# High resolution/high frequency observing and data reduction

Anita Richards UK ALMA Regional Centre Jodrell Bank Centre for Astrophysics University of Manchester thanks to ALMA and JBCA colleagues



.MA



# Weighting

- Natural weighting: Grid and FT samples
  - Weight zero in grid cells with no samples
- Uniform weighting:
  - All cells same weight
    - Makes resolution finer, but increases noise
- Robust [-2, 2]: intermediate weightings
- Taper: reduce weight of longest baselines
  - Increase sensitivity to large-scale emission
    - Increases noise if effective collecting area reduced





# Outline

- High-frequency observing
  - Hazards
- Calibration
  - Instrumental measurements
  - Astrophysical standards and phase referencing
  - Self-calibration
- Imaging
  - Clean
  - Weighting
- ALMA data products

#### ALMA, NOEMA etc.



(a) Steerable dishes
(b) Subreflector
(c) (Multiple) receivers cabin
(d) Optical fibre links to correlator
(e) Reconfigurable

#### Also SMA $\geq$ 0.45 mm; VLA $\geq$ 6 mm; ATCA $\geq$ 3mm; KVN $\geq$ 2.5mm



#### core lobe

Delay

errors

(timing)

geom

Close to AGN: Scintillation

Lobes (and ionosphere) Faraday rotation

#### Galactic CO etc.

Atmospheric  $H_2O$  vapour (also  $O_3$  etc.) Absorption – reduces S/N Refraction - phase rotation

Antenna positions, efficiency, pointing

Atmospheric emission noise System (electronic) noise System (unit of the system of the

Hazards

RFI

000

Bandpass response (

#### (sub-)mm windows & bands





Frequency (GHz)

#### **Receiver temperatures**









- Combined  $\phi$  depends on  $\delta \textbf{s}$
- Complex visibility amplitude is sinusoidal function of  $\boldsymbol{\phi}$



# Earth rotates

- Telescopes separated by baseline B<sub>geom</sub>
- Earth rotates
  - Projected separation  $b = B_{geom} \cos \theta_0$
- Samples different scales of source
- Additional geometric delay path  $\Delta$ 
  - Remove in correlator



# Earth rotation aperture synthesis



- Combined  $\phi$  depends on  $\delta s(time)$ 

(3)

(2)

core

lobe

- Complex visibility amplitude is sinusoidal function of  $\boldsymbol{\phi}$ 

1/06

1/04

# Interferometry to images

- Correlation makes complex visibilities  $V_v(u_v, v_v)$ 
  - One V,  $Ae^{i\phi}$ , per v-chan per baseline per  $\tau_{int}$  per pol.
- Fourier transform gives sky brightness distribution  $\sum V_{v}(u_{v},v_{v}) e^{[2\pi i(uvl + vvm)]} du dv = I_{v}(l,m)$

or  $V(u,v) \Leftrightarrow I_v(l,m)$  for short



- Sensitivity:  $\sigma_{rms} \propto \frac{I_{sys}}{\sqrt{N(N-1)/2} \delta v \Delta t N_{pol}}$ 
  - Number antennas

Source

- dv freq. width per image,  $\Delta t$  total time on source

# High-frequency considerations

- Same principles as any radio interferometry
- You are *unlikely* to be bothered by:
  - Ionosphere ( $\delta delay_{ionosphere} \propto \lambda^2$ )
  - Confusion ( $\Theta_{PrimaryBeam} \sim \lambda/B \sim 55''$  @ 3mm, 12-m dish)
    - Most bright extragal. sources have  $\alpha$ <0 where  $S \propto v^{\alpha}$
- You *will* suffer from:
  - Tropospheric refraction ( $\delta\phi_{troposphere} \propto 1/\lambda$  )
  - Tropospheric absorption and emission
  - Small field of view ( $\Theta_{PrimaryBeam} \sim 9'' \oplus 0.45 \text{ mm}$ )
    - Mosaicing

#### High-frequency considerations



- Small field of view ( $\Theta_{PrimaryBeam} \sim 9'' \oplus 0.45 \text{ mm}$ )
  - Mosaicing

#### Estimate observability from low-res image

- Brightness temperature  $T_{\rm b} = S_{\rm source} 10^{-26} \lambda^2 / 2k_{\rm B} \Omega$ 
  - $-\Omega$  emitting area (sr),  $\lambda$  (m), S (Jy)
    - Resolved? Use S per measured Ω
    - Unresolved by low-resolution obs. e.g. single dish? – Estimate actual source size for  $\Omega$
  - Use S per best estimate of  $\Omega$  to find  $T_{\rm b}$
  - Predict ALMA flux density  $S_{ALMA} = T_b 2k_B \theta_b^2 / 10^{-26} \lambda^2$ 
    - Substitute  $\Omega = \theta_b^2$  (ALMA synthesised beam)
  - Use **Sensitivity Calculator** need  $>5\sigma_{rms}$  on peak
- e.g. total source 1 Jy,  $\sqrt{\Omega}$ =15",  $T_{\rm b}$ ~0.07 K at  $\lambda$  1mm

#### Estimate observability from low-res image

- Brightness temperature  $T_{\rm b} = S_{\rm source} \ 10^{-26} \ \lambda^2 \ / \ 2k_{\rm B} \ \Omega$ 
  - Ω emitting area (sr),  $\lambda$  (m), S (Jy)
    - Resolved? Use S per measured Ω
    - Unresolved by low-resolution obs. e.g. single dish?
       Estimate actual source size for Ω
  - Use S per best estimate of  $\Omega$  to find  $T_{\rm b}$
  - Predict ALMA flux density  $S_{ALMA} = T_b 2k_B \theta_b^2 / 10^{-26} \lambda^2$ 
    - Substitute  $\Omega = \theta_b^2$  (ALMA synthesised beam) •

– Use **Sensitivity Calculator** - need  $>5\sigma_{rms}$  on peak

- e.g. total source 1 Jy,  $\sqrt{\Omega}$ =15",  $T_{\rm b}$ ~0.07 K at  $\lambda$  1mm
  - $-\theta_{b}=0$ ".15, flux density 1x(15/0.15)<sup>2</sup> = 0.0001 Jy/beam

#### Visibility plane coverage

- ALMA main array: – fifty 12-m antennas
- ALMA Compact Array:
  - twelve 7-m antennas
  - four 12-m antennas for total power

10

5

-5

-10

-15-15

v (k.)



#### Combining arrays



# Correlator configurations

- Two sidebands, fixed spacing depending on band
  - e.g. B7, B3 sideband centres separated by 12 GHz
    - Max. continuous b/w <4 GHz (8 GHz Bands 9, 10)



- Four basebands (BB), max. width 2 GHz
  - 128 chans per BB (dual pol) TDM
  - 4096 chans per BB (~0.5 km/s at 300 GHz) FDM
    - Useful max. 1.875 GHz (so 3840 channels usable)
- Narrower spectral windows for higher resolution
  - Factors of two down to 62.5 MHz (15.25 kHz chans)
  - Higher spectral resolution in single polarization
- See documentation and OT for full details

# Correlator configurations

- Two sidebands, fixed spacing depending on band
  - e.g. B7, B3 sideband centres separated by 12 GHz
    - Max. continuous b/w <4 GHz (8 GHz Bands 9, 10)



- Four basebands (BB), max. width 2 GHz
  - 128 chans per BB (dual pol) TDM
  - 4096 chans per BB (~0.5 km/s at 300 GHz) FDM
    - Useful max. 1.875 GHz (so 3840 channels usable)
- Narrower spectral windows for higher resolution
  - Factors of two down to 62.5 MHz (15.25 kHz chans)
  - Higher spectral resolution in single polarization
- See documentation and OT for full details

#### Instrumental & atmospheric noise

• Noise  $\sigma_{sys}$  is given by  $\sigma_{sys} = \frac{T_{sys}}{\eta_A A_{eff} \sqrt{N(N-1)/2} \Delta v \Delta t N_{pol}}$ antenna area  $A_{eff}$ , efficiency  $\eta_A$ 

N antennas, frequency span  $\Delta v$ , time span  $\Delta t$ ,  $N_{pol}$  Rx pols

• System temperature  $T_{sys}$  $T_{sys}(DSB) = \frac{1+g_{SB}}{\eta_A e^{-\tau_{atm}}} [T_{Rx} + \eta_A T_{sky} + (1-\eta_A) T_{amb}]$ 

contributions from Receiver, sky and 'ambient' e.g. hardware, ground temperature

- $g_{SB} = 1$  if there is an unwanted if sideband
- Noise increases exponentially with opacity and zenith angle

$$T_{received} = T_{source} e^{\tau_{atm}/\cos z} + T_{atm} (1 - e^{\tau_{atm}/\cos z})$$

#### Absorption and emission

The atmosphere both absorbs the astrophysical signal, and adds noise

$$T_{received} = T_{source} e^{\tau_{atm}/\cos z} + T_{atm} (1 - e^{\tau_{atm}/\cos z})$$

where the source would provide temperature T if measured above the atmosphere and z is the zenith distance

- Same source, same baselines
  - Raw amps lower at higher **P**reciptable **W**ater **V**apour



#### Phase errors cause position errors

 Averaging over phase fluctuations causes decorrelation of amplitudes

– Visibility 
$$V = V_o e^{i^{\varphi}}$$

$$\langle V \rangle = V_o \langle e^{i \varphi} \rangle = V_o e^{-(\varphi_{rms}^2)/2}$$

 $\phi_{rms}$  in radians Lose 9% amplitude for 5°  $\phi_{rms}$ 

- Fluctuations on time-scales

   of few sec: raw data position jitter<sub>0.4</sub>
- Water Vapour Radiometry
  - Measure PWV, each antenna, every ~second
  - Calculate phase delay
    - Apply corrections to all observed data



# Atmosphere

- 'Dry' component:
   Worst O<sub>2</sub>, O<sub>3</sub>
- 'Wet' component:
   H<sub>2</sub>O vapour/clouds
  - Highly turbulent layer
    - Measure PWV = precipitable water vapour
- Atmospheric depth increases at lower elevation
  - Larger zenith distances z

Column density as function of altitude





- Isoplanatic patch > sky area above single mm antenna
- Antenna 2 3 separation < W, phase stable, noise rather correlated
- 1 2, 1 3 higher but less correlated phase fluctuations
- Baselines to 4 see higher but uncorrelated noise

#### Kolmogorov turbulence



Kolmogorov prediction (Coulman'90)  $\varphi_{rms} = \frac{K}{\lambda} B^{\alpha}$ where K~100 at ALMA for  $\lambda$  in mm and  $\alpha$  depends on the length of baseline B compared with W, the thickness of the turbulent layer

- Baseline 2-3 < *W* 
  - Phase noise  $\phi_{\rm rms}$  increases as  $B^{5/6}$
- Baselines 1-2, 1-3 > W but < OS:  $\phi_{\rm rms} \propto B^{1/3}$
- Baselines 4-\* in outer scale regime:  $\phi_{rms}$  levels off

# Water Vapour Radiometry

- PWV=precipitable water vapour
  - Changes refractive index, speed of radio propagation
  - Phase error
    - $\Phi_{\rm e} \propto (2\pi/\lambda) {\rm PWV} \ / \ T_{\rm atm}$
  - 1 mm PWV
    - ~ 0.7 mm extra path
    - ~ 0.0023 ns delay
  - Each ALMA 12-m has water vapour radiometer
  - Apply bulk correction during observations
  - Use ~1Hz WVR measurements to correct residual fluctuations on-line or during data processing
- PdBI measures PWV at 22 GHz, corrects amplitudes



Blue rectangles are the production WVR filters



PWV ~0.6, Band 9 raw 0.25 - 2.5 km baselines



12:18:20 Time

#### Measure total noise T<sub>sys</sub>





- Measure T<sub>sys</sub> using hot & cold loads
  - ALMA every few min, 128 channels per 2-GHz baseband
  - NOEMA per baseband
- Tracks atmospheric & instrumental fluctuations
  - Applied during data reduction



#### Timing and antenna position errors

Source

- Signals from off-centre source reach telescopes T1, T2 with a phase shift
  - Geometric delay  $\delta$  needs to be removed and travel-time from T1 and T2 to correlator equalised to preserve astrophysical phase
    - Delays added per-antenna before correlation
  - 1 ns delay error =  $1/10^9$

T2

 $- = \pi$  rad per GHz freq error

To correlator

- Causes of delay error include:
  - Timing error in correlator
  - Bad atmospheric model
  - Antenna position error

#### Antenna positions





DA41 position corrected during data processing DA50 ????

- Antenna positions measured after every move
  - May need updating
  - DA41 ~10 cm position error ~ 0.33 ns delay error
  - Also causes time-dependent phase error
    - Incorrect model for updating geometric delay

## Delay error (rare in cycle >0)

- Phase across 2 GHz undergoes 6 full turns on one antenna
  - Delay error  $6/2 \times 10^9 = 3$  ns



#### Timing and antenna position errors





#### DA41 position corrected





# Calibration sources: flux density

- Primary flux calibration uses planets, moons, asteroids
  - Models and ephemerides available
  - Mostly negligible polarization

0.05

- Still often have to select short baselines!
- Beware planet/moon atmspheric lines
  - If no Solar System object, use monitored QSO



#### Calibration sources have lines too



# Astrophysical calibration sources

- Calibrate amplitude and phase as a function of :
  - Frequency (delay and BandPass calibration)
    - Mostly instrumental errors, stable for hours
      - Solve per-channel for TDM, or average up to 20 GHz
    - Calibrator within 45°
  - Time (phase (& amplitude) referencing applied to target)
    - NB in practice use time & bandpass calibration for all calibrators, sometimes refine interatively
- Phase error  $\phi$  radians gives fractional amp error  ${\thicksim}\phi$
- Need signal/noise  $S_{\text{calsource}}/\sigma_{\text{ant}} > 15$ 
  - ~1/15 radians ~ 7% amp decorrelation ~ 4° phase error
    - Average spw or transfer from wide to narrow if necessary

# Astrophysical calibration sources

• Most errors are antenna based

$$\sigma_{ant}(\delta t, \delta v) \approx \sigma_{array} \sqrt{\frac{N(N-1)/2}{N-3}}$$

- +  $\sigma_{\text{array}}$  is noise in all-baseline data per interval
- e.g. 36 antennas,  $\sigma_{\text{ant}} \sim$  4.5 x map noise per interval/pol.
- Band 7: rms per TDM spw 1.875 GHz ~0.5 mJy/minute
  - x  $\sqrt{2}$  per polarization ~0.7 mJy
  - x 15 (SNR) x 4.5 (per antenna) ~ 50 mJy
     In practice aim for ~ 100 mJy at B7, if possible
- More compact configs/low  $\nu :$  within 10-15 deg
  - Cycle time 5-10 min
- Extended configs/high v: within ~4 deg
  - Cycle time 1-few min

# Phase referencing

- Observe phase-ref source close to target
  - Point-like or with a good model
  - Close enough to see same atmosphere
    - ~2-15 degrees (isoplanatic patch)
  - Bright enough to get good SNR much quicker than atmospheric timescale  $\boldsymbol{\tau}$ 
    - $\tau$  10 min/30 s short/long *B* & low/high v
  - Nod on suitable timescale e.g. 5:0.5 min
    - Derive time-dependent corrections to make phase-ref data match model
    - Apply same corrections to target
  - Correct amplitudes similarly
- Self-calibration works on similar principle



#### Phase referencing



Target 3C277.1

Phase-ref 1300+580 ~3° from target Unresolved point

#### Source structure in uv plane



#### Extended source: more flux on short baselines



Visibility amplitudes

#### Baseline length in wavelengths (*uv* distance)

# Point source: same flux density on all baselines (within errors)

UVDist L

1e+06

1.5e+06

2.5e+06

3e+06

3.5e+06

#### Phase-reference visibilities





#### Corrections for point model



#### Corrected phase-ref visibilities



#### Time-dependent (self-)calibration

- ALMA data are calibrated by observatory staff
  - Pipeline or scripts
  - Instrumental corrections are usually standard
    - You can refine them but not usually necessary
  - Sample images provided
- If target bright, you can self-calibrate
  - rule of thumb: worth trying if S/N > 20 in few min
  - Starting model is the image made by applying phase-ref (and previous) corrections
- You may also want to make images differently

# Effects on imaging

-43°55'00"

10<sup>h</sup>27<sup>m</sup>56<sup>e</sup>

12000 Declinatia

to<sup>h</sup>27<sup>m</sup>54<sup>s</sup> 52<sup>s</sup> 51<sup>s</sup> 50<sup>s</sup> 32000 Right Ascension

No astrophysical calibration: no source seen

Bandpass, phase-ref etc. solutions applied: S/N 20 Anti-symmetric artefacts dominate

- residual phase errors

Phase-only self-cal: S/N 40 Residual, symmetric amplitude errors



Amp & phase self-cal: S/N 50



#### Calibration: Measurement Equation basis for calibration in CASA

 $\underline{V}_{ij} = \mathbf{M}_{ij} \mathbf{B}_{ij} \mathbf{G}_{ij} \mathbf{D}_{ij} \mathbf{F}_{ij} \mathbf{$ 

#### Vectors

 $\underline{V}$  isibility = f(u,v) $\underline{I}$  mage to be calculated

Additive baseline error

#### Scalars

 $\frac{S}{P}$  (mapping  $\underline{I}$  to observer pol.)

*l,m* image plane coords *u,v* Fourier plane coords *i,j* telescope pair

#### **Jones Matrices**

Multiplicative baseline error

Bandpass response

Generalised electronic gain

Dterm (pol. leakage)

- E (antenna voltage pattern)
- Parallactic angle
- Tropospheric effects
- Faraday rotation

# Polarization

Most (sub-)mm Rx use linear polarization X and Y combined in correlator



ALMA Band 3 ( $\lambda$ ~3mm) uses Ortho Mode Transducer

> Band 9 ( $\lambda$ ~0.45mm) uses wire grid

XX, YY for total intensity

XX, YY, XY, YX used to provide Stokes IQUV





Diagram thanks to Wikipaedia

# ALMA polarization calibration

- Calibration source has unknown linear polarization
  - Gain solutions decomposed into X and Y per antenna

$$g_x' = g_x(1+Q/I)^{0.5}$$
  $g_y' = g_y(1-Q/I)^{0.5}$ 

- Leakage between X and Y feeds ('D'-terms)
- Feeds rotate on sky as alt-az antenna tracks source
  - Parallactic angle rotates
  - $Q_{obs}$  time-dependent
- Known feed orientation
  - Directly correct pol. angle
  - Refine after leakage correction
- 3+ scans, >3hr HA coverage
  - Solve for leakage and source polarization



# Cleaning

• Fourier transform the visibilities and the uv tracks



# Cleaning

- Fourier transform the visibilities and the uv tracks
- Set a mask to include obvious emission

NGC 3256 dirty map

51<sup>\$</sup>

ടറ<sup>ട</sup>

36'

48"

00''

12"

74"

36"

48"

00"

54<sup>5</sup>



- CLEAN algorithm identifies brightest pixels
- Store e.g. 10% of each peak as Clean Component

# Cleaning

 Iteratively subtract scaled dirty beam at positions of bright pixels



#### **CLEANed** image

• Improved signal-to-noise in final image



Residual is just noise
 Note different flux scale



## CLEANed image

• Note improved signal-to-noise in image



 Final image is combination of residual and Clean Components convolved with restoring beam



# Weighting

- Natural weighting: Grid and FT samples
  - Weight zero in grid cells with no samples
- Uniform weighting:
  - All cells same weight
    - Makes resolution finer, but increases noise
- Robust [-2, 2]: intermediate weightings
- Taper: reduce weight of longest baselines
  - Increase sensitivity to large-scale emission
    - Increases noise if effective collecting area reduced





#### Cleaning and the uv plane

#### NGC 3256 model coloured by baseline over black data



- *uv* model is Fourier transform of Clean Components
  - 'Major Cycles' subtract model from uv data and remake dirty residual image
- Compare model with data to assess quality
  - Use for further rounds of self-calibration

#### ALMA data processing





- Each SB self-contained
  - Pointing, flux scale, BP, (pol.) cal scans
  - ~20-40 mins target(s)/phase ref
  - Execute (EB) until sensitivity is reached
    - same spectral & array configuration
- Initial data processing (pipeline or staff script)
  - Convert ASDM to Measurement Set
  - Calibrate
  - Combine EBs for target imaging

#### What ALMA data do you get?



Images/cubes for principal science target channels



Data processing / pipeline scripts so you could tweak imaging/ self-cal using:

Science products & scripts might be all you need or you could use CASA scripts to regenerate edited, calibrated uv data

ASDM (one per EB)

**Flag tables** 

**Calibration tables** 

- Self-calibrate if bright enough (S/N  $> \sim 20$  per scan)
- **Re-image** 
  - Weighting for higher spatial sensitivity/lower resolution or v.v.
  - Change spectral resolution, make spectral index image etc. etc.....

# ALMA observing terminology

- Scheduling Block: self-contained observation series
  - Short pointing, flux scale, bandpass, (pol.) cal scans
  - ~20-40 mins alternating between target(s)/ph ref(s)
    - Including multiple mosaic pointings
  - Execute (EB) until desired sensitivity is reached
    - All in same spectral and array configuration
    - Each EB produces one ASDM (ALMA Science Data Model) with binary data and lots of metadata
- Initial data processing (calibration, editing) per-EB
  - Convert ASDM to Measurement Set
- Combine EBs for final target imaging
- May also combine different SBs e.g. ACA+main

#### Individual scheduling block



# CASA

- Common Astronomy Software Applications
   All ALMA, VLA and similar data reduction tasks
- Python wrapper to aips++ library
  - Access to toolkit, extensions for ALMA etc. fixes
- Uses Measurement Set format
  - Can be converted to/from FITS (some caveats)
    - Export images to use specialised analysis tools
- http://casa.nrao.edu/
  - Cookbook
  - CASA Guides
- https://casaguides.nrao.edu/index.php?title=Analysis\_Utilities



# (sub-)mm Interferometry Summary

- Atmospheric refraction/absorption dominates quality
  - Cold dry sites OK  $\leq$ 370 GHz, exceptional sites  $\leq$ 1 THz
  - Troposphere affects phase & amp on  $\geq$ 1s timescales
    - Instrumental calibration (WVR,  $T_{sys}$ ) etc.
  - ALMA / NOEMA  $\mu Jy$  sensitivity, 10 / 100 mas resolution
    - Good sites, many or large antennas
  - Sub-mJy sensitivity at sub-arcsec resolution
    - Extended sources need multiple arrays/SD fill in
    - Large fields need mosaicing
- Normally observe two separate sidebands
   May have 'mirror' or noise
- Plan night-time, dry season observing at  $\lambda$  sub-mm
- ALMA delivers calibrated data, sample images
  - May want to self-calibrate, reimage changing resolution