UNIVERSITY OF ARIZONA

LBTI nulling data reduction: User's Guide

Denis Defrère
February 14, 2016

CONTENTS

Document revisions

- Version 1: January 2015 – Denis Defrère (denis@lbti.org)
- Version 2: April 2015 – DD – ORR version
- Version 3: December 2015 – DD – version released to NexSci
- Version 4: February 2016 – DD – improved according to feedback from RMG and SE

1 SOFTWARE AND SYSTEM REQUIREMENTS

The list below provides a list of minimum software and system requirements to run the pipeline:

- IDL version 5 or later (`http://www.exelisvis.com/ProductsServices/IDL.aspx`);
- Up-to-date Astrolib package (`http://idlastro.gsfc.nasa.gov/`);
- At least as much disk space as needed for the raw data. This is required to save the L1/L2 and intermediate files;
- At least 4 Gb of free RAM for standard null data reduction;

2 INSTALLATION

This section provides a few instructions to download and install the pipeline. Ideally, it is recommended to install the pipeline in your default workspace, which is usually under IDL-WorkspaceN.n (where N.n is your IDL version number). **If you do not install the pipeline in your default IDL workspace, it is necessary to switch your working directory before running the pipeline because it uses paths relative to the current workspace.** One possibility to do that is to use an IDL startup file.

- To download the last version, enter the following command from a terminal:

  ```
  svn checkout http://lbti.as.arizona.edu/svn/lbti/nodrs
  ```

  This will create a directory called *nodrs* on your local machine and create the necessary links for version control (subversion software or SVN). SVN is installed by default on any recent Mac OS system. It can also be downloaded here: `http://subversion.apache.org/`. Contact Phil Hinz to get an username and a password (`phinz@as.arizona.edu`).

- Add the pipeline and the astrolib package to your IDL path. The code is self-contained and does not require any additional routines. On an unix-based system, you can add the following lines to your ".bashrc" file:

  ```
  IDL_PATH=+/your_path/nodrs/:$IDL_PATH
  IDL_PATH='<IDL_DEFAULT>':$IDL_PATH
  IDL_PATH=+/your_path/astrolib/:$IDL_PATH
  export IDL_PATH
  ```

- To use advanced functionalities (e.g., multithread, movie creation), it is also necessary to add the following dynamic library to your IDL path: bin.darwin.x86_64 to your path. On an unix-based system, you can add the following lines to your *.bashrc* file:

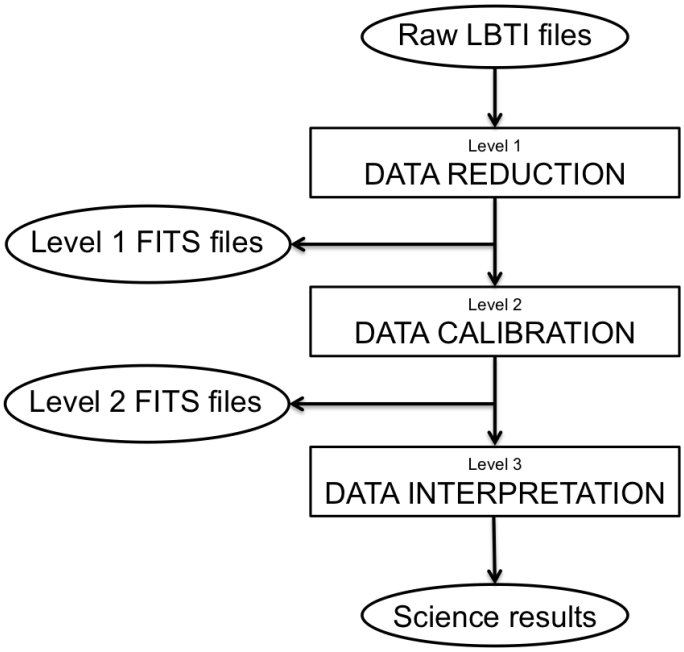

Figure 3.1: Top-level structure of LBTI nulling data reduction pipeline.

```
IDL_DLM_PATH=+/idl84/bin/bin.darwin.x86_64/:$IDL_DLM_PATH
export IDL_DLM_PATH
```

Once your ".bashrc" file is modified, don't forget to rebuild it before starting IDL (i.e., source .bashrc).

3 TOP-LEVEL STRUCTURE

3.1 FILE STRUCTURE

Currently, *nodrs* contains ~150 routines divided over 8 directories. In general, a routine contains a procedure or function with the same name as the file. Preceding may be some functions that are called in conjunction with this routine. Trailing are usually one or more test harnasses: procedures, named TEST_*, with which we have validated the functionality of the routine. The top-level structure of the pipeline is shown in Figure 3.1.

3.2 CALLING PROCEDURE

The main routine of the pipeline is called `LBTI_DRS.pro` and is located under the *main* subdirectory. It is controlled by two main mandatory inputs:

- **DATE**: String with the UT date to be reduced based on the format 'yymmdd' (e.g., '130524');

- **CGF_FILE**: String with the name of the config file containing the reduction parameters. Examples files are located in the 'cfg' directory. To run the nulling pipeline, the default file is called *hosts.cfg*. There is also a file called 'quicknull.cfg' that allows to perform a quick reduction (e.g., real-time reduction). The content of the config files is described in section 3.4.

nodrs is also controlled by a few optional keywords that superseed the automatic computation or allow to select only a fraction of the data:

- **DATA_IDX** : Two-element vector with the lower and the upper file numbers of the data files to reduce (e.g., [20,999]);
- **DARK_IDX** : Two-element vector with the lower and the upper file numbers of the dark files (e.g., [0,9]);
- **FLAT_IDX** : Two-element vector with the lower and the upper file numbers of the flat files (e.g., [10,19]);
- **BCKG_IDX** : Two-element vector with the lower and the upper file numbers of the background files (e.g., [10,19]). Generally used when BCKG_MODE = 4;
- **NOD_IDX** : 2x(number of nods) array with the lower and the upper file numbers of each nod position (e.g., [[120,239],[240,299],[300,399]]);
- **BAD_IDX** : Vector with the file number of frames to be removed;
- **DATA_PATH** : String vector pointing to the the path of scientific data (superseed the date and data_idx inputs);
- **BCKG_PATH** : String vector pointing to the the path of background data (superseed bckg_idx keyword);
- **DARK_PATH** : String vector pointing to the the path of dark data (superseed dark_idx keyword);
- **FLAT_PATH** : String vector pointing to the the path of flat data (superseed flat_idx keyword).

Finally, *nodrs* is also controlled by a few running keywords:

- **MASTERLOG** : Set this keyword to force the creation of new masterlog file (see description in section 3.6);
- **SKIP_ADI** : Set to skip ADI processing (superseed the value in the config file);
- **SKIP_FLX** : Set to skip flux computation (files restored from disk, superseed the value in the config file);
- **SKIP_NULL** : Set to skip null computation (superseed the value in the config file);

- **SKIP_RED** : Set to skip image reduction (files restored from disk, superseed the value in the config file);
- **SKIP_VIS** : Set to skip visibility computation (files restored from disk, superseed the value in the config file);
- **LOG_FILE** : Set this keyword to save the results in an external log file;
- **NO_MULTI** : Set this keyword to turn off multi-threading (useful to debug the NSC reduction). By default, the reduction is performed in parallel on 2^n CPUs where n is chosen so that 2^n is smaller than the number of available CPUs ;
- **NO_SAVE** : Set this keyword to turn off data saving;
- **PLOT** : Set this keyword to plot the data;
- **VERBOSE** : Define the level of information printed to screen: (0) completely silent execution, (1) minimum level of information, (2) nominal level of information, and (3) debugging use.

Here are below a few typical calling sequences. If you copy/paste theses lines, you may have to replace the quotes in order for it to work.

```
LBTI_DRS, '150208', 'hosts.cfg', /VERBOSE
LBTI_DRS, '150302', 'quicknull.cfg', DATA_IDX=[100,8000], /VERBOSE
LBTI_DRS, '140302', 'leech.cfg', /VERBOSE, /NO_MULTI, /LOG_FILE
```

To reduce and calibrate HOSTS data, there is a specific routine in the batch directory. Here are the calling sequences for the regular and real-time data reduction and calibration (more info can be found in the routine itself):

```
REDUCE_HOSTS, '150208', /REDUCE, /CALIB_NULL
REDUCE_HOSTS, '150208', /REDUCE, /CALIB_NULL, /QUICK
```

3.3 SETTING UP THE PATH TO THE RAW DATA

The path to the raw data is set in the path.cfg file under the 'cfg' directory *nodrs*. For instance,

```
DATA_PATH2      '/Volumes/nodrs/data/LBTI/'
DATA_PATH1      '/Volumes/LaCie/data/LBTI/'
RESULT_PATH2    '/Volumes/nodrs/nodrs/results/'
RESULT_PATH1    '/Volumes/LaCie/nodrs/results/'
```

`DATA_PATH1` is primary directory to look for data. If data are not found, the pipeline will look in the secondary directory which is `DATA_PATH2`. This `DATA_PATH` directory must contain the data of each night in a specific sub-directory named "yymmdd" (e.g., 150208 for February 2, 2015 data). This is the same format as the format used by the data acquisition computer.

3.4 STRUCTURE OF THE CONFIG FILE

The config file is a mandatory input of the pipeline. You can create your own or use an existing one (located in the 'cfg' directory). The parameters in the config file are divided into 7 main categories:

- General reduction parameters which control the computation;
- Detector and wavelength parameters;
- Image reduction parameters which control the way the frames are calibrated (e.g., flat-fielding, background subtraction mode, etc);
- Flux computation parameters which control the way the flux is computed (e.g., aperture radius size, definition of background region, etc):
- ADI/PCA parameters which control the image selection, derotation, and PCA processing (e.g., number of principal components, ADI mode, etc);
- Null computation parameters which control the null computation (e.g., classical or statistical reduction, null range, etc);
- Null calibration parameters which control the null calibration (e.g., interpolation approach, degree of polynomial);

An example of config file is shown below. Note that the description has been truncated here for clarity. More information about the parameters can be found in the existing config files or in the routine `GET_DRS.pro` which reads the config file.

data.txt

```
; Reduction parameters for HOSTS data see more info in "nodrs/cfg/get_drs.pro"
; General reduction parameters
SKIP_ADI     1                  ; Skip ADI processing
SKIP_FLX     0                  ; Skip flux computation
SKIP_NULL    0                  ; Skip null computation
SKIP_RED     0                  ; Skip image reduction
SKIP_VIS     1                  ; Skip visibility computation

; Detector and wavelength
LMIRCAM      0                  ; If LMIRCam
MIRAC        0                  ; If MIRAC
NOMIC        1                  ; If NOMIC
LAMBDA_CEN   0                  ; Superseed value derived from filter wheels
BANDWIDTH    0                  ; Superseed value derived from filter wheels
OVERLAP      0                  ; Force beam overlap imaging mode only

; Image calibration parameters
BCKG_MODE         -2                ; Background subtraction mode -2 is default for nulling
BCKG_SEL     1                  ; Background selection mode when N_FRBCK is not null
N_FRBCK      -1                 ; Maximum Number of frames to preserve in background
N_TRANS      0                  ; Number of images to remove at the beginning of each nod
```

```
NO_BPM       1                  ; Turn on/off bad pixel correction
NO_DARK      1                  ; Turn on/off dark subtraction
NO_FLAT      1                  ; Turn on/off flat computation
NO_FIND      0                  ; Turn on/off quick beam finding function
PRE_CROP         0                 ; Pre crop the frames xmin,ymin,xmax,ymax
SKIP_OPEN    0                  ; Don't read open-loop frames when appropriate

; Flux computation parameters
APER_RAD     2,4,8,10,16,24,30  ; Radius of the photmetric aperture in pixels
BCK_IRAD     19                 ; Inner radius of background region in pixels
BCK_ORAD     -1                 ; Outer radius of background region in pixels
BCK_CEN      0                  ; Two-element vector with the center of the background region
BIAS_ESTIM   1                  ; Compute the flux in a nearby empty region of the detector
FIT_MODE     5                  ; Method used to find the beam centroid
FLX_MODE     0                  ; 0: aperture photometry, 1: weighted aperture photometry
N_BIN        1                  ; Image re-binning factor in pixels
N_CLIP       128                ; Size of output images around the beam position, in pixels
SKIP_REGIS   1                  ; Don't register the frames
XCEN         0                  ; Two-element vector with the X position of the beams
YCEN         0                  ; Two-element vector with the Y position of the beams

; ADI/PCA processing parameters
ADI_MODE     0                  ; 0: regular adi, 1: PCA
BCK_LIM      0                  ; Acceptable background range after nod subtraction
N_COADD      100                ; Number of images to coadd before frame derotation
PARA_BIN     0                  ; Parallactic angle binning
PARA_RANGE   0                  ; Parallactic angle to preserve before PCA/ADI processing
PCA_DIM      0                  ; Dimensions of the ouptut image, on which sPCA will be performed
PCA_KLIP     6                  ; Number of principal components to be used in the sPCA
PCA_SMART    0                  ; Turn on/off smart PCA
PSF_FILE     'none'             ; Set to the name of the PSF file to use in the correlation
RIGHT_HANDED 0                  ; Image rotation direction must be 0 for LBTI
RIN_INIT     0                  ; Radius of the central mask to apply to the final image in FWHM
R_KEEP       1                  ; Ratio of frames to preserve based on correlation coefficient
SIG_FLX      0                  ; Flux selection in signa
SIG_POS      0                  ; Position selection in signa
SIG_SLO      0                  ; Moffat slope selection in signa
SIG_VIS      0                  ; Visibility selection
SKIP_MERGE   0                  ; Turn on/off merging the frames of all nods
SKIP_SEL     0                  ; Turn on/off frame selection
X_RANGE      0                  ; Range of X detector positions to preserve in pixels
Y_RANGE      0                  ; Range of Y detector positions to preserve in pixels

; Null computation parameters
NULL_MODE    2                  ; 0: mode, 1: best x%, 2: numerical NSC, 3: analytical NSC
KEEP_RATIO   0.05               ; Ratio of frames to preserve for null_mode = 1
MIN_FR       500                ; Minimum number of frames required per OB
N_BTSTRP     200                ; Number of bootstrap samples to compute the error bar
NSC_BFAC     1                  ; Multiplier on number of histogram bins
NSC_BINS         0                 ; Bin size for NSC reduction 0: constant, 1: variable.
NSC_CUBE     105,80,80          ; Size of the chi2 cube
NSC_OMIN     0                  ; Minimum number of occurences per bin for the fit
NULL_COR     2                  ; Set this keyword to subtract the high-freq null
NULL_RAD     8                  ; Aperture radius used for null computation
```

```
NULL_LIM      -0.01,0.20          ; Acceptable raw null range before NSC
N_FROB        0                   ; User-defined number of frames per OB
OB_MODE       0                   ; OB definition mode 0 for one OB per nod, default for nulling

; Null calibration parameters
CAL_METHOD    1                   ; Interpolation method for the transfer function
CAL_MODE          0                   ; 0: calibrate per pointing, 1: calibrate per OB
CHI2_LIM      0                   ; Limit on acceptable chi2 only with NSC
NULL_WEIGHT   1                   ; Set to 1 to weight the calibrated null computation
POLYDEG       0                   ; Degree of the polynomial for method 1
SPLIT_TF      1                   ; Set to 1 to split the TF when there is a dead time
SPLIT_NOD     0                   ; Calibrate each nod position separately
SPLIT_HOUR    0                   ; UT hour at which the TF will be split
SPLIT_TIME    2                   ; Maximum dead time between two null measurements
```

3.5 ADDING NEW OBJECTS

The pipeline requires some basic information about the observed stars to run properly. A long list of stars is already included and the code will complain if it doesn't find a star. To add a new star, it is necessary to edit manually the function `GET_TGT.pro` (see instructions there in).

3.6 OUTPUTS

The pipeline creates intermediate files after each main step of the reduction/calibration. The files are described here:

- **MASTERLOG FILE** : this file contains basic information about each frame in the data directory (e.g., file number, target name, time, wavelength, integration time, nod number, etc). It is created in the data directory every morning by a cronjob. If you reduce the data in real-time or if the file does not exist, it is created automatically before data reduction and saved in the data directory. This step can take up to 90 minutes depending on the number of frames but only needs to be executed once. Ideally, you can do this step while downloading the data to your machine with the following bash command: `idl, -e, LBTI_MASTERLOG, data_dir, /IDL`.

- **CONFIG FILE** : this file contains basic information about each instrumental configuration used during a given night (e.g., wavelength, integration time, etc). During data reduction, each file is associated with a specific configuration identification number that is used for various tasks (e.g., match background and science frames). The config file lists the main instrumental parameters for each config ID.

- **Level 0' files** : theses files contains the reduced and background-subtracted images. They are saved as one image cube per nod position under *results/instrum/l0_fits/date/* where *instrum* is the name of the instrument used (i.e., NOMIC or LMIRCAM). This step can be skipped by setting the SKIP_RED keyword.

- **DATALOG FILE** : this file contains basic information about each L0' data cube (e.g., nod id number, time, target name, position of the beams on the chip, etc). It is saved under *results/instrum/l0_fits/date/*.

- **Level 1 files** : the level 1 files consist of 9 different files:
  - UTDATE.fits contains the raw null measurements of all OBs. There is one such file per reduced night.
  - UTDATE_FILE_ID_FLAG_OBJECT_DIT_WAV_NULL.fits contains the flux measurements at null for each OB (one file per OB);
  - UTDATE_FILE_ID_FLAG_OBJECT_DIT_WAV_PHOT1.fits contains the photometry measurements for each OB for one beam (one file per OB);
  - UTDATE_FILE_ID_FLAG_OBJECT_DIT_WAV_PHOT2.fits contains the photometry measurements for each OB for the other beam (one file per OB);
  - UTDATE_FILE_ID_FLAG_OBJECT_DIT_WAV_BCKG.fits contains the background measurements for each OB (one file per OB);
  - UTDATE_FILE_ID_FLAG_OBJECT_DIT_WAV_NULL_IMG.fits contains the centered and cropped images associated to the flux measurements at null (one file per OB);
  - UTDATE_FILE_ID_FLAG_OBJECT_DIT_WAV_PHOT1_IMG.fits contains the centered and cropped images associated to the photometry measurements of one beam (one file per OB);
  - UTDATE_FILE_ID_FLAG_OBJECT_DIT_WAV_PHOT2_IMG.fits contains the centered and cropped images associated to the photometry measurements pf the other beam (one file per OB);
  - UTDATE_FILE_ID_FLAG_OBJECT_DIT_WAV_BCKG_IMG.fits contains the centered and cropped images associated to the background measurements at null (one file per OB);

  The files are saved under *results/instrum/l1_fits/date/* where *instrum* is the instrument used (i.e., NOMIC or LMIRCAM). The flux (rest. visibility) computation step can be skipped by setting the SKIP_FLX (resp. SKIP_VIS) keyword and the null computation can be skipped with the SKIP_NULL keyword.

- **Level 2 files** : these files contains the calibrated null measurements are saved under *results/nomic/l2_fits/date/*.

3.7 EXAMPLES

A few example batch routines are located under the *batch* directory. The main batch to reduce nulling data is called `reduce_hosts.pro`.

3.8 GUEST ACCOUNT

There is a guest account on the HOSTS computer at UoA. It is an Apple Mac Pro with 12 cores (2.7 Ghz Intel Xeon and 64 Go RAM) which is used for data reduction and, hence, contains all the processed files and results. In order to connect to this computer, you must be on the UoA network and type the following command `ssh hosts@ddefrere.as.arizona.edu`. The results, files, and figures are located under */Volumes/nodrs/nodrs/results/*. Contact `denis@lbti.org` if you want the password or want to use the computer to reduce the data.

4 CAVEATS AND KNOWN PROBLEMS

Roughly ordered in descending priority:

- Flux computation using PSF fitting not optimized for low SNR frames, which is generally the case for null frames.
- Background files are not valid to use with the NSC for positive background modes. Right now, the code uses the background measured in an empty region of the science frames as a best estimate for the NSC reduction.

5 FUTURE VERSIONS

Here is a list of a few items that will be implemented in the near future:

- Enable shared memory to speed up the access to the L0 fits files;
- Use the reduced images rather than the measured flux to do the null calibration;
- Improve PSF-fitting for low SNR frames;
- Implement multi-threading for flux computation;
- Interpol raw nulls based on elevation rather than time.
- Implement the analytical NSC.