ALMA Data & Pipeline

B.E. Glendenning (NRAO)
ALMA Computing Integrated Product Team Lead

ALMA Status

- Commissioning and Science Verification in progress
  - 11 (of 66) antennas in use at the high site
  - Many more are being assembled, integrated, tested, … at the low site
- Cycle 0: Call on March 30, Start of observations Sept 30
  - <1000 hours, 9 months, 16+ antennas
- The ALMA data management system was designed in from the beginning and will be delivered with the telescope


NRAO

ALMA Data

- Raw Data
  - Correlator > 10^{17} FLOPs
  - ALMA (baseline) correlator capable of 1 GB/s = 43 TB/day into the Archive (visibilities usually compressed 2x).
  - End of construction peak rate = 60 MB/s = 5 TB/day
    - Average presumed to be 10% of this = 6 MB/s = 0.5 TB/day
  - Format (complicated, ~40 table types) common between ALMA & EVLA
- Pixel Data
  - TPixel ($10,000^3$) images will not be common but will be made (0.006” finest resolution)
- Many observers have become used to radio interferometer data processing being a laptop scale problem

ALMA Software

- Developed in a large consortium (15 locations, 4 continents)
- Integrated system
  - Data models & technology
  - For example, goals and intent from Phase II flow through the observing system to the pipeline
- ~6% of the ALMA construction budget

Observatory Software

Software required by ALMA to interact with the observing community, optimize the observing process, and quality check what has been observed

- Distributed science archive
- Science Pipeline
- Phase 1 & phase 2 Observing Tool
- TAC support
- Observing project tracking
- QA metrics extraction & tracking
- User Portal
- Data packaging & distribution
- Manual data processing/ analysis (CASA)

ALMA Science Pipeline

- The science pipeline is part of the effort to make ALMA accessible to the entire astronomical community
- The science pipeline will reduce data taken in standard observing / reduction modes
- The science pipeline will produce standard data products for the science archive
- The pipeline scripts will be made available for further offline processing at the ARCs and by observers
- Pipeline images will be given to observers even in cycle 0, although the pipeline will still need training

ALMA Pipeline Heuristics

- The ALMA Pipeline Heuristics attempt to capture user intent and expert knowledge and encode it in data reduction recipes
- There is one recipe per standard reduction mode
- Currently there are heuristics recipes for
  - Single field interferometry
  - Pointed mosaics interferometry
  - Single dish data
- IF/SD combination recipe is under development (not a released mode for Cycle 0)

List of Heuristics Tasks

Interferometry

30 tasks available.

task name	description
hif_applycal	Compute and apply or apply the calibration
hif_bandpass	Compute a bandpass calibration
hif_bestbandpass	Find the best bandpass calibration method
hif_close	Close the interfometry pipeline framework
hif_comment	Writes a comment to the pipeliene web log
hif_configpaths	Configure directories for the interferometry pipeline
hif_contsubclean	Create cleaned line and continuum images
hif_delaycal	Compute the delay correction
hif_evalbandpass	Evaluate quality of the bandpass solution
hif_findedgechan	Detect the bandpass edges
hif_findlines	Find line regions in cleaned cube image sources
hif_flagcalvis	Flag calibrated visibility data
hif_flagdata	Flag known bad data
hif_flagmodelvis	Flag model visibility data
hif_flagrawvis	Flag raw visibility data
hif_gaincal	Calculate the gain and flux calibration solutions
hif_grpgaincal	Calculate the gain and flux calibration for groups of spectral windows
hif_init	Initialize the interferometry pipeline framework
hif_moscalclean	Calibrates, images and cleans pointed mosaic targets
hif_mosclean	Image and clean pointed mosaic targets
hif_plotbandpass	Plot a bandpass calibration
hif_plotcalvis	Plot flagged calibrated visibility data
hif_plotgaincal	Plot a gain calibration
hif_plotmodelvis	Plot flagged model visibility data
hif_plotrawvis	Plot flagged raw visibility data
hif_plotspectra	Plot spectra of sources in cleaned cube images
hif_regsession	Register data session/vislist with the pipeline framework
hif_sfcalclean	Cleans or calibrates and cleans single field targets
hif_sfclean	Create single field cleaned images
hif_summary	Summarize an interferometry observation

Single-Dish

24 tasks available.

Interferometry Heuristics

- The interferometry recipe implements stages to
  - Perform basic flagging, e.g. shadowing
  - Flag bad raw data (coarse flagging)
  - Bandpass, gain, and flux calibrate the data
  - Flag bad calibrated data (finer flagging)
  - Image the data
  - Detect lines and determine line regions
  - Compute line and cotinuum images

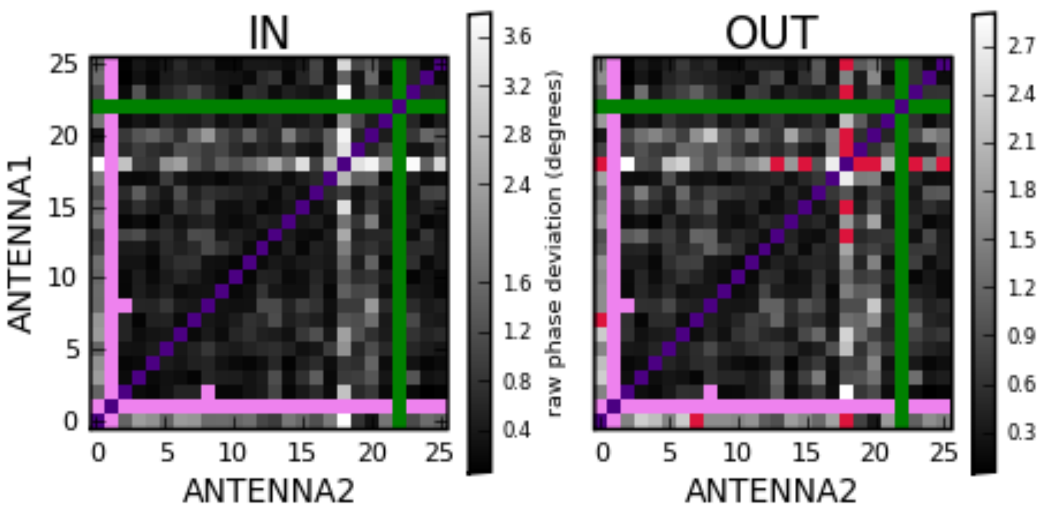
Interferometry Flagging Example

STAGE: 8 Flag calibrator baselines with noisy phases

Field:1331+305 (BANDPASS+FLUX)

SpW: 0 Pol:RR raw phase deviation

Flagging noisy calibrator phases. Each data point in the display shows a measure of the phase noise for that baseline.



No data

Flagged at previous stages:

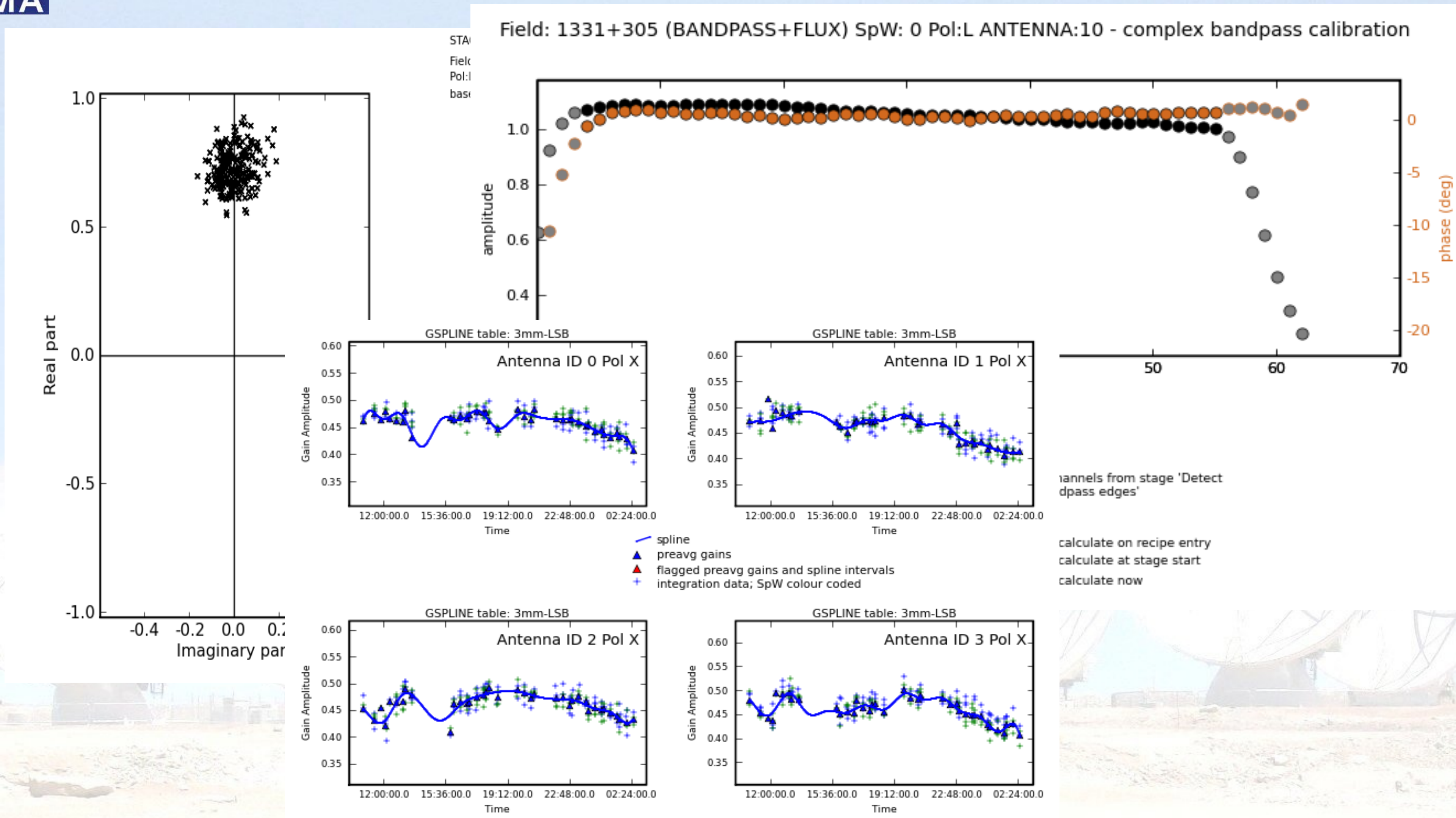
- cannot calculate on recipe entry
- cannot calculate at stage entry
- cannot calculate now

Flagged here:

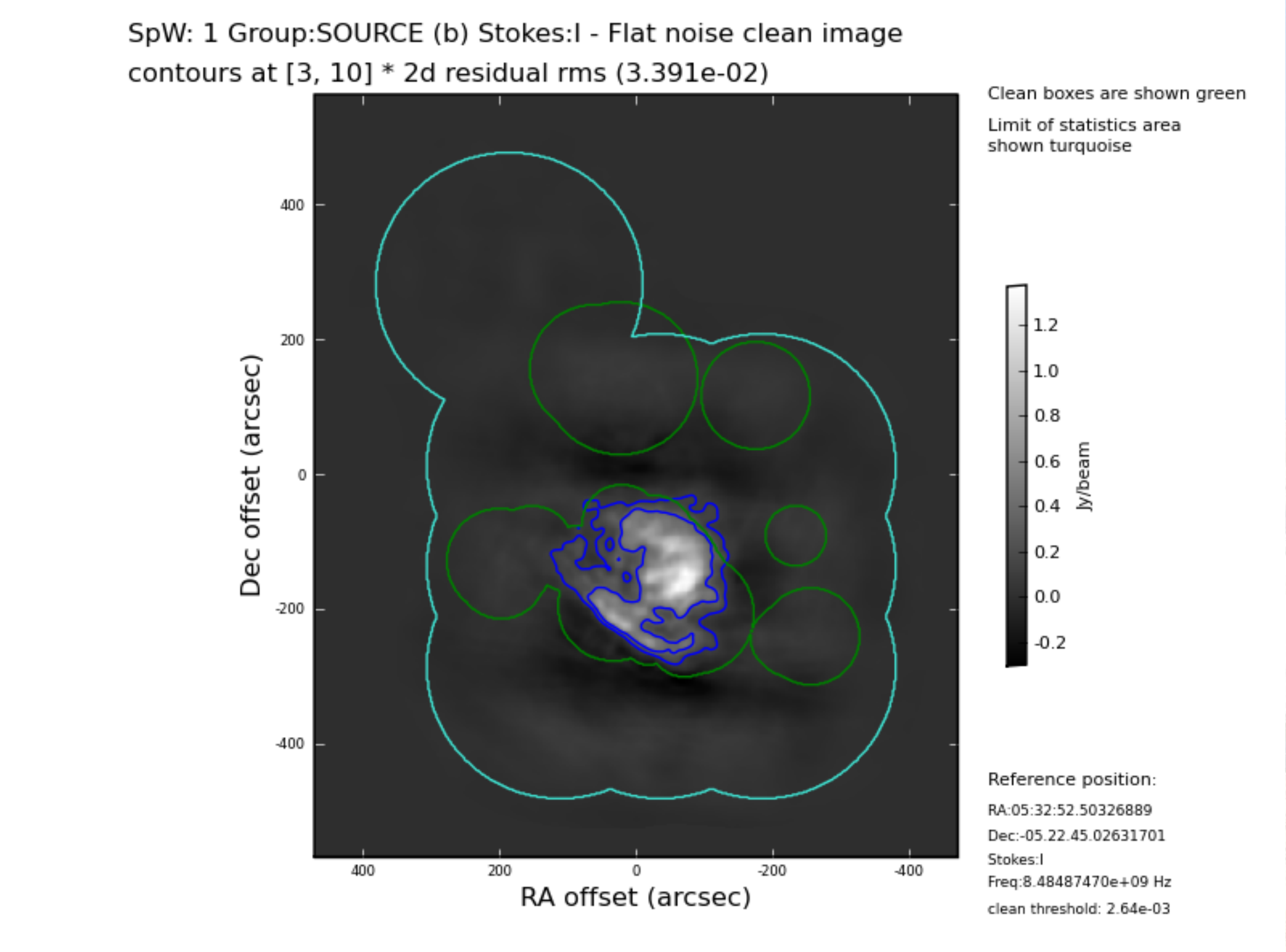
rules:

- high outlier
- ANTENNA2 axis - too many flags

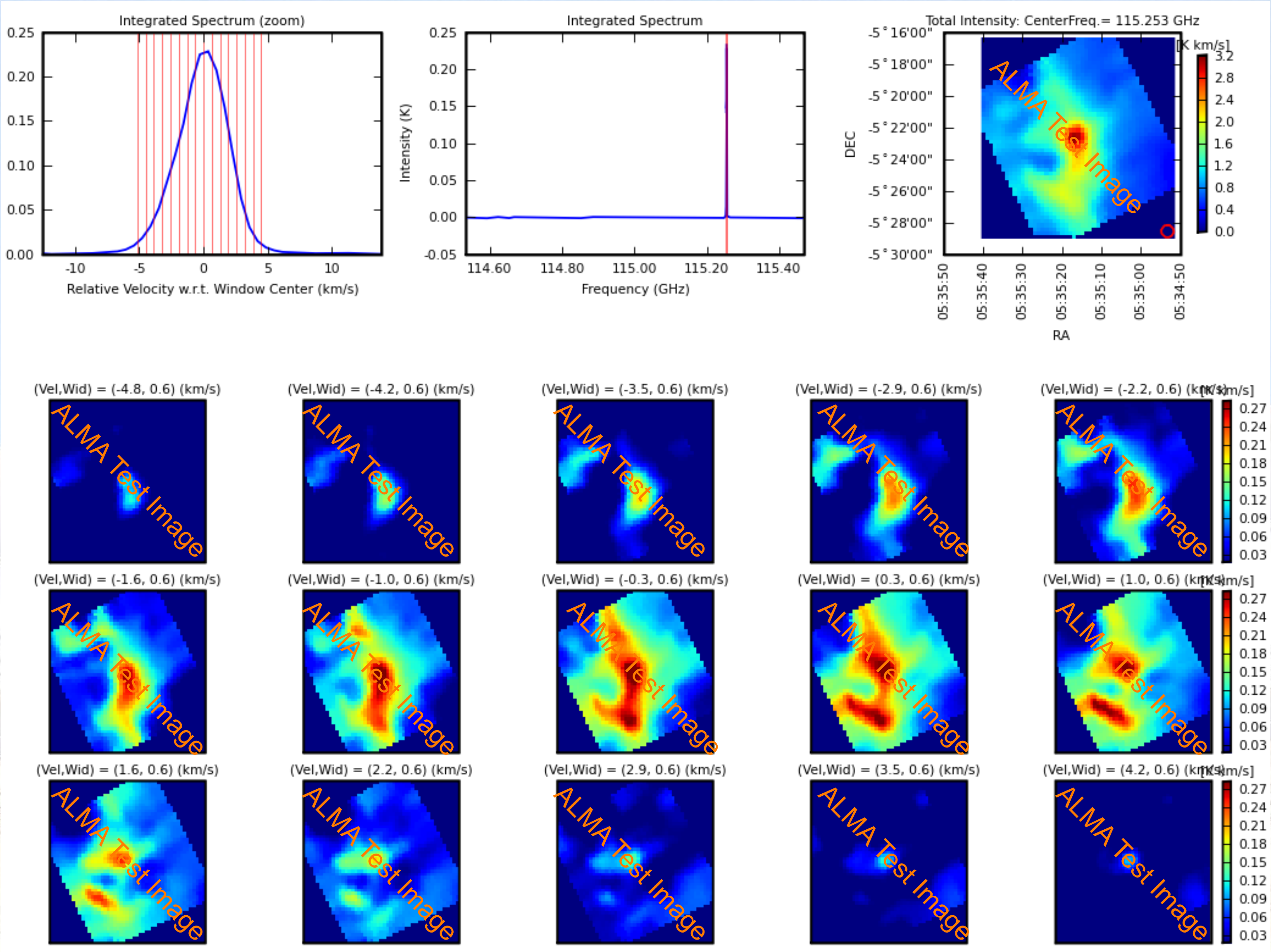
Interferometry Calibration Examples



Interferometry Imaging Example



ALMA Single Dish Test Data: Orion



ALMA Mosaic Test Data: QSO J0538-440