_ab_bkgr_grid_size_x`	Grid size in x used to calculate the background, between: 128 and 512. [577]
`eff_ab_bkgr_grid_size_y`	Grid size in y used to calculate the background, between: 128 and 512. [256]
`eff_ab_extraction_method`	Method used to extract orders. [horne]
`eff_ab_extraction_ksigma`	ksigma used to extract orders. [5.0]
`eff_ab_poly_deg`	Efficiency computation fit polynomial degree. [3]

9.6 espdr_cal_flux

9.6.1 Input

type	TAG	n	setting
raw	FLUX	1	any
ref	CCD_GEOM	1	match
ref	INST_CONFIG	1	match
ref	EXT_TABLE	1	match
ref	STD_TABLE	1	match
cdb	HOT_PIXEL_MASK	1	match
cdb	BAD_PIXEL_MASK	1	match
cdb	ORDER_TABLE_A	1	match
cdb	ORDER_TABLE_B	1	match
cdb	ORDER_PROFILE_A	1	match
cdb	ORDER_PROFILE_B	1	match
cdb	FSPECTRUM_A	1	match
cdb	FSPECTRUM_B	1	match
cdb	BLAZE_A	1	match
cdb	BLAZE_B	1	match
cdb	WAVE_MATRIX_THAR_FP_A	1	match
cdb	WAVE_MATRIX_FP_THAR_B	1	match

9.6.2 Output

ID	PRO.CATG	type	Note
0	S2D_STD_A	qc	S2D Extracted spectrum for fibre A
1	S1D_STD_A	qc	S1D Extracted spectrum for fibre A
2	S1D_ENERGY_STD_A	cdb	S1D energy for fibre A
3	S2D_BLAZE_STD_A	cdb	S2D blaze spectrum for fibre A
4	AVG_FLUX_STD_A	cdb	average flux for fibre A
5	ABS_EFF_RAW_A	cdb	absolute raw efficiency for fibre A
6	ABS_EFF_A	cdb	absolute efficiency for fibre A

9.6.3 Quality control

This recipe computes the following QC parameters (on the image extension n, order m, and read-out region i,j):

QC key name	description
QC EXTn ROXi ROYj MAX FLUX	max flux, raw image [ADU]
QC SATURATION CHECK	saturation [ADU] QC
QC ORDERm COSMIC NB	cosmics number in the order m
QC ORDERm SNR	SNR in the order m

QC CALFLUX SNR CHECK	check on SNR computed value
QC EXTn BKGR MEAN	bkgr mean in the extension [e-]
QC EXTn BKGR MIN	bkgr min in the extension [e-]
QC EXTn BKGR MAX	bkgr max in the extension [e-]
QC MIN RAW ABS EFF	min raw absolute efficiency
QC MAX RAW ABS EFF	max raw absolute efficiency
QC MEAN RAW ABS EFF	mean raw absolute efficiency
QC MIN EXTCORR ABS EFF	minimum extinction-corrected abs eff
QC MAX EXTCORR ABS EFF	maximum extinction-corrected abs eff
QC MEAN EXTCORR ABS EFF	mean extinction-corrected abs eff
QC MIN ABS EFF	minimum absolute efficiency
QC MAX ABS EFF	maximum absolute efficiency
QC MEAN ABS EFF	mean absolute efficiency
QC MIN ABS EFF CHECK	minimum absolute efficiency check
QC MAX ABS EFF CHECK	maximum absolute efficiency check
QC BERV[4]	barycentric correction [km s^{-1}]
QC BJD	barycentric Julian date (TDB) [JD]
QC BERVMAX	barycentric max [km s^{-1}]
QC CALFLUX SEEING KW CHECK	check on seeing
QC CALFLUX CHECK	overall check on recipe

9.6.4 Parameters

parameter name	description
`ovsc_sig_clip_method`	method for sigma clipping in OVSC, can be mean or median. [mean]
`ovsc_ksigma`	ksigma for sigma clipping in OVSC, must be between: 1.50 and 1000.00. [4.0]
`ovsc_max_iter`	maximal number of iterations in OVSC, must be between: 1 and 5000000. [10]
`background_sw`	if 1 (default value) then background is subtracted, 1=on 0=off. [0]
`bkgr_grid_size_x`	Grid size in x used to calculate the background, between: 128 and 512. [577]
`bkgr_grid_size_y`	Grid size in y used to calculate the background, between: 128 and 512. [256]
`flux_extraction_method`	Method used to extract orders. [horne]
`flux_extraction_ksigma`	ksigma used to extract orders. [3.5]
`flux_poly_deg`	Efficiency computation fit polynomial degree. [6]

9.7 espdr_sci_red

9.7.1 Input

type	TAG	n	setting

---

[4]Barycentric Earth Radial Velocity (BERV)

raw	OBJ_FP	1	any
ref	CCD_GEOM	1	match
ref	INST_CONFIG	1	match
ref	MASK_TABLE	11	match
ref	EXT_TABLE	1	match
ref	MASK_LUT	1	match
ref	FLUX_TEMPLATE	1	match
cdb	MASTER_BIAS_RES	1	match
cdb	HOT_PIXEL_MASK	1	match
cdb	BAD_PIXEL_MASK	1	match
cdb	ORDER_TABLE_A	1	match
cdb	ORDER_TABLE_B	1	match
cdb	ORDER_PROFILE_A	1	match
cdb	ORDER_PROFILE_B	1	match
cdb	FSPECTRUM_A	1	match
cdb	FSPECTRUM_B	1	match
cdb	BLAZE_A	1	match
cdb	BLAZE_B	1	match
cdb	WAVE_MATRIX_THAR_FP_A	1	match
cdb	WAVE_MATRIX_FP_THAR_B	1	match
cdb	DLL_MATRIX_THAR_FP_A	1	match
cdb	DLL_MATRIX_FP_THAR_B	1	match
cdb	S2D_BLAZE_THAR_FP_B	1	match
cdb	WAVE_MATRIX_LFC_FP_A	?[5]	match
cdb	WAVE_MATRIX_FP_LFC_B	?	match
cdb	DLL_MATRIX_LFC_FP_A	?	match
cdb	DLL_MATRIX_FP_LFC_B	?	match
cdb	CONTAM_FP	?	match
cdb	REL_EFF_B	1	match
cdb	ABS_EFF_A	?	match
cdb	CRH_MAP	?	match

For the observations with sky on fibre B - OBJ_SKY, the inputs for drift computation (S2D_BLAZE_THAR_FP_B) and the contamination (CONTAM_FP) are not needed. Note that there can be cross-talk from the object fibre to the sky fibre of up to 1%. This contamination is not corrected for and is thus over-subtracted in the SKYSUB frames.

The LFC products are needed if the user wishes to use the LFC wavelength calibration products. In this case s/he has to set the parameter wave_cal_source to LFC. Even if the WAVE/DLL_MATRIX_THAR_FP/FP_THAR are provided, they will not be used in this case.

For a more flexible data reduction this release allows the user to provide the CRH_MAP. The user has to make sure this frame contains two images of the pixels locations affected by cosmics in two extensions. Each image must have the same X and Y sizes and the same ARCFILE of the CCD_CORR_SCIENCE frame.

---

[5]? indicates an optional input

9.7.2 Output for OBJ_FP

ID	PRO.CATG	type	Note
0	CCF_RESIDUALS_A	qc	residuals from CCF computation
1	DRIFT_MATRIX_B	qc	drift matrix relative to fibre B
2	S2D_BLAZE_A	cdb	S2D spectrum blaze fibre A
3	S2D_A	cdb	S2D spectrum for fibre A
4	S2D_BLAZE_B	cdb	S2D spectrum blaze fibre B
5	S2D_B	cdb	S2D spectrum for fibre B
6	S1D_A	cdb	S1D spectrum for fibre A
7	S1D_FINAL_A	cdb	Table containing all above S1D products for fibre A. The format is compatible with ESO archive standard.
8	S1D_FLUXCAL_A	cdb	S1D flux calibrated spectrum for fibre A
9	S1D_B	cdb	S1D spectrum for fibre B
10	S1D_FINAL_B	cdb	Table containing all above S1D products for fibre B.
11	CCF_A	cdb	CCF for fibre A
12	CCD_CORR_SCIENCE	-	frame on which cosmics are dectected. [6]
13	CRH_MAP	-	(van-dokkum) cosmic detection map[7]
14	FCRH_MAP	-	post filtered cosmic detection map[8]

If the user does not provide `ABS_EFF_A` optional input frame the "FLUXCAL" products are not created.

The 'S2D_*' data products contain the extracted spectrum order by order or slice (depending on the mode), the associated error and quality pixel map in the extensions SCIDATA, ERRDATA, QUALDATA respectively. The 'S2D_*' products also contain the wavelength solutions and the detector pixels sizes computed in vacuum and air in the extensions `WAVEDATA_VAC_BARY`, `WAVEDATA_AIR_BARY`, `DLLDATA_VAC_BARY`, `DLLDATA_AIR_BARY` respectively.

The 'S1D_*' product tables provides the wavelength solution (in vacuum and air), the merged spectrum resampled to an equidistant grid, its error and pixel quality flag.

The `S1D_FINAL_B` product table contains the same information as the S1D products, in Phase3 compliant data format. It also includes the computed Signal to Noise ratio (SNR).

The fibre `S1D_FINAL_A` product table includes the flux calibrated (and, when available, sky subtracted) object spectrum, the SNR, and information on the flux in electrons (and eventually the one of the sky spectrum) together with the associated error and pixel quality.

The `CCF_A` data contains the computed CCF for each order and the associated error and pixel quality for each order/slice (depending on the mode) and an additional first raw for the weighted sum of the CCF.

9.7.3 Output for OBJ_SKY

---

[6]optional product to inspect cosmic detection obtained if extra_products_sw is set to True

[7]optional product to inspect cosmic detection obtained if extra_products_sw is set to True

[8]optional product to inspect cosmic detection obtained if extra_products_sw is set to True and –lacosmic.post-filter-x/y are $>= 1$

ID	PRO.CATG	type	Note
0	CCF_RESIDUALS_A	qc	residuals from CCF computation
1	S2D_SKYSUB_A	cdb	S2D spectrum sky-subtracted for fibre A
2	S2D_BLAZE_A	cdb	S2D spectrum blaze fibre A
3	S2D_A	cdb	S2D spectrum for fibre A
4	S2D_BLAZE_B	cdb	S2D spectrum blaze fibre B
5	S2D_B	cdb	S2D spectrum for fibre B
6	S1D_SKYSUB_A	cdb	S1D spectrum sky-subtracted for fibre A
7	S1D_FINAL_A	cdb	Table containing all above S1D products for fibre A. The format is compatible with ESO archive standard.
8	S1D_SKYSUB_FLUXCAL_A	cdb	S1D spectrum sky-subtracted flux calibrated for fibre A
9	S1D_A	cdb	S1D spectrum for fibre A
10	S1D_FLUXCAL_A	cdb	S1D flux calibrated spectrum for fibre A
11	S1D_B	cdb	S1D spectrum for fibre B
12	S1D_FINAL_B	cdb	Table containing all above S1D products for fibre B.
13	CCF_A	cdb	CCF for fibre A
14	CCF_B	cdb	CCF for fibre B

If the user does not provide ABS_EFF_A optional input frame the "FLUXCAL" products are not created.

9.7.4 Quality control

This recipe computes the following QC parameters (on the image extension n, order m, and read-out region i,j):

QC key name	description
QC EXTn ROXi ROYj BIAS RON	RON[ADU]
QC EXTn ROXi ROYj MAX FLUX	max flux, raw image [ADU]
QC SATURATION CHECK	saturation [ADU] QC
QC ORDERm SNR	SNR in the order m
QC ORDERm COSMIC NB	cosmics number in the order m
QC ORDERm COSMIC NB2	Additional cosmics found in each order
QC ORDERm FLUX CORR	flux correction for the order m
QC FLUX CORR MIN	min of flux correction
QC FLUX CORR MAX	max of flux correction
QC SCIRED FLUX CORR CHECK	flux correction QC
QC DRIFT DETn SLOPE_O	Drift slope across the orders per detector
QC DRIFT DETn SLOPE_O_ERR	Drift slope error
QC DRIFT DETn SLOPE_X	Drift slope along the orders per detector
QC DRIFT DETn SLOPE_X_ERR	Drift slope error
QC DRIFT DETn CHISQ	CHI2 of the drift fit
QC DRIFT DETn REJECTED	nb of rejected pixels
QC DRIFT DETn MEAN	mean drift per detector
QC DRIFT DETn MEAN_ERR	mean drift error

QC DRIFT DETn FLUX_RATIO	flux ratio per detector
QC DRIFT DETn FLUX_RATIO_ERR	flux ratio error
QC DRIFT DETn FIRST_ORDER	first order on the detector
QC DRIFT DETn LAST_ORDER	last order on the detector
QC DRIFT CHI2 CHECK	quality check on the drift
QC BERV[9]	barycentric correction [km s^{-1}]
QC BJD	barycentric Julian date (TDB) [JD]
QC BERVMAX	barycentric max [km s^{-1}]
QC EXTn BKGR MEAN	bkgr mean in the extension [e-]
QC EXTn BKGR MIN	bkgr min in the extension [e-]
QC EXTn BKGR MAX	bkgr max in the extension [e-]
QC CCF RV	radial velocity [km s^{-1}]
QC CCF RV ERROR	error on Radial velocity [km s^{-1}]
QC CCF FWHM	CCF FWHM [km s^{-1}]
QC CCF FWHM ERROR	CCF FWHM error [km s^{-1}]
QC CCF CONTRAST	CCF contrast
QC CCF CONTRAST ERROR	CCF contrast error
QC CCF CONTINUUM	CCF continuum level [e-]
QC CCF MASK	CCF mask used
QC CCF FLUX ASYMMETRY	CCF asymmetry [km s^{-1}]
QC CCF FLUX ASYMMETRY ERROR	CCF asymmetry error [km s^{-1}]
QC CCF BIS SPAN	CCF bisector span (km s^{-1})
QC CCF BIS SPAN ERROR	CCF bisector span error [km s^{-1}]
QC SCIRED CHECK	overall science check

Please note that previous list refers to the case of SCIENCE_FP. In case of SCIENCE_SKY, there is no drift computation, nor corresponding check. Moreover in that case S2D_BLAZE_THAR_FP is not a necessary input.

9.7.5 Parameters

parameter name	description
ovsc_sig_clip_method	method for sigma clipping in OVSC, can be mean or median. [mean]
ovsc_ksigma	ksigma for sigma clipping in OVSC, must be between: 1.50 and 1000.00. [4.0]
ovsc_max_iter	maximal number of iterations in OVSC, must be between: 1 and 5000000. [10]
wave_cal_source	Wavelength calibration source to be used on science fiber ('THAR' or 'LFC'). [7
ksigma_cosmic	ksigma for removing cosmics on fiber A or SKY, -1.0 - no cosmics removal. [3.5
rv_center	Approximate RV in km s^{-1}. In case of default (-9999) the pipeline will use the v FITS keyword OCS OBJ RV. [-9999.0]
rv_range	Half Range for the RV table [km s^{-1}]. [20.0]
rv_step	Range's step for the RV table [km s^{-1}] . [0.5]
mask_table_id	Mask table to be used for CCF computation, defined by it's spectral type Id, e.g. 'G8', 'G9', 'K2', 'K6', 'M0', 'M2', 'M3', 'M4', or 'M5'. In case of default ('X

---

[9]Barycentric Earth Radial Velocity (BERV)

	pipeline will read the value of the FITS keyword OCS OBJ SP TYPE and uses a correspondence table to assign the closest available mask to the spectral type giv the FITS keyword or the user. [XX] If none specified, the G2 is used. For QSO observations the CCF computation on the object spectrum is tuned off.
extraction_method	Method used to extract orders: 'horne' or 'simple' [horne]
background_sw	Background measurement activation (on/off). [on]
flux_correction_type	Flux correction: NONE, AUTO or <spectral_type>. Spectral types are the one the parameter mask_table_id. In case of AUTO the behaviour is as for mask_table_id set to 'XX'. For QSO observations the correction is perform
drift_correction_sw	Drift correction activation (on/off). [on]
bias_res_removal_sw	Flag indicating to remove or not MB residuals. [on]
sci_bkgr_grid_size_x	Grid size in x used to calculate the background, between: 16 and 10000. [577]
sci_bkgr_grid_size_y	Grid size in y used to calculate the background, between: 16 and 10000. [256]
drift_method_fp	Method adopted to compute drift. [flux_global_drift_global_sequential_fit]
drift_space	Space to compute drift (pixel/velocity). [pixel]
drift_ksigma	ksigma for computing drift, -1.0 - no ksigma clip. [50.0]
sky_sub_method	Method used to subtract the sky (pixel-by-pixel/smoothed). [pixel-by-pixel]
sky_sub_sliding_box_size	Sliding box size in smoothed sky subtraction. [50]
slit_loss	Slit loss correction in flux calibration. [-1.0]. A null or negative value calculates correction from the seeing (FWHM.IA header keyword). A positive value indica correction to apply (the flux calibrated spectrum will be divided by the value spe by the parameter) instead of the one computed by the seeing. For OBJ,SKY obse the correction is applied only to sky-subtracted products.
cosmic_detection_sw	LA Cosmic detection activation switch. [0]
lacosmic.post-filter-x	X Size of the post filtering kernel. [0]
lacosmic.post-filter-y	Y Size of the post filtering kernel. [0]
lacosmic.post-filter-mode	Post filtering mode. [dilation]
lacosmic.sigma_lim	LA Cosmic Poisson fluctuation threshold. [4.0]
lacosmic.f_lim	LA Cosmic minimum contrast. [4.0]
lacosmic.max_iter	LA Cosmic max number of iterations. [5]
extra_products_sw	Set to TRUE to create extra products to inspect LA Cosmic results. [FALSE]

10 Quick start

This section describes the most immediate usage of the HARPS pipeline recipes.

10.1 HARPS pipeline recipes

The current HARPS pipeline is based on a set of twelve stand-alone recipes involved in the data reduction cascade:

espdr_orderdef Characterises the orders as imaged on the CCD;

espdr_mflat Creates the master flat, blaze and order profiles;

espdr_wave_TH_drift Wavelength calibration with drift w.r.t. the TH lines;

espdr_cal_contam Generates a contamination frame and checks contamination level on science fibre;

espdr_cal_eff_ab Computes the relative efficiency between sky and science fibres vs. wavelength;

espdr_cal_flux Measures the absolute efficiency curve;

espdr_sci_red Performs science reduction;

Other recipes are distributed, but they refer to the reduction of ESPRESSO or NIRPS data and they are not described here.

10.2 An introduction to Reflex and EsoRex

Before being able to call pipeline recipes to process a set of data, the data must be correctly classified, and associated with the appropriate calibrations. The *Data Classification* consists of tasks such as: "What kind of data am I?", *e.g.*, BIAS, "to which group do I belong?", *e.g.*, to a particular Observation Block or observing template. *Data Association* is the process of selecting appropriate calibration data for the reduction of a set of raw science frames. Typically, a set of frames can be associated if they share a number of properties, such as instrument and detector configuration. Since all the required information is stored in the FITS headers, data association is based on a set of header keywords (called "association keywords") and the process is specific to each type of calibration. The process of data classification and association is known as *data organisation*.

An instrument pipeline consists of a set of data processing modules that can be called from different host applications, namely:

- *Reflex* is a graphical tool that helps the user to execute data reduction workflows that contain several recipes. *Reflex* takes care of grouping the different data sets, associating the calibration frames and managing the interdependencies between recipes in the calibration cascade. **Reflex is the recommended software tool for reducing your data.**

- *EsoRex* is a command line tool used to run each pipeline recipe. *EsoRex* commands can be easily scripted.

- The Paranal observatory implements automatic data management tools that trigger the execution of pipeline recipes. This aspect is not covered in this manual.

10.2.1 Using *Reflex*

Reflex is the recommended tool to reduce complete data sets that include all the calibration frames. It is an advanced, yet easy to use, tool geared towards maximum scientific return. It is based on the workflow engine *Kepler*.

Please refer to [8] for the installation procedure which also contains a detailed description of the *Reflex* application. The following provides a very brief summary of how to use Reflex.

Once installed, *Reflex* can be executed with the command:

```
user@host# esoreflex &
```
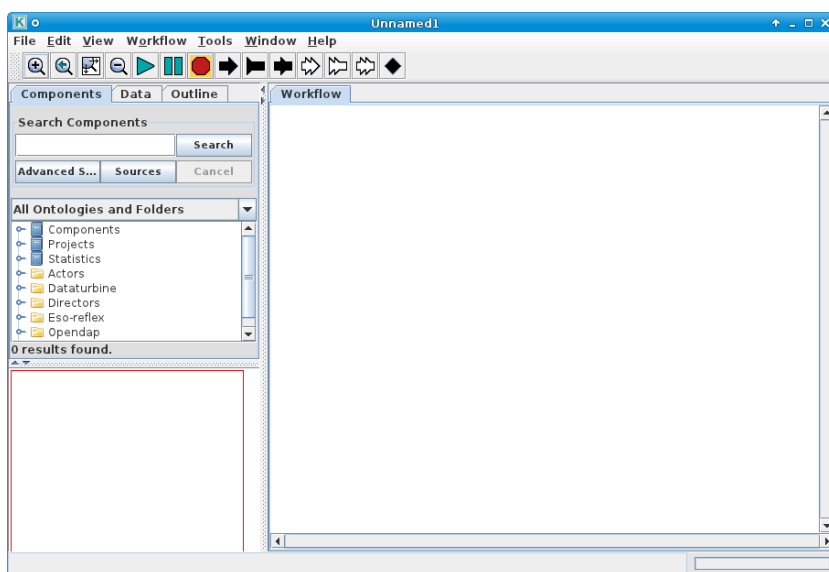

Figure 10.2.1: *Fresh Reflex canvas.*

The main concepts in *Reflex* are workflows and actors. Workflows are canvasses which show the interdependence of the pipeline recipes, allowing the user to easily obtain an overview of the reduction steps. Workflows have the advantage of requiring a small learning curve in order to get the pipeline running.

Actors are entities that perform operations. *Reflex* combines actors that correspond to pipeline recipes that perform data reduction steps, and actors useful for managing data files, such as the DataOrganizer and the FitsRouter. Each actor can be configured by right-clicking on it and selecting *Configure Actor* as shown in Figure 10.2.2. In the case of the recipe actors, the recipe parameters are part of the actor and make up the second group of parameters.
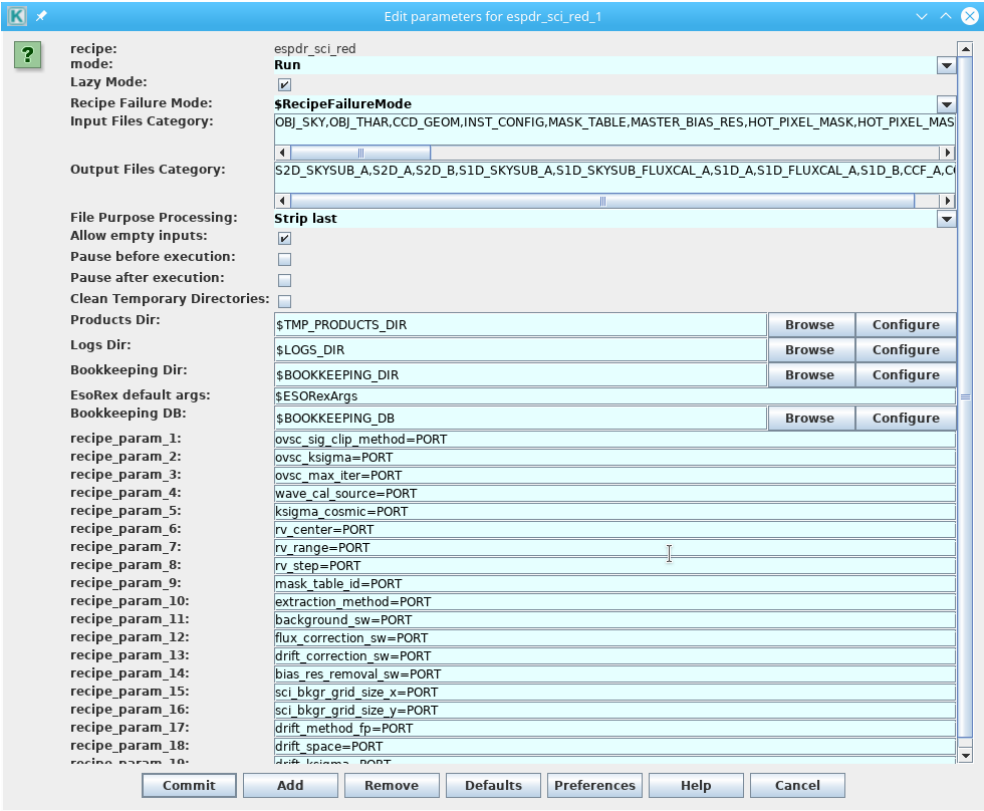
Figure 10.2.2: *Parameters of a recipe actor.*

In addition to those elements, the workflow contains variables that contain the most important settings, such as the directories where data is located and will be saved.

10.3   Quick Start: Reducing The Demo Data

We describe the steps to reduce the science data provided in the ESPRESSO demo data set supplied with the `Reflex 2.11.5` release. By following these steps, the user should have enough information to perform a reduction of his/her own data:

1. Start the `Reflex` application:

        esoreflex &

    If `install_esoreflex` was used or manual installation was performed then the start command is:
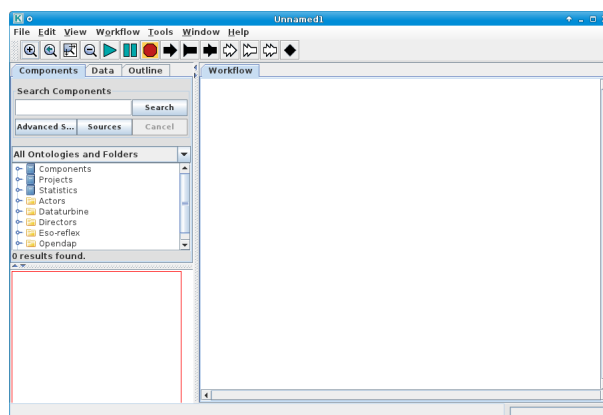
        <install_dir>/bin/esoreflex &

Figure 10.3.1: *The empty* Reflex *canvas.*

The empty Reflex canvas as shown in Figure 10.3.1 will appear.

2. Now open the HARPS, workflow by clicking on File -> Open File, selecting first espdr-3.3.12 and then the file harps in the file browser. You will be presented with the workflow canvas shown in Figure 10.3.2. Note that the workflow will appear as a canvas in a new window.

3. To aid in the visual tracking of the reduction cascade, it is advisable to use component (or actor) highlighting. Click on Tools -> Animate at Runtime, enter the number of milliseconds representing the animation interval (100 ms is recommended), and click OK.

4. Under "Setup Directories" in the workflow canvas there are seven parameters that specify important directories (green dots). Changing the value of ROOT_DATA_DIR and/or RAW_DATA_DIR is the only necessary modification if you want to process data other than the demo data[10], since the value of this parameter specifies the working directory within which the other directories are organised. Double-click on the parameter ROOT_DATA_DIR and a pop-up window will appear allowing you to modify the directory string, which you may either edit directly, or use the Browse button to select the directory from a file browser. When you have finished, click OK to save your changes.

5. Click the ▷ button to start the workflow

6. The workflow will highlight the Data Organiser actor which recursively scans the raw data directory (specified by the parameter RAW_DATA_DIR under "Setup Directories" in the workflow canvas) and constructs the DataSets. Note that the raw and static calibration data must be present either in RAW_DATA_DIR or in CALIB_DATA_DIR, otherwise DataSets may be incomplete and cannot be processed. Please also note that if the same reference file was downloaded twice to different places this creates a problem as Reflex cannot decide which one to use.

7. The Data Set Chooser actor will be highlighted next and will display a "Select Datasets" window

---

[10]If you used the install script install_esoreflex, then the value of the parameter ROOT_DATA_DIR will already be set correctly to the directory where the demo data was downloaded.

(see Figure 10.3.3) that lists the DataSets along with the values of a selection of useful header keywords[11]. The first column consists of a set of tick boxes which allow the user to select the DataSets to be processed. By default all complete DataSets which have not yet been reduced will be selected.

8. Click the `Continue` button and watch the progress of the workflow by following the red highlighting of the actors. A window will show which DataSet is currently being processed.

9. Once the reduction of all DataSets has finished, a pop-up window called *Product Explorer* will appear, showing the datasets which have been reduced together with the list of final products. This actor allows the user to inspect the final data products, as well as to search and inspect the input data used to create any of the products of the workflow. Figure 10.3.4 shows the Product Explorer window.

10. After the workflow has finished, all the products from all the DataSets can be found in a directory under `END_PRODUCTS_DIR` with the named with the workflow start timestamp. Further subdirectories will be found with the name of each DataSet.

11. Here we have described what a user should do to reduce data different from the ones of the demo data. If a user runs the workflow using the demo data, there is no need to change "Setup Directories" (step 4).

Well done! You have successfully completed the quick start section and you should be able to use this knowledge to reduce your own data. However, there are many interesting features of `Reflex` and the ESPRESSO workflow that merit a look at the rest of this tutorial.

---

[11]The keywords listed can be changed by right-clicking on the `DataOrganiser` Actor, selecting `Configure Actor`, and then changing the list of keywords in the second line of the pop-up window.
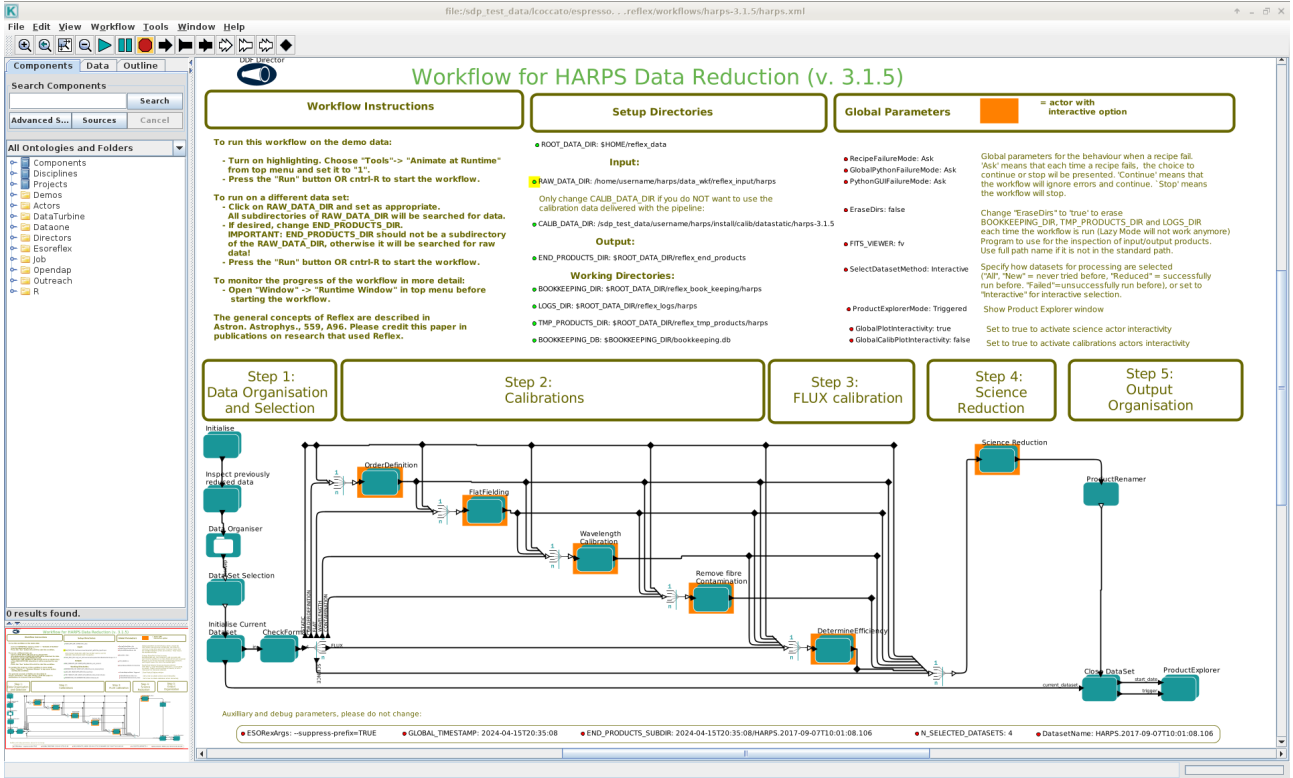
Figure 10.3.2: *HARPS workflow general layout.*

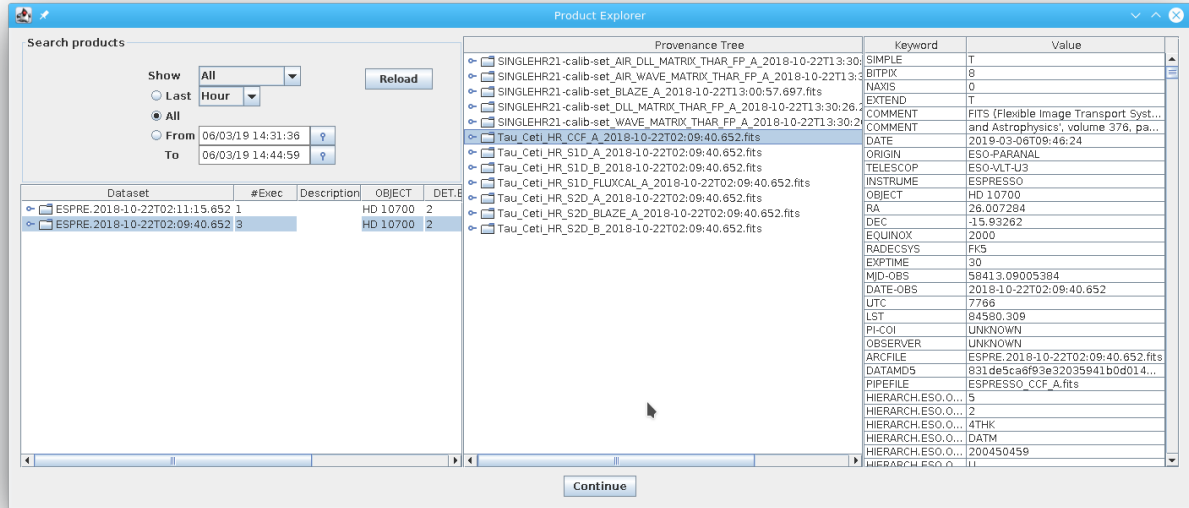Figure 10.3.3: *The "Select Datasets" pop-up window.*

Figure 10.3.4: *The Provenance Explorer shows all datasets reduced in previous executions together with the full reduction chain for all the pipeline products.*

10.3.1 Using EsoRex

EsoRex is a command line utility for running pipeline recipes. Users may embed EsoRex into data reduction scripts. However, with *EsoRex* users must manually classify and associate the data using the information contained in the FITS header keywords (see Section 6, page 18). The user has to define the input set-of-frames (SOF) and potentially set configuration parameters.

The set-of-frames: Each pipeline recipe is executed for a set of input FITS data files. When using *EsoRex* the filenames must be listed together with their Data Organiser (DO) category [12] in an ASCII file, the *set-of-frames* (SOF), that is required when launching a recipe. A Set Of Frames (SOF) is a an ASCIII text file that lists each input frame (full path), specifying its DO category. Here is an example of SOF, valid for the *espdr_orderdef* recipe

```
/file_path/HARPS.2018-09-02T10:24:23.827.fits ORDERDEF_A
/file_path/HARPS.2018-09-02T10:25:30.191.fits ORDERDEF_B

/file_path/HARPS_SINGLEHR_1x1_inst_config_2018-10-14.fits INST_CONFIG
/file_path/HARPS_1x1_CCD_geom_config.fits CCD_GEOM

/file_path/HARPS_hot_pixels.fits HOT_PIXEL_MASK
/file_path/HARPS_bad_pixels.fits BAD_PIXEL_MASK
```

---

[12]The indicated *Data Organizer category* is a label assigned to any data type after it has been classified, which is then used to identify the frames listed in the *set-of-frames*

Note that the HARPS pipeline recipes do not verify in any way the correctness of the classification tags specified by the user in the SOF. In the above example, the recipe *espdr_orderdef* will treat the frame `/file_path/HARPS.2018-09-02T10:24:23.827.fits` as a ORDERDEF_A, the frame `/file_path/HARPS_hot_pixels.fits` as a hot pixel mask, etc., even when they do not contain this type of data. The recipe will also assume that all frames are associated correctly, *i.e.*, that they all come from the same instrument mode, and detector bin setting, and that the appropriate calibration files have been specified.

EsoRex syntax: esorex [esorex_options] recipe_name [recipe_options] set_of_frames

To get more information on how to customise ESOREX (see also [5]) run the command:

esorex - -help

To generate a configuration file esorex.rc in the directory $HOME/.esorex run the command:

esorex - -create-config

A list of all available recipes, each with a one-line description, can be obtained using the command:

esorex - -recipes

All recipe parameters (aliases) and their default values can be displayed by the command

esorex - -params recipe_name

To get a brief description of each parameter meaning execute the command:

esorex - -help recipe_name

To get more details about the given recipe give the command at the shell prompt:

esorex - -man-page recipe_name

Recipe configuration file: Each pipeline recipe can be controlled by specifying its configuration parameters. One way is to define them on command line. Another way is using the ESOReflex interface. Alternatively parameters can be defined through a recipe configuration file. This is an ASCII file listing parameters, their values and and their description. Recipe configuration files are normally (by default) generated in the directory `$HOME/.esorex`, and have the same name as the recipe to which they are related, with the filename extension `.rc`. For instance, the recipe *espdr_sci_red* has its *EsoRex* generated configuration file named `espdr_sci_red.rc`, and is generated with the command:

esorex - -create-config espdr_sci_red

The definition of one parameter of a recipe may look like this:

```
# --rv_center
# Approximate RV.
espdr.espdr_sci_red.rv_center=0.0
```

In this example, the parameter `espdr.espdr_sci_red.rv_center` is set to the value `0.0`. In the configuration file generated by *EsoRex*, one or more comment lines describing the parameter parameter purpose are added, and an alias that could be used as a command line option.

The recipes provided by the HARPS pipeline are designed to implement a cascade of macro data reduction steps, each controlled by its own parameters. For this reason and to prevent parameter name clashes we

specify as parameter prefix not only the instrument name but also the name of the recipe they refer to. Shorter parameter aliases are made available for use on the command line.

The command

esorex - -create-config recipe_name

generates a default configuration file **recipe_name.rc** in the directory **$HOME/.esorex**[13].

A recipe configuration file different from the default one can be specified on the command line:

esorex - -recipe-config=my_alternative_recipe_config

The reader can find a complete description of all recipe parameters in section 9. Their purpose is described in Section 11.

For a configuration file to be used as configuration file, the file should be explicitly defined with the syntax above.

Recipe execution: A recipe can be run by specifying its name to *EsoRex*, together with the name of a set-of-frames. For instance, the following command line would be used to run the recipe *espdr_sci_red* for processing the files specified in the set-of-frames `espdr_sci_red.sof`:

esorex espdr_sci_red espdr_sci_red.sof

The recipe parameters can be modified either by editing directly the used configuration file, or by specifying new parameter values on the command line using the command line options defined for this purpose. Such command line options should be inserted after the recipe name and before the SOF name, and they will supersede the system defaults and/or the configuration file settings. For instance, to set the *espdr_sci_red* recipe *rv_center* parameter to `3.0`, the following should be typed:

esorex espdr_sci_red - -rv_center=3.0 espdr_sci_red.sof

For more information on *EsoRex*, see www.eso.org/cpl/esorex.html.

_____

[13]If a number of recipe parameters are specified on the command line, the given values will be used in the created configuration file.

11 Algorithms and recipe details

In this section we describe the main algorithms implemented in the ESPRESSO pipeline recipes. Relevant data reduction parameters are typed in **bold** face. The full name of the parameter **bias_ksigma** of the recipe `espdr_mbias` is actually

`espdr.espdr_mbias.bias_ksigma`. For convenience we omit the common prefix `espdr.espdr_mbias.` for parameter description, the common prefix `espdr_` for the recipe description as well as the step prefix name for the algorithm description.

11.1 Algorithms

11.1.1 Master frame combination using a kappa-sigma-clipped mean/median

A kappa-sigma-clipped mean or median may be used on a per-pixel basis to combine a set of frames into a master frame, thereby excluding outlier pixel values, i.e. the cosmics. In this technique, the median gives more accurate results. The user may set the method (**sig_clip_method**): mean or median, and the kappa value (**ksigma**) controlling thresholds used to clip outliers. This algorithm is used in all recipes that do frame stacking: mbias, mdark, led-ff and mflat.

11.1.2 Order definition

Orders and slices are identified on ORDERDEF frames through a clumping algorithm that groups together neighbouring pixels whose values exceed a given flux threshold. The algorithm acts recursively on all pixels meeting the flux criterion, within a given distance, until all pixels have been classified either as belonging to a clump or as background.

Clumps are then screened against several criteria to establish whether they represent valid orders/slices: minimum length in main dispersion direction and minimum width in cross-dispersion direction. The clumps that do not meet the criteria are discarded.

After that, orders and slices are assigned physical interference orders. This numbering is established by using input parameters specifying the number of slices per order, the physical number of the first detected order, the number of valid orders to be found, and the typical inter-slice distance on the detector.

Finally, order shapes and positions in cross-dispersion direction are fitted at several locations along the main dispersion using an appropriate analytical model (Gaussian or sum of Gaussians). Note that the Gaussian or multi-Gaussian fit is only needed to approximately measure order center. Then, a low-order polynomial is fitted to the Y position of each order as a function of X, where X is the main dispersion direction and Y is the cross-dispersion direction. Only the center of the order in Y is relevant at this stage, not their shape in the Y direction. The only purpose of these polynomials is to provide order center vs. X position to 0.5 pixel precision in order to define the window of extraction that will be used by the subsequent spectrum extraction procedure.

11.1.3 Background subtraction

The strategy to measure the diffuse background light on the detector is the following:

- Divide the detector into a grid of small regions (100x100 pixels in size)

- In each region, excluding the pixels belonging to spectral orders, build an histogram of (background) pixel values.

- Estimate the mode of the pixel value distribution by fitting a second-order polynomial to the histogram bins closest to the peak of the distribution. The mode of the distribution is taken as the best approximation of the local background level.

- Interpolate the background levels measured in each region over the whole detector using cubic splines. These cubic splines have low order or few nodes, i.e., are smoothly varying over the detector.

A background map is then generated and subtracted from the raw frame.

11.1.4 Spectrum Extraction

As of today, all major fiber-fed spectrographs designed for precise radial velocity work produce relatively symmetric, round slit images on the detector (e.g. ELODIE, CORALIE, HARPS, SOPHIE). This results from the use of circular fibres and simple light injection with no image or pupil slicing at spectrograph entrance. Moreover, the curvature of echelle orders on the detector is generally small. The dispersion direction is constant and along the lines, or, in other words, the slit tilt is zero accross the wavelength range. In this case, spectrum extraction can be made by simply summing the flux in cross-dispersion direction, parallel to the CCD rows/columns. The resulting S2D spectrum is sampled on the grid of extracted pixels. Note that this extraction process has the advantage of avoiding any resampling procedure.

In the case of ESPRESSO, the situation is more complex since pupil slicing and anamorphic magnification are used to achieve high spectral resolution while using a grating with the same size as that of HARPS. The slit image on the detector is made of two ellipsoidal slices aligned in cross-dispersion direction. Depending on the exact geometry of the slit images, image quality and desired spectral resolution, it may or may not be possible to extract the spectrum by summing along the CCD rows/columns in cross-dispersion direction. ESPRESSO was designed such that the Anamorphic Pupil Slicer Unit (APSU) can be optimised to obtain a slit image tilt close to zero and ensure a precise alignement of the two slices along CCD rows/columns. The residual tilts are only a few deg and slice misalignments never exceed 2-3 microns anywhere on the detector, i.e. much less than the size of a pixel (10 microns). Given the minimum sampling of 2 pixels per FWHM in the ultra-high resolution mode, this means that a simple extraction procedure parallel to CCD rows/columns will decrease the spectral resolution by only a few percent in the worst case, i.e. at CCD edges. We consider this as fully acceptable given the simplicity and advantages of the extraction along CCD rows/columns (e.g. no resampling).

Each slice is extracted separately and treated as if it was an independent interference order. This results in each physical order being represented twice in the S2D extracted spectrum.

Two spectrum extraction functions are provided in the ESPRESSO DRL: simple summation, and optimal extraction following the well-known Horne method (Horne 1986, PASP 98, 609). We investigated refinements introduced by e.g. Marsh (1989) and Donati et al. (1997), but they are generally not relevant in the case of ESPRESSO where the slit tilt on the CCD is negligible. Nevertheless, the Horne optimal extraction algorithm has been modified significantly and optimised for ESPRESSO in the following ways:

- Order profiles in cross-dispersion direction are obtained directly from spectral flat-field exposures at high SNR, which serve as empirical models. Such exposures are part of the standard calibration sets and must be obtained within 24 hours of science observations. We make use here of the high stability of ESPRESSO which guarantees that the position and shape of spectral orders remain essentially constant between calibrations and science exposures.

- If the number of hot and bad pixels does not exceed a significant fraction (0.25) of the pixels in the extraction window, hot and bad pixels are corrected by interpolating the order profile across them before extracting the spectra. In this way the knowledge and smoothness of the order profile can be used to optimally correct for these CCD defects.

- For exposures taken with simultaneous reference on fiber B (ThAr/LFC/FP), a *contamination frame* is used as additional empirical model in the optimal extraction. This means that the spectrum is modelled as the sum of a scaled order profile and a scaled contamination profile. The obtained science flux is therefore corrected from direct ThAr/comb/FP contamination. As for the *usual* optimal extraction, the model is linear in both model parameters (science and contamination fluxes), and the solution is thus obtained by analytically solving the associated linear least-squares problem.

11.1.5 Flat-Fielding

The extraction procedure described above delivers a spectrum that is still affected by variable pixel-to-pixel sensitivity along the main dispersion direction. To correct for that, the spectrum is divided by the normalised extracted flat-field, obtained from co-added FLAT frames and extracted in exactly the same way as the science spectrum.

We note that this flat-fielding procedure will remove not only pixel-to-pixel sensitivity variations, but also fringing effects, at least when the science spectrum is broadly similar to the one of the flat-field lamp. This is the case for all stellar spectra at moderate to high SNR. In the case of a very low SNR spectrum dominated by sky emission lines, the correction of fringing may in principle be less accurate. However, fringing by sky lines is expected to be a very small effect at the high spectral resolution of ESPRESSO. The calibration images of ESPRESSO reveal no fringing at all.

11.1.6 Wavelength Calibration

There are two methods used for the wavelength calibration, according to the calibration source used: ThAr lamp together with Fabry-Perot, or Laser Frequency Comb.

ThAr and Fabry-Perot Lamps

First the FP peaks are identified on the frame with both fibres illuminated with the Fabry-Perot lamp. Then the tables of these peaks are passed to the reduction of the frames in which one of the fibres is illuminated with the ThAr lamp, while the other fibre has the Fabry-Perot light for the drift calculation. Combining the detected FP peaks and the carefully chosen ThAr lines with precise wavelengths assigned, allows to compute the precise wavelength solution per extracted order. The wavelength solution fit is done on the detected Fabry-Perot peaks, which are anchored to the precise wavelength through the known ThAr lines.

ThAr spectra are characterised by a high density of emission lines with many blends and very different intensities. Identifying and properly fitting all these lines is a challenging task, especially when aiming at a repeatability of better than 1 m/s in the wavelength calibration. Variable line intensities make it necessary to use a static ThAr table to initiate the wavelength calibration process. This table indicates the approximate position of lines. Each ESPRESSO mode has its own ThAr lines table.

ThAr lamps evolve with time, and the ageing process has several consequences on their spectra. Both line intensities and wavelengths are affected. To mitigate this, a line-by-line comparison with the static ThAr table is performed to identify and reject individual lines potentially affected by fitting instabilities.

Laser Frequency Comb

Giving the previously computed wavelength solution, the Laser Frequency Comb allows to achieve a more stable and precise solution. No static table is needed to initiate the wavelength calibration process, since emission LFC lines can be searched and identified automatically thanks to the absolute accuracy of the comb pattern. Then, the wavelength of each peak is determined. To do this, the DRL first computes a list of theoretical LFC wavelengths based on the known repetition rate and anchor frequency of the LFC. Then, these wavelengths are assigned to individual peaks by comparison with a previously-obtained ThAr wavelength calibration, which provides a suitable first guess for the wavelength of the LFC lines. Since the LFC does not cover the whole spectral range, the orders with insufficient number of detected LFC lines (less than the limit fixed in the instrument configuration) have to take the solution from the ThAr method. Also, the orders, for which the LFC solution has the RMS or the CHI2 higher than limits fixed in the instrument configuration, have the ThAr wavelength solution assigned. The FITS header keywords indicate for each order where the wavelength solution is coming from. The LC solution is provided for the wavelengths from 5020 A to 7344 A, i.e. orders 81 to 158 for SINGLE modes or 41 to 78 for MULTI modes.

11.1.7 Instrumental Drift Measurement

The simultaneous reference fibre of ESPRESSO is used either for sky subtraction or instrumental drift correction. Instrumental drift is measured using wavelength calibration sources, either the ThAr lamp, the LFC or a stabilised Fabry-Perot etalon (FP). Drift is measured by comparing two similar S2D spectra on the simultaneous reference fibre: one that is acquired at the time of the wavelength calibration (with both fibre A and fibre B illuminated by a calibrator), and one that is acquired simultaneously with the science observation. The algorithm used by the pipeline to compute the drift is based on the method described in Connes (1985), ApSS 110, 211, and Bouchy et al. (2001), A&A 374, 733. The drift is simply obtained from a first-order Taylor expansion of the spectrum at each individual pixel, i.e. the measured flux difference between the two spectra at pixel X is converted into a position difference using the measured spectrum derivative at pixel X. This obviously assumes that the spectrum derivative is constant over a length X that is the shift between the two spectra. Therefore, this method only works for drifts that are small relative to the FWHM of the spectral lines, i.e. typically smaller than one pixel. The pixel size in ESPRESSO is equal to 500 m/s, while the expected instrumental drifts are smaller than 1 m/s over 24 hours. The method described here is thus applicable for corrections in between daily wavelength calibrations. It has the significant advantage that no modelling of the spectrum is needed to measure the drift, since each individual pixel independently contributes to the total drift. Note that pixels are weighted according to their Doppler information content, which is proportional to the square of the spectrum derivative.

This algorithm has been successfully used on HARPS to measure drifts of ThAr spectra. For ESPRESSO a few improvements were made to the method to adapt it to LFC/FP spectra and make it more robust in general. The

following aspects are addressed:

- Flux normalisation: the method only works if the two spectra are normalised to the same flux level. Thus a global flux normalisation on the S2D spectrum is performed.

- Local vs. global drift measurement: instrumental drifts may vary across the spectral format, i.e. a single, global value for the drift may be too coarse an approximation. It is possible to specify to the algorithm whether the drift computation and correction should be global, chip-by-chip or order-by-order.

With the algorithm described here, the instrumental drifts can be reliably measured to a precision of 1-2 cm/s globally (if not limited by photon noise).

11.1.8 Sky Subtraction

Sky subtraction in ESPRESSO can be made by observing in simultaneous sky mode and subtracting the sky spectrum obtained on fibre B from the science spectrum on fibre A. In practice, this process involves several steps:

- Measure the relative throughput of fibre B with respect to fibre A as a function of wavelength (in S2D format) by acquiring EFF_AB exposures, which are obtained by pointing the telescope at daylight sky. Note that the sky fills the entire fibre aperture uniformly, so that pointing/guiding errors or telescope pointing jitter have no effect on the amount of sky flux entering the spectrograph.

- Scale the S2D simultaneous sky spectrum using the previously-obtained relative efficiency curve.

- Rebin the scaled S2D sky spectrum to the same wavelength scale as fibre A. The resampling is done by spline interpolation of the cumulative flux vs. extracted pixel, ensuring flux conservation. Note that the wavelength scales of fibres A and B are extremely similar, so that the resampling will essentially consist of a small shift of the spectrum that is slowly varying along spectral orders.

- Subtract the scaled and rebinned S2D simultaneous sky spectrum from the S2D science spectrum.

For ESPRESSO two sky subtraction methods are available: *pixel-by-pixel* and *smoothed*. The first method subtracts the sky spectrum on a pixel-by-pixel basis. The second one performs first a sliding average of the sky spectrum and then subtracts if from the science spectrum. For very low S/N regime, readout limited or close to it, it is recommended to use the option smooth, thus reducing the readout noise contribution from the step of sky subtraction.

The precision of sky subtraction with ESPRESSO is essentially photon- and readout-noise limited, provided the sky fibre *sees* the same sky as the science fibre.

11.1.9 Barycentric Correction

In the context of high-resolution astronomical spectroscopy and high-precision radial velocity measurements, it is necessary to compute to high accuracy the projection of the velocity vector of an Earth-bound observer

along the line of sight to an astronomical target, at the time of the observation, as measured within the ICRS reference frame (centred on and at rest with respect to the Solar System barycenter). The computation should take into account both the Earth's orbital motion around the Sun and the observer's motion due to Earth rotation. For ESPRESSO a new code for this computation was developed, based on recent Solar System ephemerides produced at Institut de Mécanique Céleste et de Calcul des Ephémérides (IMCCE) in Paris.

The algorithm:

- Computes the observer's projected velocity in the direction of the target in the ICRS reference frame at the time of the observation with the accuracy of at least 1 cm/s.

- Computes the barycentric time of light arrival, i.e. the time at which the light from the target reaches the Solar System barycenter given its detection by the Earth-bound observer at the time of the observation with the accuracy of at least 1s.

- Computes an upper bound to the maximum value of the barycentric correction over one year for any given target.

Input parameters:

- Right ascension of target in decimal hours, ICRS system, epoch 2000.0

- Declination of target in decimal degrees, ICRS system, epoch 2000.0

- Proper motion of target in right ascension (mu_alpha cos delta) in arcsec/year, ICRS system

- Proper motion of target in declination in arcsec/year, ICRS system

- Date of observation in UTC (year, month, day, decimal hour)

- Longitude of observer in decimal degrees

- Latitude of observer in decimal degrees

- Altitude of observer above sea level in km

Output parameters:

- Velocity component of observer projected onto target line of sight in km/s, ICRS system (abridged as BERV for Barycentric Earth Radial Velocity)

- Total velocity of observer in km/s, ICRS system

- Barycentric time of light arrival as TDB Julian date

- Yearly maximum value of BERV (upper bound)

Step-by-step description:

- Compute Julian date from given date and time of observation

- Compute ICRS target coordinates at the time of observation (i.e. take proper motion into account)

- Compute observer's velocity vector with respect to geocenter at the time of observation, ICRS system (correct for precession and nutation)

- Compute Earth's orbital velocity vector at the time of observation, ICRS system

- Combine rotational and orbital motion into total velocity vector

- Project velocity vector onto target line of sight to obtain BERV

- Compute barycentric time of light arrival

- Compute yearly upper bound to BERV

11.1.10 Order Rebinning and Merging

The rebinning and merging process to create a S1D spectrum from the S2D spectrum is as follows:

- Define a uniform wavelength grid with a constant wavelength step that is close to the average pixel size in wavelength units.

- For each spectral order, build the cumulative flux distribution vs. extracted pixel.

- Interpolate the cumulative flux function onto the uniform wavelength grid using cubic splines.

- Build the resampled spectrum by differentiating the interpolated cumulative flux.

- Merge resampled spectral orders, computing the weighted average of rebinned pixels where spectral orders overlap.

This technique is chosen because it conserves the integrated flux within any two wavelengths of the original pixel grid. Moreover, cubic splines are well suited to ESPRESSO spectra because of the well-sampled PSF of the instrument (4 pixels FWHM in singleHR mode), which ensures that the spectrum derivatives can be numerically estimated in a reliable way.

11.1.11 Radial Velocity Computation

The ESPRESSO DRL implements a cross-correlation module that computes the cross-correlation function (CCF) of a S2D spectrum with respect to a binary template (mask) of a given spectral type. The radial velocity (RV) is then obtained from a Gaussian fit to the CCF. This technique has been successfully used on the ELODIE, CORALIE, HARPS, SOPHIE and HARPS-N spectrographs (see Baranne et al. 1996, A&AS 119, 373, and Pepe et al. 2002, A&A 388, 632). One of its main advantages is that CCFs can be computed in an automatic way using a line mask. Line masks are simply lists of central wavelengths and depths of spectral lines, and can be created for various spectral types. We note here that the CCFs of slowly-rotating stars are extremely well approximated by Gaussian profiles with a flat continuum. However, the particular fitting function does not matter much; the crucial aspect is to fit the CCFs of a given star with always the same function to avoid

systematic effects on the derived radial velocities. As such, it is fundamental to reduce all the spectra with the same mask and CCF parameters, as comparison between reductions with different parameters will be affected by systematics.

The main steps of the algorithm are:

- Compare the global flux distribution in the S2D spectrum to a static flux template that approximately corresponds to the spectral type of the star. The S2D flux is scaled accordingly to match the flux distribution of the template. In this way, spectra of any given star are always brought to the same flux distribution, which ensures that variable atmospheric conditions will not induce systematic effects in the CCF computation.

- Shift the wavelength scale of the S2D spectrum to the Solar System barycenter using the barycentric correction.

- Define a uniform radial velocity grid that is approximately centered on the radial velocity of the star.

- For a given RV value in the grid, shift the line mask by the corresponding Doppler shift, project the line mask onto the S2D spectrum using a specified line width (about one pixel), and sum the S2D flux that goes through the so-defined mask *holes*. The flux from partial pixels is computed via simple linear interpolation. The sum is actually a weighted sum, using line depths as weights to optimally extract the Doppler information. During this process, the S2D spectrum is locally blaze-corrected to remove any continuum slope around spectral lines. This produces one point of the CCF.

- Loop over all RV values in the grid.

- Fit a Gaussian profile to the CCF to derive RV, FWHM and contrast.

Note that, by construction, CCFs are simply co-added spectral lines in velocity space, weighted by their depth and continuum flux and corrected of the transmission of the grating (blaze function). As such they can be considered as a *master* spectral line for the star.

11.1.12 Removal of cosmic rays

.

Cosmic rays are identified via a ksigma rejection when extracting the one-dimensional spectrum of one order. In each recipe that does it, the clipping is regulated by a dedicated parameter. In the science recipe, `espr_sci_red` the clipping is regulated by the parameter **ksigma_cosmic**. A value of -1 turns off the ksigma clipping.

The science recipe has also an additional algorithm to detect cosmic rays, that exploits LACosmic (van Dokkum 2001, PASP, 113, 1420). The algorithm is applied to the so-called detector cleaned frames, i.e. the raw science where the overscan regions have been trimmed and their contribution subtracted. Cosmics identified this way are masked during the ectraction of the one-dimensional spectra.

Both algorithms can be run simultaneously.

The LA Cosmic algorithm is regulated by the following paramters: **cosmic_detection_sw**.

- **cosmic_detection_sw**. It turns LA Cosmic on (1) or off (0). By default, its value is taken by the input instrument configuration table. The algorithm is turned on for SKY-mode and it is turned off for FP mode observations.

- **lacosmic.post-filter-mode** : dilage/dilute. It specifies whether cosmic ray mask has to be dilated (i.e. expanded) or diluted (i.e. shrinked). It takes effect only if the post-filter-x/y values are positive.

- **lacosmic.post-filter-x** : X Size of the post filtering kernel. The mask is dilated or diluted by this amount along x direction. Defaut: 0.

- **lacosmic.post-filter-y** : Y Size of the post filtering kernel. The mask is dilated or diluted by this amount along y direction. Default: 0.

- **lacosmic.sigma_lim** : LA Cosmic Poisson fluctuation threshold. It is the minimum value of the fluctuation image, obtained by dividing the Laplacian image (a second order derivative of the original image along x and y) and the noise model. High values of **sigma_lim** find the more intense cosmics and have less risk to detect false positives. Low values are more efficient in finding also faint cosmics, but have higher risk to detect false positives.

  The default values are taken by the input instrument configuration table: 5 (SHR-SKY, 2x1), 7 (SHR-FP, 2x1), 7 (SHR 4x2), 8 (SUHR 1x1), 10 (MHR 4x2), 8 (MHR 8x6).

- **lacosmic.f_lim** : LA Cosmic minimum contrast between the Laplacian image (see above) and the fine-structure image (created from the original image by a combination of median filters). High values of **f_lim** find the more intense cosmics and have less risk to detect false positives. Low values are more efficient in finding also faint cosmics, but have higher risk to detect false positives.

  The default values are taken by the input instrument configuration table: 5 (SHR-SKY, 2x1), 5 (SHR-FP, 2x1), 5 (SHR 4x2), 8 (SUHR 1x1), 4 (MHR 4x2), 5 (MHR 8x6).

- **lacosmic.max_iter** : LA Cosmic max number of iterations. [5]

- **extra_products_sw** : Set to TRUE to create extra products to inspect LA Cosmic results. [FALSE]

11.1.13 Error and Bad Pixel Propagation

S2D Spectra

All significant error sources are propagated through the pipeline using the standard error propagation formulae. We list below the noise sources that are taken into account in the science reduction cascade producing the S2D spectra:

- Detector readout noise

- Detector dark current noise

- Total photon noise on science fiber (science target + diffuse background)

- Flat-fielding noise

- Sky noise if applicable (simultaneous sky mode)

Error propagation for S2D spectra is relatively straightforward since there is no resampling step in the process. Extracted pixels remain independent of each other.

If not too many, bad pixels and hot pixels will be ignored during the spectrum extraction process. The empirical, high-SNR extraction profile will be interpolated across those pixels, so that the extracted science spectrum is effectively corrected. The presence of bad/hot pixels will be reported in the quality maps associated to science products; they are therefore traceable throughout the reduction chain.

S1D Spectra

The resampling process that is necessary to generate S1D spectra from S2D spectra inevitably introduces correlations between adjacent rebinned pixels in the S1D spectra. To calculate the errors on the S1D spectra one propagates errors in the usual way to the S1D spectrum, while including in the error computation an additional array containing the correlation factor of each rebinned pixel. This factor is the quadratic sum of the relative contributions of the original pixels to the rebinned pixel, considering that a maximum of two original pixels contribute to the rebinned pixel. In this way we believe all the information on noise properties is conveyed into the S1D spectrum. We note however that the existence of correlations makes the use of the S1D spectrum non-trivial for science purposes if a rigorous treatment of the noise is required. It is up to the user to understand how to deal with this issue, e.g. when performing least-squares fitting of models to the S1D spectrum.

In general, the use of the S2D spectrum is recommended whenever possible to avoid correlations between adjacent data points, which are unavoidable as soon as some resampling is performed.

Cross-Correlation Functions

Uncertainties on the CCF data points are obtained by propagation of S2D error maps through the cross-correlation stage, which is a simple additive process (fluxes from many spectral lines are co-added). Finally, an estimate of the radial velocity uncertainty is obtained by converting CCF flux errors into RV errors using the measured CCF derivative, as described in Boisse et al. (2010), A&A 523, 88 (appendix).

11.2   Recipes

An overview of the data reduction cascade is given in Section 8.

11.2.1   orderdef

The recipe detects the orders position in the image. The input frames are two flat images: one for fibre A and one for fibre B. Each input raw frame is checked for the saturation and cleaned from the CCD signature (bias, dark current, gain). A FITS keyword QC CHECK is added and set to 1 in case of correct and 0 for bad quality results. Than the orders detection is performed:

- For each pixel, decide if it belongs to an order or to the background.

- Run the clumps detection algorithm, which marks the pixels belonging to the same clump.

- Detect which of the clumps are valid orders.

- Attribute interference / physical order numbers.

- Fit the valid orders position.

- Mark the orders position on the ORDER_TABLE products.

- Save the coefficients of the fit in the 2nd extension of the ORDER_TABLE products.

11.2.2   mflat

The recipe generates the extracted flat-field, blaze function and the order profile. Recipe steps:

- Check the raw frame for saturation.

- Remove the detector signatures (correct for overscan level).

- Measure and subtract inter-order background.

- Perform optimal extraction of orders for both fibres.

- Compute the blaze function.

- Correct the blaze function for spectral energy distribution of the source and instrumental efficiency by fitting the flux distribution at blaze peak.

- Save the QC KWs in the FITS header of the products: max flux per detector output, background min, max and mean per extension (blue and red), number of removed cosmics per order, SNR per order, RMS of the flat per order, QC CHECKS

- Save the ORDER_PROFILE, the FLAT and the BLAZE for each fibre

11.2.3   cal_contam

The recipe generates a contamination frame and checks the contamination level on science fibre. Recipe steps:

- Check the raw frame for saturation.

- Remove the detector signatures (correct for overscan level).

- Measure and subtract inter-order background.

- Perform optimal extraction of orders for both fibres.

- Correct the flat-field.

- Measure the maximum contamination level in extracted spectrum and compare it with the specified threshold.

- Save the contamination frame with QC KWs.

11.2.4 cal_eff_ab

The recipe computes the relative efficiency between sky and science fibres vs. wavelength. Recipe steps:

- Check the raw frame for saturation.

- Remove the detector signatures.

- Measure and subtract inter-order background.

- Perform optimal extraction of orders for both fibres.

- Correct the flat-field.

- Divide reference spectrum by science spectrum and fit a low-order polynomial across each order.

- Save the relative efficiency frame with QC KWs.

11.2.5 cal_flux

The recipe measures the absolute efficiency curve. Recipe steps:

- Check the raw frame for saturation.

- Remove the detector signatures.

- Measure and subtract inter-order background.

- Perform optimal extraction of orders for both fibres.

- Correct the flat-field.

- Assign wavelength solution to the spectrum.

- Rebin and merge the orders.

- Convert S1D spectrum from detected photoelectrons to physical units (erg/s/cm2/A).

- Measure and retrieve the flux of the spectrophotometric standard star, and compute absolute efficiency at the reference wavelengths.

- Interpolate efficiency measurements onto the S1D wavelength scale using cubic splines.

- Save the absolute efficiency frame with QC KWs.

11.2.6 sci_red

The recipe performs the science reduction. Recipe steps:

- Check the raw frame for saturation.

- Remove the detector signatures.

- Identify cosmic rays.

- Measure and subtract inter-order background.

- Perform optimal extraction of orders for both fibres.

- Correct the flat-field.

- Assign wavelength solution to the spectrum.

- In simultaneous sky mode:

  - Scale and rebin sky spectrum.
  - Subtract sky.

- In simultaneous reference mode:

  - Measure instrumental drift.
  - Correct instrumental drift.

- Compute barycentric correction using FITS header information.

- Correct wavelength solution from BERV, shifting it to the barycenter of the SS.

- Rebin and merge orders.

- Convert S1D flux into physical units.

- Flux calibrate S1D spectrum (with and without sky subtraction) using absolute efficiency curve.

- Measure fibre centring on integrated guiding image and perform quality control based on specified tolerances.

- Measure flux-weighted mid-exposure time on exposure-meter data and perform quality control.

- Filter residual cosmic hits from S2D spectrum.

- Correct flux distribution in S2D spectrum.

- Compute the radial velocity.

- Compute the CCF bisector.

- Save the S2D, S1D and CCF products with QC KWs.

A Installation

ESO pipelines can be installed via several methods, depending on your OS, most of which facilitate easy installation, upgrade and removal. Please see the "ESO Data Reduction Pipelines and Workflow Systems" page (https://www.eso.org/pipelines).

A.1 System Requirements

The processing of ESPRESSO data is very demanding in terms of computing resources. In particular it requires a machine with sufficient memory installed. Less critical but still important is the number of available CPU core sand the amount of available disk space. Because of the memory constraints, the ESPRESSO pipeline is only supported on 64-bit platforms. The recommended platform is a powerful workstation with a recent 64-bit Linux system. The minimum system configuration is:

- 32 GB of memory

- 4 CPU cores (physical cores)

- 1 TB of free disk space

- GCC 8.3.1 (or newer)

We recommend the user to have more resources:

- 64 GB of memory

- 8 or more CPU cores (physical cores)

- 12 TB of free disk space

The peak memory consumption is with the flat recipe, of 12 GB for each data set, in case of 20 input flats. Using Reflex as data processing tool, data sets will be processed sequentially, thus a single user will not require a large amount of memory. More useful can instead be to have a large disk space if a user processes without using interactivity a lot of data. If a user process the data using esorex, for example with a customised system of scripts, it is possible that several esorex based recipes execution in parallel, and consequently memory requirements are higher.

A.2 Installing the Pipeline KIT

The ESPRESSO pipeline with all the required tools is available from www.eso.org/pipelines, and depending on user's platform we recommend respectively to install via RPM (Linux) or MacPorts (macOS) and if the user's platform is not one of the supported Linux or mac OS, but is a recent Linux or Mac OS X, an installer script is available.

However, the recommended target platform for using the ESPRESSO DRS is a 64-bit Linux system, since Mac OS X imposes certain restrictions when it comes to running the pipeline. To install the pipeline unpack the kit in a temporary location, go to the top level directory of the unpacked distribution package and execute the installer script as shown in the following example.

Note: The installation script uses the compiler which is found first in the path! If more than one compiler are installed on the system one should make sure that an appropriate 64-bit compiler will be found first when the installation script is executed!

```
tar -zxf espdr-kit-X.Y.Z.tar.gz
cd espdr-kit-X.Y.Z.tar.gz
./install\_pipeline
```

Then follow the instructions on the screen. Once the script finishes successfully and the path variables have been set, the installation of the ESPRESSO pipeline is complete.

ESPRESSO provides also a Data Analysis Software. This can be installed similarly.

```
tar -zxf espda-kit-X.Y.Z.tar.gz
cd espda-kit-X.Y.Z.tar.gz
./install\_pipeline
```

A.3   Installing the Software including ESOReflex

Installation via RPM or MacPorts is recommended. In case the user platform is not one for which RPM and MacPorts are provided the user may use the install_esoreflex script.

ESOReflex can be installed as:

```
./install_esoreflex
```

Then follow the instructions on the screen (selecting HARPS). Once the script finishes successfully and the path variables have been set, the installation of the HARPS ESOReflex workflow and the pipeline are complete.