Laboratory Data, Line Confusion and Other Unique Opportunities and Challenges for ALMA Line Surveys

John Pearson JPL

Introduction

- ALMA will revolutionize the way molecular astrophysics is performed
  - 8 to 16 GHz of SSB spectra per LO setting
  - Tiny beam on the sky
  - Previously unimaginable sensitivity
- ALMA is big science that requires better tools and collaborations
  - Spectroscopy, chemistry, data bases and modeling
- Fully exploiting ALMA requires a paradigm shift in spectral analysis
  - Applies to both astronomical and laboratory data
  - Requires new tools and techniques
- A new approach will be proposed to address the coupled problems of laboratory data, line confusion and ALMA

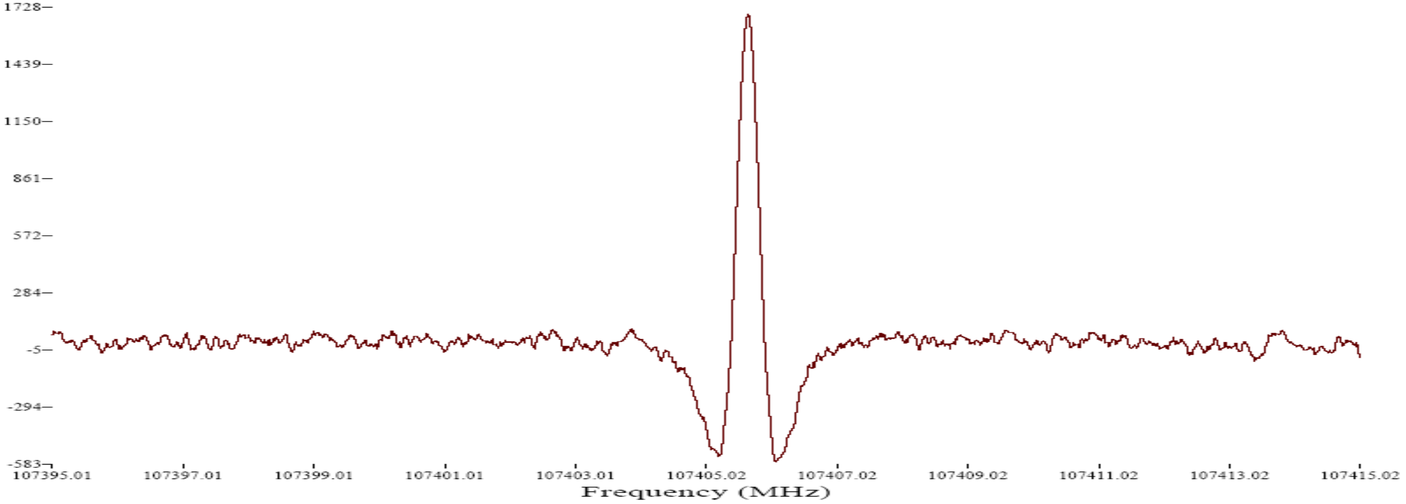
Key ALMA Questions

- ALMA will try to understand the details of star and planet formation
  - Cloud collapse
  - Planet formation processes
  - Distribution of organic material
  - Transport of organic material
- Ultimately we would like to address questions about life and habitability
  - Need to understand molecular complexity
  - Need to detect more complex (biogenic) molecules
- Submillimeter observations probably hold keys to
  - Identification of PAH bands to individual molecules
  - Understanding the PAH formation mechanism(s)
  - Identification of carriers of DIBs

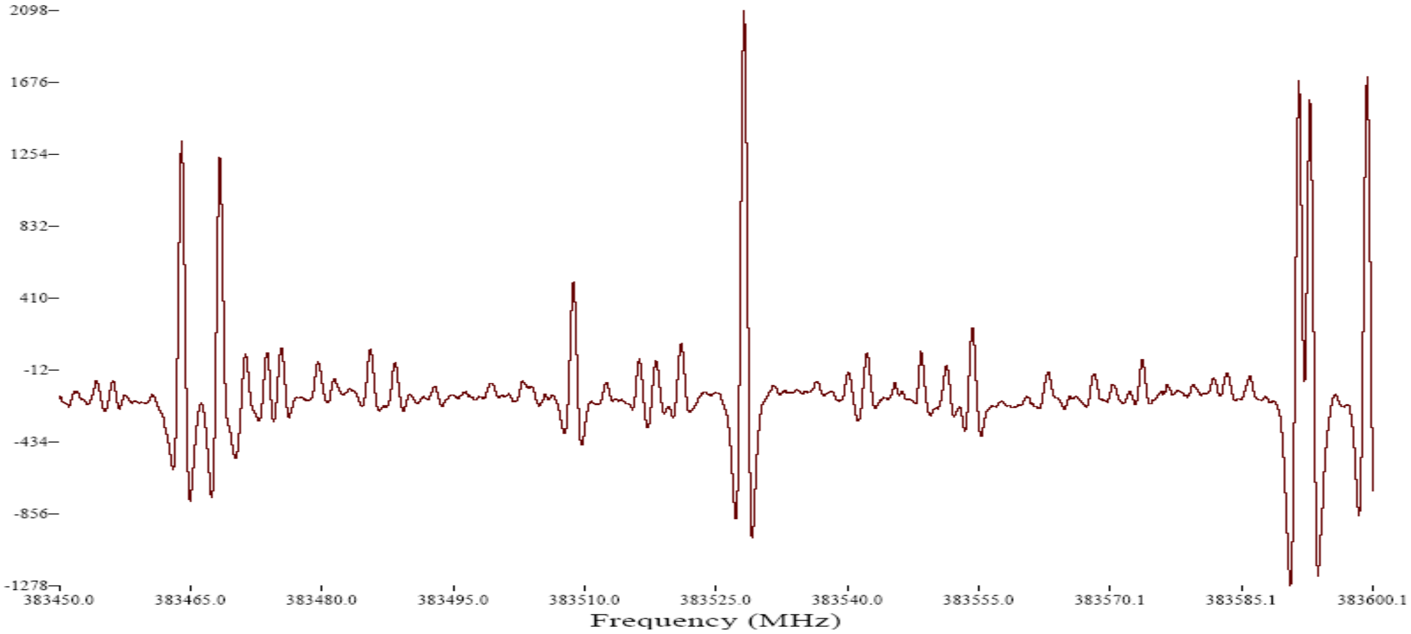
Submillimeter Spectroscopy

- Historically laboratory spectroscopy measured a few lines and fit these to a Hamiltonian model to predict the remainder of the spectrum
  - Measurements one at a time
  - Technical difficulties with radiation sources
- Historical astronomical spectroscopy measured a few (strong) lines of a molecule and assigned it to a molecule
  - Technical difficulties with LO sources
  - Narrow IF receivers
  - Poor system temperatures
- Problem is strong lines are mostly assigned and weak lines are everywhere
  - Lab spectra of complex molecules is equally if not more confused
  - A few lines are insufficient to identify a complex molecule

Good Old Lab Spectra

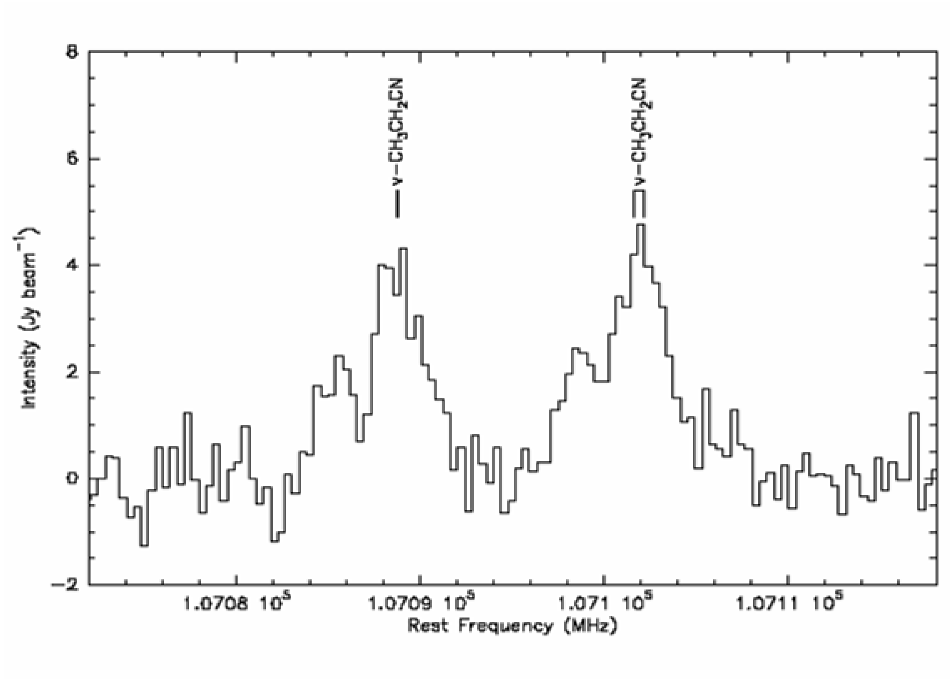


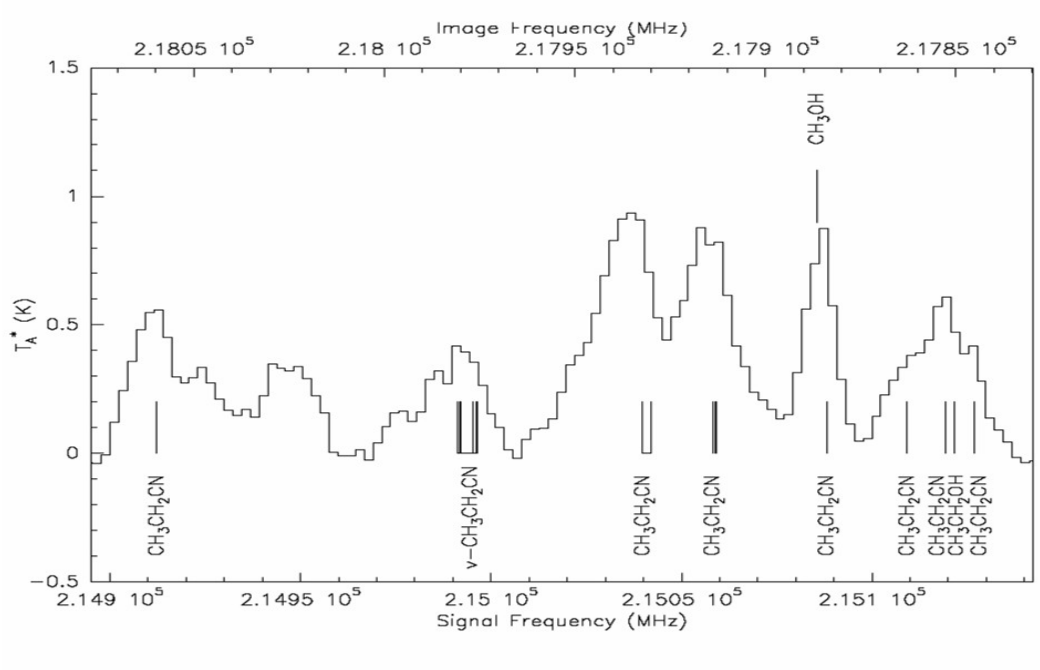
20 MHz
Narrow band
Typical to measure
1 Line
JPL (Klystron)



150 MHz
"wideband"
Searching for lines
JPL (Klystron)

Traditional Astronomical Spectra

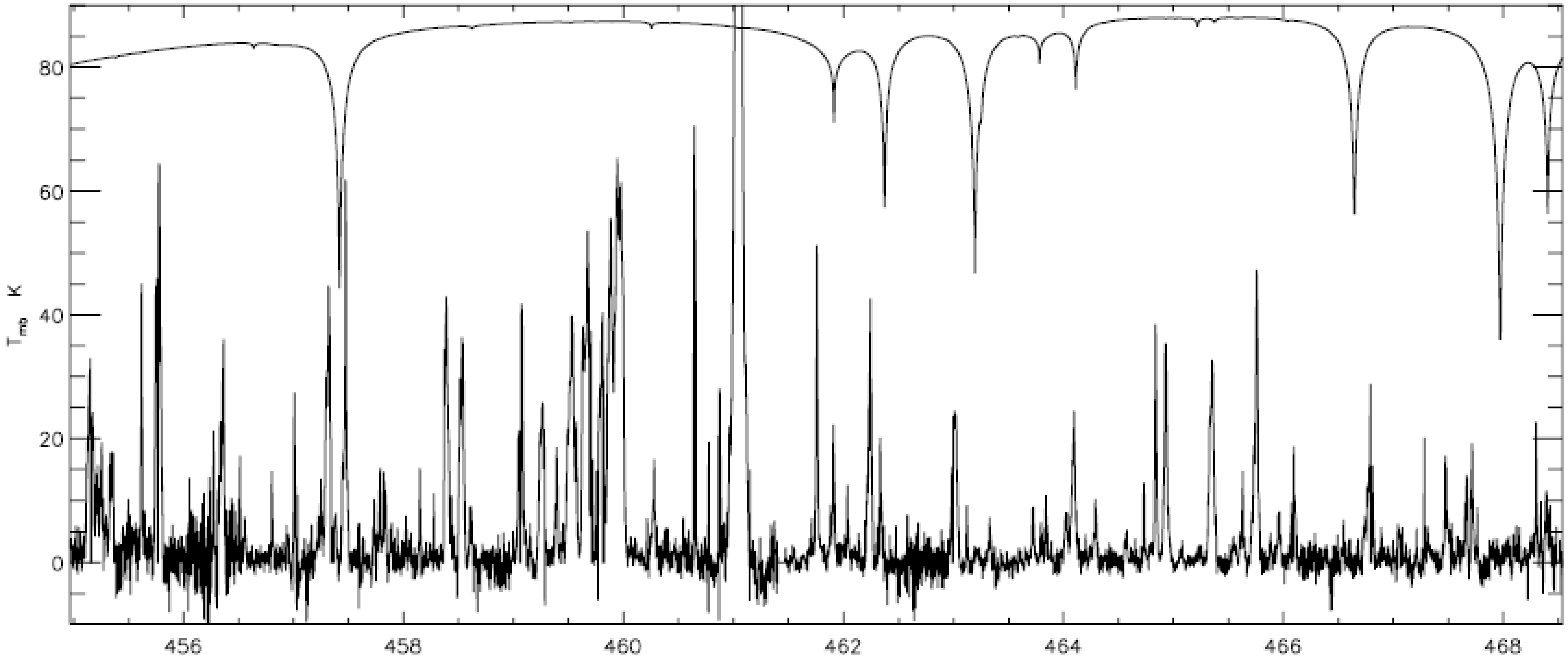




Narrow band Spectrum
~50 MHz from BIMA

“Wide band” Spectrum
~250 MHz from CSO

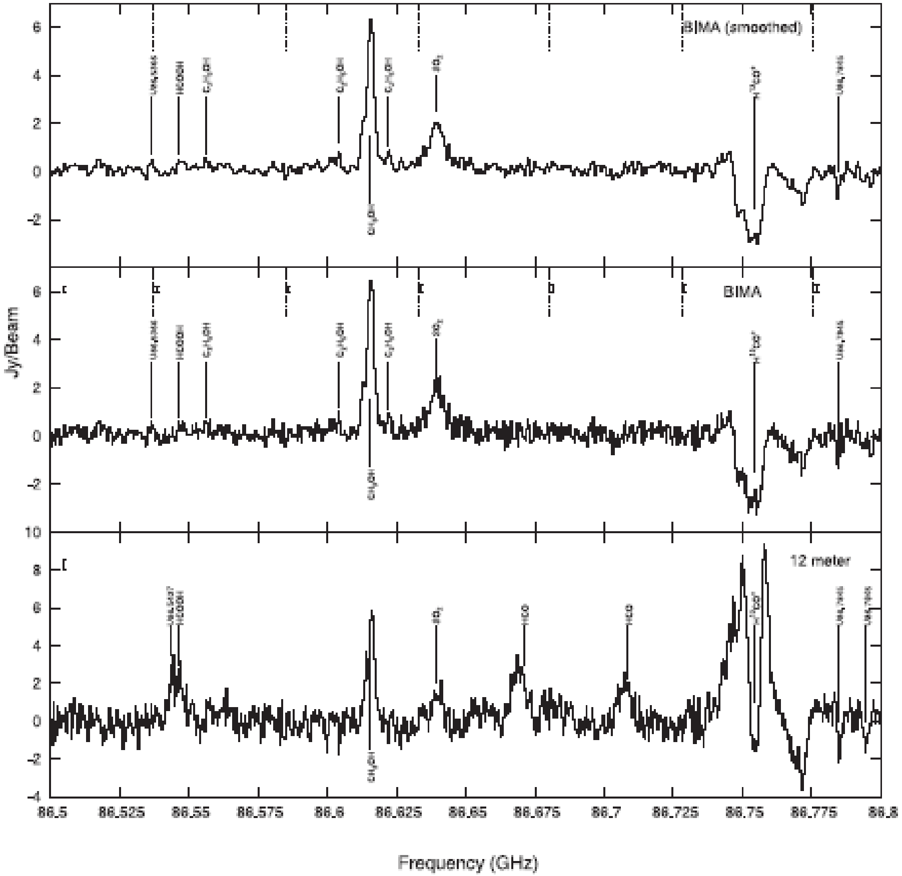
The ALMA challenge (1/2)

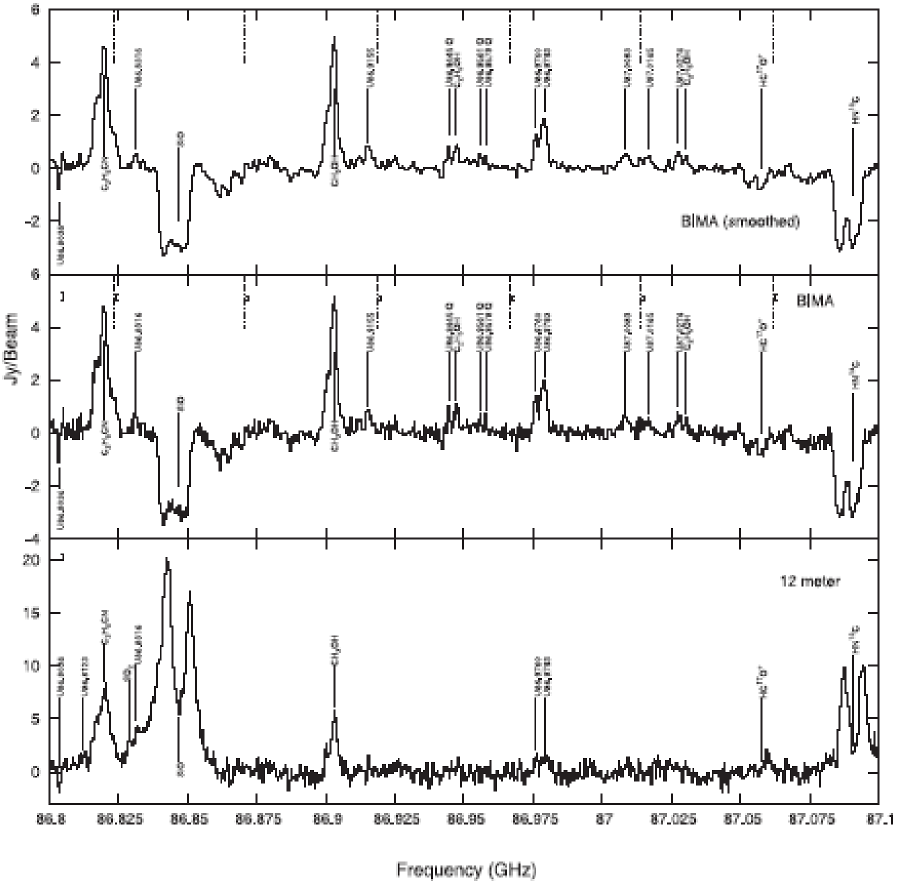


Orion Survey G. White et al., 2003, A&A 407, 589 Total of 13.5 GHz

But the ALMA beam is smaller and the noise floor will be a lot lower!
There will be 100’s of lines in each spectra of any rich source

The ALMA Challenge (2/2)





Friedel et al., 2004, ApJ 600, 234

BIMA vs 12 meter spectra of Sgr B2 (600 MHz total)
ALMA will go much deeper and obtain much more spectra at each LO setting

Dealing with 1000’s of Lines

- Line by Line identification of ALMA will not be efficient
  - Ignoring of weaker lines will be common
  - Develop tools to “automate” assignment (not in current plan)

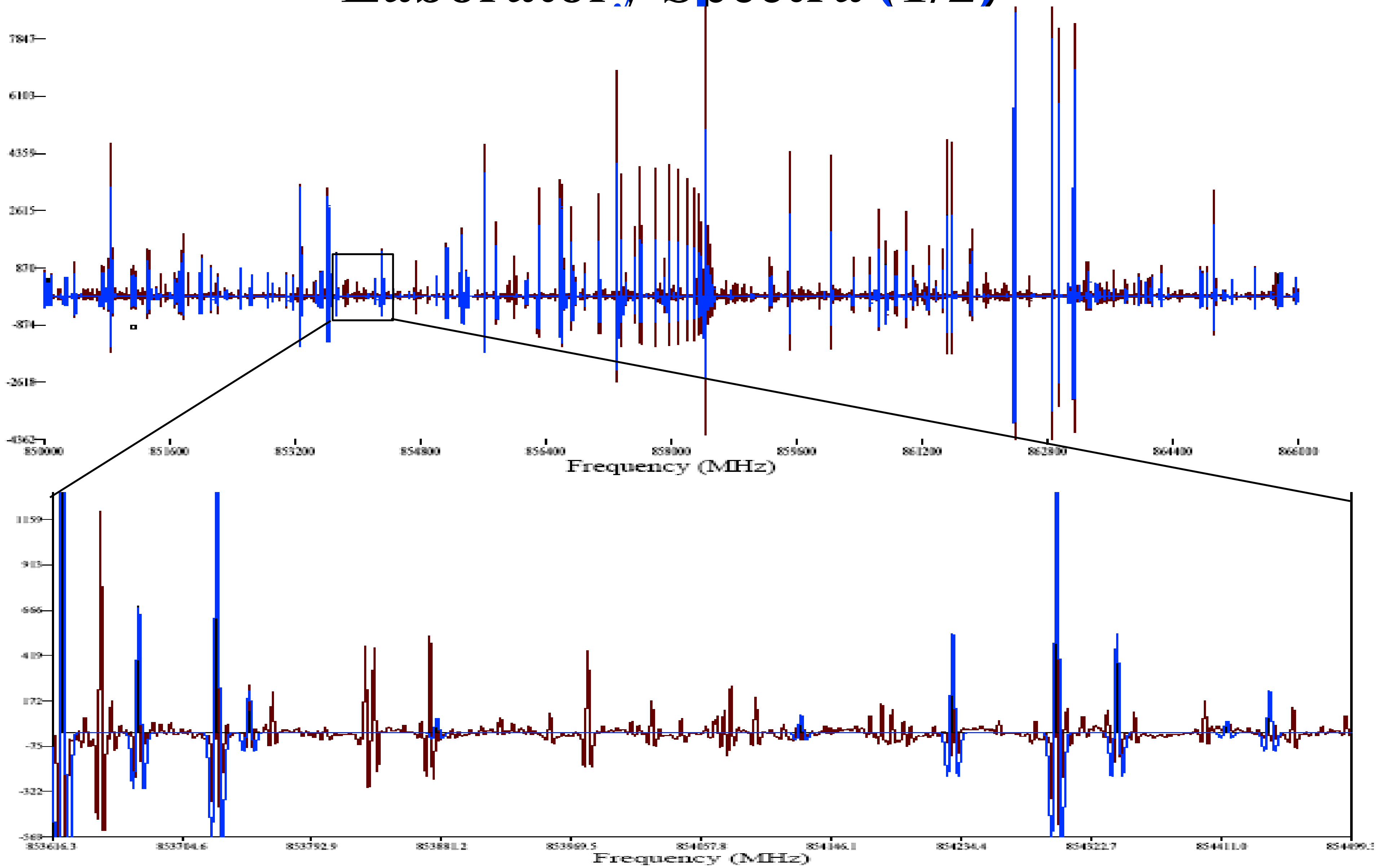
- Fundamental input for assignments are complete and reliable catalogs
  - Ground state main isotopes to modes quantum numbers are in pretty good shape
  - Isotopic variants, vibrational states and larger quantum numbers generally need help
  - Analysis of JPL and CDMS data bases ~850 pieces of spectroscopy to complete
    - Just the known molecules
    - About 60 pieces of spectroscopy to solve worst offenders

- Molecule-by-molecule assignment tool is needed
  - Needs excitation model and all isotopes of each molecule
  - Should globally do all lines including those in the confusion
  - Isotope ratios should be a parameter (all will be there at some level)
  - Unwise to make assignment of weaker features unless multiple lines

Catalog Extrapolations

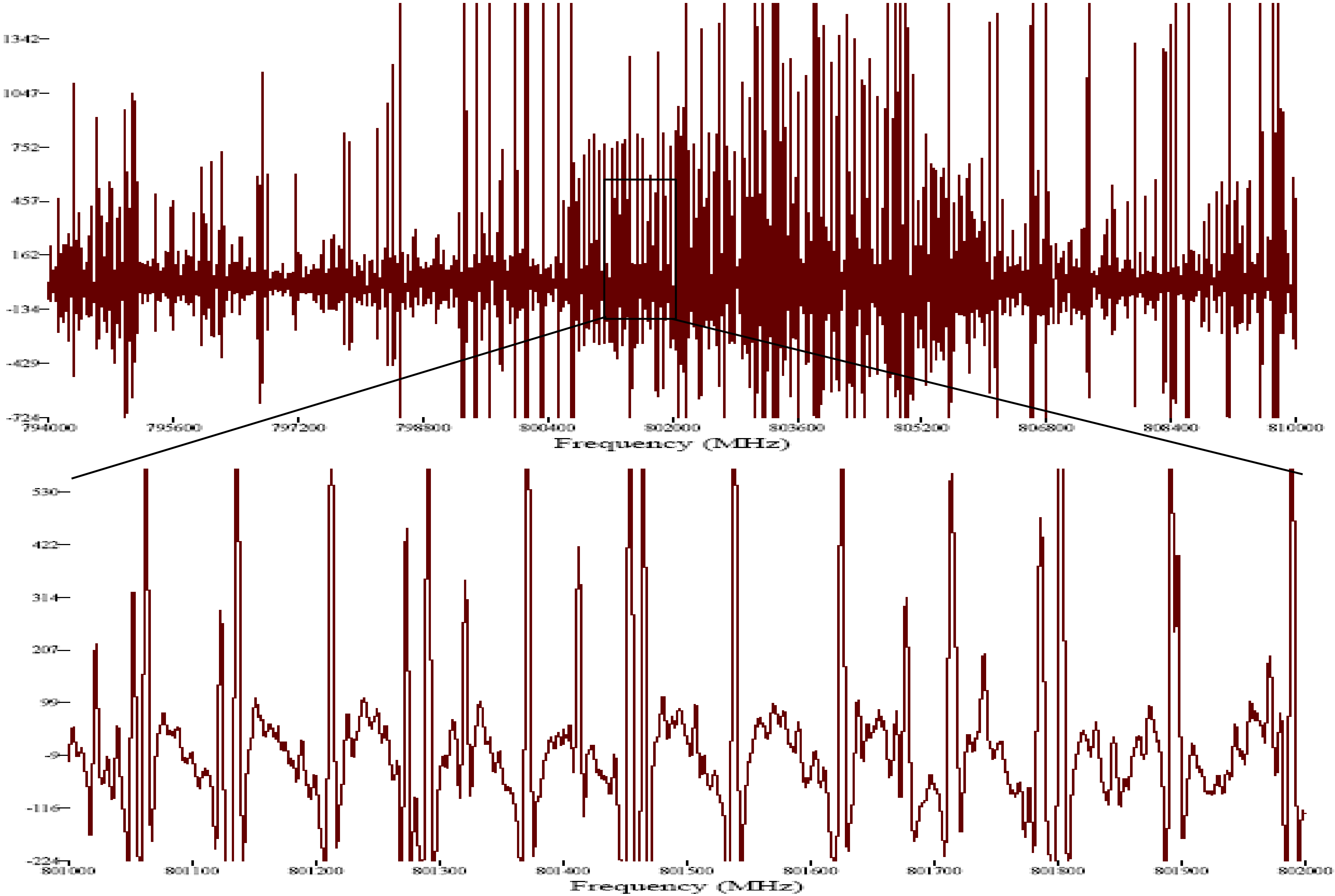
- Catalog Calculations determine distortion constants
  - Magnitude related to molecular potential
  - Hamiltonians parameterizes as even powers of quantum numbers (J^{2n}, K^{2n})
  - Undetermined constants set to zero with zero uncertainty (not reality)
  - Extrapolations are generally have overly optimistic errors
- Non-rigid molecules have worse problems
  - Power series non convergence (e.g. CH_2, NH_2, H_2O)
  - Methyl internal rotation (e.g. CH_3OH, CH_3OCHO, $(CH_3)_2O$)
    - Series in K has limited (or no) predictive power
  - Vibrational interactions (perturbations) matrix element is proportional to $1/\Delta E$
    - No predictive power unless in model
- Spectroscopy and cataloging must be done or data will not be reliable
  - Currently no ALMA contribution to molecular data

Laboratory Spectra (1/2)



16 GHz of CH_3CH_2OH spectra, blue is model brown is data with many U-lines

Laboratory Spectra (2/2)



CH_3CH_2CN almost line confused but not quite

How many states are needed (1/2)?

- **Column is divided up into populated states**

$$Q = Q_{el} Q_{vib} Q_{rot} Q_{spin}$$

$$Q_{rot} \cong \left(\frac{5.34 \times 10^6}{\sigma} \right) \left(\frac{T^3}{ABC} \right)^{1/2}$$

$$Q_{vib} = \left(\sum_{v_1} e^{-v_1 h\omega_1 / kT} \right) \left(\sum_{v_2} e^{-v_2 h\omega_2 / kT} \right) \left(\sum_{v_3} e^{-v_3 h\omega_3 / kT} \right) \ldots$$

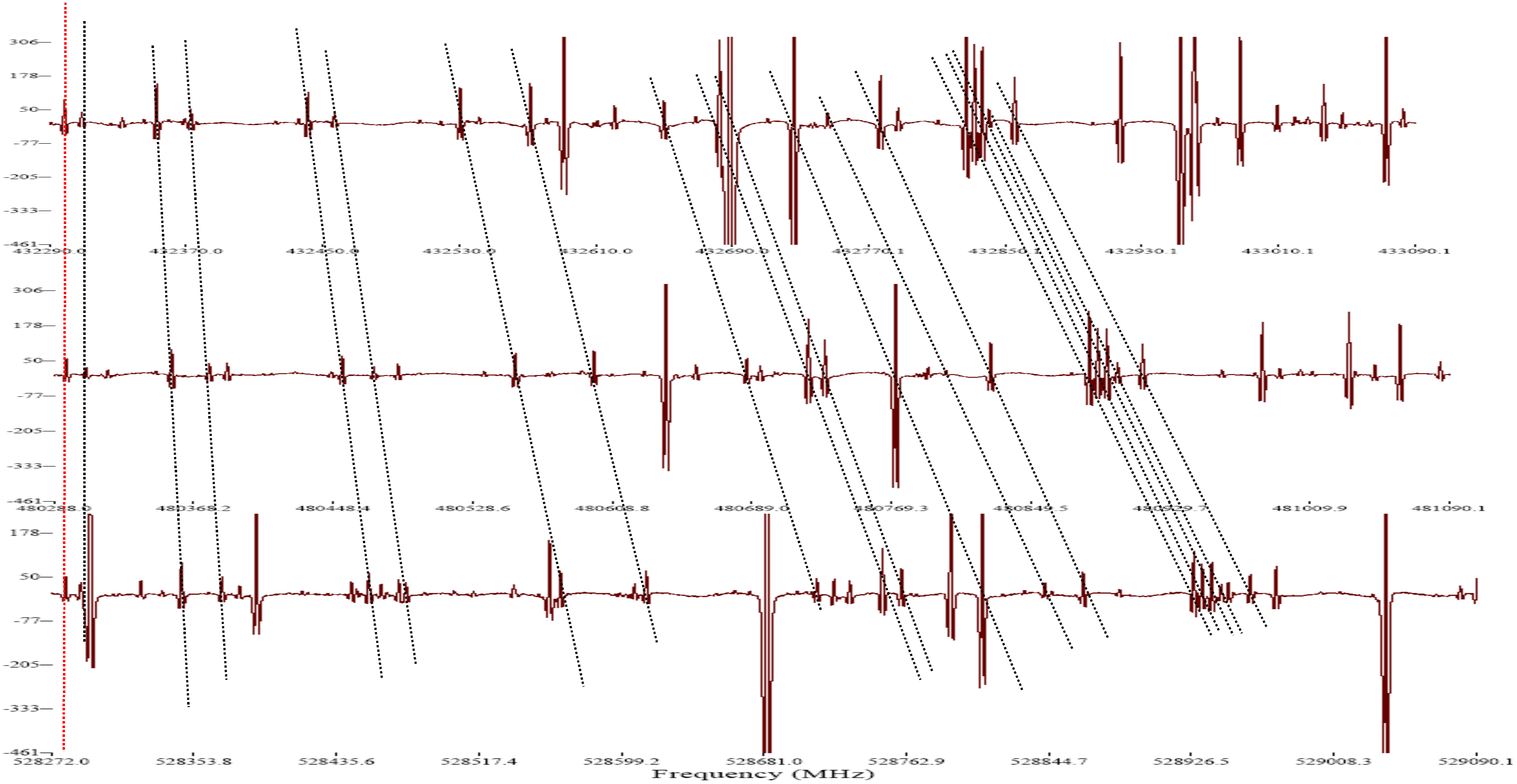
- **In thermal equilibrium strongest vibrational lines are $e^{-h\omega/kT}$ weaker**
  - **CH_3CH_2CN torsion and in-plane bend are 200 cm^{-1} and out of plane bend is at 378cm^{-1}**
  - **For a 200K source**
    - **torsion and in-plane bend are 23.7% as strong**
    - **Out-of-plane bend is 7.5%**
    - **Triad of 2 in-plane bend, 2-torsion, in-plane with torsion are 5.6%**
    - **Next set of four states at 600cm-1 are 1.3%**

- **If strongest GS lines are 1 K and noise floor is 1 mK ALMA will easily see states at 600cm^{-1} (13mK on the strongest lines)**

How many states are needed (2/2)?

- Isotopes are similar problem (assuming Earth abundance)
  - $^{13}CH_3CH_2CN$, $CH_3{}^{13}CH_2CN$ and $CH_3CH_2{}^{13}CN$ are 1.1% or 11mK
  - In-plane bend and torsion of these would be ~4mK
  - Out-of-plane bend is at detection limit
  - $CH_3CH_2C^{15}N$ is 3.6mK
  - In-plane bend and torsion at the detection limit
  - D species (two equivalent) are 0.45 and 0.3 mK and would not be detectable
- Similar scary calculations can be done for any molecule
- Reality for ALMA is probably worse than scary calculation
  - Small beam will resolve hottest gas
  - High sensitivity will surely reach line confusion limits
  - Lab data (and molecular physics) is not close to being ready for ALMA
- ALMA will have to deal with line confusion

Laboratory Line Identification



Red is v_t=3 K=0 A state lines of methanol

Dotted lines go through more ^{a}R branches in Methanol

Lines are identified by branch, visual techniques help enormously

Large pieces of spectra greatly simplify problem

Astronomical Line Identification

- Astronomical U-line problem is very similar to lab identification
  - Visualization tools could be used in the same way
  - Lab is LTE and LTE is where you start for astronomy
  - Both need a tool that allows the best available molecular data to be imported

- One tool if written properly could work for all high resolution sub-millimeter spectroscopy
  - Modularity will be needed for unique problems of astrophysics
  - Line tools (other than derivative line shapes) are the same
  - Collectively ALMA, Herschel, SOFIA and laboratory spectroscopy might have the expertise and resources
    - Good starts in labs and by astronomy groups
    - A vision and organization is needed to produce a tool
  - Unfortunately not in any official plan but there is a serious need

Tool Requirements (1/4)

- Ability to import lab and astronomical data (probably in several formats)
  - SSB and DSB
- Ability to import molecular data (Catalog files)
  - By molecule, all isotopes and vibrational states
  - Isotope abundance parameterized
  - Red shift automatically
- Calculate LTE excitation of molecular data
- Ability to import arbitrarily complex source model
  - Calculate simulated spectra with different relative abundances for components
  - Built in uniform density and temperature of fully filled beam
- Ability to calculate line shapes
  - Gaussian setable width (along with derivatives)
  - Voigt for Laboratory (along with derivatives)

Tool Requirements (2/4)

- Graphical Interface capability
  - Overlay modeled and actual spectra
    - Including previously assigned species color coded
    - Click on accepted assignment of molecule (locks all in source model)
    - Click on accepted assignment of line
  - Display Sorting
    - Sort molecules by branch P,Q,R
    - Sort molecules by K (aR branch)
  - Calculate Obs-Calc and display residuals
    - Frequency error looks like derivative
    - Improper width gives two humps
    - Improper excitation gives systematic deviation with energy
    - Blended lines will have obvious residuals
    - Source component errors will be systematic
    - Must display on sorted data

Tool Requirements (3/4)

- Fitting and export
  - Export of global and by molecule obs-calculated to source models
    - Run off line with possibility to re-import updated version
  - Export of frequencies and quantum numbers to molecular model
    - Run off line with possibility to re-import improved catalog
  - Should fit line shapes including known blends
    - Should use Temperature to extract turbulance
    - Should fit optically thick shapes with a single click
    - Graphical display of residuals, export of residuals
  - Automatic export of assigned lines
    - Export of flux (line confused) by molecule and in total
  - Export of determined model parameters
  - Export of global spectra and obs-calculated

Tool Requirements (4/4)

- U-line Tools
  - Spectral parsing tool (Loomis-Wood)
    - Line up in strips (B+C)/2
    - Knob to fine tune spacing
    - Point and click tentative assignment
  - Fortrat branch analysis
    - Parse and display predicted branch
    - Knob to tune origin
    - Point and click tentative assignment
  - Bruit force combination differences for u-lines
    - Rydberg-Ritz combination differences
      - Calculates combination differences of all U-lines
      - Approximate calculation of A,B,C from loops
  - General signal processing (real and model data)
    - Box car average for smoothing
    - FFT for baseline
    - Numerical derivatives

Conclusions

- ALMA is big science that will generate an enormous amount of spectroscopic data
  - An enormous improvement in the existing spectroscopic data is required
  - An enormous improvement in the existing data catalogs is required
  - Old line by line analysis astronomical analysis will be prohibitively expensive
  - New tools and techniques based on a whole spectrum approach must be developed
    - Identification based on a few lines is dangerous under the best conditions
- A cross discipline approach is required to develop the tools and techniques needed
  - Observers, Modelers, Molecular Physicists and Programmers are required
  - Some understanding of cross discipline subjects will be necessary
- There is no unified effort and no real resources for this development
  - The few best effort attempts need to be unified and incorporate spectroscopic expertise
  - The resources need to be found if ALMA is maximize its science return

Acknowledgement

- This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration
- Thanks to the Herschel project for supporting my attendance
- Special thanks to Dr. Brian Drouin for providing spectra and his spectral line tool, SMAP, available at http://spec.jpl.nasa.gov/ftp/pub/calpgm/