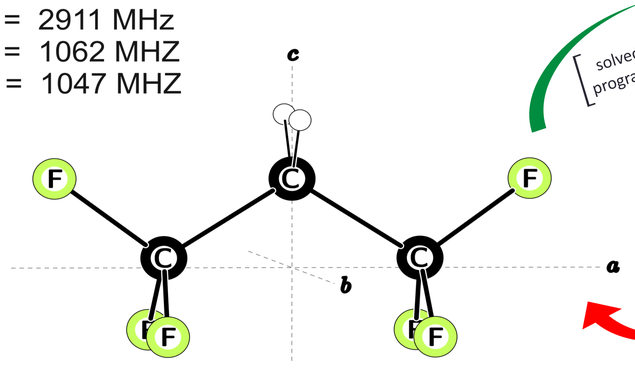
# Robust Autonomous Assignment of Rigid Rotors

Lia Yeh, Dave Patterson
Department of Physics, University of California, Santa Barbara

## MICROWAVE SPECTROSCOPY IS A TOOL FOR MIXTURE ANALYSIS

### Rotational constants

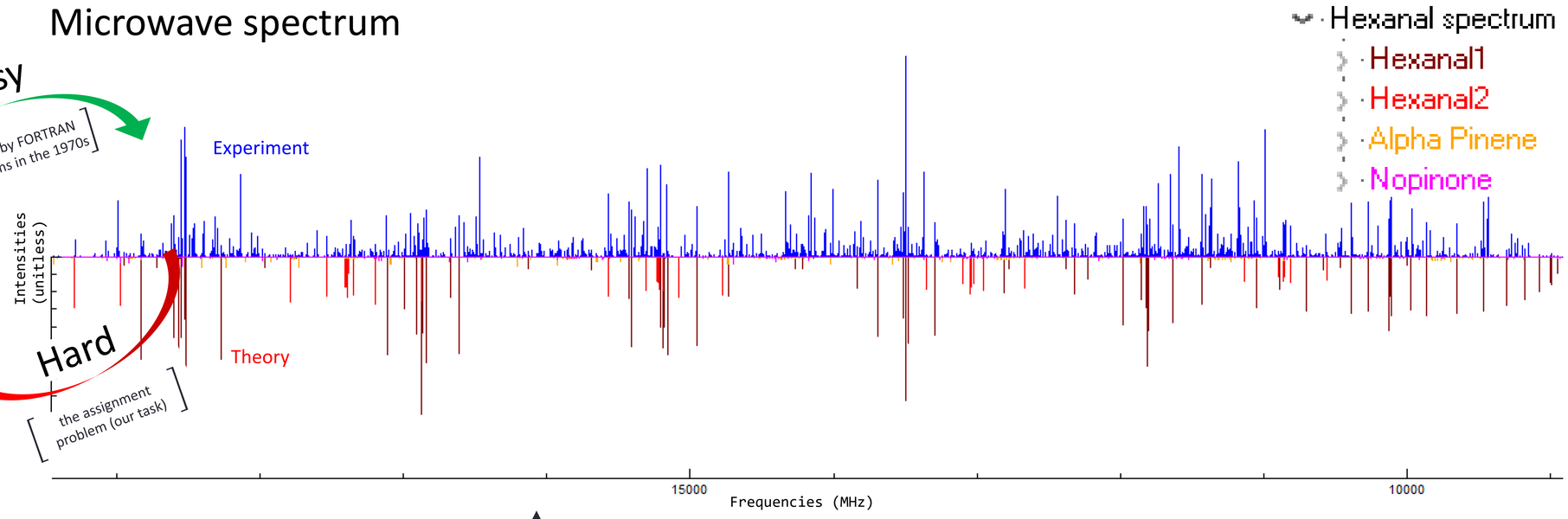
*A* = 2911 MHz
*B* = 1062 MHZ
*C* = 1047 MHZ




Easy
solved by FORTRAN programs in the 1970s

Hard
the assignment problem (our task)

### Microwave spectrum




**Cold molecules are a new frontier** for quantum physics and chemistry. In the microwave regime, we observe discrete spectra of the molecules' rotational transitions, with 100 to 5,000 distinguishable lines in a typical spectrum.

The assigned microwave spectrum provides an accurate measurement of the rotational constants A,B, and C of a molecule*; these constants are related to the principal moments of inertia $I_x$, $I_y$, and $I_z$ via

$$A = \frac{h}{8\pi I_x} \qquad B = \frac{h}{8\pi I_y} \qquad C = \frac{h}{8\pi I_z}$$

If we know a molecule's rotational constants, generating the molecule's theoretical spectra is a simple matter of running decades-old software.

**The challenge is,** if we have a spectra of an unknown mixture, how do we figure out which rotational constants belong to the molecules in the mixture? As rotational constants uniquely identify a molecule, this is the same as asking what molecules are in the mixture.

This problem, called the assignment problem, has for decades eluded all but a few human experts.

We are designing and implementing an algorithm to solve the above hard problem, generalized to any mixture of molecules. For the above mixture, for instance, our code identified four molecules: two conformers of hexanal and two contaminants from spectra taken previously.

**ALGORITHM**

We present the first fast, context-free solution to spectral assignment, needing no prior knowledge of the molecules other than the spectra itself.

**ACCURACY**

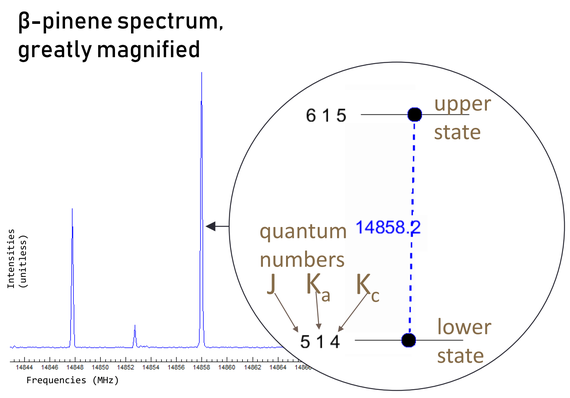
Molecules are very diverse in shape, size, rigidity, and magnetic properties. We use an efficient strategy to target each major case.

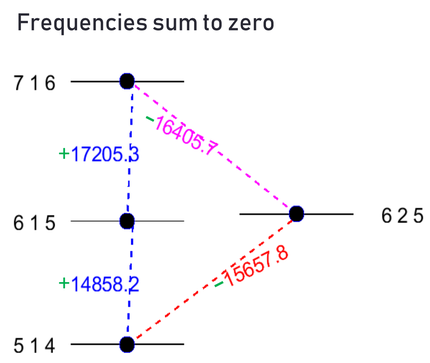
**APPLICATION**

We push the boundaries of quantum chemistry by performing mixture analysis on microwave spectra with the press of a button.

## THE SPECTRAL ASSIGNMENT PROBLEM

**The heart of the assignment problem** is to correctly "guess" a set of line assignments. An assignment is a pairing between a frequency and the quantum numbers J, $K_a$, and $K_c$ of its two states. To find the three rotational constants, we must guess at least three lines to provide a solvable system.




We vastly reduce the number of guesses required through pattern finding. As an example, consider the 14858.2 MHz line of β-pinene above. Note that 14858.2 + 17205.3 – 16405.7 – 15657.8 = 0.

By only considering four lines that add up to approximately zero, we have significantly narrowed our search space. These four lines are actually part of a larger structure, which we've termed a scaffold:

## THE BIGGER PICTURE


**Given a mixture** of chemicals, we aim to non-destructively measure the structure of each chemical. Our algorithm is capable of identifying traces of molecules too faint for the human eye to ascertain. This provides us with a powerful tool capable of bringing us closer than ever before to:

- assign more molecules, faster
- identify contaminants in our samples
- determine molecules that two spectra have in common
- given a molecule's rotational constants, find its isotopologues and conformers

We are working towards the ability to calculate the x-, y-, and z- positions of atoms in the molecular backbone. This would have far-reaching ramifications for atomic, molecular and optical physics; astrochemistry; molecular biology; pharmaceuticals; forensics; and many more scientific fields.

## HOW IS OUR CODE SO FAST?

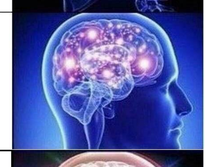
**Our runtime** vs other approaches:

The number of plausible assignments is enormous for even the smallest spectra of 100 lines. This search space increases exponentially with the number of lines in the spectra.

When we first approached this problem, for each correct guess of a set of line assignments, we also found thousands of false positives.

We resolved this by:

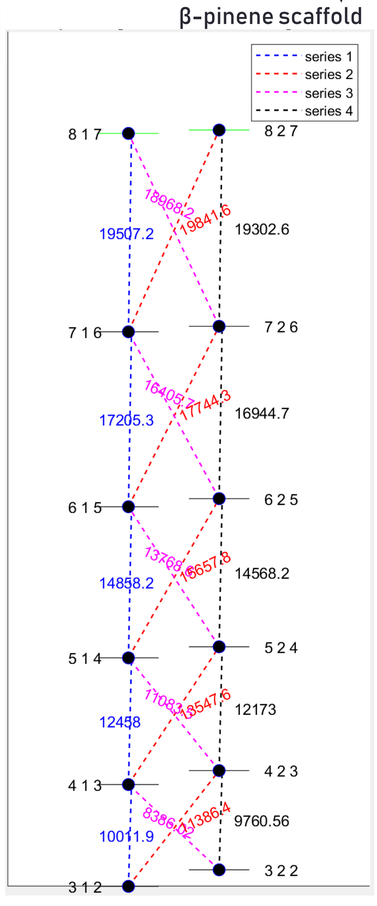
- Finding R-branch, a- & b-type series
- Searching by probability of success
- Weeding out unlikely guesses

| | |
|---|---|
| Brute force **> 3 years** |  |
| Human expert **> 4 hours** |  |
| Human record **~ 10 minutes** |  |
| RAARR **~ 10 seconds** |  |

*RAARR's limitations beyond those inherent to microwave spectroscopy:

- It only works on rigid rotors – molecules that don't distort much as they rotate. >80% of the molecules around us are rigid rotors anyway
- It currently misses ~30% of cases; some we suspect aren't rigid rotors
- It prefers spectra with broader range and higher signal-to-noise

β-pinene scaffold

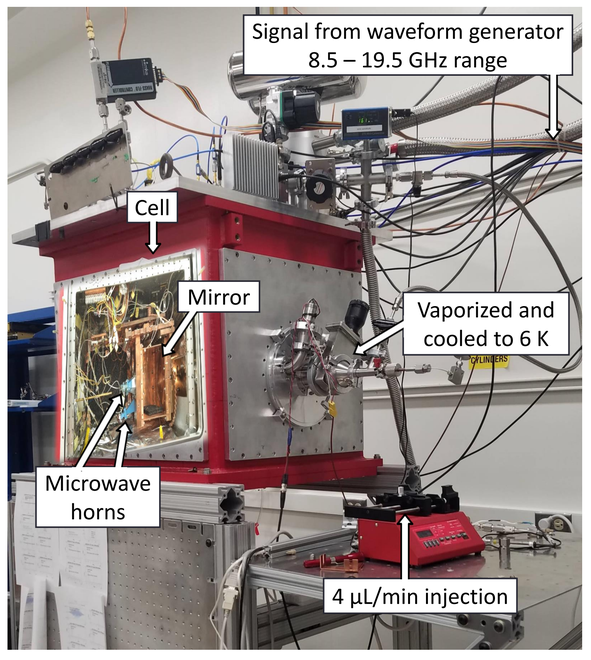



## HIGHER RESOLUTION SPECTRA

Taking spectra:

1. We steadily inject our mixture so it is a cold gas by the time it enters the cell.
2. Microwaves enter from one horn, bounce off the mirror, and are collected in another horn, bombarding all molecules in their way.
3. The received signals are Fourier transformed, giving this method its namesake, Fourier Transform Microwave (FTMW) spectroscopy.
4. We process this data by calibrating the intensity values to account for known hardware limitations, finding the frequency peaks, and subtracting the noise floor.

We are expanding our broadband range to reach ~30 GHz*. This will improve our ability to assign our spectra, particularly for very small molecules.




## ACKNOWLEDGMENTS

Dave Patterson  Sathya Guruswamy  Samantha Davis  Jin Kim