Spectral line data analysis

Mark Lacy, NRAO (NAASC)

Fourteenth Synthesis Imaging Workshop

2014 May 13– 20

#SIW2014

NRAO

Associated Universities, Inc.

New Mexico Tech
SCIENCE · ENGINEERING · RESEARCH · UNIVERSITY

The University of New Mexico

IGPPS
Los Alamos

Outline

- Why spectroscopy?
- What happens in the correlator
- Data volumes
- Special calibration considerations
- Doppler shifts, and how to correct
- 3D imaging
- Visualization and analysis of data cubes
- Those units…

Technical details in orange boxes

Why spectroscopy?

- Molecular and atomic transitions inform on the physical conditions in, and the chemistry of, the emitting/absorbing source.
  - Density – different transitions have different critical densities, also specific tracers of dense gas (optically thinner than CO) like HCN, CS, Ammonia
  - Temperature – look for different energy transitions of the same species (e.g. Ammonia), also water vapor/ice transition
  - Shocks (SiO, methanol)
- Even continuum observations are usually taken in spectroscopic correlator modes to reduce bandwidth smearing.

Example 1: detection of complex molecules in protostars

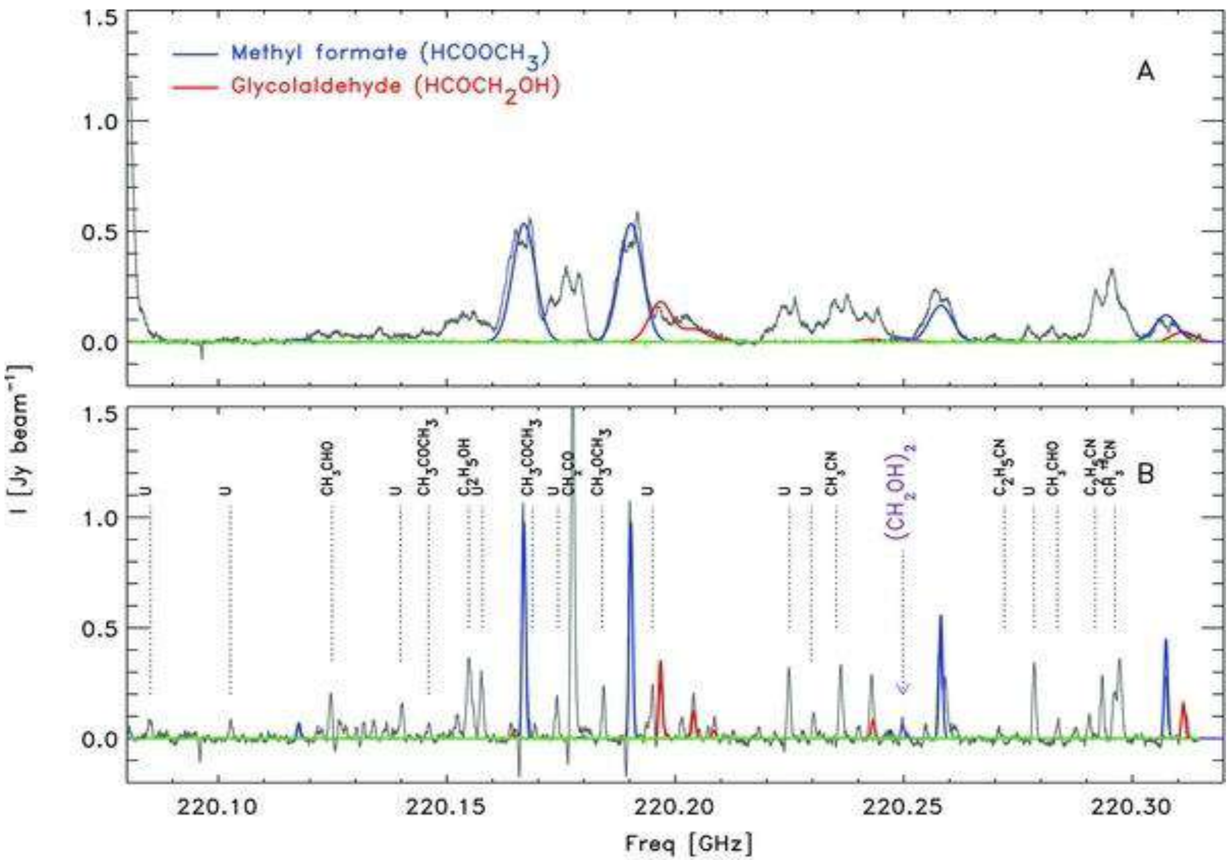
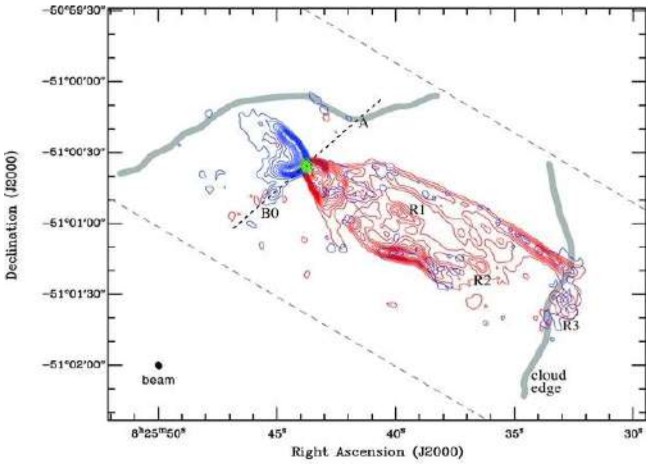
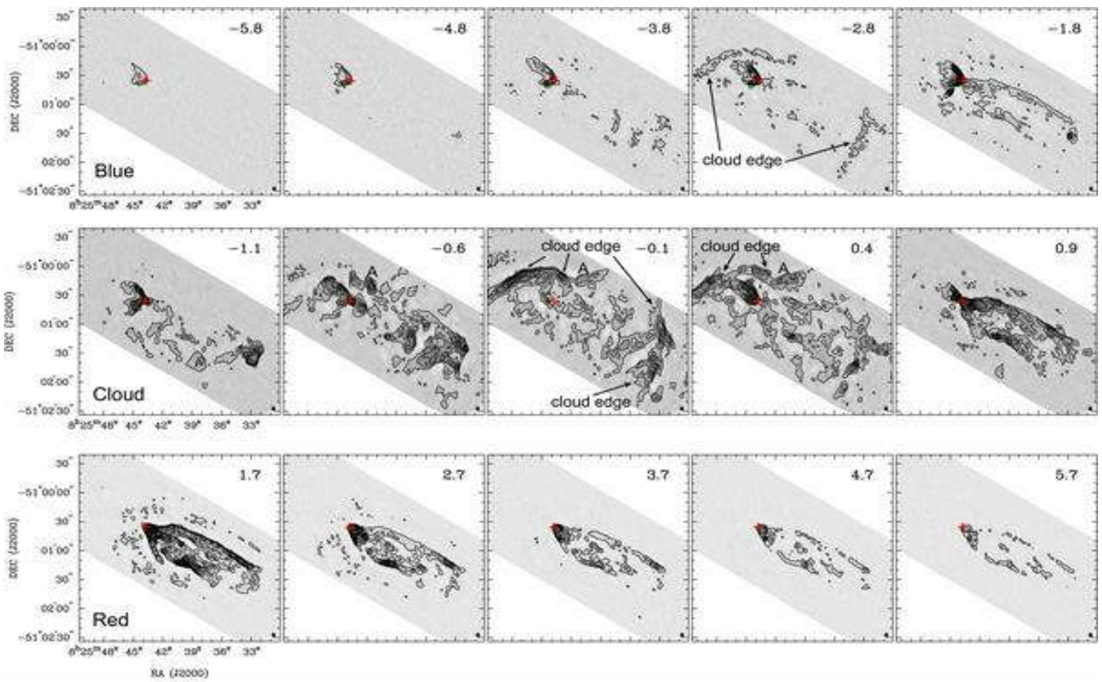


Figure 1 from Detection of the Simplest Sugar, Glycolaldehyde, in a Solar-type Protostar with ALMA
Jes K. Jørgensen et al. 2012 ApJ 757 L4 doi:10.1088/2041-8205/757/1/L4

Example 2: dynamical information

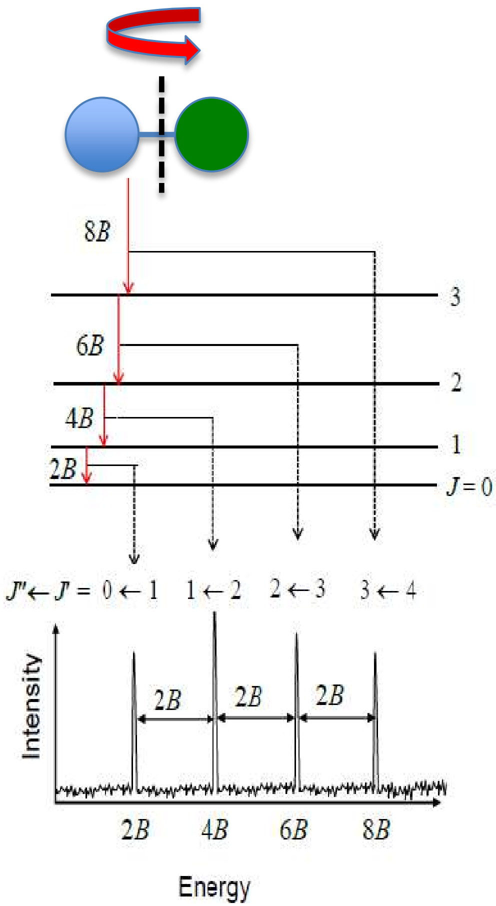




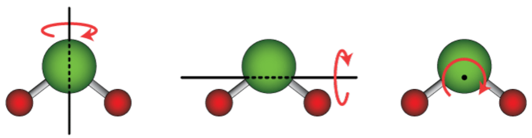
Figures 1&3 from ALMA Observations of the HH 46/47 Molecular Outflow
Héctor G. Arce et al. 2013 ApJ 774 39 doi:10.1088/0004-637X/774/1/39 (CO 1-0 line)

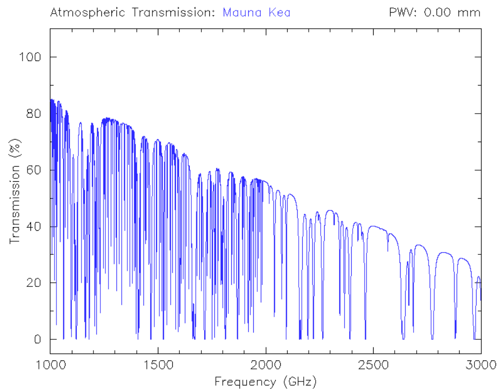
Molecular spectroscopy

- The richness of a given molecule's spectrum depends on its shape.
- Molecules can have one, two or three rotation axes with dipoles.
- CO just has a single axis
  - ➢ Evenly-spaced ladder of transitions
- Most molecules have more than one axis, and much more complex spectra.



Simple rigid rotor (CO)





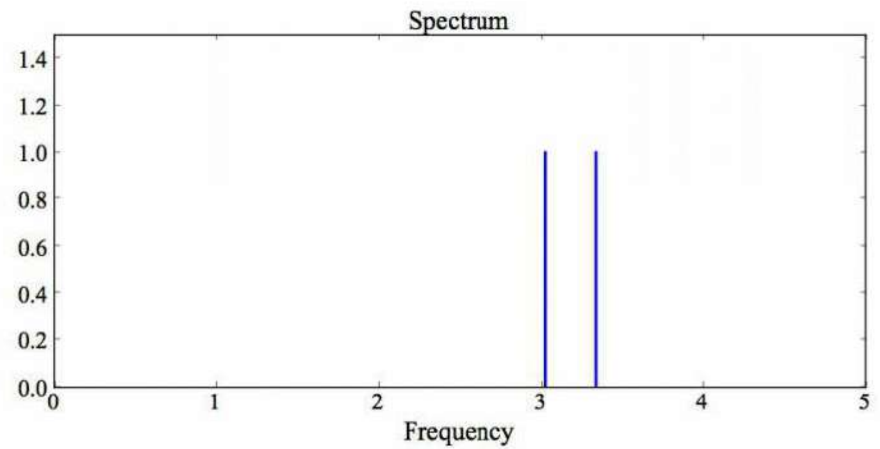
Water (in absorption) – asymmetric top

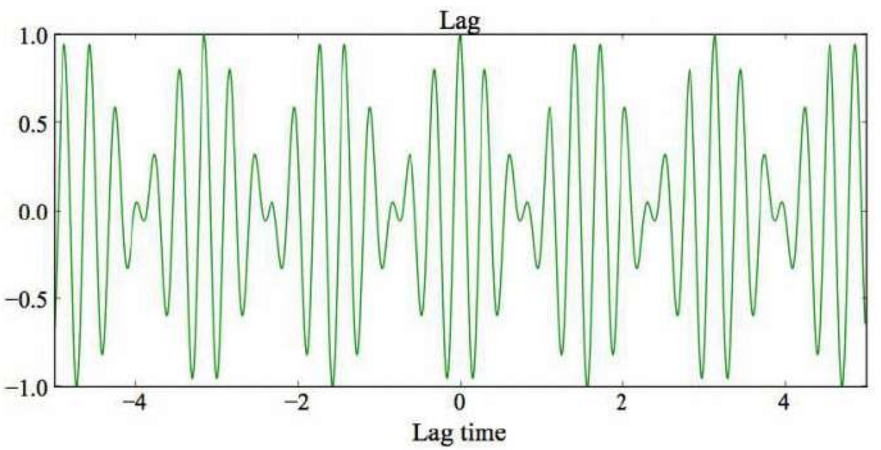
What happens in the (XF) correlator

- XF correlators (ALMA, VLA) do X-correlation (actually a convolution over a lag window) first, then Fourier transform.
- Instead of inefficiently scanning through frequencies, uses fact that time and frequency are Fourier pairs.
- Fourier transform the signal as a function of convolution lag to get the spectrum.
- Resolution set by longest lag.
- Bandwidth is set by minimum time interval.

- Use a lag interval, Δt
- Total bandwidth = $1/(2\Delta t)$
- For N spectral channels, have to measure 2N lags from $-N\Delta t$ to $+(N-1)\Delta t$
- Spectral resolution $1/(2N\Delta t)$ (Nyquist sampling)
- Can adjust Δt and N to suit science

Example of two narrow lines

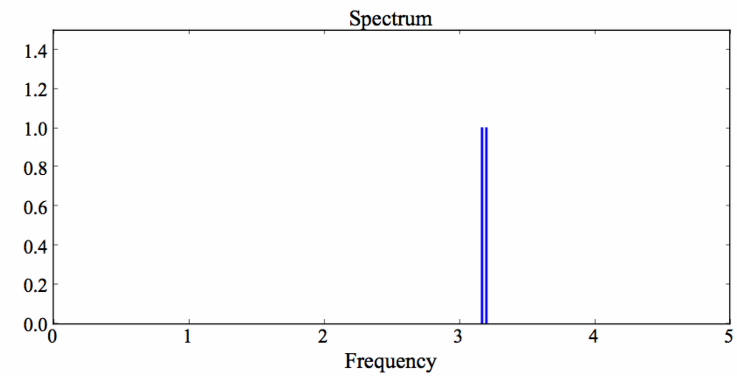


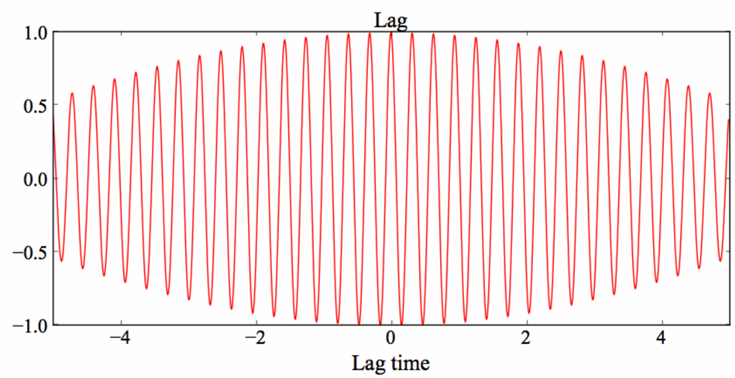


- FT of $\delta(\omega_1)+\delta(\omega_2)$ is proportional to $\cos([\omega_1+\omega_2]/2)*\cos([\omega_1-\omega_2]/2)$
- i.e. a beat pattern with a beat period determined by the difference in the frequencies of the two lines.

Just resolved when one full period fits in the lag window

Much closer and it's hard to tell, e.g. two lines from a single broad line (FT of a Gaussian is a Gaussian)



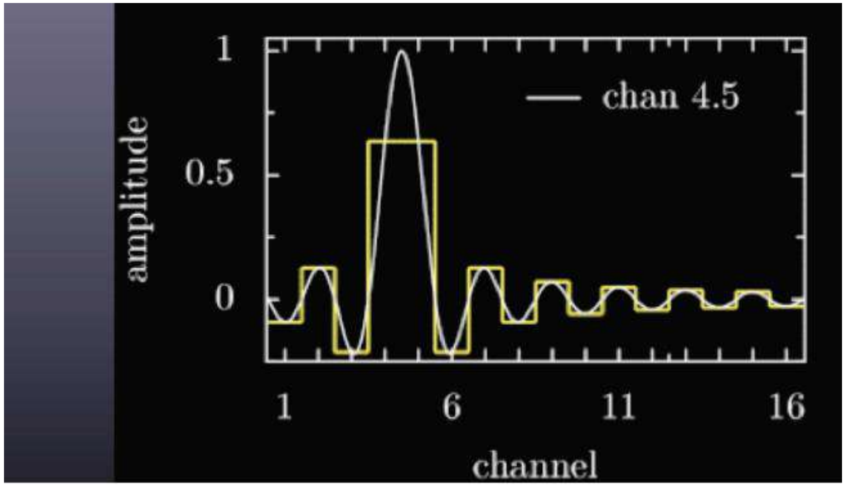


NRAO

New Mexico CONSORTIUM

Gibbs ringing and spectral smoothing

- Can't measure all the lags (correlator runs out of memory), so cutoff after 2N lags
- Fourier transform of a top-hat lag window is a sinc function – 22% sidelobes – "Gibbs phenomenon" (cf. diffraction rings in optical telescopes).
- Usual solution is to taper off the lag signal at long delay times with a smoothing function.
- Reduces spectral resolution, but fixes sidelobes.

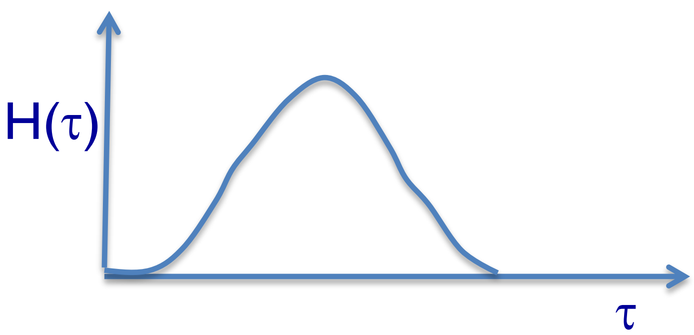


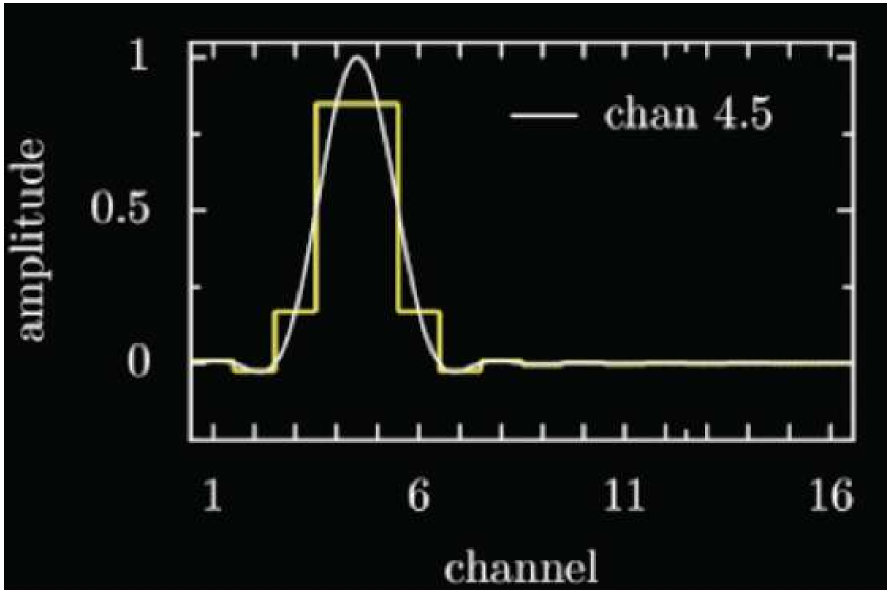
Hanning smoothing

Hanning smoothing often applied to the lag spectrum:

$H(\tau)=0.5(1+\cos(\pi\tau/[N\Delta t]))$

Results in lower spectral resolution (2 channel).



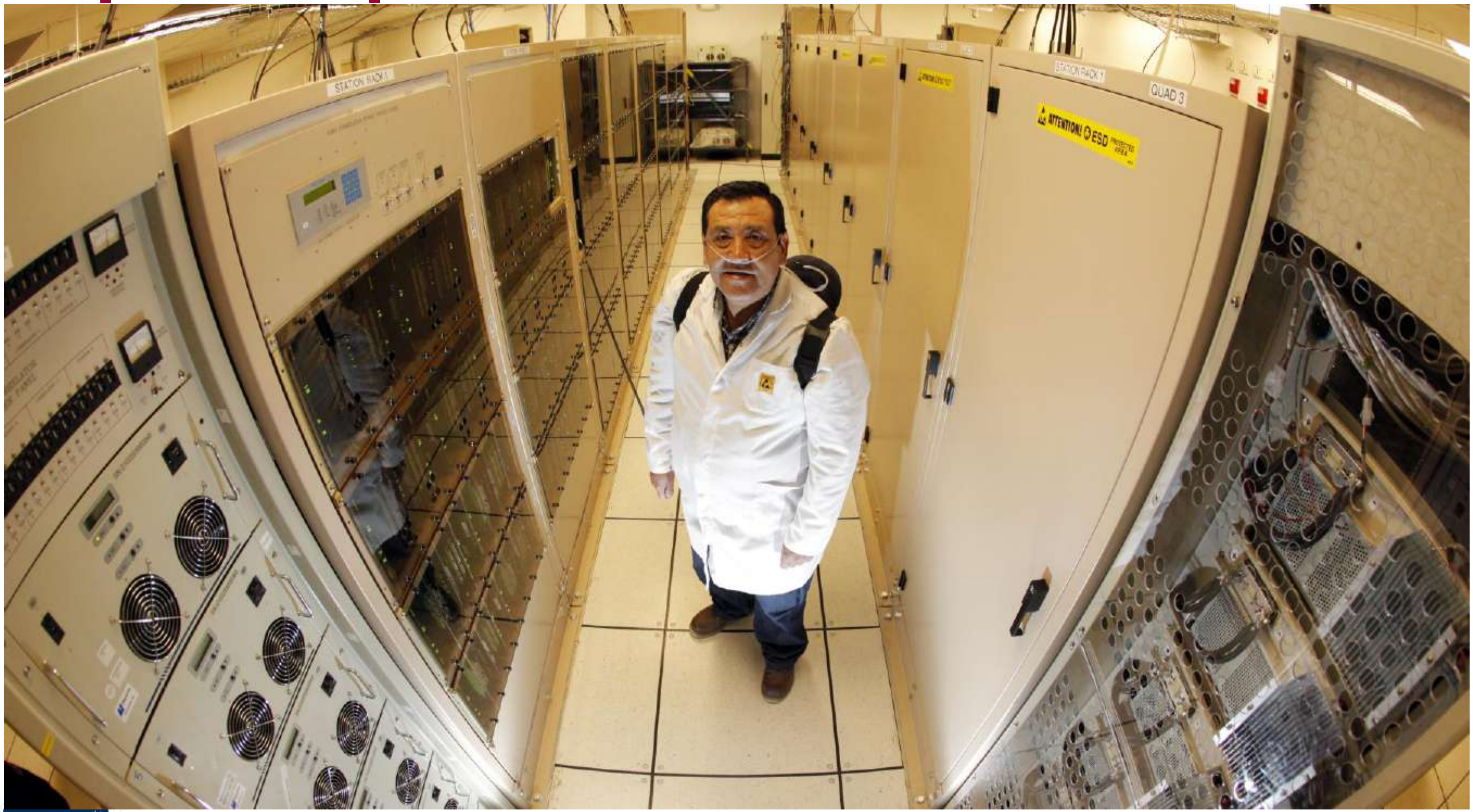


Spectroscopy with ALMA

- Usually use the "Frequency Division Mode" (FDM)
- Spectral coverage 84-163GHz (Bands 3 & 4); 211-373GHz (Bands 6 & 7); 385-500GHz (B8) and 602-720GHz (B9)
- 4x2GHz basebands
- Up to 32 Spectral Windows (spws) per baseband, each spw can be 60MHz-1875MHz wide.
- Highest resolution is 15kHz (Hanning smoothed, single polarization).

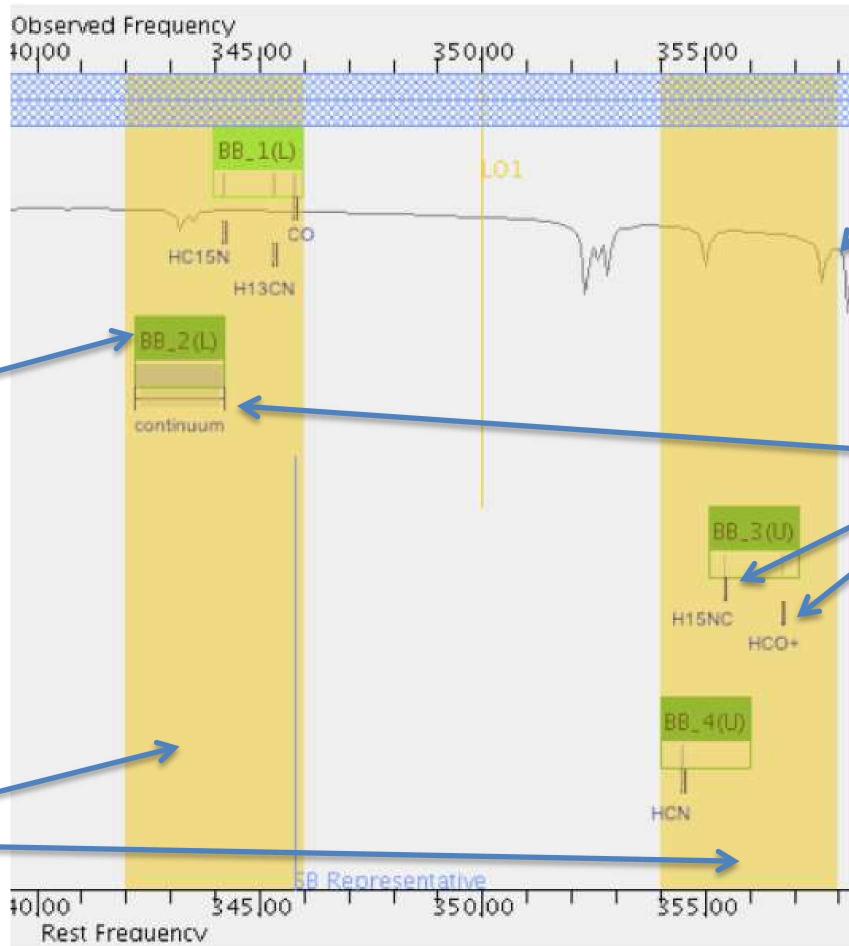
- Within a baseband, the total number of channels is fixed at 7680/N_{pol}, and are distributed amongst the assigned spws. (N_{pol}=2 usually)
- At the time of writing, all spws in a baseband must have the same resolution.
- For low resolution spectroscopy, can smooth the FDM settings or use the Time Division Mode (TDM).
- Note that the spectra come pre-divided by the autocorrelation – edges of the bandpass are "corrected" for dropoff but very noisy.
- See the ALMA Technical Handbook (available from www.almascience.org) for details.

The ALMA correlator – world's highest supercomputer



Example ALMA spectroscopic setup

ALMA OT Spectral setup display



Atmospheric transparency

Basebands shown in green

Individual spws

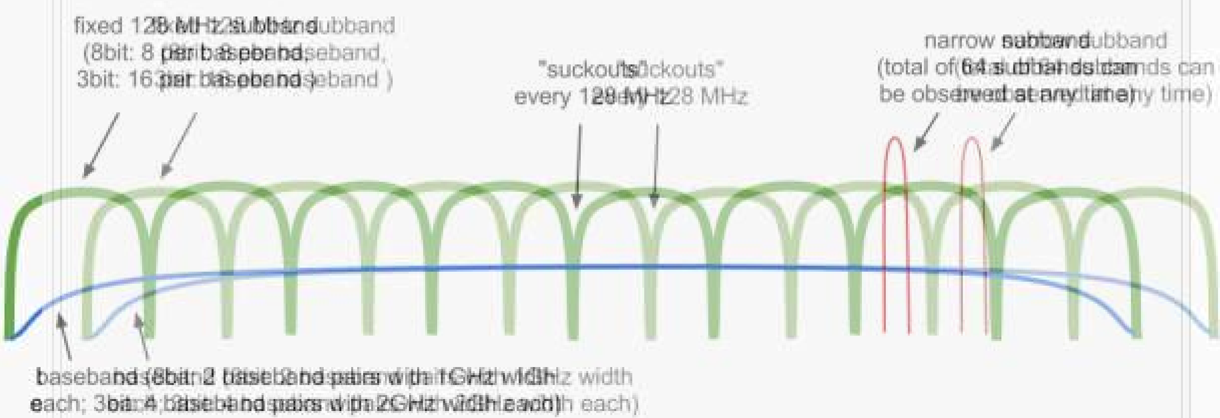
IF ranges shown in yellow

Spectroscopy with the VLA

- Also up to 8GHz of total bandwidth.
- Up to 64 sub-bands within each baseband, independently configurable wrt number of channels, polarizations and bandwidths.
- Note that the edges of each sub-band have low responses – avoid putting your lines in there (interleave basebands if you need continuous coverage).

- The VLA samplers can operate in 8-bit or 3-bit mode. 3-bit is about 15% noisier and requires more setup time, but allows up to 8GHz of bandwidth in 4x2GHz basebands compared to only 2GHz in 2x1GHz basebands in 8-bit mode.
- Number of channels per subband limited to $16384/N_{pol}$

Schematic VLA spectral setup



WIDAR correlator baseband with subbands

Data rates

- Modern correlators are capable of producing more data than the storage and processing facilities can deal with.
- Data rates are a consideration for both VLA and ALMA.
  - Match spectral resolution to science.
  - Time average if possible (but beware of time smearing in wide fields).

- For the VLA, data rate, R:

$$R = 160\mathrm{GB/hr}\frac{N_{chan}N_{pol}}{16384}\frac{N_{ant}(N_{ant}-1)}{26\times 27}\frac{1\mathrm{s}}{\Delta t}$$

- ~32 bits/visibility
- Proportional to ~N_{ant}^2
- ALMA in full resolution mode is similar (but more antennas; calculation is done for you in the ALMA Observing Tool).

Doppler shifts

- The earth is not a stationary observing platform. Doppler shifts arise from:
  - Earth rotation
  - Earth moving around the sun
  - Sun moving around the galactic centre
  - Galaxy moving in local group
  - Local group moving wrt Cosmic Microwave Background.
- LSRK most commonly used for non-solar system observations

- A Doppler correction needs to be applied at the time of observation, ("Doppler setting") and sometimes afterwards in software (tasks cvel/mstransform in CASA, or via the outframe parameter in clean).
- Doppler corrections are dependent on sky position, so always specify a phasecenter/field.
- Beware! two different ways to specify velocities:
  - $v_{radio}=c(\nu_{rest}-\nu_{obs})/\nu_{rest}$
  - $v_{optical}=c(\nu_{rest}-\nu_{obs})/\nu_{obs}$
  - $v_{optical}=cz$ corresponds to redshift.
  - v_{radio} often used though as Δv linear with $\Delta\nu_{obs}$

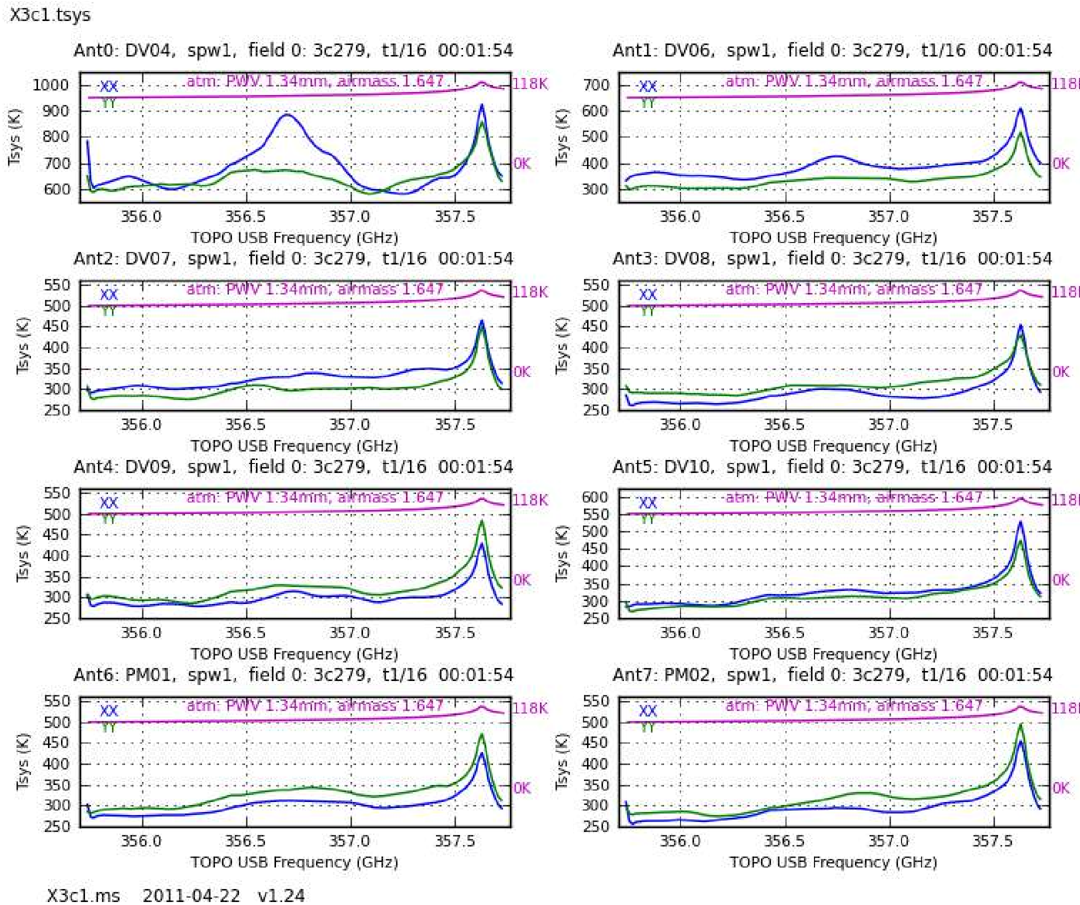
Velocity frames

Rest Frame Name	Rest Frame	Corrects for	Max. Amplitude (km/s)
Topocentric*	Telescope	Nothing	0
Geocentric	Earth Center	Earth Rotation	0.5
Earth-Moon Barycentric	Earth+Moon center of mass	Motion around earth-moon center of mass	0.013
Heliocentric	Center of sun	Earth's orbital motion	30
Barycentric	Earth+sun center of mass	Earth+sun center of mass	0.012
Local Standard of Rest [Kinematic] (LSRK)	Center of mass of local stars	Solar motion relative to nearby stars	20
Galactocentric	Center of Milky Way	Milky Way rotation	230
Local Group Barycentric	Local Group center of mass	Motion of Milky Way within Local Group	100
Virgocentric	Center of the local Virgo Supercluster	Local Group motion	300
Cosmic Microwave Background	CMB	Everything else	600

*Beware! Topocentric is the ALMA default

Calibration considerations - system temperature calibration

- When working above about 80GHz, a system temperature (T_{sys}) calibration is carried out using observations of the sky (and sky+object) and internal calibration loads.
- T_{sys} is usually dominated by the atmosphere, and is a function of frequency as atmospheric lines are common.
- For the VLA (<50GHz), an opacity correction is usually adequate.

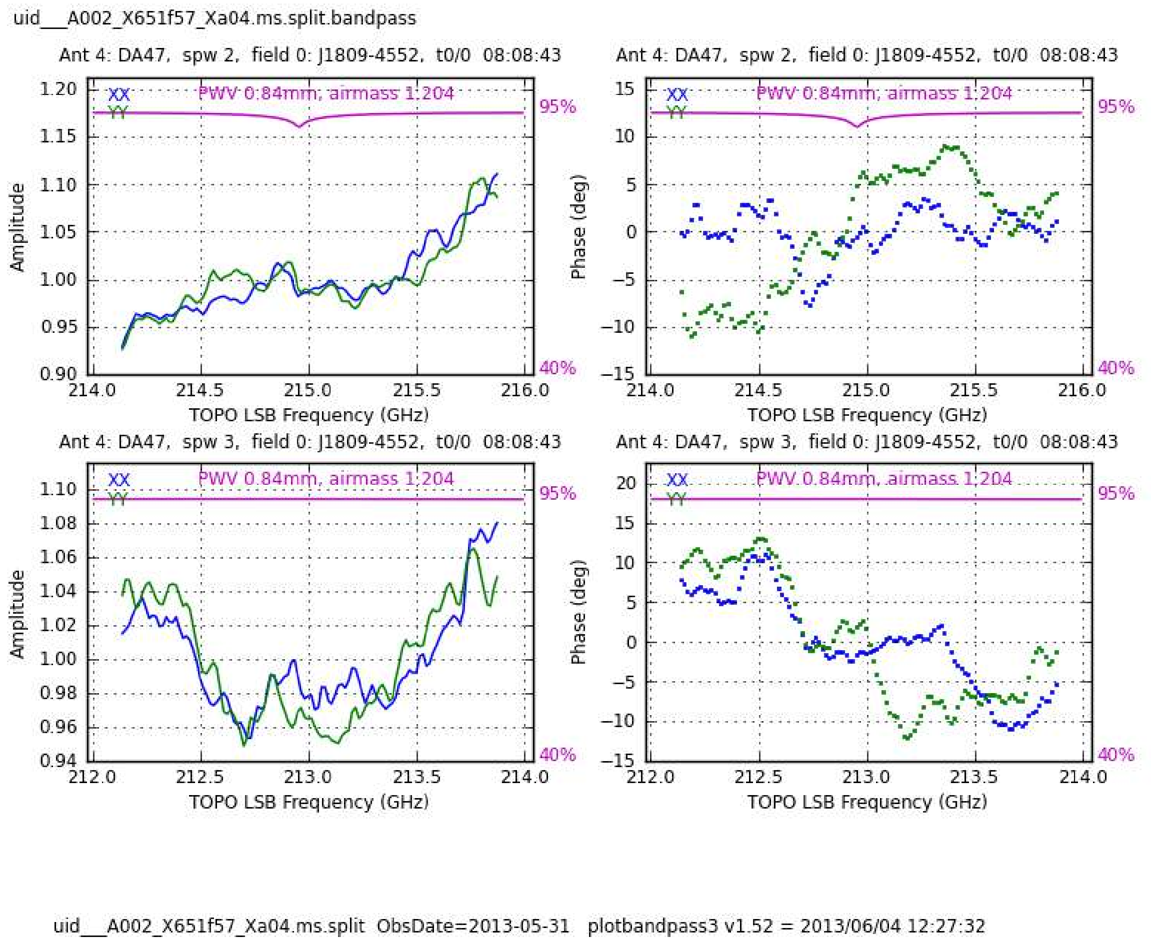


Calibration - bandpass calibration

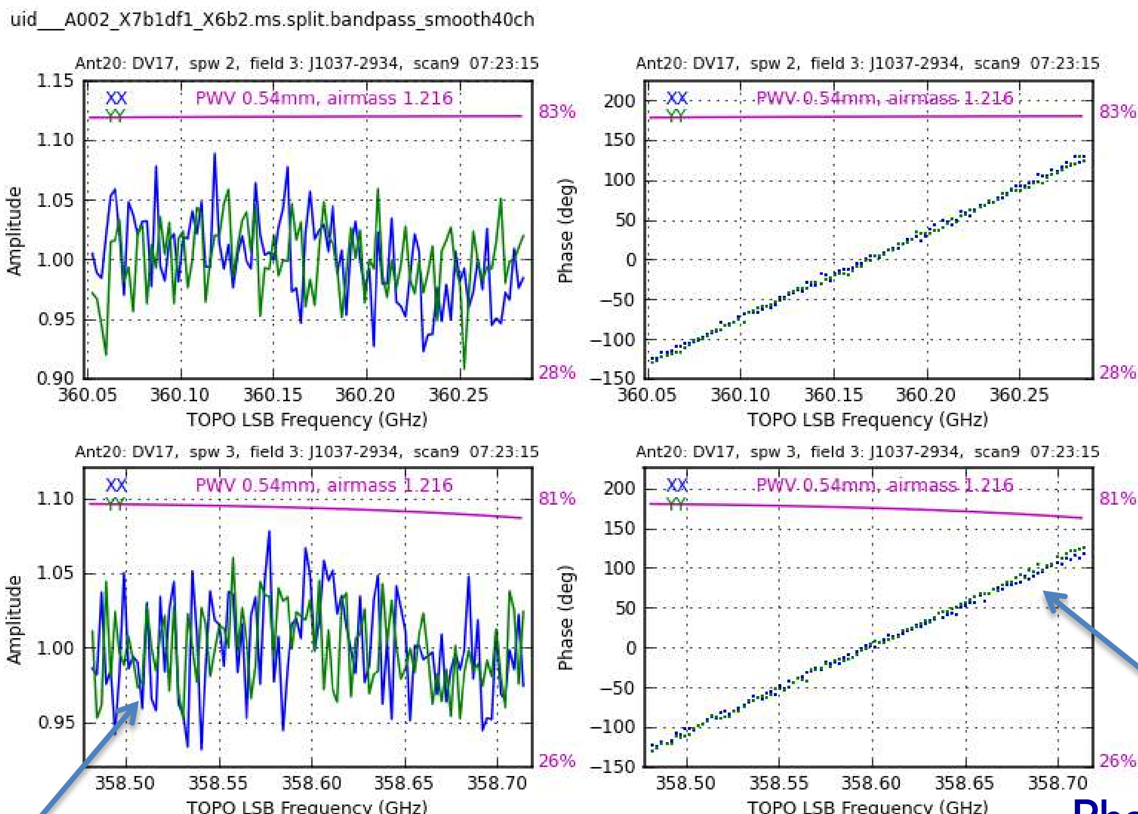
- Response of system is not uniform as a function of frequency across the band. Affected by:
  - Atmosphere (should be taken out by Tsys, but…)
  - Front-end system
  - Antenna position errors (delays)
- Observe a bright calibrator source that has no spectral features to correct this.
- Bandpass calibration varies slowly with time, so only one bandpass calibrator observation is typically needed per tuning.
- Solve on a per-antenna basis

- Correct both amplitude and phase – phase errors can mimic position changes as a function of frequency: $\theta/\theta_B \sim \Delta\phi/360$deg
- For good spectral dynamic range, need to have S:N on the bandpass calibrator >> on the source.
- Time on bandpass cal, $t_{BP}>9(S_T/S_B)^2 t_T$, where t_T is the time on target and S_T and S_B are the target and bandpass fluxes.
- Even more time if looking for faint lines on strong continuum

Example bandpass calibration from ALMA



Bad Bandpass!



uid__A002_X7b1df1_X6b2.ms.split ObsDate=2014-02-23 plotbandpass3 v1.128 = 2014/04/11 19:12:41

What's wrong with this? (At least a couple of things…)

Poor S:N

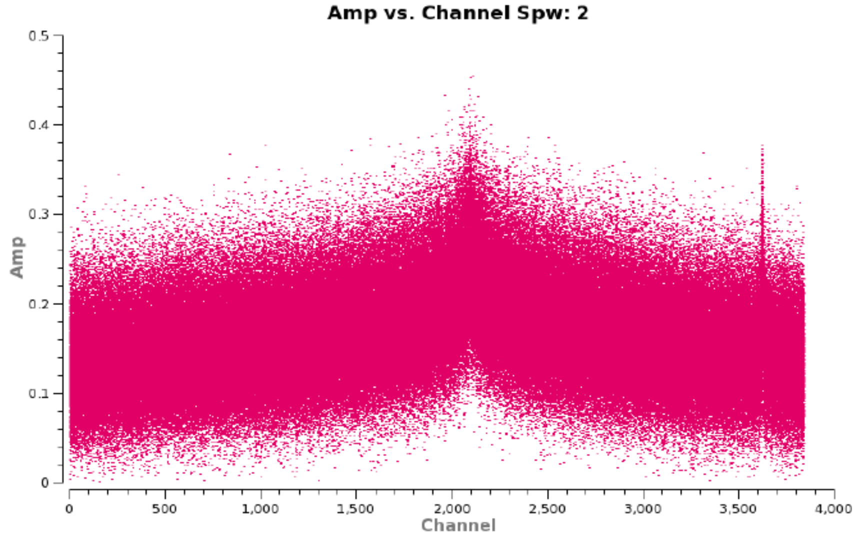
Phase error – most likely delay error due to bad antenna position

NRAO

Flux calibration

- If using a solar system object with an atmosphere for calibration (Jovian or Saturnian moons, for example), be aware that these objects often have absorption lines.
- Check your calibrator spectrum carefully, particularly if your object is not redshifted.
- You should exclude affected channels (or even basebands if the absorption is broad).

- To apply a flux calibration from one spectral window to another in CASA use the refspwmap parameter in fluxscale.
- For example, if spw 2 (of 4) has a poor calibration due to a strong absorption line, setting refspwmap=[0,1,3,3] will use the calibration of spw3 on spw2 (in the 2nd slot).



Imaging: making your image cube

- Imaging process is the same as for a continuum map, but you are making a cube (axes RA, Dec, frequency [or velocity]).
- Same considerations for beam weighting etc apply to each plane.
- Beam may vary significantly along the frequency axis (if so, CASA will make a beam per plane).
- Line and continuum probably have different spatial structures – bear in mind while cleaning.

- When making the map, you may wish to bin the channels to speed things up if you don't need the full velocity resolution.
- Don't forget to set the output velocity frame(!).
- If the source is bright enough to self-calibrate, pick whichever is the higher signal-to-noise of line or continuum (usually continuum), and self-calibrate it. Then apply the calibration to both line and continuum data.

Making a continuum image

- The primary product of spectroscopic data is an image cube.
- Continuum images can also be made, and indeed can be the primary goal.
- Multifrequency synthesis (MFS) effectively combines maps made in each channel. Gives better signal-to-noise and uv-coverage than a single channel map, and reduces bandwidth smearing in a wide-band, wide field continuum map

- MFS can also be used to obtain spectral indices for sources in the continuum image.
- Beware that the primary beam can vary from one side of the band to the other if the fractional bandwidth is large.
- Special case: Faraday Synthesis – instead of producing an image cube with the 3rd dimension velocity, it is possible to produce a cube of polarized intensity whose plane spacings scale with Faraday depth (λ^2) – see polarization lecture.

Smoothing

- Smoothing (spatial or spectral) may be applied at the mapping stage, or afterwards, in the image plane.
- Spatial smoothing during imaging can be done by changing the weight function (natural weighting gives the lowest noise), or applying a uvtaper
- For spatial smoothing in the spatial dimension in the image plane use imsmooth.

- The imaging process for an interferometer is effectively a spatial filter.
- Smoothing the data may help with showing up large-scale structure, but will not be able to recover structure larger than the largest angular scale sampled during the observations.
- The effect of smoothing is to remove data in the uv-plane. Thus the true noise of an image can rise if heavily smoothed.
- Spatial smoothing is useful though if you want to match observations made with different beam sizes.

Spectral smoothing

- Hanning smoothing is applied online to ALMA data, and can be applied to VLA data (and is by default in the pipeline).
- Further spectral smoothing in CASA currently best done by setting the "width" parameter in clean, but could also be done in the image plane.

- Raw spectral data has an effective resolution of 1.2 channels
- Hanning-smoothed data has an effective resolution of 2 pixels.
- Channel-to-channel correlations mean that smoothing needs to be accounted for when calculating an RMS over multiple channels.
- E.g. for Hanning smoothed data, if the measured noise per channel is σ, then the noise in N channels is $\sigma/\sqrt{(N/2)}$.

Continuum subtraction

- Analysis of lines is easier if underlying continuum is subtracted.
- Can be done before imaging (uvcontsub in CASA), or afterwards (imcontsub).
- Generally uv-plane continuum subtraction preferred in challenging cases, but either valid.

- In both image and uv-plane fitting, pick line-free regions to define the continuum.
- The fit order should be as low as necessary to remove the continuum (usually 0^{th} or 1^{st} order).
- Check your results – should not be either residual emission or a "hole" at the position of the continuum source in line-free channels.

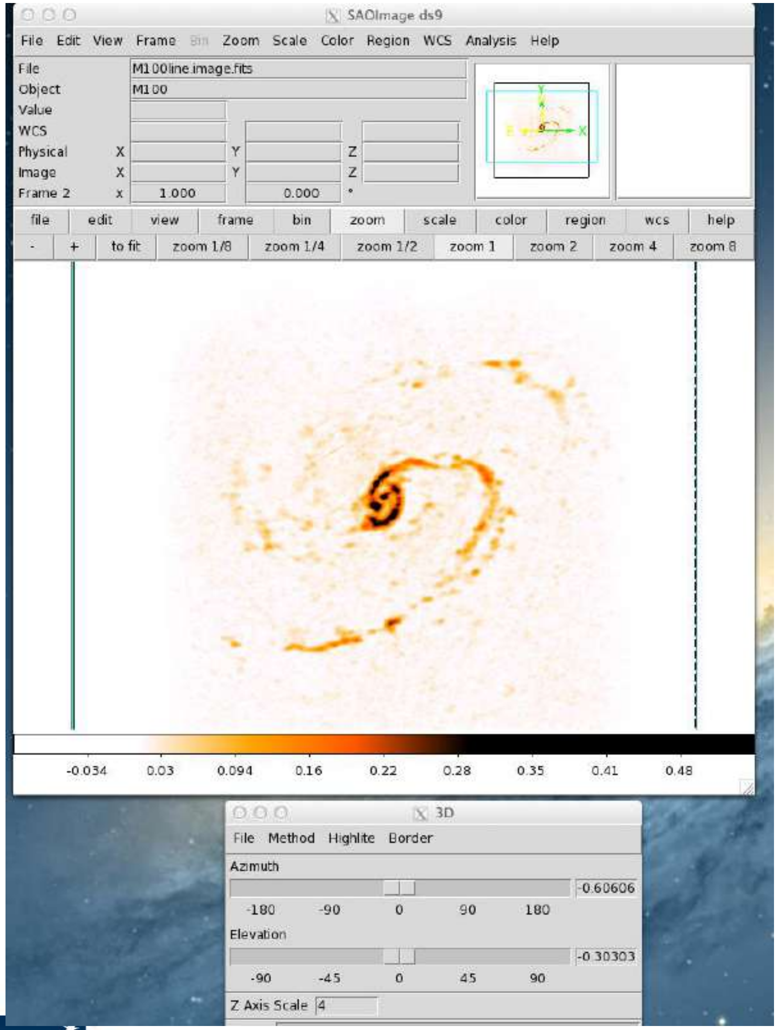
3D (volume-rendered) visualization

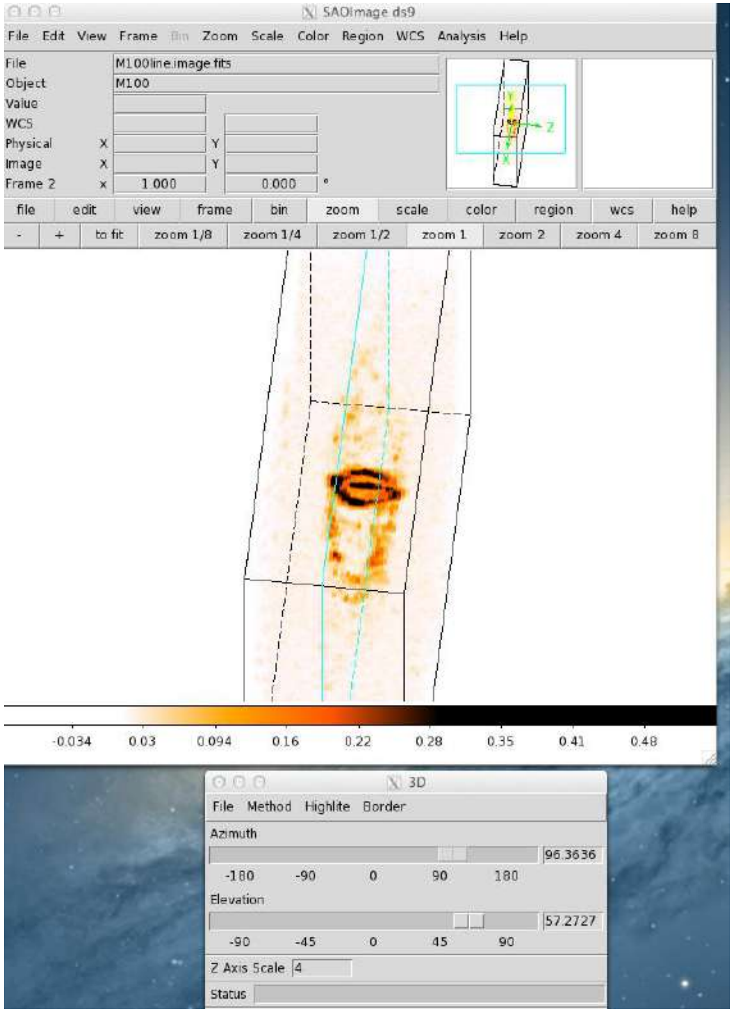
- Some software allows 3D visualization of image cubes (for example, SAOimage ds9).
- Can be useful for inspection of cube, and to find features that conventional moment analyses would miss.
- Can be hard to interpret though – velocity/frequency axis is not a spatial axis.

3D display software

- SAOimage ds9 is available from http://ds9.si.edu
- GAIA is available from star-www.dur.ac.uk/~pdraper/gaia/gaia.html (long-term support unclear)
- Karma kvis is available (though no longer updated) from http://www.atnf.csiro.au/computing/software/karma/
- Other, not observational astronomy specific 3D rendering packages are also available (ParaView, VisIt, yt….). Drawback is lack of understanding of astronomical coordinate systems.

SAOimage ds9 renderings of M100





Moment maps

- A popular way of reducing a 3D line to 2D
- Calculate the "Moment 0" image (line map), "Moment 1" image (velocity map) and, if sufficient signal-to-noise, the "Moment 2" image (velocity dispersion map).
- Moment images are usually masked at low signal-to-noise.
- Higher order moments (skew, kurtosis…) not usually useful

- Moment 0 (integration of flux density *S* is carried out over the full width of the line profile):

$$\int S \mathrm{d}v$$

- Moment 1 = <v>:

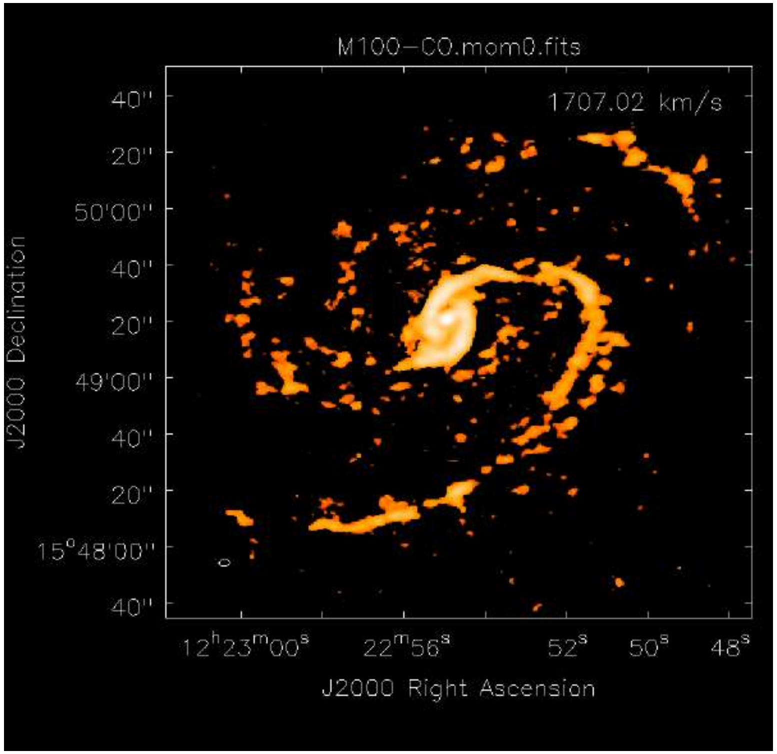
$$\frac{\int vS\mathrm{d}v}{\int S\mathrm{d}v}$$

- Moment 2:

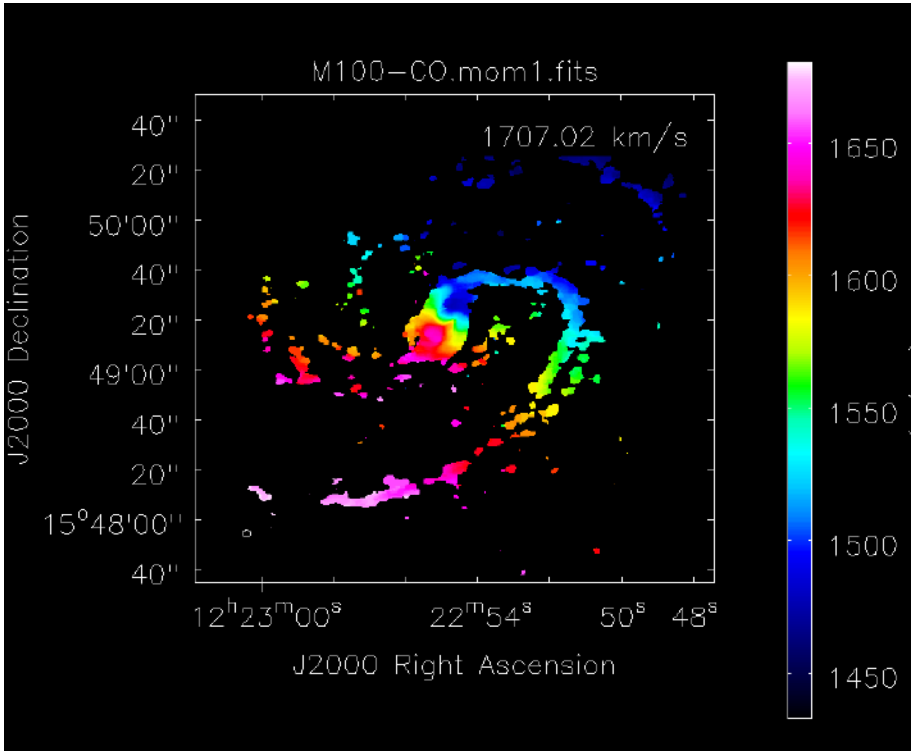
$$\frac{\int (v - <v>)^2 S\mathrm{d}v}{\int S\mathrm{d}v}$$

Example moment maps (M100)

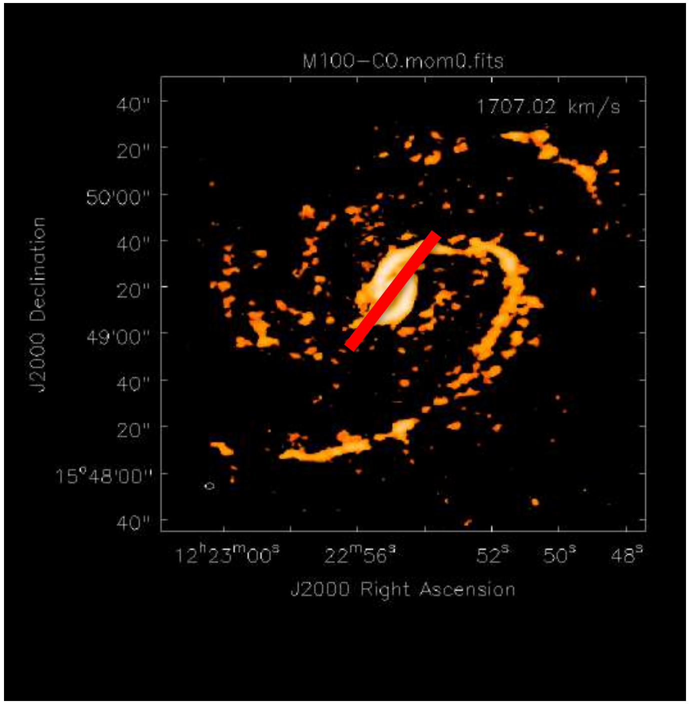
Moment 0

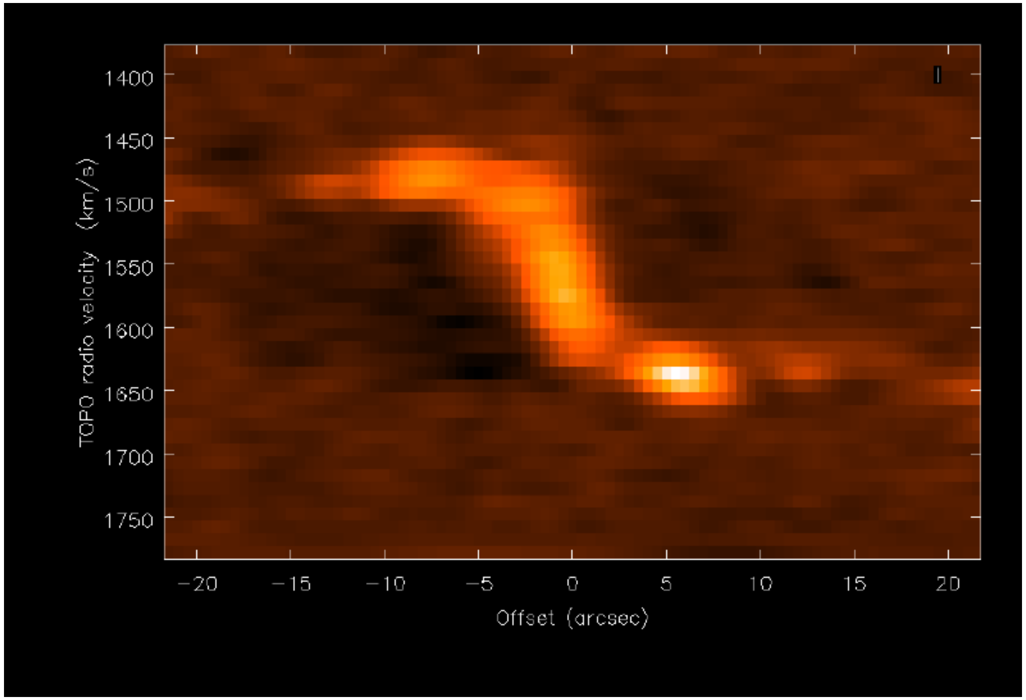


Moment 1



Position-velocity diagrams



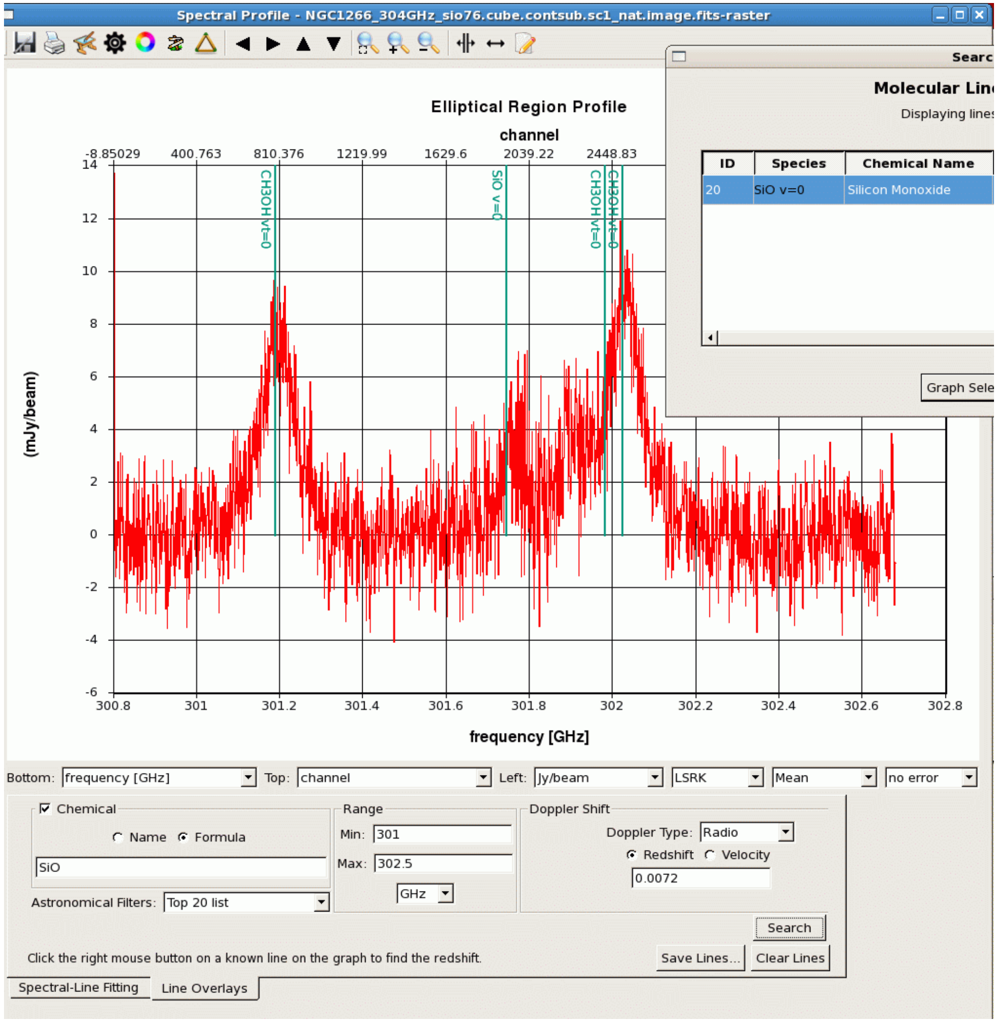


Spectral extraction and line identification

- Can use e.g. the CASA Viewer to extract a 1-D spectrum from a specified region.
- Line identification also now be done in the Viewer, and/or with the aid of Splatalogue.
- Usually good to filter by species or common astrophysical lines to avoid getting too many results.

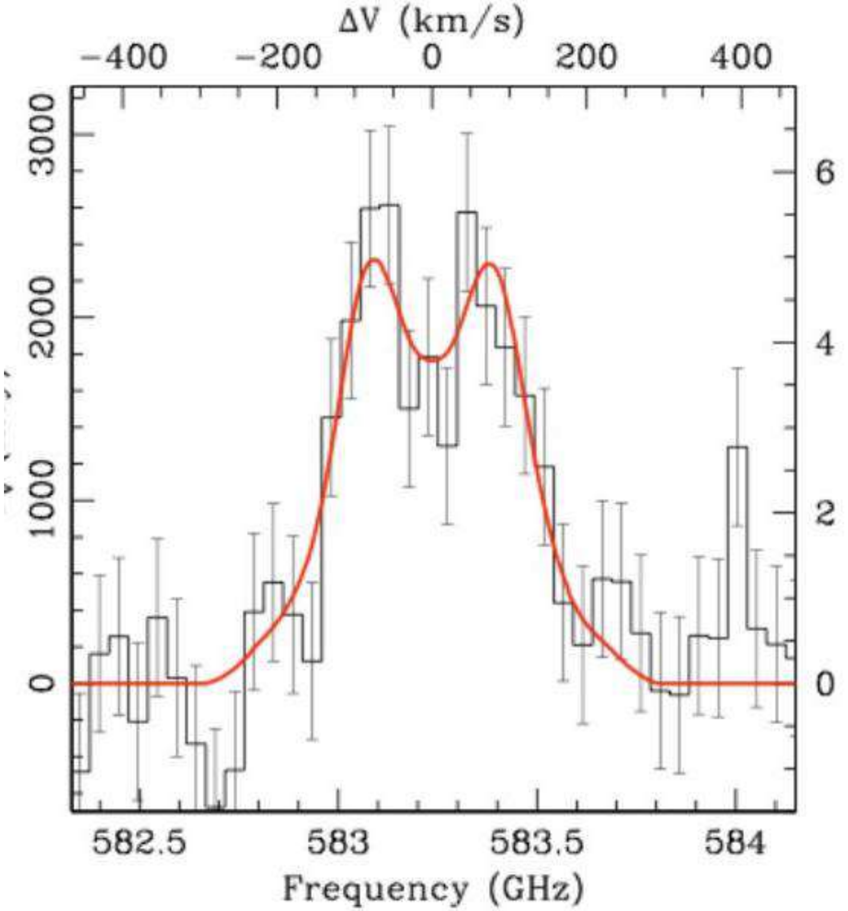
- Splatalogue (www.splatalogue.net) provides a web interface to a molecular line database maintained at NRAO. (Other databases also exist, e.g. JPL, Cologne)
- Most transitions you will be looking at are in the ground vibrational state (v=0). E.g. CO (1-0) v=0 is the usual 115.27GHz line.
- The CASA viewer allows fitting of gaussian profiles to lines.

CASA Viewer line finding example



Line profile example

- These profiles are commonly seen in extragalactic radio surveys, are they:
  A. Two galaxies merging?
  B. An instrumental artifact caused by bad calibration?
  C. The line profile of a disk-dominated galaxy?



New and automated algorithms

- Current techniques focus on reducing a 3D cube to a 1D spectrum or a 2D moment map or position-velocity diagram.
- Spatial, chemical and velocity information should really be analyzed together to get the most out of a dataset.
- “Clump finding” algorithms address this by characterizing emission in the 3D cube directly.

- Johnathan Williams’ clump-finding algorithm (IDL) http://www.ifa.hawaii.edu/users/jpw/clumpfind.shtml
- Starlink’s CUPID package contains several clump-finding algorithms, including the Williams one (http://starlink.jach.hawaii.edu/starlink/CUPID).

NRAO

A note on units

- Line fluxes can be expressed in several different ways:
  - Wm^{-2}
  - $erg\ s^{-1}cm^{-2}$
  - $Jy\ kms^{-1}$
- Also in terms of surface brightness:
  - $K\ kms^{-1}$
- Similarly, units of luminosity vary:
  - W
  - $ergs^{-1}$
  - $K\ kms^{-1}pc^2$

- Conversions:
  - 1 W = 10^7 erg
  - 1 Jy = 10^{-26} Wm^{-2}
  - Velocity integrated flux to integrated flux:

$$\frac{S}{[\mathrm{W\ m^{-2}}]} = \frac{1}{\lambda_o} \frac{S^V}{[\mathrm{Jy\ km\ s^{-1}}]} \times 10^{-23}$$

where λ_o is the observed wavelength in meters.

More flux and luminosity conversions

- Conversions continued:
  - For a source subtending solid angle Ω sr with a velocity-integrated brightness temperature T_B K kms^{-1} (all constants in SI units):

$$\frac{S^V}{[\mathrm{Jy\ km\ s^{-1}}]} = \frac{2k_B\nu_o^2\Omega}{c^2}\frac{T_B}{[\mathrm{K\ km\ s^{-1}}]} \times 10^{-26}$$

  - For luminosities,

$$\frac{L}{[\mathrm{W}]} = 4\pi\left(\frac{D_L}{[\mathrm{m}]}\right)^2\frac{S}{[\mathrm{Wm^{-2}}]}$$

$$\frac{L^T}{[\mathrm{K\ km\ s^{-1}pc^2}]} = 3.255\times 10^7\left(\frac{\nu_o}{[\mathrm{GHz}]}\right)^{-2}\left(\frac{D_L}{[\mathrm{Mpc}]}\right)^2(1+z)^{-3}\frac{S^V}{[\mathrm{Jy\ km\ s^{-1}}]}$$

Where ν_o is the observed frequency and D_L the luminosity distance (the same as the source distance at non-cosmological distances).

Further resources

- The ALMA CASAguides (http://casaguides.nrao.edu/index.php?title=ALMAguides) (especially the TWHydra guide) provide an excellent resource for understanding calibration and imaging of ALMA spectral line data.
- Similarly, for the VLA the http://casaguides.nrao.edu/index.php?title=EVLA_high_frequency_Spectral_Line_tutorial_-_IRC%2B10216 CASAguide is very useful
- Appendix A of Obreschkow et al. 2009 provides a very useful guide to the units and measures of spectral line emission used in radioastronomy.
- The ALMA Technical Handbook (from www.almascience.org) contains detailed information on ALMA spectroscopy (and is also more generally useful).