

SPHERE DATA CENTER : ACCESS, DATA DESCRIPTION, REDUCTION DETAILS

Introduction and acknowledgements

SPHERE-DC is an observing service provided by the French community that targets users of the ESO instrument SPHERE. We have implemented a data base and data reduction workflow facility pipeline based on the SPHERE ESO pipeline, complemented by additional routines (e.g. improved centering routines, production of SPARTA adaptive optics information, automatic frame selection, analysis routines) to systematically process public data, which are then made available to the community. Reduced public data are available from two entry points, depending on the data level: at the processing center in Grenoble (data levels 0-2) and through the DIVA+ database in Marseille (subset of level 2 and above levels, and also for instruments other than SPHERE). We also perform reduction on request from PI/CoIs of SPHERE observing programs with the same pipeline.

In this document, you will find a general description of the available data and how to access the data, followed by more technical information about the outputs and the reduction routines.

Website: <https://sphere.osug.fr/spip.php?rubrique16&lang=en>

Contact : sphere-dc-req@univ-grenoble-alpes.fr

ACKNOWLEDGEMENTS: your acknowledgements are vital to allow this service to continue

If you use DC-processed data for any publication, please add the following sentence in the acknowledgments

'This work has made use of the the SPHERE Data Centre, jointly operated by OSUG/IPAG (Grenoble), PYTHEAS/LAM/CESAM (Marseille), OCA/Lagrange (Nice), Observatoire de Paris/LESIA (Paris), and Observatoire de Lyon. '

and refer to

Delorme et al. 2017: <http://cdsads.u-strasbg.fr/abs/2017sf2a.conf..347D> **for all SPHERE reduced data**

Galicher et al. 2018 : <http://cdsads.u-strasbg.fr/abs/2018A%26A...615A..92G> **and for P99 and following semesters (i.e. after 2017-04-01)**

Maire et al. 2016 : <https://ui.adsabs.harvard.edu/abs/2016SPIE.9908E..34M/abstract> **for IRDIS and IFS reduced data at level 2 and above**

List of acronyms

AO	Adaptive Optics
col	co-Investigator
DC	Data Center
DIVA	Direct Imaging Virtual Archive
ESO	European Southern Observatory
IFS	Integral Field Spectrograph
IRDIS	InfraRed Dual-band Imager and Spectrograph
IPAG	Institut de Planétologie et d'Astrophysique de Grenoble
LAM	Laboratoire d'Astrophysique de Marseille
OCA	Observatoire de la Côte d'Azur
OSUG	Observatoire des Sciences de l'Univers de Grenoble
PI	Principal Investigator
SPARTA	Standard Platform for Adaptive optics Real Time Applications
SPHERE	SpectroPolarimetric High-contrast Exoplanet REsearch
SPHERE-DC	SPHERE Data Center
ZIMPOL	Zurich Imager and POLarimeter

Introduction and acknowledgements	1
List of acronyms	2
1. How to access reduced data from the SPHERE Data Center	3
1.1 Public data	3
1.1.1 What is available in the SPHERE-DC public releases?	3
1.1.2 Access to data	4
Login in to the SPHERE-DC:	4
Data available at the SPHERE-DC	4
How to efficiently browse data on the SPHERE-DC client	6
Recommended browsing habits to select relevant datasets	7
Transferring data from the SPHERE-DC to the user	8
1.1.3 Access to data available at DIVA+	10
1.2 Proprietary data	11
2. Understanding the data delivered by the SPHERE-DC	11
2.1 Download script directory architecture for SPHERE data	11
2.2 Output description for SPHERE data	12
2.2.1 Reduced calibration files	12
2.2.2 Level 1 reduced data : IFS and IRDIS output description	13
2.2.3 Level 2 reduced data available at the SPHERE-DC	16
Sortframe informations	16
Residual image production	16
2.2.4 Useful header keywords added at the end of DC-processed fits headers	18
2.2.5 Level 2 reduced data available at DIVA+	19
2.3 Output description for NaCO/VLT, Keck, Hubble data	20
3. Reduction Information	20
3.1 IRDIS and IFS summary	20
3.1.1 IRDIS-only steps	20
Star centering	21
In case of suspected issue with centering:	22
3.1.2 IFS-only steps	24
3.1.3 Final steps common to IRDIS and IFS	25
References	25

1. How to access reduced data from the SPHERE Data Center

1.1 Public data

1.1.1 What is available in the SPHERE-DC public releases?

We mostly provide SPHERE reduced data, from all SPHERE instruments (IRDIS, IFS, and ZIMPOL). The most common modes are currently supported. We are reducing progressively each ESO period. PIs are contacted prior to reduction of a given semester to check if they have on-going second epoch observations or similar constraints before releasing their public reduced data. Reduction of SPHERE data is performed by our team at the processing center.

The SPHERE data products correspond to different levels:

- Level 0: reduced calibrations
- Level 1: processed master cubes from science observations, and associated files
- Level 2: final image from science observations, and associated files
- Level 3: products resulting from the analysis of the final images, for example detections (candidates) including their astrometric properties

More details here: <https://sphere.osug.fr/spip.php?article77&lang=en>, including updates on the dates for which reductions are or will be shortly available. The content of the output files corresponding to these levels are described below (Section 2).

In addition, we provide access to reduced survey data from other instruments (NACO/VLT, KECK, HST) which have been processed outside our center in a homogeneous manner (see below).

1.1.2 Access to data

Login in to the SPHERE-DC:

The connection to the SPHERE-DC is done through a java-based client called the SPHERE-DC client:

1. Go to <http://sphere-dc.osug.fr:8080/sphere-server/>
2. Download the zip
3. Unzip It

Once installed, **you can launch the client and connect by clicking on the “Public Login” button** or by using the following login info:

Login: public_user
Password : public_user



Figure 1: Window of the SPHERE-DC client just after log-in

Data available at the SPHERE-DC

Once you are connected to the client, you first need to select the workspace: in practice, public users have access to “calibrations_reduites” (to get reduced calibrations) and “PUBLIC_reduced_science” (to get reduced science data!) only as relevant choices. Then you can browse data as follows.

Level 0:

SPHERE calibrations are currently automatically retrieved from the ESO archives (every 24 hours), reduced (after 7pm, CET) and put online. They are then available to all. Click on “Data Browse”, which allows you to access all available reduced calibrations

Level 1 and 2:

The reduced science data are in the workspace “PUBLIC_reduced_science”. You have two ways to browse available data :

- Similarly, click on “Data Browse” to access all available *individual* files
- Click on “Process Browse” to access outputs coming from each process (results of one routine on a given data set, so this usually provides a self-consistent set of data): they may include *several files*, described in more detail below.

How to efficiently browse data on the SPHERE-DC client

You can filter data similarly to what you can do on the ESO archive, just always think to click on the « submit » button after doing your selection.

For instance, if you want to see all the IRDIS reduced master cubes available for GJ285 during the night of 2015-12-25 (which, like all DC nights goes from 2015-12-25 at 20:00 UT to 2015-12-26 at 19:59 UT, so that the morning calibrations are within the same night as their observation) you need to:

1. set the « Observation night » to 2015-12-25
2. set the « frametype » to IRD_SCIENCE_REDUCED_MASTER_CUBE
3. set the « Target » to "GJ_285" or "GJ 285" (star names are matched using various catalogs).

If you want to see all reduced IFS and IRDIS data for this target you leave the « frametype » empty and you put the « raw/reduced » tab to « reduced ». You can see more information by just hovering your mouse on any field in the « filtering » tab to get dedicated info on this filtering parameter. Also, for most parameters, you can enter partial information and the DC will propose the relevant completion, so it is better to look for GJ_285 by typing "285" and selecting GJ_285 rather than typing GJ285, without the "_" and getting nothing that matches! There is one field "Object" which comes from the fits header, hence that was manually written in the OB. Some typos can have been made during the observation, so it is a better practice to use the box "Target", which will look in the database for the Simbad resolved object name. Note that you can only see files for which your user has access to.

The DC interface shows by default only the 1000 first results. If your selection leads to more than 1000 results, you need to use the right click button on the list of results (followed by a click on "Load next") to show the 1000 next files. It might be more convenient to work with smaller selections, so in case your query leads to more than 1000 results, you might prefer to repeat it with narrower criteria (i.e. shorter date range), to get less than 1000 results.

The screenshot shows the SPHERE client interface with a data browse table and a detailed file information panel. The table lists files with columns for ID, File name, Size, Frame type, Obs. date, RA, DEC, Filter, Exp. time, Ext., Type, Ra., Sta., and Sta... The file information panel shows details for a specific file, including File ID, Reduced data ID, Filename, File size, Directory, File creation date, Standard, Archive, Data type, Raw / reduced, Observation night, Observation date, Process date, Observation ID, Frame type, FITS extensions, Processed, Status, User, and Specific FITS Key... (OBJECT, TARGET ID, EXPTIME, ESO DET SEQU DIT, ESO INS COMB FLT, ESO INS COMB ICOR, ESO INS4 COMB IND, ESO DET READ CURL., ESO DPR TYPE, ESO DPR TECH, ESO OBS PRGID ID, ESO OBS PI-COI NAMEUNKNOWN, ESO OBS ID).

ID	File name	Size	Frame type	Obs. date	RA	DEC	Filter	Exp. time	Ext.	Type	Ra.	Sta.	Sta...
2241...	ird_specal_dc-IRD_SPECAL_CONTRAST_CURVE_TAB...	101.2 KB	IRD_SPECAL_CONTRAST_CURVE_TABLE	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	2	1			
2241...	ird_specal_dc-IRD_SPECAL_SNRMAP_STACK...	5.2 MB	IRD_SPECAL_SNRMAP_STACK	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	1	1			
2241...	ird_specal_dc-IRD_SPECAL_MEDIAN_RESIDUAL_CU...	10.3 MB	IRD_SPECAL_MEDIAN_RESIDUAL_CUBE	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	1	1			
2241...	ird_specal_dc-IRD_SPECAL_ALLCOMB_TARGZ-spec...	28.8 MB	IRD_SPECAL_ALLCOMB_TARGZ	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	1	1			
2241...	ird_specal_dc-IRD_SPECAL_MEDIAN_RESIDUAL_STA...	5.2 MB	IRD_SPECAL_MEDIAN_RESIDUAL_STACK	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	1	1			
2241...	ird_specal_dc-IRD_SPECAL_PNS-reduced_image.png	470.4 KB	IRD_SPECAL_PNS	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	1	1			
2241...	ird_specal_dc-IRD_SPECAL_JSON-reduction.json	474 B	IRD_SPECAL_JSON	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	1	1			
2241...	ird_specal_dc-IRD_SPECAL_MSC_TARGZ-specal_R...	1.3 GB	IRD_SPECAL_MSC_TARGZ	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	1	1			
2241...	ird_specal_dc-IRD_SPECAL_CONTRAST_CURVES_TA...	39.4 KB	IRD_SPECAL_CONTRAST_CURVES_TARGZ	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	1	1			
2241...	ird_specal_dc-IRD_SPECAL_CONTRAST_CURVE_TAB...	101.2 KB	IRD_SPECAL_CONTRAST_CURVE_TABLE	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	2	1			
2241...	ird_specal_dc-IRD_SPECAL_SNRMAP_STACK...	5.2 MB	IRD_SPECAL_SNRMAP_STACK	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	1	1			
2241...	ird_specal_dc-IRD_SPECAL_MEDIAN_RESIDUAL_CU...	10.3 MB	IRD_SPECAL_MEDIAN_RESIDUAL_CUBE	2015-05-01 02:53:3...	199.644...	-51.28569	DB_K12	8.0	1	1			

Figure 2: Data browse example to access all reduced data from observing nights from 2015-04-29 to 2015-05-15

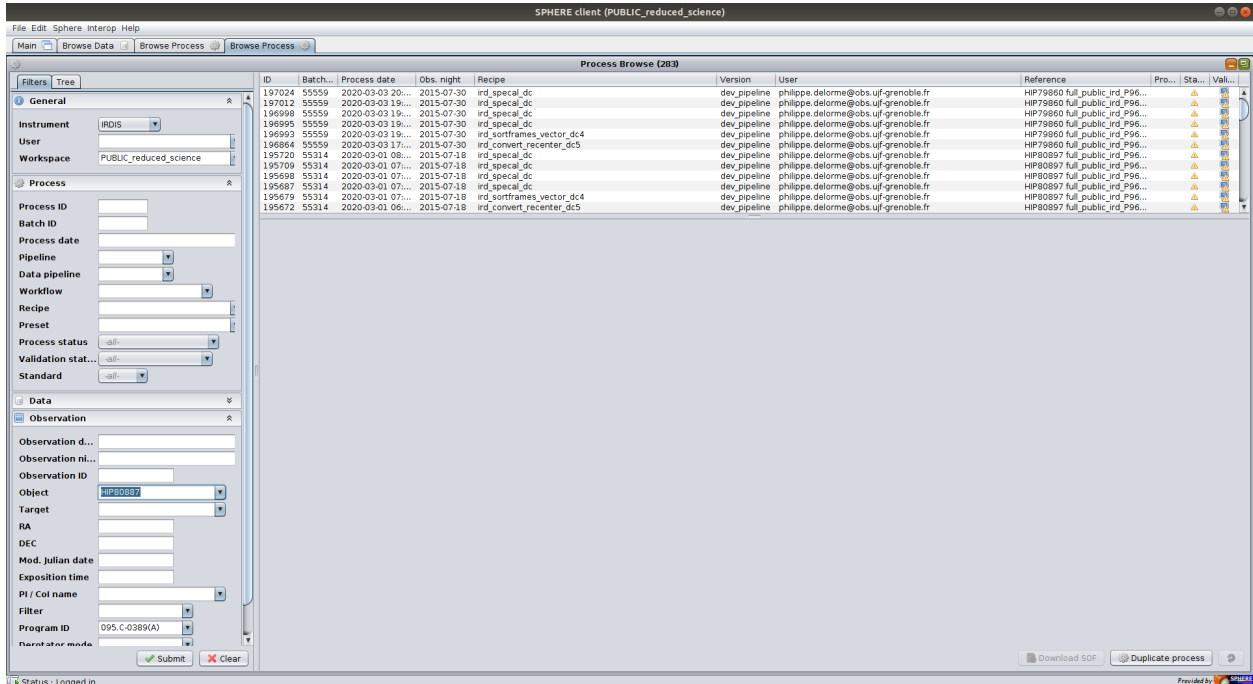


Figure 3: Process browse example to get all IRDIS reductions related to target HIP80887 within ESO run 095.C-0389(A)

Recommended browsing habits to select relevant datasets

First, most DC client fields have a help information box that appears if you hover your mouse for a few seconds on them : use this to get relevant info on the spot !

For “level 0” data (calibrations), we recommend to use “Browse data”, and then use the “instrument” and “observing night” range that are relevant for you.

For “level 1” data (reduced master cube), we recommend to use “Browse Process”, and use the “Target”, and/or ESO “Program ID” and “observing night” range that are relevant for you. Then type “convert” into the “Recipe” field and select “IFS_CONVERT_DC2” to select all IFS level 1 data and “IRD_CONVERT_RECENTER_DC5” to select all IRDIS level 1 data.

For “level 2” data (final residual images), we recommend to use “Browse Process” as well, and use the “Target”, and/or ESO “Program ID” and “observing night” range that are relevant for you. Then type “specal” into the “Recipe” field and select “IFS_SPECAL_DC” to select all IFS level 2 data and “IRD_SPECAL_DC” to select all IRDIS level 2 data.

For level 1 and 2 data, using “Browse Process” instead of “Browse Data” indeed allows you to see and therefore retrieve all output files obtained when processing a given target at this

level, i.e. not only the main scientific output but also calibration or SPARTA files which will be necessary for you to exploit the data.

Transferring data from the SPHERE-DC to the user

Once you have browsed the data or the processes, you can conveniently and efficiently download hundreds of Go if necessary, by right clicking on your selection and select the « download script » option. You can either download all files that are displayed by your filtering or select a subset with the regular ctrl or shift clicks to do your homemade subselection.

If you are using “browse data”, you will get the selected data files, and if you are using “browse process”, you will get the choice of downloading either all inputs and outputs of each selected process or only its outputs. We recommend to use “browse process” (that provides self-consistent data sets because of the way the processes are run) and download “outputs only”.

Saving the download script is almost instantaneous since it is only a txt file. Copy it in the repository where you want to download your data, make it executable (chmod u+x sphere_dl_script.sh), and then execute it (./sphere_dl_script.sh) so that all your files are efficiently downloaded, like with an ESO download script, through wget. You can also mail the script to somebody you want to share the data with, instead of sending loads of data, but note that the script has a time-limited validity (typically 1 month).

Note: To download a single file or a few files of small size (<50Mo), you might want to just right click on each file and select « save as ». Since this goes through java this is not at all an optimized download so please use it only for small files.

ACKNOWLEDGEMENTS: your acknowledgements are vital to allow this service to continue

If you use DC-processed data for any publication, please add the following sentence in the acknowledgments

'This work has made use of the the SPHERE Data Centre, jointly operated by OSUG/IPAG (Grenoble), PYTHEAS/LAM/CESAM (Marseille), OCA/Lagrange (Nice), Observatoire de Paris/LESIA (Paris), and Observatoire de Lyon. '

and refer to

Delorme et al. 2017: <http://cdsads.u-strasbg.fr/abs/2017sf2a.conf..347D> for all SPHERE reduced data

Galicher et al. 2018 : <http://cdsads.u-strasbg.fr/abs/2018A%26A...615A..92G> for IRDIS and IFS reduced data at level 2 and above

1.1.3 Access to data available at DIVA+

The second entry point to our reduced public data is through the Reduced high contrast imaging archive data (CeSAM/LAM located in Marseille) through *the DIVA+ archive*. Access is provided through <http://cesam.lam.fr/diva/>. Various high-contrast imaging surveys are available (not only from SPHERE), including releases of SPHERE data (for instance the first release of the SHINE data, called F150, the 150 targets reduced/validated and published in <https://doi.org/10.1051/0004-6361/202038806>). Click on the desired target, then follow Details -> Observations -> Reductions

This archive provides:

- All the reduced data will be available in the HCI-FITS format (<http://cesam.lam.fr/diva/index/format>)
- Possibility to sort and download the reduced data by many criteria: stellar age, mass, metallicity, planets detected, etc.
- Tools to compute the orbital parameters of the planets detected at several epochs from one or several surveys
- Scientific Plots : orbital parameters, color-magnitude, statistics, etc.

ACKNOWLEDGEMENTS: your acknowledgements are vital to allow this service to continue

If you use DC-processed data for any publication, please add the following sentence in the acknowledgments

'This work has made use of the the SPHERE Data Centre, jointly operated by OSUG/IPAG (Grenoble), PYTHEAS/LAM/CESAM (Marseille), OCA/Lagrange (Nice), Observatoire de Paris/LESIA (Paris), and Observatoire de Lyon. '

and for data from instruments other than SPHERE:

'This research has made use of the Direct Imaging Virtual Archive (DIVA), operated at CeSAM/LAM, Marseille, France.'

and refer to

Delorme et al. 2017: <http://cdsads.u-strasbg.fr/abs/2017sf2a.conf..347D> for all SPHERE reduced data

Galicher et al. 2018 : <http://cdsads.u-strasbg.fr/abs/2018A%26A...615A..92G> for IRDIS and IFS reduced data at level 2 and above

Vigan et al. 2017 : <https://ui.adsabs.harvard.edu/abs/2017A%26A...603A...3V/abstract> for data from instruments other than SPHERE

1.2 Proprietary data

If you request a reduction of your PI data to the SPHERE-DC, the DC staff will create a workspace corresponding to your program and reduce your data within it. They will directly provide you with a download script containing level 1 and level 2 data as described above. However, if you want to have a closer look at the details of the reductions on your data and select a subset of them, you can request direct access to your private workspace. For this you need to have a valid DC-client login, that you can request if needed by clicking on the “?” button near the login interface in Fig. 1. Note that the password you enter in the form will be your dedicated password and that it cannot be retrieved by DC staff because it is crypted : don't forget it. When logged to your private workspace you will be able to browse data and processes as described above for public data. **Of course no public user (nor other private users) will be able to see your data.**

2. Understanding the data delivered by the SPHERE-DC

2.1 Download script directory architecture for SPHERE data

When you execute a SPHERE-DC download script your data are downloaded within the following repository architecture:

SPHERE_DC_DATA/TARGET_filter_night_processname_PID/filenames

SPHERE_DC_DATA is the name of the main directories within which all your SPHERE-DC download will arrive, so that your download do not get disseminated everywhere within your local system

TARGET_filter_date_processname_PID is a data-dependent -long- name of the directory within which the downloaded files are stored:

TARGET is the name of the target as found after a query to Simbad at the RA/DEC of the pointing. *Filter* is the filter used for the observation, e.g. DB_H23. *Night* is the observing night of the observation (Remember a DC night called e.g. 2015-12-25 goes from 2015-12-25 at 20:00 UT to 2015-12-26 at 19:59 UT, so that the morning calibrations are within the same night as their obs). *Processname* is the name of the process that produced this data (typically *ird_convert* or *ifs_convert* for level 1 data from IRDIS or IFS, and *ird_specal* or *ifs_specal* for level 2 data). *PID* is a unique number identifying the process on the DC, so **if you have any remark, query or problem related to your data, please use this PID number so that we can identify quickly which dataset you are referring to, and send a mail to sphere-dc-req@univ-grenoble-alpes.fr**

filename is the name of each file you requested. Note that given the data structure within the SPHERE-DC, the names of files of the same frame type (such as flats or residual maps, etc...) are all identical. **This means that the directory structure is necessary to easily understand what a given file is about. This also means that if you move files from one directory to another, files of the same frame type will be overwritten.**

2.2 Output description for SPHERE data

2.2.1 Reduced calibration files

Several types of reduced calibration files are available. The complete description can be found in the ESO SPHERE pipeline user manual ([latest version here](#)). The available frametype are the following (names are self-explanatory in most casea, and a brief definition is given otherwise):

For the IRDIS instrument:

- IRD_FLAT_FIELD: flat field
- IRD_STATIC_BADPIXELMAP: Badpixel map
- IRD_MASTER_DARK: dark, internal background or sky background calibration

For the IFS instrument:

- IFS_IFU_FLAT_FIELD: Flat field, including optics (e.g. dispersing elements) and detector, so full instrument
- IFS_WAVECALIB: Pixel to wavelength calibration
- IFS_MASTER_DFF_LONG*: Detector only flat field (* can be 1, 2, 3, 4 or BB for the different calibration lamps of the IFS unit)
- IFS_STATIC_BADPIXEL_MAP: Badpixel map
- IFS_MASTER_DARK: dark, internal background or sky background calibration

2.2.2 Level 1 reduced data : IFS and IRDIS output description

These output files have been produced using the `ird_convert_recenter_dc5` for IRDIS (or its newest version `ird_convert_dc2021`) and `ifs_convert_dc2` for the IFS, See Section 3. It produces different types of outputs files:

Main science file:

- Files `*center_im.fits` : ***master reduced cube in units of ADU per second***. The values are divided by the neutral density transmission (if any ND filter is used). The 4 dimensions are : [X, Y (detector coordinates), number of frames (temporal dimension), number of wavelength channels (spectral dimension)]. ***On IRDIS, the star is centered on 513,513 in image coordinates (as in visualization tools like ds9/SAO Image, with the center of the first pixel of images being located at [1.0,1.0]) at the center of a pixel. For 0-based indexing systems (as used in programming languages such as IDL or Python), this center corresponds to the coordinates [512,512], when arrays start with [0,0] at the center of the first pixel. On the IFS the star is centered on 146,146 in image coordinates or 145,145 in array coordinates.***

→ Corresponding frametype: IFS_SCIENCE_REDUCED_SPECTRAL_MASTER_CUBE (for the IFS instrument) and IRD_SCIENCE_REDUCED_MASTER_CUBE (for the IRDIS instrument)

Scientific calibration files:

- Files *median_unsat.fits (for flux calibration) : units in ADU per second. Those values are divided by the neutral density transmission (if any ND filter is used). The 4 dimensions are : [X, Y (detector coordinates), number of PSF (often one before and one after the science sequence), number of wavelength channels (spectral dimension)].

→ Corresponding frametype: IFS_SCIENCE_PSF_MASTER_CUBE (for the IFS instrument) and IRD_SCIENCE_PSF_MASTER_CUBE (for the IRDIS instrument)

Note that when no PSF is available the pipeline default is to use a typical SPHERE PSF (useful for PSF-fitting purpose), that cannot be used for photometric calibration. Values in units of contrast are still provided but have been divided by a factor 10^9 , to make obvious they are meaningless. These dummy PSF have minimal headers that explicitly say the PSF are not fit for photometric calibration.

- Files *lam.fits: array containing the central wavelength for each spectral channel.

→ Corresponding frametype: IFS_SCIENCE_LAMBDA_INFO (for the IFS instruments) and IRD_SCIENCE_LAMBDA_INFO (for the IRDIS instruments)

- Files *rotnth. fits : array containing the angle to rotate each frame to have the true north UP. A clockwise rotation applying the values (in degree, without changing the sign) contained in this vector will bring the north up. Corrected using true north calibration (TN=-1.75+/-0.08°) from [Maire et al. 2016 SPIE 9908, id. 990834](#).

→ Corresponding frametype: IFS_SCIENCE_PARA_ROTATION_CUBE (for the IFS instrument) and IRD_SCIENCE_PARA_ROTATION_CUBE (for the IRDIS instrument)

- Files *ATMO_CONDITIONS-seeing.txt contains information on MASS-DIMM seeing and coherence time, extracted from the header of each observed cube (containing NDIT frames), so with a sampling of 1 per cube. This file also provides the average value of seeing and coherence time along all cubes in the observing sequence.

- File *timestamp.fits

SPARTA files:

Level 1 data also include processed meta-data from the AO system 'SPARTA', including strehl ratio, turbulence coherence time, seeing and other. These values are natively sampled every 20s by SPARTA, and resampled here to provide one value per frame. Although the relative evolution of the seeing and coherence time estimated by SPARTA best

reflects the evolution of the turbulence conditions during the observations and can be for instance reliably used for frame selection, there are in absolute terms discrepancies with the MASS-DIMM seeing and coherence time. We recommend to cite the MASS-DIMM values in publication. The Strehl can be trusted for stars brighter than $G < \sim 12$. The related files are described below:

- **timestamp.fits** : dates of each frame, days, with an offset value to apply to find the Julian_day in the ESO convention. This value can be found in the header ('SUBTRACT' keyword). Alternatively the first value in the table **sampled_sparta_data.fits** is the time in JD already corrected by this constant.
- **sampled_sparta_data.fits**: table of sparta data resampled at the time of frame, of shape [Nframes,7], where the 2nd dimension correspond to the 7 following variables:
 - fulltime : time in JD
 - r0frames : r0 (Fried diameter, m)
 - windframes : effective wind velocity ($\text{m}\cdot\text{s}^{-1}$) See Cantalloube+2020 for more details)
 - srframes : Strehl Ratio (@ 1.6 micron), as estimated by the AO system
 - seeingframes: seeing (arcsec, estimated at 500 nm)
 - nb_sparta_frames: This frame by frame information is interpolated from the coarser sampling available for SPARTA data. This row gives the information on the availability of sparta data it was interpolated from: -1 if no data, 0 if no data within DIT but some value present within 1 minute from the center of DIT, N if $N > 0$ data present within DIT
 - neighbour_distance: -1 if no data, 0 if nb_sparta_frames > 0, delay (in seconds) to the closest extracted sparta in case nb_sparta_frames = 0
- **psf_sparta_data.fits**: sparta information at the time of the PSF observations. if npsf PSF files exists, this is a [3, npsf+1,4] array for reduction before 2021-03-11, dimensions are [3, npsf+1,5] after because timestamp for each PSF as been added
 - dim 1 : min, max, avg
 - dim 2 : 0: npsf-1 : for each PSF file. Last line = stat combining all frames of all files
 - dim 3 : r0, effective wind velocity, strehl ratio, seeing and time after 2021-03-11

Reminder: The SPARTA Strehl ratio is provided at 1.6 micron and the turbulence parameters (r0, seeing) at 500nm. Note that the coherence time τ_0 can be derived from r0 and effective wind velocity from Roddier 1981: $\tau_0 = 0.314 r_0 / v_{\text{eff}}$. However this value is also over optimistic (same comment as above for DIMM Vs SPARTA seeing).

- **all_contr_curve_1sigma.fits** : normalized flux profiles of coronographic science images ("raw contrast") format = [3, n_separations, n_frames, 2 spectral channels]
 - output(0,*,k, l) : distance to the central pixel
 - output(1,*,k, l) : profile of azimuthal stdev, normalized by the maximum of the PSF, of frame k (1 sigma), hence in contrast units.

-
- $\text{output}(2,*,k, l)$: profile of azimuthal mean, normalized by the maximum of the PSF , of frame k , hence in contrast units.

2.2.3 Level 2 reduced data available at the SPHERE-DC

We first apply an automatic frame selection and sorting procedure. Then several routines are used to produce a residual image whenever possible, for IFS and IRDIS observations. Each algorithm produces a series of files, described below.

Sortframe informations

Bad quality frames are regularly present in data cubes and may degrade the quality of the final image. For coronagraphic images, they can be due to major problems with the AO (for example star outside the coronagraph mask for one or more frames) or to bad atmospheric conditions. We have implemented an automatic method allowing us to perform two levels of selection for coronagraphic images: a soft selection and a more aggressive selection, which is applied to all public reduction. For the public data analysis, we have chosen to use, after extensive tests, the following selections:

- Soft selection for noADI, cADI analysis, PCA (selection after step 1 only)
- Aggressive selection for TLOCI (selection after step 1+ step 2)

See Section 3.2.8 for a description of the criteria and steps used in the analysis. The three output files are:

- main output file (FRAME_SELECTION_VECTOR frame type): Vector containing the frame number to be kept, for two selection levels (soft and aggressive). Indexes start at 1.
- A pdf file (frametype IRD_IMAGES_CRITERIA for IRDIS and IFS_IMAGES_CRITERIA for IFS), to visualize images in a synthetic ways (48 plots per page, extraction of image center only, 120x120 pixels, superposition of information: selection status, step2 criteria)
- PNG file (frametype IRD_CRITERIA_VS_FRAME for IRDIS and IFS_CRITERIA_VS_FRAME for IFS), showing the various criteria used to performed the selection and additional information (for example adaptive optics information / SPARTA) to understand the selection.

Residual image production

The final images at this level have been produced by the `ird_specal` (for IRDIS) and `ifs_specal` (for IFS) described in Galicher et al. (2018)

Flux in the reduced images: ***The pixel values in the reduced images are the flux ratio between the residual flux in the coronagraphic image and the estimated maximum of the star PSF. Thus, it gives the contrast ratio in the reduced images.*** The maximum of the star PSF is derived from the best fit of the PSF by a function that reproduces the SPHERE PSF (close to a Gaussian function). The coronagraphic image results from the speckle intensity minimization. ***The normalization factors are saved in the normalization_vector.fits file.***

Most files of interests are available in the *allcomb.tar file and after extraction in the Reduction_000/Images/allcomb/ directory. These files are the following:

- *reduced_image_median.fits: Final reduced image stacking both channel together
- *cube_reduced_image_median.fits : Final reduced images, one stack per channel
- *reduced_image_corrthput.fits: Final reduced image corrected from throughput (files with throughput correction are only available for two of the algorithms, see below, cADI and TLOCI)
- *reduced_image_median.fits and reduced_image_median_corrthput.fits: same but using a median-combination instead of an average of the data cube
- *snrmap.fits and snrmap_unsharp13.fits: 2D SNR map before/after a 13x13 pixel unsharpened mask.
- detection_map.fits: map of a local 1-sigma standard deviation worked out in boxes of 2 FWHM radial size and a total area of 5 FWHM.
- normalization_vector.fits contains the value to convert the reduced images that are in contrast units into ADU/s units. Contains one value per wavelength channel.
- Note that the files cube_* contains the same images but one per wavelength instead of being stacked.

Other files of interests are in the *contrast_curves.tar file and include:

- Various plots in .eps format
- The contrast curve (contrast versus the angular separation) both in fits and ascii format. The fits file is called *contrast_curve_tab.fits, it is a list stored as a binary table in the first first data unit. The list contains 3 elements for IRDIS and 40 for the IFS corresponding to the individual spectral channels (2 for IRDIS, 39 for the IFS) plus the combination of all channels. Each of those elements is itself a list of respectively:
 - the target name (stored as a string)
 - the wavelength (be careful, also stored as a string)
 - the pixel scale (be careful, also stored as a string)
 - the confidence level expressed as a number of gaussian sigma, typically 5 (be careful, also stored as a string)
 - the repository where the data is stored in the internal SPHERE-DC architecture
 - the separation in arcsec stored as an array of floats
 - the contrast also stored as an array of floats

The algorithms used to perform the analysis can be identified by the field "Reference" in the process browser (see FAQ) and in the directories after retrieval. They are:

- NoADI: simple stacking of the master cubes
- cADI : Classical Angular Differential Imaging (Marois et al. 2006, Lafrenière et al 2007)
- TLOCI : Flavor of Angular Differential Imaging based on Locally Optimised Combination of Images (Marois et al. 2010, Galicher et al. 2018)

- RadProfile: Simple subtraction of the median flux at each given radius. Effectively removes the circularly symmetric component of the PSF.

2.2.4 Useful header keywords added at the end of DC-processed fits headers

These keywords provide you details of the effective exposure time, parallactic rotation and various observing conditions during the observations, for interpretation and publication purposes. Some of these keywords might be missing for processes run before July 20th 2020. Don't hesitate to contact us if you need some information missing for older processes.

TN_CORR	-1.80200 / deg True North correction: Astrometric calibration applied
TN_ERROR	0.0420000 / Error on TN_CORR in deg
PIXTOARC	7.46000 / mas per pixel(default)
PIX_ERR	0.0200000 / Error on PIXTOARC in mas
CALPIXTN	'NA ' / date of calibration for pix scale
CALTN	'2017-06-02' / date of calibration for TN
SR_MIN	0.606077 / Strehl ratio from SPARTA WFS, min of full dataset
SR_MAX	0.654006 / Strehl ratio from SPARTA WFS, max of full dataset
SR_AVG	0.621446 / Strehl ratio from SPARTA WFS, avg of full dataset
SEEI_MIN	1.28428 / Seeing from SPARTA WFS, min of full dataset Not as accurate in absolute as DIMM values below
SEEI_MAX	1.45104 / Seeing from SPARTA WFS, max of full dataset Not as accurate in absolute as DIMM values below
SEEI_AVG	1.37323 / Seeing from SPARTA WFS, avg of full dataset Not as accurate in absolute as DIMM values below
WIND_MIN	11.8809 / Effective wind velocity (m.s-1) from SPARTA WFS, min of full dataset Not as accurate in absolute as DIMM values below
WIND_MAX	14.2776 / Effective wind velocity (m.s-1) from SPARTA WFS, max of full dataset Not as accurate in absolute as DIMM values below

WIND_AVG	13.1817 / Effective wind velocity (m.s-1) from SPARTA WFS, avg of full dataset Not as accurate in absolute as DIMM values below
OBS_STA	'2017-06-13T01:26:16.56' / Starting date of observation
OBS_END	'2017-06-13T01:45:53.45' / End date of observation
HIERARCH ESO TEL AIRM MEAN	1.060 / UPDATED average airmass during full observing sequence
HIERARCH ESO TEL AMBI FWHM MEAN	2.720 / UPDATED average MASS-DIMM FWHM during full observing sequence
HIERARCH ESO TEL TAU0 MEAN	0.00100000 / UPDATED average MASS-DIMM TAU0 during full observing sequence
LEVELSEL	2 /Agressiveness of selection vector used when removing bad frames: 0: NO selection, 1 Selection applied
EFF_NFRA	59.00 /UPDATED effective number of frames used in dataset, after selection
EFF_ETIM	944.0 /UPDATED effective exposure time used in dataset, after selection
BIN_MISS	'Binning info missing' /Assumed no temporal binning. EFF_ETIM might be underestimated if this keyword is present.
SCPIPE	'SpeCal ' / All "SC" keywords below refer to values used/determined by the reduction algorithm, usually SpeCal
SCVERS	'2020-09-09' / Version of the package
SCDATE	'Thu Feb 4 10:12:12 2021' / Date of the data reduction
SCFOVROT	9.63456 / Total fov rotation(deg)
SCPSFVAR	'Alert ' / Strong flux variation for >1 PSF
SC_FWHM	5.00000 / FWHM in pixels used by the algo
SC_FLRMS	0.0456585 / variation of the star flux (sequence + PSF)
SC_MODE	'NO ADI ' / algorithm

SC ADI	'ADI ' / accounts for ADI only
SCRMIN	1.00000 / min radius, in FWHM,
SCRMAX	40.0000 / max radius in FWHM
SC TYPE	'NOADI REDUCED MED' / algo applied, here Median of median images

Note that for files processed after July 20th 2020, in addition to these keywords placed at the end of each header, some ESO keyword in the bulk of the header are also updated, like the START/END values that are modified so that they represent value at the start and end of the full dataset, and not only one individual datacubes.

2.2.5 Level 2 reduced data available at DIVA+

The files available are a subset of those available at the Processing Center (described above), and are currently in the DIVA+ database:

- The final reduced image using a median-combination (equivalent to reduced_image_median.fits above in fits, and png format for a quick look)
- The contrast curve (5-sigma contrast versus the angular separation) in ascii format.
- The 2D SNR map after a 13x13 pixel unsharpened mask, stacked in wavelength (equivalent to the snrmap_unsharp13.fits above)
- The 1-sigma detection limit as a 2D map :cube_detection_map.fits
- The cubes in lambda (planned):
 - cube_reduced_image_median_corrthput.fits
 - cube_reduced_image_median.fits (when the _corrthput is not available)
 - cube_snrmap.fits

2.3 Output description for NaCO/VLT, Keck, Hubble data

The Output description for NaCO/VLT, Keck, Hubble data is available here:

<http://cesam.lam.fr/diva/index/surveys>

Depending the surveys, only one or several high-level science products (PSF and reduced image, detection limit, candidates, etc.) are available. But, all the data are downloadable in the HCI-FITS format (<http://cesam.lam.fr/diva/index/format>).

3. Reduction Information

3.1 IRDIS and IFS summary

3.1.1 IRDIS-only steps

For IRDIS data, the first reductions steps (dark/background, flat, and bad pixel correction) rely on the SPHERE Data Reduction and Handling (hereafter DRH) pipeline (Pavlov et al. 2008) provided by ESO. However, we do not use daytime background calibrations to subtract the dark and background because we noted a significant difference in K-band with a systematic, spatially variable, offset (typically of the order of 100~ADU for 64s exposure time) between day-time dark background calibrations and the sky background within science frames. A similar effect is visible in the H-band, but at a much lower level, of a few ADU. We corrected this limitation in the background subtraction by systematically using instead sky background frames acquired just after the science observations, in both H and K bands, when available.

Star centering

Since most of SPHERE observations use the pupil-tracking observation mode a very accurate determination of the star center is needed in order to successfully apply all the angular differential imaging and spectral differential imaging methods. For IRDIS dual-band imaging, the star center needs to be found in both IRDIS channels so that the full dataset can be stacked. By default, we use the DRH *sph_ird_star_center* routine to find these star center positions, using the waffle images acquired for this purpose just before and after the science observations. This very fast routine provides an accurate centering in many cases, but we found that it could fail in the case the waffle spots are too weak, especially in the K-band, for which the much stronger thermal background noise could hide them. We therefore designed an automated way to check the quality of the DRH centering by comparing the 2 possible center positions that can be derived out of the 4 waffles. When the distance between these 2 possible center positions was found to be greater than 0.9 pixel, we used a SPHERE-DC made IDL routine, *ird_star_center_3waffles_dc*, that is more robust to identify weak waffle spots. This routine is able to detect weaker waffles because it looks for them only at small circular apertures located at the expected (wavelength-dependent) position of each waffles, uses a high-pass spatial filtering in addition of the sky-background subtraction and median stacks all waffle images available to increase SNR. When no waffle frames are available within a routine, we use the homemade routine *ird_star_center_nowaffle_dc*, that attempts to find a center on any non saturated, non coronagraphic science frame in the dataset. If none of these exist, the routines determine a rough (accuracy ~1pixel) center on the Arago/Poisson spot at the center of the coronagraphic image. This last resort strategy is there to allow the pipeline to proceed,

because it needs a centering file, but the left and right channels will probably not use the same absolute centering, leading to sub-optimal stacking, see below.

Some SPHERE datasets, notably those targeted at confirmed substellar companions, use continuous waffle mode observations, meaning that the waffle spots are activated during the entire science observation. In this specific case, we perform an individual recentering of each frame in the sequence using the SPHERE-DC made star-centering IDL routine described above. This slightly improves the images quality by correcting any drift or jitter of the targeted star behind the coronagraph, and improves the quality of the resulting astrometric measurements by removing these sources of error.

In case of suspected issue with centering:

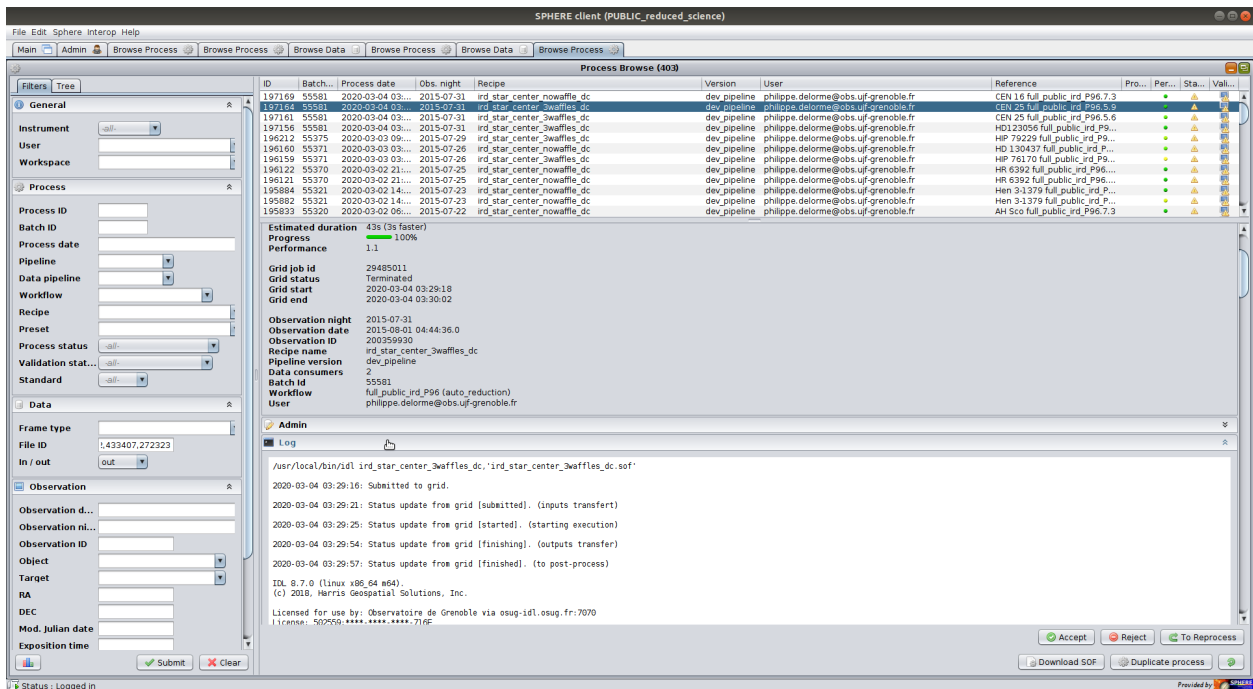
For the IRDIS pipeline, the star centering step is a very tricky step, for multiple reasons:

1- it is mandatory because IRDIS uses 2 detectors, whose centers need to be set to the exact same position for any post-processing application.

2- Depending on the science case, the centering relies on different observations such as waffle images (standard for SPHERE data), non saturated science images, or even no data dedicated to finding the center exists (notably for low contrast, resolved solar system objects), so the pipeline needs to be adaptable.

3- The quality of the raw centering frames is not always good, especially in K band, and sometimes waffles are very difficult to find. Binary stars and low contrast binaries can also confuse the centering routines. **When the centering fails on one channel, our routine purposely produces an obviously wrong centering on this channel, that should be obvious to notice, see example below.** In this case level 2 data are not relevant , but level 1 data can be used as a basis for a user-tuned centering. This is notably the case for solar system resolved objects for which our automatic pipeline cannot find the center but that some specialist users might be able to derive via a carefully tuned cross-correlation approach.

While our automated pipeline does find a correct center in most cases, there are cases when it failed and we did not notice it (we do not inspect manually all the products of the automated pipeline). If you notice anything weird with your final data, we advise you to look at the IRD_STAR_CENTER file associated with a given dataset (use browse data, IRD_STAR_CENTER frametype, and target and observing night corresponding to your dataset). This unique file can be produced by various routines (depending on the context described above) and is a fits table containing the coordinates of the center for IRDIS left and right detectors. Using the SPHERE-DC client, double-click on this file, browse down on the resulting window, and then double click on the process shown in the tab "output of recipe processes:". This will open another window, on which you click on the "log" tab (see figure).



At the bottom of the log, you will find the following information:

Image centers are located here:

Left: 479.473 524.280

Right: 1506.03 510.793

Corresponding waffle positions:

435.080 568.664

524.594 568.942

434.823 480.084

523.898 479.863

1458.69 558.109

1554.00 558.231

1458.42 463.704

1553.21 463.647

distance between center estimates on left detector is: 0.477676

distance between center estimates on right detector is: 0.536144

When waffle image is available, the DC used the 4 waffles for each channel to find 2 center solutions for each detector and compare them, which provide a robust sanity check on the center quality, as well as an estimate of the error on the center (or rather of $\sqrt{2}$.error). In this example the error is $0.5/\sqrt{2}=0.35$ pixels, or 4.3 mas, which is indicative of a correct centering.

When the recipe fails this log looks like this:

Corresponding waffle positions:

435.610	568.354
525.275	568.910
435.519	480.081
385.000	385.000
1409.00	385.000
1409.00	385.000
1458.88	463.563
1409.00	385.000

Fit completely failed on at least one right channel waffle

distance between center estimates on left detector is: 84.6131

distance between center estimates on right detector is: 1412.80

Left quadrant center estimation is dirty, try to find which waffle is outlier

Try to identify the good waffles

only one waffle is outlier, determine center using the 3 others

435.610	
525.275	
435.519	
568.354	
568.910	
480.081	
435.610	435.519
525.275	

center is: 480.420 524.357

Right quadrant center estimation is dirty, try to find which waffle is outlier

Try to identify the good waffles

only one waffle is outlier, determine center using the 3 others

1409.00	
1409.00	
1409.00	
385.000	
385.000	
385.000	

center is: 1409.00 385.000

In this case, the left channel center is correct (while only 3 waffles were found), while only one of the right channel waffles could be found, and the routine returns an obviously wrong center. For this dataset, the level 2 full stack (**reduced_image.fits*) is useless, but the level 2 channel by channel stacks (**cube_reduced_image.fits*) provide correct residuals stacks for the left channel.

*In case of any remaining doubt on the quality of the centering, **don't hesitate to directly ask us at sphere-dc-req @ univ-grenoble-alpes.fr**. Please provide the process ID number (PID in the download script repertory architecture) of the corresponding dataset so that we can easily identify your issue.*

3.1.2 IFS-only steps

For IFS data, we use the Data Reduction and Handling (DRH, Pavlov et al. 2008) pipeline but complement it with additional steps implemented at the SPHERE Data Center (Delorme et al. 2017) that improve the wavelength calibration, apply a correction for cross-talk, and improve the handling of bad pixels. The improvement to the wavelength calibration uses a cubic fit whose coefficients are obtained from the shift with wavelength of the spots generated in the star-center calibration, as that should scale linearly with wavelength. The cross-talk correction is an iterative procedure that corrects for the spectrograph PSF, using coefficients derived using appropriate tests performed in the laboratory during instrument assembly. Bad pixels are corrected using a sky observation acquired at the end of each exposure. This produces more accurate results than that based on the flat field observations used within DRH.

3.1.3 Final steps common to IRDIS and IFS

After these cosmetics steps, both IRDIS and IFS datasets are corrected from the anamorphism of the instrument and the astrometric solution (pixel scale and True north solution) found by the calibration described in [Maire et al. 2016 SPIE 9908, id. 990834](#). are applied to each image. The result is a reduced master cube combining all frames obtained during a given observation, that can be used as input of any Angular Differential (Marois et al. 2006, Lafreniere et al. 2007) algorithms. We also associate this master cube with the vector of accurate de-rotation angles for each frame, using the accurate timing of each frame, to obtain a frame-by-frame determination of the parallactic angle.



References

Delorme, P., Meunier, N., Albert, D., Lagadec, E., Le Coroller, H., Galicher, R., Mouillet, D., Boccaletti, A., Mesa, D., Meunier, J. -C., Beuzit, J. -L., Lagrange, A. -M., Chauvin, G., Sapone, A., Langlois, M., Maire, A. -L., Montargès, M., Gratton, R., Vigan, A., Surace, C., **The SPHERE Data Center: a reference for high contrast imaging processing**, SF2A-2017: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, held 4-7 July, 2017 in Paris. Eds.: C. Reylé, P. Di Matteo, F. Herpin, E. Lagadec, A. Lançon, Z. Meliani and F. Royer, pp.347-361 <http://cdsads.u-strasbg.fr/abs/2017sf2a.conf..347D>

Galicher, R., Boccaletti, A., Mesa, D., Delorme, P., Gratton, R., Langlois, M., Lagrange, A. -M., Maire, A. -L., Le Coroller, H., Chauvin, G., Biller, B., Cantalloube, F., Janson, M., Lagadec, E., Meunier, N., Vigan, A., Hagelberg, J., Bonnefoy, M., Zurlo, A., Rocha, S., Maurel, D., Jaquet, M., Buey, T., Weber, L., **Astrometric and photometric accuracies in high contrast imaging: The SPHERE speckle calibration tool (SpeCal)**, *Astronomy & Astrophysics*, Volume 615, id.A92, 11 pp. <http://cdsads.u-strasbg.fr/abs/2018A%26A...615A..92G>

Lafrenière D., Marois C., Doyon R., Nadeau D., Artigau E., **A New Algorithm for Point-Spread Function Subtraction in High-Contrast Imaging: A Demonstration with Angular Differential Imaging**, *The Astrophysical Journal*, Volume 660, Issue 1, pp. 770-780., <https://ui.adsabs.harvard.edu/abs/2007ApJ...660..770L/abstract>

Pavlov, Alexey; Möller-Nilsson, Ole; Feldt, Markus; Henning, Thomas; Beuzit, Jean-Luc; Mouillet, David **SPHERE data reduction and handling system: overview, project status, and development**. *Advanced Software and Control for Astronomy II*. Edited by Bridger, Alan; Radziwill, Nicole M. *Proceedings of the SPIE*, Volume 7019, article id. 701939, 12 pp. (2008). <https://ui.adsabs.harvard.edu/abs/2008SPIE.7019E..39P/abstract>

Marois C., Lafrenière D., Doyon R., Macintosh B. Nadeau D. **Angular Differential Imaging: A Powerful High-Contrast Imaging Technique**, *The Astrophysical Journal*, Volume 641, Issue 1, pp. 556-564 <https://ui.adsabs.harvard.edu/abs/2006ApJ...641..556M/abstract>

Marois C., Macintosh B., Véran J.-P., **Exoplanet imaging with LOCI processing: photometry and astrometry with the new SOSIE pipeline**, *Proceedings of the SPIE*, Volume 7736, id. 77361J (2010). <https://ui.adsabs.harvard.edu/abs/2010SPIE.7736E..1JM/abstract>

Maire, Anne-Lise; Langlois, Maud; Dohlen, Kjetil; Lagrange, Anne-Marie; Gratton, Raffaele; Chauvin, Gaël; Desidera, Silvano; Girard, Julien H.; Milli, Julien; Vigan, Arthur; Zins, Gerard; Delorme, Philippe; Beuzit, Jean-Luc; Claudi, Riccardo U.; Feldt, Markus; Mouillet, David; Puget, Pascal; Turatto, Massimo; Wildi, François **SPHERE IRDIS and IFS astrometric strategy and calibration** *Proceedings of the SPIE*, Volume 9908, id. 990834 12 pp. (2016). <https://ui.adsabs.harvard.edu/abs/2016SPIE.9908E..34M/abstract>

Vigan et al. , Bonavita M., Biller B., Forgan D., Rice K., Chauvin G. et al., **The VLT/NaCo large program to probe the occurrence of exoplanets and brown dwarfs at wide orbits. IV.**

Gravitational instability rarely forms wide, giant planets, 2017, A&A, 603, A3
<https://ui.adsabs.harvard.edu/abs/2017A%26A...603A...3V/abstract>