**Introduction:** The chemistry of the interstellar medium is relatively unknown, due to the nature of the large volume and low temperatures of “empty” space. Molecules and ions can exist in unstable states for significant periods of time. Therefore, there are compounds foreign to Earth that we can’t spectroscopically analyze. To understand the characteristics of these molecules, we need to use computational methods.

One useful tool researchers have to identify compounds is infrared spectra, which correspond to the vibrational motions of a molecule. We can measure the wavelengths of light present in certain regions of space, so if we look specifically at the IR region of EM radiation, then we can understand which types of molecules are present there. By using certain levels of computational theory, we can calculate the vibrational modes of the molecule based on its geometry. There are some methods (CCSD(T)) that take absurd amounts of time to generate very accurate data, and there are some semi-empirical methods (AM1, MNDO, PM3, PM6, PM7, RM1) that compute quickly but are not accurate.

Re-optimized semi-empirical calculations may help with the Polycyclic Aromatic Hydrocarbon World Hypothesis, which claims that more than 20% of all carbon in the universe exists in the form of large (C\(_{100}\)) sheets of conjugated carbon, whose vibrational spectra remain relatively unknown.

**Goal:** The purpose of this research is to find a set of semi-empirical method parameters that will yield accurate rovibrational spectra for any class of hydrocarbon.

**Methods:** We chose semi-empirical methods to improve for a few reasons: 1) they’re extremely fast, especially compared to benchmark methods, and 2) they’re easy to work with, because they use parameters from an input file to do all of their calculations. By reading and changing these parameters, we can easily affect the output energies and geometries of our calculations. Parameters are atom-specific, so we are limiting our research to hydrocarbons, specifically the 6 molecules and ions shown on this poster.

Previous research done in this lab was done fitting semi-empirical energies to high-level CCSD(T) \textit{ab initio} data for only one molecule at a time, but in the interest of having transferable parameters, we decided that optimizing our parameters while comparing data from multiple molecules would grant a certain degree of applicability for more molecules. Therefore, the key to this research was moving from optimizing over one molecule at a time to multiple molecules at the same time.

We compare our data with high-level methods by generating a potential energy surface (PES) for each molecule we are interested in. This process involves slight displacements in the equilibrium cartesian coordinate geometry. Each displacement increases the conformational energy of the molecule. We can use the time-independent Schrödinger equation, \( \hat{H} = E \), to find the scalar energy (E) of the system when we operate on the wavefunction (\( \Psi \)) with the Hamiltonian (\( \hat{H} \)). The computationally difficult part of the Hamiltonian is finding the electron-electron repulsion energy. Semi-empirical methods bypass these difficult calculations in favor of using set parameters to find these repulsive energies.

Once we’ve calculated the molecule’s PES, we can turn that into a Quartic Force Field (QFF), which is a 4th-order polynomial that approximates the PES we obtained. This QFF allows us to calculate anharmonic vibrational modes of the molecule, and those modes are the approximate IR spectra.

**Discussion:** We have looked at certain classes of molecules so far, and with a total training size of 6 molecules, there have been a few interesting observations. According to what we have done so far, when trying to optimize parameters over more molecules, it gets harder to get final results to fit known or \textit{ab initio} spectroscopic data. Another issue is that we don’t yet have the ability to directly test our best parameters with large molecules. This makes it hard to know if our parameters would scale in the first place.

There’s also no clear cut-off as to when “good enough” is, because our usual metrics for determining if an optimization is complete hardly ever indicate a useful set of output spectral data, for reasons still being investigated. The resulting spectral data we generate often gets certain vibrational modes very accurate, but there are certain vibrational modes that are so inaccurate that it outright invalidates that set of parameters from being justified as accurate for any \textit{ab initio} calculation.

**Future Work:** More molecules’ \textit{ab initio} PESs are still being calculated for us to add to our list of 6 central molecules, as we are hoping to gain further insight into whether or not any set of parameters would be transferrable. There are a variety of detailed optimization steps that have yet to be explored, like optimizing certain parameters before others, weighing certain regions of data more heavily to ensure that they are accurate, optimizing over different types of data, or trying to fix the worst relative errors first.

**Acknowledgements:**

Members of the Layfield Lab

Dr. Ryan Forthenberry, Georgia Southern University of St. Thomas Chemistry Department

UST Grants and Research Office

NASA GRANT NNX17AH15G