Factor pipeline

Facet Calibration for LOFAR (T6)

Marco Iacobelli (ASTRON)

6th LOFAR data processing school
March 22/26 2021

Outline

Lecture

- DDE calibration/imaging & Factor in a nutshell
- The Factor parset & directions file
- The selfcal strategy
- Running Factor

Tutorial handout

Outline

Lecture

- DDE calibration/imaging & Factor in a nutshell
- The Factor parset & directions file
- The selfcal strategy
- Running Factor

Tutorial handout

Introduction: DDE cal/img

Direction-Dependent Effects in LOFAR data are primarily caused by:

- The ionosphere (see D2,T1) → mostly phase effects (vary quickly in time)
- The LOFAR beam (see L3,L4) → mostly amplitude effects (vary slowly in time)

Direction-Dependent calibration attempts to correct for these effects

Introduction: Factor

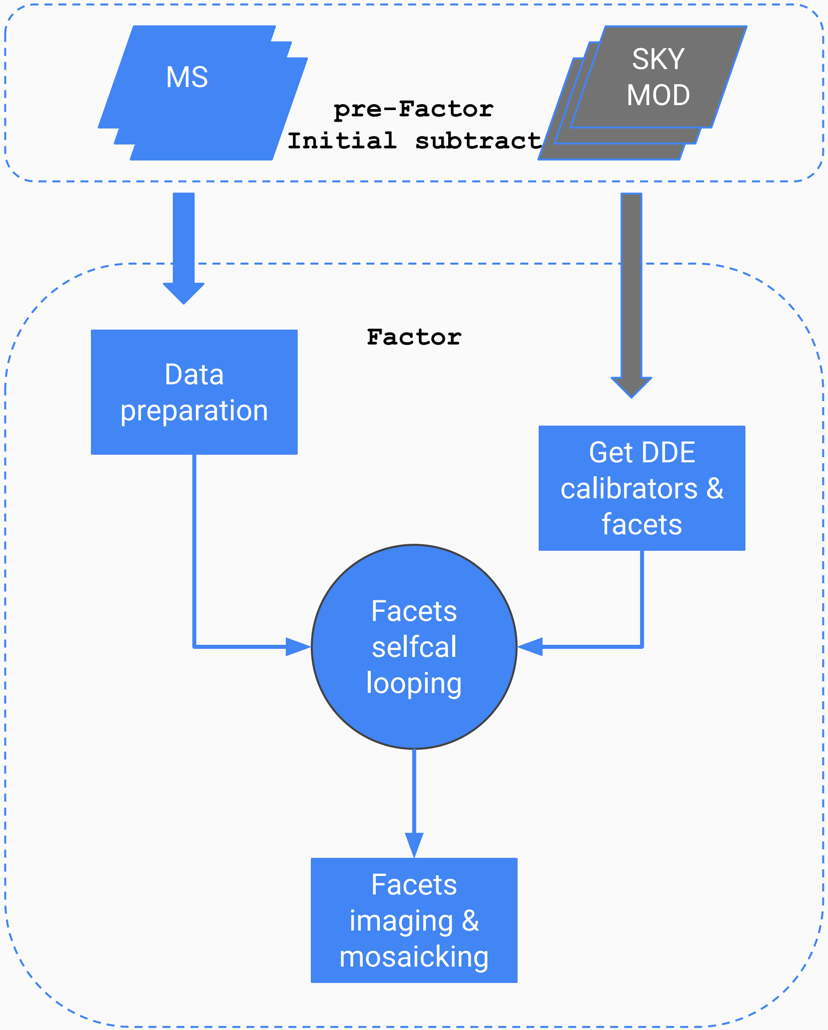
Factor performs Direction-Dependent calibration and imaging

- Based on the facet-calibration scheme of van Weeren et al. (2016)
- Supports multi-epoch datasets (interleaved or multiple nights)
- Designed to operate on HBA data only
- Recommended for targeted exposures
- Sub-optimal for:
    - full-field of view processing possible
    - very extended (e.g. >=1 deg) target sources (i.e. >1 facet needed)

Introduction: Factor

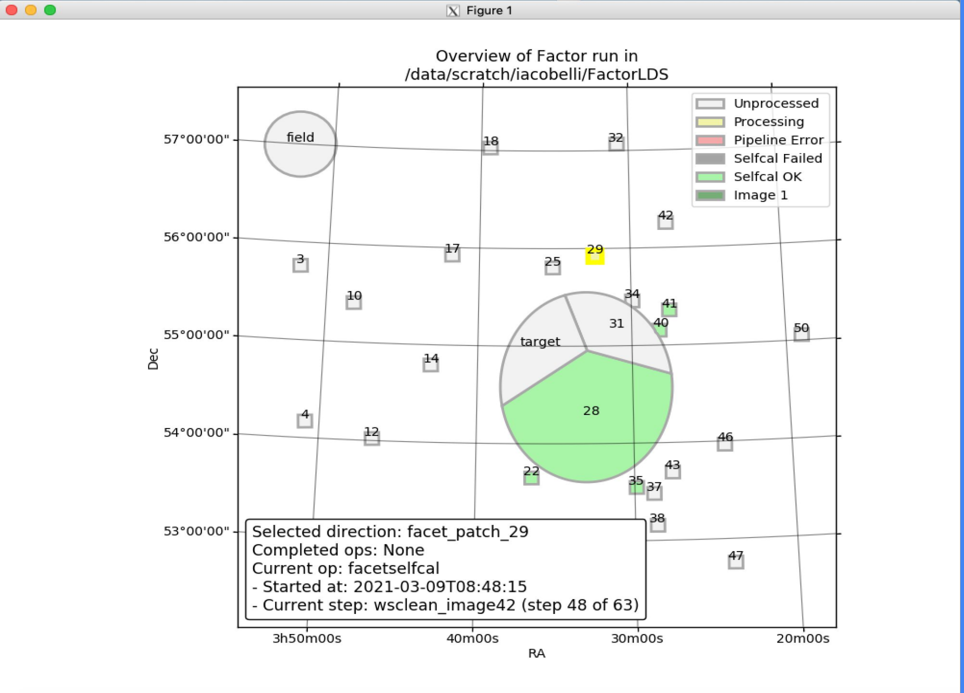
Factor performs Direction-Dependent calibration and imaging

- Uses LOFAR generic pipeline framework as backend:
  - Allows distribution over cluster nodes
  - Allows resuming of interrupted jobs
- Available from GitHub at https://github.com/lofar-astron/factor
- Relatively modest resources required: 2 CPUs, 32 GB memory, 1TB of disk space



Introduction: Factor

- subtract best model of the sky (pre-Factor Initial-Subtract, see T2)
- divide field into facets based on bright DD calibrators
- cycle over facets:
  - add back the calibrator source
  - self calibrate the calibrator sources (`facetselfcal`)
  - add back the rest of the facet and apply calibration solutions
  - improve the subtraction with new model and calibration (`facetsub`)
- image the facets (`facetimage`)
- make a mosaic of all facets & correct for the primary beam attenuation (`fieldmosaic`)




```
Singularity lds-img-cep3.sif:/data/scratch/iacobelli/FactorLDS> checkfactor factor2.0.parset
/opt/lofarsoft/lib/python2.7/site-packages/lofarpipe/support/utilities.py : Using default
subprocess module!
/usr/local/lib/python2.7/dist-packages/matplotlib/cbook/deprecation.py:107:
MatplotlibDeprecationWarning: The mpl_toolkits.axes_grid module was deprecated in version 2.1.
Use mpl_toolkits.axes_grid1 and mpl_toolkits.axisartist provies the same functionality instead.
  warnings.warn(message, mplDeprecation, stacklevel=1)
INFO - factor:progress - Plotting directions...
INFO - factor:progress - Left-click on a direction to select it and see its current state
INFO - factor:progress - Right-click on a direction to deselect it
INFO - factor:progress - (In both cases, pan/zoom mode must be off)
INFO - factor:progress - Press "c" to display calibrator selfcal images for selected direction
INFO - factor:progress - Press "i" to display facet image for selected direction
INFO - factor:progress - Press "v" to display facet verify image for selected direction
INFO - factor:progress - Press "t" to display TEC solutions for selected direction
INFO - factor:progress - Press "g" to display Gain solutions for selected direction
INFO - factor:progress - Press "u" to update display (display is updated automatically every
minute)
INFO - factor:progress - Press "h" to repeat these instructions on this terminal
```

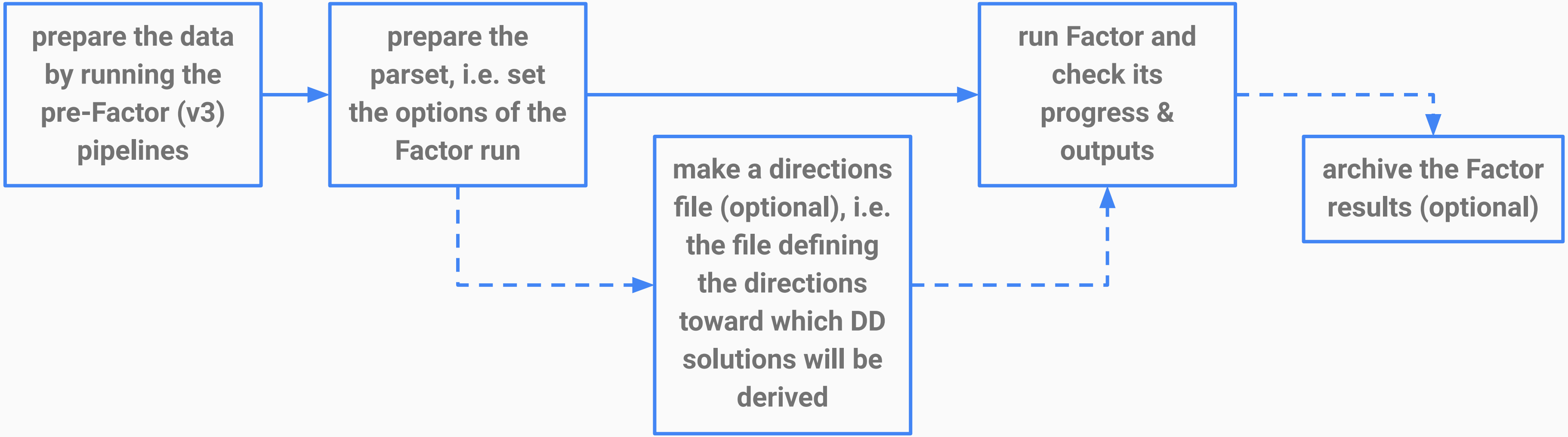
Introduction: Factor

Factor workflow can be monitored interactively with `checkfactor`:

```
$ checkfactor <parset>
```

Workflow of a Factor run

Using Factor on a dataset generally involves the following basic steps:



Data preparation

First calibrate data via pre-Factor pipelines https://github.com/lofar-astron/prefactor

- input MS files (uv data) consist of concatenated bands of e.g. 10 subbands each (i.e. = 2 MHz); typical observations will have 24 bands (= 48 MHz total)

After calibration, the Initial-subtract pre-Factor pipeline must be run to:

- image the field at medium and low resolution to make initial sky models
- subtract these models of the sources from the uv data

`L644395_SB000_uv_12BAF0A46t_163MHz.pre-cal.merge` → sky model for 2-MHz band at 163 MHz

`L644395_SB000_uv_12BAF0A46t_163MHz.pre-cal.ms` → concatenated MS for 2-MHz band at 163 MHz

Setting up the Factor pipeline

Make a directory in your area to hold the parset and Factor output:

```
$ mkdir /your/FactorTutorial/dir
```

For this tutorial, copy the tutorial parset to the directory you made above:

```
$ cd /your/FactorTutorial/dir
```

Edit your copy of the parset:

```
$ emacs /your/FactorTutorial/dir/factor.parset
```

The Factor parset

The parset is divided into sections to control and specify the workflow:

`[global], [calibration], [imaging], [directions], [cluster], [checkfactor], [<input ms>]`

- See http://www.astron.nl/citt/facet-doc/parset.html for a full description.
- An example parset that you can use as a basis for your reduction is available at https://www.astron.nl/citt/facet-doc/parset.html
- A customized parset for this tutorial is also shared on the Slack channel. For most parameters, the default values will be fine.

The Factor parset

The parset is divided into sections to control and specify the workflow:

`[global] [calibration] [imaging] [directions] [cluster] [checkfactor] [<input ms>])`

Section	Option	Value	Notes
global	dir_working	`/your/FactorTutorial/dir`	
global	dir_ms	`/your/FactorTutorial/dir`	
global	min_fraction_per_band	`0.8`	A large value set to maximise the S/N
global	use_compression	`True`	For less storage usage & faster processing

The Factor parset

The parset is divided into sections to control and specify the workflow:

`[global]` `[calibration]` `[imaging]` `[directions]` `[cluster]` `[checkfactor]` `[<input ms>])`

Section	Option	Value	Notes
calibration	max_selfcal_loops target_max_selfcal_loops	5 5	For faster processing
calibration	preaverage_flux_Jy	1.0	To maximise the S/N
calibration	multires_selfcal	True	To improve convergence of self-cal

The Factor parset

The parset is divided into sections to control and specify the workflow:

`[global]` `[calibration]` `[imaging]` `[directions]` `[cluster]` `[checkfactor]` `[<input ms>])`

Section	Option	Value	Notes
imaging	wsclean_nchannels_factor	`2`	For faster processing
imaging	fractional_bandwidth_selfcal_facet_image	`0.75`	Large bandwidth for a better subtraction
imaging	wsclean_use_idg idg_mode	`True` `cpu`	For faster processing

The Factor parset

The parset is divided into sections to control and specify the workflow:

`[global]` `[calibration]` `[imaging]` `[directions]` `[cluster]` `[checkfactor]` `[<input ms>])`

Section	Option	Value	Notes
directions	max_radius_deg	3	For faster processing
directions	flux_min_for_merging_Jy size_max_arcmin separation_max_arcmin flux_min_Jy	0.4 1.0 7.0 0.4	Constraints to select calibrators automatically (rule of thumb: bright and compact)
directions	minimize_nonuniformity	True	To make size of facets more uniform in size

The Factor parset

The parset is divided into sections to control and specify the workflow:

`[global]` `[calibration]` `[imaging]` `[directions]` `[cluster]` `[checkfactor]` `[<input ms>])`

Section	Option	Value	Notes
directions	faceting_radius_deg	`0.8`	For faster processing

directions	target_ra
	target_dec
	target_radius_arcmin
	target_has_own_facet
		`03h35m23.407`
		`+55d04m07.124`
		`4.0`
		`True`
			To specify a target area to ensure that a facet edge does not cross it. For this tutorial, we specify a region around a faint extended source

directions	minimize_nonuniformity	`True`	To make size of facets more uniform in size

The Factor parset

The parset is divided into sections to control and specify the workflow:

`[global]` `[calibration]` `[imaging]` `[directions]` `[cluster]` `[checkfactor]` `[<input ms>])`

Section	Option	Value	Notes
cluster	ncpu	`all`	Tune it based on your machine features !
cluster	wsclean_fmem	`0.9`	Tune it based on your machine features !
cluster	ndir_per_node	`1`	Tune it based on your machine features !

The Factor parset

The parset is divided into sections to control and specify the workflow:

`[global]` `[calibration]` `[imaging]` `[directions]` `[cluster]` `[checkfactor]` `[<input ms>])`

Section	Option	Value	Notes
checkfactor	facet_viewer	`ds9`	
checkfactor	ds9_load_regions	`False`	Set to True for inspection
checkfactor	image_display	`eog`	

The directions (facets) file

Factor records selected DDE calibrators in a file called `factor_directions.txt`

```
$ more factor_directions.txt
# name position atrous_do mscale_field_do cal_imsize solint_ph solint_amp dynamic_range region_selfcal region_facet
peel_skymodel outlier_source cal_size_deg cal_flux_mJy

# LSMTool history:
...

facet_patch_41 3h27m31.1093s,55d21m30.5782s empty empty 0 0 0 LD empty empty empty False 0.1050687533 36442.853868
facet_patch_35 3h29m52.7506s,53d32m43.7471s empty empty 0 0 0 LD empty empty empty False 0.0128485275297 10059.9407256
facet_patch_22 3h36m28.7034s,53d38m56.1925s empty empty 0 0 0 LD empty empty empty False 0.0112901740997 8669.60507632
facet_patch_40 3h28m11.8884s,55d09m08.5092s empty empty 0 0 0 LD empty empty empty False 0.0110088789021 7315.00279903
...
```

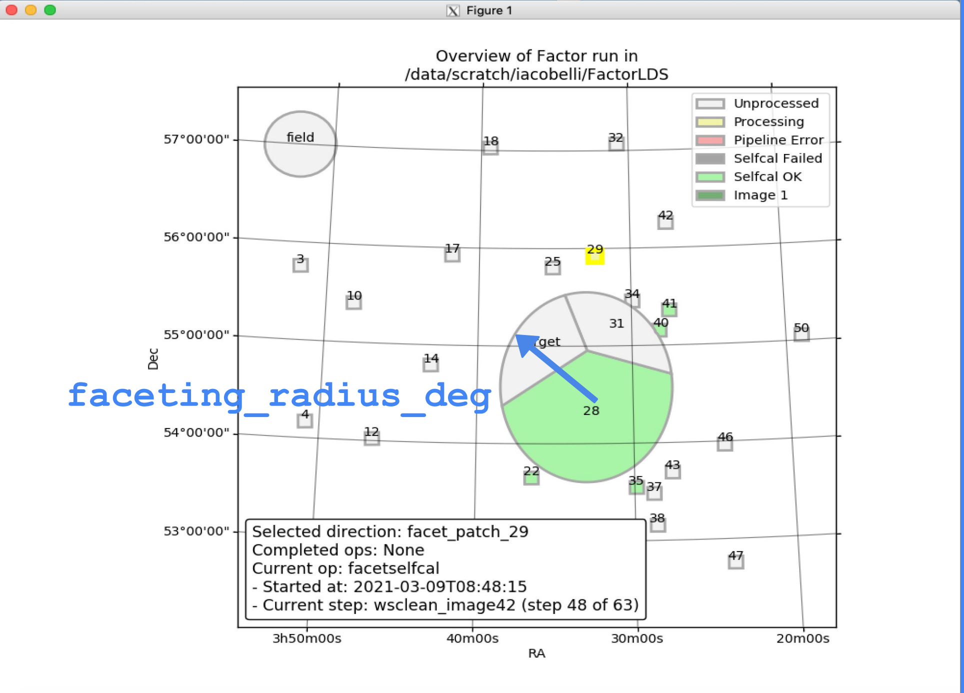
Direction (facet) name

Direction (facet) RA

Direction (facet) Dec

Dynamic range

Approximate flux density



Factor will divide the field into facets (one for each DDE calibrator) inside of `faceting_radius_deg` (under the `[directions]` section of the parset).
Outside of this radius, boxes are used instead (to speed up processing).

The directions (facets) file: editing

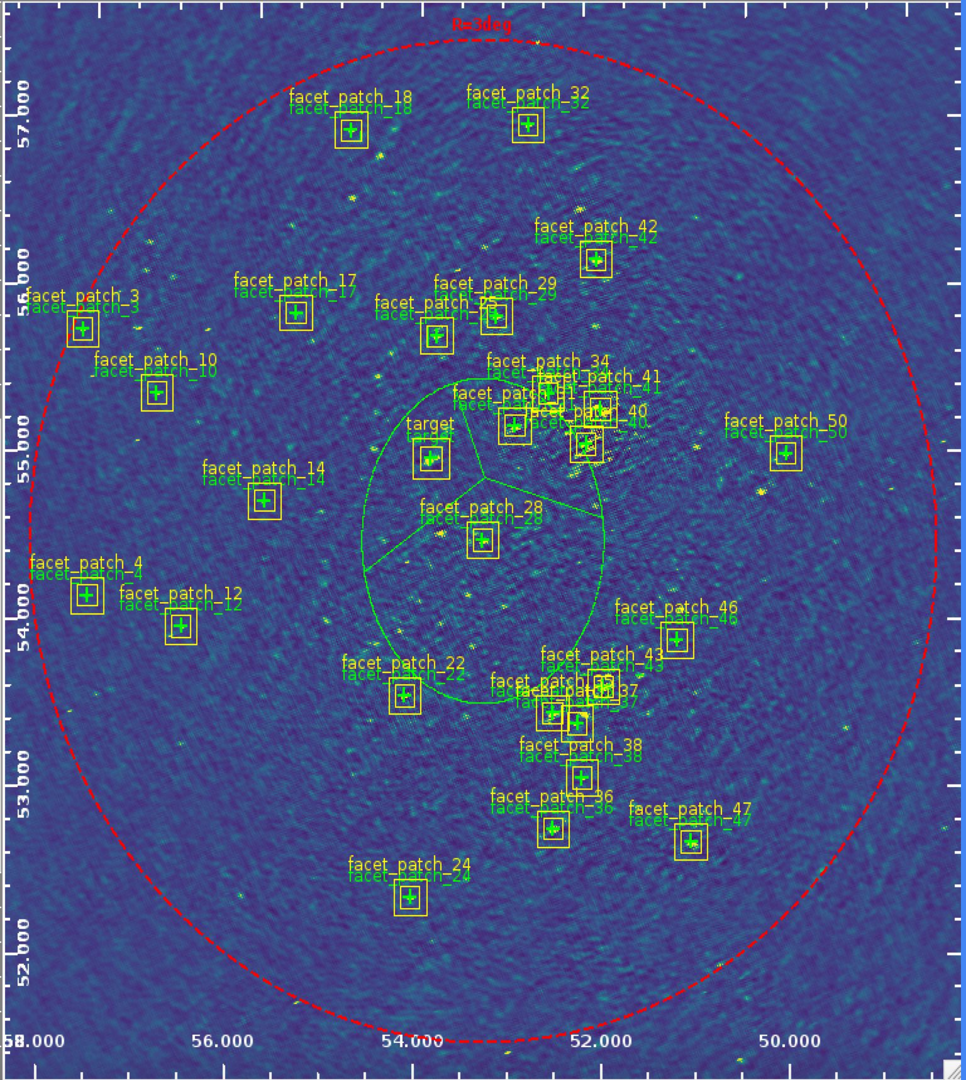
Directions can be:

- added or removed to improve facet layout
- reordered (e.g., in order to process only a subset of directions)

Other, advanced options can be activated on a per-direction basis (see online doc)

Warning

- after the first direction has completed for a run, changes to the facet layout will result in incorrect results!



Verification of DDE calibrators

To check pre-selected calibrators and facets use auxiliary files saved in `/your/FactorTutorial/dir/regions`:

- load a wide-field image of the field in ds9 (e.g., one made by the Initial Subtract pre-Factor pipeline)
- load (under the "Region" menu) files named
  - `calimages_ds9.reg`
  - `facets_ds9.reg`

Warning

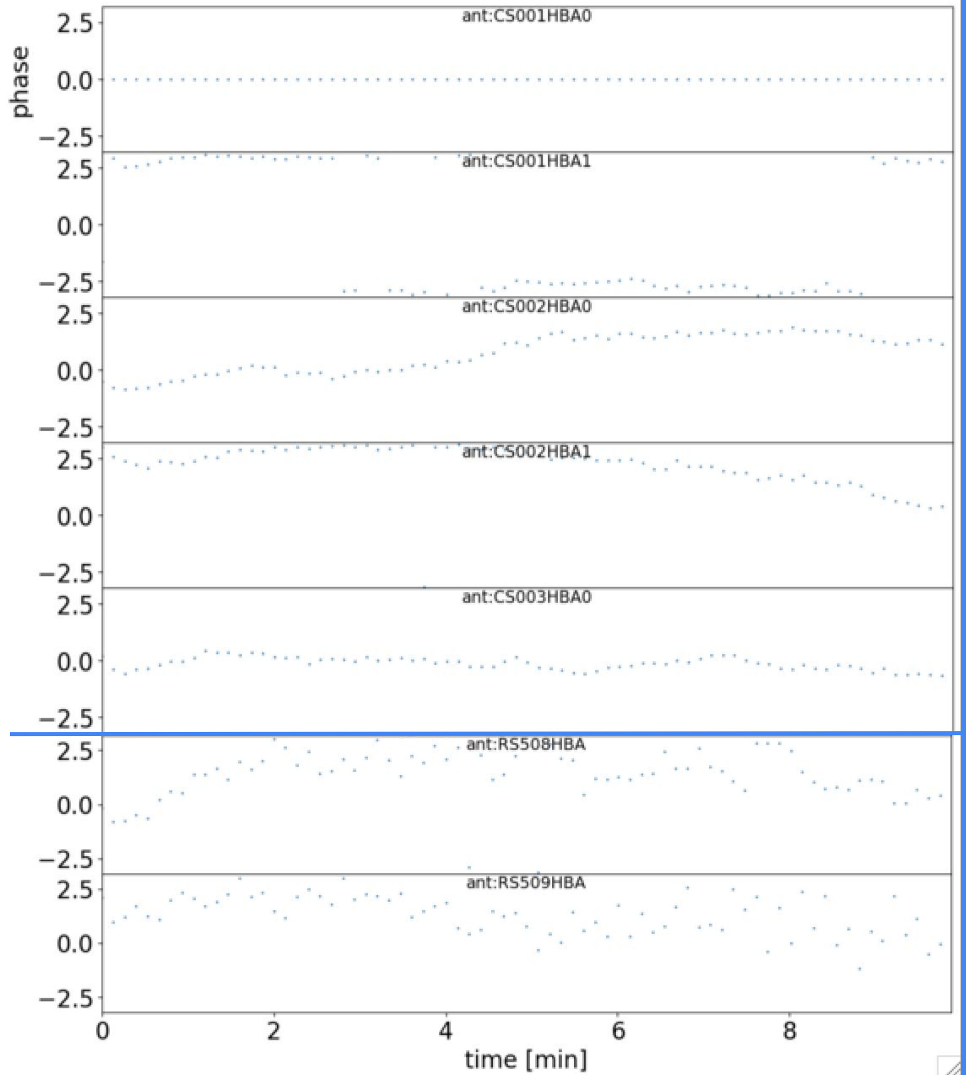
- To check & edit DDE calibrators list set in the **`[global]`** section of the parset `interactive=True`

The selfcal strategy

Selfcal is done on each calibrator to improve its model and DDE corrections

Factor minimizes the number of free parameters solved for during selfcal in order to avoid overfitting:

- **Fast phases (TEC)**: one solution every ~10–20 seconds and ~10 MHz to track rapid changes due to ionosphere
- **Slow gains (amp+phase)**: one solution every ~10–20 minutes and ~2 MHz to correct for beam effects

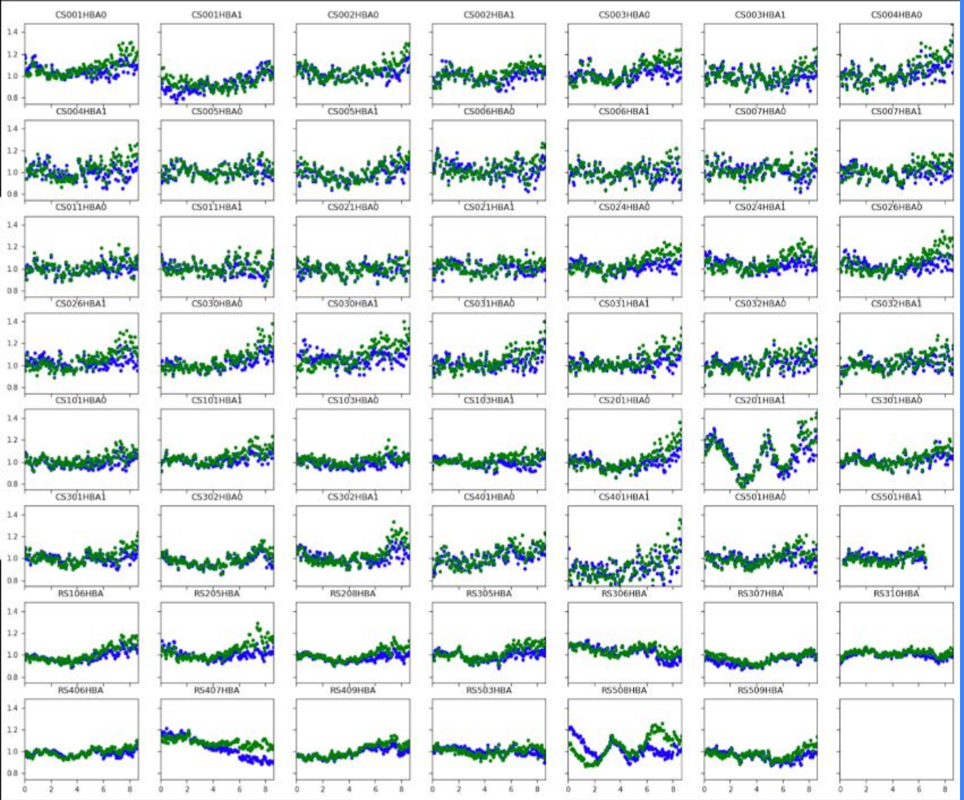


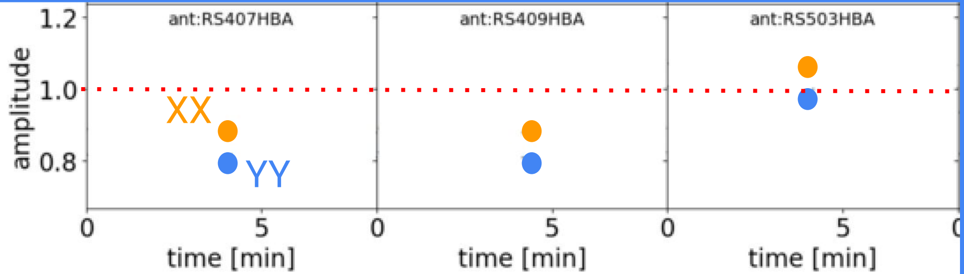
The selfcal strategy

TEC (from fast-phases): correction for ionospheric errors

- Tracking of phase changes has different S/N for different DDE calibrators as well as stations
  - The brightest and compact the DDE calibrator is the better results expected
- Quicker phase changes expected for stations further away from the reference one
- Noisy periods or jumps indicate problems !

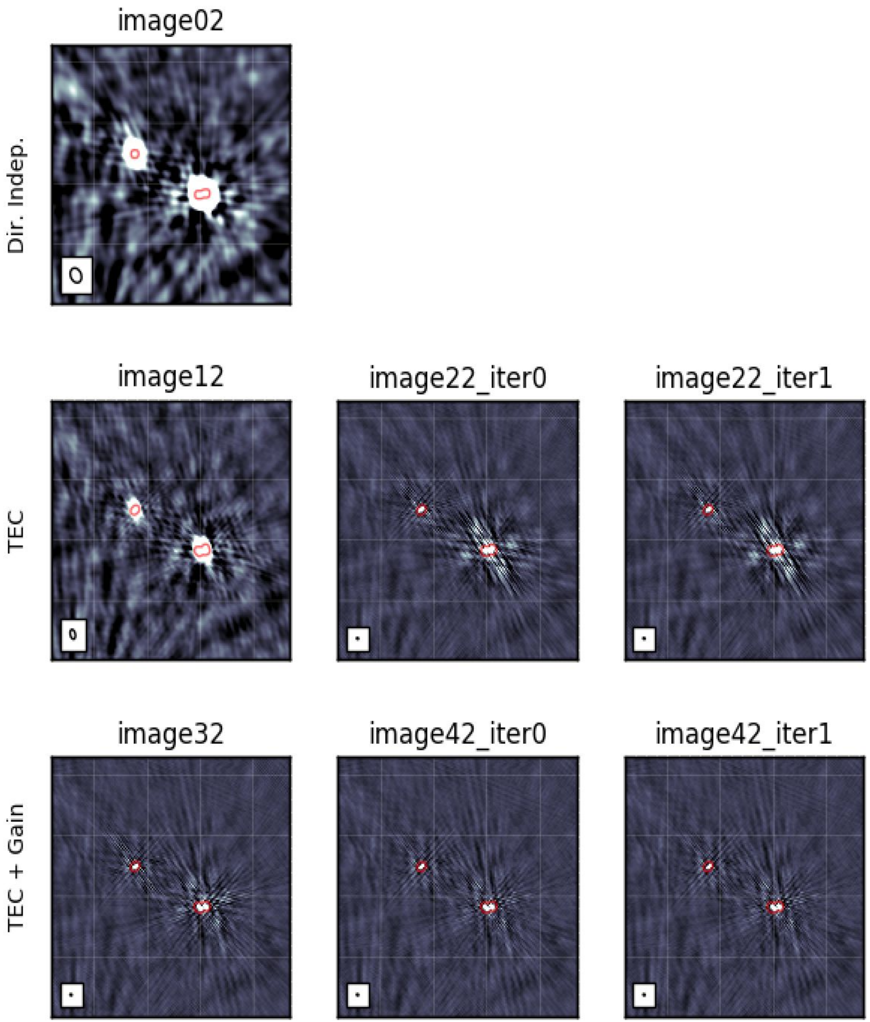
The selfcal strategy





Amplitudes (from slow-gains): correction for beam effects

- Tracking of amplitude changes has different S/N for different DDE calibrators as well as stations
  - The brightest and compact the DDE calibrator is the better results expected
- Amplitude correction values for XX and YY should be close to 1; systematic deviations suspicious:
  - resolved source structure,
  - Instrumental systematics



The selfcal strategy

Results of iterative selfcal (DIE, TEC, amplitude and TEC) corrections on the image plane are provided as useful diagnostic per DDE calibrator (see next section)

Running Factor

A run of Factor is started via the command

```
$ runfactor -v <your factor>.parset
```

Because the expected pipeline runtime is >2 days run the pipeline in a screen session!

Running Factor

Once Factor has started, you can check the progress of a run with the checkfactor tool:

```
$ checkfactor <your factor>.parset
```

```
INFO - factor:progress - Plotting directions...
INFO - factor:progress - Left-click on a direction to select it and see its current state
INFO - factor:progress - Right-click on a direction to deselect it
INFO - factor:progress - (In both cases, pan/zoom mode must be off)
INFO - factor:progress - Press "c" to display calibrator selfcal images for selected direction
INFO - factor:progress - Press "i" to display facet image for selected direction
INFO - factor:progress - Press "v" to display facet verify image for selected direction
INFO - factor:progress - Press "t" to display TEC solutions for selected direction
INFO - factor:progress - Press "g" to display Gain solutions for selected direction
INFO - factor:progress - Press "u" to update display (display is updated automatically every minute)
INFO - factor:progress - Press "h" to repeat these instructions on this terminal
```

After a short time, a window should appear showing the facet layout

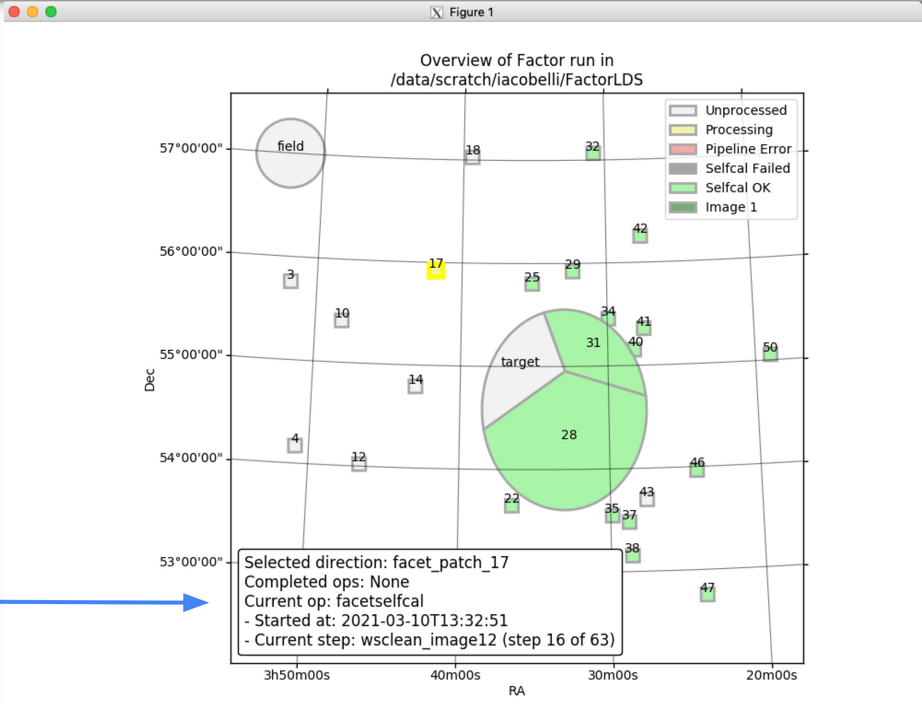
Running Factor

Click on a facet

- yellow: the one being processed
- green: the one completed

Hit the “c” key to see the calibrator images made during selfcal

Summary of processing for selected facet



Running Factor

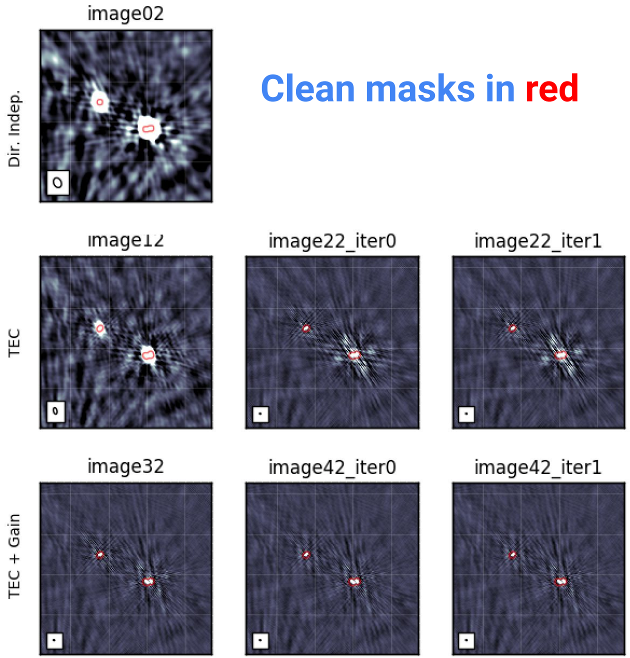
Click on a facet

- yellow: the one being processed
- green: the one completed

Hit the “c” key to see the calibrator images made during selfcal

Image before any direction dependent calibration

Clean masks in red



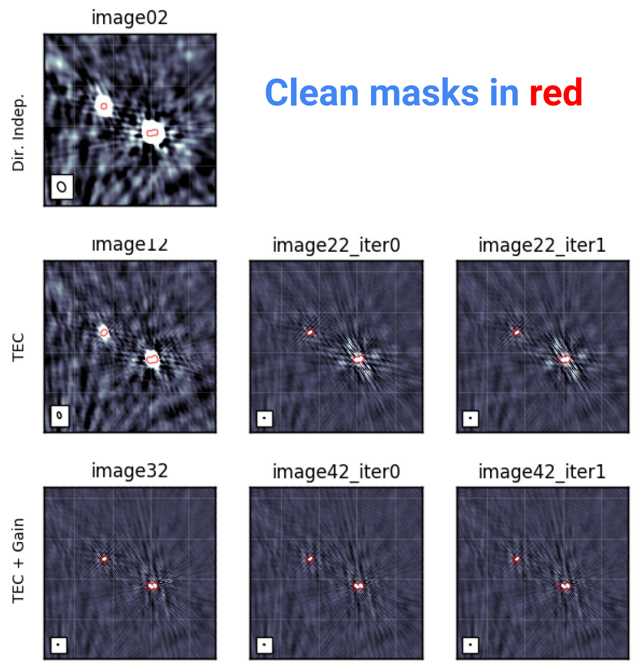
Running Factor

Click on a facet

- yellow: the one being processed
- green: the one completed

Hit the “c” key to see the calibrator images made during selfcal

Images after fast phases calibrations



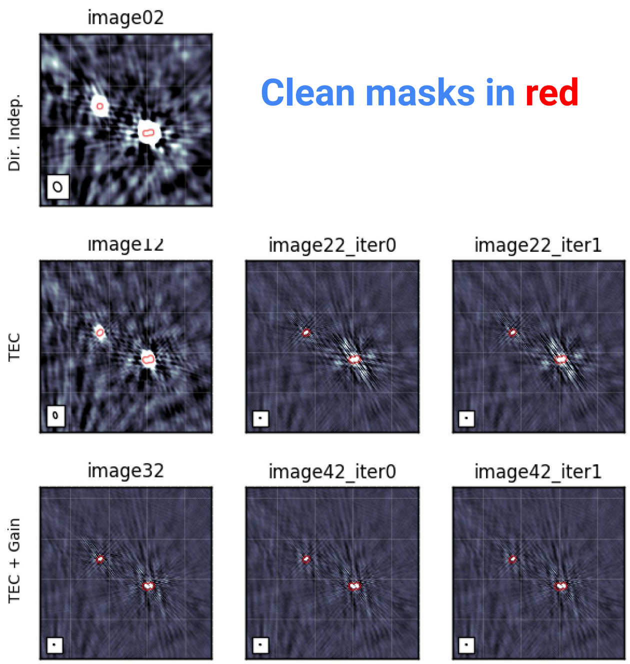
Running Factor

Click on a facet

- yellow: the one being processed
- green: the one completed

Hit the “`c`” key to see the calibrator images made during selfcal

Images after slow gain calibrations



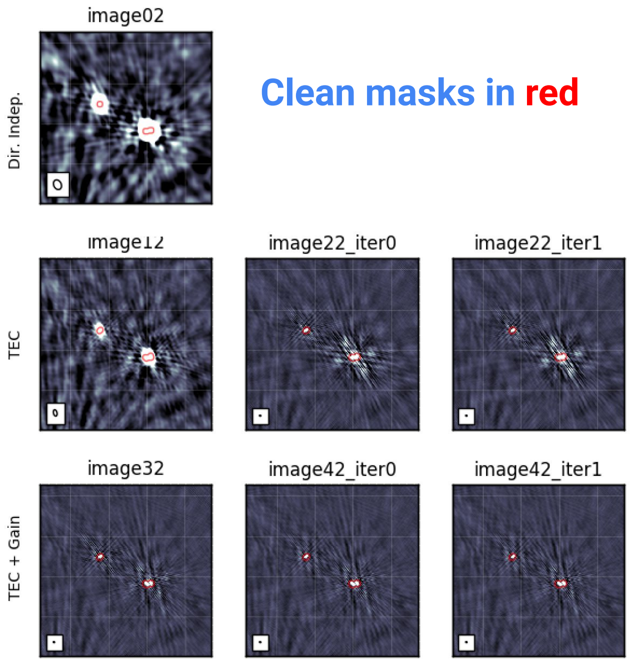
Running Factor

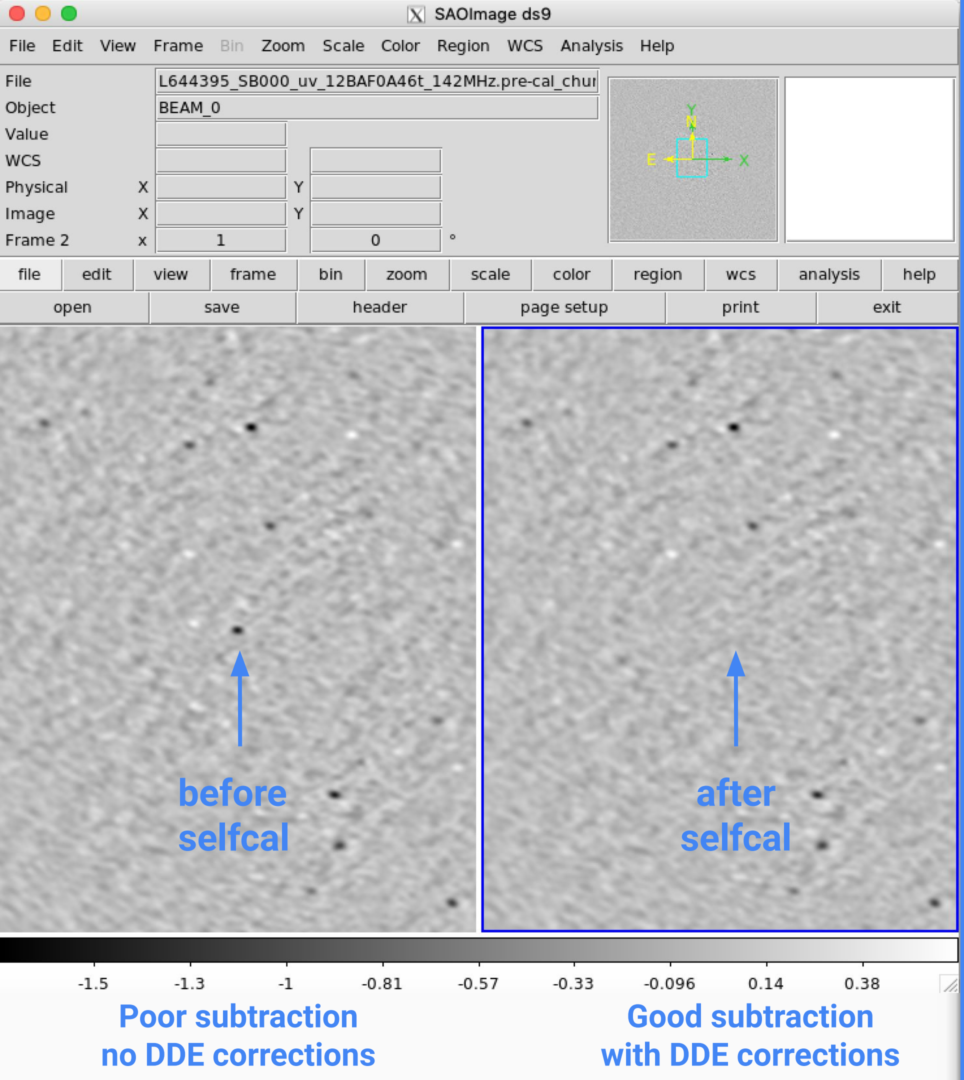
Click on a facet

- yellow: the one being processed
- green: the one completed

Hit the "`c`" key to see the calibrator images made during selfcal

The tutorial run setup was with `max_selfcal_loops = 5`
Factor stops after no more improvement is seen !





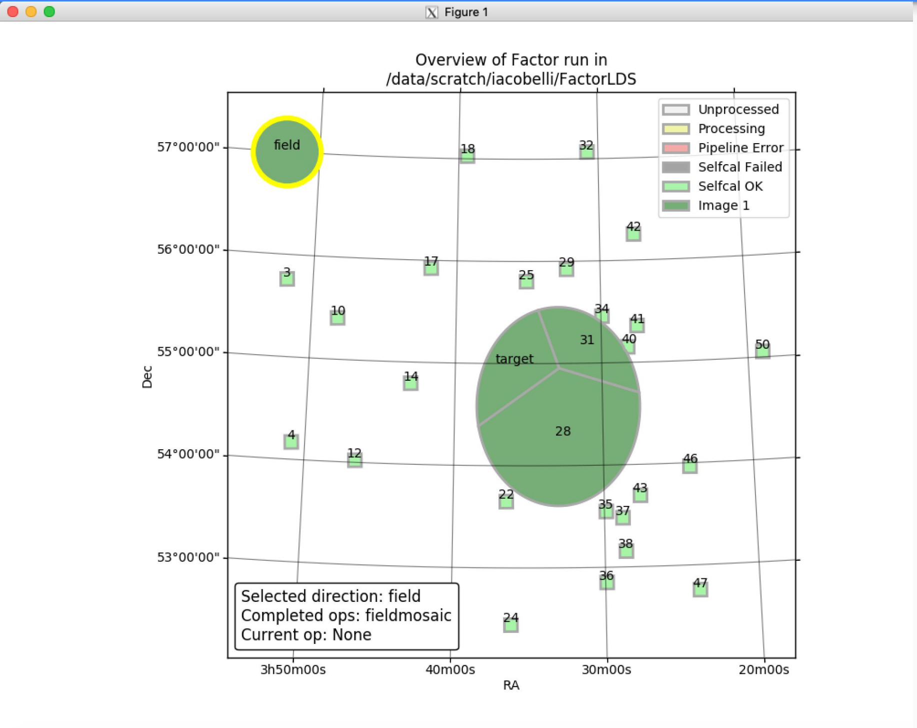
Running Factor

Click on a facet

- yellow: the one being processed
- green: the one completed

Hit the "`v`" key to open the selfcal verification (i.e. the residual, source subtracted) images in `ds9` (the facetselfcal operation must be finished first). This may take a few seconds..

Running Factor

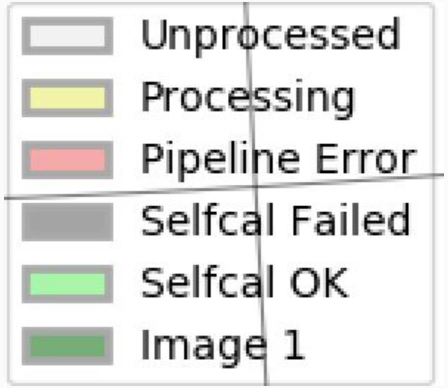


Once selfcal has completed for all the facets, Factor re-images each facets using the full bandwidth.

- These are the final images, but must be corrected for the primary beam attenuation !

Once the reimaging is complete, Factor will mosaic the images together and correct them for the primary beam

Running Factor



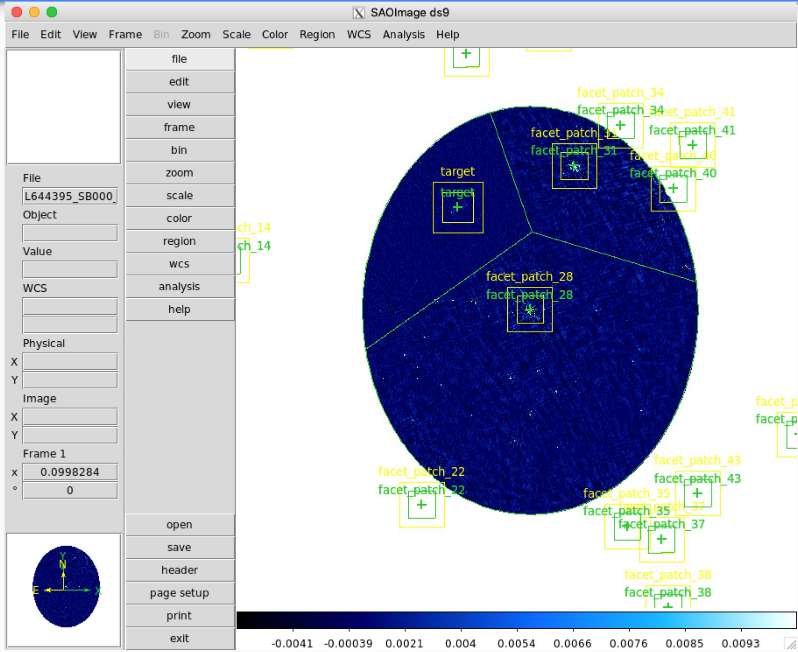
For each facet Factor will make one or more images, as set by the WSClean imaging parameters in the **`[imaging]`** section of the parset, e.g.

```
facet_cellsize_arcsec = [ 1.5, 5.0]

facet_robust          = [-0.5, -1.0]

facet_taper_arcsec    = [ 0.0, 15.0]

facet_min_uv_lambda        =   [80.0, 80.0]
```

Running Factor

Select the “field” direction (in the top-left corner)

- Hit the “`i`” key to list the available field images
- Hit the “`1`” key to open the full- resolution image in `ds9` (this may take a few seconds...)



Facets regions can be loaded:

- via the `[checkfactor]`section of the parset.
- via the `ds9` “Region” menu (in `Factor_output/regions/facets_ds9.reg`)

Running Factor

Two mosaics are made for each set of imaging parameters:

- Non-primary-beam corrected ("flat noise"), loaded by checkfactor (for inspection) `Factor_output/results/fieldmosaic/field/*.correct_mosaic.`**`pbcut`**`.fits`
- Primary-beam corrected image needed for flux measurements: `Factor_output/results/fieldmosaic/field/*.correct_mosaic.`**`pbcor`**`.fits`

Due to problems with the LOFAR beam model, the flux scale can be off by up to 15%. A check against e.g. TGSS catalog can be done (see T11)

Outline

Lecture

DDE calibration/imaging & Factor in a nutshell
The Factor parset & directions file
The selfcal strategy
Running Factor

Tutorial handout

To get started

- Make a directory in your area to hold the
  - parset,
  - Factor input data (10GB) & output (+10GB)
- Open in your web browser the online full documentation https://www.astron.nl/citt/facet-doc/

Input data & parset

- Collect input data in a single directory
  - Input data are 23 subbands (= 2 MHz) of a snapshot (10-min) HBA observation of the PSR0329+54
  - Averaged by preFactor to 0.05 MHz per channel and 8 seconds per timeslot
- Create & edit your parset file (see Lecture notes under “The Factor parset”):

```
$ touch factor2.parset

$ emacs factor2.parset
```

Starting Factor

- Enter a screen session

```
$ screen [+enter]
```

- Initialise your Singularity environment, mounting your Factor directory

```
$ singularity shell --bind /your/FactorTutorial/dir/, /your/Singularity.sif
$ source /opt/lofarsoft/lofarinit.sh
```

- Run Factor in the screen session with the parset you made

```
$ runfactor -v factor2.parset
```

- Exit from the screen (CTRL-A CTRL-D)

Check DDE calibrators & facets

- When the factor_directions.txt file is made, initialise your Singularity environment, mounting your Factor directory

```
$ singularity shell --bind /your/FactorTutorial/dir/, /your/Singulatity.sif
```

```
$ source /opt/lofarsoft/lofarinit.sh
```

- Load in `ds9` (1) the provided wide-field image of the field (obtained by the Initial Subtract preFactor pipeline) and (2) Factor auxiliary `*.reg` files (see Lecture notes under “Verification of DDE calibrators”)

```
$ ds9 L644395_SB000_uv_12BAF0A46t_163MHz.pre-cal.wsclean_high-image.fits
```

- Next check the progress of the run (see Lecture notes under “Running Factor”)

6th LOFAR data processing school
March 22/26 2021